Lecture Notes in Electrical Engineering 516

Qilian Liang
Xin Liu
Zhenyu Na
Wei Wang
Jiasong Mu
Baoju Zhang *Editors*

# Communications, Signal Processing, and Systems

## Proceedings of the 2018 CSPS Volume II: Signal Processing

Springer

# Lecture Notes in Electrical Engineering

## Volume 516

### Series Editors

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

**China**
Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

**India**
Swati Meherishi, Executive Editor (swati.meherishi@springer.com)
Aninda Bose, Senior Editor (aninda.bose@springer.com)

**Japan**
Takeyuki Yonezawa, Editorial Director (takeyuki.yonezawa@springer.com)

**South Korea**
Smith (Ahram) Chae, Editor (smith.chae@springer.com)

**Southeast Asia**
Ramesh Nath Premnath, Editor (ramesh.premnath@springer.com)

**USA, Canada:**
Michael Luby, Senior Editor (michael.luby@springer.com)

**All other Countries:**
Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)
Christoph Baumann, Executive Editor (christoph.baumann@springer.com)

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Web of Science and Springerlink ****

More information about this series at http://www.springer.com/series/7818

Qilian Liang · Xin Liu ·
Zhenyu Na · Wei Wang ·
Jiasong Mu · Baoju Zhang
Editors

# Communications, Signal Processing, and Systems

## Proceedings of the 2018 CSPS Volume II: Signal Processing

Springer

*Editors*
Qilian Liang
Department of Electrical Engineering
University of Texas at Arlington
Arlington, TX, USA

Zhenyu Na
School of Information Science
and Technology
Dalian Maritime University
Dalian, China

Jiasong Mu
College of Electronic
and Communication Engineering
Tianjin Normal University
Tianjin, China

Xin Liu
School of Information
and Communication Engineering
Dalian University of Technology
Dalian, China

Wei Wang
College of Electronic
and Communication Engineering
Tianjin Normal University
Tianjin, China

Baoju Zhang
College of Electronic
and Communication Engineering
Tianjin Normal University
Tianjin, China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-13-6503-4 ISBN 978-981-13-6504-1 (eBook)
https://doi.org/10.1007/978-981-13-6504-1

Library of Congress Control Number: 2019930976



This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

# Contents

# Part I
# Image and Video Processing

# A Length and Width Feature Extraction Method of Ship Target Based on IR Image

Yan Chen(✉), Shuhua Wang, Weili Chen, Jingli Wu, Junwei Li, and Shilei Yao

Science and Technology on Optical Radiation Laboratory, Beijing 100854, China
chenyan340206@163.com

**Abstract.** Length and width feature of ship target is usually used as the initial criterion for ship type. A length and width feature extraction method of ship target based on IR image is proposed in this paper. At first, the preprocesses such as denoise and contrast enhancement are carried out, then the Hough transform is employed to detect the sea-sky-line, and the target potential area is determined, then edge detection and expansion and hole filling are used to obtain the whole connected region of the target. Finally, the minimum enclosing rectangle of the connected region is obtained according to the minimum area criterion, and the length and width of the minimum enclosing rectangle is the length and width of the ship target. The experimental results show that the method can effectively extract the length and width feature of ship target in complex sea-sky background, then with other auxiliary information can realize ship target recognition.

**Keywords:** Feature extraction · Ship target · IR image · Area detection · Minimum enclosing rectangle

## 1 Introduction

Feature extraction of IR image is the key step for target recognition, which directly affects the recognition result. Because of its advantages of simple intuition, high efficiency and easy extraction, length and width feature is often used as the initial decision of ship type. In foreign laboratories, size, area, and other geometric features are usually as the criterion for ship recognition. For different ship types, the size and structure embodied in IR image are different. Therefore, according to the extracting result of length-width ratio, the ship type can be judged preliminarily, and then the ship target can be identified effectively with other features [1]. Length-width ratio can be a ratio of length to height or width to height, and this paper refers to the ratio of ship deck length to ship height above waterline.

In the research of length and width feature extraction, Askari and Zerr [2] proposed to use the length-width ratio as a preliminary indicator of ship type in 2000.

Yan Chen: Mainly research on optical image processing, automatic target recognition and so on.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_1

Subsequently, Gao [3], Wu et al. [4] carried out the length and width feature extraction of ship target in SAR images. For IR image, since the detector is always affected by itself noise, system transfer function, relative motion between detector and ship target, complex sea environment and so on, the signal-to-noise ratio and contrast of IR ship image are not ideal, and the length and width feature extraction of ship is difficult. This paper studies a length and width feature extraction method of ship target based on IR image. Firstly, the IR image denoising and contrast enhancement are carried out, and then edge extraction, expansion and hole filling are used to obtain the whole connected region of the target, finally, the minimum enclosing rectangle of the connected region is obtained according to the minimum area criterion, and the length and width of the minimum enclosing rectangle is the length and width of the ship target.

## 2 Preprocessing of IR Image

IR detector always affected by itself noise, system transfer function, relative motion between detector and ship target, background clutter and atmospheric transmission and other factors, IR image always shows the characteristics of large noise, low contrast and blurred edge, which seriously affect the accuracy of length and width feature extraction of ship target, thus preprocessing is essential.

### 2.1 Denoising

The traditional image denoising algorithms such as low-pass filter, middle filter, and average filter always smooth the high-frequency information while denoising, which means smoothing the target edge information. Aiming at the image features of sea surface target, the 3D block-matching (BM3D) algorithm [5, 6] is employed to denoise the image. The flowchart of BM3D algorithm is shown in Fig. 1.

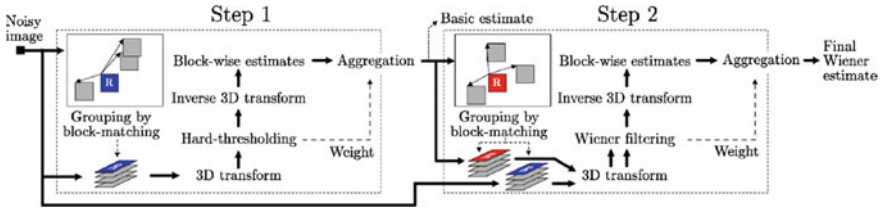



**Fig. 1.** Flowchart of BM3D

BM3D algorithm not only uses the similarity and redundancy of image, but also combines the threshold method of transform domain, thus it is a multi-scale and non-local denoising algorithm. Firstly, the image is divided into blocks, and then 2-D image blocks are constructed according to the similarity of blocks. Finding blocks that are similar to the reference one (block-matching) and stacking them together to form a 3-D

array(group), then perform collaborative filtering to the 3-D array, at last inverse transform is performed to get the denoised image. BM3D algorithm can effectively maintain the target edge information while effectively removing noise.

### 2.2 Contrast Enhancement

Due to the dynamic range of IR detector, IR images have poor contrast and fuzzy edge, especially for the ship target. Due to the influence of sea surface environment, the image seems to be covered with fog, which results in poor image contrast. The contrast stretching algorithm is used to enhance the contrast of the ship target image, and the formula is as follows [7]:

$$s = T(r) = \frac{1}{1 + (m/r)^E} \tag{2.1}$$

where, $r$ represents the brightness of the input image, and $s$ is the corresponding brightness of the output image. $m$ usually takes the average value of the whole image, and $E$ is used to control the slope of the function, and it can adjust itself according to the image stretching effect.

## 3 Area Detection of Ship Target Based on Morphological Processing

### 3.1 Sea-Sky-Line Detection

According to the characteristic that remote sea surface target always appearing near the sea-sky-line, through detecting the sea-sky-line position to determine the target potential area can greatly reduce the search scope, and eliminate the disturbance of cloud and sea clutter, then increase the probability of target detection in single frame image, and reduce the false detection rate.

Considering that Hough transform is robust to image noise and line segment discontinuities [8], it is employed in this paper for sea-sky-line detection. The basic idea of Hough transform is the duality of point to line, through detecting the local maximum of parameter space, the corresponding line can be found in the image [9]. In order to detect the longest line in the image, the maximum value of the parameter space is detected.

After the sea-sky-line is detected, the upper and lower σ region of the sea-sky-line is deemed to target potential area, which greatly reduces the computation of subsequent image segmentation and target recognition, at the same time restrains the disturbance of external area. Because the sea-sky-line detection is only to determine the target potential area, even if the location of the sea-sky-line is offset, it will not affect the target detection result.

### 3.2 Area Detection of Ship Target

Extracting the edge contour of the target effectively is the key to automatic target detection, a variety of edge detectors are commonly used for edge extraction, such as Sobel operator, Prewitt operator, Canny operator, and so on. Edge detectors usually detect the discontinuity of image brightness, thus the detected edge is usually not continuous, which may contain background region with wide range brightness, and may also include a lot of subregions of the target, but not the whole edge contour of the target. Thus, on the base of edge detection, morphological method is used to remove the edge fragment and connect the correlative regions, and finally extract the edge contour of target, thus realize the target automatic detection. The flowchart is shown in Fig. 2 [10].

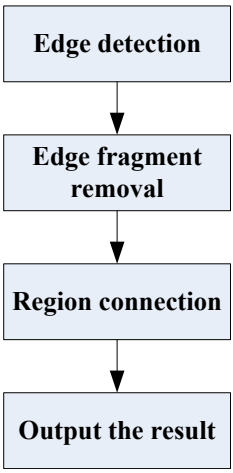



**Fig. 2.** Flowchart of automatic detection based on morphological processing

The main steps are as follows:

(a) Edge detection. Edge detection is easily disturbed by noise, and some weak edge is not effectively detected, thus, Canny detector with threshold is used for edge detection, which not only can effectively restrain noise disturbance, but also can availably detect the strong edge and weak edge.
(b) Edge fragment removal. In edge detection, the region with wide range brightness in the background is always mistaken for the target region, these regions are always small and are not connected to the real target region, which we call edge fragment. By using the connectivity of pixels, the detected edges are divided into several connected parts, and the area of each part is calculated, and the areas smaller than the certain threshold is removed.
(c) Region connection. The detected target region may be composed of a lot of unconnected small regions, and the dilation operation in morphology is employed to extend the edge. The dilation operation is as follows:

$$g = f \oplus B \tag{3.1}$$

where $f$ represents the result image after edge fragment removal, $B$ is the structure element, and $\oplus$ is the dilation operation in morphology.

On the basis of edge extension, the hole filling in morphology is adopted to fill these small regions into a whole target region, that is, the connectivity of the target regions is achieved. The pixel values of the image after region connection contain only 0 and 1, the target region is 1, and other region is 0. So far, the area detection of ship target is realized.

## 4 Length and Width Feature Extraction of Ship Target Based on Minimum Enclosing Rectangle

Minimum enclosing rectangle is the maximum range of some two-dimensional shapes which represented with two-dimensional coordinates, such as point, line, polygon, and so on. That is the rectangle determined by the maximum abscissa, the smallest abscissa, the maximum ordinate, and the minimum ordinate in the vertex of a given two-dimensional shape. The enclosing rectangle can be obtained for each ship target in every direction, in which conforms to the judgment criterion is the minimum enclosing rectangle. Typical criteria such as maximum target-background pixel number ratio, minimum circumference of enclosing rectangle, minimum area of enclosing rectangle. In this paper, the minimum area is used as the criterion, and the length and width of the minimum enclosing rectangle is the length and width of the ship target. The specific steps are as follows:

(a) According to the detected region of the ship target, finding the number and coordinates of the pixel points on the boundary of the region;
(b) Finding the coordinates of the convex hull point of the boundary pixels;
(c) Rotating the boundary of convex hull points with a small angle step in 90°, in each rotation, recording the maximum and minimum coordinate values of the enclosing rectangle boundary points in direction of coordinate system;
(d) Calculating the area of these rectangles according to the coordinate value, and the rectangle with smallest area is the minimum enclosing rectangle, and its length and width is the length and width of the ship target, thus the length-width ratio of the ship target can be obtained.

## 5 Experimental Results

A representative IR image with ship target and sea-sky background is selected to validate the proposed method. Figure 3 is the real navigating ship IR image sized 240 × 320 with obvious noise and low contrast. Before detecting the sea-sky-line, preprocessing operation which contains BM3D and contrast stretching is executed, and the pretreated result is shown in Fig. 4. Compared with the original image, in the

**Fig. 3.** Original IR ship image

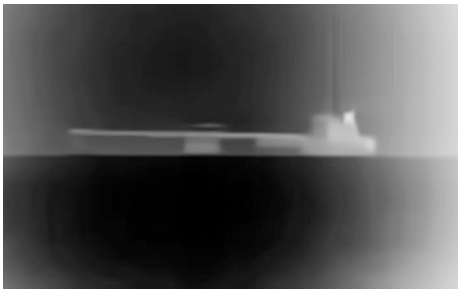


**Fig. 4.** Pretreated result

pretreated image, the noise is mostly removed, and the image contrast is enhanced obviously, which is beneficial to sea-sky-line detection.

Hough transform is used to detect the sea-sky-line of the pretreated image. The result is shown in Fig. 5.

In order to avoid the disturbance of sea clutter, the whole region above the sea-sky-line and the following n rows below the sea-sky-line ($n = 10$ in this paper) are extracted as the target potential region, and the rest regions are set to 0. After edge detection and edge fragment removal, the result is shown in Fig. 6.

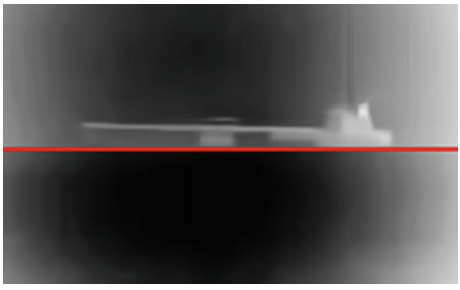


**Fig. 5.** Result of sea-sky-line detection

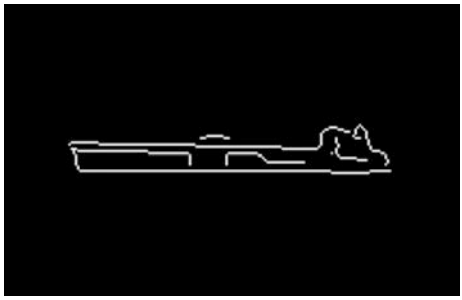


**Fig. 6.** Result of edge detection and edge fragment removal

The image in Fig. 6 does not include background region that is mistaken for the target, but the target is divided into several disconnected small regions, the dilation operation and hole filling are executed to connect these small regions, and the result is shown in Fig. 7.

**Fig. 7.** Result of region connection

The obtained minimum enclosing rectangle is shown in Fig. 8 according to the minimum area criterion. The long axis of the ship target is 153 pixels, the short axis is 26 pixels, and the length-width ratio is 5.88. Then combining distance, course and other

**Fig. 8.** Minimum enclosing rectangle of ship target

information, the size of the ship target can be obtained, which is beneficial to ship target recognition.

## 6 Conclusions

The length and width feature of ship target is often used as the initial criterion of ship type. In this paper, a length and width feature extraction method of ship target based on IR image is studied. Based on denoising, contrast stretch and sea-sky-line detection, edge detector is used for edge detection, through morphological processing such as expansion and hole filling, the whole connected region of ship target is obtained, at last, seeking the minimum enclosing rectangle according to minimum area criterion, and the length and width of the minimum enclosing rectangle is the length and width of the ship target. The experimental results show that the method is adaptable, and it can effectively extract the length and width feature of ship target in complex sea and sky background, and then combine other auxiliary information to achieve effective recognition of ship target.

## References

1. Li D, Meng X-w, Zhang H, Liu L. Research of ship feature extraction technology from SAR image. Ship Sci Technol. 2016;38(3):115–9.
2. Askarif F, Zerrb B. An automatic approach to ship detection in spaceborne synthetic aperture radar imagery: an assessment of ship detection capability using RADARSAT. Italy: SACLANT Undersea Research Centre; 2000.
3. Gao G, Kuang G-y, Li D-r. High-resolution SAR image segmentation and target's feature extraction. J Astronaut. 2006;27(2):238–44.
4. Wu F, Wang C, Zhang B, et al. Study on vessel classification in SAR imagery: a survey. Remote Sens Technol Appl. 2014;29(1):1–8.
5. Dabov Kostadin, Katkovnik Vladimir, Egiazarian Karen. Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans Image Process. 2007;16(8):2080–95.
6. Tan P, Jiang C. Analysis of several kinds of image denoising algorithm. Geomatics Spatial Inf Technol. 2014;37(7):39–42.
7. Gonzalez RC, Woods RE, Eddins SL. Digital image processing using MATLAB. Publishing House of Electronics Industry; 2006. p. 300–4.
8. Liu S-t, Zhou X-d, Wang C-g. Robust sea-sky-line detection algorithm under complicated sea-sky background. Opto-Electron Eng. 2006;33(8):5–10.
9. Junwei LV, Chenggang W, Xiaodong Z, Songtao L. A sea-sky-line detection algorithm based on fractal feature and Hough transform. J Naval Aeronaut Eng Instit. 2006;21(5):545–8.
10. Yan C, Shuhua W, Guangping W, Weili C, Jingli W, Shilei Y. Automatic detection method of sea surface target based on morphological processing. LIDAR imaging detection and target recognition 2017. In: Proceedings of SPIE vol. 10605, 106053W-1-106053W-7.

# Lunar Image Matching Based on FAST Features with Adaptive Threshold

You Zhai[1(✉)], Shuai Liu[2], Xiwei Guo[1], Peng He[1], and Zhuanghe Zhang[1]

[1] Shijiazhuang Campus, Army Engineering University, Shijiazhuang, China
youyoul952@sina.com
[2] Technical Division, 66132 Troops, Beijing, China

**Abstract.** The contrast of lunar images is low, and few features can be extracted. Therefore, lunar images can be hardly matched with high accuracy. A lunar image matching method based on features from accelerated segment test (FAST) feature and speeded-up robust features (SURFs) descriptor is presented. First, entropy of image is adopted to automatically compute threshold for extracting FAST features. Second, SURF descriptors are used to describe candidate features, and then initial matches with nearest neighborhood strategy are obtained. Third, outliers are rejected from initial matches by RANSAC-based model estimation strategy and homography constraint. Experimental results show that the proposed method can get enough image correspondences and the matching errors are less than 0.2 pixels. It indicates that the proposed method can automatically achieve high-accuracy lunar image matching and lay good foundation for subsequent lunar image stitching and fusion.

**Keywords:** Image matching · Moon image · FAST · SURF

## 1 Introduction

With the progress of rocket propulsion, sensor technology, and image acquisition, more and more high-resolution lunar images are captured, which can be used to obtain useful information about the 2D and 3D information about the surface of the moon. Image matching is the premise for getting 2D and 3D data. Image matching is the process to find the corresponding points on different images of the same spatial point, which is the basis for further image processing, such as image mosaic and image fusion.

In recent years, the state-of-the-art image matching methods are feature-based matching methods. Common features used for image matching include corner features, such as Harris, CSS, and SUSAN [1–3], as well as blob features, such as DoG [4, 5] and FH [6–8], and zone features such as MSER [9]. SIFT and SURF are two commonly used feature extraction and description methods. SIFT algorithm includes DoG detection operator and SIFT descriptor, and SURF algorithm includes FH detection operator and SURF descriptor. They can achieve scale, rotation, and illumination invariance and partly viewpoint invariance. Compared to blob features and zone features, corners are easy to detect and fast, which usually appears near the actual corner points. Corner features have strong physical meaning and are still widely used in image


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_2

matching. Especially in recent years, with the advent of fast corner detector, such as FAST [10, 11], AGAST [12], and BRISK [13], corners play more important role in image matching.

Lunar remote sensing images are different with common images. The distribution of gray levels is compact resulting in low image contrast. Therefore, when carrying image matching with the above features and usual threshold, it can hardly get enough features leading to inaccurate matching precision. One alternative method is to manually adjust the threshold for feature extraction, but this will greatly affect the automatic level of image matching. Therefore, lunar image matching based on FAST features with adaptive threshold is proposed. Contrast thresholds for image matching are automatically computed based on information entropy for matching image pair, respectively. The thresholds are used to extract FAST features. Then a dominate orientation is assigned to each feature, and SURF descriptor is computed along this dominate orientation to achieve rotation invariance. Finally, RANSAC-based homography estimation method is used to reject false matches, which can effectively improve the accuracy of the model estimation.

## 2 Adaptive Threshold Strategy

### 2.1 Overview of FAST Algorithm

FAST detects corners according to the self-dissimilarity of the candidate pixels with its surrounding pixels [10]. The FAST algorithm operates by considering a circle of sixteen pixels around the corner candidate $p$. Point $p$ is classified as a corner if there exists a set of $n$ contiguous pixels in the circle which are all brighter than the intensity of the candidate pixel $I_p + t$, or all darker than $I_p - t$, as illustrated in Fig. 1. Continuous $n$ brighter or darker points mean the dissimilarity with its surroundings of $p$ is low and then it is determined as corners, or reject. There exist different choices of $n$, such as 9, 10, 11, 12, resulting in different versions of FAST (FAST-9, FAST-10, FAST-11, and FAST-12). FAST-ER is a new version of FAST-$n$ with higher repeatability [11], but low computing efficiency. Comparison studies show FAST-9 is best in different versions of FAST including FAST-ER. FAST-9 is fast, and its

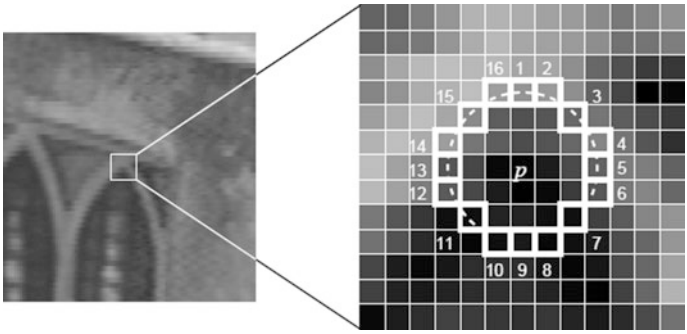



**Fig. 1.** Illustration of corner detection in an image patch

repeatability is almost the same with FAST-ER, even better in some cases. The computing efficiency of FAST-9 is more than 100% above FAST-ER. Therefore, in the following of the paper FAST-9 is chosen to detect features.

### 2.2 Analysis of FAST to Images with Different Contrast

As mentioned above, FAST detects corners with a contrast threshold $t$, which affects the adaptability of the FAST to images with different contrast. Literatures [10, 11] recommend that t is chosen between 35 and 40. In fact, the gray scale and contrast of images captured in different illumination conditions vary a lot. There is no uniform threshold t for all images. The Leuven images in Graffi standard test image set are images of the same scene under different illumination conditions, including six images, as shown in Fig. 2, representing different illumination conditions and contrast changes [14].

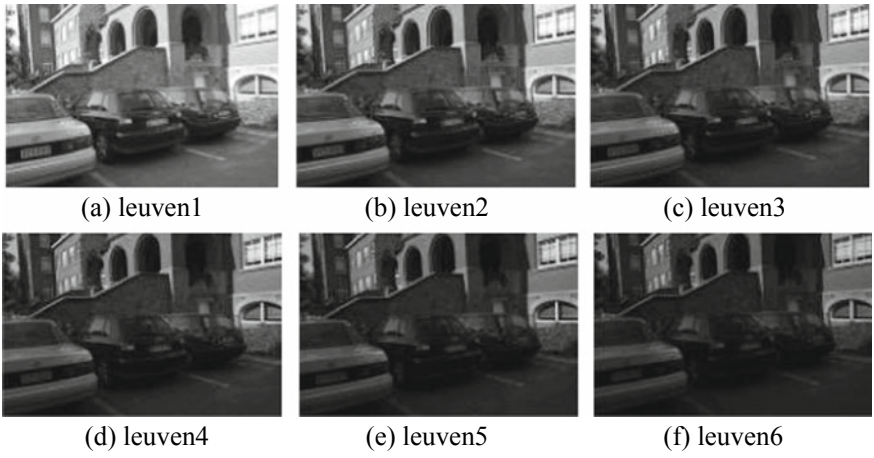
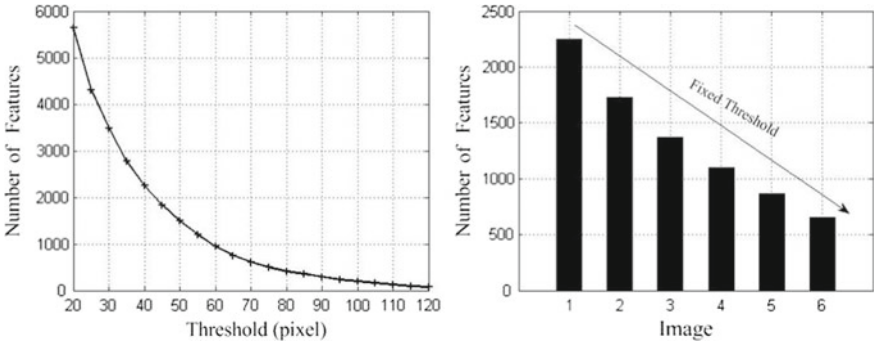



**Fig. 2.** Different images of the same scene under different illumination conditions

Two experiments are carried out to analyze the effects of threshold $t$ and image contrast change on amount of features. First, detect corners to Fig. 2a with varying threshold from 20 pixels to 120 pixels with interval 5 pixels. Second, detect corners with fixed threshold $t = 40$ to all the six images shown in Fig. 2. The resolution of the images used in the experiment is 500 * 900, and the experimental results are shown in Fig. 3.

Figure 3a shows feature detection results with different thresholds on image leuven1. From the experiment results, it is known that the number of features decreases with the increase of the threshold $t$. When $t = 40$, 2249 features are detected and when $t = 40$, only 407 features are detected. Figure 3b shows feature detection results with fixed threshold to images from leuven1 to leuven6. It can be concluded that the number of features detected also decreases with the decreasing of image contrast. For example, when $t = 40$, the number of detected features of the leuven1 is about 3.5 times that of the leuven6. Therefore, the number of feature detected is related to the image contrast




(a) feature detection results to leuven1 (b) results to leuven 1 to leuven 6

**Fig. 3.** Feature detection results

and threshold $t$. So the threshold $t$ should be selected according to the change of image contrast to ensure the algorithm can adapt to images with different contrast.

### 2.3 Adjusting Contrast Threshold Using Entropy

Different images have different scenes, gray scales, and ambiguity, and therefore, there is no universal threshold suitable for all images. The adaptive threshold-adjusting strategy proposed here is only applicable to lunar images or images similar to lunar images. The remote sensing lunar images have narrow gray scale, and edges and corners of the image are not apparent. Besides, there are obviously bright and dark areas, as shown in Fig. 5. We need to calculate the contrast threshold based on the characteristics of the lunar image. The key to compute the contrast threshold automatically is how to measure the contrast of the lunar image. Image contrast is related to the distribution of image gray scale. Generally speaking, image contrast varies with the distribution of the gray scale. When the distribution of the gray scale is narrow, the image contrast and the entropy of image are low. When the distribution of the gray scale is even, the image contrast and the entropy of image are high. Therefore, the image entropy can be used to describe the contrast of the image. In order to measure the entropy in the range [0, 1], the image entropy is normalized by its possible maximum value, as shown in Eq. (1):

$$norm\_image\_entropy = \frac{-\sum_{i=0}^{255} p_i \log 2(p_i)}{\log 2(L)} \tag{1}$$

where $p_i$ represents the proportion of pixels with gray level equating to $i$, and $i$ varies from 0 to 255 with integer value. $L$ represents the number of the gray level of the image, and it is 256. $\log2(L)$ represents the maximum entropy when the gray scale is evenly distributed to all 256 gray levels. The values of normalized entropy of leuven1

are 0.96, 0.91, 0.88, 0.84, 0.80, and 0.76, respectively. It can be inferred that the normalized entropy varies with the image contrast. The following steps are used to determine the adaptive threshold of image based on image entropy.

First, determine the reference threshold $t_0$. The reference threshold is the basis of the subsequent adaptive threshold calculation. It is the threshold when the image normalized entropy is greater than 0.9. The original FAST algorithm recommends that $t_0$ should be selected between 35 and 40 empirically.

Second, compute normalized entropy for matching image pair using Eq. (1).

Third, determine the adaptive contrast threshold *t_auto* based on Eq. (2). The exponent function is used to nonlinearly map the reference threshold $t_0$ to (0, $t_0$).

$$t\_auto = \left(2^{norm\_image\_entropy} - 1\right) \times t_0 \tag{2}$$

## 3 Feature Description and Matching

After feature detection, the classic SURF descriptors are used to describe each feature for matching. SURF descriptors are fast, have invariance to image scale, rotation, illumination, and certain viewpoint changes, and have been widely used in stereo image matching. The detailed construction procedure can be referred in [6]. Feature description and matching are as follows.

First, determine the dominate orientations with the method of SURF. The dominate orientation is the starting direction of the local neighborhood coordinate system for each feature. The dominate orientation is used as a reference direction for matching, and therefore, the rotation invariance can be realized. Feature scale represents the size of local neighborhood for computing feature descriptor. When lunar satellites scan the moon surface, the flight height changes little. So the image scales of the same region have little change, and so feature scale is not considered. Scale factors are set to 1.

Second, build SURF descriptor for each feature centered at the FAST feature, taking the dominate orientation as reference and scale as 1. Theoretically, the SURF descriptors corresponding to the same space point have a higher similarity. This is the basis for false matches rejecting.

Third, KD tree searching strategy is used to get the nearest matching point and the second nearest matching point. When the ratio of nearest matching point and the second nearest matching point is less than the given threshold, the candidate matching is considered to be the correct match. According to the literature [4], the threshold is determined to be 0.7.

Fourth, RANSAC-based strategy and homography constraint are used to eliminate mismatches. The fluctuation of the moon's surface is much smaller than the height of the lunar satellite. Therefore, we can assume that the surface of the moon is planar. The transformation induced by images of the same scene is homography. Homography is a strong constraint, and under the constraint of homography, matching points from different images are mapped one to one, as shown in Eq. (3), where that $x_r$ and $x_t$ are image points belonging to the same spatial point. If more than four pairs of matches are obtained, the homography transformation matrix $H$ can be solved.

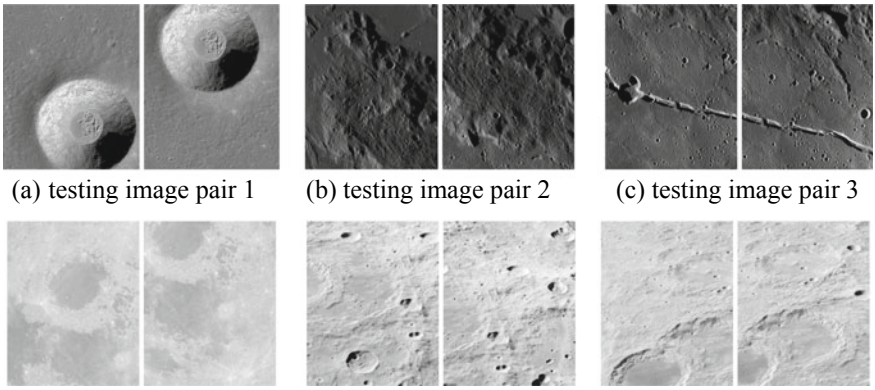
$$\boldsymbol{x}_t = \boldsymbol{H}\boldsymbol{x}_r \tag{3}$$

After the third step, large number of matches can be obtained. However, there are still many mismatches. In order to further reject false matches (outliers), RANSAC-based strategy is used to estimate homography matrix and eliminate the false matches. Two main parameters, random sampling times $N$ and threshold for determining inliers (correct matches) and outliers, are needed. The random sampling times $N$ depends on the ratio of the inliers in initial matching set. After the initial match, the ratio of the inliers is usually higher than 85%. In order to ensure the random selected 4 matches are inliers, at least 10 random sampling is needed. To improve the robustness, the number of random sampling is increased. $N$ is set to 50, and after 50 times of random sampling, the probability of error occurrence is about $10^{-17}$. The sum of the mutual Euclidean distance between the transformation point and the matching point is used as the error function as shown in Eq. (4). The threshold value is set to 0.2 pixels. When the error is less than 0.2, the matches are determined as inliers that are correct matches. Or the matches are determined as outliers.

$$\boldsymbol{e} = \|\boldsymbol{x}_t - \boldsymbol{H}\boldsymbol{x}_r\|_2 + \left\|\boldsymbol{x}_r - \boldsymbol{H}^{-1}\boldsymbol{x}_t\right\|_2 \tag{4}$$

## 4 Experiments

Experiments are carried out with actual lunar image. The images used are from the International Planetary Society [15], who collects large dataset of lunar images taken by different nations, including Chinese Chang'e mission. The testing images have certain degree of rotation, illumination and blur changes, and small viewpoint changes. The resolution of Fig. 4a–f is 1000 * 800.



(a) testing image pair 1 (b) testing image pair 2 (c) testing image pair 3

(d) testing image pair 4 (e) testing image pair 5 (f) testing image pair 6

**Fig. 4.** Testing images

The gray scale of lunar image is narrow. Therefore, the reference threshold $t_0$ is set to 25, which is lower than the recommended value. The experiment results and the auto-computed threshold are shown in Table 1. From Table 1, it can be seen that the above strategies can automatically calculate the threshold for lunar images with different contrast. There are more than 3000 matches for Fig. 4a, c, and the transformation error is less than 0.1 pixels. There are more than 700 matches for Fig. 4b, e, f, and the matching error is less than 0.15 pixels. There are far less matches for Fig. 4d, but the error after matching is still less than 0.06, which indicates that the accuracy of homography estimation is accurate.

**Table 1.** Experiment results with adaptive threshold

| Figure | Adaptive threshold | | | Fixed threshold | |
|---|---|---|---|---|---|
| | No. of matches | Trans. errors | Adaptive threshold | No. of matches | Trans. errors |
| 4a | 3008 | 0.06 | 20.4, 20.6 | 1551 | 0.04 |
| 4b | 1018 | 0.02 | 18.4, 19.1 | 281 | 0.03 |
| 4c | 3660 | 0.03 | 19.4, 20.3 | 1694 | 0.02 |
| 4d | 279 | 0.06 | 14.3, 14.0 | 9 | 0.08 |
| 4e | 791 | 0.14 | 20.3, 20.3 | 197 | 0.13 |
| 4f | 753 | 0.14 | 19.0, 19.2 | 114 | 0.12 |

For comparison, matching experiment with fixed threshold is also carried out here. The threshold is 35, which is the original algorithm recommended. The experiment results of fixed threshold are shown in the right two column of Table 1. It can be concluded that the original FAST algorithm recommends a higher contrast threshold and is not suitable for lunar image matching. The adaptive threshold-adjusting strategy is suitable for lunar image matching. The matching results of Fig. 4b, d, f with lines connecting the matching image pairs are shown in Fig. 5.

## 5 Conclusion

In order to achieve automatic lunar image matching, a matching method based on FAST features with adaptive threshold is proposed. The proposed method adopts entropy and reference threshold to automatically compute the contrast threshold for extracting FAST features, which can effectively promote the adaptability of FAST features. When matching, SURF descriptors and RANSAC-based homography-estimating strategy are used to improve the accuracy of matching. The experiment results show that the method is effective and suitable for lunar image matching. The automatic level and accuracy of matching are both improved, which lay good foundation for subsequent image mosaic and fusion. The disadvantage of this method is that the adaptability of adaptive threshold-adjusting strategy needs further improved to adapt to a wider variety of images.

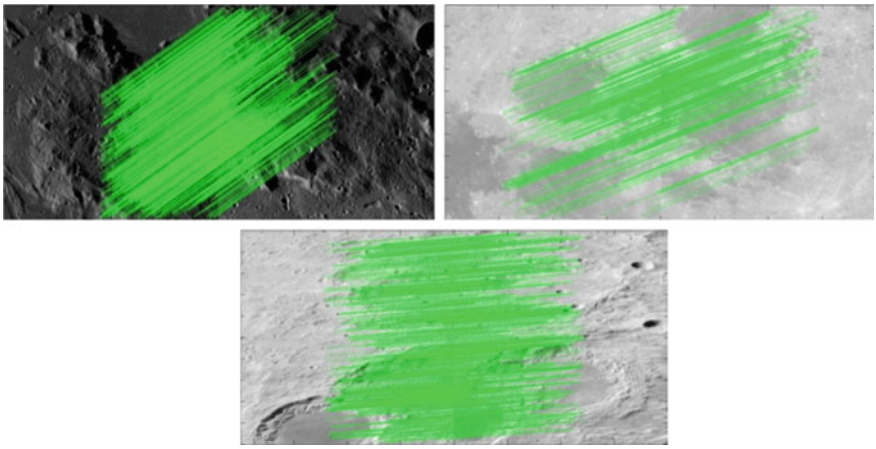


**Fig. 5.** Part of the image matching result

## References


1. Cheng C, Wang X, Li X. UAV image matching based on surf feature and Harris corner algorithm. In: Proceedings of 4th international conference on smart and sustainable city; 2017. p. 1–6.
2. He XC, Yung NHC. Curvature scale space corner detector with adaptive threshold and dynamic region of support. In: Proceedings of international conference on pattern recognition, vol. 2; 2004. p. 791–4.
3. Huiqing Z, Lin G. Image registration research based on SUSAN-SURF algorithm. In: Proceeding of the 26th Chinese control and decision conference; 2014. p. 5292–6.
4. Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vis. 2004;60(2):9l–1.
5. Ma W, Wen Z, Wu Y, et al. Remote sensing image registration with modified SIFT and enhanced feature matching. IEEE Geosci Remote Sens Lett. 2017;14(1):3–7.
6. Bay H, Ess A, Tuytelaars T, et al. Speeded up robust features. Comput Vis Image Underst. 2008;110(3):346–59.
7. Cheon SH, Eom IK, Moon YH. Fast descriptor extraction method for a SURF-based interest point. Electron Lett. 2016;52(4):274–5.
8. Wei Q, Qiao L, Cheng H, Yaohong X, et al. Strategies of improving matching accuracy about SURF. In: Proceedings of first international conference on electronics instrumentation & information systems; 2017. p. 1–6.
9. Matas J, Chum O, Urban M, et al. Robust wide baseline stereo from maximally stable extremal regions. Image Vis Comput. 2004; 22(10):761–7.
10. Rosten E, Drummond T. Machine learning for high-speed corner detection. In: Proceedings of European conference on computer vision, vol. 1; 2006. p. 430–43.
11. Rosten E, Porter R, Drummond T. Faster and better: a machine learning approach to corner detection. IEEE Trans Pattern Anal Mach Intell. 2010;32(1):105–19.
12. Mair E, Hager GD, Burschka D, et al. Adaptive and generic corner detection based on the accelerated segment test. In: Proceedings of ECCV 2010, volume 6312 of Lecture Notes in Computer Science; 2010. p. 183–96.

13. Stefan L, Chli M, Siegwart RY. BRISK: binary invariant scalable keypoints. In: Proceedings of the IEEE international conference on computer vision; Barcelona, IEEE; 2011. p. 2548–55.
14. Tuytelaars T, Mikolajczyk K. Local invariant feature detectors: a survey. Found Trends Comput Graph Vis. 2007;3(3):177–280.
15. http://www.planetary.org/.

# Image-Based Detecting the Level of Water Using Dictionary Learning

Jinqiu Pan, Yaping Fan, Heng Dong(✉), Shangang Fan, Jian Xiong, and Guan Gui

College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
heng.dong@163.com, guiguan@njupt.edu.cn

**Abstract.** This paper proposes a novel method to detect the water level of a river or reservoir. Images of the ruler which is used to measure the water level are obtained easily from a camera installed on the bank. Based on the property of the images captured by the camera, the problem of water level calculation can be transformed to the problem of classifying each image into two classes of ruler and water. As dictionary learning model has shown, its ability and efficiency in image classification problems, it is utilized in this paper to solve the problem of water level detection.

**Keywords:** Water level detection · Dictionary learning · Image classification

## 1 Introduction

As a big country of water resources, the problem of water security seriously affects the sustainable development in China. Rainy season makes it really important to prevent flooding by mastering water level information of rivers and reservoirs. However, research has found that most of the rivers and reservoirs in China still rely on manual observation of water levels.

In order to liberate the workforce, automatic water level detection has been studied for a long time. At present, most common non-artificial water level measurement methods use sensors to automatically collect analog signals, and then convert them to water level in different ways. In terms of the way it senses water levels, these water level detection methods utilizing sensors can be divided into two types as non-contact type [1, 2] and contact type [3–5].

Nevertheless, these methods with sensors usually have several disadvantages. Both non-contact and contact type sensors have a prominent problem of high cost as well as their sensitivity to the surrounding environment. What's more, most of the non-contact type sensors are not universal while contact type sensors are vulnerable because of the continuous water pressure. In case of these problems, several water level detection methods using image processing instead of water sensors are proposed [6]. In [6], an additional reference indicator needed to be installed and this will no doubt consume extra cost.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_3

As dictionary learning models have shown their great performance in many distinct tasks, including the task of image classification [7, 8], which is the main goal of water level detection, this paper proposes a water level detection algorithm using ruler and camera which are already installed in most of the rivers and reservoirs based on the model of dictionary learning. It directly calculates the level of water by the images of ruler captured from the camera, requiring no other references. For the purpose of 24 h of continuous monitoring, light is needed to be projected on the ruler during night. The method proposed in this paper is real time, 24-h continuous, inexpensive, and not affected by the environment and weather.

## 2 Algorithm of Water Level Detection

As the images obtained by the camera can be apparently divided into ruler region and water region, the main idea of the water level detection algorithm is classifying each image into two classes of ruler and water, using the method of dictionary learning. By searching the boundary of these two classes, the water level can be easily calculated. Briefly, the water level detection algorithm can be divided into three procedures in Fig. 1.

First of all, convert all the training images into a training matrix of $\mathbf{Y}$, in which each column represents a training sample and corresponds to a specific label. Then, input the training matrix into the dictionary learning model to output a learned dictionary $\mathbf{D}$. Finally, with the discriminative and compact $\mathbf{D}$, classify the images captured from the camera into ruler and water real time to get the water level. In the next part, these three procedures will be presented in detail.

### 2.1 Preprocessing of Training Images

In this section, several training images are converted into a training matrix of $\mathbf{Y}$, in which each column represents a training sample extracted from the training images. Figure 2 is an example of image captured by the camera.

Considering not all information in a training image is useful, such as the large regions of water around the ruler, regions of interest (ROI) should be chosen. In order to calculate the water level, ROI should be around the interface of the ruler and the water.

Although the ROI can be roughly classified into two classes as ruler and water, different parts of the ruler are not exactly the same and should be divided into more subclasses based on its characteristic of structure. Likewise, the part of water should be divided into more subclasses based on its characteristic of texture for further distinction.

Region in the red box is chosen as the ROI, and the ROI is classified into two classes of ruler and water by the black line. Region of ruler in the ROI is divided into $C_1$ subclasses by the blue lines and region of water in the ROI is divided into $C_2$ subclasses by the yellow lines. Totally, ROI for each training image is classified into $C = C_1 + C_2$ subclasses.

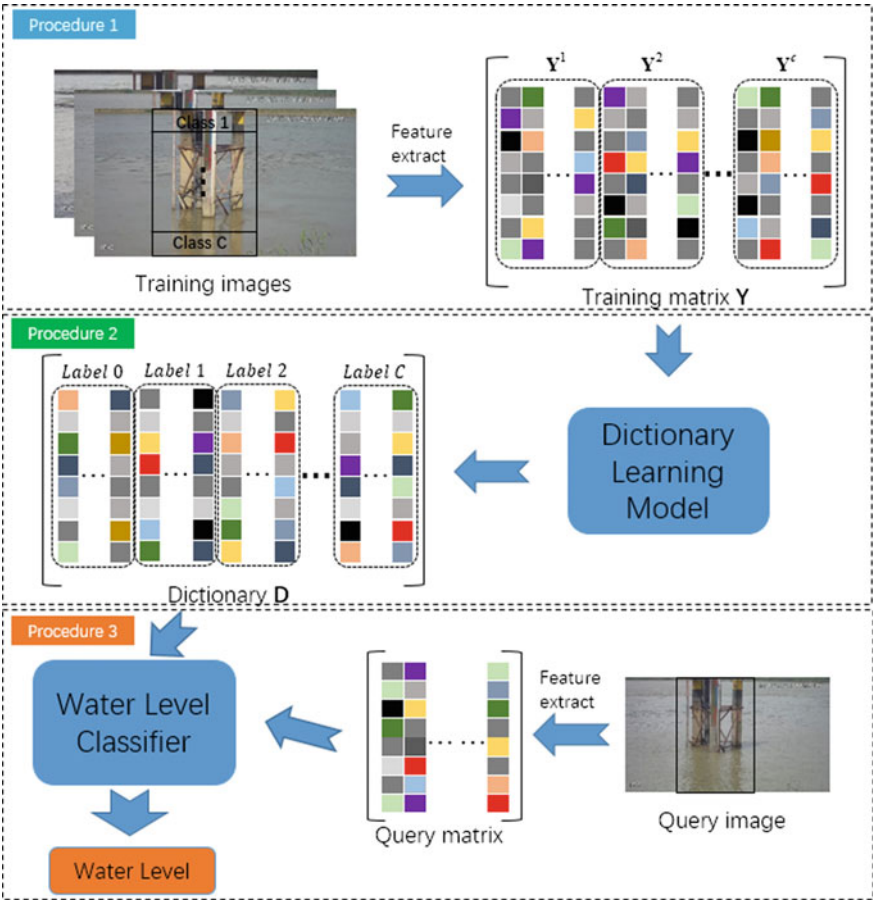



**Fig. 1.** Three procedures of the proposed water level detection algorithm.

In order to preserve the correspondence between rows, a sliding window of $m \times n$ pixel is used to acquire grayscale values in each subclass. Slide the window by step of s pixel and transform each $m \times n$ matrix into a column vector, several training samples in the form of column vectors are collected. In addition, each column vector is related to a subclass label $C_i$. Finally, combine all the training samples extracted from each subclass of each training image together, all the training images are converted into a training matrix in which each column of training sample is related to a subclass label.

## 2.2 Training for a Discriminative Dictionary

The training matrix $\mathbf{Y}$ obtained in the last procedure can be formed as $\left[\mathbf{Y}^1, \mathbf{Y}^2, \ldots, \mathbf{Y}^C\right]$, with every sub-matrix $\mathbf{Y}^c$ collected from the $c$th class of training samples. The dictionary can be received by inputting the training matrix into a dictionary learning model along with its subclass labels.

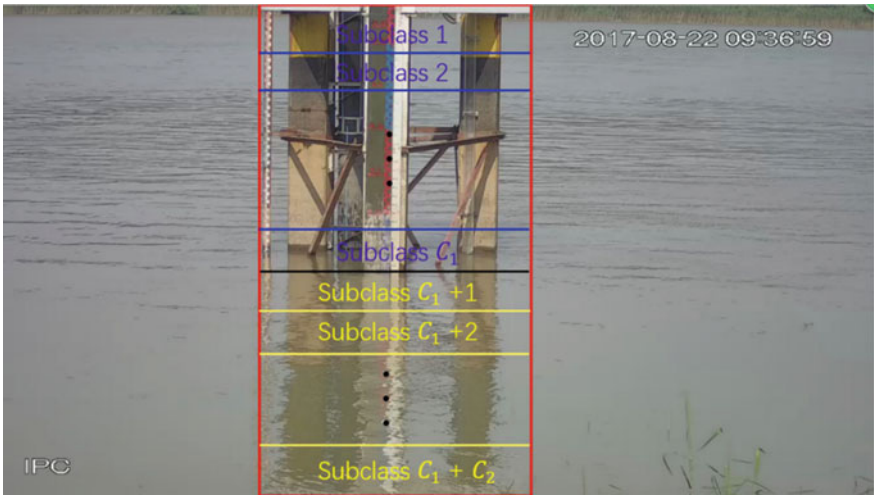



**Fig. 2.** Image captured by the camera.

The method to detect water level uses a discriminative dictionary to classify the image into ruler and water and the model chosen in this paper is a dictionary learning model with cross-label suppression and group regularization [6].

In such dictionary learning model, to obtain the property of discrimination and construction, each atom of the learned dictionary is associated with a specific label which is corresponded to a particular class. As a result, the dictionary $\mathbf{D}$ can be presented as $\left[\mathbf{D}^0, \mathbf{D}^1, \ldots, \mathbf{D}^C\right]$. Each $\mathbf{D}^c$ represents a block of the dictionary related to the $c$th class of samples. Additionally, to describe the common features between different classes such as analogical background, an extra block in the dictionary is proposed as $\mathbf{D}^0$, related to all the classes. In this case, each atom of the dictionary has its essential semantic meaning associated with specific class in the high dimension space while each column of the training matrix is related to particular class in the low dimension space. With this discriminative property of the dictionary, representation coefficients of the $c$th class samples will mainly locate on the block of $\mathbf{D}^c$ and $\mathbf{D}^0$, while the coefficients over other blocks will be suppressed to nearly 0. In such case, samples of different classes will have distinct representations and samples of the same class will have similar representations over the dictionary so that these samples can be simply classified.

While the structure of the dictionary is pre-defined, the dictionary learning model used in this paper is shown as following,

$$\min_{\mathbf{D},\mathbf{X}} \sum_{c=1}^{C} \left\{ \|\mathbf{Y}^c - \mathbf{D}\mathbf{X}^c\|_F^2 + \beta \sum_{j=1}^{N^c} \left\|\mathbf{x}_j^c\right\|_2^2 + \lambda \|\mathbf{P}^c\mathbf{X}^c\|_F^2 + \gamma \mathrm{tr}\left(\mathbf{X}^c \tilde{\mathbf{L}}^c (\mathbf{X}^c)^{\mathrm{T}}\right) \right\} \quad (1)$$

$$\text{s.t. } \|\mathbf{d}_k\|_2 = 1, \quad \forall k$$

where $c = 1, \ldots, C$ and $\mathbf{d}_k$ is an atom of $\mathbf{D}$. $\mathbf{Y}^c$ is the $c$th block of training matrix in $\mathbf{Y}$, $\mathbf{X}^c$ is the representation matrix of the training samples in $\mathbf{Y}^c$ over the dictionary $\mathbf{D}$,

$N^c$ is the number of training samples in the sub-matrix $\mathbf{Y}^c$, and $\mathbf{x}_j^c$ is the $j$th representation vector in $\mathbf{X}^c$.

The first term in formula (1) denotes the reconstruction error and the second term represents the representation regularization. $\ell_2 - norm$ is used here to accelerate the learning process and classification.

Moreover, the above model has the third term represents the cross-label suppression and the fourth term denotes the group regularization.

In the cross-label suppression term, the matrix $\mathbf{P}^c$ is an extracting matrix which extracts the representation coefficients locating at other label-related atoms, in the form of,

$$\mathbf{P}^c(m,n) = \begin{cases} 1, m = n \text{ and } m \in \mathcal{I} \backslash \left(\mathcal{I}^0 \cup \mathcal{I}^c\right) \\ 0, \text{ elseswhere} \end{cases} \tag{2}$$

where $c = 1, \ldots, C$ and $\mathbf{P}^c(m,n)$ denote the $(m,n)$th element of the matrix $\mathbf{P}^c$. $\mathcal{I}^0$ is a set of index of the shared atoms in the dictionary, $\mathcal{I}^c$ is a set of index of the atoms related to the $c$th class in the dictionary, and $\mathcal{I}$ is a set of index related to all the atoms in the dictionary.

With the extracting matrix extracting the representation coefficients locating at other label-related atoms instead of the coefficients locating at the $c$th label-related and shared atoms, the cross-label suppression term will suppress the large coefficients' appearance over other label-related atoms in the dictionary, pushing the coefficients mainly assembling at the locations which are related to the $c$th class atoms in the dictionary. Since this constraint of representation coefficients mainly gathering at the label-related atoms, the representation matrix becomes sparse representation.

The fourth term of group regularization is used to keep the label property which means that the representation coefficients of samples from the same class are kept similar. In this term, the matrix $\tilde{\mathbf{L}}^c$ represents the normalized Laplacian matrix of a graph whose vertexes are representations of training samples from the $c$th class. Besides, tr() is the trace operator to calculate the variation of the graph. As a property of graph, constrain the variation small will keep the vertexes similar which means the representations of the same class are kept similar.

### 2.3 Water Level Calculating

As mentioned above, the problem of water level detection has been transformed into a problem of image classification. Using the sliding window extracting samples which has been used on the training images in the first procedure, the query image is transformed into a matrix of query samples. The goal is to judge the labels ranging from 1 to $C$ of each query sample in the matrix of query samples. By searching the boundary of the two classes of ruler and water, in other words, the boundary of the label $C_1$ and $C_1 + 1$, the water level in the image will be detected. Finally, it will be easy to transform the water level of the image into the water level in real world.

With the query sample $\mathbf{y}$ and the dictionary $\mathbf{D}$ learned in the last procedure, the representation of $\mathbf{y}$ can be calculated by the formula (3),

$$\begin{aligned}\hat{\mathbf{x}} &= \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \beta\|\mathbf{x}\|_2^2 \\ &= \left(\mathbf{D}^{\mathrm{T}}\mathbf{D} + \beta\mathrm{I}\right)^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{y}\end{aligned} \tag{3}$$

Based on the discriminative and constructive property of the learned dictionary $\mathbf{D}$, the representation coefficients of the sample belonging to the $c$th class mainly gather at the locations which are related to the $c$th class atoms and the shared atoms over the dictionary and the coefficients over the other label-related atoms are very small. As a result, the label of the query sample $\mathbf{y}$ can be obtained by the formula (4),

$$\mathrm{label}(\mathbf{y}) = \arg\min_{c} \frac{\left\|\mathbf{y} - \sum_{k\in\mathcal{I}^0\cup\mathcal{I}^c} \mathbf{d}_k\hat{\mathbf{x}}(k)\right\|_2^2}{\sum_{k\in\mathcal{I}^0\cup\mathcal{I}^c} |\hat{\mathbf{x}}(k)|} \tag{4}$$

Judging the label of each query sample in the query matrix obtained from the image, each $m \times n$ block of the image can be classified into the class of either ruler or water. Searching the boundary of these two classes and its corresponding coordinate $l_r$ in the $\mathbf{y}$-axis is the water level in the image.

Considering 1 m of the ruler corresponds to L pixels in the image and the level of w meter on the ruler corresponds to the vertical coordinate of $l_w$ in the image, the water level in the real world can be calculated by the formula (5),

$$R = w + \frac{l_w - l_r}{L} \tag{5}$$

Thus, the water level in the real world can be calculated.

## 3 Experiment

Samples from six different scenes have been tested in our experiment. All of these samples are images of 1280 × 720 pixels. Typical examples of these six scenes are demonstrated in Fig. 3.

It can be clearly seen that different samples from different scenes have different ROIs, but all these samples have the property that their ROIs can be easily divided into ruler and water. As a result, the water level algorithm proposed in this paper should be helpful. As to different scenes, different ROIs are chosen based on the different locations of ruler in the images and different subclasses are chosen based on the different structure and texture of different rulers. Consequently, different dictionaries are learned separately and they are appropriate for different scenes.

In order to evaluate the water level detection algorithm proposed in this paper, 100 images from each sample are chosen to be tested. These images are extracted every 15 min so that 100 images are extracted from 24 h of the video. The real water level of each image is given by hand for evaluation. In the test, performance of the algorithm is

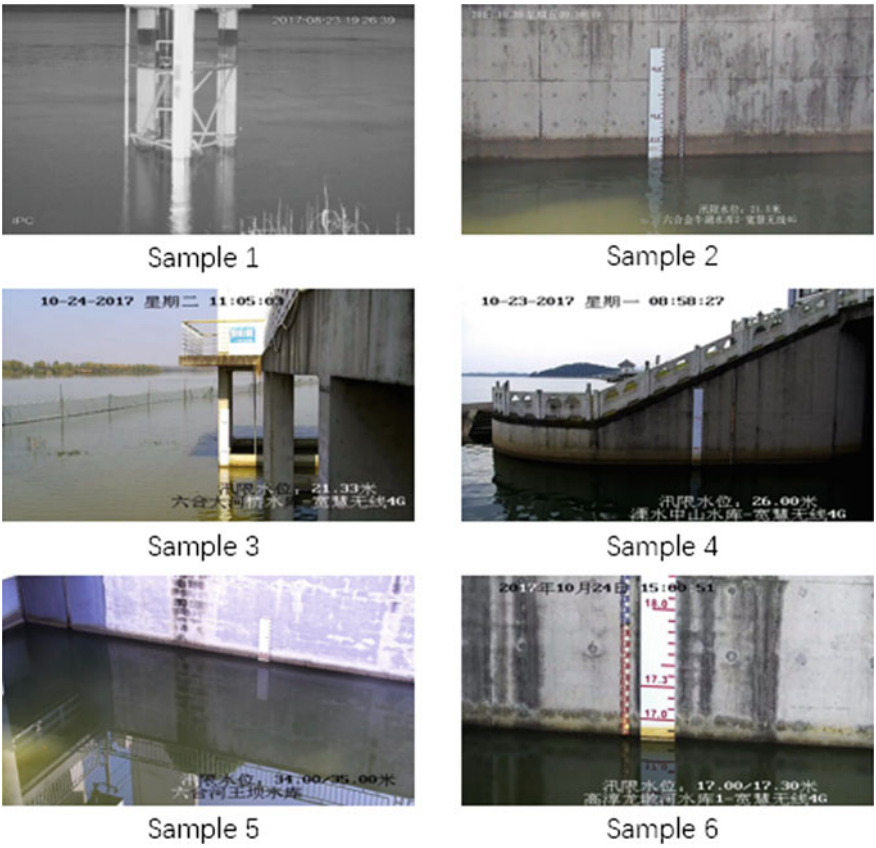



**Fig. 3.** Samples of six different scenes tested in the experiment.

evaluated by detection error in average and recognition rate. If the detection error of an image is larger than 5 cm, it is said that the algorithm made a wrong decision and recognition rate means the percentage of right decision by the algorithm.

Table 1 indicates the performance of the algorithm under six different scenes.

As can be seen from the table, all six scenes achieve a recognition rate of 100% and the average detection errors are less than 3 cm. Taking the ripples into consideration, the precision of the proposed algorithm is excellent.

## 4 Conclusion

A novel water level detection algorithm using camera monitoring the river or reservoir is proposed in this paper. This algorithm is based on the idea of dictionary learning, converting the problem of water level detection into the problem of classifying the images captured by the camera into two classes of ruler and water. As experimented,

**Table 1.** Performance of the proposed algorithm under six different scenes.

| | Total number of images | Right decision | Average detection error (cm) | Recognition rate (%) |
|---|---|---|---|---|
| Scene 1 | 100 | 100 | 1.23 | 100 |
| Scene 2 | 100 | 100 | 2.12 | 100 |
| Scene 3 | 100 | 100 | 2.05 | 100 |
| Scene 4 | 100 | 100 | 2.64 | 100 |
| Scene 5 | 100 | 100 | 1.78 | 100 |
| Scene 6 | 100 | 100 | 2.11 | 100 |

the recognition rate of the algorithm can achieve 100% and the average error is really tiny. What's more, the algorithm proposed in this paper is real time and can work 24 h a day.

## References

1. Sulistyowati R, Agus Sujono H, Khamdi Musthofa A. Design and field test equipment of river water level detection based on ultrasonic sensor and SMS gateway as flood early warning. In: AIP conference proceedings, vol. 1855; 2017. p. 50003.
2. T. Mantoro T, Istiono W. Saving water with water level detection in a smart home bathtub using ultrasonic sensor and Fuzzy logic. In: 2017 second international conference on informatics and computing (ICIC); 2017. p. 1–5.
3. Breuer E, Heitsenrether R, Fiorentino L, Hensley W, Krug W. Error analysis of water level measured from a bottom mounted ocean platform. In: OCEANS 2017—Anchorage; 2017. p. 1–8.
4. Loizou K, Koutroulis E, Zalikas D, Liontas G. A low-cost capacitive sensor for water level monitoring in large-scale storage tanks. In: 2015 IEEE international conference on industrial technology (ICIT); 2015. p. 1416–21.
5. Poovizhi S, Premalatha M, Nivetha C. Automatic water level monitoring and seat availability details in train using wireless sensor network. In: 2017 international conference on computation of power, energy information and communication (ICCPEIC); 2017. p. 321–4.
6. Kim K, Lee N-K, Han Y, Hahn H. Remote detection and monitoring of a water level using narrow band channel. J. Inf. Sci. Eng. 2010;26:71–82.
7. Skretting K, Husøy J. Texture classification using sparse frame-based representations. EURASIP J Adv Sig Process. 2006;2006(1):52561.
8. Wang X, Gu Y. Cross-label suppression: a discriminative and fast dictionary learning with group regularization. IEEE Trans Image Process. 2017;26(8):3859–73.

# Application of Improved FCM Algorithm in Brain Image Segmentation

Manzhuo Yin, Jinghuan Guo(✉), Yuankun Chen, and Yong Mu

College of Information Science and Technology,
Dalian Maritime University, Dalian City, Liaoning Province, China
{yinmanzhuo,guojinghuan}@dlmu.edu.cn

**Abstract.** Aiming at the problems of fuzzy c-means clustering (FCM) and its improved algorithms for MRI image segmentation, this paper proposes a new FCM algorithm based on neighborhood pixel correlation. The algorithm works out the influence degree of the neighborhood pixels on the central pixel by the correlation of the gray-level difference between the domain pixel and the center pixel. Then, the distance between the neighborhood pixel and the cluster center is used to control the membership of the center pixel, the improved algorithm will solve the existing influence factors of unification, ignoring the difference between pixels, resulting in inaccuracy of segmentation results. At last, this algorithm is implemented by MATLAB tool and compared with FCMS and FLICM algorithms. The feasibility of the presented algorithm and the accuracy of the segmentation result are verified by evaluating the algorithm and the experimental results according to the relevant evaluation criteria.

**Keywords:** Brain MRI image · FCM improvement algorithm · Pixel gray correlation · Dissimilarity coefficient

## 1 Introduction

In the field of medical imaging technology, MRI images, that is, magnetic resonance imaging technology is one of the most representative imaging techniques [1]. The technology has a variety of imaging processes, can be multi-angle, multi-parameter imaging, and imaging quality without bone artifacts. In addition, the study of brain MRI image segmentation, also the segmentation and extraction of the brain accurately. It has great practical significance to attack brain disease and develop the modern surgery.

Due to the fact that the current magnetic resonance imaging technology is not yet fully mature and some defects inherent in imaging equipment, it has brought great difficulties to medical image segmentation [2]. Among these, volume effect [3] is the most difficult part of the solution, and it also leads to the phenomenon of inconsistent strength [4]. Therefore, solving the volume effect is one of the important ways to improve the segmentation accuracy. There is such a difference


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_4

in the brain, that is, the composition and structure of different brain tissues are different. This difference is usually expressed in grayscale of brain MRI images. Therefore, for MRI images of the brain, the difference in gray level is usually the same as the same brain tissue and vice versa. In addition, the diversity of the individual creates different brain structures, which also affects the segmentation results of the image.

The medical images are often ambiguous because of the imaging techniques and the defects of the scanning instruments themselves. Aiming at this problem, some scholars introduced fuzzy theory into medical image processing [5] and proposed a new segmentation algorithm. Among them, the most famous is the fuzzy c-means clustering algorithm (FCM) proposed by Dunn. After the algorithm is proposed and popularized by Bezdek et al., it has been successfully applied in many fields and has become a classic clustering algorithm [6]. FCM algorithm has the characteristics of unsupervised, simple operation and fast operation, which has made a great contribution to image segmentation and has become a widely used algorithm by researchers. However, this algorithm also has the following two problems.

First, the FCM algorithm only considers a single pixel, ignoring the relationship between pixels in the segmentation process, and does not make use of the information such as space and grayscale of the pixel. Therefore, many scholars and researchers put forward an improved method based on spatial information, combined with pixel gray or spatial information in target image to accurately segment results, which is an important direction of FCM algorithm improvement [7]. Among them, the FCMS algorithm [8] introduces the distance information of neighborhood pixels in the objective function and makes the membership degree of the target pixels more accurate. However, the FCMS algorithm requires many iterations, the computational complexity is huge, and the efficiency is low. FCMS1 and FCMS2 algorithm [9] filter the image information and calculate the neighborhood information of pixels. In this way, the amount of computation is reduced, and the time required for the algorithm to run is reduced. The FLICM algorithm proposes an improved algorithm using fuzzy factors instead of determining parameters [10]. The algorithm improves the segmentation effect to some extent. Second, the essence of the FCM algorithm is the continuous iteration of the target function, which requires a large amount of computation, leads to the low efficiency of the FCM algorithm. Aiming at this problem, researchers have proposed algorithms such as EnFCM algorithm [11], FGFCM algorithm [12] and so on.

In summary, improving FCM algorithm and applying it to the brain MRI images are worthy of further study.

## 2 Existing FCM Improvement Algorithm

Ahmed proposed the FCMS algorithm. The FCMS algorithm proposes the influence factor of the neighborhood pixel to the center pixel [13], the degree of influence of each neighborhood pixel on the central pixel is considered to be the

same, this led to the edge fuzzy segmentation result, appeared the phenomenon of excessive segmentation.

Stelios Krinidis and others proposed the FLICM algorithm based on neighborhood pixels [14]. The FLICM algorithm introduces the fuzzy factor into the objective function, and the segmentation image retains the better edge details. But the algorithm is sensitive to noise.

The analysis of FCMS and FLICM algorithm shows that the improvement of the FCM algorithm from the neighborhood pixel is the key to the utilization of neighborhood pixels. Therefore, this paper will focus on the analysis of how to modify the FCM algorithm by using the neighborhood pixel.

## 3 Improvement of FCM Algorithm Based on Neighbor Pixel Correlation

### 3.1 Pixel Correlation

Pixel correlation is the degree to which pixels interact [15]. In image segmentation, the correlation of pixels is determined by the location of pixels and the grayscale information [16].

Contrast between the pixel gray-level difference by calculation, this paper puts forward dissimilarity coefficient to express the degree of interaction between the pixel dissimilarity coefficient is inversely proportional to the pixel correlation degree, the smaller the dissimilarity coefficient, the greater the correlation. In the brain MRI image, when the pixel belongs to the same site or organ, the correlation between the pixels is larger, at this time their gray value of the pixel is smaller. Therefore, the gray difference between pixels can be used as a measure of their relevance. The grayscale difference between pixels is defined as follows.

For pixel i, $\Omega_i$ represents the neighborhood pixel set of i, which exists in pixels $j \in \Omega_i$, $\partial(i)$ and $\partial(j)$ represent the gray value of the i and j, $\partial_{i,j}$ represents the gray value of the center pixel i and the neighborhood pixel j, then there is formula (1):

$$\partial_{i,j} = \partial(i) - \partial(j) \quad (j \in \Omega_i)\,. \tag{1}$$

The dissimilarity coefficient $G_{ij}$ is defined as follows:

$$G_{ij} = \frac{1}{1 + \exp(\frac{-(\partial - \mu)}{\sigma})}, \tag{2}$$

in particular, when $\partial(i) = \partial(j)$, $G_{ij} = 0$.

Where $\partial = |\partial(i) - \partial(j)|$ and $\partial(i) \neq \partial(j)$ represent the gray difference of the adjacent pixel j and the center pixel i. $\sigma$ controls the rate of change of the dissimilarity coefficient. The parameter $\mu$ is the average gray difference of neighborhood pixel j and center pixel i in neighborhood pixel set $\Omega_i$. The parameter $\mu$ calculation formula is as follows:

$$\mu = \frac{1}{|\Omega_i|} \sum_{j \in \Omega_i} \partial_{i,j}. \tag{3}$$

When $\sigma = 0$, $G_{ij}$ can write formula (4):

$$G_{ij} = \begin{cases} 1, \partial > \mu \\ 0, \partial < \mu \end{cases}, \tag{4}$$

this is an extreme case of dissimilarity coefficient, indicating that the relation between the neighborhood pixel and the center pixel is only related or not related. When $\partial < \mu$, $G_{ij} = 0$, shows that there is a correlation between pixels. When $\partial > \mu$, indicating that the pixel is not relevant.

### 3.2 Improved FCM Algorithm Based on Neighborhood Gray Correlation Difference

Through the above analysis, it is the key to calculate the degree of influence of the neighboring pixel to the center pixel in the process of improving the FCM algorithm using the neighborhood pixels. This paper proposes a method for calculating the degree of influence based on the correlation of pixel grayscale, and its specific principles are as follows:

(1) Calculate the correlation between the neighborhood pixel and the center pixel.
According to formula (2), the dissimilarity coefficient of each neighborhood pixel and center pixel is obtained by calculating the gray difference of pixel and center pixel in the domain.
(2) Bring the correlation between pixels to the objective function.

Each neighborhood pixels to the influence degree of the center pixel membership degree is decided by their relevance, through formula (5) enlarges impact on membership center pixel neighborhood pixels, and vice versa. The calculation formula of the objective function of the improved algorithm is presented as follows:

$$J_{new} = \sum_{k=1}^{c} \sum_{i=1}^{n} \mu_{ik}^{m} D_{k,i}, \tag{5}$$

$$D_{k,i} = \frac{1}{N_{\Omega_i}} \sum_{j \in \Omega_i} [d_{ki}^2 G_{ij} + d_{kj}^2 (1 - G_{ij})], \tag{6}$$

where, $N_{\Omega_i}$ is the number of neighborhood pixels, $\sum_{k=1}^{c} u_{ik}^{m} = 1$, $d_{ki}$ is the distance between the cluster center $v_k$ and the center pixel i, $d_{kj}$ is the distance between the clustering center $v_k$ and the neighborhood pixel j. Compared with other improved algorithms, the distance between the neighborhood pixels and the clustering center is added to the objective function, and the distance calculation formula of the original algorithm is modified into the formula (6).

In order to minimize the objective function, this paper uses the Lagrange operator method to calculate, and the formula is as follows:

$$F = \sum_{k=1}^{c} \sum_{i=1}^{n} u_{ik}^{m} D_{k,i} + \sum_{k=1}^{n} \lambda_i \left( \sum_{k=1}^{n} u_{ik}^{m} - 1 \right). \tag{7}$$

Make $\frac{\partial F}{\partial u_{ik}} = 0$, available:

$$u_{ik} = [\sum_{l=1}^{c} (\frac{D_{k,i}}{D_{l,i}})^{\frac{1}{m-1}}]^{-1}. \tag{8}$$

Make $\frac{\partial F}{\partial v_k} = 0$, and $\sum_{k=1}^{c} u_{ik}^m = 1$ get:

$$v_k = \frac{\sum_{i=1}^{n} u_{ik}^m x_i}{\sum_{i=1}^{n} u_{ik}^m}. \tag{9}$$

The membership degree $u_{ik}$ and the clustering center $v_k$ are calculated by the above formula.

The specific implementation of the proposed algorithm is as follows:

(1) Input image I, threshold $\varepsilon$.
(2) Traversing the image I, pixel i is the pixel in image I, $\Omega_i$ is the $3 \times 3$ neighborhood set of pixel i, pixel j satisfies $j \in \Omega_i$, $\partial(i)$ and $\partial(j)$ are their gray value. The average gray degree $\mu$ of the neighborhood set is calculated. Traversing the neighborhood of the pixel i, by calculating each neighborhood pixels and the center pixel gray-level difference, and calculate the dissimilarity coefficient between them, put them into the dissimilarity coefficient matrix U.
(3) Initialize the pixel i membership $u_{ik}$.
(4) The membership of pixel i was introduced into formula (9) and the clustering center $v_k$ was obtained.
(5) Let $u_{ik}^{old} = u_{ik}$ and calculate the new membership $u_{ik}^{new}$ according to Eq. (8).
(6) When $max\left\{\left|u_{ik}^{old} - u_{ik}^{new}\right|\right\} < \varepsilon$, the iteration ended and the image segmentation was completed. Otherwise return (4) to continue the calculation.

## 4 Experimental Results and Analysis

### 4.1 Experimental environment

This paper selects non-functional MRI brain images as an experimental image. The experimental images used in this paper are from 60 MRI images of the brain from radiology department of the fourth hospital of Jilin University. The image resolution is $512 * 512$. The parameters of this algorithm are set as follows: threshold = 0.00001, m = 1.75. The maximum number of iterations is 100. The computer used in this paper is Inter (R) Core CPU T4300, the main frequency is 2.60 GHZ, the memory is 8.0 GB, the software environment is Microsoft's Windows 7, the simulation software is MATLAB R2014a.

### 4.2 Evaluation Criteria

(1) $V_{pc}$: a good cluster of $V_{pc}$ values should be as large as possible; $V_{XB}$: for a good cluster, the $V_{XB}$ value should be slightly smaller.

$$V_{pc} = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^2, V_{XB} = \frac{\sum_{i=1}^{c} \sum_{j=1}^{n} \left[u_{ij}^2 x_j - v_i^2\right]}{nmin\left\{v_j - v_k^2\right\}}. \tag{10}$$

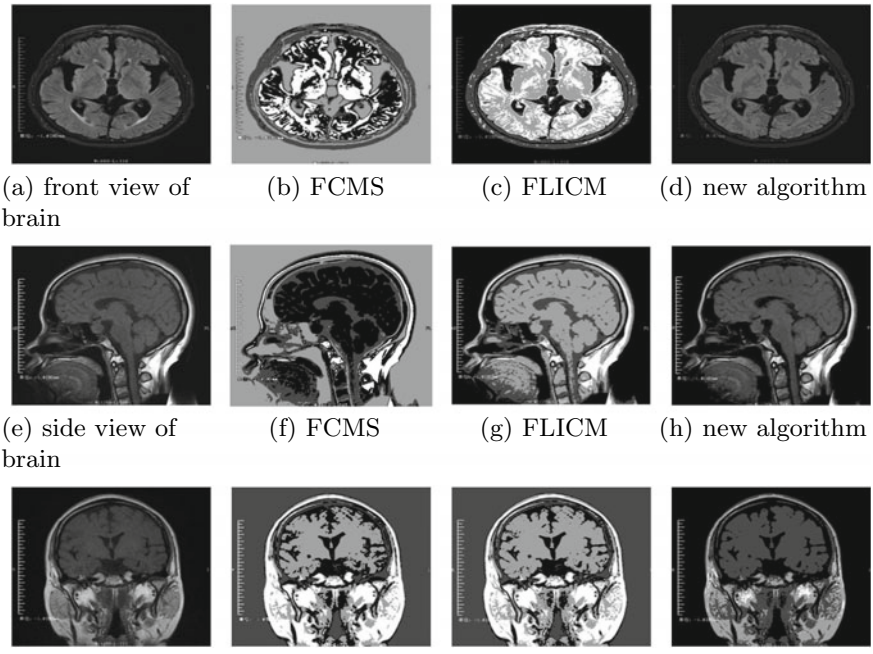
(2) SA: Classification accuracy. $A_i$ is a collection of pixels containing the i$th$ cluster, and $C_i$ is the set of pixels contained in the i$th$ cluster in the standard segmentation image;
$V_{RE}$: the reconstruction error rate, a good image segmentation algorithm should have small reconstruction error rate, where $I^{n(i)}$ is the grayscale value of the ith pixel in the reconstructed image is defined as follows:

$$SA = \sum_{i=1}^{c} \frac{card\,(A_i \cap C_i)}{\sum_{j=1}^{c} cardC_j}. \tag{11}$$

$$V_{RE} = \frac{1}{n} \sum_{i=1}^{n} \left\| I^n(i) - I(i) \right\|^2, I^n(i) = \frac{\sum_{k=1}^{C} u_{ki}^m I(i)}{\sum_{k=1}^{C} u_{ki}^m}. \tag{12}$$

### 4.3 FCMS Algorithm and FLICM Algorithm Comparison Experiments

Compared with FCM algorithm and FLICM algorithm, the experimental results are shown in Fig. 1.The experimental data from algorithm design and FCMS



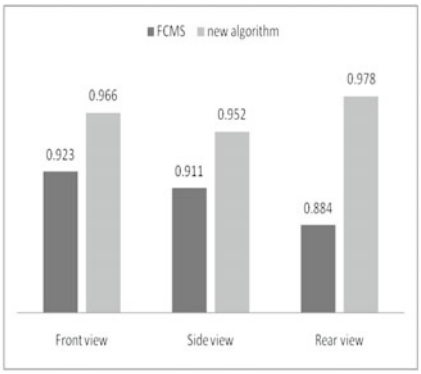
(a) front view of brain (b) FCMS (c) FLICM (d) new algorithm

(e) side view of brain (f) FCMS (g) FLICM (h) new algorithm

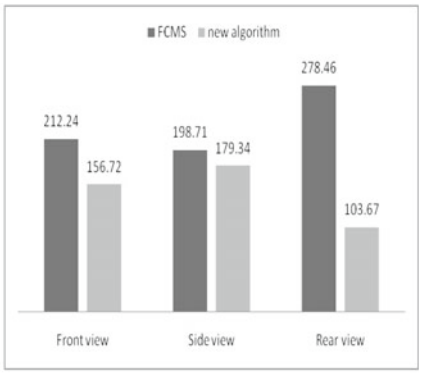
(i) rear view of brain (j) FCMS (k) FLICM (l) new algorithm

**Fig. 1.** Comparison with FCMS algorithm and FLICM algorithm experimental results

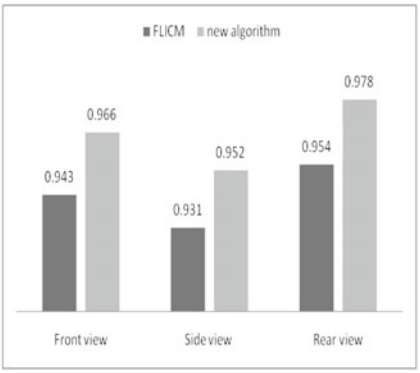
algorithm, and FLICM algorithm are shown in Table 1. The experimental data comparing the FCMS algorithm and FLICM algorithm from the segmentation results are shown in Fig. 2.



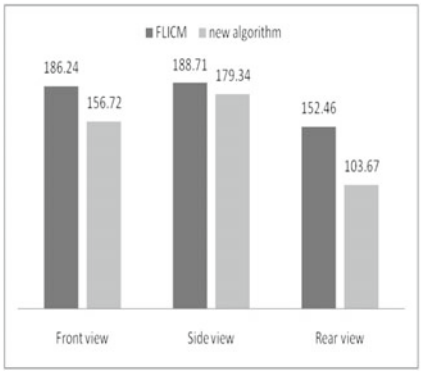

(a) Contrast with FCMS of SA




(b) Contrast with FCMS of $V_{RE}$




(c) Contrast with FLICM of SA




(d) Contrast with FLICM of $V_{RE}$

**Fig. 2.** Comparison with FCMS algorithm and FLICM algorithm experimental data

From the above diagram and table, it can be seen that the pixel membership of the algorithm and the pixel distance coefficient of clustering are better than that of FCMS, the algorithm to extract image edges more accurate and indicators are better than that of FCMS algorithm. The segmentation effect of the edge is slightly better than the FLICM algorithm, the noise processing is better than the FLICM algorithm, and the performance of this algorithm is slightly better than that of FLICM algorithm.

**Table 1.** Experimental data comparison of FCMS algorithm and FLICM algorithm

| Algorithm | $\mathbf{V}_{pc}$ | $V_{XB}$ |
|---|---|---|
| FCMS | 0.753245 | 1.926275 |
| FLICM | 0.842252 | 1.643263 |
| Algorithm of this paper | 0.867822 | 1.401027 |

## 5 Concluding Remarks

Aiming at the problem of brain MRI image segmentation, this paper proposes an improved FCM algorithm based on the neighborhood pixel gray difference correlation. The algorithm using neighborhood pixels and the center pixel gray-level difference to calculate the correlation coefficient of the two, to determine the connection degree of pixels in each neighborhood pixels and center, and put the process into the objective function, eventually making the center pixel membership degree calculation results more accurate. The experimental results show that the method can accurately segment the brain region from the MRI image, solve the problem of poor extraction of the edge of the image by the FCMS algorithm and the sensitivity of the FLICM algorithm to noise. The evaluation index obtained by related experiments is superior to both. It verifies the feasibility and effectiveness of the proposed algorithm.

**Acknowledgments.** Fundamental Research Funds for the Central Universities (No. 3132018194). Project name: Research on Ship Scheduling Method Based on Swarm Intelligence Hybrid Optimization Algorithm.

## References

1. Wang L, Chitiboi T, Meine H, Gnther M, Hahn HK. Principles and methods for automatic and semi-automatic tissue segmentation in mri data. Magn Reson Mater Phys Biol Med. 2016;29(2):95–110.
2. Makropoulos A, Gousias IS, Ledig C, Aljabar P, Serag A, Hajnal JV, Edwards AD, Counsell SJ, Rueckert D. Automatic whole brain mri segmentation of the developing neonatal brain. IEEE Trans Med Imaging. 2014;33(9):1818–31.
3. Roche A, Forbes F. Partial volume estimation in brain MRI revisited. In: International conference on medical image computing and computer-assisted intervention; 2014. p. 771–8.
4. Iglesias JE, Sabuncu MR, Leemput KV. A unified framework for cross-modality multi-atlas segmentation of brain mri. Med Image Anal. 2013;17(8):1181.
5. Namburu A, Samay SK, Edara SR. Soft fuzzy rough set-based MR brain image segmentation. Appl Soft Comput. 2016;54
6. Bezdek JC, Hall LO, Clarke LP. Review of MR image segmentation techniques using pattern recognition. Med Phys. 1993;20(4):1033.
7. Mohamed NA, Ahmed MN, Farag A. Modified fuzzy c-mean in medical image segmentation. In: Proceedings of the international conference of the IEEE engineering in medicine and biology society, 1998, vol. 3;1999. p. 1377–80.

8. Lei X, Ouyang H. Image segmentation algorithm based on improved fuzzy clustering. Cluster Computing. 2018;62(1):1–11.
9. Zhang DQ, Chen SC. A novel kernelized fuzzy c-means algorithm with application in medical image segmentation. Artif Intell Med. 2004;32(1):37–50.
10. Krinidis S, Chatzis V. A robust fuzzy local information c-means clustering algorithm. IEEE Trans Image Process. 2010;19(5):1328–37.
11. Cai W, Chen S, Zhang D. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. Pattern Recogn. 2007;40(3):825–38.
12. Zhu L, Chung FL, Wang S. Generalized fuzzy c-means clustering algorithm with improved fuzzy partitions. IEEE Trans Syst Man Cybern Part B Cybern A Publ IEEE Syst Man Cybern Soc. 2009;39(3):578–91.
13. Pei HX, Zheng ZR, Wang C, Li CN, Shao YH. D-fcm: Density based fuzzy c-means clustering algorithm with application in medical image segmentation. Procedia Comput Sci. 2017;122:407–14.
14. Zhang X, Song L, Lei P. Improvement of flicm for image segmentation. J Comput Inf Syst. 2014;10(21):9429–36.
15. KNTHJMDS MM. Weighted neighborhood pixels segmentation method for automated detection of cracks on pavement surface images. J Comput Civil Eng. 2016;30(2):04015021.
16. Ding S, Du M, Sun T, Xu X, Xue Y. An entropy-based density peaks clustering algorithm for mixed type data employing fuzzy neighborhood. Knowl-Based Syst. 2017;133.

# Generating Pedestrian Images for Person Re-identification

Zhong Zhang[1,2(✉)], Tongzhen Si[1,2], and Shuang Liu[1,2]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
{zhong.zhang8848,tjsitongzhen,shuangliu.tjnu}@gmail.com
[2] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China

**Abstract.** Person re-identification (re-ID) is mainly used to search the target pedestrian in different cameras. In this paper, we employ generative adversarial network (GAN) to expand training samples and evaluate the performance of two different label assignment strategies for the generated samples. We also investigate how the number of generated samples influences the re-ID performance. We do several experiments on the Market1501 database, and the experimental results are of essential reference value to this research field.

**Keywords:** Person re-identification · GAN · Generated samples

## 1 Introduction

Person re-identification (re-ID) is a technology which searches the target pedestrians under different cameras. Due to the potential application values, such as video surveillance and safety precaution [1,2], this technology has attracted more attention in the research field. The technology contains two fundamental problems, i.e., metric learning [3,4] and feature representation [5–7]. Recently, deep learning has demonstrated great potential for solving these two problems and obtains good performance in person re-ID [8–10]. Since training a deep model requires a large number of training samples and the number of each pedestrian images is relatively small, the trained model has poor generalization performance.

Recently, the generative adversarial network (GAN) [11] has attracted much attention and gradually becomes an important research direction. The important function of GAN is to generate similar virtual samples using training samples. The GAN is applied to generate pedestrian images for person re-ID and obtains promising performance [12–14]. Afterward, Radford et al. [15] proposed deep convolutional generative adversarial network (DCGAN) to increase the stability during the training process.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_5

In this paper, our goal is to evaluate the performance of using samples generated by DCGAN with different label assignment strategies. We hope this evaluation could provide an important reference for this research field. We employ two ways to label the generated samples. The first method is that we label the generated samples as the $(k+1)$th class, where $k$ is the total number of class in the training set. The second method is that we randomly assign each generated sample to one of the $k$ classes. We evaluate the two label assignment strategies with different generated samples on the Market1501 database [16] and obtain referenced results.

The rest of the paper includes the following sections. In Sect. 2, we present the GAN and the implementations of the two methods. In Sect. 3, we show the evaluation results. Finally, we make a summary.

## 2 Approach

We design two kinds of methods to assign labels for the generated samples in order to expand the training data. In this section, we show the principle of GAN and the process of assigning generated sample labels.

### 2.1 Generative Adversarial Network

The GAN contains two parts, i.e., a generator and a discriminator. We utilize the generator to generate samples and utilize the discriminator to determine whether the samples are true or generated samples. The flowchart of GAN is shown in Fig. 1. By training the two networks, the generator produces images that approximate the training set. The DCGAN is proposed to improve the stability during training, and its generator is shown in Fig. 2. We randomly generate a vector of 100-dim with normal distribution as the input of generator. In order to enlarge the vector, five deconvolution functions are applied with a kernel size of $4 \times 4$. The first four deconvolution functions are followed by a batch normalization and a ReLU function. The last deconvolution function is followed by a tanh function. Finally, the generator generates samples with the size of $64 \times 64 \times 3$. Some samples generated by the DCGAN are shown in Fig. 3.

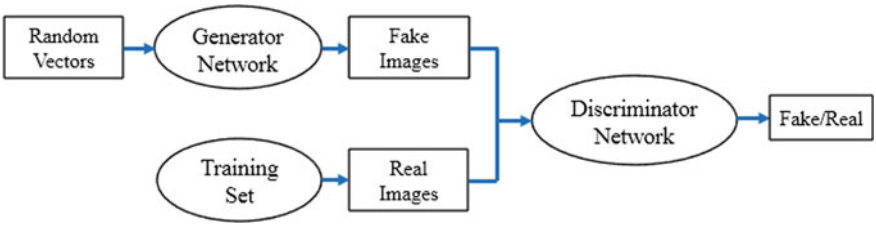



**Fig. 1.** The flow chart of GAN.

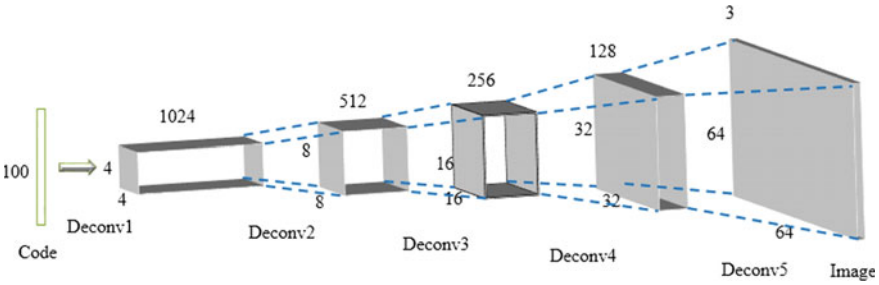



**Fig. 2.** The structure of the generator in DCGAN.

**Fig. 3.** The generated samples using the DCGAN model trained on the Market1501 database.

Both the real samples and the generated samples as the input of the discriminator. Similarly, the discriminator utilizes five convolutional layers to classify and ultimately determines whether the input samples are generated or not.

### 2.2 Label Assignment Strategies for Generated Samples

The first label assignment strategy is to assign the generated samples as the $(k+1)$th class. For example, there are 751 classes in the Market1501 training set, and therefore, the generated samples are assigned to the 752th class. The label setting is formulated as:

$$F_i = k + 1 \tag{1}$$

where $F_i$ represents the $i$th generated sample label and $k$ represents the total number of classes in the training set.

In addition, the other method is that we randomly assign each generated sample to a label. The distribution of labels is written as:

$$F_i \in [\, 1, k \,] \tag{2}$$

In the training process, we utilize the cross-entropy loss function.

$$Loss = -\sum_{i} p_i \log y_i \tag{3}$$

where $y_i$ is the prediction probability of a sample belonging to label $i$, and $p_i$ is the target probability. If the input label is $n$, then $p_n = 1$, otherwise $p_i = 0$.

## 3 Experiments

In this section, we introduce the experimental setup and the network structure. We implement the experiments on the Market1501 database [16], which effectively demonstrates the impact of generated samples.

### 3.1 Experimental Setup

We take use of the ResNet-50 [17] which is trained on the ImageNet [18] as the pretraining network. The network structure of our training model is shown in Fig. 4. In the training process, we resize all pedestrian images into 288 × 144, and then, each pedestrian image is randomly cropped to 256 × 128 as the input. We add the fully connected layer to classify pedestrian images and take $f \in \mathbb{R}^{2048\times 1}$ as pedestrian features. We set the maximum value of epochs to be 50 and the momentum to be 0.9. The weight decay is 0.0005. In addition, we assign the learning rate of the fully connected layer to be 0.1 and other layers to be 0.01. We set batch size to be 32 and utilize stochastic gradient descent (SGD) to optimize the parameters. We generate different numbers of generated samples to evaluate the impact on the classification, and the results are shown in Sect. 3.2. In the test phase, we extract the features of pedestrian images and then flip pedestrian images horizontally to extract the features. The two features are added as the final feature for each pedestrian image.

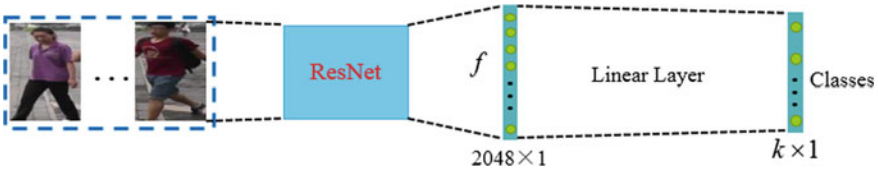



**Fig. 4.** Training model structure. We take $f$ as the features of pedestrian images and $k$ is the total number of classes in the training set.

### 3.2 The Results of Two Label Assignments

Market1501 [16] is a large-scale database containing a total of 32,668 pedestrian samples of 1501 identities. This database was obtained by a total of six cameras,

including five HD cameras and one low-resolution camera. According to the specific allocation of the database, it contains 12,936 pedestrian samples for training and 19,732 pedestrian samples for testing. Here, we utilize the training samples to train the DCGAN. In addition, the query setting contains 3368 pedestrian samples, and the number of pedestrian categories is the same as the test setting. All pedestrian samples in the database are cropped by the deformable part-based model (DPM) [19]. Hence, these pedestrian samples are close to real-life setting.

The experimental results of the first label assignment strategy are shown in Table 1. We can see that the best results are achieved when we utilize 4000 samples. The accuracy decreases when utilizing too many or too few generated samples. The experimental results of the other label assignment strategies are listed in Table 2. From the results, we can see that the results are better than others when the number of generated samples is 500. We obtain better results when the number of generated samples is small. This is because the generated images may be interfered by the incorrect labels.

**Table 1.** The rank-$i$ precious (%) and mAP precious (%) of assigning the generated samples as the ($k$+1)th class on the Market1501 database.

| No. | Rank = 1 | Rank = 5 | Rank = 10 | mAP |
|---|---|---|---|---|
| 500 | 76.75 | 90.17 | 93.37 | 55.18 |
| 4000 | **81.26** | **91.45** | **94.95** | **61.05** |
| 8000 | 78.27 | 91.03 | 94.09 | 58.13 |
| 12000 | 73.55 | 87.11 | 91.63 | 52.71 |
| 16,000 | 77.58 | 88.95 | 92.25 | 57.44 |
| 24,000 | 74.61 | 87.97 | 91.69 | 53.49 |
| 36,000 | 73.37 | 87.59 | 91.39 | 52.91 |

**Table 2.** The rank-$i$ precious (%) and mAP precious (%) of assigning each generated sample to a random label on the Market1501 database.

| No. | Rank = 1 | Rank = 5 | Rank = 10 | mAP |
|---|---|---|---|---|
| 500 | **81.26** | **91.45** | **94.95** | **61.05** |
| 4000 | 77.49 | 90.59 | 93.61 | 57.24 |
| 8000 | 76.54 | 89.34 | 92.70 | 56.06 |
| 12,000 | 78.41 | 89.49 | 92.73 | 58.67 |
| 16,000 | 74.61 | 87.98 | 91.69 | 53.49 |
| 24,000 | 72.06 | 85.48 | 89.76 | 50.62 |
| 36,000 | 73.51 | 87.38 | 90.62 | 52.89 |

## 4 Conclusion

In this paper, we have utilized GAN-generated images to expand training data, so that our trained model has better generalization performance and can provide a good reference value for others. We have verified the impact of generated samples on the Market1501 database. The experiment results have given us an important inspiration that we can take the appropriate number of generated samples to optimize our training model. In the future, we will study other ways to utilize the generated samples to enhance training data.

**Acknowledgments.** This work was supported by National Natural Science Foundation of China under Grant No. 61501327 and No. 61711530240, Natural Science Foundation of Tianjin under Grant No. 17JCZDJC30600 and No. 15JCQNJC01700, the Fund of Tianjin Normal University under Grant No.135202RC1703, the Open Projects Program of National Laboratory of Pattern Recognition under Grant No. 201700001 and No. 201800002, the China Scholarship Council No. 201708120039 and No. 201708120040, and the Tianjin Higher Education Creative Team Funds Program.

## References

1. Zhang Z, Wang C, Xiao B, Zhou W, Liu S, Shi C. Cross-view action recognition via a continuous virtual path. In: IEEE conference on computer vision and pattern recognition. Portland; 2013. p. 2690–7.
2. Zhang Z, Wang C, Xiao B, Zhou W, Liu S. Action recognition using context-constrained linear coding. IEEE Sig Process Lett. 2012;19(7):439–42.
3. Liao S, Hu Y, Zhu X, Li ZS. Person re-identification by local maximal occurrence representation and metric learning. In: IEEE conference on computer vision and pattern recognition. Boston; 2015. p. 2197–206.
4. Koestinger M, Hirzer M, Wohlhart P, Peter M, Horst B. Large scale metric learning from equivalence constraints. In: IEEE conference on computer vision and pattern recognition. Providence; 2012. p. 2288–95.
5. Bazzani L, Cristani M, Murino V. Symmetry-driven accumulation of local features for human characterization and re-identification. Comput Vis Image Underst. 2013;117(2):130–44.
6. Ma B, Su Y, Jurie F. Local descriptors encoded by fisher vectors for person re-identification. In: European conference on computer vision. Firenze ; 2012. p. 413–22.
7. Zhang Z, Wang C, Xiao B, Zhou W, Liu S. Attribute regularization based human action recognition. IEEE Trans Inf Forensics Secur. 2013;8(10):1600–9.
8. Zhang Z, Si T. Learning deep features from body and parts for person re-identification in camera networks. EURASIP J Wirel Commun Network. 2018;52.
9. Zheng Z, Zheng L, Yang Y. A discriminatively learned cnn embedding for person re-identification. ACM Trans Multimedia Comput Commun Appl. 2017;14(1):13.
10. Sun Y, Zheng L, Yang Y, Tian Q, Wang S. Beyond part models: person retrieval with refined part pooling; 2017. arXiv preprint arXiv:1711.09349.
11. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Advances in neural information processing systems. Montreal; 2014. p. 2672–80.

12. Zheng Z, Zheng L, Yang Y. Unlabeled samples generated by Gan improve the person re-identification baseline in vitro; 2017. arXiv preprint arXiv:1701.07717.
13. Zhong Z, Zheng L, Zheng Z, Li S, Yang Y. Camera style adaptation for person re-identification. In: IEEE conference on computer vision and pattern recognition; 2018.
14. Deng W, Zheng L, Kang G, Yang Y, Ye Q, Jiao J. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification; 2017. arXiv:1711.07027.
15. Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks; 2015. arXiv preprint arXiv:1511.06434.
16. Zheng L, Shen L, Tian L, Wang S, Wang J, Tian Q. Scalable person re-identification: a benchmark. In: IEEE international conference on computer vision. Chile; 2015. p. 1116–24.
17. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition. Las Vegas; 2016. p. 770–8.
18. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M. Imagenet large scale visual recognition challenge. Int J Comput Vis. 2015;115(3):211–52.
19. Felzenszwalb P, Girshick R, McAllester D, Ramanan D. Object detection with discriminatively trained part-based models. IEEE Trans Pattern Anal Mach Intell. 2010;32(9):1627–45.

# The Image Preprocessing and Check of Amount for VAT Invoices

Yue Yin, Yu Wang, Ying Jiang, Shangang Fan(✉), Jian Xiong, and Guan Gui

College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 21003, China
{ql40l030l, fansg, guiguan}@njupt.edu.cn

**Abstract.** With the continuous development of the social economy, the problem of low efficiency of invoice reimbursement has received more and more attention from companies, universities, and governments in China. In this paper, based on the recognition of invoices by OCR, we use Hough transform to preprocess the scanned image of invoices and creatively introduce the idea of checking the amount of money. We proofread the uppercase and lowercase amounts in the OCR recognition results. Using this method, the accuracy rate of OCR recognition increased from 95 to 99%, which greatly reduced the employees' reimbursement time.

**Keywords:** Image processing · Hough transform · Check amount · VAT invoice

## 1 Introduction

With the rapid popularization of network technologies, especially Internet technologies, various network applications are developing rapidly in China, such as online payment, e-commerce, and network office. It greatly facilitates the work of corporate employees. However, the problem of low efficiency of invoice reimbursement remains to be solved. Social economy is developing continuously, and more and more attention has been paid to the work of smart invoice reimbursement. Currently, the reimbursement of invoices basically depends on financial officers' business capabilities. The efficiency of financial reimbursement personnel directly affects the length of employees' time spent on reimbursement. Invoice reimbursement has become a stumbling block to the development of enterprises. Invoice reimbursement wastes a lot of manpower and material resources. In order to reduce costs, we try to use the Internet and artificial intelligence technologies for smart invoice reimbursement. At present, some related products have appeared, but the accuracy of invoice recognition is not high enough which results they cannot be promoted.

Nowadays, some experts and scholars have made their own contributions in related fields. Yu et al. designed a general commercial machine-to-invoice identification system [1], which has a recognition accuracy of 90.36% for randomly selected invoices. Liu et al. made a certain study on the identification algorithm of motor vehicle sales unified invoices and designed three simple feature extraction methods to complete the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_6

digital identification [2]. Experts are interested in check recognition [3–5]. Sneha et al. propose a technique of text word recognition based on template matching technique using correlation coefficient [6]. They developed a database of 61 words, combination of which can represent any legal amount written in words in Indian bank check, which achieved 76.4% word recognition accuracy. Liu et al. present a complete courtesy amount recognition system for Chinese bank checks [7]. The read rate of the system is around 82% when the substitution rate is set to 1%, which corresponds to that of a human operator. Jayadevan et al. propose a lexicon driven segmentation-recognition scheme for the recognition of legal amount words from Indian bank checks written in English [8]. They use a database consisting of 5400 words, collected from 50 writers to test the system and an accuracy of 97.04% is observed. In addition, OCR has many other applications, including enhancing web security, filtering Chinese image spa and recuperating video [9–20]. OCR is also used to recognize car license plate [21, 22].

In this paper, we focus on image preprocessing and the verification of the amount of each VAT invoice to improve the accuracy of invoice reimbursement. The rest of this paper is organized as follows. Section 2 introduces the use of Hough transform to preprocess images. Section 3 gives the method of identifying uppercase and lowercase numbers and explains the methods for checking each formula. Section 4 summarizes the whole paper (Fig. 1).


**Fig. 1.** A typical image of VAT invoice

## 2 Hough Transform

In general, there will be a certain angle deviation when the invoice is placed in the scanner, resulting in the entire image being tilted, as shown in Fig. 2. Observing the characteristics of the invoice, we found that there is a rectangular box in the image that is parallel to the edge of the invoice. Therefore, we determined that the image was tilt corrected with the Hough transform.


**Fig. 2.** Tilted image after scanning

In the existing tilt correction algorithm, the Hough transform is a very mature algorithm used. The Hough transform is one of the basic methods for identifying geometric shapes from images in image processing. It is widely used and there are many improved algorithms. It is mainly used to separate geometric shapes (e.g., lines, circles) from an image that have a certain characteristic. The basic idea is to use the duality of points and lines to transform the linear detection problem of image space into a local maximal value search problem of parameter space ($\rho$, $\theta$). The coordinate transformation formula is:

$$\rho = x \cdot \cos\theta + y \cdot \sin\theta \tag{1}$$

where $\rho$ is the vertical distance from the straight line to the origin, $\theta$ is the angle from the $x$-axis to the vertical line, $x$ and $y$ are the rectangular coordinates of the pixel.

The concrete realization process of the Hough transform is as follows.

In the first step, a discrete parameter space is established between the appropriate maximum and minimum values of $\rho$, $\theta$.

In the second step, the parameter space is quantized into $m \times n$, $n$ is the number of equal parts of $\theta$ and the accumulator matrix is set.

In the third step, an accumulator $A(i, j)$ is allocated to each cell of the parameter space, and the initial value is set to zero.

In the fourth step, take the point in the Cartesian coordinate system into the coordinate transformation formula, and calculate $\rho$ with the quantized $\theta$ value.

In the fifth step, find the corresponding unit of $\rho$, $\theta$ in the parameter space, and add the accumulator of the unit. 1, that is, $A(i, j) = A(i, j) + 1$.

In the sixth step, when all the points in the Cartesian coordinate system have gone through the third step and the fourth step, detecting the parameter space for each accumulator. The value, $\rho$, $\theta$ corresponding to the maximum unit is the parameter of the linear equation in the Cartesian coordinate system (Fig. 3).


**Fig. 3.** Geometrically corrected image

## 3 Recognition and Verification

The method of character recognition is basically divided into three categories: statistics, logical judgment, and syntax. Common methods include template matching and geometric feature extraction.

① **template matching method** Matches the input text with the given standard texts (templates) of each category, calculates the degree of similarity between the input text and each template, and takes the category with the highest degree of similarity as the recognition result. The disadvantage of this method is that as the number of identified categories increases, so does the number of standard text templates. This aspect will increase the storage capacity of the machine, on the other hand, will also reduce the accuracy of the recognition, so this method is suitable for identifying fixed font print text. This method has the advantage of using the entire text for similarity calculations, so it has a strong ability to adapt to text defects and edge noise.

② **geometric feature extraction method** Extract some geometric features of the text, such as the endpoints of the text, the bifurcation points, the concave and convex sections, and the line segments, closed loops, etc. in each direction of the horizontal, vertical, and oblique directions, and perform logical combination judgment based on the positions and mutual relations of these features. Recognition result. This type of recognition is also applicable to texts that have been greatly deformed as handwritten characters because of the use of structural information.

Since the text in the invoice is printed, we chose to use a template matching method that is more suitable for printing text.

We use OCR software to identify information which is useful to invoices reimbursement, including invoice codes, invoice numbers, billing dates, buyers, sellers, and prices of goods. OCR is an Optical Character Recognition technology, which refers to an electronic device (such as a scanner or a digital camera) that checks characters printed on paper, determines its shape by detecting dark and light patterns, and then translates the shape into a character recognition method. The process of computer text;


**Fig. 4.** Invoice image is entered in OCR

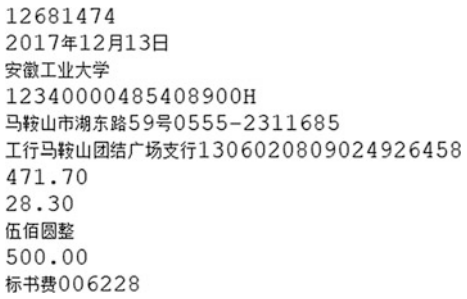



**Fig. 5.** OCR outputs recognition result

that is, the process of scanning the text data, analyzing and processing the image file, and obtaining the text and layout information (Figs. 4 and 5).

If the amount verification is not carried out, the financial reimbursement personnel need to check the invoice one by one to prevent the reimbursement amount obtained by the employee from being wrong. After the OCR outputs the uppercase and lowercase numbers, we use the programming language to convert uppercase numbers to lower case and compare the converted lowercase numbers with the recognized lowercase numbers. If the two are equal, then it is determined that the invoice is successfully identified.

The idea of programming is as follows: Define different arrays to store uppercase numbers and multiples of keywords, multiple traverse multiple times to select the multiple keywords in the input characters and process them, and then convert uppercase numbers in front of the multiple keywords to lowercase numbers. Compare the converted lowercase numbers with the recognized lowercase numbers.

When the amount check is not used, the accuracy rate of the OCR recognition invoice text is 95%. The financial reimbursement personnel need to check whether the invoice identification result is correct. Through the verification of the amount of money, the accuracy of the OCR recognition of the invoice text can reach 99%, and the financial reimbursement personnel can only manually reimburse the failed verification invoice (Fig. 6).

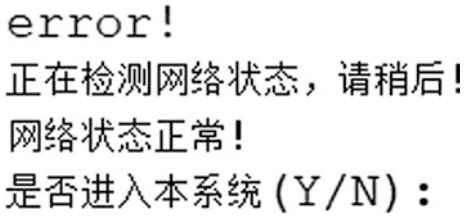



**Fig. 6.** OCR outputs the result which is recognized wrong after verification of amount

## 4 Conclusion

We preprocessed the input OCR image, and the geometrically corrected image input OCR can be correctly identified. When the OCR directly outputs the amount verification result without verifying the amount, the recognition accuracy rate is 95%. Financial reimbursement personnel need to check whether each invoice is identified correctly, otherwise, the employee may get money not equal to the invoice amount. OCR can not significantly reduce personnel work. Through our verification of the amount of money, it is possible to filter out invoices with incorrect OCR identification and ensure that the accuracy of invoice identification can reach 99%. Financial reimbursement personnel only use manual reimbursement of selected invoices, which can significantly reduce the workload and time spent on financial and general staff.

## References

1. Yu F, You-guo PI. Research and implementation on common machine-printed commercial invoice recognition system. Inf Technol. 2013;37(6):36–40.
2. He W. Studying vehicle sales invoice recognition algorithm. Guangzhou: South China University of Technology; 2010.
3. Liu D, Chen Y. A courtesy amount recognition system for Chinese bank checks. In: 2012 international conference on frontiers in handwriting recognition; 2012. p. 710–5.
4. Gorski N, Anisimov V, Augustin E. A2iA check reader: a family of bank check recognition systems. In: Proceedings of the fifth international conference on Bangalore; 1999. p. 523–6.
5. Singh S, Kariveda T, Gupta JD. Handwritten words recognition for legal amounts of bank cheques in English script. In: 2015 eighth international conference on advances in pattern recognition (ICAPR); 2015. p. 1–5.
6. Sneha S, Kariveda T, Gupta JD, Bhattacharya K. Handwritten words recognition for legal amounts of bank cheques in English script. In: 2015 Eighth international conference on advances in pattern recognition (ICAPR); 2015. p. 1–5.
7. Liu D, Chen Y. A courtesy amount recognition system for Chinese bank checks. In: 2012 international conference on frontiers in handwriting recognition; 2010. p. 710–5.
8. Jayadevan R, Pal U, Kimura F. Recognition of words from legal amounts of Indian Bank Cheques. In: 2010 12th international conference on frontiers in handwriting recognition; 2010. p. 166–71.
9. Kaur R, Pooja. A non OCR approach for math CAPTCHA design based on Boolean algebra using digital gates to enhance web security. In: 2016 international conference on wireless communications, signal processing and networking (WiSPNET); 2016. p. 862–6.
10. Wankhede PA, Mohod SW. A different image content-based retrievals using OCR techniques. In: 2017 international conference of electronics, communication and aerospace technology (ICECA); 2017. p. 155–61.
11. Alghamdi MA, Alkhazi IS, Teahan WJ. Arabic OCR evaluation tool. In: 2016 7th international conference on computer science and information technology (CSIT); 2016. p. 1–6.
12. Ekram MAU, Chaudhary A, Yadav A, Khanal J, Asian S. Book organization checking algorithm using image segmentation and OCR. In: 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS); 2017. p. 196–9.

13. Ganai AF, Lone FR. Character segmentation for Nastaleeq URDU OCR: a review. In: 2016 international conference on electrical, electronics, and optimization techniques (ICEEOT); 2016, p. 1489–93.
14. Rawls S, Cao H, Sabir E, Natarajan P. Combining deep learning and language modeling for segmentation-free OCR from raw pixels. In: 2017 1st international workshop on Arabic script analysis and recognition (ASAR); 2017. p. 119–23.
15. Jain V, Dubey A, Gupta A, Sharma S. Comparative analysis of machine learning algorithms in OCR. In: 2016 3rd international conference on computing for sustainable global development (INDIACom); 2016. p. 1089–92.
16. Xu B, Li R, Liu Y, Yan H, Li S, Zhang H. Filtering Chinese image spam using pseudo-OCR. Chin J Electron. 2015;24(1):134–9.
17. Budig B, van Dijk TC, Kirchner F. Glyph miner: a system for efficiently extracting glyphs from early prints in the context of OCR. In: 2016 IEEE/ACM joint conference on digital libraries (JCDL); 2016. p. 31–4.
18. Chiron G, Doucet A, Coustaty M, Visani M, Moreux JP. Impact of OCR errors on the use of digital libraries: towards a better access to information. In: 2017 ACM/IEEE joint conference on digital libraries (JCDL); 2017. p. 1–4.
19. Kakani BV, Gandhi D, Jani S. Improved OCR based automatic vehicle number plate recognition using features trained neural network. In: 2017 8th international conference on computing, communication and networking technologies (ICCCNT); 2017. p. 1–6.
20. Pranali B, Anil W, Kokhale S. Inhalt based video recuperation system using OCR and ASR technologies. In: 2015 international conference on computational intelligence and communication networks (CICN); 2015. p. 382–6.
21. Xu S, Smith D. Retrieving and combining repeated passages to improve OCR. In: 2017 ACM/IEEE joint conference on digital libraries (JCDL); 2017. p. 1–4.
22. Omran SS, Jarallah JA. Iraqi car license plate recognition using OCR. In: 2017 annual conference on new trends in information & communications technology applications (NTICT); 2017. p. 298–303.

# Detection of White Gaussian Noise and Its Degree in Image Processing Using Generative Adversarial Nets

Wentao Hua, Jian Xiong(✉), Jie Yang, and Guan Gui

College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
l3773573962@163.com, {jxiong,guiguan}@njupt.edu.cn

**Abstract.** Since the theory of generative adversarial nets (GANs) put forward in 2014, various applications based on GANs have been developed. Most of the applications focused on generator network ($G$) of GANs to solve the daily challenges. However, rare of them had been aware of the great value of the discriminator network ($D$). In this paper, we propose a new method of detecting white Gaussian noise and its degree by the discriminator of generative adversarial nets. The results of our experiments show the feasibility of detecting white Gaussian noise (WGN) and evaluating its degree through generative adversarial nets.

**Keywords:** Generative adversarial nets · Gaussian white noise detection · Deep learning

## 1 Introduction

With the development of digital multimedia, digital image has gradually become an important form of information carrier. In the process of obtaining, recording, converting, transmitting, and amplifying, the digital image signal is generally polluted by white Gaussian noise (WGN). The performance of most image processing algorithms can be greatly affected by WGN. Therefore, numerous researchers have been working on detecting and removing the WGN to reduce the affection. These researchers concentrate on eliminating the WGN through spatial domain filters which include local filters [1], non-local mean filter [2], and 3$D$ filtering method [3]. There is no doubt that these image denoising algorithms will play a more important role in image processing.

However, these denoising algorithms are designed for the images polluted by WGN, without considering that the processed image may be noise-free. Therefore, it is very meaningful to design an algorithm that can automatically detect white Gauss noise. In recent years, with the development of neural networks, the deep learning method has been widely used in image processing field like image compression, image recognition, image denoising, and other more. In [4], the authors apply convolutional neural network (CNN) [5] to Gaussian noise detection and its level. However, the CNN-based algorithms all belong to supervised learning algorithm. The supervised learning characteristic of CNN requires labeling all the images before training the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_7

neural network. On the one hand, it will take researchers a lot of time to prepare training data before training. On the other hand, labeling all the data is usually difficult to realize in the process of practical experiment.

In this paper, a Gaussian noise detection algorithm for image processing is proposed based on GANs. The main idea of the algorithm is fitting the distribution of real images by adversarial learning, and evaluating the differences between the WGN distorted images and the real distribution by using discriminator. When the model is trained completed, we use the trained discriminator to detect white Gaussian noise and evaluate its distortion degree.

Compared with traditional machine learning detection algorithms, we make the following contributions:

- An unsupervised detection algorithm is proposed based on GANs, which can be trained with unlabeled data. Only lossless data without labels are needed during the training process. This will save researchers countless hours preparing for training data.
- The discriminator network is applied to anomaly detection field. And this idea will provide strong support for the applications of unsupervised learning.

## 2 Detect White Gaussian Noise and Its Level with DCAGN

Compared with previous machine learning algorithms, GNAs has distinct advantages in many respects. For example, it can generate more distinct and higher-quality images than other generative algorithms. What's more, GANs can be trained to adapt to any kind of generative tasks without setting any specific function. More important, GANs can realize unsupervised learning [6]; in other words, the model can be trained with unlabeled data. Although GANs have advantages in many aspects, there are still some defects for GANs. Firstly, GANs face non-convergence problem. Training a GAN requires finding the Nash equilibrium in the two-play game; sometimes gradient descent does this, but sometimes it does not. Secondly, in the training process, GANs face the collapse problem. This will lead to the degeneration of generator network, generating numerous similar images. Alec Radford et al. proposed DCGAN which combines supervised CNNs with unsupervised GANs to improve the above shortcomings. Therefore, we apply DCGAN to build our model. The architecture of the two neural networks (the $G$ and $D$) can be designed in a variety of ways. In the next section, we will provide detailed information about the architecture of the $G$ and $D$.

The generator netwrok which consisted of serveral deconvolutional neutral networks aims to generate fake images according to the distribution of real data. And try to make the discriminator network cannot distinguish the difference between real data and the fake data generated by itself. The structure of the generator network is as shown in Fig. 1. We define a random vector $Z \sim P_z(z)$ which is a 100-dimensional uniform distribution valued $[-1, 1]$ as the input of the generator network. Next, we change the shape of $Z$ in order to fulfill the demand of deconvolutional operation. Through the deconvolutional process, the weight and the height of input will be twice of its used size. Meanwhile, the depth will reduce by half. Between the two deconvolutional layers, an important operation cannot be

ignored, that is batch normalization (BN) [7]. Through the experiment, we find the improved algorithm avoids overfitting of the neural networks and improves the convergence speed and accuracy of the whole network. Furthermore, the activation function is used to add the nonlinear factor, because the expression of the linear model is not enough. We choose rectified linear unit (RELU) [8] activation function to activate neurons of networks, in our model. After five deconvolutional layers, a fake image with the shape of 16 * 16 * 1 is generated by the generator network.

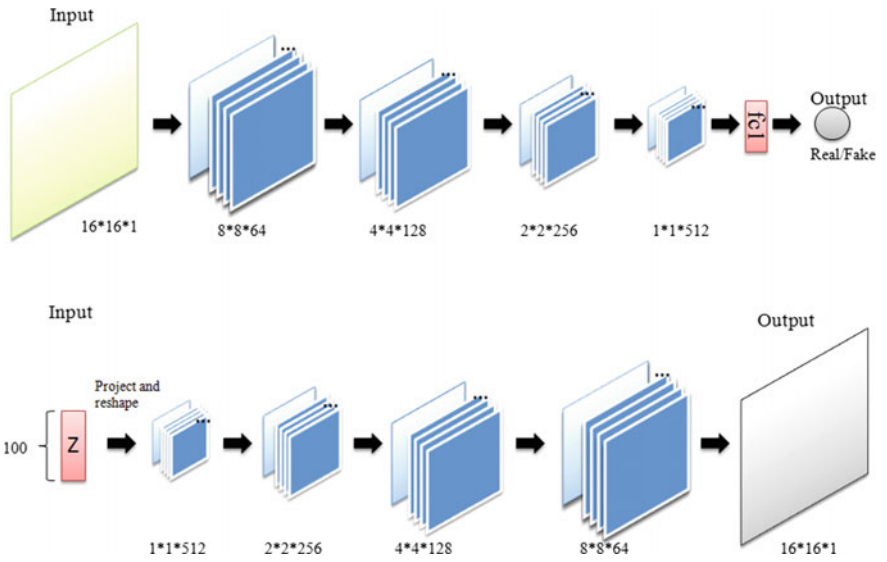



**Fig. 1.** Architecture of discriminator and generator

The discriminator network, as the competitor of generator network, focuses on distinguishing the difference between the real data $x$ and the fake data $G(Z)$ generated by $G$. The architecture of $D$ is as shown in Fig. 1. The discriminator network consists of four convolutional layers and one fully connected layer (FC). The architecture starts with taking an image with the shape of 16 * 16 * 1 as the input. There are two kinds of input for $D$, the real image $x$ from training dataset and the fake data $G(Z)$ generated by $G$. The discriminator attempts to export 1 when the input is real data $x$, and 0 for fake data $G(Z)$. In each convolutional layer, the input images obtain low-level features using convolutional filters. The parameters of these filters will be determined by back-propagation algorithm in training process. As well as to avoid vanishing gradient problem, a batch normalization (BN) layer is added behind the convolutional layer to normalize the data. The last layer is fully connected layer (FC), which is used to print out the judgment of discriminator (1 for real data $x$, 0 for fake data $G(Z)$). In fully connected layer, we choose sigmoid function as the activation function.

The cost function that we used to train $G$ is:

$$V = E_{x\sim p_{g(x)}}[\log(1 - D(x))] \tag{1}$$

$$V = -E_{x\sim p_{\text{data}(x)}}[\log(D(x))] - E_{x\sim p_{g(x)}}[\log(1 - D(x))] \tag{2}$$

During the training process, we must pay attention to keeping a balance between the training process of $G$ and $D$. Usually, we train the $D$ for $k$ times before we train the $G$ for one time. In the experiment, we use Adam optimizer to update the parameters of $G$ and $D$.

## 3 Experiment

The target of our experiment is to detect Gaussian noise and its level through the discriminator of DCGAN. In this section, we will offer more details about the preparation of training dataset, the process of detection, and the performance of our model.

### 3.1 Datasets

**LIVE**: The LIVE dataset contains 29 original pictures. And it contains 982 distorted images which are created by above-mentioned 29 reference images with five distorted types: JPEG2000, JPEG, Fastfading, White Gaussian Niose (WGN) and Guassian Blur (GBLUR). Each type of distortion contains five images with different degrees of distortion. In the experiment, we only consider using the 29 original pictures and the WN distorted pictures generated by them.

**CSIQ**: This dataset contains 30 original pictures with the same size. Meanwhile, 866 distorted pictures which are divided into six different distorted types are collected. Similar to LIVE, we only consider using the original pictures and the WN distorted pictures generated by them in our experiment.

### 3.2 Training Process

In Sect. 3, the structures of $G$ and $D$ are introduced. In this section, we will offer more details about the training process of the detection model. Firstly, we chose the 59 original images from databases LIVE and CSIQ to build the training datasets. In order to extract the structure information of the image, we changed the picture format from RGB to YUV (luma–chroma model) and only kept the luminance information of the images. Then, we cut the treated images into small pieces which size is 16 * 16 * 1. By doing this, we got the training datasets which contain 71,009 image blocks. The size of training epoch was set to 1000, and the mini-batch size was 64. Meanwhile, the noise, with 100 dimensions, was put into the $G$ to generate 64 fake images according to the distribution of training data. The $D$ took the training images and the fake images as the input and gave its identification results. According to Eq. 1, we calculated the loss function of $G$ and $D$. At last, we updated the parameters of $G$ and $D$ through the optimizer which is mentioned in Sect. 3 until the model converged.

### 3.3 Detection Process

Since we only trained the model with original pictures, the trained discriminator can easily detect the picture which is polluted by white Gaussian noise and its degree. And the discriminator will give its detection results in the form of score. The lower score we got, the higher Gaussian noise pollution degree is. Based on the RGB to YUV algorithm, we first extracted luminance information of the images. Then, we cut the treated images into small pieces with the size of 16 * 16 * 1, took them as the input of discriminator, and got the detection results. Finally, we added the score of each image block and normalized the total score to make the range of scores between 0 and 1.

### 3.4 Evaluation Criteria

In order to evaluate the detection performance of the model we proposed, we choose the following evaluation criteria:

- Detection rate (DC): It is used to evaluate whether the model can detect white Gaussian noise effectively. And it is defined as follows:

$$\mathrm{DC} = \frac{\mathrm{img}_{\mathrm{wn}}}{\mathrm{img}_{\mathrm{all}}} \tag{3}$$

where $\mathrm{img}_{\mathrm{all}}$ is the total number of the images with white Gaussian noise, and $\mathrm{img}_{\mathrm{wn}}$ represents the number of the images which can be detected accurately by our model. In the test process, we set a threshold to help determine whether the detected image is polluted by white Gaussian noise or not.

- Pearson product-moment correlation coefficient (PLCC): This is a nonparametric algorithm for measuring the linear correlation.

$$\mathrm{PLCC} = \frac{\sum_i (q_i - q_m)(\hat{q}_i - \hat{q}_m)}{\sqrt{\sum_i (q_i - q_m)^2}\sqrt{\sum_i (\hat{q}_i - \hat{q}_m)^2}} = \frac{\mathrm{cov}(q_i, q_m)}{\delta_{q_i}\delta_{q_m}} \tag{4}$$

where $q_i$ is truly score and $q_m$ represents the predicted scores generated by discriminator.

### 3.5 Experiment Results

We test our model in public database LIVE and CSIQ. Firstly, the quality scores of all the test images are calculated by the discriminator. In order to test whether our algorithm can detect the pictures which are polluted by white Gaussian noise accurately. We use the detection rate (DC) to measure it according to Eq. 3. The threshold is set to be 0.98. We consider that the image is polluted by WN if its quality score is lower than the threshold and vice versa. As shown in Table 1, the detection rates based on two public databases are both higher than 95%. And that means, our algorithm has excellent

performance in white noise detection. On the other hand, the ability of our algorithm to detect the degree of WN is measured by PLCC. In other words, we measure it through calculating the correlation between the quality scores (evaluation) and the subjective scores (loss). The value of PLCC is shown in Table 1. The scatter plots of subjective scores and quality scores on two public datasets are shown in Fig. 2.

**Table 1.** Experiment results

| Database | Detection rate (%) | PLCC |
|---|---|---|
| LIVE | 96.5 | 0.825 |
| CSIQ | 97.3 | 0.834 |

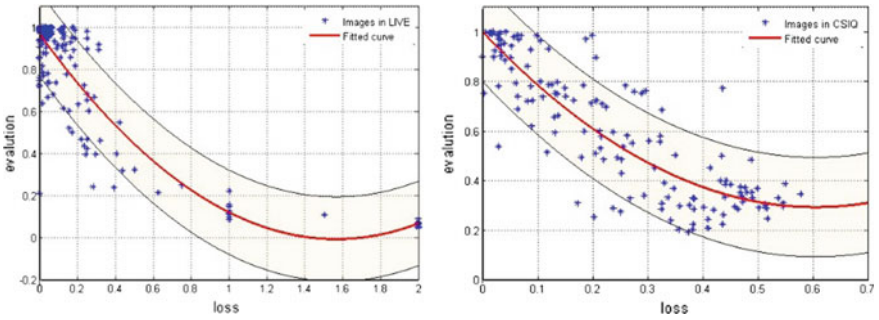



**Fig. 2.** Objective scores by our algorithm

## 4 Conclusion

In this paper, we propose a new deep learning method for the detection of white Gaussian noise and its degree using DCGAN. Through the experiment, we find it is feasible to detect the Gaussian noise and its level by the model we proposed. It also makes it more convenient for research to choose suitable filters to eliminate Gaussian noise from images. Meanwhile, different from the former applications of GANs, we turn our attention to the discriminator ($D$) instead of generator ($G$) and apply the $D$ to the anomaly detection filed. This application will also inform researchers of great value of discriminator. However, there are still several defects in our algorithm to be perfected. The detection accuracy of the level of Gaussian noise needs to be improved, and the training time of our model needs to be reduced.

## References

1. Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Sixth international conference computer vision (IEEE Cat. No.98CH36271). p. 839–46.
2. Buades A, Coll B, Morel J-M. A non-local algorithm for image denoising. In: IEEE computer society conference computer vision pattern recognition, 2005. CVPR 2005, vol. 2; 2005. p. 60–5.

3. Dabov K, Foi A, Katkovnik V. Image denoising by sparse 3D transformation-domain collaborative filtering. IEEE Trans Image Process. 2007;16(8):1–16.
4. Chuah JH, Khaw HY, Soon FC, Chow C. Detection of Gaussian noise and its level using deep convolutional neural network; 2017. p. 2447–50.
5. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst. 2012;25:1–9.
6. Heidemann G. Unsupervised Learning.pdf. Image Vis Comput. 2005;23:861–76.
7. B. N. Transform and B. N. Transform, "Batch Normalizing".
8. Hara K, Saito D, Shouno H. Analysis of function of rectified linear unit used in deep learning. In: Proceedings of international joint conference neural networks, vol. 2015–Sept 2015.

# Super-Resolution Imaging Using Convolutional Neural Networks

Yingyi Sun(✉), Wenhua Xu, Jie Zhang, Jian Xiong, and Guan Gui

College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
{1017010607, guiguan}@njupt.edu.cn

**Abstract.** Convolutional neural networks (CNN) have been applied to many classic problems in computer vision. This paper utilized CNNs to reconstruct super-resolution images from low-resolution images. To improve the performance of our model, four optimizations were added in the training process. After comparing the models with these four optimizations, Adam and RMSProp were found to achieve the optimal performance in peak signal to noise ratio (PSNR) and structural similarity index (SSIM). Considering both reconstruction accuracy and training speed, simulation results suggest that RMSProp optimization in the most scenarios.

**Keywords:** Super-resolution imaging · Convolutional neural networks · Gradient descent · Adam · RMSprop

## 1 Introduction

Recent years, with the development of optimized algorithms, mass data, and the increased performance of hardware, deep learning has received a lot of attentions in many aspects, such as super-resolution imaging, image recognition, etc. Among these topics, super-resolution imaging has been deeply studied. The goal of super-resolution imaging is to rebuild high-resolution images from the low-resolution images, which is one of the classic problems in computer vision.

Dong et al. proposed a relative new method for super-resolution imaging based on deep learning [1]. They establish connections between the images with the low and high resolution correspondingly. It is an end-to-end method that generates one high-resolution image as the output from only one low-resolution image as the input. To realize the connections between the inputs and the correspondingly outputs, they use deep convolutional neural networks (CNN) to obtain the mapping between the inputs and outputs.

In this paper, we further study the CNN model [1] and then train our own CNN models with partial optimizations. We find that gradient descent algorithm, which is the core of CNN's weight update, can be optimized by some other strategies in deep learning field. We improve one of the best models which have good performance by these optimizations, such as stochastic gradient descent (SGD), adaptive gradient (AdaGrad), adaptive moment estimation (Adam), and RMSprop. Among these


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_8

optimizations, we find that RMSprop and Adam have the best performance on peak signal to noise ratio (PSNR), and AdaGrad and SGD are relatively poorer.

The remainder of this paper is organized as follows. Section 2 presents the related works about super-resolution imaging. In Sect. 3, we utilize some optimization algorithms to improve our proposed model. Experimental results are conducted to confirm the proposed method in Sect. 4.

## 2 Related Works

As a classic problem in computer vision, super-resolution has been studied in various methods. To solve this underdetermined inverse problem, the early methods are mostly based on sampling theory [2–4]. And the limitations that the detailed textures are not as good as expected are then discovered.

Recently, deep learning has been used in various fields, and no exception in super-resolution imaging. Convolutional neural networks have been invented for decades and many computer vision fields, such as pedestrian recognition [5], image classification [6, 7], object detection [8], and face recognition [9].

Dong et al. [1] use their deep CNNs models to rebuild high-resolution images. According to their results, 9-1-5 model has the best score considering performance and running time, under contrast of bicubic interpolation. Their deep CNNs have three convolution layers, each of which followed by an activation layer. The rectified linear unit (ReLU) [10] is used as the activation layer. ReLU makes the model converge faster in the case of high performance. The good performance of the deep CNNs models is shown in their research, with a lot of experimental results as support.

Figure 1 shows the structure of the model. This model has three convolutional layers (yellow), each of which is followed by an activation layer (green) named ReLU, except the third convolutional layer. Mean squared error (MSE) is used as loss function (orange) to calculate the loss between data and label. The light blue blocks represent the input and output data.

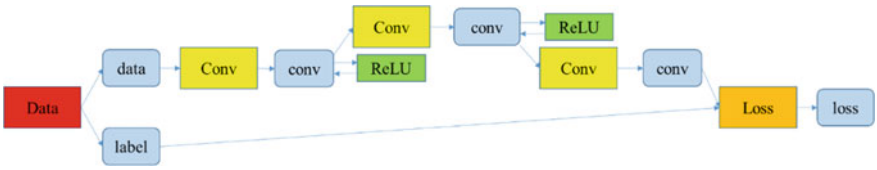



**Fig. 1.** Model structure of CNNs

## 3 Optimization Methods

### 3.1 Model Structure

Dong et al. [1] proposed their 9-1-5 model to train the convolutional neural networks for super-resolution imaging. As shown in Fig. 1, the model contains three

convolutional layers, each of which is connected to an activation layer named ReLU except the third convolutional layer. The operation of ReLU is:

$$O = \max(0, I) \tag{1}$$

where $O$ denotes the output of the ReLU as well as $I$, the input. Thinking of a convolutional layer and a corresponding activation layer as a pair, the operation of the first two pairs is:

$$Y_i = \max(0, W_i * X_i + B_i) \tag{2}$$

where $i$ equals to 1 or 2 denoting the first or second pair and $X_i$ or $Y_i$ represent the input or output of each pair, respectively. Here, $W_i$ corresponds to the weight coefficient of the filter, as well as $B_i$ the bias coefficient. These parameters like $W_i$ and $B_i$ are achieved by minimizing the loss function, which calculates the mean squared error (MSE) of the reconstructed image and the corresponding high-resolution image. The formula for MSE is:

$$L = \frac{1}{n}\sum_{i=1}^{n} \|X_{ri} - X_{hi}\|^2 \tag{3}$$

where $n$ is the number of training samples, and $X_{ri}$ or $X_{hi}$ represent the reconstructed image or the corresponding high-resolution image, respectively. $L$ is the function of the filter parameters, such as weight and bias coefficients.

## 3.2 Some Optimization Algorithms

### 3.2.1 Stochastic Gradient Descent (SGD) Algorithm

Gradient descent (GD) is widely used in deep learning field. In this paper, it can help the CNNs determine the most suitable parameters of the filters as soon as possible, under the premise of good performance. As a result of GD's easily falling into a local optimal solution, this paper replaces GD with SGD. SGD in this article refers to mini-batch SGD. The difference of GD and SGD is only the quantity of the training samples. The training sample in GD is the whole data set, while SGD uses only a part, which calculates some training samples once a time instead of all. Samples for training in SGD are randomly selected at each iteration.

For example, the weight coefficients are updated as

$$W_{i+1} = W_i - \eta \frac{\partial L}{\partial W_i} \tag{4}$$

Here, $\eta$ is learning rate, $L$ denotes the loss function, and $W$ represents the weight coefficients. The iteration number is denoted by $i$.

### 3.2.2 Momentum Algorithm

Momentum algorithm is such a method that it simulates the concept of momentum in physics to optimize the SGD algorithm. The most notable feature is that the last gradient is taken into account in current gradient calculation. It can be expressed as

$$W_{i+1} = \mu W_i - \eta \frac{\partial L}{\partial W_i} \tag{5}$$

Here, the added $\mu \in (0, 1)$ is momentum factor, which controls the effect of the last gradient on the current gradient. The general value of $\mu$ is 0.95.

### 3.2.3 AdaGrad Algorithm

The parameters in algorithms above, such as momentum algorithm, are needed to be set based on experience. Therefore, the determination of specific values of these parameters is seem to be fixed, at most in several common values. Obviously, such setting method cannot be convinced.

AdaGrad algorithm as well as the following methods can set parameters adaptively. After calculating the gradient, AdaGrad algorithm calculates the cumulative squared gradients for the next updating.

### 3.2.4 RMSProp Algorithm

Considering to the non-convex properties of the CNNs and RMSProp algorithm's adaptive to non-convex, RMSProp algorithm can be used for the optimization. Based on the momentum and AdaGrad algorithms, RMSProp combines the superiority of both two algorithms. It takes into account both the second-order moments of the gradient and the historical effects.

### 3.2.5 Adam Algorithm

Considering the advantage of AdaGrad and RMSProp on dealing with sparse gradients and non-stationary objectives, Adam algorithm can be seen as a combination of AdaGrad and RMSProp. It makes full use of first and second moments of the gradients to adjust the learning rate of each parameter, respectively.

Adam is superior to the other optimization algorithms in terms of performance. Yet it calculates both the first and second moments of the gradients, the training speed is relatively slower. In spite of this, Adam is still found to be robust in problems with large datasets or parameters with high dimension and is widely applied.

## 3.3 Comparison of Several Optimization Algorithms

In Sect. 3.2, this paper introduces some optimization algorithms for traditional gradient descent methods. The following part we will discuss these algorithms, such as SGD, AdaGrad, RMSProp, and Adam.

### 3.3.1 GD and SGD

In GD algorithm, the loss functions need to traverse the entire training data, which is waste of time and is more possible to fall into local optimal solutions. Therefore, in deep learning field like this super-resolution imaging problem, the CNNs model falls into the unwanted solutions easily.

Compared to GD, the samples for training in SGD are randomly selected at each iteration. In this way, we can avoid the local optimal solutions and save much running time.

### 3.3.2 AdaGrad, RMSProp, and Adam

As shown in the algorithms, AdaGrad makes the use of cumulative squared gradients to minimize the loss functions. As a result, it can make the learning rate adaptive. However, the training speed of AdaGrad is relatively slower.

RMSProp optimizes AdaGrad using the Newton iteration method in the process of accumulating squared gradient. The added coefficient in RMSProp is used to control the acquisition of historical squared gradients.

Adam combines the advantages of both AdaGrad and RMSProp. It can calculate different adaptive learning rates for different parameters and is suitable for big or high-dimensional data. Beside these, Adam needs smaller requirements for memory.

The detailed performance comparisons of these algorithms are discussed in Sect. 4.

## 4 Experiments

In this section, we first discuss the determination of training data and label. Next, we add the four optimizations discussed in Sect. 3 to the model and compare the performance of them, respectively. The last but not least, we examine the training loss and speed of the model with four optimizations.

### 4.1 Training Data

To train our model for super-resolution imaging, we need to determine the data and label in training sets. The purpose of our model is to rebuild the images, so the super-resolution images generated from low-resolution images by CNNs are seen as data. In order to continuously improve the performance of the model, it needs the label as contrast. The corresponding high-resolution images are used as label, which is the same as data as the input of the loss layer. Mean square error (MSE) is used as a loss function in this paper.

To realize the process above, the size of data and label should be calculated ahead of time. For example, the sizes of these three types of convolution kernels in the 9-1-5 model are set to 9 by 9, 1 by 1, and 5 by 5, respectively. With the stride of 1, the size of data and label can be 33 by 33 and 21 by 21, respectively.

### 4.2 Comparison of the Performance

Figure 2 shows the performance of bicubic reconstructed image and another four images with optimizations. It displays one original image and five reconstructed images as contrast. The first image is the original image, which is grayed out. The reconstructed image below the original one is using bicubic interpolation, which is obviously fuzzy. The right four images are clearer, which reflects on the effects of these optimizations.

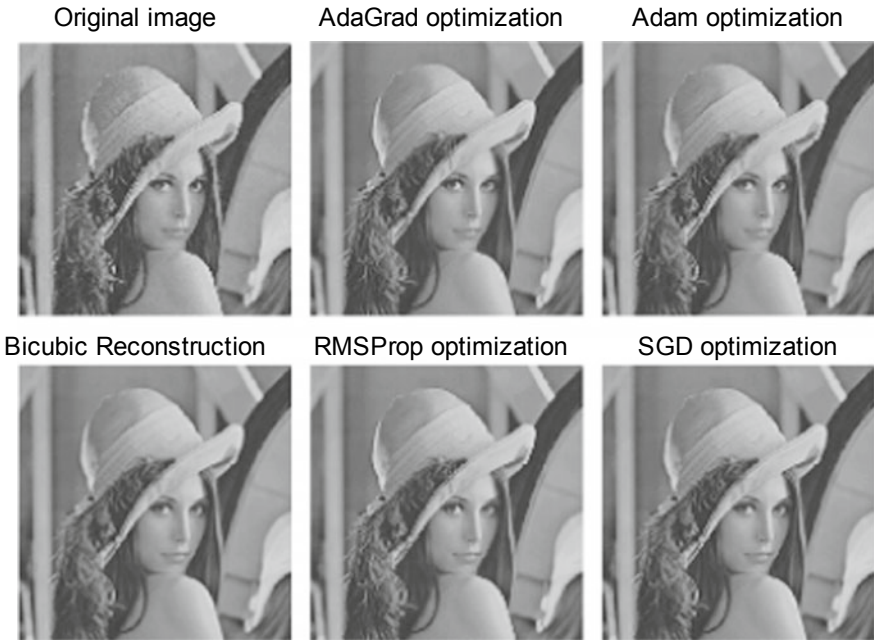



**Fig. 2.** Original and five reconstructed images

Table 1 displays the specific PSNR and SSIM of five reconstructed images. It is clear that Adam and RMSProp achieve the highest PSNR 33.30228 and 33.10496 dB, respectively. As a result, Adam is the best choice in terms of performance and the

**Table 1.** PSNR and SSIM of five reconstructed images

| Type | PSNR (dB) | SSIM |
|---|---|---|
| Bicubic | 31.67762 | 0.859749 |
| AdaGrad | 32.20875 | 0.867244 |
| Adam | 33.30228 | 0.882372 |
| RMSProp | 33.10496 | 0.879238 |
| SGD | 32.42305 | 0.874023 |

performance of RMSProp is slightly worse. It shows that Adam has the highest scores of SSIM with 0.882372 while RMSProp ranks second with 0.879238, followed by SGD and AdaGrad, respectively.

### 4.3 Comparison of the Training Process

Section 4.2 discusses the performance of the model with four optimizations, respectively. Yet considering performance alone to choose the most suitable model cannot be convinced. Thus, the training process is also recorded to be the basis for selecting the most suitable model.

Figure 3 describes the changes in training loss of the models with optimizations like AdaGrad, Adam, RMSProp, and SGD during the training process. It is obvious that AdaGrad and Adam have the minimal training loss after convergence.

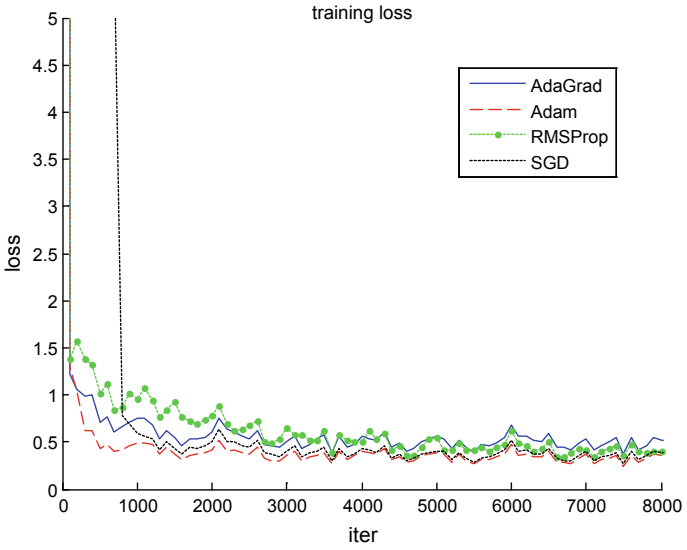



**Fig. 3.** Comparison of the training loss of the four models

In addition, Adam has another advantage that it makes the model convergences faster. As a result, considering the performance and effect of convergence, Adam is superior to other optimizations.

Table 2 displays the accurate values of running time for every 100 iterations. As described in Fig. 3, the fastest training speed for RMSProp and SGD is about 52.1928 and 51.85337 seconds for every 100 iterations. In contrast, Adam's training speed is two to three times slower. As a result, considering the performance, speed of convergence, and training time, the model with RMSProp optimization is the most suitable.

**Table 2.** Average time required for 100 iterations

| Type | Time (s/100 iterations) |
|---|---|
| AdaGrad | 272.9645 |
| Adam | 130.0681 |
| RMSProp | 52.1928 |
| SGD | 51.85337 |

## 5 Conclusion

In this paper, we first utilized the convolutional neural networks to reconstruct the super-resolution images from the corresponding low-resolution images. Then, we introduced four optimization algorithms and discussed the relationships between them. At the end, we analyzed the performance and training process of the models with these four optimizations, and we found that in the case of all these factors, RMSProp may be the most suitable optimization algorithm for the super-resolution imaging.

## References

1. Dong C, Chen CL, He K, Tang X. Image Super-Resolution Using Deep Convolutional Networks. IEEE Trans Pattern Anal Mach Intell. 2016;38(2):295–307.
2. J. Allebach and P. W. B. T.-I. C. on I. P. Wong 1996. Proceedings, "Edge-directed interpolation," *Int. Conf. Image Process.*, vol. 3, no. 3, pp. 707–710, 1996.
3. Li X, Orchard MT. New edge-directed interpolation. IEEE Trans Image Process. 2001;10 (10):1521–7.
4. Zhang L, Wu X. An edge-guided image interpolation algorithm via directional filtering and data fusion. IEEE Trans Image Process. 2006;15(8):2226–38.
5. Ouyang W, X. B. T.-I. I. C. on Wang CV. Joint deep learning for pedestrian detection. In: IEEE international conference on computer vision; 2014. p. 2056–63.
6. He K, Zhang X, Ren S, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell. 2015;37(9):1904–16.
7. Krizhevsky A, Sutskever I, GEBT-IC on Hinton NIPS. ImageNet classification with deep convolutional neural networks. In: International conference on neural information processing systems; 2012. p. 1097–105.
8. Ouyang W, Zeng X, Wang X, Qiu S, Luo P, Tian Y, Li H, Yang S, Wang Z, Li H. DeepID-Net: deformable deep convolutional neural networks for object detection. IEEE Trans. Pattern Anal. Mach. Intell. 2014;99:2403–12.
9. Sun Y, Wang X, Tang X. Deep learning face representation by joint identification-verification, vol. 27; 2014. p. 1988–96.
10. Nair V, GEBT-IC on IC on Hinton ML. Rectified linear units improve restricted boltzmann machines. In: International conference on international conference on machine learning; 2010, p. 807–14.

# Image Super-Resolution Based on Multi-scale Fusion Network

Leping Lin[1,2], Huiling Huang[2], and Ning Ouyang[1,2(✉)]

[1] Key Laboratory of Cognitive Radio and Information Processing (Guilin University of Electronic Technology), Ministry of Education, Guilin 541004, Guangxi, China
lin_leping@163.com, ouyangning@guet.edu.cn
[2] School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, Guangxi, China
834597430@qq.com

**Abstract.** It is important and necessary to obtain high-frequency information and texture details in the image reconstruction applications, such as image super-resolution. Hence, it is proposed the multi-scale fusion network (MCFN) in this paper. In the network, three pathways are designed for different receptive fields and scales, which are expected to obtain more texture details. Meanwhile, the local and global residual learning strategies are employed to prevent over-fitting and to improve reconstruction quality. Compared with the classic convolutional neural network-based algorithms, the proposed method achieves better numerical and visual effects.

**Keywords:** Multi-pathways · Residual learning · Multi-scale · Receptive field · Texture details

## 1 Introduction

Deep learning (DL)-based methods are well applied in super-resolution reconstruction [1, 2]. The advantage of the methods is the new nonlinear mapping learning idea. As the pioneer network model for SR, super-resolution convolutional neural network (SRCNN) [3] established the nonlinearlow resolution–high resolution (LR–HR) mapping via a fully convolutional network, and significantly outperformed the classical non-DL methods. However, the method loosed many textures, and converged slowly. Later, the accelerated version of SRCNN, namely FSRCNN [4], is proposed, where the de-convolution layer instead of the bicubic interpolation is made used to improve the reconstruction quality. Based on it, Shi proposed a more efficient sub-pixel convolution layer [5] to fulfill the upscaling operation, which is performed at the very end of the network.

Though the available models based on deep neural networks have achieved success both in reconstruction accuracy and computational performance, the high-frequency information and texture details of images are ignored. In this paper, it is proposed the image super-resolution method based on the multi-scale fusion network (MCFN), which enlarged the reception field [13] of the network by multiple pathways and


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_9

obtained more texture details. The network is composed of multiple pathways and trained by the residual strategies. The advantages of this network are twofolds:

- Three pathways that employed different sizes of convolution layers are designed to get different size of receptive fields, which could extract multi-scale features of the input image. This network design is expected to preserve more image details after the extracted features are fused for the reconstruction.
- The residual strategy is composed of two types: local and global residual learning. The former, which linked the input and the convergence end of the pathways, aimed at optimizing this network. The local residual learning, which linked the input and the end of the network, aimed at preventing overfitting. The use of residual learning could further improve the reconstruction quality.

The method will be validated based on the publicly available benchmark datasets, and compared with the available DL-based methods. By the experiments, it is shown that the proposed method achieves better results.

## 2 Multi-scale Fusion Network

The proposed multi-scale fusion network is composed of three parts: the block extraction and representation, the nonlinear mapping, and the reconstruction part. As shown in Fig. 1, the nonlinear mapping composed three pathways, which could obtain more high-frequency information and texture details than single pathway. The residual learning has two patterns: local residual learning and global residual learning, which could prevent overfitting training, and further improve the quality of reconstruction.

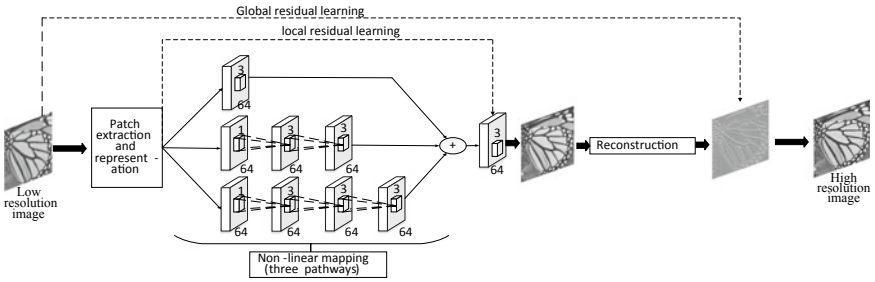



**Fig. 1.** Framework of the proposed MCFN method

### 2.1 Multiple Pathways Design

The three pathways in the nonlinear mapping part are designed to obtain feature details of diverse scales, which are followed by the fusion operation for reconstruction. The three pathways perform 1, 2, and 3 cascaded $3 \times 3$ convolution operations, respectively. Besides, a $1 \times 1$ filter is put on the beginning of the second and third pathway [6] to reduce the amount of parameters. The design is motivated by the GoogLeNet [7], where multiple paths are adopted to obtain various receptive fields. Besides, based on

GoogLeNet, Kim [8] proposed to remove the pooling operations, and to replace the small convolutional kernels by the big ones, which achieved good performance in image restoration. The paralleled pathways could easily be applied the parallel computations.

Figure 2 shows the features abstracted by the three pathways. It can be seen that different receptive fields and features of different scales are obtained from different paths. The first pathway learns a smooth feature, the second pathway learns texture details, and the third pathway learns contour information.

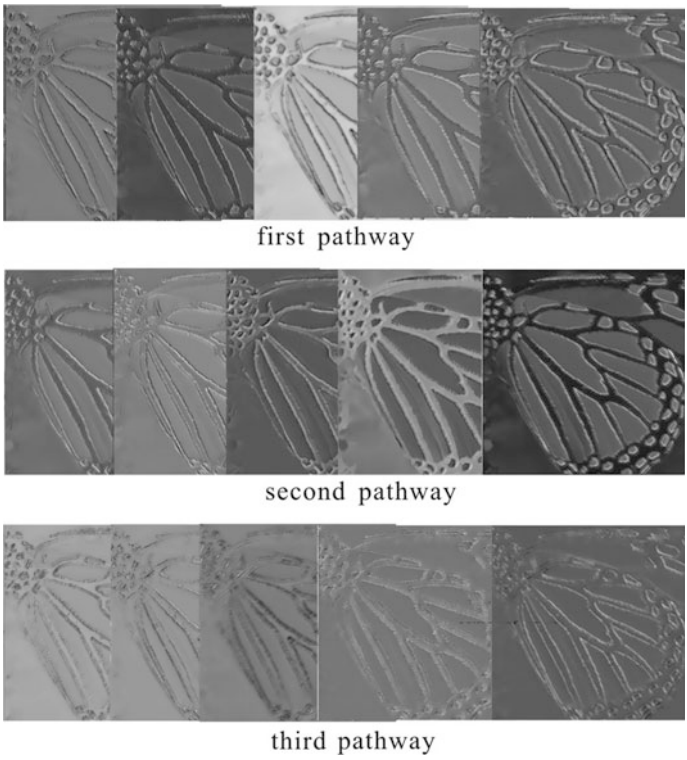



**Fig. 2.** Different texture details abstracted by three pathways, respectively

The output formulas of these pathways as follows:

$$F_1 = \sigma(w_1 x) \tag{1}$$

$$F_2 = \sigma(w_4 \sigma(w_3 \sigma(w_2 x)) \tag{2}$$

$$F_3 = \sigma(w_8 \sigma(w_7 \sigma(w_6 \sigma(w_5 x)))) \tag{3}$$

where $w_1$, $w_2$, $w_3$, $w_4$, $w_5$, $w_6$, $w_7$, $w_8$ represent the filters of the convolution layers, and $\sigma$ denotes the rule operation. $F_1$, $F_2$, and $F_3$ are the output of the first, second, and third pathway, respectively.

### 2.2 Residual Learning

In order to optimize the training of the nonlinear mapping part, and to improve the convergence speed, a residual connection is designed between the input and output of the pathway. Since the LR image and the HR image compose the same structures, and share the same low-frequency information, it would be advantageous to model the connection [9–11]. Similarly, a residual connection is made between the input and the output of the network. This connection has proven to be an effective network structure.

The residual learning structures are shown in Fig. 1. The two parts adopt the structure of residual learning, which is achieved through the forward neural network and the residual connection. The connection between the convolutional layers is a sequential connection using batch normalization (BN) and rectified linear units (Relu). This is a forward-activated structure proposed by He [12], which is easy to train and easier to fit. At the same time, the residual connection is equivalent to simply performing the equivalent mapping, without generating additional parameters or increasing computational complexity to prevent overfitting of the network.

In the nonlinear mapping stage, it is achieved through three pathways. In this part of the input and output residual connection, this connection structure is different from Resnet, as shown in the following formula:

$$y_l = h(x_l) + F(x_l, w_l) \tag{4}$$

$$x_{l+1} = f(y_l) \tag{5}$$

where $x_l$ and $x_{l+1}$ are the input and output of the first layer, respectively. $F$ is a residual function. $h(x_l) = x_l$ is an identity mapping and $f$ represents the Relu. The activation function of this connection structure $x_{l+1} = f(y_l)$ has an impact on two aspects, as shown in the following equation:

$$y_{l+1} = f(y_l) + F(f(y_l), w_{l+1}) \tag{6}$$

This is an asymmetrical approach that allows the activation function $\hat{f}$ to affect $l$ arbitrarily. Only the $F$ path can be affected, given by:

$$y_{l+1} = y_l + F(\hat{f}(y_l), w_{l+1}) \tag{7}$$

$F(\hat{f}(y_l), w_{l+1})$ is the sum of the nonlinear mappings of the three pathways, as shown as following:

$$F(\hat{f}(y_l), w_{l+1}) = f\{F_1 + F_2 + F_3) \tag{8}$$

The first row on the right side of the equation represents the output of the first pathway. The second and third rows of the equation represent the second pathway and the third pathway, respectively.

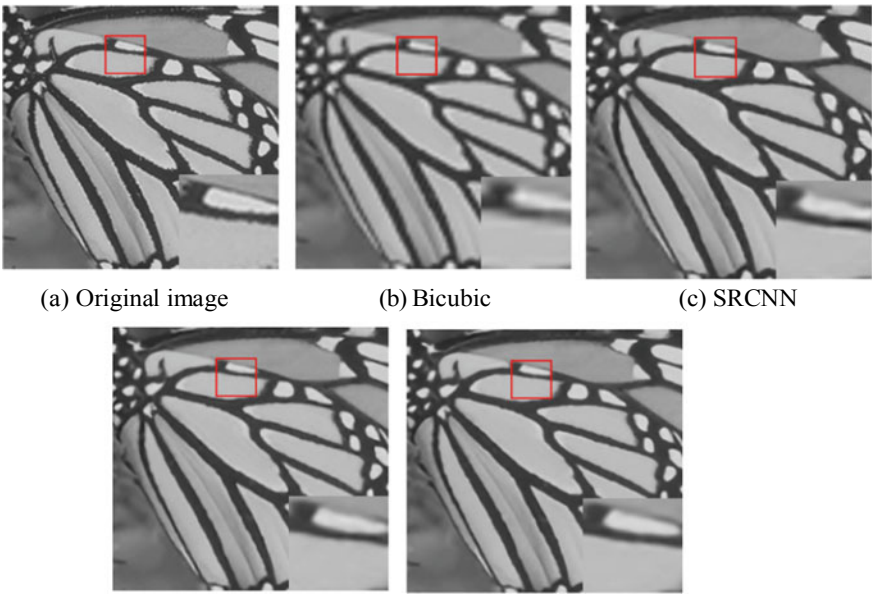
To optimize and estimate the network model, the mean squared error (MSE) is used as a loss function to estimate the network parameters, given by:

$$L(\theta) = \frac{1}{2N}\sum_{l=1}^{N}\left\|\tilde{x}^{(l)} - D(x^{(l)})\right\|^2 \tag{9}$$

$\{x^{(l)}, \tilde{x}^{(l)}\}_{i=1}^{N}$ is the training set, in which $N$ is the number of training fast, and $\tilde{x}^{(l)}$, $x^{(l)}$ are the ground truth value corresponding to the LR block. $\theta$ is a set of parameters to evaluate.

## 3 Experiments and Analysis

In order to make a comparison with the existing methods, a widely used training set of 91 images was used, and a magnification of 3 was used to train the data. A random gradient tailored training mode was used. The algorithms and comparison algorithms are on the same experimental platform (Intel CPU 3.30 GHZ and 8G memory), and MATLAB R2014a and Caffe are applied. The four common datasets are evaluated separately: Set5, Set14, BSD100, and Urban100. Quantitative assessment is used by



(a) Original image (b) Bicubic (c) SRCNN

(d) ESPCN (e) MCFN

**Fig. 3.** Comparison of reconstruction results of original HR and each pathway of butterfly

PSNR and SSIM, because humans are more sensitive to changes in brightness than colors. Following most existing methods, super-resolution reconstruction of luminance channels is performed only in the YCbCr color space in order to display patterns, and the other two channels are simply performed by bicubic interpolation.

As shown in Fig. 3, compared with classic networks, the reconstruction image of MCFN is more clear, which have more detailed textures which improves the reconstruction effect.

The comparison of the average numerical results is shown in Table 1. Through the comparative experimental analysis, it can be seen that MSFN has higher PSMR and SSIM, intuitively indicating that the information extracted by the MSFN is conducive to reconstruction.

**Table 1.** PSNR (dB) comparison of test image reconstruction results

| Dataset | Enlargement factor | Bicubic | SRCNN | ESPCN | MSFN |
|---|---|---|---|---|---|
| Set5 | 3 | 30.39 | 32.39 | 32.55 | 33.19 |
| Set14 | 3 | 27.54 | 29.00 | 29.08 | 29.48 |
| Urban100 | 3 | 27.05 | 27.94 | 28.12 | 28.40 |
| BSD100 | 3 | 27.57 | 28.45 | 28.65 | 28.96 |

## 4 Conclusions

This paper proposes a super-resolution reconstruction method based on multi-scale fusion network, which is composed multiple pathways and adopted residual learning strategies. The multiple pathways could capture the texture information of different scales, which increases the diversity of features, and improves the accuracy of reconstruction. Using the residual learning strategies can not only enhance the ability of the network model to fit features, but also optimize network model. The residual links at the input and convergence ends of the three pathways, as well as at the input and output ends of the network. By the experiment results, the method is shown effective.

**Acknowledgments.** This work is partially supported by the following foundations: the National Natural Science Foundation of China (61661017); the China Postdoctoral Science Fund Project (2016M602923XB); the Natural Science Foundation of Guangxi province (2017GXNSFBA198212, 2016GXNSFAA38014); the Key Laboratory Fund of Cognitive Radio and Information Processing (CRKL160104, CRKL150103, 2011KF11); Innovation Project of GUET Graduate Education (2016YJCXB02); the Scientific and Technological Innovation Ability and Condition Construction Plans of Guangxi (159802521); the Scientific and Technological Bureau of Guilin (20150103-6).

## References


1. Jiao LC, Liu F, et al. The neural network of seventy years: review and prospect. Chin J Comput. 2016;39(8):1697–716.
2. Jiao LC, Zhao J, et al. Research progress of sparse cognitive learning calculation and recognition. Chin J Comput. 2016;39(4):835–51.
3. Dong C, Loy CC, He K, et al. Image super-resolution using deep convolutional networks. IEEE Trans Pattern Anal Mach Intell. 2016;38(2):295–307.
4. Dong C, Loy CC, Tang X. Accelerating the super-resolution convolutional neural network. In: European conference on computer vision. Springer International Publishing; 2016. p. 391–407.
5. Shi W, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 1874–83.
6. Lin M, Chen Q, Yan S. Network in network [EB/OL]. (2013-1-31)[2018-3-25] (2013) https://arxiv.org/abs/1312.4400.
7. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. p. 1–9.
8. Kim Y, Hwang I, Cho NIA. New convolutional network-in-network structure and its applications in skin detection, semantic segmentation, and artifact reduction [EB/OL]. (2017-1-22) [2018-3-25] https://arxiv.org/abs/1701.06190.
9. Huang J, Yu ZL, Cai Z, et al. Extreme learning machine with multi-scale local receptive fields for texture classification. Multidimension Syst Sig Process. 2017;28(3):995–1011.
10. Kim J, Kwon Lee J, Mu Lee K. Deeply-recursive convolutional network for image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 1637–45.
11. Kim J, Kwon Lee J, Mu Lee K. Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 1646–54.
12. He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks. In: European conference on computer vision. Springer International Publishing; 2016. p. 630–45.
13. Luo W, Li Y, Urtasun R, et al. Understanding the effective receptive field in deep convolutional neural networks. In: Advances in neural information processing systems; 2016. p. 4898–906.

# Quadratic Segmentation Algorithm Based on Image Enhancement

Ying Jiang, Heng Dong(✉), Yaping Fan, Yu Wang, and Guan Gui

Key Laboratory of Broadband, Wireless Communication and Sensor Network Technology (Nanjing University of Posts and Telecommunications), Ministry of Education, Nanjing 210003, China
{dongh, guiguan}@njupt.edu.cn

**Abstract.** Due to the indefinite position of the characters in the invoice and the difference of the color shades, which greatly increases the difficulty of intelligent identification and thus it is difficult to meet practical applications. In order to solve this problem, this paper proposes a quadratic segmentation algorithm based on image enhancement. Specifically, we firstly enhance the color of the image based on gamma transformation and then separate the machine-printing character from the blank invoice based on the color analysis of the machine-printing character. Then according to the open operation in the image processing field and the boundingRect algorithm, the pixel information of the machine-playing character is obtained, which is convenient for getting the character information. The algorithm can achieve effective extraction of machine-playing characters and also reduce the difficulty of invoice identification and improving the accuracy of invoice identification. Simulation results are given to confirm the proposed algorithm.

**Keywords:** Quadratic segmentation algorithm · Image processing · Invoice identification · Color determination · Image processing

## 1 Introduction

With the rapid development of social economy, the use of invoices in China is required in financial management. At present, China has used hundreds of millions of invoices for annual reimbursement, and it has shown an upward trend. However, most of the reimbursement of invoices is done manually. The manual reimbursement of invoices has many disadvantages such as complex reimbursement procedures, long manual processing time and high processing error rate. In other words, the manual reimbursement of invoices not only aggravates the financial staffs' workload, but also takes up a lot of extra energy from the reimbursement. From financial market perfective, the manual reimbursement spends extra labor cost and thus increases the product or management costs [1].

In recent years, with the rapid development of image processing and computer vision technology, high-precision, high-efficiency, and low-cost text recognition technology has been realized. Many domestic and foreign scientific research institutes and scientific researchers have introduced the emerging technology of computer vision into related fields such as invoice identification, and conducted rigorous and profound analysis on the feasibility of this technology. Hence, it is becoming more and more urgent to find an


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_10

effective and practical invoice processing method. VAT notes are printed by dot matrix printers, and the position of the invoices printed by different printers is indefinite, and the shades of color are different, which is one of the main reasons that lead to a serious decline in the quality of invoice information extraction. Therefore, it is significant to study image enhancement secondary segmentation in practical applications [2].

At present, the color of the machine-used characters for the invoice is divided into blue and black. The depth of character color directly affects the effect of image segmentation and recognition. Extracting too much information at a time also affects the recognition. There are still some deficiencies in the existing method for extracting characters from the machine. According to the extraction of the frame, the content to be recognized is too much, which greatly reduces the recognition accuracy.

This paper proposes a secondary segmentation extracted algorithm, which can be applied in the actual invoice identification system. The proposed algorithm first performs color enhancement [3] on the image. Then, based on the color analysis of the machine characters, the first segmentation is performed to separate the machine-printing characters from the blank invoices. The pixel information of the player is then obtained, and the secondary segmentation is performed so as to realize the extraction of the driver characters. Experiment results are given to validate the proposed algorithm.

## 2 Extraction of the Machine Character

The character extraction method of this paper is mainly composed of image enhancement and secondary segmentation. Figure 1 is the scan invoice image. Figure 2 shows the system flow.


**Fig. 1.** An example of an original invoice image. The above picture is an example of a Chinese VAT invoice. VAT is a kind of consumption tax. In some countries, it is also called "the goods and services tax is a regressive tax" or "consumption tax". It is an indirect tax that is taxed based on the value added of goods or services, which is printed and produced by the State Administration of Taxation.

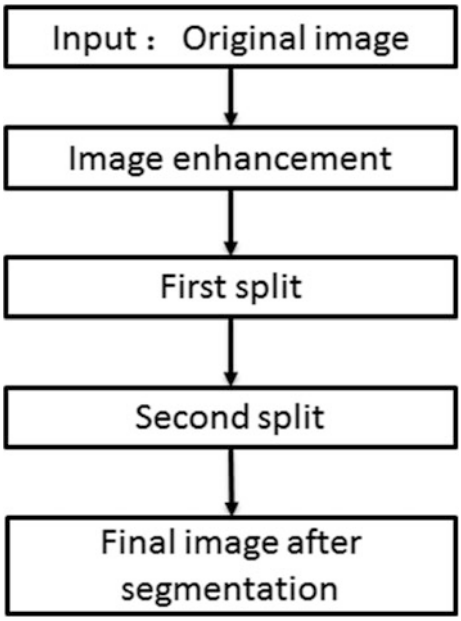



**Fig. 2.** Flow chart of image segmentation

First, image enhancement is performed on the acquired invoice color image to obtain a clearer image $\boldsymbol{I}$. A color judgment is made on the password area of the invoice, the color of the machine-playing character is detected, and the first blank invoice and the machine-used character are segmented to obtain the image $\boldsymbol{I_1}$ with only the machine-playing character. The black and white conversion of $\boldsymbol{I_1}$, corrosion, and the acquisition of the rectangle of the outside world are performed. The pixel information of the machine-played character is obtained, and after the secondary division, the machine extracts the character block $\boldsymbol{G}$.

### 2.1 Image Enhancement

The color of invoices produced by different merchants due to the difference of printers. Therefore, the image enhancement is required for invoices to make them clear. We use a method of image enhancement based on gamma transformation [4]. The enhancement effect is shown in Fig. 3.

The gamma conversion is mainly used for image correction and the correction of the image with too high grayscale or low grayscale, which is aim to enhance the contrast. The transformation formula does a multiplication of each pixel value on the original image:

$$S = c\gamma^{\gamma} \quad \gamma \in [0, 1] \tag{1}$$


**Fig. 3.** An example of the enhanced image

The correction effect of the gamma transformation on the image is actually achieved by enhancing the details of low gray level or high gray level.

The $\gamma$ value is demarcated by 1. The smaller the value is, the stronger the expansion effect is on the low-gradation part of the image. The larger the value is, the stronger the expansion effect on the high-gradation part of the image is. By changing the gamma value, the effect of enhancing the details of low gray levels or high gray levels can be achieved.

## 3 Secondary Segmentation

### 3.1 First Split

The machine-playing characters on the general invoice are divided into blue and black. Only the color of the machine-playing character of the value-added ticket is not fixed. Therefore, it is necessary to determine the color of the character played in the ticket. The character color that is the color other than white can be judged by using RGB in the password area. According to the judged color of the machine, the corresponding color is divided.

There is a fixed frame line for the value-added tax ticket, we can firstly use the mouse to take the function ginput() [5], which can manually cut out the part to be extracted according to the fixed frame of the blank invoice, and generate the Excel file of the location information, such as the buyer, the seller, and the password area. Then import Excel in the code, you can initially split the invoice, reduce the follow-up workload, and improve the extraction rate.

The function ginput provides a cross cursor so that we can more accurately select the position we need and return the coordinate value. The function call form is:

[$x$, $y$] = ginput($n$), which enables you to read $n$ points from the current coordinate system and return the $x$, $y$ coordinates of these $n$ points, all of which are $n$X1 vectors.

Importing Excel here to segment the password area is used to color determination. Assuming that the color of the machine-playing character is blue, which is shown in Fig. 3, the invoice image is first converted from RGB space to HSV space [6], where $H$ represents hue, $S$ represents saturation, and $V$ represents brightness. Then, we use the in Range function to segment the image region we need by adjusting the $H$, $S$, and $V$ regions to obtain a white binary image $\boldsymbol{I_1}$ that satisfies the interval condition, which is shown in Fig. 4. After testing, $H$: 100–140, $S$: 30–255, $V$: 0–255, we can have the best effect.


**Fig. 4.** An example of the image after first split

### 3.2 Second Split

The black and white transformation of $\boldsymbol{I_1}$ is performed, and then the etching [7] is performed in a block operation to obtain an image $\boldsymbol{I_2}$ which is shown in Fig. 5 after that we can, respectively, press the imported Excel to separate the parts $\boldsymbol{I_3}$ and $\boldsymbol{I_4}$ except the password area in the images $\boldsymbol{I}$ and $\boldsymbol{I_2}$. $\boldsymbol{I_3}$ and $\boldsymbol{I_4}$ are, respectively, shown in Figs. 6 and 7. The boundingRect [8] function is used to calculate the minimum rectangle of the vertical boundary of the contour for $\boldsymbol{I_4}$, where the rectangle is parallel to the upper and lower boundaries of the image, and then the corrosion block boundary information is obtained. For visual convenience, the rectangle boundary is drawn using the rectangle function, as shown in Fig. 8. Then split the segmented part a second time.

We can suppose that the horizontal position of the invoice is $X$ and the longitudinal direction is $Y$. The pixel information of the top and bottom points of the rectangular

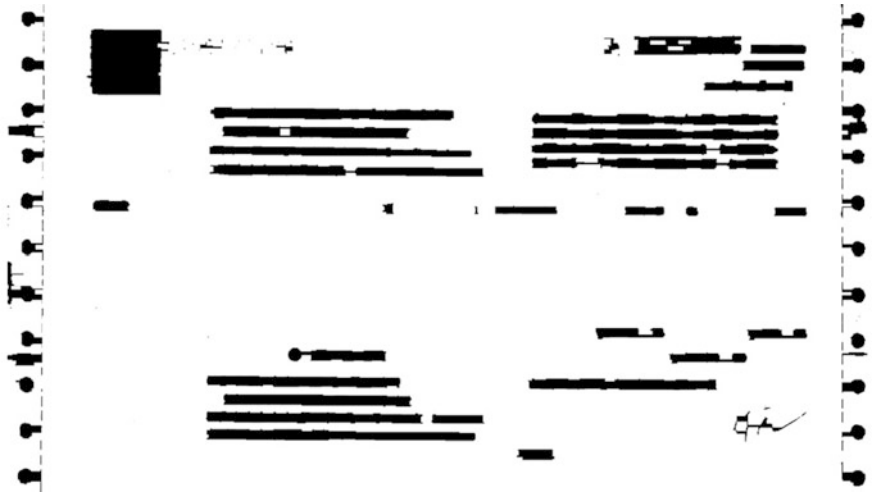


**Fig. 5.** An example of the eroding image


**Fig. 6.** An example of the cut of the enhanced image

**Fig. 7.** An example of the cut of the eroding image

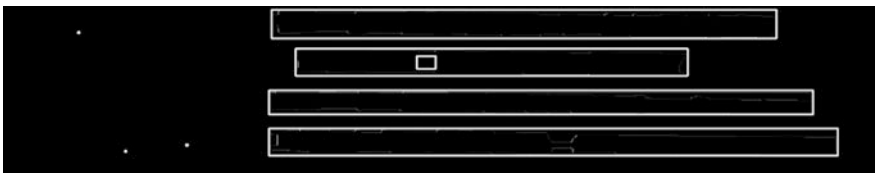


**Fig. 8.** An example of the image with rectangular block

block is stored in a two-dimensional matrix in the order of vertex *Y*, bottom point *Y*, vertex *X*, and bottom point *Y*.

The two-dimensional matrix is sorted in ascending order of vertex *Y*, and the array of vertex *Y* is traversed. If the interval of vertex *Y* is smaller than *k*, the range of *k* is between [5, 30], and the same row is determined. The rectangular block in the same row is sorted by two-dimensional matrix from small to large according to the vertex *X*. The row of rectangular blocks is cut out from left to right, and the image *G* shown in Fig. 9 is extracted by analogy.


**Fig. 9.** An example of the results after segmentation algorithm

## 4 Conclusion and Future Work

This paper proposes a quadratic segmentation algorithm based on image enhancement. The algorithm can achieve effective extraction of machine-playing characters. Through the secondary division of the enhanced invoice image, the pixel information of the machine-played character is finally obtained, and a small block image of the machine-playing character is obtained. The automatic extraction process of invoice information is conducive to the promotion, application, and reality of the invoice intelligent reimbursement system, which has broad application prospects.

This article is only a preliminary study on the extraction of invoice image information. In the future work, we will continue to conduct in-depth studies to achieve better results and more convenient operations. It is also necessary to increase the number of experimental samples and conduct large-scale tests to further increase the effectiveness.

## References

1. Zhu J, He HJ, Tang J, Yan HZ. Tax-control network invoice machine management platform based on socket. In: 2016 2nd IEEE international conference on computer and communications (ICCC); 2016. p. 2343–8.

2. Cai D, Tan M, Cai J. The VAT tax burden warning model and modification based on CTAIS system data. In: 2011 2nd international conference on artificial intelligence, management science and electronic commerce (AIMSEC); 2011. p. 2653–6.
3. Li T, Zhang H. Digital image enhancement system based on MATLAB GUI. In: 2017 8th IEEE international conference on software engineering and service science (ICSESS); 2017. p. 296–9.
4. Kumari A, Thomas PJ, Sahoo SK. Single image fog removal using gamma transformation and median filtering. In: 2014 annual IEEE India conference (INDICON); 2014. p. 1–5.
5. Song SH, et al. Developing and assessing MATLAB exercises for active concept learning. IEEE Trans. Educ. 2018. p. 1–9.
6. Indriani OR, Kusuma EJ, Sari CA, Rachmawanto EH, Setiadi DRIM. Tomatoes classification using K-NN based on GLCM and HSV color space. In: 2017 international conference on innovative and creative information technology (ICITech); 2017. p. 1–6.
7. Palekar RR, Parab SU, Parikh DP, Kamble VN. Real time license plate detection using openCV and tesseract. In: 2017 international conference on communication and signal processing (ICCSP); 2017. p. 2111–5.
8. Sharma T, Kumar S, Yadav N, Sharma K, Bhardwaj P. Air-swipe gesture recognition using OpenCV in Android devices. In: 2017 international conference on algorithms, methodology, models and applications in emerging technologies (ICAMMAET); 2017. p. 1–6.

# Real-Time Vehicle Color Recognition Based on YOLO9000

Xifang Wu[1,2,3(✉)], Songlin Sun[1,2,3], Na Chen[1,2,3], Meixia Fu[1,2,3], and Xiaoying Hou[1,2,3]

[1] School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China
wuxifang_bupt@163.com

[2] Key Laboratory of Trustworthy Distributed Computing and Service (BUPT), Ministry of Education, Beijing University of Posts and Telecommunications, Beijing, China

[3] National Engineering Laboratory for Mobile Network Security, Beijing University of Posts and Telecommunications, Beijing, China

**Abstract.** In this paper, we proposed a real-time automated vehicle color recognition method using you look only once (YOLO)9000 object detection for intelligent transportation system applications in smart city. The workflow in our method contains only one step which achieves recognize vehicle colors from original images. The model proposed is trained and fine tuned for vehicle localization and color recognition so that it can be robust under different conditions (e.g., variations in background and lighting). Targeting a more realistic scenario, we introduce a dataset, called VDCR dataset, which collected on access surveillance. This dataset is comprised up of 5216 original images which include ten common colors of vehicles (white, black, red, blue, gray, golden, brown, green, yellow, and orange). In our proposed dataset, our method achieved the recognition rate of 95.47% and test-time for one image is 74.46 ms.

**Keywords:** Vehicle color recognition · YOLO9000 · Intelligent surveillance

## 1 Introduction

Recently, real-time vehicle information recognition system has attracted much attention in intelligent surveillance, which is essential in the rapid development of smart city. In order to identify and track the targeted vehicle, the license plate, color, and type become the indispensable factors. Colors, as one of the basic attributes of vehicles, play a significant role in vehicle recognition. However, it is still a challenge to recognize the colors of vehicles in original images due to the complex impacts of illuminations, weather conditions, barriers, and image capture qualities.

Yet despite the difficult nature of this task, there has been a recent surge of research interest in attacking the vehicle color recognition problem [1–6]. Prior to large-scale application of deep neural networks, using hand-crafted features combined with classifiers to recognize the vehicle colors was dominant [1–3], which contained color


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_11

histogram, color collection, and color correlogram. The whole workflow consists of manually designing multiple features of vehicle colors and using these features to train the classifiers including KNN, SVM, and ANN. These methods tend to show higher execution speed but weaker generalization and lower recognition accuracy. Rather than use heuristic knowledge to adopt manually designed features, convolutional neural networks (CNN) allow to obtain implicitly vector representation of images [7, 8]. Inspired by the advantage of CNN, the approaches based on deep-learning algorithms make remarkable advances recently. In [4, 5], CNN combined with spatial pyramid strategy was employed to adaptively learn representation which is more effective for vehicle color recognition. These methods extract the feature representation of the images using deep neural networks and then apply the feature in a combination of the traditional classifiers to obtain vehicle colors. Another part of deep-learning-based methods involved the end-to-end architecture, in which the feature extraction and classifier were integrated into a unified deep neural network framework through joint optimization. Rachmadi and Purnama [6] introduced a parallel end-to-end CNN to learn the vehicle color recognition model from big data. These methods based on deep-learning displayed stronger generalization capability and higher recognition performance.

However, vehicle color recognition is greatly affected by the background, light, and weather, which brings a challenge on how to improve the robustness of vehicle color recognition system. In this paper, we propose a new robust real-time vehicle color recognition system based on the YOLO9000 object detection [9]. In the traditional workflow, vehicle location is firstly needed to reduce interference of complex background, and then it comes to color recognition. Yet the method proposed simplifies the workflow of automatic vehicle color recognition, which transforms the two steps of vehicle location and color recognition into one, as shown in Fig. 1. Furthermore, aimed at usual and different real-world scenarios, we specifically introduce VDCR dataset which collected the original vehicle images captured from the cameras in multiple conditions. In our proposed dataset, the end-to-end method we proposed achieved the recognition rate of 95.47% and 74.46 ms per image.

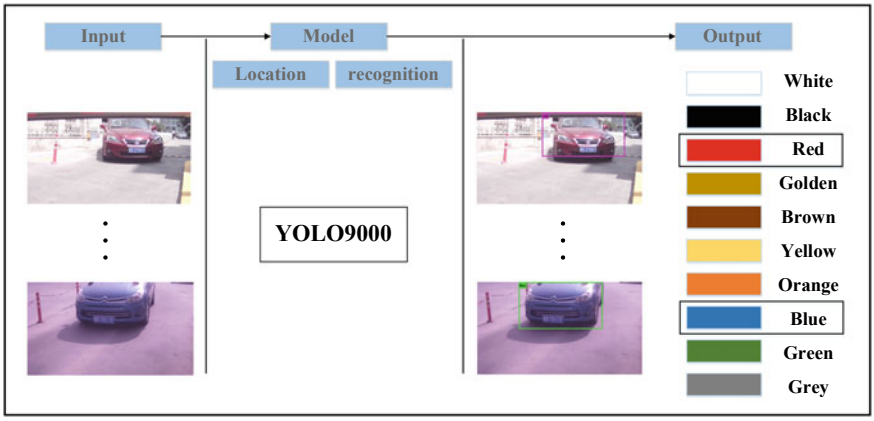



**Fig. 1.** Architecture of real-time automated vehicle color recognition system

## 2 The Proposed Method

Recently, there is a rich body of works concerning vehicle color recognition [1–6]. With the development of artificial intelligence (AI), vehicle color recognition is closely combined with deep learning [4–6]. Besides, a series of algorithms for object detection has been reported [9–11]. In this section, we introduce an end-to-end model for vehicle color recognition using YOLO9000.

As shown in Fig. 2, the original picture is resized to 416 × 416 as the input of the model. Then it is sent to Darknet-19 which has 19 convolutional layers and 5 max-poling layers. To some extent, the Darknet-19 is quite similar to VGG-19 [12]. However, what is different from VGG models is that Darknet-19 uses 3 × 3 filters and 1 × 1 filters in the convolutional layer and double the channels after every pooling layer. Especially, 1 × 1 filters were added between 3 × 3 convolutions in order to compress the feature representation as well as help to eliminate the effect of size on the recognition result. Furthermore, YOLO9000 added batch normalization to significantly improve the convergence of model and eliminate the need for other forms of regularization. After 16 layers of the model, the 26 × 26 × 512 feature map is changed into the 13 × 13 × 256 feature map similar to the identity mappings in ResNet [13]. The model removed full connected layers and instead adding on three 3 × 3 convolutional layers with 1024 filters each followed by a final 1 × 1 convolutional layer for detection. After the last convolutional layer, we obtain 13 × 13 × 75 feature map as the output of the Darknet-19 network. Then the model uses the region proposal network to predict five bounding boxes at each cell of the feature map by using k-means clustering instead of hand-picking anchor boxes.

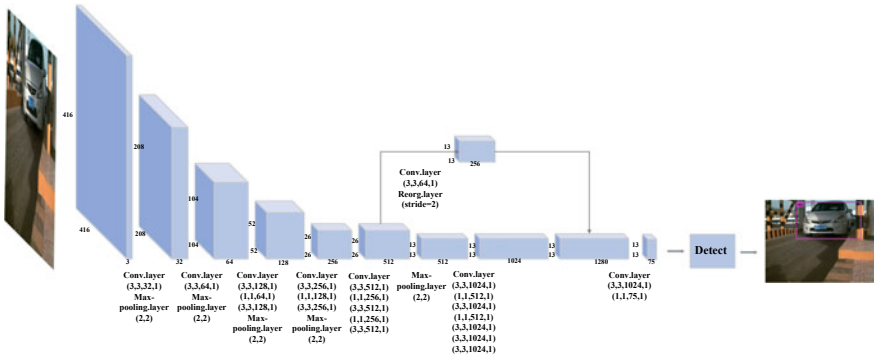



**Fig. 2.** Structure of vehicle color recognition model based on YOLO9000

In bounding box regression and classification, the output dimension for each bounding box is $(N + C)$ in which $N$ is the number of classes and $C$ is the number of the coordinates. The number of classes is ten which depend on the number of the vehicle colors for recognition. And the network predicts each bounding box with five coordinates ($t_x$, $t_y$, $t_w$, $t_h$, $t_o$). Here, $t_x$, $t_y$, $t_w$, and $t_h$ are the offset between the bounding

box and ground-truth bound, $t_o$ is the confidence score of the bounding box. And ($c_x$, $c_y$) is the offset of the cell from the top left corner of the image, ($p_w$, $p_h$) is the width and height of the bounding box. Then the new four coordinates of bounding box correspond to:

$$b_x = \delta(t_x) + c_x \tag{1}$$

$$b_y = \delta(t_y) + c_y \tag{2}$$

$$b_w = p_w e^{t_w} \tag{3}$$

$$b_h = p_h e^{t_h} \tag{4}$$

Once the bounding box can be fixed, the predictions are computed by:

$$p_r(\text{object}) * \text{IOU}(b, \text{object}) = \delta(t_o) \tag{5}$$

Additionally, YOLO9000 randomly chooses a new image dimension size every ten-batches training, which includes {320, 352, …, 608} with the model down-samples by a factor of 32. It improves detection performance of the model for different size of objects.

## 3 Experiments

The robust real-time color recognition system is implemented by NVIDIA Ge-Force GTX 1080Ti with 11 GB memory.

### 3.1 Dataset

The dataset called VDCR dataset contains 5216 images captured from the cameras in different weather (fog, day, rainy, etc.) and different illumination, which pose a great challenge to vehicle color recognition. The images in VDCR dataset was collected in access surveillance scenarios, such as high way entrance, parking lot and residential district entrance. Besides, various vehicle types, such as cars, trucks, vans, and buses, were contained in the dataset in order to ensure the diversity of the vehicles. Especially, the VDCR dataset includes images of sheltered vehicles targeting a more realistic scene. Sample images of the VDCR dataset are shown in Fig. 3.

The images were acquired with different cameras and are available in the Joint Photographic Experts Group (JPEG) format with the size of 1920 × 1080 pixels and 2048 × 1536 pixels. Images obtained with different cameras do not necessarily have the same quality, although they have the same resolution and frame rate. This is due to different camera specifications, such as autofocus, bit rate, focal length, and optical image stabilization.

**Fig. 3.** Sample images of the VDCR dataset. The first three rows show the variety in backgrounds, lighting conditions, as well as vehicle positions and types. The fourth row shows examples of sheltered or distorted vehicles

The VDCR dataset contains 5216 vehicle images, including white, black, gray, red, blue, golden, brown, yellow, green and orange, totally ten color categories. The number of vehicle images for each color is shown in Table 1, in which the minimum proportion is the orange vehicle with 120 images and the maximum proportion is the gray vehicle with 883 images. All of the images in the VDCR dataset come from oblique front images captured from road monitoring system. Some examples of ten vehicle colors in the VDCR dataset are shown in Fig. 4. In the experiments, the dataset is split as follows: 80% for training, 20% for testing.

**Table 1.** Number of images for each color

| Color | White | Black | Gray | Red | Blue | Golden | Brown | Yellow | Green | Orange |
|---|---|---|---|---|---|---|---|---|---|---|
| Number | 786 | 759 | 883 | 707 | 742 | 447 | 335 | 193 | 244 | 120 |

### 3.2 Experiment Result

In the experiments, we input 4173 training images to our end-to-end model, which includes ten vehicle colors. Then we train the model using pre-trained weight file for vehicle location regression and vehicle color classification. As is shown in Table 2, for the test dataset the recognition accuracy of ten vehicle colors is 95.47%. The testing time of YOLO9000 is about 74.46 ms per image.

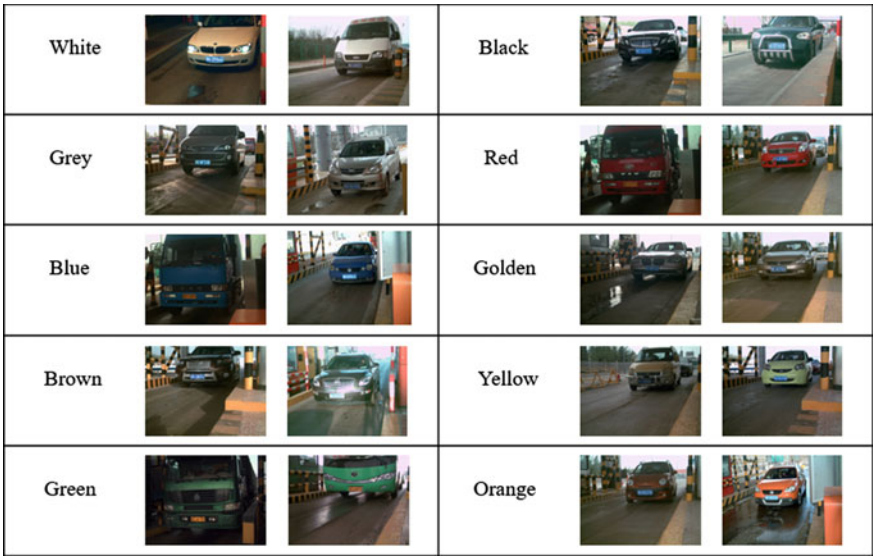



**Fig. 4.** Examples of ten vehicle colors in VDCR dataset

**Table 2.** Experiment result on vehicle color recognition

| Color | White | Black | Gray | Red | Blue | Golden | Brown | Yellow | Green | Orange | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Accuracy | 0.9934 | 0.9802 | 0.9322 | 0.9929 | 0.9729 | 0.8989 | 0.9104 | 1 | 0.8571 | 0.8333 | 0.9547 |

In order to evaluate the result of vehicle location, we use Intersection over Union (IOU) as the standard for the accuracy of the object detection. IOU is the quotient of the overlap area and the union area between the ground-truth boxes and the bounding box. In our test dataset, the average IOU reached 90.64% which shows high performance of vehicle location in our method. As we can see in Fig. 5, the model shows strong robustness for recognizing the correct location and category when the vehicle is incomplete and in the complex background.

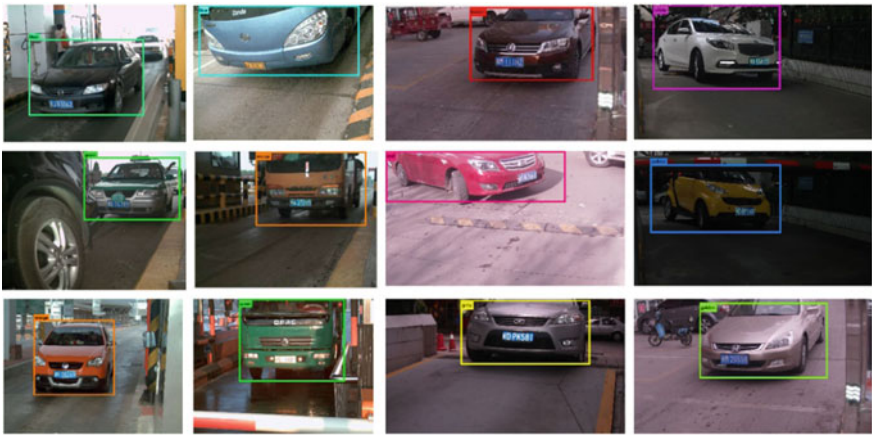


**Fig. 5.** Examples of recognizing the correct vehicle color

## 4 Conclusion

In this paper, a robust real-time end-to-end vehicle color recognition method based on YOLO9000 object detection is proposed, which achieves vehicle location and color recognition in one step. The method proposed reached the 95.47% recognition accuracy and the testing time is merely 74.46 ms per image. We also introduced VDCR dataset for vehicle localization and color recognition that includes 5216 fully annotated images (with ten common vehicle colors) in access surveillance scenarios. The experiment result demonstrates the feasibility and real time of our proposed method.

**Acknowledgments.** This work is supported by National Natural Science Foundation of China (Project 61471066) and the open project fund (No. 201600017) of the National Key Laboratory of Electromagnetic Environment, China.

## References

1. Chen P, Bai X, Liu W. Vehicle color recognition on urban road by feature context. IEEE Trans Intell Transp Syst. 2014;15(5):2340–6.
2. Yang M, Han G, Li X, et al. Vehicle color recognition using monocular camera. In: International conference on wireless communications and signal processing; 2011. p. 1–5.
3. Dule E, Gökmen M, Beratoğlu MS. A convenient feature vector construction for vehicle color recognition. 2010.
4. Hu C, Bai X, Qi L, et al. Vehicle color recognition with spatial pyramid deep learning. IEEE Trans Intell Transp Syst. 2015;16(5):2925–34.
5. Zhang Q, Zhuo L, Li J, et al. Vehicle color recognition using multiple-layer feature representations of lightweight convolutional neural network. Sig Process. 2018;147:146–53.
6. Rachmadi RF, Purnama IKE. Vehicle color recognition using convolutional neural network. 2015.

7. Wang D, Zhang M, Li J, et al. Intelligent constellation diagram analyzer using convolutional neural network-based deep learning. Opt Express. 2017;25(15):17150–66.
8. Wang D, Zhang M, Li Z, et al. Modulation format recognition and OSNR estimation using CNN-based deep learning. IEEE Photonics Technol Lett. 2017;29(19):1667–70.
9. Redmon J, Farhadi A. YOLO9000: better, faster, stronger. 2016.
10. Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector. In: European conference on computer vision. Springer, Cham; 2016. p. 21–37.
11. Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017;39(6):1137–49.
12. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Comput Sci. 2014.
13. He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks. 2016:630–45.

# Improved Video Reconstruction Basing on Single-Pixel Camera By Dual-Fiber Collecting

Linjie Huang(✉), Zhe Zhang, Shaohua Wu, and Junjun Xiao

College of Electronic and Information Engineering, Harbin Institute of Technology, Shenzhen 518055, Guangdong Shenzhen, China
{huanglinjie, zhangzhe}@stu.hit.edu.cn
{hitwush, eiexiao}@hit.edu.cn

**Abstract.** The single-pixel camera is a new architecture of camera proposed in recent years. The difference between a traditional camera and a single-pixel camera is that one image can be reconstructed by acquiring less amount of data with the latter. Most existing single-pixel cameras only collect data for one light path. In this paper, in order to reduce the impact of measurement noise, we adopt a way of dual-fiber acquisition to collect data. We compared the result of traditional single-fiber acquisition with our proposed dual-fibers acquisition. For video reconstruction, we use a dual-scale matrix as the image measurement matrix which can restore images with two different spatial resolutions as needed. We use the low-resolution video as a preview to acquire optical flow, and then we reconstruct a better-quality video by using the optical flow as a restrictive condition. We built an actual single-pixel camera hardware platform based on dual-fiber acquisition, and we show that our high-quality video can be restored by collecting data from our single-pixel camera.

**Keywords:** Single-pixel camera · Dual-fiber acquisition · Dual-scale matrix · Optical flow · Low frame rate

## 1 Introduction

Compressive sensing (CS) [1,2] can discretely sample well-below Nyquist rate by exploiting the signal of the sparse nature, and then the signal is perfectly reconstructed by a nonlinear reconstruction algorithm. CS can take measurements far fewer than the number of pixels of original image, then the image can be fully restored, and CS has an obvious advantage that it can reduce the complexity of the encoder side by increasing the computational complexity at

Linjie Huang and Zhe Zhang—These authors contributed equally to this work.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_12

the decoder. CS has many applications in our life, such as wireless communication and biomedical engineering. Single-pixel camera (SPC) [3] proposed by Rice University is an imaging system that relies on CS. For video reconstruction, we should consider the tradeoff between spatial resolution and temporal resolution [4]. Now there are two different types of cameras [5] for video sensing systems. Spatial multiplexing cameras can use a low-resolution sensor to improve spatial resolution. SPC belongs to spatial multiplexing cameras. Temporal multiplexing cameras can improve time resolution, and its main application is in high-speed imaging. In our experimental platform, we adopt the single-pixel camera which acquires random measurements by using spatial light modulator. For algorithm of video reconstruction, Ref. [6] reconstructed video by frame to frame in 3D wavelets. Reference [7,8] proposed a block-based model for video recovery. Reference [9] used linear dynamic system methods to recover video . Besides, the multi-scale video reconstruction algorithm is proposed, which first reconstructs a low spatial resolution preview that is used to obtain motion estimation, and then reconstructs a better-quality video by exploiting motion estimation as a restrictive condition [10]. Lately, Ref. [11] adopted deep fully connect networks to reconstruct video for compressive sensing. In this paper, we use dual fibers to collect compressive measurements to reduce the impact of measurement noise, and we apply a dual-scale matrix as a sensing matrix which can reconstruct two different spatial resolutions by utilizing a group of data. Finally, the algorithm which we use to reconstruct the video is a total variation regularization [12]. In this paper, we can build a single-pixel camera based on dual-fiber acquisition, and we can restore a high-quality video by improving the traditional single-pixel camera platform and using the property of dual-scale matrix.

In this paper, the overall structure of the article is following. Section 2 introduces compressive sensing theory and dual-scale matrix. Section 3 presents dual-fiber acquisition imaging system and system analysis. Section 4 displays our results and analysis in our experiment. The last section is conclusion.

## 2 Video Compressive Sensing

### 2.1 Compressive Sensing

CS acquiring compressive measurements from $x \in \mathbb{R}^N$ is following.

$$y = \Phi x + e, \tag{1}$$

where $y \in \mathbb{R}^M$ is the compressive measurements, and $\Phi$ is the observation matrix. $e$ represents measurement noise. Restoring the signal $x$ from the compressive measurements is an ill-posed problem since the equations $y = \Phi x + e$ is underdetermined. If we want to recover a K-sparse vector, we need to take at least $M = 2K$ measurements; at the same time, there is a fundamental theory that a stably restoring the vector of $x$ must satisfy Eq. 2.

$$M \sim Klog(N/K) \tag{2}$$

In addition, the observation matrix $\Phi$ meets the restricted isometry property (RIP) [13] with high probability. The following is a SPC sampling model. We can acquire the compressive measurements $y_t$ from a SPC sampling at the time $t = 1, ..., T$ as the following.

$$y = \langle \Phi_t, x_t \rangle + e_t, \tag{3}$$

where $x_t \in \mathbb{R}^N$ is an object scene at the sample instant $t$ , $\Phi_t \in \mathbb{R}^N$ represents the measurement vector, $e_t \in \mathbb{R}$ is the measurement noise, and $T$ is the total number of sampling for an image. If we make the vector be the 2-D scene consists of $n \times n$ spatial pixels, we can get the result of $N = n^2$. We utilize the continuous compressive measurements $W \leq T$ as a group to restore an image. The equation is following.

$$y_{1:W} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_W \end{bmatrix} = \begin{bmatrix} \langle \phi_1, x_1 \rangle + e_1 \\ \langle \phi_2, x_2 \rangle + e_2 \\ \vdots \\ \langle \phi_W, x_W \rangle + e_W \end{bmatrix} \tag{4}$$

### 2.2 Dual-Scale Matrix

Dual-scale matrix is a matrix of two different scales. The construction of a dual-scale matrix satisfies Eq. 5. The elements of matrix $\Phi_1$, $\Phi_2$, $\Phi_3$ are $\{0, 1\}$ since single-pixel camera can only support binary data formats.

$$\Phi_3 = \Phi_1 U + \Phi_2, \tag{5}$$

where $\Phi_3 \in \mathbb{R}^{1024 \times 4096}$ is a dual-scale matrix, and $\Phi_1 \in \mathbb{R}^{1024 \times 1024}$ is a hadamard matrix. $U \in \mathbb{R}^{1024 \times 4096}$ is an upsampling operator. $\Phi_2 \in \mathbb{R}^{1024 \times 4096}$ is an auxiliary matrix, and it satisfies uniform distribution as Eq. 6.

$$\Phi_{2(i,j)} = \begin{cases} 1 & a \geqslant 0.85 \\ 0 & a < 0.85 \end{cases} \qquad a \in U(0, 1), \tag{6}$$

where $a > 0.85$ can change. In order to use the maximize of the auxiliary matrix, the auxiliary matrix $\Phi_2$ is used as a switch function to influence the dual-scale matrix $\Phi_3$. The relationship between them is following. When $\Phi_{2(i,j)} = 0$ $\Phi_3 = \Phi_1 U$. When $\Phi_{2(i,j)} = 1$, $\Phi_1 U = 0$, $\Phi_3 = 1$, and $\Phi_{2(i,j)} = 1$, $\Phi_1 U = 1$, $\Phi_3 = 0$. Finally, we can sum up the relationship between them as shown in Eq. 7.

$$\Phi_3 = mod((\Phi_1 U + \Phi_2), 2) \tag{7}$$

Figure 1 is the process of generating rows of dual-scale matrix.
If we use a dual-scale matrix as a sensing matrix, we can reconstruct two different spatial resolutions from a group of data. Firstly, we can fast restore a low spatial resolution video as a preview by using the low-dimension matrix of the dual-scale matrix, and then we can acquire optical flow by utilizing preview. Finally, we can restore a high-quality video by using the high-dimension matrix of the dual-scale matrix and optical flow [14].

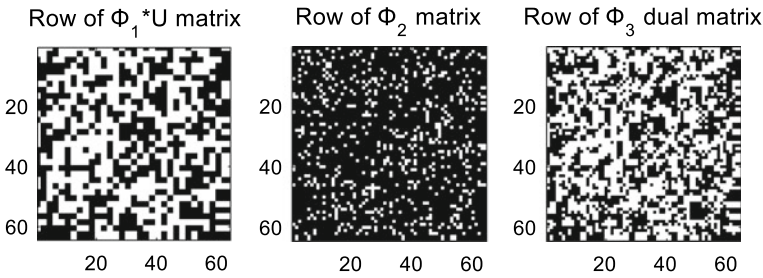



**Fig. 1.** Generating dual-scale matrix masks.

### 2.3 Recovery Algorithm

The process of reconstruction video without optical flow constraints is a convex-optimization problem. It can be solved by Eq. 8.

$$\hat{x} = arg\min_{x} TV(x) \qquad s.t \quad ||y - \Phi x||_2 \leq \epsilon, \tag{8}$$

where the $TV(x)$ represents sparse gradient, the operator $TV(x)$ can be defined as Eq. 9.

$$TV(x) = \sum_{i} \sqrt{(D_x x(i))^2 + (D_y x(i))^2}, \tag{9}$$

where $D_x x$ and $D_y x$ represent the spatial gradient in the $x$ direction and $y$ direction of the 2-D image. Next, the algorithm that we use to reconstruct the video is TVAL3 [12].

If we restore video with optical flow constraints, we need to restore a preview video by TVAL3, then we can acquire optical flow constraints by the preview video. Finally, we can reconstruct a high-quality video by using the optical flow constraints.

**Optical flow.** We can acquire the optical flow between any two frames $\hat{f}^i$ and $\hat{f}^j$ since we have preview mode. The preview video can be reconstructed by the low dimension of dual-scale matrix, and then we can restore a high-quality video by using optical flow as a restriction condition. For estimating optical flow, there are two restriction conditions. (i) For gray scale, two frame images need to have constant brightness, (ii) pixels between two frames of image produce only slight motion. The constraints of optical flow between $\hat{f}^i$ and $\hat{f}^j$ can be written as

$$\hat{f}^i(x, y) = \hat{f}^j(x + u_{x,y}, y + v_{x,y}), \tag{10}$$

where $\hat{f}^i(x, y)$ represents the pixel $(x, y)$ int the image $\hat{f}^i$, and $u_{x,y}$ and $v_{x,y}$ response to the translation of the pixel $(x, y)$ between frame $i$ and frame $j$. In order to acquire robustness, we perform forward and backward optical flow to the reference image.

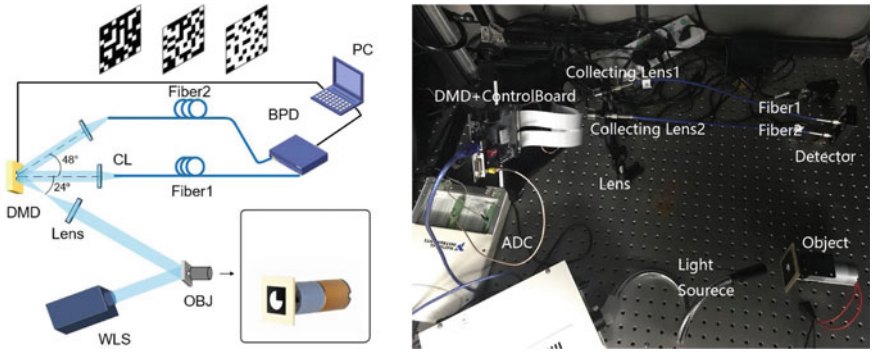
# 3 Dual-Fiber Acquisition Imaging System and System Analysis

## 3.1 Dual-Fiber Acquisition Imaging System

Our SPC hardware is shown in Fig. 2b. It consists of light source, lens, and spatial light modulator (XD-ED01N) which can control digital mirror device (DMD), two collecting lens, two fibers, balanced detector, and data collecting card.

A complete sampling process for a single-pixel camera is as follows. Firstly, light source illuminates at object, and then the object is focused at DMD through the lens. A row of sensing matrix is preloaded on DMD; at the same time, a synchronous pulse signal is sent to data collecting card. Finally, balanced detector collects light intensity by Fiber1 and Fiber2, and data collecting card converts analog voltage to digital voltage.

The DMD consists of $1024 \times 768$ mirrors. Actually, we only use $768 \times 768$ mirrors of DMD. The DMD masks consist of the row of dual-scale matrix $\Phi_3 \in \mathbb{R}^{1024\times4096}$. The column of the dual-scale matrix corresponds to 1024 masks. One mask has 4096 elements which are resized $64 \times 64$ mask, so corresponding to one image spatial resolution is $64\times64$ pixels, and we group $12\times12$ mirrors together on the DMD as one pixel. The rate of DMD masks changes 3000 times per second. For one DMD mask, the data collecting card can collect 16 measurements, and we average 13 of these values as one measurement to reduce measurement noise.




(a) SPC light reflection (b) SPC hardware platform

**Fig. 2. a** Shows the light reflection of DMD and **b** shows the single-pixel camera in our hardware platform.

In our SPC hardware, we exploit the dual fibers to collect the light intensity instead of using lens to collect, and we can acquire a much more light intensity than the lens by reducing the path loss, and then we can reconstruct a better-quality video.

### 3.2 System Analysis

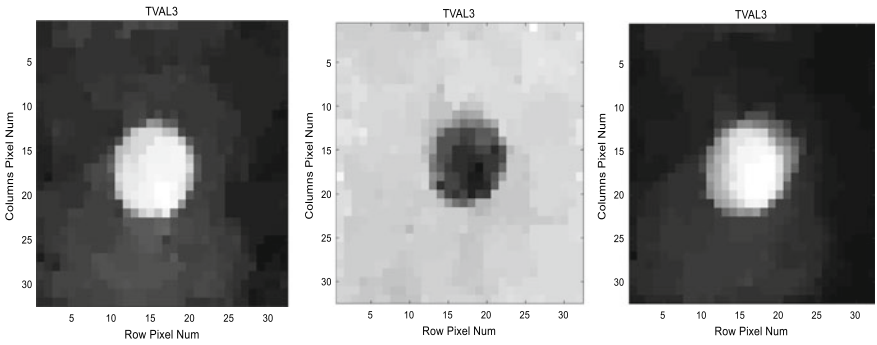
DMD consists of $1024 \times 768$ micro-mirrors. There are two different flip angles that are $+12$ and $-12$ of every micro-mirror so that the reflection has two roads. In order to eliminate the effect of measurement noise, we can respectively set the collection lens at different reflection angles [14]. Reflecting light is following as Fig. 2a. From Fig. 2a, Fiber1 acquires light intensity of the positive reflection, and Fiber2 acquires light intensity of the opposite reflection.

$$\begin{cases} y_1 = \Phi_1 x + e \\ \overline{y}_1 = \overline{\Phi}_1 x + e \end{cases} \qquad y = y_1 - \overline{y}_1 \tag{11}$$

Where $y_1$ acquires compressive measurements of the positive reflection as Fiber1 from Fig. 2a, and $\overline{y}_1$ acquires compressive measurements of the opposite reflection as Fiber2 from Fig. 2a. $y$ is the balanced compressive measurements.

## 4 Results and Discussion

From Sect. 3 instruction, we take three groups of data from our SPC. The first group of data is acquired from Fiber1, and the second group of data is acquired from Fiber2. The last group of data is acquired from balanced detector. Figure 3 is the reconstruction result by using TVAL3 algorithm. From Fig. 3, the spatial resolution of an image is $32 \times 32$, and the compression ratio is 0.3, and then we can find Fig. 3c reconstruction quality better than Fig. 3a and b since Fig. 3c can balance the effect of measurement noise. Figure 3a has better quality than Fig. 3b since Fiber1 can collect more light intensity reflection than Fiber2 collecting. Next, we will display restore video on real data from our single-pixel camera. The video frame rate can be achieved 30 fps in our life, but our reconstruction video cannot achieve 30 fps since our DMD operates at 3000 mirror flips per

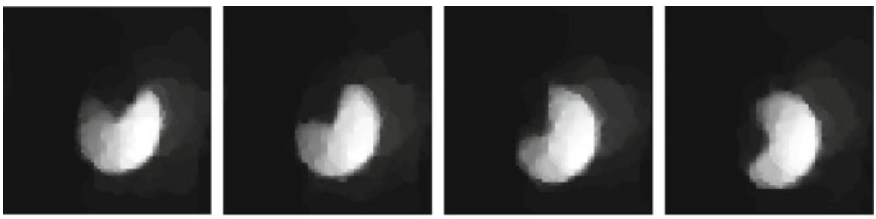



(a) Fiber1 result (b) Fiber2 result (c) BPD result

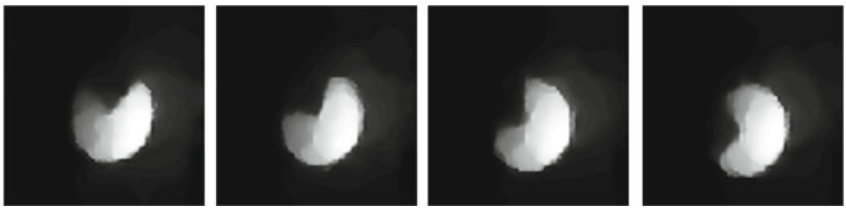
**Fig. 3.** The different reconstruction quality of an image by acquiring the different collecting strategy.

second. If the spatial resolution of an image is $64 \times 64$ and the compression ratio is 0.25, we need to together 1024 compressive measurements as a group to reconstruct an image, and the frame of reconstruction video is about 3 fps. If the spatial resolution of an image is $32 \times 32$ and compression ratio is 0.25, we need to together 256 compressive measurements as a group to reconstruct an image, and the frame of reconstruction video is about 12 fps.

In our SPC platform, we use the dual-scale matrix as the sensing matrix so that we can fast reconstruct the preview video by using the low dimension of the dual-scale matrix, and then we can acquire optical flow by using the preview video, Finally, we can reconstruct a better-quality video by exploiting optical flow as a restriction condition. We take about 34 s to acquire compressive measurements of each video and correspond to M = 102,400 measurements from our hardware platform. Figure 4a shows the reconstruction video without adding optical flow, and Fig. 4b shows the reconstruction video with adding optical flow.



(a) restore video without optical flow constraints



(b) restore video with optical flow constraints

**Fig. 4.** The reconstruction video of two different strategies, where the spatial resolution of the image is $64 \times 64$, and the compression ratio is 0.25. We select four images from the reconstruction images.

From Fig. 4, we can compare the performance of two different reconstruction strategies for a dynamic, moderately rotating target. Restoring video with optical flow has a better quality than reconstructing video without optical flow. Using optical flow motion can make the texture of the reconstructed image more clear and reduce motion blur.

## 5 Conclusion

In this paper, we show our single-pixel camera hardware platform based on dual-fiber acquisition. Our proposed single-pixel camera can greatly reduce the impact of measurement noise so that we can restore a high-quality image. We show how to reconstruct a dual-scale matrix, and we can restore a high-quality video by collecting data from our single-pixel camera and using the dual matrix as a measurement matrix.

**Acknowledgments.** The authors sincerely thanks to the financial support from National Natural Science Foundation of China (61371102, 61001092) and in part by the Shenzhen Municipal Science and Technology Plan under Grant JCYJ20170811160142808, JCYJ20170811154309920.

## References

1. Baraniuk RG. Compressive sensing [lecture notes]. IEEE signal processing magazine. 2007;24(4):118–21.
2. Donoho DL. Compressed sensing. IEEE Transactions on information theory. 2006;52(4):1289–306.
3. D. Takhar, J. N. Laska, M. B. Wakin, et al. "A new compressive imaging camera architecture using optical-domain compression." *Computational Imaging IV*. Vol. 6065. International Society for Optics and Photonics, 2006.
4. Park JY, Michael BW. Multiscale algorithm for reconstructing videos from streaming compressive measurements. Journal of Electronic Imaging. 2013;22(2):021001.
5. Baraniuk RG, Goldstein T, Sankaranarayanan AC, et al. Compressive video sensing: algorithms, architectures, and applications. IEEE Signal Processing Magazine. 2017;34(1):52–66.
6. M. Wakin, J. Laska, M. F. Duarte, et al. "Compressive imaging for video representation and coding." *Picture Coding Symposium*. Vol. 1. 2006.
7. S. Mun, and J. E. Fowler. "Residual reconstruction for block-based compressed sensing of video." *Data Compression Conference (DCC)*, 2011. IEEE, 2011.
8. J. E. Fowler, S. Mun, and E. W. Tramel. "Block-based compressed sensing of images and video." *Foundations and Trends? in Signal Processing* 4.4 (2012): 297-416.
9. A. C. Sankaranarayanan, P. k. Turaga, et al. "Compressive acquisition of linear dynamical systems." *SIAM Journal on Imaging Sciences* 6.4 (2013): 2109-2133.
10. Sankaranarayanan AC, Xu L, Studer C, et al. Video compressive sensing for spatial multiplexing cameras using motion-flow models. SIAM Journal on Imaging Sciences. 2015;8(3):1489–518.
11. J. Kim, L. J. Kwon, and L. K. Mu. "Deeply-recursive convolutional network for image super-resolution." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
12. Li C. An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing. Diss: Rice University; 2010.
13. Candes EJ. The restricted isometry property and its implications for compressed sensing. Comptes rendus mathematique. 2008;346(9-10):589–92.
14. Horn BK, Schunck BG. Determining optical flow. Artificial intelligence. 1981;17(1-3):185–203.

# A Wavelet Based Edge Detection Algorithm

Qingfeng Sun(✉)

Department of Electronic Engineering, Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241000, Anhui, China
sunqife@l26.com

**Abstract.** Edge is in the place where image gray scale changes severely, it contains abundant image information. Image edge detection is a hot and difficult research field. Compare and analyze several classic edge detection method, aim at the advantages and disadvantages, respectively, propose a multi-scale edge detection algorithm based on the wavelet. The simulation shows the algorithm obtains an ideal effect in edge location and noise suppression.

**Keywords:** Edge · Gray scale · Edge detection · Multi-scale · Wavelet transform

## 1 Introduction

The commonly used definition of edge detection is that edge detection is the process of describing the change of image gray scale based on the physical process that causes the change of image's gray scale [1]. The edge of the image contains most of the image information. It also contains the important features of the object while characterizing the basic outline of the object. By extracting the key features of the target object, it can generalize the target object with less information and then use these features to carry out the image calculation and processing. In the image segmentation, matching, fusion, recognition, and computer vision, edge detection can be used as an important link in the implementation of the algorithm and has great research value.

## 2 Classical Edge Detection Algorithm

The edge of an object usually means the appearance of change. As a partial feature of the image, it appears discontinuous. In a sense, edge detection is the calibration of such regional changes. Generally, amplitude and orientation are regarded as two characteristic parameters of image edges, and the overall construction information can be better represented by edge features of images.

Generally speaking, for one-dimensional continuous signals and two-dimensional continuous images, the way of defining edge features is different. For one-dimensional signal, the edge is often defined as the place where the first derivative is maximum or the derivative of the two order is zero. In the two-dimensional image, the gradient information of the gray distribution of the image is commonly used to reflect the change of the image gray level, thus the edge features of the image are expressed.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_13

Usually, edge detection contains four main steps [2]: (1) filtering, smoothing the noise of the image, (2) strengthen and highlight the point of significant change, (3) detection, determine which point is the edge point, (4) positioning and determining the edge position.

Since the theory of edge detection should be used for image processing, many scholars have proposed many algorithms. In many algorithms, differential and optimal operator methods are more common and classic. The differential method uses the first order differential and two order differential to characterize the image gradient, and approximated by the difference operator or the region template convolution. The edge points are selected from the threshold, and then the edge of the image is detected. Based on the classical differential method, the relevant scholars put forward Roberts [3] operator, Sobel operator, Prewitt operator, Laplacian operator [4], etc. The optimal operator method is the optimization and development of differential operators, including LOG [5] operator and Canny operator [6, 7].

For the continuous function $I(x, y)$, at the point $(x, y)$, the gradient is a vector, defined as:

$$\nabla I(x,y) = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \tag{1}$$

The amplitude and direction of the gradient:

$$\begin{aligned} |G(x,y)| &= \sqrt{G_x^2 + G_y^2} \\ \varphi(x,y) &= \arctan \left|\frac{G_x}{G_y}\right| = \arctan \left|\frac{\partial I}{\partial y} / \frac{\partial I}{\partial x}\right| \end{aligned} \tag{2}$$

Generally speaking, all kinds of edge detection operators have their own merits and demerits in specific application scenarios. In the first order differential operator, the Roberts edge detection operator is relatively simple, with high positioning accuracy but poor noise immunity. The Sobel edge detection operator enhances the smoothness of the noise, but increases the computation and the positioning accuracy is not high. In the two order differential operator, the Laplacian operator is beneficial to the removal of the pseudo-edge and the edge positioning because of the two order difference, but doubles the noise [8] at the same time. In view of the sensitivity of Laplacian operator to noise, the LOG operator is proposed by D. Marr and E. Hildreth. First, the filter function is used to smooth the original image, and then the Laplacian operator is used to detect the edge. The Canny operator also introduces the Gauss function, smoothing the filter to deal with the problem of losing the edge information and not preserving a large number of image details [9]. It can be seen that the traditional edge detection algorithms have some problems in different applications because of the diversity of images and their own limitations. It is the two major directions for many researchers to improve the existing algorithms, integrate new features, or introduce new ideas and propose new algorithms. At present, with the in-depth application of fuzzy mathematics, neural network, and multi-scale theory, the development of edge detection algorithm is constantly promoted.

# 3 Research on Wavelet Transform and Its Edge Detection Algorithm

As a representative of multi-scale theory, wavelet transform is widely used in image processing because of its completeness of mathematical description. In the image edge detection, the multi-scale property of the image can be well combined with the traditional operator. It has achieved good results and has great research value.

## 3.1 Wavelet Transform

The wavelet transform has good time frequency characteristics. Through the scale function and translation operation, the high-frequency components and low-frequency components of the image can be separated, and the image processing speed and diversity can be improved by image decomposition and reconstruction. Wavelet decomposition of images at different scales will form a Pyramid structure. The wavelet decomposition of the image is not the deletion of the original image information, but the low-frequency details and the high-frequency information, respectively, in order to facilitate the processing and utilization of the related information. In the framework of multiresolution analysis, Mallat's fast wavelet decomposition algorithm [10] is often adopted. As shown in Fig. 1a, 4 new images are generated from the original image in each layer of the wavelet decomposition, the size of the original image is 1/4, the low-frequency information of the original image is stored in the LL, for the next layer of decomposition, the HH is stored in the high-frequency information, the LH and HL store the edge features of the horizontal and vertical direction, respectively. Figure 1b is the actual image. The two layer wavelet decomposition.

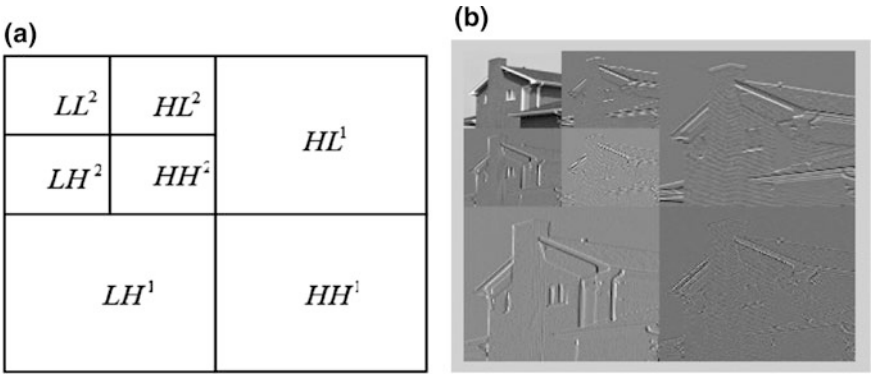



**Fig. 1.** Image wavelet decomposition

## 3.2 Algorithm of Edge Detection Based on Wavelet Transform

It is generally believed that the discontinuous change of the gray level in the image can be regarded as the edge, which is a high-frequency signal, which is similar to the noise

signal. When the image contains a lot of noise, the traditional edge detection method based on differential operation is more sensitive to the high-frequency signal, and it is difficult to accurately distinguish the edge and noise containing a large number of high-frequency components. Therefore, it is difficult to accurately detect edge features in noisy environment, but noise resistance is an important performance guarantee for image processing and related links, and the adaptability of noise directly affects the final effect of the algorithm. Based on this, Mallat and other scholars use the good time frequency analysis ability of wavelet transform to detect the high-frequency mutation of the local area of the image and carry on the processing, and then apply the wavelet transform to the edge detection.

Since wavelet transform is applied to edge detection, there are roughly two main ideas. One is to combine wavelet transform to improve the classical edge detection operator; the other is to improve the wavelet decomposition algorithm to improve the edge detection by incorporating new properties or constructing a new wavelet structure. Literature [9, 11] is a new method based on classical Canny edge detection operator combined with wavelet transform. Duan [9] first improves the Canny operator by replacing traditional Gauss filter by bilateral filtering and then improves the image high-frequency details by wavelet transform to realize edge detection. In document [11], Canny algorithm and wavelet modulus maximum method are used to extract edge, and then two edge images are weighted and optimized to get the final edge image. In document [12, 13], by combining the interpolation wavelet sampling theory and the omnidirectional wavelet method, the traditional Mallat wavelet transform edge detection algorithm is improved, and a better and more adaptive edge detection algorithm is obtained. Many studies have shown that the introduction of wavelet transform has good effect on improving the traditional edge detection operator, improving the adaptability of the scale and the noise resistance. It is of practical significance to study the edge detection algorithm based on the wavelet transform transformation.

### 3.3 Algorithm in this Paper

Using the multi-scale characteristics of the wavelet transform and its characteristics of approximate and generalization in image decomposition, it is applied to the edge detection of the image. The wavelet transform is applied on the different scales, and the high-frequency information under the large-scale decomposition is suppressed and the noise is suppressed. The results are obtained according to the gradient direction modulus maximum and threshold. On the edge of each scale, we should grasp the details and the general picture and then get the edge image. The following steps are as follows:

1. For image decomposition using wavelet transform, four coefficients will be obtained at different scales, using $A^i$, $HL^i$, $HV^i$, and $HD^i$, respectively. Among them, $i$ represents scale factor, A is approximate coefficient, and HL, HV, and HD, respectively, represent horizontal, vertical, and diagonal direction high-frequency coefficients.

2, The horizontal and vertical high-frequency coefficients (HL, HV) of each scale are processed to get their respective modulus images and amplitude images.
3, Seek the local maximum value of image and amplitude image, select the threshold, and get the edge points.
4, Connect the edge points, get the edge features of each scale, reconstruct the series of edge images, and determine the final image edge.

## 4 Experimental Results

The experimental results of this paper are simulated under MATLAB2013b. First, the image of the roof of Fig. 3 is selected and the images of salt and pepper noise and Gauss white noise are added, respectively, as experimental images. The experiments are carried out with Roberts operator, Sobel operator, Canny operator, LOG operator and the algorithm proposed in this paper, and the results are compared and analyzed. Then the algorithm is applied to other images, and the adaptability of the experiment is tested.

As shown in Fig. 2, the left side of figure (a) is the original image, the middle is the image after adding the salt and pepper noise with $\delta = 0.01$, and the image on the right is added $\delta = 0.01$ Gauss noise, and the following graphs are the results of the corresponding method in the corresponding case, and the graph (b) is the Roberts operator detection result, which is sensitive to the noise and some edges are missing. In graph (c), the Sobel operator has some improvements to the detection of the edge position and direction relative to the Roberts operator and is more adaptable to the noise of the image, but it is still not enough for the accuracy of the edge location; in figure (d), the Canny operator is more noisy when the noise distribution is dense; in graph (e), the LOG operator can smooth. Some noise is removed, the edge detection effect is better, but the edge is still missing, and the noise immunity needs to be strengthened. Figure (f) is an edge detection algorithm based on wavelet transform, which makes full use of the multi-scale characteristics of wavelet, and has achieved good results in the anti-noise and accurate positioning.

As shown in Fig. 3, the experimental results of the wavelet detection algorithm on the other two test charts show the original image, the image after adding the $\delta = 0.02$ Gauss noise and the corresponding wavelet edge detection images. The results show that the wavelet transform edge detection algorithm proposed in this paper can adapt to other test images and has a certain adaptability and good noise immunity.

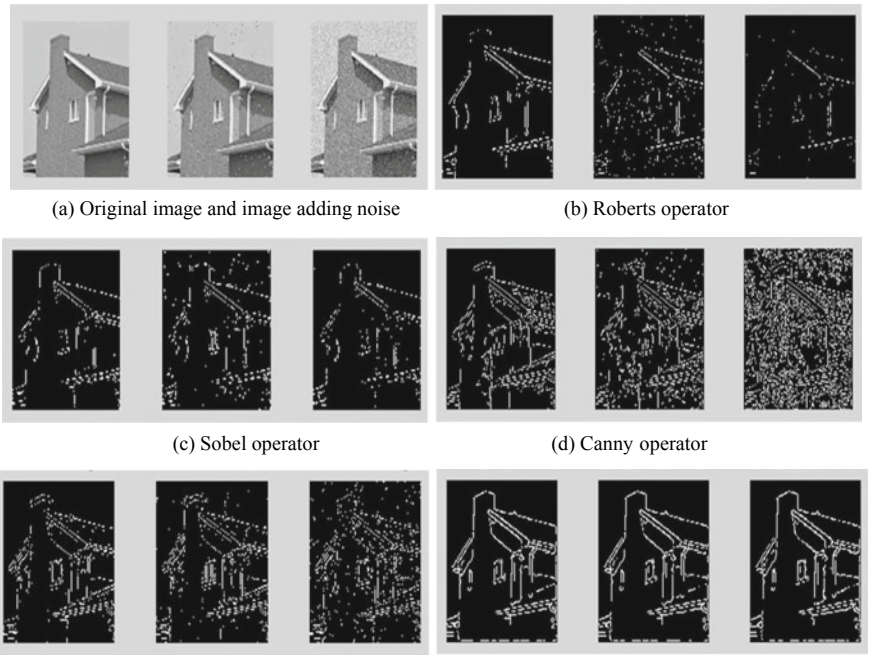



**Fig. 2.** Roof test image and its edge detection results

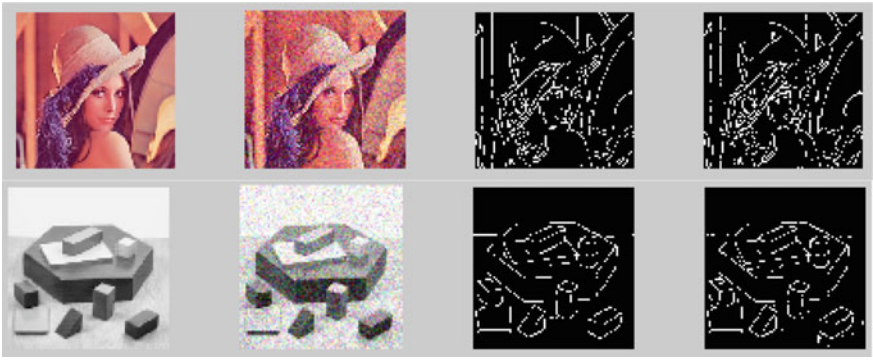


**Fig. 3.** Other test images and their edge detection result

## 5 Conclusion

The edge detection based on wavelet transform makes use of the adaptive characteristics of the noise and positioning accuracy in large and small scales after the decomposition of the wavelet image, which solves the contradiction between the edge location and the noise, and has a certain improvement effect compared to the traditional

edge detection operator. The following aspects can be further studied in scale space theory, adaptive threshold, and in combination with practical application of image edge detection in production practice.

**Acknowledgments.** Project found: (1) Young teachers development and support program of Anhui Technical College of Mechanical and Electrical Engineering (project number: 2015yjzr028); (2) Anhui Province Quality Engineering Project "Exploration and Practice of Innovative and Entrepreneurial Talents Training Mechanism for Applied Electronic Technology Specialty in Higher Vocational Colleges" (project number: 2016jyxm0196); (3) Anhui Quality Engineering Project "Industrial Robot Virtual Simulation Experimental Teaching Center" (project number: 2016xnzx007).

## References

1. Torre V, Poggio T. On edge detection.IEEE Trans Pattern Anal Mach Intell. 1986;8:147–61.
2. Gao Z, Zhang T, Qu Y. Progress in image edge detection. Sci Technol Rev. 2010;28(20): 112–7.
3. Roberts LG. Machine perception of three-dimension solids. Optical and electro-optimal information processing. Cambridge: MIT Press; 1965. p. 157–97.
4. Gobzalez RC, Woods RE. Digital image processing. 2nd ed. Bei Jing: Publishing House of Electronics Industry; 2007.
5. Mary D, Hildreth E. Theory of edge detection. Proc Roy Soc. 1980;B(27):187–217.
6. Deriche R. Using Canny's criteria to derive a recursively implemented optimal edge detection.Int J Comput Vis. 1987;12(1):167–87.
7. Canny JFA. Computational approach to edge detection.IEEE Trans Pattern Anal Mach Intell. 1987;8(6):679–98.
8. Duan R, Li Q, Li Y. A survey of image edge detection methods. Opt Technol. 2005;03:415–9.
9. Duan H, Shao H, Zhang S, et al. An improved algorithm of image edge detection based on Canny operator. J Shanghai Jiao Tong Univ. 2016;12:1861–5.
10. Mallat S, Huang WL. Singularity detection and processing with wavelets. IEEE Trans Inf Theor. 1992;38(2):617–43.
11. Zhao J, Yang H. Edge detection operator integrating Canny operator and wavelet transform. Comput Simul. 2017;6:277–80.
12. Zhang Z, Zheng X, Lan Jing C. Image edge detection based on interpolated wavelet pyramid decomposition algorithm. Comput Sci. 2017;6:164–8.
13. Li Z, Li R, Sakai O. Image edge detection algorithm based on omnidirectional wavelet transform. Electron J. 2012;12:2451–5.

# An Improved Harris Corner Detection Algorithm

Qingfeng Sun(✉)

Department of Electronic Engineering, Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241000, Anhui, China
sunqife@l26.com

**Abstract.** The traditional Harris corner detection algorithm is sensitive to noise, and Corner is prone to drift at different image resolution. Combined with the multi-scale features of wavelet transform, propose a corner detection algorithm based on the wavelet transform. The algorithm maintains the advantages of Harris corner detection algorithm in image scaling, rotation or gray scale change, improves its disadvantage of scaleinvariance, and has strong anti-noise and real-time performance. It has good anti-noise and real-time performance.

**Keywords:** Corner detection · Wavelet transform · Multi-scale

## 1 Introduction

Corner is an important local feature of images, which contains a wealth of information. At present, there are many statements about the definition of corner points. Generally speaking, its foothold is basically the two aspects of the sharp change point of two-dimensional image brightness and the point of maximum curvature on the edge curve of the image. While preserving the important features of the image, the corner point has rotation invariance and is almost unaffected by the illumination conditions [1]. It can effectively reduce the amount of information data, effectively improve the speed of the calculation, and is beneficial to the reliable matching of the image, making the real-time processing possible. It plays a very important role in the field of computer vision, such as 3D scene reconstruction, motion estimation, target tracking, target recognition, image registration, and matching.

## 2 Harris Corner Detection

Harris corner detection method [2] is proposed by Harris and Stephens. The basic theory of the classical Harris corner detection method is based on the Moravec operator [3], which introduces the autocorrelation function theory in signal processing. It combines the corner detection with the local autocorrelation function of the image and determines whether the detection point is the corner point by the eigenvalue analysis.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_14

In the original Harris corner detection algorithm, the corner response function used is:

$$R = \mathrm{Det} - KTr^2 \tag{1}$$

Only when the $R$ value is greater than the threshold and the local maximum is found in the eight directions around, this point is considered to be the corner point. The $K$ in the response function is a constant factor, which is an empirical value between 0.03 and 0.15. When the classical Harris corner detection method is used, the selection of $K$ often brings some unnecessary errors. In the subsequent improvement, the response function is generally not used.

The corner detection method of Harris operator is a better image feature point extraction algorithm at present. The computation is relatively small. When the image has rotation, view point change, or gray change, the angle point extraction effect is still good. However, it is sensitive to noise, and it is easy to extract a large number of false corner points under the interference of noise. In addition, the traditional Harris corner detection operator does not have scale invariance, and the resolution of the image is different.

## 3 Corner Detection Algorithm Based on Wavelet Transform

### 3.1 Wavelet Transform

The theory of wavelet analysis originated in the early twentieth century. Haar proposed the Haar wavelet normalized orthogonal basis. In the 1980s, Morlet and Mcyer proposed the concept of wavelet analysis and the basic theoretical framework of multiresolution analysis. Then, Mallat established the fast wavelet transform method and proposed the theory [3, 4] of multiresolution signal decomposition, which realized the transformation of wavelet analysis from mathematical expression to technology. With the development of Daubechies and Swelden, wavelet transform has been widely applied.

As shown in Fig. 1, at each level of the wavelet decomposition, 4 new images, which are 1/4 of the original image, are generated from the original image. These new images are LL, LH, HL, and HH [5], respectively. Among them, the LL image preserves the low-frequency information of the original image, and the HH contains high-frequency information and typical noise. LH and HL contain both the horizontal edge features and the vertical edge features, respectively. In the wavelet decomposition of images, only LL images are used to produce the next level decomposition.

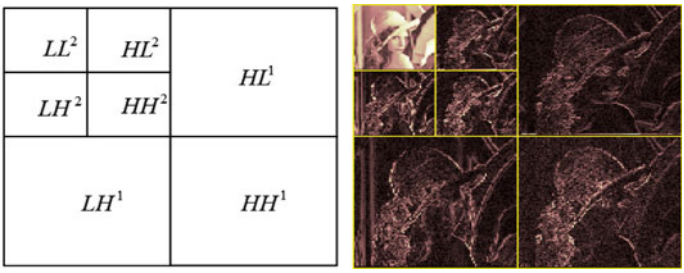



**Fig. 1.** Image wavelet decomposition

### 3.2 Algorithm of Corner Detection Based on Wavelet Transform

Considering the shortcomings of the traditional Harris corner detection operator in the field of anti-noise, a corner detection operator based on wavelet transform is proposed. The steps are as follows:

1. To decompose the image with stationary wavelet transform [6], four series of coefficients are obtained on each decomposition layer, which are the high-frequency coefficients of the approximate coefficient and the other three directions, respectively.
2, Build an autocorrelation matrix based on the high-frequency coefficients of horizontal direction ($I_x$), vertical direction ($I_y$), and diagonal direction ($I_{xy}$) obtained by wavelet decomposition.
3, The improved Harris corner response function is used to extract the feature points from the image obtained from the previous step.

$$R = \left(I_x^2 * I_y^2 - I_{xy}^2\right) \Big/ \left(I_x^2 + I_y^2 + \text{eps}\right) \tag{2}$$

## 4 Experimental Results

The experimental results of this paper are simulated under MATLAB2013b. The standard Lena map and building block test chart of 256 * 256 pixels are selected as experimental images. Combined with the superposition of Gauss white noise and scale, the experiment is carried out with this algorithm in many cases, and the comparison analysis is made with the traditional Harris corner detection algorithm.

The left of Fig. 2 is Lena image, and the Lena image of Gauss white noise is superimposed on the right. Figure 3 is the corner point obtained by the traditional Harris corner detection algorithm in two cases of Fig. 2. It can be seen that the traditional Harris corner detection algorithm has more corners detected, but it is sensitive to noise. In addition, in actual simulation, because of its relatively large amount of computation, it takes about 0.5 s. Figure 4 is the application of this algorithm in Fig. 2. In the two images, the number of corners obtained before and after the addition of noise

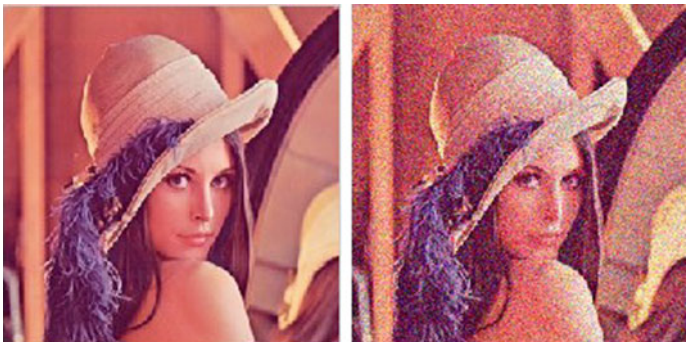


**Fig. 2.** Original image

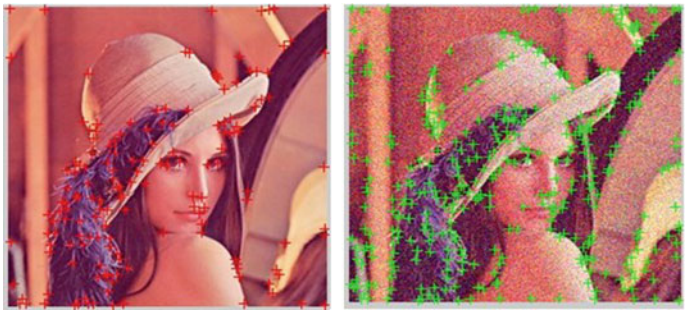
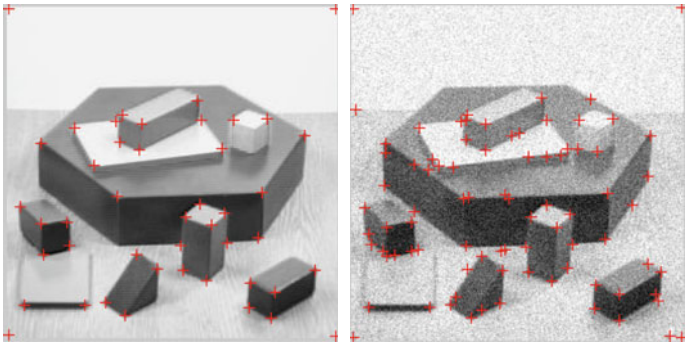


**Fig. 3.** Traditional Harris corner detection

**Fig. 4.** Algorithm in this paper

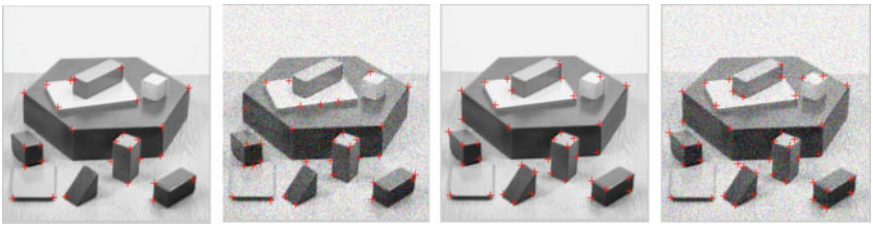
is basically equal, and its corner location is basically consistent and the stability is good. In addition, the time of extracting corner points is about 0.2 s. Based on the corner detection algorithm, this method has better matching effect in subsequent image registration.

Figure 5 is the result of extracting corner points from block diagrams under different conditions using traditional algorithms. Figure 6 is the result of using this algorithm to extract corners of the block diagram after adding scale conditions, in which figures (a) and (c) are the corners detected at different scales (scale = 1, 2) when no noise is added. The results show that, at different scales, the position and number of the corners are approximately equal, which obviously improves the disadvantage of the traditional Harris corner detection that does not have the scale invariance; figures (b) and (d) are the corner points detected on the same scale after adding Gauss noise (=0.01), although they are still disturbed by the noise, but compared with the original algorithm, it is resistant to the original algorithm. The noise ability is obviously improved.



(a) No noise (b) $\delta$ =0.01

**Fig. 5.** Corner points detected by the traditional Harris corner detection method



(a) scale=1 (b) scale=1, $\delta$ =0.01 (c) scale=2 (d) scale=1, $\delta$ =0.01

**Fig. 6.** Corner points detected at different scales using this method

## 5 Conclusion

The multi-scale characteristics of wavelet transform can obtain the general features of images at large scales, and the details of the images can be obtained at small scales. By approximation of detail and general appearance, noise suppression in large scale and accurate location of corner points in small scale, the angle detection algorithm of

wavelet transform is proposed in this paper, which balances the contradiction between noise and location, and the effect of corner detection is ideal. However, the selection of wavelet and the setting of some parameters are related to the final detection results. The next step, we should study in depth and improve the algorithm.

**Acknowledgments.** Project found: (1) Young teachers development and support program of Anhui Technical College of Mechanical and Electrical Engineering (project number: 2015yjzr028); (2) Anhui Quality Engineering Project "Industrial Robot Virtual Simulation Experimental Teaching Center" (project number:2016xnzx007); (3) nhui Province Quality Engineering Project "Exploration and Practice of Innovative and Entrepreneurial Talents Training Mechanism for Applied Electronic Technology Specialty in Higher Vocational Colleges" (project number:2016jyxm0196); (4) Anhui Quality Engineering Project (MOOC) "Off-line Programming of Industrial Robots" (project number:2017 mooc091).

## References

1. Schimid C, Mohr R, Bauckhage C. Evaluation of interest point detectors. Int J Comput Vis. 2000;37(2):151–72.
2. Harris C, Stephens MJ. A combined corner and edge detector. In: 4th Alvey vision conference; 1988. p. 147–52.
3. Moravec HP. Towards automatic visual obstacle avoidance. Proceedings of the 5th International Joint Conference on Artificial Intelligence.1977;584–596.
4. Mallat SG. A theory for multi-resolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell. 1989;11:673–94.
5. Gobzalez RC, Woods RE. Digital image processing. 2nd ed. Bei Jing: Publishing House of Electronics Industry; 2007.
6. Mallat S, Huang WL. Singularity detection and processing with wavelets. IEEE Trans Inf Theor. 1992;38(2):617–43.

# Color Correction Method for Digital Camera Based on Variable-Exponent Polynomial Regression

Yingjie Zhou[1], Kun Gao[1(✉)], Yue Guo[1], Zeyang Dou[1], Haobo Cheng[1], and Zhuoyi Chen[2]

[1] School of Optoelectronics, Beijing Institute of Technology, Beijing 100081, China
594836904@qq.com, gaokun@bit.edu.cn
[2] Beijing Institute of Spacecraft System Engineering, Beijing 100094, China
blacksnow54ll@163.com

**Abstract.** Subject to the response uniformity of photoelectric sensors, the captured raw images always have serious chroma distortions. How to determine the mapping matrix between *RGB* and *XYZ* color spaces is important for the color distortion correction. However, the commonly used algorithms cannot give consideration to the precision and the adaptability. A more reasonable mapping algorithm based on variable-exponent polynomial regression is proposed to evaluate the mapping matrix coefficients. Variable-exponent regularization with the $L_\rho$-norm ($1 < \rho < 2$) combines the features of lasso regression and ridge regression methods, owning both the sparsity and smoothing properties. The optimal solution for the variable-exponent regularization is given using lagged fix-point iteration method. Data from the standard color correction experiments are used to test the variable-exponent, lasso, ridge, and least-squares regression algorithms with different polynomial regression models. The results demonstrate that the proposed algorithm has the best performance.

**Keywords:** Color correction · Polynomial regression · Regularization · Variable-exponent

## 1 Introduction

With advances in digital imaging technology, accurate color acquisition and display have been given a much higher priority. Human visual system has amazing color constancy properties, unlike imaging devices that are vulnerable to the influence of illumination or sensor responses. To obtain the same colors as those perceived by our visual system directly, reasonable color calibration processing is vitally important.

The color difference caused by camera sensors is mainly due to the deviation of spectral sensitivities to *RGB* from the cones of the human eye [1]. To effectively reduce the difference, a large number of color correction algorithms based on learning for


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_15

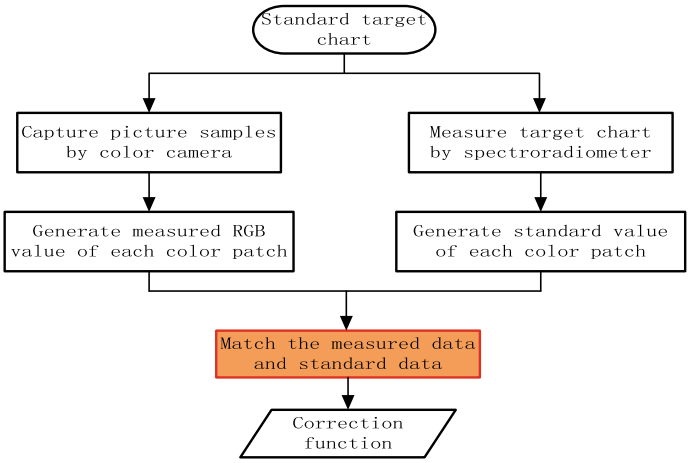



**Fig. 1.** Color correction process

digital cameras have been introduced. Most of these color correction algorithms rely on the principle shown in Fig. 1. One typical example is CIEXYZ, which is determined only by the reflectance property of the observed object [2, 3].

These algorithms specifically include polynomial regression methods based on least squares with different polynomial orders, 3D lookup tables with interpolation methods, and neural network methods, etc. However, most algorithms are either time-consuming or unable to derive a unique model [4–6]. Polynomial regression methods require only that the number of test samples be greater than the number of terms of the polynomial model and have a simple principle with low computational complexity [7]. Therefore, these methods are always used in colorimetric characterization and color correction for image acquisition devices.

Linear polynomial regression is the earliest used regression method for color correction. Although it owns a better adaptability to illumination, it has a high error. To reduce the error, high-order polynomial was developed to color correction. It is more powerful but not stable. Then, Graham D. Finlayson and Michal Mackiewicz proposed root-polynomial regression, which can scale with exposure [8].

Theoretically, the relationship between CIE colorimetric values $X$, $Y$, $Z$ and camera response values $R$, $G$, $B$ is linear as the relative spectral power distribution is definite. While in real scene, it tends to be some nonstandard or separated models. High-order polynomial brings high precision; meanwhile, it brings the danger of overfitting and lowers the adaptability of the model.

To compensate the weakness, variable-exponent polynomial regression algorithm is presented herein. The algorithm utilizes variable-exponent regularization term to constrain the loss function, which possesses prominent fitness to complex data relationship like color space mapping.

## 2 Variable-Exponent Polynomial Regression

The specific color correction algorithm workflow in this paper is illustrated in Fig. 2. The common-used linear polynomial regression, quadratic polynomial regression, and root-polynomial regression were chosen as the original fitting models. On the basis of these models, $L_1$ regularization, $L_2$ regularization, and variable-exponent regularization were introduced to calculate the mapping relationship.

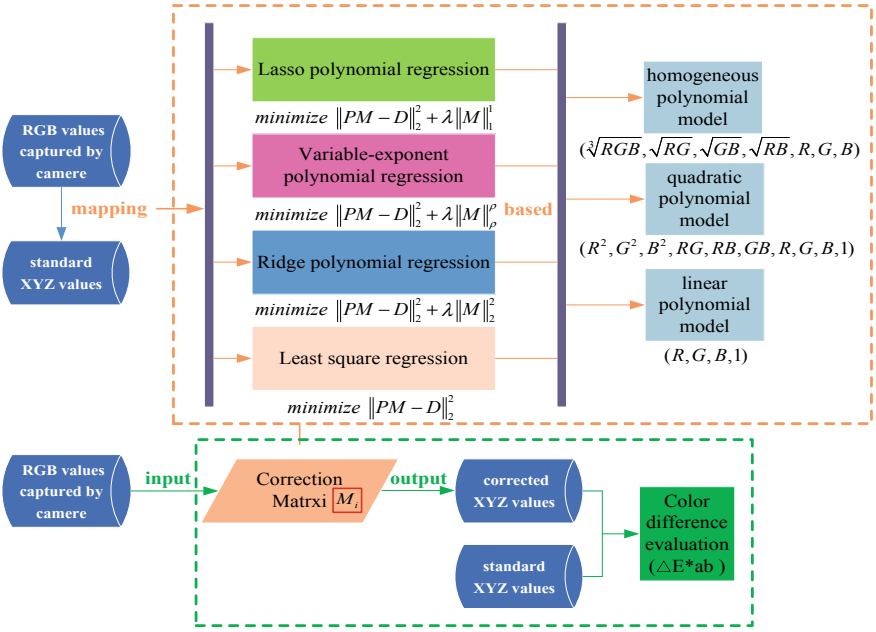



**Fig. 2.** Workflow of the color correction algorithm

All the models shared the same test dataset. The input variable was the combination of measured $R$, $G$, and $B$ values. The target variable was the standard colorimetric values $X$, $Y$, $Z$, denote as $\mathbf{D}$. It can be represented by Eq. (1):

$$\mathbf{P}_i\mathbf{M}_i = \mathbf{D}, \quad i = 1, 2, 3 \tag{1}$$

where $\mathbf{M}$ is the coefficient vector that needs to be solved. The matrix $\mathbf{P}_i$ represents different polynomial regression model.

### 2.1 The Evolution of Variable-Exponent Regularization

For the sake of fitting precision, some innovations were introduced to the conventional polynomial regression, such as the $L_1$-norm penalty and the $L_2$-norm penalty, to yield

lasso regression and ridge regression [3]. Taking account of the color correction dataset, $L_1$-penalized regression and $L_2$-penalized regression can be described as:

$$\min\|\mathbf{PM}-\mathbf{D}\|_2^2 \ + \ \lambda\|\mathbf{M}\|_1^1 \tag{2}$$

$$\min\|\mathbf{PM}-\mathbf{D}\|_2^2 \ + \ \lambda\|\mathbf{M}\|_2^2 \tag{3}$$

$L_1$ regularization (lasso) yields a convex problem, which is an attractive trait for computation [9]. $L_2$ regularization has the intuitive interpretation of heavily penalizing peaky weight vectors. It has the smoothing property, but the final weight vectors usually tend to be diffuse, small numbers [10]. In practice, if explicit feature selection is not a major concern, $L_2$ regularization can be expected to give superior performance to $L_1$ [11]. To combine the advantages of both lasso regression and ridge regression, variable-exponent regularization is proposed in this paper to choose regular method with self-adapting, as defined in Eq. (4):

$$E(\mathbf{M}) = \min\|\mathbf{PM}-\mathbf{D}\|_2^2 \ + \ \lambda\|\mathbf{M}\|_\rho^\rho \tag{4}$$

where

$$\rho(W_j) = 1+\frac{1}{1+\frac{k}{W_j^2}} \tag{5}$$

where $\lambda$ is the coefficient of regularization term, deciding the impact degree that the regularization brings. $\rho$ is a function of the final desirable coefficient $W_j, j \in (1,n)$, $n$ is the number of polynomial terms, $k$ is a constant, which can be adjusted with different datasets.

The $L_\rho$ penalty function has a self-adaptation characteristic, in which the $L_\rho$-norm $(1<\rho<2)$ adjusts the $\rho$ value according to the coefficient $W_j$. When the coefficient is large, the $L_\rho$-norm tends toward the $L_2$-norm, and when the coefficient is small, the $L_\rho$-norm tends toward the $L_1$-norm. With the proper selection of penalty function, the optimal fitting model for the dataset can be defined.

### 2.2 The Solution to Variable-Exponent Regularization

To obtain the desirable solution, the derivative of the loss function in Eq. (4) was calculated. The result is shown as follow:

$$\partial_{\mathbf{M}} = \mathbf{P}^T(\mathbf{PM}-\mathbf{D}) + \lambda\rho(W_j)|\mathbf{M}|^{\rho(W_j)-2}\mathbf{M} \tag{6}$$

The optimal solution can be acquired by setting the above gradient to zero and solving the nonlinear equation system (6) using lagged fix-point iteration method [12]. Specifically, the coefficient $\mathbf{M}$ in the denominator is lagged behind one iteration, making the nonlinear equation system change into a linear equation system. The coefficient vector $\mathbf{M}$ then can be formulated to Eq. (7).

where

$$\mathbf{M} = (\mathbf{P}^T\mathbf{P} + \lambda\mathbf{diag})^{-1}\mathbf{P}^T\mathbf{D} \tag{7}$$

$$\mathbf{diag} = \begin{pmatrix} \rho(W_j)|M_1|^{\rho(W_j)-2} & & & \\ & \rho(W_j)|M_2|^{\rho(W_j)-2} & & \\ & & \ddots & \\ & & & \rho(W_j)|M_j|^{\rho(W_j)-2} \end{pmatrix} \tag{8}$$

In practice, one critical problem is that we don't know the real coefficient $W_j$. To fix that question, one alternative way is to adopt the idea that dynamic approximation to the real coefficient [13]. Gauss-Seidel iteration method can realize that validly. Firstly, the unknown coefficient $\mathbf{W}$ is given an initial value $\mathbf{M}^0$. The vector $\mathbf{M}$ is continually updated by the last consequence. Then repeat updating the vector $\mathbf{M}$ until converge. During the iteration, the coefficient would gradually get closer to the real coefficient vector $\mathbf{W}$.

The algorithm process is shown in Table 1. As there are two different parameters for different regression models, some empirical values are recommended here. Specifically, for linear polynomial, $k = 30$, $\lambda = 4.7$; for quadratic polynomial, $k = 0.017$, $\lambda = 3$; for root-polynomial, $k = 40$, $\lambda = 22$.

**Table 1.** Algorithm process

| Algorithm (Variable-exponent regularization) |
|---|
| Input: Matrix $\mathbf{P}$, vector $\mathbf{D}$ |
| Initialization $\mathbf{M}^0 = 0$, $\mathbf{W}^0 = \mathbf{M}^0$, $k = 1$, $\lambda = 0.1$ |
| While $\left\|\mathbf{M}^l - \mathbf{M}^{l-1}\right\|_2 > \varepsilon$, |
| (1) Update $\mathbf{M}^l$ with given $\mathbf{M}^{l-1}$ and $\boldsymbol{\rho}^{l-1}$ by Eq. (7) |
| (2) If $E^l > E^{l-1}$, |
| Set $\boldsymbol{\rho}^l = \boldsymbol{\rho}^{l-1}$; |
| Else |
| Update $\rho_j^l = 1 + \frac{1}{1+\frac{k}{(M_j^l)^2}}$, as defined in Eq. (5). |
| (3) $l = l + 1$ |
| End |
| Output: the final $\mathbf{M}$ |

## 3 Color Correction Experiment

### 3.1 Experimental Process

The dataset was obtained by a standard color correction experiment. According to the color correction principle shown in Fig. 1 and the ISO17321-1 standard [14], the experiment design is illustrated in Fig. 3. And the equipments used in the experiment are listed in Table 2.

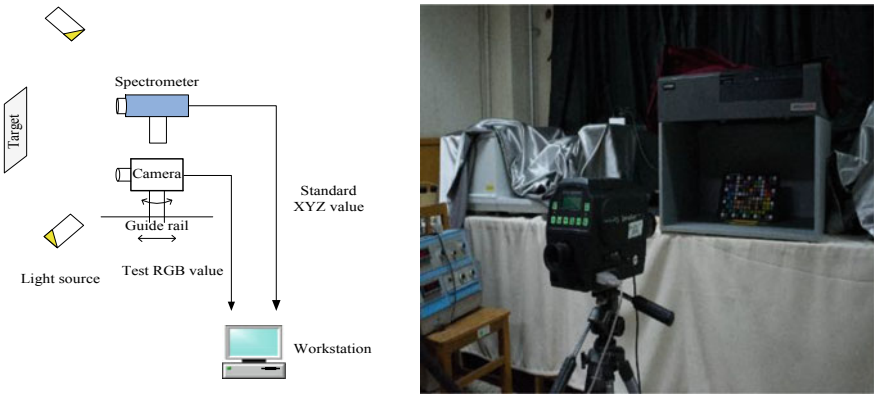



**Fig. 3.** Illustration and experimental scene of the color correction experiment

**Table 2.** Equipment list

| Instrument | Type |
|---|---|
| Spectrometer | SpectraScan PR 715 |
| Color camera | Nikon D750 |
| Target color chart | X-Rite Color Checker Digital SG 140 |
| Light source | Multi-light lightbox |
| White-point measurement | Standard white block |
| Illuminometer | TES-136 chroma meter |
| Others | 45° chart support |
| | Tripod |

In the specific experiment, the parameters of the color camera were fixed except for the exposure time. The exposure time was varied from 1/10 to 1/40 to simulate different illumination intensities. When the exposure time was 1/20, the picture was taken under normal conditions. The fixed parameters were approximately ISO 200, F/8, and the focal length was set to 160 mm. Half the data were chosen as training samples, and the other half were used for testing.

### 3.2 Experimental Result and Analysis

All the experimental results were calculated using the Lab1976 color difference equation [15], as shown below.

The color differences after corrected by different models are listed in Table 3. The bold data shows that the variable-exponent regression algorithm has the smallest difference compared to lasso regression and ridge regression in any original polynomial model. This prove our method has a better correction capability and manifests the superiority.

In the case of color correction, the quadratic polynomial is preferable to describe the correction relationship. This indicates that the response of the sensors in the color

**Table 3.** CIELAB color differences of test pictures taken at 1/20 exposure

| Original model regression method | Linear polynomial | Quadratic polynomial | Root-polynomial |
|---|---|---|---|
| Least-squares regression | 8.0275 | 3.4095 | 7.2759 |
| Lasso regression | 7.2459 | 3.4090 | 7.2569 |
| Ridge regression | 6.8616 | 3.3957 | 7.2644 |
| Variable-exponent regression | **6.7481** | **3.3914** | **7.2331** |
| Before correction | 56.7608 | 56.7608 | 56.7608 |

camera has nonlinearity error. To demonstrate the universality of the variable-exponent regularization, the same fitting experiment was also conducted using test samples at other exposure times. The results are shown in Fig. 4.

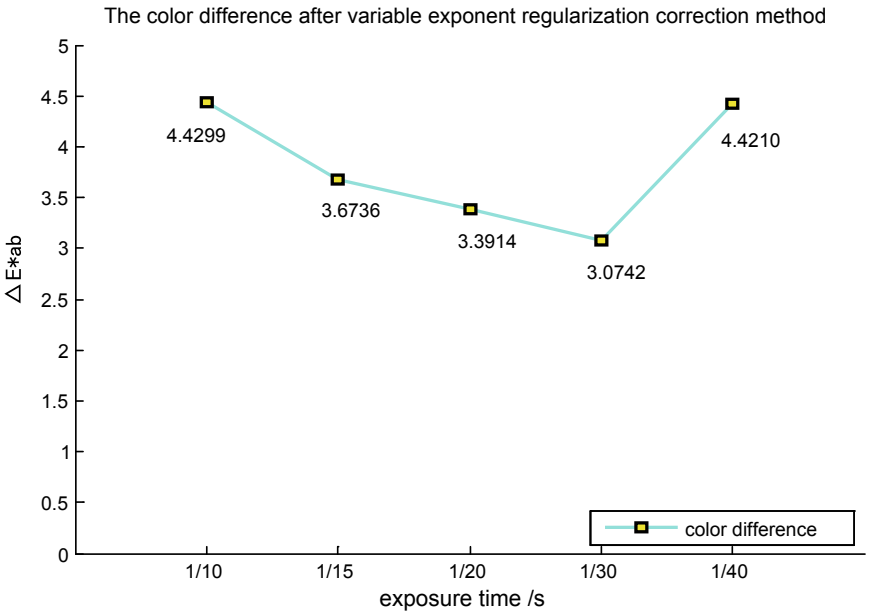



**Fig. 4.** CIELAB color difference after correction by variable-exponent polynomial regression

The color difference is less than 5 CIELAB units at all exposure times, which can fulfill the requirement in the color correction field. The exposure range included both underexposed and overexposed conditions. Therefore, it can be concluded that the variable-exponent polynomial regression algorithm can achieve color correction under different illuminations.

## 4 Conclusions

In this paper, we present a new fitting algorithm, variable-exponent regularization polynomial regression, which has the characteristic of self-adaptation. We applied the algorithm to three common polynomial models, linear, quadratic, and root polynomial. From comparison of the results, the proposed algorithm clearly performs better than original models. With the variable-exponent regularization, the basic model obtains better fitting precision and stability. Therefore, this algorithm has advantages in describing complicated color correction relationship. In addition, variable-exponent regularization can achieve effective correction under different illuminations, proving the practicality of the algorithm.

## References

1. Zhang XS, Gao SB, Li RX, et al. A retinal mechanism inspired color constancy model. IEEE Trans Image Process. 2016;25(3):1219–32.
2. Wang Y, Xu H. Colorimetric characterization for scanner based on polynomial regression models. Acta Opt Sin. (in Chinese). 2007; 6:035.
3. Kim SJ, Koh K, Lustig M, Boyd S. An interior-point method for large-scale-regularized least squares. IEEE J Sel Top Sig Process. 2007;1(4):606–17.
4. Zhang J, Yang Y, Zhang J. A MEC-BP-Adaboost neural network-based color correction algorithm for color image acquisition equipment. Optik. 2016;127(2):776–80.
5. Wang X, Zhang D. An optimized tongue image color correction scheme. IEEE Trans Inf Technol Biomed. 2010;14(6):1355–64.
6. Zhuo L, Zhang J, Dong P, et al. An SA–GA–BP neural network-based color correction algorithm for TCM tongue images. Neurocomputing. 2014;134:111–6.
7. Finlayson GD, Drew MS. Constrained least-squares regression in color spaces. J Electron Imaging. 1997;6(4):484–93.
8. Finlayson GD, Mackiewicz M, Hurlbert A. Color correction using root-polynomial regression. IEEE Trans Image Process. 2015;24(5):1460–70.
9. Xu Z, Zhang H, Wang Y, Chang X, Liang Y. L1/2 regularization. Sci China Inform Sci. 2010;53(6):1159–69.
10. Li F, Justin J, Yeung S. Regularization (L2/L1/Maxnorm/Dropout), http://cs231n.github.io/neural-networks-2/#reg 2017.
11. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. J R Stat Soc. 2011;73(3):273–82.
12. Galligani E. Lagged diffusivity fixed point iteration for solving steady-state reaction diffusion problems. Int J Comput Math. 2012;89(8):998–1016.
13. Dou Z, Gao K, Zhang B, Yu X, Han L, Zhu Z. Realistic image rendition using a variable exponent functional model for retinex. Sensors. 2016;16(6):832.
14. International Organization for Standardization. Graphic technology and photography-color characterization of digital still cameras (DSCs). ISO17321-1.2012.
15. Schmittmann O, Schulze LP. A true-color sensor and suitable evaluation algorithm for plant recognition. Sensors. 2017;17(8):1823.

# Single-Image Super-Resolution: A Survey

Tingting Yao[1], Yu Luo[1], Yantong Chen[1], Dongqiao Yang[1], and Lei Zhao[2(✉)]

[1] College of Information Science and Technology, Collaborative Innovation Research Institute of Autonomous Ship, Dalian Maritime University, Dalian, China
luoyurl@l26.com, yttl030@dlmu.edu.cn

[2] Institute of Environmental Systems Biology, College of Environmental Science and Engineering, Dalian Maritime University, Dalian, China
zhaol@dlmu.edu.cn

**Abstract.** Single-image super-resolution has been broadly applied in many fields such as military term, medical imaging, etc. In this paper, we mainly focus on the researches of recent years and classify them into non-deep learning SR algorithms and deep learning SR algorithms. For each classification, the basic concepts and algorithm processes are introduced. Furthermore, the paper discusses the advantages and disadvantages of different algorithms, which will offer potential research direction for the future development of SR.

**Keywords:** Single-image super-resolution · Learning-based · Deep learning · Survey

## 1 Introduction

Single-image super-resolution (SISR) is the process of generating one high-resolution (HR) image from one low-resolution (LR) image input with signal and image processing techniques. As artificial intelligence has become a predictable trend in recent years, and more image-related applications such as medical imaging, face recognition have penetrated in our daily life. It is of great significance to produce HR images to facilitate the follow-up processing. Hence, it is necessary and essential to make researches on super-resolution.

The proposed SR algorithms can be mainly divided into three categories: interpolation-based algorithms, reconstruction-based algorithms, and learning-based algorithms. Interpolation-based algorithms reconstruct the input LR image via interpolating the missing pixels that are calculated by the known pixels, and common algorithms include bicubic interpolation, edge detected interpolation, and so on. And reconstruction-based algorithm is a process that utilizes a certain prior knowledge to recover the LR image. Plenty of priors have been exploited and formulated into framework such as maximum a posteriori and variational Bayes approach. Compared to the interpolation-based methods, it can generate a sharper edge, but the reconstructed HR image degrades rapidly when applying to large-scale factor problem. Therefore, learning-based algorithms proposed afterward become mainstream methods within the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_16

three algorithms as they show the best performances. The basic idea of learning-based SR methods is to obtain the prior by calculating the relationships between testing sample image patches and training image patches. And the algorithms could be further divided into four subcategories: example-based, neural network, manifold, and compressive sensing. Considering that the superior performance of deep learning application in image processing in recent years, more researches incorporate deep learning, which attach to neural network methods, into SISR algorithms [14]. According to the connection methods, we classify deep neural networks into feed-forward deep networks, feedback deep networks and bidirectional-forward deep networks.

Therefore, focusing on the learning-based algorithms mainly, we review the papers under the classification framework of non-deep learning SR algorithms and deep learning SR algorithms that published on the topic of single-image super-resolution. Furthermore, we discuss the strengths and limitations of these algorithms.

The rest of the paper is organized as follows: In Sect. 2, we address in the non-deep learning SR algorithms and mainly introduce three kinds of prevalent methods: example-based methods, neighbor embedding methods, and compressive sensing. Then we will review deep learning SR algorithms in Sect. 3. Finally, the paper comes to a conclusion in Sect. 4.

## 2 Non-deep Learning SR Algorithms

In this section, we mainly focus on three kinds of common non-deep learning SR algorithms: example-based, manifold, and compressive sensing. These non-deep learning methods learn fine details that correspond to different image regions seen at low-resolution samples and then use those learned relationships to predict fine details in other images.

### 2.1 SR Based on Example-Based Methods

The example-based algorithm is presented by Freeman et al. in [5]. The basic process could be generalized into two steps: In training step, the image patches are firstly paired according to its relationship between HR image patches and LR image patches, then prior knowledge is learned to establish the model. In reconstruction stage, LR image input is interpolated and divided into patches. After that, the high-frequency patches which match best are searched in the training set. Finally, we could obtain HR output by adding the interpolated image patches and the searched high-frequency patches.

Based on the algorithm, many recent researches have been proposed to make improvements. According to the example types of training set, we can divide example-based methods into three groups: external example-based, internal example-based, and joint both internal and external examples methods.

Most early algorithms based on external database, in which the prior between LR image patches and HR image patches is learned from external natural database and hence the images in training database are the main influencing factor. In [16] A+ algorithm preserves some key facets such as anchored regressors, features, and time complexity. It learns regressors via full training material to reconstruct input image.

Then an algorithm [15] is proposed to make further improvement based on A+ algorithm. It improves the performance of example-based methods from seven ways: augmentation of data, large dictionary and hierarchical search, iterative back projection, cascade of anchored regressor, enhanced prediction, self-similarity, and reasoning with context.

Afterward researches demonstrate that internal example-based methods still show good performance when applies to specific images whose patches repeat (or across different scales). The algorithm proposed in [18] defines repeated patches with a new perspective, as it exploits the 3D scene geometry and patch search space expansion to improve the self-exemplar search. It mainly works on the images with repeated 3D transformed patches across different scales and hence owns a strong pertinence though performs well.

Furthermore, there are researches on how to combine these two methods to make them complement each other and achieve a better performance on SR. In [17], a joint both internal and external examples algorithm is proposed to combine the two methods by defining two loss functions, one is sparse coding-based external examples while the other is epitomic matching-based internal examples. Two functions are concatenated by a corresponding adaptive weight to automatically balance their contributions according to their reconstruction errors.

Example-based methods are able to learned fine high-frequency information via training set, but the quality of reconstruction image is highly depended on the quality of training set.

### 2.2 SR Based on Manifold

According to manifold learning theory, the LR image patch and its HR counterpart can form a manifold with same local geometric structure in feature space. And it is useful for nonlinear dimensionality reduction of image feature space. Inspired by it, [3] puts forward neighbor embedding method. And the process can be summarized as following steps: In training step, manifold could be obtained through HR images and corresponding LR images in training database. In testing step, first find K-nearest neighbors in LR images of training database, then compute the reconstruction weights of each neighbor that minimize the error of reconstruction error, and using the HR features of K-nearest neighbors as well the reconstruction weights to reconstruct the target HR image.

For neighbor embedding methods, it is significant to maintain the neighborhood relationship. There are two possible ways to realize it. One is to choose a more plausible representation of image patch feature, and the other is to choose a better reconstruction function for given HR image patch neighborhoods. Approaches presented recently mainly based on the former one.

Yang et al. [20] focus on dealing with the nonlinear high-dimensional feature of LR image input. It incorporates projection method and projects the high-dimensional feature into two subspaces, kernel principal component analysis (KPCA) subspace and manifold locality preserving projection (MLPP) subspace, to acquire the global and local structures, respectively. In [21], LR image input is decomposed into several directional sub-bands using designed steerable filters in frequency domain and each sub-band stores correlated structure information of different directional frequencies.

These methods preserve the structures well so it is beneficial to recover more details. Moreover, they can search for k-nearest neighbors with higher matching precision in low-dimensional subspace when applying NE algorithm.

### 2.3 SR Based on Compressive Sensing

According to the signal sparsity theory, compressive sensing methods reconstruct HR images with the help of a set of sparse representations and over-complete dictionaries. Yang et al. [7] first attempt to apply sparse signal representation into super-resolution. The algorithm assumes that there are a set of dictionaries learned jointly from LR to HR training image patches. In reconstruction step, the LR image input is divided into many overlapping patches and their sparse representation coefficients can be calculated through LR dictionary, and then HR patches are estimated using these coefficients and the HR dictionary.

Most of recent methods focus on how to cluster the training images into groups according to different image geometric characters to learn better dictionaries. In the meanwhile, a mapping function with sparse representation can be obtained to reconstruct target HR image in training phase. In reconstruction phase, an approximate clustered dictionary and mapping function are selected based on some certain paradigm.

Ahmed and Shah [1] cluster the training set into nine directional clusters according to the correlation between image patches. For LR training images, dictionaries are learned via K-SVD algorithm while learning high-resolution dictionary is a pseudo-inverse problem. And in [11], nonlocal similar image patches are divided into groups. The group sparse representation incorporates nonlocal similarity redundancy to improve the accuracy of sparse coding, and the clustered dictionaries are organized as the combination of many orthogonal sub-dictionaries which trained based on the clusters. Other characteristics are still being exploited such as edge sharpness and multiscale redundancy of image structure feature proposed in [2, 13], respectively.

Compressive sensing algorithms can achieve a good quality of reconstruction images. However, the complexity of algorithms is relatively large as the sparse coding and reconstruction procedure need to be iterated for many times.

## 3 Deep Learning SR Algorithms

In this section, we review on the deep learning SR algorithms from three classifications according to different connection methods of the deep networks: They are, respectively, feed-forward deep network, feedback deep network and bidirectional-forward deep network.

### 3.1 SR Based on Feed-Forward Deep Network

Feed-forward deep network is the earliest but also the simplest deep network structure, in which each neuron starts from the input layer, receives the previous input and outputs it to the next level until it reaches the output layer. There is no feedback

throughout the whole network and can be demonstrated with a directed acyclic graph. And in terms of super-resolution task, the feed-forward network extracts the local patterns of input images through convolution algorithm in the input layer, then passes them to the hidden layer in one direction. We could acquire deeper level features while the hidden layers are getting deeper. Finally, the output layer would output the reconstructed image.

The typical network of feed-forward deep network is convolution neural network (CNN). According to the network structures, we will divide it into super resolution using convolution neural network (SRCNN) and very deep networks for super resolution (VDSR).

Dong et al. first put forward SRCNN in [4], using three deep convolutional neural networks to achieve three stages of SR: image patches extraction, nonlinear mapping, and image reconstruction. In training step, overlapped LR patches samples are abandoned to avoid increasing the computation amount, and mean square error is adopted as loss function and minimized with stochastic gradient descent method. Liang et al. [12] further incorporate the image prior information to improve the performance of SRCNN. SRCNN-Pr adds a feature extraction layer at the end of SRCNN to extract the edge information of image via Sobel filter and finally fuses all information to rebuild the image. CSCSR [6] is also an expansion of SRCNN network. It focuses on the entire image instead of overlapped patches. And the algorithm first utilizes a set of filters to achieve three steps of SR: decompose LR image into LR sparse feature maps, predict the HR feature maps from the LR ones, and reconstruct the HR images from the predicted HR feature maps.

Deeper than the SRCNN network, VDSR [8] designs a deeper network that contains 20 convolutional layers. The filters in each layer are more and smaller than those in SRCNN, which help increase the local receptive fields to acquire the context of LR and HR images. In training step, VDSR adopts residual learning method and adjustable gradient clipping strategy to solve gradient vanishing and gradient explosion problems.

### 3.2 SR Based on Feedback Deep Network

Back-forward deep network is also called recursive network, in which each neuron feeds its own output signal as an input signal to other neurons while it still feeds forward the information. And it takes a while for the deep network to reach stability. In SR issue, feedback deep network learns convolutional features through filter group and finally combines the information to obtain the reconstructed image. On the other hand, feed-forward deep network could be regarded as a coding procedure of input signal while feedback deep network is regarded as a decoding procedure.

The feedback deep network can be divided into deconvolutional networks (DN), recursive, and recurrent neuron networks (RNN).

SR based on fast deconvolution (FD) [10] method has three steps: initial upsampling, gradient prior computation, and fast deconvolution. Firstly, initial LR image is obtained by upsampling, then horizontal and vertical gradients are computed in Fourier domain to improve the efficiency of deconvolution and the HR image is recovered by fast deconvolution method finally.

SR based on deeply recursive convolutional network (DRCN) [9] consists of embedding network, inference network, and reconstruction network. First, image features are extracted through embedding network. Then the feature information could be applied to LR-HR mapping in inference network, which has plenty of recursive layers and shares the same reconstruction layer. Finally, the reconstruction network recovers the HR image by weighted averaging all the outputs of inference network. In training strategy, the network utilizes skip connection and recursive supervision methods to resolve gradient vanishing and explosion problems. And considering the prior information, deep edge guided recurrent residual (DEGRR) [19] incorporates the image edge prior knowledge to guide the feature mapping learning.

### 3.3 SR Based on Bidirectional-Forward Deep Network

Bidirectional-forward deep network has the traits of feed-forward deep network and feedback deep network. The information transmission between different layers in the deep network is bidirectional. It includes deep boltzmann machines (DBM), deep belief networks (DBN) and stacked auto-encoders (SAE). Under such propagation strategy, the network can train the weights better so that it manages to achieve a better performance.

## 4 Conclusion

This survey paper reviews recent papers published on the topic of super-resolution and introduces these SR algorithms under the proposed framework. Besides giving the details of the methods, it mentions the pros and cons of them. And observing the traits of non-deep learning SR algorithms and deep learning SR algorithms, we find that some deep learning algorithms could be regarded as a migration of non-deep learning SR algorithms but own stronger feature representation and learning ability, which could enlighten our future research direction of SR.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (No. 31700742), the Young Elite Scientist Sponsorship Program by CAST (2017QNRC001) and the Fundamental Research Funds for the Central Universities (No. 3132018306, 3132018180, 3132018172).

## References

1. Ahmed J, Shah MA. Single image super-resolution by directionally structured coupled dictionary learning. Eurasip J Image Video Process. 2016;1:36.
2. Ahmed J, Klette R. Coupled multiple dictionary learning based on edge sharpness for single-image super-resolution. In: International conference on pattern recognition. IEEE; 2017.
3. Chang H, Yeung DY, Xiong Y. Super-resolution through neighbor embedding. Proc Comput Vis Pattern Recogn. 2004;1:I-275–82.
4. Dong C, et al. Image super-resolution using deep convolutional networks. IEEE Trans Pattern Anal Mach Intell. 2016;38(2):295–307.

5. Freeman WT, Jones TR, Pasztor EC. Example-based super-resolution. Comput Graph Appl. 2002;2:56–65 (IEEE22.).
6. Gu S, et al. Convolutional sparse coding for image super-resolution. In: IEEE international conference on computer vision. IEEE; 2015. p. 1823–31.
7. Huang JB, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. Comput Vis Pattern Recogn. 2015:5197–206.
8. Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. Comput Vis Pattern Recogn. 2016:1646–54.
9. Kim J, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution. Comput Vis Pattern Recogn. 2016:1637–45.
10. Krishnan D, Fergus R. Fast image deconvolution using hyper-Laplacian priors. In: International conference on neural information processing systems Curran Associates Inc.; 2009. p. 1033–41.
11. Li X, et al. Single image super-resolution via subspace projection and neighbor embedding. Neurocomputing. 2014;139:310–20.
12. Liang Y, et al. Incorporating image priors with deep convolutional neural networks for image super-resolution. Neurocomputing. 2016;194:340–7.
13. Song S, et al. Joint sub-band based neighbor embedding for image super-resolution. In: IEEE international conference on acoustics, speech and signal processing. IEEE; 2016. p. 1661–5.
14. Sun X, Xiao-Guang LI, Jia-Feng LI, et al. Review on deep learning based image super-resolution restoration algorithms. Acta Automatica Sin. 2017;43(5):697–709.
15. Timofte R, Rothe R, Gool LV. Seven ways to improve example-based single image super resolution; 2015. p. 1865–73.
16. Timofte R, Smet VD, Gool LV. A+: Adjusted anchored neighborhood regression for fast super-resolution. In: Asian conference on computer vision. Springer, Cham; 2014. p. 111–26.
17. Wang Z, Yang Y, Wang Z, et al. Learning super-resolution jointly from external and internal examples. IEEE Trans Image Process. (A Publication of the IEEE Signal Processing Society). 2015;24(11):4359.
18. Yang J, et al. Image super-resolution via sparse representation. IEEE Trans Image Proces. 2010;19(11):2861–73.
19. Yang W, et al. Deep edge guided recurrent residual learning for image super-resolution. IEEE Trans Image Process (A Publication of the IEEE Signal Processing Society). 2016; 26(12):5895–907.
20. Yang W, et al. Image super-resolution via nonlocal similarity and group structured sparse representation. Vis Commun Image Process. 2016:1–4.
21. Zhang Y, et al. Image super-resolution based on, structure-modulated sparse representation. IEEE Trans Image Process. 2015;24(9):2797–810.

# Research on Infrared Sequential Image Generation Techniques in Pure Background

Changyun Ge[1], Haibei Zhang[3], Ti Han[1], Jiaqi Li[1(✉)], Xiujie Zhao[1], Baowei Lin[2], Shuhan Yan[1], and Baoxin Huang[1]

[1] Department of Electronic Engineering, Dalian Neusoft University of Information, Dalian 116023, Liaoning, China
{gechangyun, lijiaqi}@neusoft.edu.cn
[2] Ali, Beijing, China
[3] Department of Foreign Language, Dalian Neusoft University of Information, Dalian 116023, Liaoning, China

**Abstract.** Infrared image generation technology is an important part of infrared target guidance simulation. This paper studies the method of environment modeling, introduces the method of constructing SNR model and the method of coordinate transformation. The infrared image sequence is generated by SNR, coordinate information, and projectile information. The infrared image generated has three different target models: sky, sea, and land. The background is pure, the target is clear, and the experiment shows that it is feasible.

**Keywords:** Infrared image generation · SNR · Coordinate transformation

## 1 Introduction

In the process of infrared-guided weapon development and performance evaluation, the test of the infrared-guided system is usually conducted via the external field and the internal field, namely the field live-fire target shooting experiment and the simulation test. Although the former excels in test results, it has disadvantages such as large capital cost and long test period. In addition, the external field is unable to traverse all test condition due to the weather condition of test spot and harmful to the environment as well as possible to cause accidents. The latter is featured with a number of advantages, for instance, being able to traverse all inspection conditions, low cost, no restriction on test sites, and environmental protection. Therefore, the second method is often implemented in the research and development experiments of weapon. However, due to certain deviations in the simulation test, the common international practice is to develop and design the weapon system after performing simulation experiments, which aim to obtain accurate performance data. After the weapon system is developed, a live-fire test will be performed to verify the accuracy of its guidance [1].


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_17

## 2 Introduction of the Related Principles

### 2.1 Modeling of SNR Model

According to the definition of SNR, the SNR of the detector's output voltage can be calculated by Eq. 1.

$$\mathrm{SNR} = \frac{U_1 - U_2}{U_3} \tag{1}$$

$U_1$ is the output voltage of the detection unit part corresponding to the target, $U_2$ is the output voltage of the detection unit part corresponding to the background $U_3$ is the equivalent noise voltage of the infrared detector's radiation receiver [2, 3].

According to the definition of each parameter in the formula, Formula 1 can be derived as Formula 2

$$\mathrm{SNR} = \frac{A_1 \tau_1 \tau_2 D^*}{\pi f^2} \left(\frac{A_2}{\Delta f_n}\right)^{1/2} \int_{\lambda_2}^{\lambda_1} \varepsilon_1 M(\lambda, T_1) - \varepsilon_2 M(\lambda, T_2) \mathrm{d}\lambda \tag{2}$$

1. $A_1$ is the area of the optical system in the infrared detector,the value of which can be calculated through the equation: $A_1 = (1/4)\pi D^2\ (\mathrm{cm}^2)$.
2. $\tau_1$ is the infrared transmittance radiation of the atmosphere. The value is calculated by LOWTRAN 7 and stored in SQL Table.
3. $\tau_2$ is the infrared transmittance of optical system and the general value is 0.8.
4. $D^*$ is peak wavelength detectivity, which values $102 \times 10^{11}\ \mathrm{cm\,Hz^{1/2}\,W^{-1}}$.
5. $A_2$ is the area of the detection unit, which values $25 \times 25$.
6. $M(\lambda, T)$ is Planck radiation emission intensity of the target or background in IR detector.
7. $f$ is the focal length for optical system and the range of value is 50–200 mm.
8. $D$ is entrance pupil aperture and the range of value is 50–200 mm.
9. $\Delta f_n$ is equivalent width of noise.
10. $\varepsilon_1$ is the emissivity of the target; $\varepsilon_2$ is the emissivity of the background.

$$\int_{\lambda_2}^{\lambda_1} |\varepsilon_1 M(\lambda, T_1) - \varepsilon_2 M(\lambda, T_2)| d\lambda. \tag{3}$$

Formula 3 is integral from $\lambda_1$ to $\lambda_2$, which can be calculated by searching relative radiance out function $F(\lambda, T)$ table using Formula 4

$$M_{\lambda_1-\lambda_2} = [F(\lambda_2 T) - F(\lambda_2 T)]\sigma T^4 \tag{4}$$

$\sigma$ is Sterling Boltzmann constant valued $2.6703 \times 10^{-8}\,\mathrm{Wm^{-2}K^{-4}}$.

## 2.2 Conversion of Latitude and Longitude Coordinate System into Geocentric Coordinate System

In the coordinate conversion, the target latitude and longitude coordinate system is first converted into a geocentric coordinate system and the method of which can be calculated by Eq. 5.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} (N+H)\,\cos B\,\cos L \\ (N+H)\,\cos B\,\sin L \\ [N(1-e^2)+H]\,\sin B \end{bmatrix} \tag{5}$$

In the formula, $B$ represents latitude; $L$ represents longitude; $H$ represents height; $X$, $Y$, and $Z$ are geocentric rectangular coordinates, and the radius of the turn is calculated as Eq. 6

$$N = \frac{a}{\sqrt{1 - e^2 \sin B^2}} \tag{6}$$

$e$ represents eccentricity and the formula is as shown in Eq. 7, where $a$ is the long semi-axis of the earth and $b$ is the short semi-axis of the earth [4].

$$e^2 = \frac{a^2 - b^2}{a^2} \tag{7}$$

## 2.3 Conversion of Geocentric Coordinate System into the Geodetic Coordinate System

The base point of the ground coordinate system could be any point on the ground, and the base point could be the launch spot of the infrared detector carrier, the coordinates of which are $X'$, $Y'$, $Z'$. The coordinate of the ground coordinate system is $X''$, $Y''$, $Z''$ The coordinates of the geocentric coordinate brick can follow the Eq. 8. Calculation.

$$\begin{bmatrix} X'' \\ Y'' \\ Z'' \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \tag{8}$$

## 2.4 Conversion of Ground Coordinate System into Detector Coordinate System

The relationship between the detector coordinate system and the ground coordinate system is usually determined by the pitch angle $\theta$, yaw angle $\varphi$, and tilt angle $\gamma$.

Pitch angle $\theta$, the angle between the longitudinal axis ($Ox_1$ axis) of the detector and the horizontal plane ($Axz$ plane). $\theta$ is positive when the longitudinal axis of the detector pointing above the horizontal plane, otherwise, $\theta$ is negative when pointing below the horizontal plane.

Flight angle $\varphi$: the angle between the projection of the longitudinal axis of the detector in the horizontal plane ($Ax'$ axis) and the $Ax$ axis of the ground coordinate system. Observed from the $Ax'z$ plane, if $Ax$ turned to $Ax'$ counterclockwise, $\varphi$ is positive; if $Ax$ turned to $Ax'$ clockwise, $\varphi$ then is negative.

Tilt angle $\gamma$: The angle between the $Oy_1$ axis of the detector coordinate system and the $Ax'y'$ plane. From the front view of the trailing axis, if $Oy_1$ axis is on the right side of $Ax'y'$, $\gamma$ is positive. While if it is on the left side, $\gamma$ is negative.

In coordinate transformation, the origin and coordinate axes of the detector coordinate system and the ground coordinate system should be overlapped in a corresponding manner, and then the rotation matrix is rotated around the corresponding coordinate axes according to the corresponding angle parameters (Fig. 1).

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = L(\gamma, \theta, \varphi) \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{9}$$

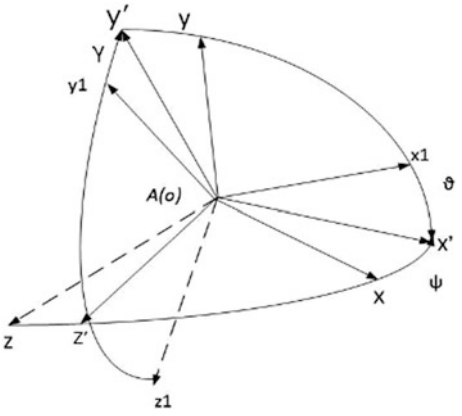



**Fig. 1.** Relationship between ground coordinate system and detector coordinate system

### 2.5 Conversion of Detector Coordinate System into Image Coordinate System

$$\begin{bmatrix} z_3 \\ y_3 \end{bmatrix} = x_3 \begin{bmatrix} \tan\theta \\ \tan\sigma \end{bmatrix} \tag{10}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{n_1}{2} \\ \frac{n_2}{2} \end{bmatrix} + \begin{bmatrix} \frac{fz_3}{dx_3} \\ \frac{fy_3}{dx_3} \end{bmatrix} \tag{11}$$

$x$, $y$ represent pixel coordinate $n_1$, $n_2$ represent the number of horizontal and vertical pixels of the infrared image. $f$ is the focal length of the optical system, $d$ is the physical actual distance between pixel points in the image, the value is 0.025, $\theta$ is the azimuth angle of the target in the detector coordinate system, and $\sigma$ is the pitch angle of the target in the detector $x_3$, $y_3$, $z_3$ represent the distance of the target in the detector coordinate system, which can be calculated by distance formula (Fig. 2).

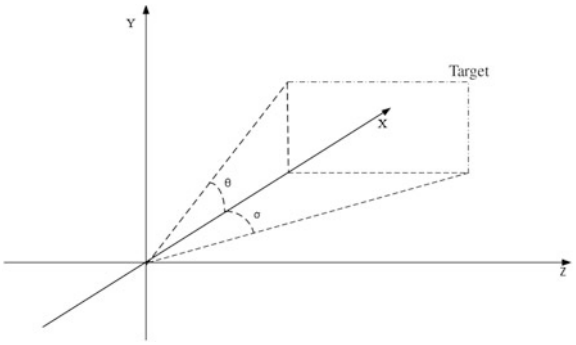



**Fig. 2.** Target coordinate in detector coordinate system

## 3 Generation of Simulated Infrared Image

The infrared image is generated on the basis of current target model and current signal-to-noise ratio as well as background and coordinate information.

Three typical background models of infrared images are sky background, ocean background, and land background, and the respective targets are aircraft, ships, and tanks, respectively. The target of the sky is a fighter, the actual length of which is 20 m; the ocean target is a large ship, and the actual length is 300 m; the ground target is a tank, the actual length of which is 9 m. Noise is inserted into the target background of infrared image with the intention of interfering recognition and tracking, and thus affecting the accuracy of the angle measurement (Fig. 3).

The signal-to-noise ratio of infrared images is calculated based on atmospheric information, coordinate information, and detector information. The specific methods

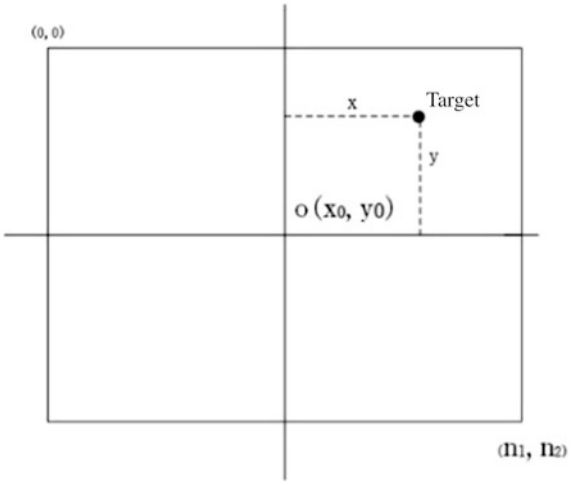



**Fig. 3.** Position of the target in the infrared image

are manifested in Sect. 2.1. The following method is implemented in converting the signal-to-noise ratio to the standard deviation of image noise.

$$U_{\mathrm{m}} = \frac{(U_{\mathrm{c}} - U_{\mathrm{b}})}{\mathrm{SNR}} \tag{12}$$

In this equation, $U_{\mathrm{c}}$ refers to the average gray value corresponding to the target; $U_{\mathrm{b}}$ refers to the average gray value corresponding to the background; $U_{\mathrm{m}}$ refers to the value of the standard deviation of the noise to be added [5].

The target position in the infrared image is set according to the positional relationship between the current target and the infrared detector. Therefore, a large number of coordinate transformations will be involved, including: latitude and longitude coordinate transforming into ground coordinates, ground coordinate transforming into missile coordinates, and projectile coordinate transforming into azimuth elevation angle in the optical axis coordinate system and azimuth pitch angle transforming into coordinate in the axis coordinate system.

The functions involved in the generation of infrared images include target fill function, background noise addition function, target noise addition function, azimuth elevation pitch transforming into image coordinate function, image coordinate transforming into azimuth elevation angle function, latitude and longitude transforming into azimuth elevation angle function, image filter function, and addition transmittance function. The flow chart is shown in Fig. 4.

Since the real infrared images are inaccessible with simulation system, infrared images are all generated by software. The typical target is generated from a real infrared image, varying brightness and additional noise. The background of the infrared image is generated by adding noise and image filtering ensuring the resemblance to the real infrared image. The position of the infrared target in the image is obtained by

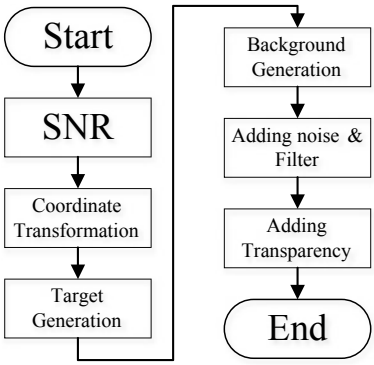



**Fig. 4.** Flowchart for the generation of infrared image

coordinate conversion, and the coordinate conversion obtains a single coordinate point of the infrared target, and the infrared target is a surface target with pixel size. Therefore, in the process of generating infrared image, the target image is pasted into the background image.

The infrared image generation module converts the SNR which is calculated by SNR model module into a Gaussian noise standard deviation of the image, and then the infrared image is generated in combination with the background model selected by the user. Figures 5, 6, and 7 show the infrared image generation process.

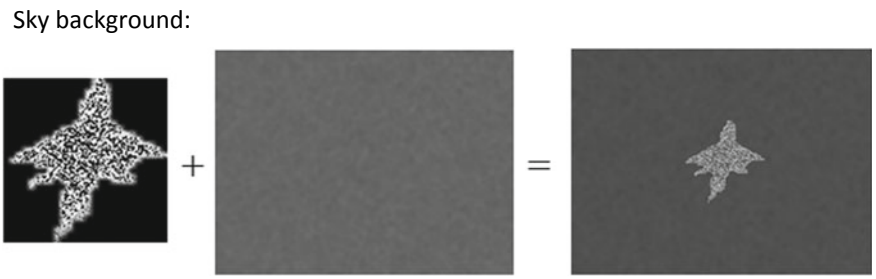



**Fig. 5.** Generation process of infrared image of sky image background

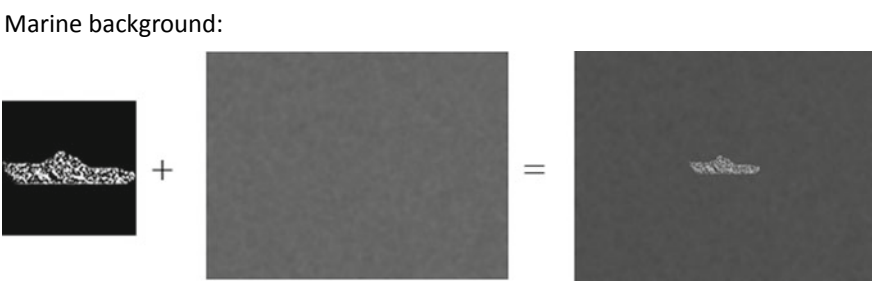



**Fig. 6.** Generation process of infrared image of marine background

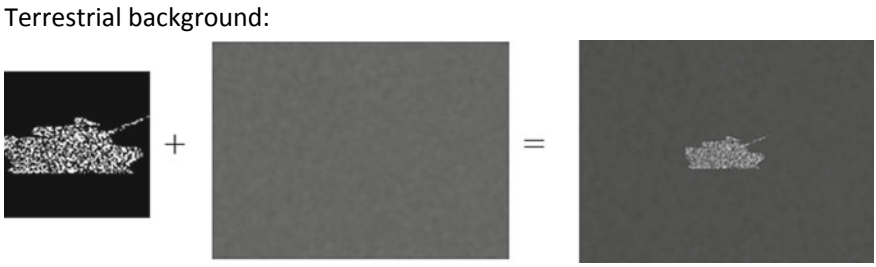



**Fig. 7.** Generation process of infrared image of terrestrial background

In accordance with Figs. 5, 6 and 7, the generated infrared images are slightly different from the original infrared images. However, the background and the target can be clearly identified in the generated infrared images which distinctively manifested the three background modules, the sky, the ocean and the land.

**Acknowledgments.** Project supported by Research on real-time 3D detection algorithm for complex scene change, Dalian youth technology star, Dalian science and technology project (ZX2016KJ013).

## References

1. Ge C. Design of infrared target tracking simulation platform. Harbin: Harbin Engineering University; 2015.
2. Lu X, Shen Z. Calculation and analysis of image signal-to-noise ratio output by infrared focal plane detector. Syst Eng Electron Technol. 2002. 1001-506X:111-113.
3. Razeghi M, Pour SA, Huang EK, et al. Type—II InAs/GaSb photodiodes and focal plane arrays aimed at high operating temperature. Opto-Electron Rev. 2011;19(3):261–9.
4. Liu M, Li P, Chen Y, Wang Z. Study on the algorithm of geogeocentric right Angle coordinate and geodetic coordinate conversion. Electron Technol. 2010;23(2):18–20.
5. Qian X, Lin R, Zhao Y. Missile flight mechanics. Beijing: Beijing University of Technology Press; 1990. p. 30–6.

# Nonlinear Image Enhancement Based on Non-sub-sampled Shearlet Transform and Phase Stretch Transform

Ying Tong[1,2], Kaikai Li[1,2], Jin Chen[1,2(✉)], and Rong Liu[1,2]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
cjwoods@163.com
[2] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China

**Abstract.** In this paper, non-sub-sampled shearlet transform (NSST) multi-scale analysis is combined with phase stretch transform (PST) to nonlinearly enhance images. The components of different scales after NSST multi-scale decomposition are processed by nonlinear models with different thresholds, and the noise is well suppressed while enhancing the detail features. The thresholds of the enhanced model are determined by the local standard deviation of PST feature map. Experiments on Matlab platform show that the proposed algorithm has improved image distortion, cleared details, and enhanced image contrast.

**Keywords:** Image enhancement · Non-sub-sampled shearlet transform · Phase stretch transform · Nonlinear function

## 1 Introduction

Image enhancement is a crucial pre-treatment step for many image processing applications. Image enhancement algorithms can be roughly grouped into two kinds [1]: space domain image enhancement and frequency domain image enhancement. Space domain image enhancement algorithm is mainly based on grayscale transformation, histogram processing, mean filtering, and other methods. Frequency domain image enhancement is also called transform domain image enhancement. Nowadays, the transform domain image enhancement algorithm has gradually become a research hotspot with the continuous development of multi-scale analysis tools [2–6]. Common multi-scale analysis tools for image enhancement include wavelet transform, contourlet transform, shearlet transform, and so on.

Image phase information plays a crucial role in the procedure of human perception of images. Phase information can reflect the space information of images and has a strong ability to interpret images [7–9].

From the above analysis, it is supposed that if we can combine the accuracy of multi-scale analysis with the stability of phase analysis in the process of image enhancement, we will get a more ideal enhancement effect.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_18

## 2 Design of the Algorithm

### 2.1 NSST Multi-scale Analysis

Non-sub-sampled shearlet transform (NSST) is a multi-scale and multi-directional analysis tool. It is an optimal approximation, which is a "true" sparse representation of the image in all directions and at various scales. In addition, NSST does not perform down-sampling when performing multi-scale decomposition of an image. It combines a non-down-sampling tower transform with a non-sub-sampled directional filter bank and transforms. The size of the directional sub-bands of each scale is the same as that of the original image. The transform has shift-invariance and can overcome the pseudo-Gibbs phenomenon when the image is reconstructed.

### 2.2 Feature Extraction with PST

PST is a new image processing method proposed by Asghari and Jalali. It is said that PST has superior performance in feature extraction, especially edge extraction. It can even be applied to detect features in low-contrast visually impaired images [10]. The PST processing can be divided into several steps: Firstly, smooth the image with a low pass filter to reduce noise. Secondly, apply the PST kernel which has a nonlinear frequency dependent transfer function to the image. The output of the transform is the phase in space domain. The applied phase is frequency dependent with the original image, which is higher amount of phase applied to higher frequency features of the image. Finally, find image sharp transitions by thresholding the phase and apply binary morphological operations, so as to enhanced the features and clean the transformed image.

### 2.3 Nonlinear Enhancement Model

The enhancement function selected in this paper is presented by A. F. Laine in 1996. After that, many scholars applied this function to image enhancement processing and achieved good results. The enhancement function is shown in formula 1.

$$f(x) = a[\mathrm{sigm}(c(x-b)) - \mathrm{sigm}(-c(x+b))] \tag{1}$$

where $a = \frac{1}{\mathrm{sigm}(c(1-b))-\mathrm{sigm}(-c(1+b))}$, sigm is defined as $\mathrm{sigm}(x) = \frac{1}{1+\mathrm{e}^{-x}}$.

In this function, there are two parameters of $b$ and $c$, where b is used to control the enhancement range, and its value is usually between 0 and 1, while c is used to control the enhancement strength, and it usually uses the fixed value between 20 and 50. However, our algorithm does not preset fixed values for $b$ and $c$. Instead, we set two thresholds of the function to determine the shape of the function curve, and then obtain the values of the $b$, $c$ parameters.

### 2.4 Algorithm Flow and Steps

First, the NSST transform is applied to decompose the original image. After that, the NSST coefficients of different scales and directions can be obtained. Since image detail information often exists in high-frequency coefficients, we choose high-scale coefficients for processing, while low-frequency coefficients remain unchanged.

Considering that PST algorithm has obvious advantage in extracting features of image, the proposed method extracts the details of the original image by PST and uses its output as the image feature map. This feature map can clearly distinguish the region where the detail information is concentrated, so it can be used as the reference of image enhancement. Figure 1 is the feature map of the Lena image processed by PST.

**Fig. 1** Feature map of the Lena image processed by PST

After PST, we calculate the local standard deviation of each pixel in the feature map. The local standard deviation reflects the contrast change of the local area in an image. Therefore, the pixels with larger local standard deviation in feature map will be rich in the detail texture around it, so they should be enhanced, while the region with the smaller value of the local standard deviation can be mildly enhanced or preserved.

The enhancement of the NSST high-scale coefficients depends on the output of the nonlinear enhancement model. In order to facilitate the calculation, the values of the input need to be normalized. The nonlinear enhancement function of Fig. 1 contains two thresholds. One threshold is the first intersection of the enhancement function and the line of $y = x$. It is the threshold that characterizes the feature coefficients from noise coefficients. We call it $T1$. If the abscissa value of a certain point on the function curve is less than $T1$, then the value of its ordinate will be reduced nonlinearly, that is suppression (applicable to noise information); while if the abscissa value of a certain point on the function curve is greater than $T1$, the value of its ordinate will be nonlinearly stretched, that is enhanced (applicable to detail information). The other threshold is first intersection of the enhancement function and the line of $y = 1$. It is the

threshold that determines the extent of enhancement. We call it $T2$, if the abscissa of a certain point on the enhancement function curve is larger than $T2$, it is considered as strong edge, so this part will be enhanced greatly.

As shown in Fig. 2, the abscissa value in the graph is set to be the local standard deviation of the corresponding pixels in the PST feature map. We choose the $k1$ times of the minimum value in the local standard deviation as $T1$ ($k1$ is a positive integer greater than 1) and $k2$ times of the maximum as $T2$ ($k2$ is a positive integer greater than 1). After that, if we bring the pixel's local standard deviation in the PST feature map into the enhancement function, we can get the enhancement ratio at the pixel in the enhancement curve. It is the ratio of the ordinate value to the abscissa value. Finally, the NSCT coefficients of the N-scale can be multiplied with the enhancement ratio. After the processed NSCT coefficients are inversely transformed, the enhanced image is formed.

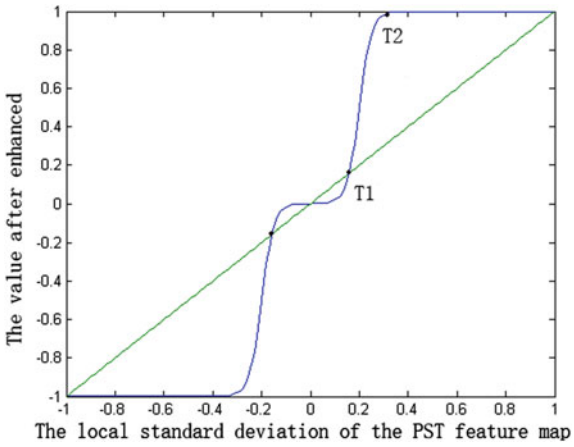



**Fig. 2** Nonlinear enhancement function

## 3 Experiments and Simulation

### 3.1 The Experiment Settings

We carried out experiments on two images, they were Lena and Barbara. As a contrast, wavelet nonlinear enhancement, NSCT nonlinear enhancement, the NSST nonlinear enhancement proposed in the literature [5], and the algorithm proposed in this paper enhanced four images, respectively. All experiments were carried out on Matlab R2016b platform.

### 3.2 Subjective and Objective Evaluation

The enhancement results are shown in Figs. 3 and 4, among which (a) is the original image, (b) is acquired using wavelet nonlinear enhancement, the NSCT nonlinear

enhancement result is presented in (c), (d) is the enhancement result with the method proposed by literature [5], and (e) is the image enhanced by our algorithm. Through subjective evaluation, our algorithm has significant effect on image contrast stretching and edge enhancement. The enhanced images have no pseudo-Gibbs phenomenon, and the details are clear without obvious distortion.


**Fig. 3** Enhancement results of Lena

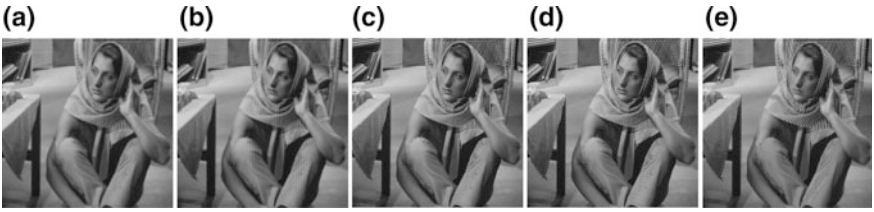



**Fig. 4** Enhancement results of Barbara

Generally, it is difficult for human beings to perceive the mild differences among the enhancement results of various algorithms. Hence, the paper adopts two widely used objective evaluation parameters, Edge Preserved Index (EPI) and Contrast Improvement Index (CII), to evaluate the image enhancement effects. The greater values of EPI and CII indicate the better quality of the image. It can be seen from Table 1 that the effects of our algorithm are more ideal than the other algorithms.

**Table 1** Objective evaluation parameters for Lena and Barbara image sets

| Lena | CII | EPI | Barbara | CII | EPI |
|---|---|---|---|---|---|
| Wavelet | 1.0295 | 1.0771 | Wavelet | 1.0189 | 1.0172 |
| NSCT | 1.0639 | 2.0942 | NSCT | 1.1463 | 1.8084 |
| NSST | 1.0734 | 2.3440 | NSST | 1.1961 | 2.0732 |
| NSST+PST | 1.0823 | 2.5283 | NSST+PST | 1.3666 | 3.1209 |

## 4 Conclusion

This paper describes an approach for accomplishing a nonlinear image enhancement algorithm that combines NSST and PST and takes the advantages of multi-scale analysis and image phase analysis to perform nonlinear enhancement of images. The algorithm can enhance image details and edge features without amplifying noise. Experiments on Matlab platform show that the algorithm is better than other algorithms in the experiments in terms of accurate edge preservation and adequate contrast enhancement.

**Acknowledgments.** This work was supported by National Natural Science Foundation of China (Grant No: 61701344), Tianjin Edge Technology and Applied Basic Research Project (14JCYBJC15800) in China, Tianjin Normal University Application Development Foundation (52XK1601), Tianjin Normal University Doctoral Foundation (52XB1603, 52XB1713), and Tianjin Higher Education Creative Team Funds Program in China.

## References

1. Lee S-Y, Kim D, Kim C. Ramp distribution-based image enhancement techniques for infrared images. IEEE Sig Process Lett. 2018;25(7):931–5.
2. Kim K, Kim S, Kim KS. Effective image enhancement techniques for fog-affected indoor and outdoor images. IET Image Proc. 2018;12(4):465–71.
3. Dong L, Bing Y, Mei Y, et al. Image enhancement based on the nonsubsampled contourlet transform and adaptive threshold. Acta Electronica Sin. 2008;36(5):527–30.
4. Huang L, Zhao W, Wang J, et al. Combination of contrast limited adaptive histogram equalisation and discrete wavelet transform for image enhancement. Image Process Lett. 2015;9(10):908–15.
5. Tong Y, Chen J. Nonlinear adaptive image enhancement in wireless sensor networks based on non-sub-sampled shearlet transform. EURASIP J Wirel Commun Network. 2017;2017 (1):46.
6. Tong Y, Chen J. Compressive sensing image fusion in heterogeneous sensor networks based on shearlet and wavelet transform. EURASIP J Wirel Commun Network. 2017;2017(1):52.
7. Dong Y, Long T, Jiao W, et al. A novel image registration method based on phase correlation using low-rank matrix factorization with mixture of Gaussian. IEEE Trans Geosci Remote Sens. 2017;56(1):446–60.
8. Fan J, Wu Y, Wang F, et al. SAR image registration using phase congruency and nonlinear diffusion-based SIFT. IEEE Geosci Remote Sens Lett. 2014;12(3):562–6.
9. Cakir S, Kahraman DC, Cetin-Atalay R, Cetin AE. Contrast enhancement of microscopy images using image phase information. IEEE Access. 2018;6:3839–50.
10. Suthar M, Asghari H, Jalali B. Feature enhancement in visually impaired images. IEEE Access. 2017;6:1407–15.

# Hardware Implementation of Convolutional Neural Network-Based Remote Sensing Image Classification Method

Lei Chen[1,2], Xin Wei[1,2], Wenchao Liu[1,2], He Chen[1,2], and Liang Chen[1,2(✉)]

[1] Radar Research Lab, School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
chenl@bit.edu.cn
[2] Beijing Key Laboratory of Embedded Real-Time Information Processing Technology, Beijing Institute of Technology, Beijing 100081, China

**Abstract.** The convolutional neural networks have achieved very good results in the field of remote sensing image classification and recognition. However, the cost of huge computational complexity with the significant accuracy improvement of CNNs makes a huge challenge to hardware implementation. A promising solution is FPGA due to it supports parallel computing with low power consumption. In this paper, LeNet-5-based remote sensing image classification method is implemented on FPGA. The test images with a size of 126 × 126 are transformed to the system from PC by serial port. The classification accuracy is 98.18% tested on the designed system, which is the same as that on PC. In the term of efficiency, the designed system runs 2.29 ms per image, which satisfies the real-time requirements.

**Keywords:** CNN · Remote sensing image · FPGA · Classification

## 1 Introduction

The convolutional neural networks have performed very well in the field of remote sensing image processing [2, 3, 8]. In recent years, with the development of remote sensing technology, the remote sensing image classification puts forward the requirements of real-time, high precision, and low power consumption. Therefore, there is an urgent need for a real-time, reliable, and low-power remote sensing image classification scheme. At present, due to the large amount of data and the changing network architecture in convolutional neural networks, the convolutional neural networks are mainly implemented on CPUs or GPUs [1, 6, 7]. However, CPU can hardly make full use of the characteristics of parallel computing in convolutional neural networks. The computational efficiency is low, and it is difficult to meet the requirements of real-time processing. Similarly, although GPUs have high computational speeds, they have high power consumption and it is difficult to meet the applications requiring low power consumption.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_19

As a field-programmable gate array, FPGA can realize parallel computing and has low power consumption. At the moment, researchers have previously tried to speed up CNN algorithms by using FPGAs which is a promising way [5, 9, 10]. Zhang et al. [9] proposed an analytical design scheme using the roofline model to overcome the problem that the computation throughput may not well match the memory bandwidth provided by an FPGA platform. Zhang et al. [10] presented a deeply pipelined multi-FPGA architecture that expands the design space for optimal performance and energy efficiency. However, most of the implementations have limitations such as complex and difficult to achieve. In the context of remote sensing image classification, this article aims at real-time, accuracy, and low-power requirements. Therefore, the paper studied the deep pipeline structure between convolution and pooling layers, the cache optimization of intermediate calculation results, the efficiently read and write design of parameters and image data, and the problem of FPGA storage and processing bandwidth hardly matching. The convolutional neural network hardware implementation scheme for remote sensing classification based on FPGA was proposed in this paper.

The main work of this paper is as follows: (1) A remote sensing image classification hardware system based on FPGA was designed. (2) In order to weak bandwidth limitation, a read/write structure of data was designed. (3) Analyzing the relationship between convolution calculations and pooling calculations in LeNet-5, a deep flow structure between convolution calculations and pooling calculations was designed. (4) We test 2425 images of M-STAR test set on the designed hardware system. Each image takes 2.29 ms. The classification accuracy is 98.18% without accuracy loss, which is basically the same as the classification accuracy of the PC.

## 2 Network Architecture and Hardware System Architecture

The LeNet-5 network [4] is adopted to classify the remote sensing image in this paper. This network consists of two convolution layers, two pooling layers, and three full connection layers. In this paper, the pooling layer adopts mean pooling. The ReLu function is selected as activation function. M-STAR data sets are used in network training and testing. The image number in training set and test set is 2747 and 2425, respectively. The size of each image is cropped to $126 \times 126$.

Our design contains off-chip part and on-chip part of FPGA. The FPGA on-chip part consists of serial port module, image data processing module, convolution module, convolution bias and intermediate cache module, convolution weight parameter module, pooling module, DDR control module, and full connection module. The off-chip part consists of host computer and the off-chip memory. The overall system design block diagram is shown in Fig. 1.

For the off-chip part, the host computer is mainly responsible for the image data and some parameters transfer, as well as the task of displaying and saving the classification results uploaded by the FPGA. The off-chip memory is responsible for the storage of weight parameters. For the on-chip part, the serial port module receives the weight parameters and image data whose formats are single precision floating point; the DDR control module writes the parameter data received by the serial port module to the off-chip memory DDR3 through the Mig core; the image data processing module is

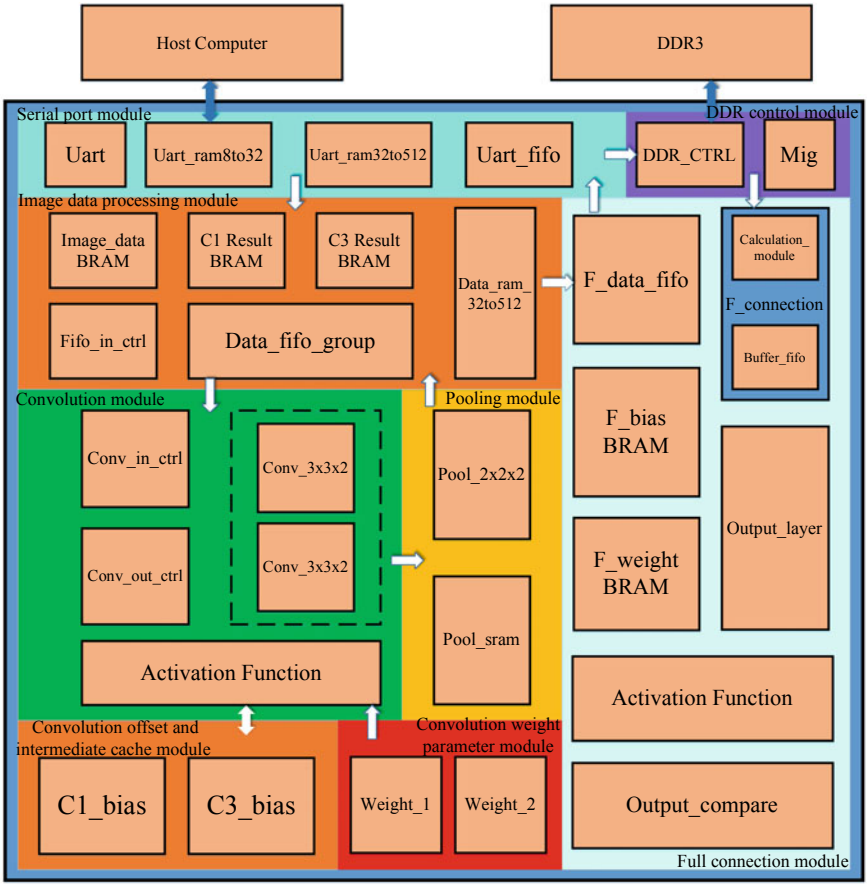



**Fig. 1.** Remote sensing image classification system architecture

responsible for buffering the image data calculated by the system and sorting the image data in a corresponding order; the convolution offset and the intermediate cache module and the convolution weight parameter module are responsible for the storage of the parameters in the convolution calculation and the caching of the intermediate results. The convolution module, the pooling module, and the full connection module complete the network's convolution calculations, pooling calculations, and full connection calculations, respectively. The implementation methods of image data processing module, convolution module, pooling module, and full connection module are crucial and will be introduced in detail in the next chapter.

## 3 Key Module Design

The system adopts mean pooling in the pooling layer with a stride of 2 and a sliding window size of 2. The hardware implementation is relatively direct. We mainly introduce the deep flow calculation between the pooling module and the convolution module. The implementation of activation function ReLu only needs to determine the sign bit of the input data. In addition, the system design selects 32-bit single precision floating point as the system design data type.

### 3.1 Image Data Processing Module

The image data processing module first receives the test image data from the serial port module and caches it in the Image_data BRAM. When the BRAM is full, the system start-up signal is pulled high. In the system calculation process, the image data processing module also completes the ordering of the image data and caches the results of the intermediate results.

In order to meet the design requirements for highly parallel flow between the convolution calculation and the pooling calculation, we use a parity row design to read and write image data. In this design, the image data of the four adjacent lines must be ready before the convolution operation. This task is accomplished through the Data_fifo_group module. The Data_fifo_group module consists of six Fifo_groups, each of which is combined with six fifos. They are numbered FIFO_1-FIFO_6. If it is specified that the time for writing one line of data in the fifo is an image data writing cycle, the time for reading one line of data is an image reading cycle. For the input of the C1 layer, the system locks the first Fifo_group, and each image data write cycle writes two lines of data into two fifos. After three cycles, the first Fifo_group is full and the image data output task starts. Each image data readout cycle reads six lines of data from six fifos at the same time, four lines of data in FIFO_1-FIFO_4 are output to the convolution module, and two lines of data in FIFO_3 and FIFO_4 are written back to FIFO_1 and FIFO_2. Two lines of data in FIFO_5 and FIFO_6 are written in FIFO_3 and FIFO_4. The implementation flow is shown in Fig. 2. The system then spends one image data write cycle, writing the new two lines of data into FIFO_5 and FIFO_6. Then repeat the above process to complete the next data read operation. For the convolution input of the C3 layer, considering the order of the output results of the previous pooling layer and maximize the parallel flow calculation of the system, the system first locks the first Fifo_group and writes two lines of data into the first two fifos, respectively, in an image data write cycle. Then, the system locks the second Fifo_group and writes two lines of data into the first two fifos, respectively, in an image data write cycle. By analogy, the first two fifos in the six Fifo_groups are filled after six cycles. Then, the system relocks the first Fifo_group, writes two lines of data into FIFO_3 and FIFO_4, and repeats the above process until the image data output task starts. The data output process is similar to the C1 layer, except that each image data readout cycle reads out six lines of data from six fifos in a Fifo_group at the same time.

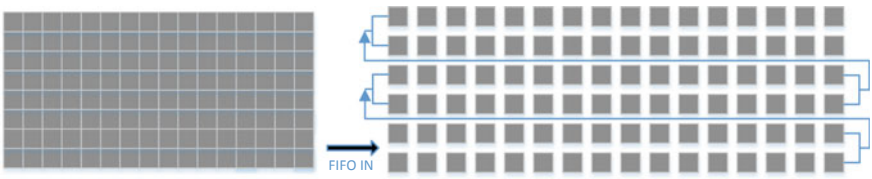



**Fig. 2.** A Fifo_group data readout process

### 3.2 Convolution Module and Pooling Module

The convolution calculation is to multiply the input image data $x$ with the weight parameter $w$ and then add the bias $b$, as shown in Eq. (1).

$$y_{i,j} = \sum_{m=0}^{N} \sum_{n=0}^{N} x_{i+m,j+n} w_{m,n} + b_{i,j} \tag{1}$$

The kernel size of convolution layers is $3 \times 3$ in this paper. Thus $N$ is 3. The traditional method of convolution is to take an image and a convolution kernel, then slide the kernel on the image to get a result. Then, the same operation is performed after switching the convolution kernel. The feature maps are accumulated after all the feature maps are obtained, as shown on the left of Fig. 3. But under normal circumstances, the size of the image data is much larger than the size of the convolution kernel. If the convolution kernel is fixed in this way and the image data is frequently switched, complicated address jumps are required for hardware logic, and the time cost for system reading the data is also increased. Even with the complexity of the network, such an implementation can seriously limit the data processing efficiency of the system. Therefore, this paper uses the implementation method shown on the right of Fig. 3. First take a sliding window of the image and then switch the convolution kernel until all the convolution kernels and the sliding window data of the image are calculated. Then switch the next sliding window of the image and calculate all the convolution kernels with it again. Although this implementation method cannot directly obtain the result of convolution calculation, it eliminates the complicated operation of frequent image replacement, which is beneficial to reduce the complexity of system control, shorten the time for system reading data, and reduce the pressure of system data processing bandwidth.

The system in this paper is designed to achieve deep flow between convolution calculations and pooling calculations. Considering that each pooling operation requires four adjacent points of the feature map, while fully utilizing the data processing bandwidth of the current architecture, we instantiate two convolution calculation modules in our design. Each convolution calculation module inputs two convolution kernels and one image sliding window data at the same time. That is, each convolution calculation module includes two convolution calculation units. Each convolution calculation unit performs a convolution of one image sliding window data and one convolution kernel to calculate a point of the feature map. Combined with the order of the image data processing module to input the image data of the convolution module,

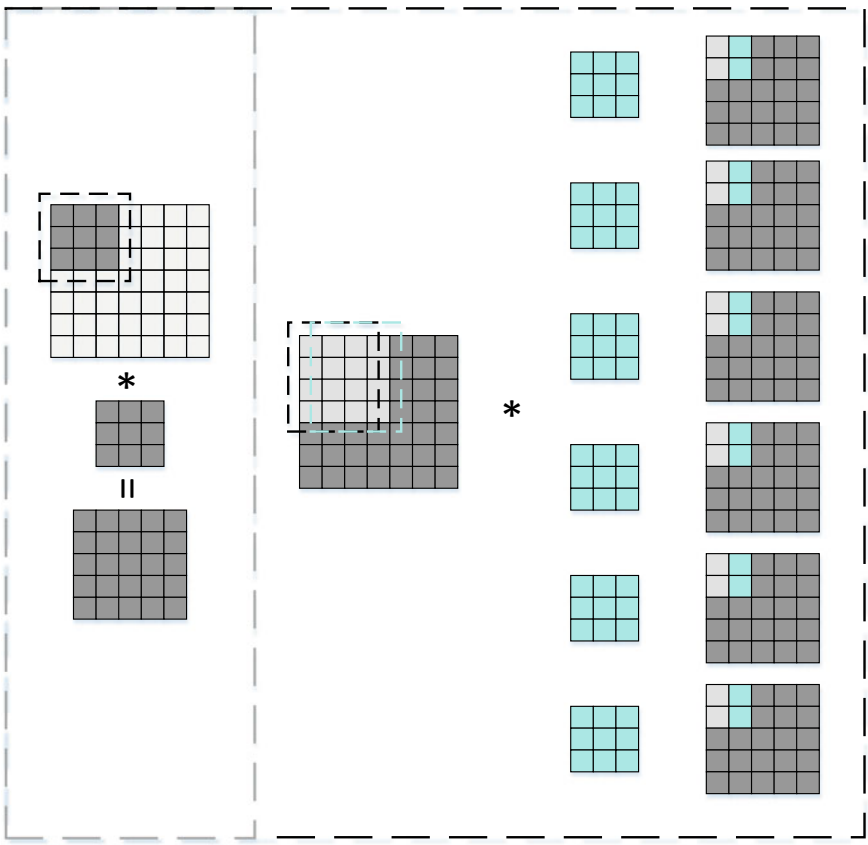


**Fig. 3.** Convolution calculation process diagram

it can be known that the convolution module outputs four points each time when a convolution calculation is completed. But these four points are not in the same feature map. Take the first two convolution kernels in Fig. 3 as an example, specifically, the convolution module completes a convolution calculation, in which a convolution calculation module outputs the first gray point of the first feature map and the second feature map. Another convolution calculation module outputs the second gray point of the first feature map and the second feature map. Therefore, these four points cannot be calculated directly to obtain the results of a pooling operation. It is necessary to wait until the convolution module completes the second convolution calculation to obtain four blue points in order to get the pooling results. This obviously does not allow convolution and pooling operations to achieve deep flow. Thus, this paper uses a clever design to solve this problem.

In this paper, the pooling module divides the pooling computing unit into an addition part and a division part. First, the system divides the four points sent by the convolution module for the first time into two groups, which are, respectively, sent to the adding part for addition, and temporarily buffer the calculation results. Then wait

until the convolution module sends four more points of the corresponding convolution kernel, and these four points complete the addition part calculation, the pooling module sends the output value of the corresponding addition part and the output value of the last buffer to the division part at the same time, and calculates two points of the two result feature maps in the pooling result.

### 3.3 Full Connection Module

The full connection operation needs to multiply the image data $x$ and the weight parameter $w$ and add the bias $b$, as shown in Eq. (2).

$$y_i = \sum_{j=0}^{N} x_j w_{j,i} + b_i \tag{2}$$

Therefore, the full connection layer needs a large number of multiply-add operations, and the calculation method is relatively simple, but it is necessary to focus on solving the problem that the bandwidth of the FPGA is difficult to match with the processing speed of the large data volume. In this design, high-speed Mig cores are used to achieve bidirectional transmission of parameter storage to achieve fast calculations. At the same time, the system design fully utilizes the data bandwidth of DDR3. The first- and second-tier full connection calculations use 16 multipliers as a full connection computing unit. This full connection computing unit receives 512 bits of data from DDR3 for processing in one clock cycle. That is to receive 16 numbers at the same time. The specific implementation process is shown in Fig. 4.

**Fig. 4.** Schematic diagram of full connection calculation

The input of the adder of the first multiplier inputs the bias of the full connection calculation, and then, the input of the adder of each multiplier inputs the output result of the previous multiplier. This design fully utilizes the data bandwidth of DDR3 and can realize the computing flow of full calculation, greatly improving the system's data processing efficiency. At the same time, it speeds up the speed of the full connection calculation and completes a large number of multiply-accumulate calculations efficiently.

## 4 Result Analysis

This paper designed a FPGA-based CNN hardware system for remote sensing image classification. First, the parameter data and the image data are sent from the host computer. Then, the system starts to calculate and return the classification results. The host computer receives the classification result and delivers the next test image data after saving the classification result. Finally, the classification results are evaluated by the host computer.

In this paper, the CNN is trained on PC with the assistance of TITAN Xp GPU. The test classification accuracy on the designed system is 98.18%, which is basically the same as the classification accuracy on PC.

In addition, in this study, the system uses a single precision floating point data type for calculation. The on-chip FPGA processing part is based on the Xilinx xc7k325t-ffg900 platform for hardware implementation. FPGA operating frequency is 100 MHz, DDR3 read and write frequency is 1600 MHz, so the FPGA to complete an image calculation time is 228,854 clock cycles, that is 2.29 ms or so, which fully proves that the design of the system meets the requirements of real time. The index parameters of the system are described in Table 1. The on-chip FPGA resource consumption is given in Table 2.

**Table 1.** Designed hardware implementation system partial index parameters

| | |
|---|---|
| Frequency | 100M |
| Classification accuracy | 98.18% |
| FPGA calculation time per picture | 228,804 cycles |

**Table 2.** FPGA resource occupation of the designed hardware implementation system

| | Resource consumption | Available | Occupancy rate (%) |
|---|---|---|---|
| LUT | 55,745 | 203,800 | 27.35 |
| FF | 45,561 | 407,600 | 11.18 |
| BRAM | 150.5 | 445 | 33.82 |

Table 2 shows that the FPGA resource occupation is in an acceptable range. However, according to FPGA area and speed conversion principles, FPGA hardware implementation design still has room for further optimization.

**Acknowledgments.** This work was supported by the Chang Jiang Scholars Programmed under Grant T2012122 and the Youth Science and Technology Innovation Leader of National Innovation Talent Promotion Program under Grant No. 2013RA2034.

## References

1. Garea AS, Heras DB, Argüello F. Caffe CNN-based classification of hyperspectral images on GPU. J Supercomputing. 2018;3:1–13.
2. Hu F, Xia GS, Hu J, et al. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. Remote Sens. 2015;7(11): 14680–707.
3. Liu W, Ma L, Chen H. Arbitrary-oriented ship detection framework in optical remote-sensing images. IEEE Geosci Remote Sens Lett. 2018;99:1–5.
4. Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proc IEEE. 1998;86(11):2278–324.
5. Qiu J, Wang J, Yao S, et al. Going deeper with embedded FPGA platform for convolutional neural network. In: Acm/sigda international symposium on field-programmable gate arrays. ACM; 2016. p. 26–35.
6. Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks. In: International conference on neural information processing systems. MIT Press; 2015. p. 91–9.
7. Vedaldi A, Lenc K. MatConvNet: convolutional neural networks for MATLAB; 2015:689–92.
8. Yang Y, Zhuang Y, Bi F, et al. M-FCN: effective fully convolutional network-based airplane detection framework. IEEE Geosci Remote Sens Lett. 2017;14(8):1293–7.
9. Zhang C, Li P, Sun G, et al. Optimizing FPGA-based accelerator design for deep convolutional neural networks. In: Acm/sigda international symposium on field-programmable gate arrays. ACM; 2015. p. 161–70.
10. Zhang C, Wu D, Sun J, et al. Energy-efficient CNN implementation on a deeply pipelined FPGA cluster. In: International symposium on low power electronics and design. ACM; 2016. p. 326–31.

# Deep 2D Convolutional Neural Network with Deconvolution Layer for Hyperspectral Image Classification

Chunyan Yu[1], Fang Li[1(✉)], Chein-I Chang[1,2], Kun Cen[1], and Meng Zhao[1]

[1] Dalian Maritime University, Dalian 116026, China
lifang@dlmu.edu.cn
[2] Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD, USA

**Abstract.** Feature extraction and classification technology based on hyperspectral data have been a hot issue. Recently, the convolutional neural network (CNN) has attracted more attention in the field of hyperspectral image classification. To enhance the feature extracted from the hidden layers, in this paper a deconvolution layer is introduced in the deep 2DCNN model. Analyzing the function of convolution and pooling to determine the structure of the convolutional neural network, deconvolution is used to map low-dimensional features into high-dimensional input; the target pixel and its pixels in a certain neighborhood are input into the network as input data. Experiments on two public available hyperspectral data sets show that the deconvolution layer can better generalize features for the hyperspectral image and the proposed 2DCNN classification method can effectively improve the classification accuracy in comparison with other feature extraction methods.

**Keywords:** Deep learning · Convolutional neural network · Hyperspectral image classification

## 1 Introduction

Due to the rich spectral information, hyperspectral images are widely used in the field of precision agriculture, forestry monitoring, and mining mapping, and classification is the most fundamental research of hyperspectral images. In the past, there are many methods to classify and extract features [10], such as nearest neighbor classifier [8], linear dimensionality reduction based on principal component analysis, support vector machine classifier, etc., but these classification methods all belong to shallow feature extraction [4]. For deep learning, it allows the computer to automatically extract deep features and extract more abstract features to improve the accuracy of the classification [9]. Wang [6] proposed a deep convolutional neural network classification model based on spectral information and spatial information of hyperspectral images simultaneously. Guo [3] proposed extreme learning machine to extract spatial information adaptively. However, these methods are a black box operation for feature extraction,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_20

and the effect of different hidden layers on the classification accuracy is not explored, and the complex model structure taken too long training time [1].

In this paper, a classification method based on 2DCNN with deconvolution layer is introduced. Through a series of operations such as convolution, pooling, deconvolution, and full connection to the input hyperspectral data, the network can obtain the approximate structure of the target function and better represent the image characteristics. The network structure of the CNN and the mathematical theory base used in each layer are introduced in Sect. 2. The related operation of the experiment and further analysis of the data in the experiment are listed in Sect. 3. Finally, conclusions are drawn in Sect. 4.

## 2 2DCNN with Deconvolution Layer Classification Method

The flow of the 2DCNN classification method based on the space spectrum combination can be simply summarized as: Firstly, in order to improve the accuracy and the convergence speed of the model, the input hyperspectral data is needed to be normalized, then the feature map is extracted through convolution layer and the nonlinear feature is obtained after activation, then the deconvolution layer is used to enhance the feature down sampled by pooling layer, finally the two fully connection layers finish the feature mapping to the softmax classifier.

### 2.1 Network Structure of CNN

The proposed model of CNN consists of seven layers, including two convolution layers (C1, C4), two pooling layers (P2, P5), one deconvolution layer (D3), and two full connection layers (F6, F7). For a sample, the input of the model is the spectral vector of the sample, and the output is the category of the sample. The convolution layer is used for feature extraction, and the pooling layer is used to reduce the sampling, thus reducing the amount of calculation. At last, the full connection layer and the softmax classifier are used to classify the hyperspectral data. The hierarchical structure diagram of CNN is shown in Fig. 1.

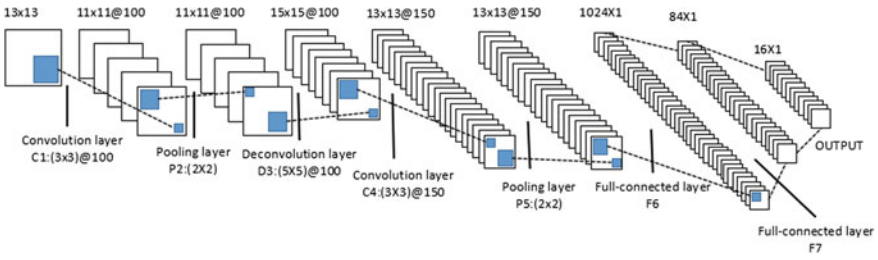



**Fig. 1.** Hierarchical structure diagram of convolutional neural network

### 2.2 Convolution and Pooling

Convolution is one of the most important concepts in depth learning. In convolution, certain parameter values are given to the convolution kernel, the CNN constantly updates these parameters in order to improve the feature extraction of the convolution kernel. This process is also called feature learning. The calculation equation for the convolution layer forward is defined as follows

$$\mu_{ij}^{xy} = b_{ij} + \sum_{p=1}^{P_i} \sum_{q=1}^{Q_i} w_{ij}^{pq} x_i^{(x+p)(y+q)}$$

where $\mu_{ij}^{xy}$ represents the value of the location $(x, y)$ of the $j$ feature graph on the $i$ level of the network after the convolution operation. $P_i$, $Q_i$ represent the size of the convolutional kernel space dimension. $w$ represents the convolution kernel, $x_i$ represents the data of the $i$ layer before the convolution operation.

For deep CNN, each hidden layer needs to connect a nonlinear activation function. ReLU is a nonsaturated nonlinear function, and its value is max $(0, x)$, it automatically regulates the number of nonzero values in the network through the gradient training [2]. The definition of ReLU function is listed as below

$$y_{ij}^{xyz} = \max(0, u_{ij}^{xyz})$$

For the pooling layer, the number of the input features and output features does not change. The function is to improve the robustness of the network, reduce the number of parameters, and prevent the occurrence of overfitting. Experiment selects the max-pooling function. The equation can be written as follows

$$\text{pool}_{\max}(R_k) = \max_{a_i \in R_k}(a_i)$$

where $R_k$ is the $k$ non-overlapping subregions.

### 2.3 Deconvolution

Deconvolution is also called transposition convolution, which is usually used to map low-dimensional features into high-dimensional input, and is contrary to convolution operations. The forward operation of the convolution layer can be expressed as a matrix multiplication in the TensorFlow, that is, the output $Y = CX$, then the deconvolution layer can be expressed $X = C^T Y$. It can be seen that the forward propagation process of the convolution layer is the reverse propagation process of the deconvolution layer. The forward and backward calculations of the convolution layers are multiplied by $C$ and $C^T$, the forward and backward calculations of the deconvolution layers are multiplied by $C^T$ and $(C^T)^T$, so their forward and reverse propagations are just exchanged. Deconvolution is also a process of upper sampling. Its definition is expressed as

$$\hat{y}_1^c = \sum_{k=1}^{K_1} z_{k,1} * f_{k,1}^c$$

where $Z$ represents the feature graph, and the $f$ represents the convolution kernel.

### 2.4 Full Connection and Softmax Classifier

After the alternation of the convolution layer and the pooling layer, the full connection layer is usually added to compress the extracted features, and the classifier is used to match the relationship between the eigenvalues and the tag values [5].

Suppose the number of sample sets is $m$, the input is $x$, the output is $h_{w,b}(x)$, the expected output is $y$, $a_i^{(l)}$ represents the output value of the $i$ unit of the $l$ layer, and the relationship between them can be written as follows

$$a_i^{(l)} = f(\sum_{j=1}^{s_{l-1}} w_{ij}^{(l)} x_j + b_i^{(l)})$$

where $s_{l-1}$ represent the number of neurons in the $l - 1$ layer.

Because the hyperspectral images classification is a multi-classification problem, it is necessary to connect a softmax layer after the full connection layer. The input data and output data of the softmax layer are both $N * 1$ vectors, but each value of the output vector must be between 0 and 1. The equation can be written as follows

$$S_i = \frac{e^{Z_i}}{\sum_k e^{Z_k}}$$

where $S_i$ represents the output of the $i$ neuron.

## 3 Experiment Result and Analysis

### 3.1 Experimental Setup

The experiment platform is listed as follows: PC (Intel Core i7 GPU, 8G RAM), the operating system is Ubuntu 16.04, the implementation framework is TensorFlow, and all the algorithms are implemented in Python.

The hyperspectral image databases used in the experiment are Indian Pines and PaviaU. Indian Pines database was gathered over the Indian Pines test site in Northwestern Indiana and consists of 145 × 145 pixels and 220 spectral reflectance bands. The main content is natural vegetation, which contains 16 types of ground material information. Pavia University image was acquired over the urban area of Pavia. The image consists of 610 × 340 pixels and 103 spectral reflectance bands [7].

In the experiment, the sample images are divided into training set, verification set, and test set according to a certain proportion, the proportion of training set is 10%. The size of the input data is 13 × 13 × 220. Table 1 shows the parameters setting of each layer for proposed network structure.

**Table 1.** Parameters setting of 2DCNN

| Class | C1 | P2 | D3 | C4 | P5 | F6 | F7 |
|---|---|---|---|---|---|---|---|
| Kernel size | 3 × 3 | 2 × 2 | 5 × 5 | 3 × 3 | 2 × 2 | – | – |
| Feature size | 11 × 11 | 11 × 11 | 15 × 15 | 13 × 13 | 13 × 13 | – | – |
| Feature number | 100 | 100 | 100 | 150 | 150 | 1024 | 84 |

### 3.2 Experiment Result and Analysis

In order to explore the effectiveness of CNN for feature extraction of hyperspectral data, we apply the extracted features to classification tasks and compare several other feature extraction methods, including SVM, ordinary CNN (there are three layers of network, including a convolution layer, a pooling layer, and a full connection layer.), and CNN without deconvolution layer. These feature extraction methods are selected because they are all commonly used methods for hyperspectral data processing and have achieved good results. Figure 2 shows the experiment result of the Indian Pines data set with the proposed method. Compared with the results of SVM, the feature extraction method based on CNN has obvious advantages, especially 2DCNN model with deconvolution layer; in addition to a few errors in the boundary, the classification of the ground objects is basically consistent with the real objects and shows a clearer classification boundary of the ground objects. Table 2 shows the classification accuracy of each class of the Indian Pines data set. We explored the effect of deconvolution layer on the model classification results. The classification accuracy of the CNN with deconvolution layer is higher than that of the CNN without deconvolution layer by two percent. In the same way, the experiment result of the PaviaU is shown in Fig. 3 and Table 3. It can be seen that the classification accuracy of the most categories obtained by the 2DCNN with deconvolution layer is higher than 99%, some other categories even up to 100%, only the accuracy of the seventh class bitumen is 95%.

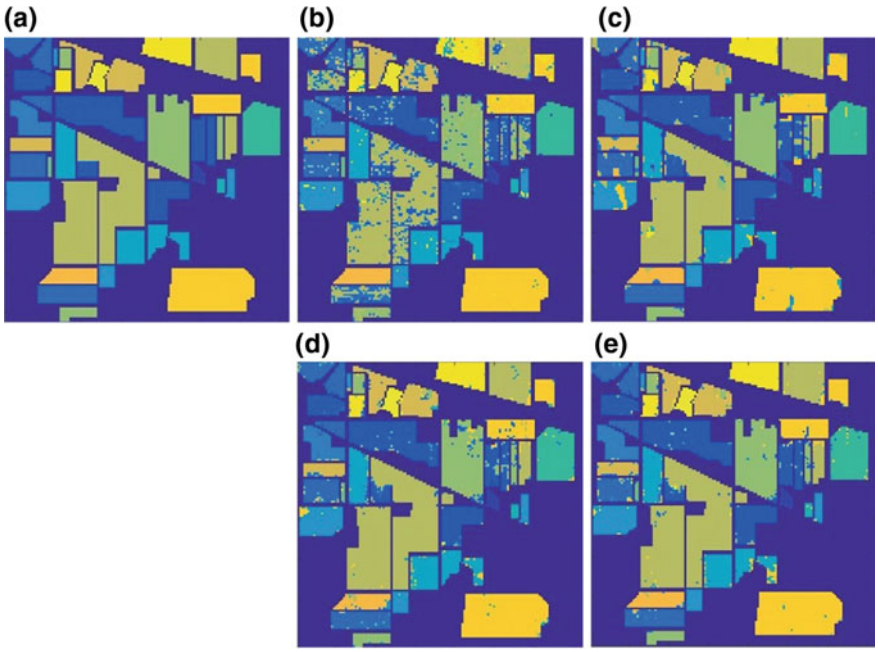



**Fig. 2.** **a** Original ground-truth image; **b** the result of SVM; **c** the result of ordinary CNN; **d** the result of CNN without deconvolution layer; **e** the result of CNN with deconvolution layer

**Table 2.** Classification accuracy of each class of Indian Pines

| CLASS | SVM | Ordinary CNN | 2DCNN without deconvolution layer | 2DCNN with deconvolution layer |
|---|---|---|---|---|
| 1 | 89.13 | 93.47 | 1 | 1 |
| 2 | 67.09 | 92.43 | 92.78 | 95.58 |
| 3 | 63.49 | 90.60 | 93.85 | 92.77 |
| 4 | 88.19 | 97.46 | 93.67 | 98.73 |
| 5 | 88.20 | 84.26 | 93.16 | 96.89 |
| 6 | 93.84 | 92.32 | 88.63 | 87.67 |
| 7 | 92.86 | 92.85 | 85.71 | 85.71 |
| 8 | 95.40 | 98.53 | 97.69 | 96.65 |
| 9 | 75.00 | 80.00 | 80.00 | 90.00 |
| 10 | 68.21 | 93.51 | 90.32 | 94.13 |
| 11 | 71.85 | 94.66 | 97.02 | 97.10 |
| 12 | 70.99 | 86.84 | 88.36 | 89.71 |
| 13 | 99.51 | 91.21 | 84.87 | 94.14 |
| 14 | 89.72 | 94.70 | 96.28 | 98.33 |
| 15 | 67.10 | 90.93 | 94.81 | 97.15 |
| 16 | 100.00 | 82.79 | 97.84 | 95.69 |
| OA | 76.91 | 92.70 | 93.76 | 95.22 |

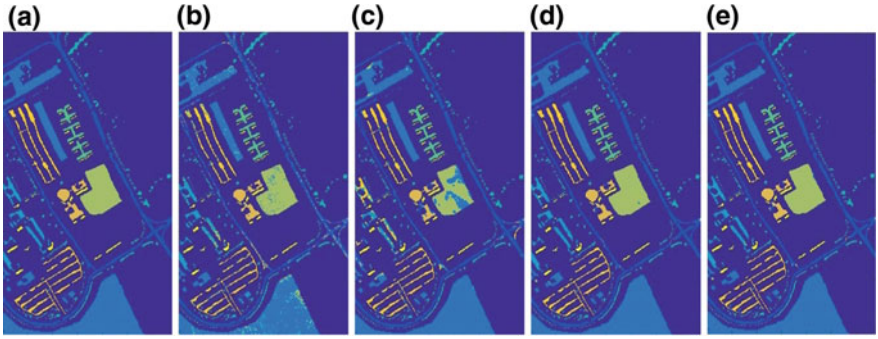



**Fig. 3.** **a** Original ground-truth image; **b** the result of SVM; **c** the result of ordinary CNN; **d** the result of CNN without deconvolution layer; **e** the result of CNN with deconvolution layer

**Table 3.** Classification accuracy of each class of PaviaU

| CLASS | SVM | Ordinary CNN | 2DCNN without deconvolution layer | 2DCNN with deconvolution layer |
|---|---|---|---|---|
| 1 | 89.11 | 95.06 | 98.83 | 99.17 |
| 2 | 93.30 | 98.91 | 99.86 | 99.89 |
| 3 | 86.90 | 48.64 | 97.85 | 98.47 |
| 4 | 97.32 | 98.49 | 99.31 | 99.54 |
| 5 | 99.85 | 98.88 | 100 | 100 |
| 6 | 93.56 | 73.31 | 99.62 | 99.22 |
| 7 | 95.94 | 81.80 | 94.96 | 95.26 |
| 8 | 88.38 | 93.18 | 99.42 | 99.37 |
| 9 | 99.89 | 99.26 | 99.47 | 99.47 |
| OA | 92.66 | 91.79 | 99.35 | 99.41 |

## 4 Conclusion

The classification technology of hyperspectral data has great application value and development space. In order to obtain better classifier, a CNN model with deconvolution layer is proposed in this paper. By analyzing the data processing methods of each layer network structure and discussing the function of the deconvolution layer for feature extraction, a network model is constructed. Experiments on two HSI data sets show that the accuracy of classification is improved obviously, which verifies the effectiveness of deconvolution for feature extraction.

## References

1. Chen X. Hyperspectral image classification using deep learning method. China University of Geosciences; 2016.
2. Guo K, Li N. Research on classification of architectural style image based on convolution neural network. In: IEEE, information technology and mechatronics engineering conference. IEEE; 2017. p. 1062–6.
3. Guo LL. Research on image classification algorithm based on deep learning. Doctoral dissertation, China University of Mining and Technology; 2016.
4. Lu X, Chen Y, Li X. Hierarchical recurrent neural hashing for image retrieval with hierarchical convolutional features. IEEE Trans Image Process. 2017;99:1.
5. Song M, Chang CI. A theory of recursive orthogonal subspace projection for hyperspectral imaging. IEEE Trans Geosci Remote Sens. 2015;53(6):3055–72.
6. Wang Q. Classification for hyperspectral remote sensing image based on deep learning. Doctoral dissertation, Huaqiao University; 2016.
7. Wang Y, Lee LC, Xue B et al. A posteriori hyperspectral anomaly detection for unlabeled classification. IEEE Trans Geosci Remote Sens. 2018:1–16.
8. Yu C, Xue B, Song M, et al. Iterative target-constrained interference-minimized classifier for hyperspectral classification. IEEE J Sel Top Appl Earth Obs Remote Sens. 2018;99:1–23.
9. Yu C, Lee LC, Chang CI et al. Band-specified virtual dimensionality for band selection: an orthogonal subspace projection approach. IEEE Trans Geosci Remote Sens. 2018;99:1–11.
10. Zhao M, Zhang J, Tao C. Land use classification in remote sensing images based on deep convolution neural network. In: Asia-Pacific computational intelligence and information technology conference. Shanghai; 2017.

# Research on Video Compression Technology for Micro-Video Applications

Dongna Cai(✉), Yuning Li, and Zhi Li

College of Art and Design, Beijing Forestry University, Beijing 100083, China
caidongna@bjfu.edu.cn

**Abstract.** Micro-video has fragmented propagation mode and short, flat, and fast features. The convenience of its shooting and dissemination is widely used by the public. In addition to the development of Internet technology, the video compression technology has also played an important role. In this paper, three kinds of micro-video compression coding techniques are selected and analyzed; they are MPEG-4, AVC/H.264, and HEVC/H.265. The three video compression techniques are compared by the film trailer compression experiment, and the conclusion is applied to the micro-video compression of scientific research projects. Combining the actual advantages and disadvantages of the scheme, it is beneficial to the effective application of micro-video coding technology in practical projects.

**Keywords:** Micro-video · Video compression · MPEG-4 · H.264 · H.265 · Effective application

## 1 Introduction

Micro-video refers to video recording or editing using cameras, digital cameras, or smartphones for no more than five minutes. The concept of micro-video can be traced back to the mobile micro-video social applications published by Viddy in the USA in 2011. In recent years, the popularity of micro-video applications has gradually attracted the attention of scholars. In 2014, Oskar Juhlin et al. mentioned the rapid development of micro-video on the Internet in the research agenda article on the video interaction [1] and Ma Xiaoqiang et al. studied the video sharing mode of class YouTube social [2]. Li Xiaotong used a popular micro-video mobile phone 2017 as an example to analyze the communication strategy and effect of micro-video [3]. And in 2018, Li Junzuo combed the timeline and popular reasons for the development of micro-video [4]. The above documents analyze the micro-video from the perspective of network technology, information dissemination, and mode development.

The micro-video meets the needs of mass communication in the Internet era; the upgrading of video compression and decoding technology has played a key role in the popularization of micro-videos. For the development and application of video compression technology sequence of thought, in 2012, a new method of compressing video data for two Bezier curves was confirmed by Khan [5]. In 2013, Sambuddha Kumar et al. studied color video compression based on color mapping to grayscale video frames [6]. In summary, the above literature is the research on the macrovideo


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_21

compression technology in recent years. While for the specific compression technology, in 2015, Rowayda Sadek et al. elaborated on IEEE802.11e wireless Mesh MPEG-4 video transmission based on dynamic cross-layer method [7]. In 2017, Chih-Heng Ke et al. improved the layered packet pre-discarding method and overloading MPEG-4 video transmission wireless network [8]. In 2009, Song Li et al. published a research article on improving the intra-coding method of H.264/AVC [9]. An efficient intra-frame rate control algorithm for H.264/AVC video was proposed by Dandan Zhao in 2012 [10]. In 2015, Grzegorz Pastuszak obtained the research results for the flexible structure design of H.265/HEVC inverse transformation [11]. In 2018, S. Gnanavel and S. Ramakrishnan explored for uhB high-definition video transmission based on H.265 encoder and ANFIS rate controller [12].

Through the literature review and analysis, the author found that academic research on micro-video and video compression is common, but research based on micro-video compression is rare. After network research, the iOS system, OS system, Android system, Instagram, WeChat, and Twitter all use AVC/H.264 video compression and decoding technologies. As a representative of advanced video decoding technology in recent years, HEVC/H.265 is widely used in network video live broadcasting and real-time monitoring. MPEG-4 technology has been available for many years and is still active in online video and transcoding software. According to the above network research and academic literature,we selected the three micro-video compression techniques currently frequently used, namely MPEG-4, AVC/H.264, and HEVC/H.265. Compare and analyze these three video compression coding methods in the film trailers micro-video and gain the advantages and disadvantages of three coding techniques through experiments, meanwhile applying the experimental results to scientific research projects.

## 2 Frequently Used Micro-Video Compression Technology

Since 1980, The International Organization for Standardization, The International Electrotechnical Commission, and the International Telecommunication Union have developed an international series of image coding video compression standards, including MPEG-4, AVC/H.264, and HEVC/H.265 [13].

### 2.1 MPEG-4 Compression Coding Technology and Its Working Principle

MPEG-4 is a fourth-generation dynamic image decompression protocol developed by ISO and IEC in 1998 and officially became an international standard in early 2000. It is specially designed for playing high-quality video of streaming media in order to obtain the best image quality with the least amount of data, numbered ISO/IEC 14496.

The MPEG-4 video decoder consists of five modules: an entropy decoder, inverse quantization, inverse discrete cosine transform, motion compensation module, and video frame buffer. The working principle of the MPEG-4 video decoder can be interpreted as follows: First, the input stream is entropy-decoded, and then the header information of the frame determines the type of the frame. For each macroblock, after entropy decoding, IQ is first passed, and IDCT is used to obtain the value of the empty domain. For the reference frame (R_Frame), since no motion compensation is required,

the transformed result is directly output, and it is also stored in the video frame buffer and left to the following prediction frame (P-Frame) for motion compensation. For the prediction frame, the motion vector is first obtained by entropy decoding; after the corresponding reference frame is searched according to the motion vector, the IDCT-transformed prediction difference is added to synthesize the final prediction frame image (Fig. 1).

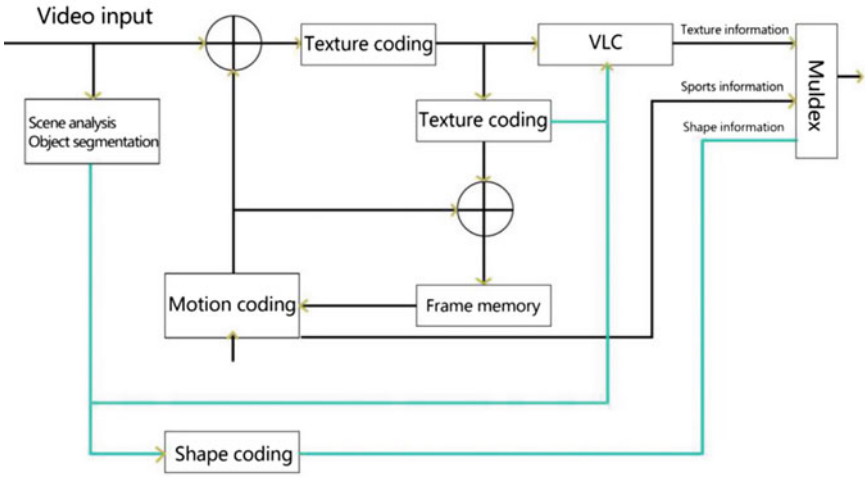



**Fig. 1.** Basic structure flowchart of MPEG-4 coding

## 2.2 AVC/H.264 Compression Coding Technology and Its Working Principle

AVC/H.264 is digital video coding standards designed for multimedia transmission defined by ITU-T and ISO.

The working principle of H.264 encoding can be interpreted as follows: First, a macroblock ($F_n$) to be encoded is taken from the current input video image and the macroblock is encoded in an intra-frame or inter-frame manner to generate a predicted macroblock $P$. Under inter-mode coding, $P$ is derived from one or more reference pictures for motion compensated (MC) prediction ME. The current macroblock ($F_n$) subtracts the predicted macroblock $P$ to obtain a residual block $D_n$, an integer transform $T$ is performed on the residual block $D_n$, a set of coefficients $X$ is obtained after quantization $Q$, and the coefficient $X$ is further reordered and entropy encoded. The entropy coded code stream plus some information required for macroblock decoding, such as prediction mode, quantization step size motion vector information, etc., constitutes a compressed code stream of the macroblock. The code stream adds the slice header information to compose the coded code stream of the slice and then transmits or stores through the network abstraction layer NAL (the image reference set PPS and the sequence parameter set SPS are separately transmitted). The reconstruction branch is

mainly used for the prediction of the encoding and the confirmation of the related information at the time of decoding. Its flow: the macroblock coefficient $X$ is inverse quantized and inverse transformed, and the approximate value of the residual macroblock $D'_n$ is obtained. Predicting block P plus $D'_n$ results in an unfiltered reconstructed macroblock uFn′, and do loop. Add $D'_n$ to get the unfiltered reconstructed macroblock $uF'_n$, and do loop filtering to reduce the blockiness, to get the final reconstructed macroblock $uF'_n$. After all the macroblocks in the current image are reconstructed, a reconstructed image is formed (Fig. 2). H.264 the standard encoding idea is the mix coding mode, It uses inter-frame and intra-frame prediction to remove the redundant components of space and time, and transform and quantization coding to remove frequency domain redundancy components. The H.264 standard enhances the adaptability to various channels, which is applicable to different transmission and storage scenarios with different rates and resolutions.

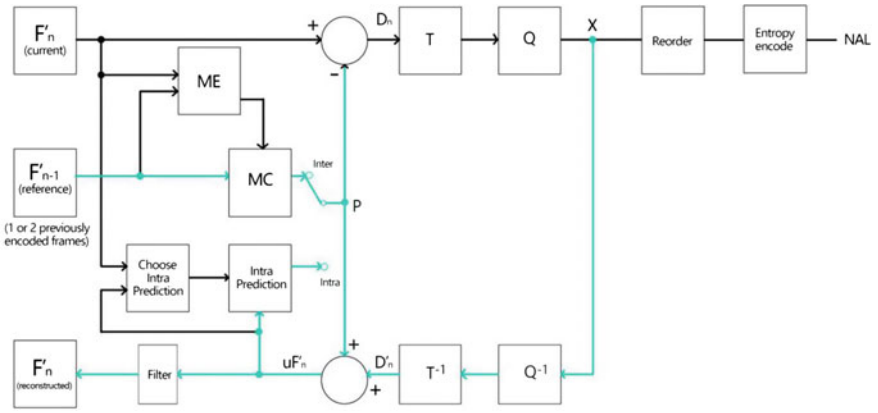



**Fig. 2.** Basic structure flowchart of AVC/H.264 coding

## 2.3 HEVC/H.265 Compression Coding Technology and Coding Structure

The coding structure of H.265 (Fig. 3) mainly includes intra-prediction, inter-prediction, transform, quantization, deblocking filter, entropy coding, and so on. However, in the HEVC coding architecture, the whole is divided into three basic units: coding unit (CU), prediction unit (PU), and transformation unit (TU). Compared with AVC/H.264, the coding architecture of HEVC/H.265 is roughly similar to that of AVC/H.264, but it is improved at all major technical points and is more complex than H.264 (comparing Figs. 1 and 2), HEVC/H.265 provides more different tools to reduce the code rate, in terms of coding units, each macroblock/MB size in H.264 is a fixed 16 * 16 pixels, and H.265 coding unit can be chosen from the smallest 8 * 8 to the largest 64 * 64 pixels, which shows that the key coding of the image reduces the overall code rate and the coding efficiency increases accordingly especially the adaptive quadtree partitioning of the image block is introduced, and a series of variable-scale

adaptive coding techniques for image texture features are used to greatly improve the compression capability of the flat region of the image (Fig. 3).

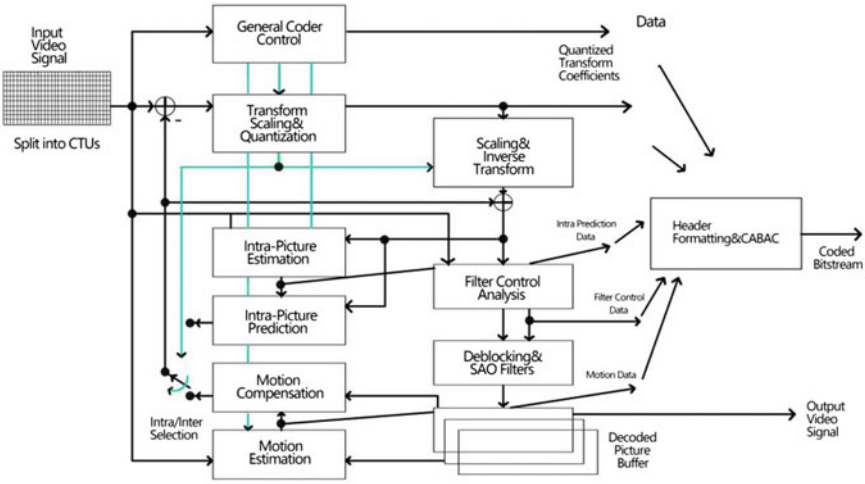



**Fig. 3.** Basic structure flowchart of HEVC/H.265 coding

## 3 Experiment and Comparative Analysis

Through the intuition comparison of the above three common video compression methods, this chapter has performed two sets of experiments to analyze them in detail.

### 3.1 Film Trailer Compression Experiment Process and Result

In order to take into account the characteristics and propagation factors of micro-videos, the average time of the selected material does not exceed 3 min, and the sample of the experiment was selected for the Iqiyi network video platform by random selection of 50 film trailer of resolution ratio 1280 * 720, using MPEG-4, H.264, and H.265 coding for re-coding and compression. The average size and video bit rate of original videos are calculated. The average video volume is 13.6 MB, the frame rate is 25 FPS, the video bit rate is 1050.76 kb/s, and the format is QSV. Second, recompress and record the time required for encoding to obtain three sets of data. Calculate the average size of the three groups of data and the average video bit rate, and check whether the frame rate and resolution have changed (Table 1).

**Table 1.** Source video information (average)

| Number | Video volume (MB) | Frame rate (FPS) | Resolution ratio (PX) | Video bit rate (kb/s) | Format |
|---|---|---|---|---|---|
| 50 | 13.6 | 25 | 1280 * 720 | 1050.76 | QSV |

The formula used in the experiment is

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

$$\overline{x} = \frac{w_1 x_2 + w_2 x_2 + \cdots + w_n x_n}{w_1 + w_2 + \cdots + w_n} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \tag{2}$$

### 3.2 Comparison and Analysis of Compression Results of Film Trailers

After re-encoding, compressing, and calculating the source of the image, the comparison and the data in Table 2 of Fig. 4 are obtained. If we only use the naked eye to distinguish, the difference between the three coding methods of experimental results is not easy to find, but through the comparison of the data, we can find that when the encoding method is MPEG-4, the video size increases by an average of 10.3 MB than the original video and the transcoding takes a short time. When the 50 video transcodes are shared, the average frame rate has decreased and the resolution has not changed. The video bit rate is higher than the original video code rate and H.264 and H.265 encoding. So the MPEG-4 re-encoded video file is very close to the original file in terms of accuracy, but the video volume is slightly larger, and it cannot meet the demand of large-area compression volume. After re-encoding and compressing the film trailers, the data in Table 2 and Fig. 4 are obtained. Its frame rate and resolution have

**Table 2.** Source video after transcoding through MPEG-4, H.264, and H.265 (average)

| Encoding method | Video volume (MB) | Time (S) | Frame rate (FPS) | Resolution ratio (PX) | Video bit rate (kb/s) | Format |
|---|---|---|---|---|---|---|
| MPEG-4 | 23.9 | 834 | 24.4 | Unchanged | 1728.1 | Mp4 |
| H.264 | 18.67 | 1640 | 25 | Unchanged | 1294.2 | Mp4 |
| H.265 | 13 | 14,700 | 25 | Unchanged | 870.84 | Mp4 |

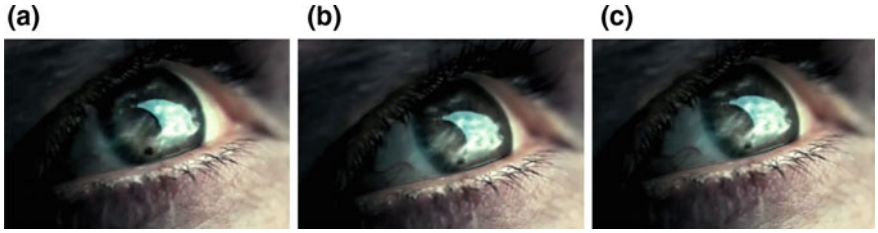



**Fig. 4.** Micro-video picture quality contrast after compression by three coding technologies. **a** MPEG-4 compress picture quality, **b** H.264 compression picture quality, and **c** H.265 compression picture quality

not changed, and the video bit rate is 433.9 kb/s less than the average MPEG-4 encoding. Compared with MEPG-4 encoding, the advantages of H.264 lie in the size of the video, as much as possible to restore the video accuracy. The H.265 encoding method is undoubtedly the most time-consuming one of these three encoding methods. Its advantage is maximizing the size of compressed video.

### 3.3 Experimental Research Based on Micro-Video Compression of Flowers

After the experimental analysis of the compression and decoding of the film trailer, three kinds of common micro-video compression techniques are given a specific application method, and the method is applied to the practical research project. This section mainly selects 10 micro-videos of flower material taken by mobile phone of iOS11 system. The average size and video rate of the material are calculated and arranged. The frame rate and other basic data are collected to get the resolution of 1920 * 1080, the average size is 116.1 MB, the time length is 62.8 s, the frame rate is 29.98 FPS, and the video bit rate is a bit rate. For 15,361.8 kb/s, the format is MOV. Then, the 10 videos are coded and compressed using MPEG-4, H.264, and H.265 codes, respectively (Table 3).

**Table 3.** Filming the original information of the flower material (average)

| Number | Video volume (MB) | Frame rate (FPS) | Resolution ratio (PX) | Video bit rate (kb/s) | Format |
|---|---|---|---|---|---|
| 10 | 116.11 | 29.98 | 1920 * 1080 | 1050.76 | MOV |

### 3.4 Analysis of Micro-Video Compression Experiments on Flowers

After the video material is re-encoded, compressed, and calculated, the comparison in Fig. 5 and the data in Fig. 6 are obtained. Through data comparison, when the encoding method is MPEG-4 the video size is 86.83 MB lower than the original video on average, the transcoding takes a short time, and 10 video transcodings share 341 s. The average frame rate has decreased, while the resolution has not changed. The video bit rate is higher than the original video bit rate and H.264, H.265 encoding, and the processed file is close to the original file. After re-encoding and compressing the flower micro-video, the data in Figs. 5 and 6 are obtained. Its frame rate and resolution have not changed, and the video bit rate is 1032.6 kb/s H.265 less than the average MPEG-4 encoding. The H.265 encoding mode is still the most time-consuming one among these three encoding methods, but its advantage is to maximize the volume of compressed video. Also demonstrated in this experiment, it maximizes the compression of the video volume, the average size of each video is only 14.64 MB, an average of 101.47 MB smaller than the original video material, and its code rate is higher than the original video code rate.

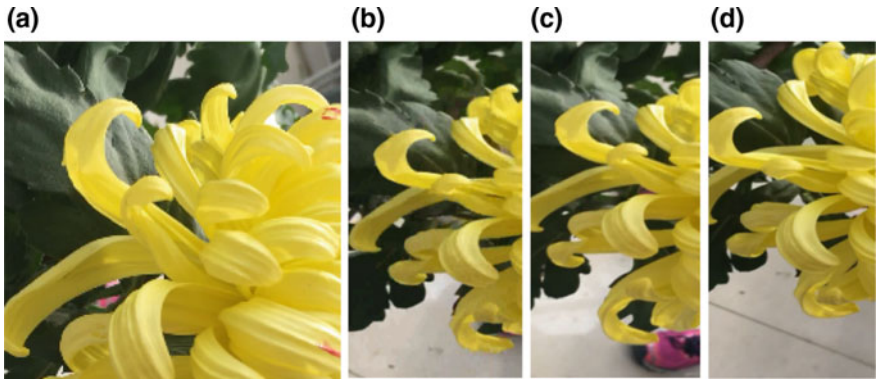



**Fig. 5.** Comparison of picture quality after micro-video compression of Chrysanthemum **a** material quality, **b** MPEG-4 compression picture quality, **c** H.264 compression picture quality, and **d** H.265 compression picture quality

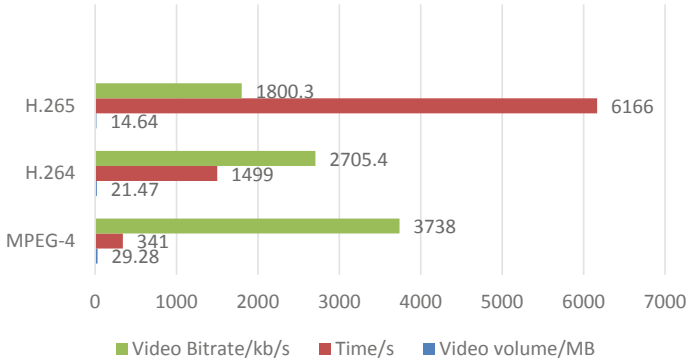



**Fig. 6.** Source video is transcoded by MPEG-4, H.264, and H.265

## 4 Conclusion

With the wide application of micro-video, the video decoding and compression method has become an important technical element for the development of micro-video transmission. This paper introduces three common micro-video compression methods: MPEG-4, H.264, and H.265 and a new decoding method AV1 that may become a trend in the future. Through experimental data analysis and practical application of three common video compression methods, it is concluded that the MPEG-4 decoding compression method compresses the size of the video volume yet it has little advantage compared with H.264 and H.265. Its advantage lies in the restoration of video accuracy and decoding speed, so it is suitable for projects and people who are pursuing video quality, purely converted formats, and short-term expectations. The H.264 decoding mode is in a balanced position among the three decoding modes, and its advantage lies in balancing the relationship between the size of the micro-video volume and the video

accuracy. At the same time, the decode time is relatively short. It is the most compromised choice at present and is suitable for the network or UGC platform application. The advantage of the H.265 decoding method lies in the fact that the size of the video is compressed while maintaining the video accuracy; this is the most advantageous method for compressing large video files and the storage and project of micro-video material in the after effect. Creating a 3D model with reference to a micro-video provides a better solution for the use of micro-video storage cloud platforms, micro-video sharing platforms, and high-tech electronic devices. The disadvantages of H.265 video compression coding are excessive time and excessive consumption of equipment which is a problem that its users must face.

**Acknowledgments.** The author would like to thank Beijing Forestry University for support. This research is supported by the funded projects: Fundamental Research Funds for the Central Universities (2016JX08).

## References

1. Juhlin O, et al. Video interaction: a research agenda. Pers Ubiquit Comput. 2014;18(3):685–92.
2. Ma X, et al. Exploring sharing patterns for video recommendation on YouTube-like social media. Multimedia Syst. 2014;20(6):675–91.
3. Li XT. Research of short video communication strategy and effect-YiTiao TV as an example. Master's thesis of Anhui University; 2017. p. 69.
4. Li JZ. The rise and development of short video. Youth Journalist. 2018;05:95–6.
5. Khan MA. A new method for video data compression by quadratic Bézier curve fitting. SIViP. 2012;6(1):19–24.
6. Kumar S, Gupta S. Color video compression using color mapping into textured grayscale video frames. Pattern Anal Appl. 2014;17(4):809–22.
7. Sadek R, Youssif A, Elaraby A. MPEG-4 Video transmission over IEEE 802.11e wireless mesh networks using dynamic-cross-layer approach. Nat Acad Sci Lett. 2015;38(2):113–9.
8. Ke C, Yang C, Chen J. Hierarchical packet pre-dropping approach for improved MPEG-4 video transmission in heavily loaded wireless networks. Mobile Netw Appl. 2017;22(1):30–9.
9. Song L, et al. Improved intra-coding methods for H.264/AVC. EURASIP J Adv Sig Proces. 2009(1).
10. Zhao D, et al. An efficient intra-frame rate control algorithm for H.264/AVC video coding. Wuhan Univ J Nat Sci. 2012;17(3):243–8.
11. Pastuszak G. Flexible architecture design for H.265/HEVC inverse transform. Circ Syst Sig Process. 2015;34(6):1931–45.
12. Gnanavel S, Ramakrishnan S. HD video transmission on UWB networks using H.265 encoder and anfis rate controller. Cluster Comput. 2018;21(1):251–63.
13. Meng B. Analysis of main video compression technology development in mainland China. China Integr Circ. 2006(10):62–72.

# An Adaptive Iteratively Weighted Half Thresholding Algorithm for Image Compressive Sensing Reconstruction

Qiwei Peng[1], Tongwei Yu[2], Wang Luo[1(✉)], Tong Li[2], Gaofeng Zhao[1], Qiang Fan[1], Xiaolong Hao[1], Peng Wang[1], Zhiguo Li[1], Qilei Zhong[1], Min Feng[1], Lei Yu[1], Tingliang Yan[1], Shaowei Liu[1], Yuan Xia[1], Bin Han[1], Qibin Dai[1], Yunyi Li[3], Zhenyue Zhang[3], and Guan Gui[3]

[1] NARI Group Corporation/State Grid Electric Power Research Institute, Nanjing, China
luowang@sgepri.sgcc.com.cn
[2] State Grid Liaoning Electric Power Supply Co. Ltd., Shenyang, China
[3] Nanjing University of Posts and Telecommunications, Nanjing, China
guiguan@njupt.edu.cn

**Abstract.** The $L_{1/2}$ regularization has been considered as a more effective relaxation method to approximate the optimal $L_0$ sparse solution than $L_1$ in CS. To improve the recovery performance of $L_{1/2}$ regularization, this study proposes a multiple sub-wavelet-dictionaries-based adaptive iteratively weighted $L_{1/2}$ regularization algorithm (called MUSAI-$L_{1/2}$), and considering the key rule of the weighted parameter (or regularization parameter) in optimization progress, we propose the adaptive scheme for parameter $\lambda_d$ to weight the regularization term which is a composition of the sub-dictionaries. Numerical experiments confirm that the proposed MUSAI-$L_{1/2}$ can significantly improve the recovery performance than the previous works.

**Keywords:** $L_{1/2}$ regularization · Multiple sub-wavelet-dictionaries · Enhancing sparsity · Adaptive · Iteratively weighted

## 1 Introduction

With the explosive growth of Internet of things (IoT) devices and the growing communication requirements, we need more bandwidth to transmit large amounts of data. However, the traditional Nyquist sampling method will cause a large amount of samples; thus, it is inefficient and challenging in the future. Compressive sensing (CS) is a new technique to the reconstruction and acquisition of high-dimensional signal from a lower one, which can be used to tackle some challenges in information processing system, including efficient transmission, low-complexity encoder, and less network congestion [1]. In the CS framework, we are interested in recovering signal/image from the $M \ll N$ linear measurement system $\mathbf{y} = \mathbf{\Phi x} + \mathbf{n}$, where $\mathbf{\Phi} \in R^{M \times N}$ is a given matrix called the sensing matrix, $\mathbf{y} \in R^{N \times 1}$ is the measured data, and


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_22

$\mathbf{n} \in R^{N\times 1}$ denotes the noise. Such ill-posed problem formally called the underdetermined inverse problem, which has arisen in image processing [2], wireless communications [3], and array signal processing [4]. The problem can be resolved well theoretically called $L_0$ regularization described as follows:

$$\hat{\mathbf{x}}_{l_0} = \arg\min_{\hat{\mathbf{x}}} \left\{ \gamma \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0 \right\} \tag{1}$$

where $\|\mathbf{x}\|_0 = |\mathrm{supp}(\mathbf{x})|$ denotes $L_0$-norm to count the number of nonzero components and $\mathrm{supp}(\mathbf{x}) = \#\{i, x_i \neq 0\}$ denotes support set of nonzero components in signal vector $\mathbf{x}$; $\gamma > 0$ is a tuning parameter that can reflect the noise level; and $\lambda > 0$ is a regularization parameter and plays a role in balancing the two terms of the square error and the sparsity term. However, solving the $L_0$ regularization problem in (1) is a NP (non-deterministic polynomial) problem [5]. As a relaxed method for approximately computing (1), the $L_1$ regularization is one of the most popular and efficient methods by solving:

$$\hat{\mathbf{x}}_{l_1} = \arg\min_{\hat{\mathbf{x}}} \{ \gamma \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 + \boldsymbol{\lambda} \|\mathbf{x}\|_1 \} \tag{2}$$

where $\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |x_i|$, denotes $L_1$-norm operator. The $L_1$ regularization (2) is a typical convex optimization problem. Hence, it can be resolved easily via linear program, which can be considered as a relaxation of the $L_0$ regularization (1). In other words, the solution of (2) can be approximated coinciding with the solution of $L_0$ regularization satisfying certain conditions such as [6]. However, the $L_1$-norm penalized regularization can lead to the sub-optimal solution in many sparse recovery cases. Theoretical analysis of CS implies that better performance can be obtained by taking advantage of sparser information in many systems especially in the presence of strong noise interference.

According to the theoretical analysis [7–9], the $L_p$ regularization will recover sparse signal with fewer measurements (lower sampling ratio) and may enhance the robustness and the stability under noise scenarios. Recently, the $L_{1/2}$ regularization is adopted as a typical $L_p$ regularization method and can yield a stronger sparsity solution than the $L_1$ regularization. A well-known half proximal thresholding-based algorithms has been proposed in [10]. The $L_{1/2}$ regularization problem is defined as below:

$$\hat{\mathbf{x}}_{l_{1/2}} = \arg\min_{\hat{\mathbf{x}}} \left\{ \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 + \lambda_{1/2} \|\mathbf{x}\|_{1/2}^{1/2} \right\} \tag{3}$$

where $\|\mathbf{x}\|_{1/2}^{1/2} = \sum_{i=1}^{N} |x_i|^{1/2}$ denotes $L_{1/2}$ quasi-norm and $\lambda_{1/2}$ stands for the regularization parameter.

According to the fact that the sparsity of a certain signal $\mathbf{x}$ will vary across dictionaries, multiple dictionaries sparse representation strategy combining with the iterative reweighted method have been proposed to solve the $L_1$ minimization problem for reconstructing the sparse images and can obtain a significant improvement [11]. In order to improve the existing $L_{1/2}$ regularization algorithm and obtain the sparser

solution, in this study, we propose a new MUSAI-$L_{1/2}$ algorithm by designing multiple sub-wavelet-dictionaries combining with an adaptive update parameter. In our proposed framework, the signal structure will be different across sub-dictionaries; then, we can appropriately weight the contribution from every dictionary by adjusting suitable regularization parameters $\lambda_d$ for each $\|\mathbf{\Psi}_d \mathrm{x}\|_{1/2}^{1/2}$, this difference varies sub-dictionaries that can be exploited as a prior knowledge for improving recovery.

## 2 Multiple Sub-wavelet-Dictionaries-Based $L_{1/2}$ Regularizer

According to the sparse representation theory, if the raw signal $\mathbf{x}$ is not sparse, then we can conduct a sparsifying transform, such as the wavelet transform (WT) and discrete cosine transform (DCT), which can be expressed as $\mathbf{x}_s = \mathbf{\Psi x}$, where $\mathbf{\Psi} \in R^{N_1 \times N_2}$ is a given analysis dictionary, $\mathbf{x}_s$ presents the sparse coefficients. Hence, the general analysis $L_{1/2}$ regularizer can be described as $R_G = \|\mathbf{x}_s\|_{1/2}^{1/2} = \|\mathbf{\Psi x}\|_{1/2}^{1/2}$.

It is the fact that the signal structure (e.g., sparsity) of a given signal $\mathbf{x}$ will vary with different dictionaries; if this difference can be employed as prior knowledge for recovery improvement, the key issue is that how to exploit and utilized theses prior knowledge. To exploit more prior knowledge of the signal structure for sparse recovery improvement, we utilize the multiple sub-wavelet-dictionaries strategy by the following steps. We first construct the $(nN) \times N$ sparse transform operator (or dictionary) by

$$\mathbf{\Psi} = [\boldsymbol{\psi}_1 | \boldsymbol{\psi}_2 | \ldots | \boldsymbol{\psi}_{nN}] \in R^{(nN) \times N} \tag{4}$$

in which $\boldsymbol{\psi}_i = [\psi_{i1}, \psi_{i2}, \ldots, \psi_{iN}] \in R^{1 \times N}, \quad i = 1, 2, \ldots, nN$ denotes the $i$th row of $\mathbf{\Psi}$, and the operator of '|' denotes 'next row.' Then, the sub-dictionaries $\mathbf{\Psi}_d, d = 1, 2, \ldots, n$ can be obtained by the form

$$\mathbf{\Psi_d} = \left[\boldsymbol{\psi}_{(d-1)N+1} | \boldsymbol{\psi}_{(d-1)N+2} | \ldots | \boldsymbol{\psi}_{dN}\right] \in R^{N \times N} \tag{5}$$

According to (4)–(5), we have the relationship of $\mathbf{\Psi}$ and each sub-dictionaries $\mathbf{\Psi}_d$

$$\mathbf{\Psi} = [\mathbf{\Psi}_1 | \ldots | \mathbf{\Psi}_d | \ldots | \mathbf{\Psi}_n] \in R^{(nN) \times N} \tag{6}$$

Then, these prespecifed sub-dictionaries can be employed for following non-convex $L_{1/2}$ regularizers of $\|\mathbf{\Psi}_1 \mathbf{x}\|_{1/2}^{1/2}, \|\mathbf{\Psi}_2 \mathbf{x}\|_{1/2}^{1/2}, \ldots, \|\mathbf{\Psi}_d \mathbf{x}\|_{1/2}^{1/2}$; then, we design a non-convex regularizer model $R_{\mathrm{MUSAI}}$ by linear combination of these non-convex $L_{1/2}$ regularizers, which can be expressed as

$$R_{\mathrm{MUSAI}} = \sum_{d=1}^{D} \lambda_{d,1/2} \|\mathbf{\Psi}_d \mathbf{x}\|_{1/2}^{1/2} \tag{7}$$

where $\lambda_{d,1/2}, d = 1, 2, \ldots, D$ is the weighting parameter (or regularization parameter), which makes a trade-off between the fidelity term and regularization terms. Moreover, we can appropriately weight the contribution of each sub-dictionary by adjusting suitable regularization parameters $\lambda_d$ for each $\|\boldsymbol{\Psi}_d \mathrm{x}\|_{1/2}^{1/2}$. Hence, these differences vary sub-dictionaries which can be exploited as a prior knowledge for improving recovery.

## 3 Proposed Multiple Sub-wavelet-Dictionaries-Based Adaptive Iteratively Weighted $L_{1/2}$ Regularization Algorithm

Considering the following optimization problem and setting $\gamma = 1$, the proposed $L_{1/2}$ multiple sub-wavelet-dictionaries-based regularization problem is formulated as

$$\hat{\mathbf{x}}_{l_{d,1/2}} = \arg\min \left\{ \boldsymbol{F}(\mathbf{x}) = \|\mathbf{y} - \boldsymbol{\Phi}\mathbf{x}\|_2^2 + \sum_{d=1}^{D} \lambda_{d,\frac{1}{2}} \|\boldsymbol{\Psi}_d \mathbf{x}\|_{1/2}^{1/2} \right\} \tag{8}$$

where the first term of the objective function $\boldsymbol{F}(\mathbf{x})$ is a $L_2$-norm, and the second term is the composite of several non-convex, non-smooth, and non-Lipschitz $L_{1/2}$ regularizers. Commonly, it is hard because there are $D$ regularization terms; benefiting from the designed sub-wavelet-dictionaries, this problem is easily be resolved [11]. Here are the main steps to solve the optimization problem in (8).

### 3.1 Solving the Analysis Problem

We first consider the analysis optimization problem by replacing the regularization term with $\boldsymbol{\Psi}' \in R^{N \times N}$; then, the analysis problem can be described as

$$\hat{\mathbf{x}}_{l_{1/2}} = \arg\min \left\{ \boldsymbol{F}(\mathbf{x}) = \|\mathbf{y} - \boldsymbol{\Phi}\mathbf{x}\|_2^2 + \lambda_{1/2} \|\boldsymbol{\Psi}'\mathbf{x}\|_{1/2}^{1/2} \right\} \tag{9}$$

According to the half thresholding algorithm theory [10], firstly, the first-order optimality condition of $\mathbf{x}$ is obtained as

$$\nabla f(\mathbf{x}) = 2\boldsymbol{\Psi}'\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Phi}\mathbf{x} - \mathbf{y}) + \lambda_{1/2} \nabla(\|\boldsymbol{\Psi}'\mathbf{x}\|_{1/2}^{1/2}) \tag{10}$$

where the operator $\nabla(\cdot)$ denotes the gradient of the objective function. Let $\nabla f(\boldsymbol{x}) = 0$, then we have

$$\boldsymbol{\Psi}'\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Phi}\mathbf{x} - \mathbf{y}) = \frac{1}{2} \nabla \lambda_{1/2} \|\boldsymbol{\Psi}'\mathbf{x}\|_{1/2}^{1/2} \tag{11}$$

Then, multiplying by any parameter $\tau$ and adding $\boldsymbol{\Psi}\mathbf{x}$ in both sides of (11), then we have

$$\boldsymbol{\Psi}'\mathbf{x}+\tau\boldsymbol{\Psi}'\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{\Phi}\mathbf{x}-\mathbf{y})=\boldsymbol{\Psi}'\mathbf{x}+\frac{\tau}{2}\nabla\left(\lambda_{1/2}\|\boldsymbol{\Psi}'\mathbf{x}\|_{1/2}^{1/2}\right) \tag{12}$$

To this end, there is the resolvent operator [10] that exists

$$H_{\lambda_{d,\frac{1}{2}}}(\cdot)=\left(\mathbf{I}+\frac{\lambda_{1/2}}{2}\nabla(\|\cdot\|_{1/2}^{1/2}\right)^{-1}. \tag{13}$$

Then, we imply

$$\begin{aligned}\boldsymbol{\Psi}'\mathbf{x}^{n+1}&=\left(\boldsymbol{I}+\frac{\lambda_{1/2}}{2}\nabla(\|\cdot\|_{\frac{1}{2}}^{\frac{1}{2}}\right)^{-1}\left(\boldsymbol{\Psi}'\left(\mathbf{x}^{n}+\tau\boldsymbol{\Phi}^{T}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x}^{n})\right)\right)\\&=H_{\lambda_{d,1/2}}\left(\boldsymbol{\Psi}'\left(\mathbf{x}^{n}+\tau\boldsymbol{\Phi}^{T}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x}^{n})\right)\right)\end{aligned} \tag{14}$$

Then, the desired image can be reconstructed by

$$\mathbf{x}^{n+1}=(\boldsymbol{\Psi}')^{-1}H_{\lambda_{d,1/2}}\left(\boldsymbol{\Psi}'\left(\mathbf{x}^{n}+\tau\boldsymbol{\Phi}^{T}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x}^{n})\right)\right) \tag{15}$$

where the $\tau>0$ is the step size (e.g., $\tau=\frac{0.99}{\|\Phi\|_2}$) to control the step size in each iteration. The component-wise thresholding operator $\mathrm{H}_{\lambda_{d,1/2}}(\cdot)$ is defined as [10]

$$\mathrm{H}_{\lambda_{1/2}}(\mathbf{x})=\left(h_{\lambda_{1/2}}(x_1),h_{\lambda_{1/2}}(x_2),\ldots,h_{\lambda_{1/2}}(x_N)\right)^{\mathrm{T}} \tag{16}$$

where the $h_{\lambda_{1/2}}(x_i)$ is nonlinear and defined by:

$$h_{\lambda_{1/2}}(x_i)=\begin{cases}\varphi_{\lambda_{d,1/2}}(x_i), & |x_i|>T\\ 0, & \text{otherwise}\end{cases} \tag{17}$$

where the $T=\frac{3\sqrt[3]{2}}{4}\left(\lambda_{1/2}\cdot\tau\right)^{2/3}$ is the threshold value, $\lambda_{1/2}$ is the parameter mentioned above, and $\varphi_{\lambda_{1/2}}(x_i)=\frac{2}{3}x_i\left(1+\cos\left(\frac{2\pi}{3}-\frac{2}{3}\cos^{-1}\left(\frac{\lambda_{1/2}\tau}{8}\left(\frac{|x_i|}{3}\right)^{-\frac{2}{3}}\right)\right)\right)$.

### 3.2 Updating the Adaptive Iteratively Weighting Parameter $\lambda_{d,1/2}$

Considering the optimization problem of (8), the multiple sub-wavelet-dictionaries-based regularization term consists of $D$ regularizers by multiplying with $D$ parameters $\lambda_{d,1/2}, d=1,2,\ldots,D$. These parameters not only make the trade-off between the prior knowledge and the error, but also control the contribution of each sub-dictionary. In order to achieve this goal, this paper proposes an adaptively updating parameter scheme by

$$\lambda_{d,1/2}^{(t+1)} = \frac{L_d}{\epsilon + \|\mathbf{\Psi}_d \mathbf{x}^{(t)}\|_{1/2}^{1/2}} \tag{18}$$

where $L_d$ denote the number of the rows of each sub-dictionary. So we can use the parameter $\lambda_{d,1/2}$ to counteract the influence of the signal sparsity on the $L_{1/2}$ regularization by this formulation.

Hence, for a series of parameters $\lambda_{d,1/2}^{(t+1)}$ given by (18), the problem (8) can be converted into (9) and can be resolved using the above analysis iterative half thresholding algorithm easily, because these $D$ regularizers are finally equivalent to the popular regularized model (similar work see [11]).

## 4 Experiments and Discussions

In this section, we demonstrate the efficacy of the proposed MUSAI-$L_{1/2}$ algorithm in compressive imaging in vehicular systems and compare the performance with prior work, including the Co-L1 [11] and the well-known half thresholding algorithm for $L_{1/2}$ [10]. We select the well-known 'cameraman' figure as typical image in vehicular system as shown in Fig. 1a, in order to reduce the computation; we choose a partial image for the experiment shown in Fig. 1b. We obtain the measurement where $\mathbf{\Phi} \in R^{M \times N}$ is the measurement matrix, which is constructed by the spread spectrum, and $\mathbf{n}$ denotes the white Gaussian noise; in this paper, we define the measurement SNR (mSNR) as $\text{mSNR} = \frac{\|\mathbf{y}\|_2^2}{M\sigma^2}$ to represent the noise level, where $M$ present the number of measurement $\mathbf{y}$ and $\sigma^2$ is the variance of the white Gaussian noise. All the experiments were conducted on a personal computer (2.21 GHz, 16 Gb RAM) in MATLAB (R2014a) platform.

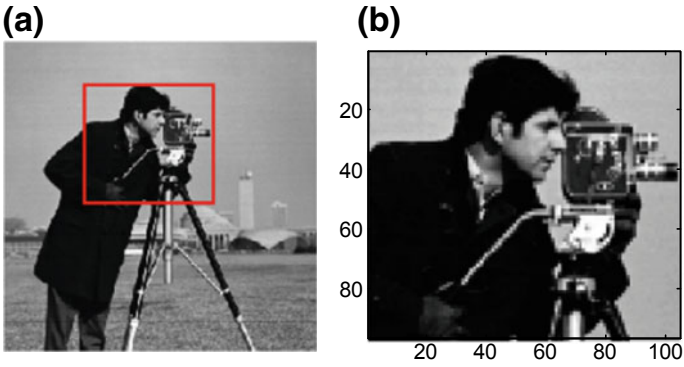



**Fig. 1.** Clear cameraman image. Left: the cameraman image of size $N = 256 \times 256$; right: the selected part image of size $N = 96 \times 104$

We first assess the convergence by the relative error versus the number of the iteration times under low sampling ratio of 0.2 and a higher sampling ratio of 0.4 with mSNR $= 40$ dB, and the number of wavelet decomposition level $N = 2$. The relative error is defined as $\|\mathbf{x} - \hat{\mathbf{x}}\|_2 / \|\mathbf{x}\|_2$, where $\hat{\mathbf{x}}$ and $\mathbf{x}$ denote the recovery signal and original signal, respectively. Figure 2a, b presents the relative error curves of the three algorithms when the sampling ratio is 0.2 and 0.4, respectively. Both results show that our proposed MUSAI-$L_{1/2}$ can obtain the less relative error than Co-L1 in both low sampling ratio and higher sampling ratio scenarios. Moreover, our proposed MUSAI-$L_{1/2}$ can obtain not only faster convergences but also less errors than the corresponding $L_{1/2}$, especially when sampling ratio is very low, e.g., 0.2.

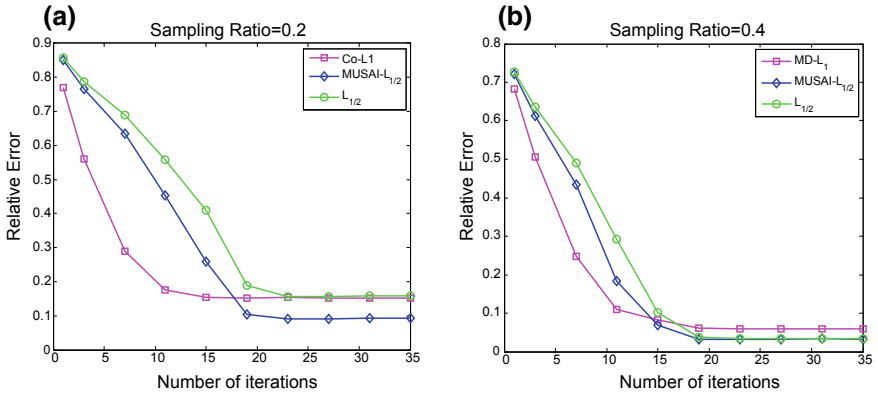



**Fig. 2.** Relative error versus the number of iteration by Co-L1, MUSAI-$L_{1/2}$ and $L_{1/2}$. **a** Sampling ratio = 0.2; **b** sampling ratio = 0.4

Then, we consider the influence of different levels of noise and the sampling ratio on reconstruction performance. We evaluate the performance by the RSNR versus different mSNR values of 10, 12.5, 15, 20, 25, 30, 35, and 40 dB. The mathematical definitions of RSNR are $-20\log\left(\frac{\|\mathbf{x}-\hat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}\right)$, where $\hat{\mathbf{x}}$ and $\mathbf{x}$ denote the reconstruct image and the desired image, respectively. The result is represented in Fig. 3a. From the result, we can find that the proposed MUSAI-$L_{1/2}$ can obtain a higher RSNR, and with the increment of the signal mSNR, the advantage of the proposed algorithm becomes more and more significant. To evaluate the performance versus sampling ratio, we set mSNR $= 40$ dB, and the decomposition level $N$ of wavelet transform is 2. Figure 3b shows the recovery SNR versus sampling ratio for the noisy image which tells us that the proposed MUSAI-$L_{1/2}$ can improve the recovery significantly than the Co-L1 algorithm in [11], especially at 0.2–0.5 of the sampling ratio, and it also outperforms the single dictionary $L_{1/2}$ algorithm. Both figures show that the proposed MUSAI-$L_{1/2}$ algorithm outperforms the other two. Finally, we present the reconstruction effect of our proposed algorithm and other two algorithms under a low sampling ratio of 0.2 in Fig. 4.

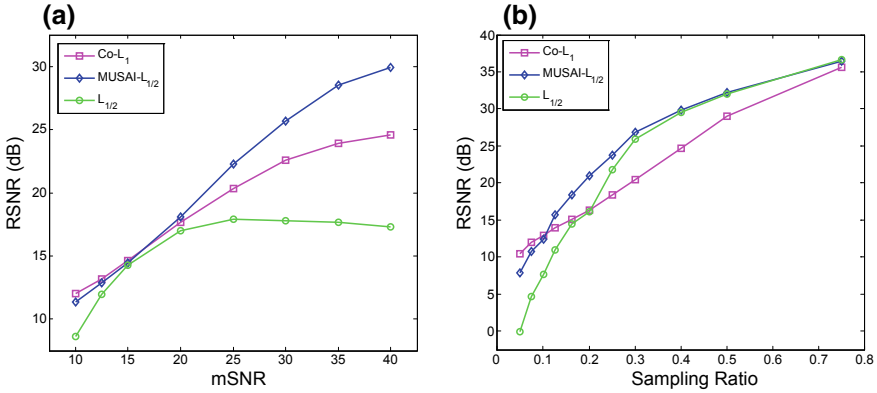



**Fig. 3.** **a** Recovery SNR versus the measurement SNR (mSNR) level of 10, 12.5. 15, 20, 25, 30, 35, 40 (dB), with the sampling ratio 0.4; **b** RSNR of three algorithms under different sampling ratio for the 'cameraman' image

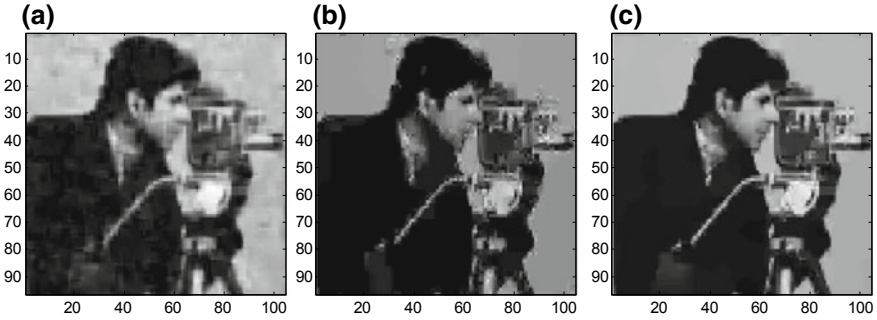



**Fig. 4.** Reconstruction effect by conducting the Co-L1, $L_{1/2}$, MUSAI-$L_{1/2}$ algorithms with sampling ratio 0.20. The RSNRs are: **a** 16.3405 dB; **b** 16.1024 dB; **c** 20.9655 dB

## 5 Conclusions

In this paper, we proposed an adaptive iteratively weighted $L_{1/2}$ regularization algorithm for image compressive sensing reconstruction based on the prespecified multiple sub-wavelet-dictionaries strategy (MUSAI-$L_{1/2}$), and we also proposed an adaptive parameter $\lambda_{d,1/2}^{(t+1)} = L_d / \left(\epsilon + \left\|\boldsymbol{\Psi}_d \mathbf{x}^{(t)}\right\|_{1/2}^{1/2}\right)$ to make a trade-off between the prior knowledge of sparse signal and the data error and, moreover, to iteratively weight the contribution of each dictionaries. The performance results demonstrate that our prespecified multiple sub-dictionaries can effectively exploit the structure of desired signals for prior knowledge, and the proposed MUSAI-$L_{1/2}$ algorithm may perform better than the previous work of $\sum_{d=1}^{D} \lambda_{d,1} \|\boldsymbol{\Psi}_d \boldsymbol{x}\|_1^1$ and $\lambda \|\boldsymbol{\Psi} \boldsymbol{x}\|_{1/2}^{1/2}$ penalized regularization approaches.

**Acknowledgments.** This research was funded by State Grid Corporation Science and Technology Project (named 'Research on intelligent patrol and inspection technology of substation robot based on intelligent sensor active collaboration technology').

## References

1. Guo J, Song B, He Y, Yu FR, Sookhak M. A survey on compressed sensing in vehicular infotainment systems. IEEE Commun Surv Tutorials. 2017;19(4):2662–80.
2. Luo Y, Wan Q, Gui G, Adachi F. A matching pursuit generalized approximate message passing algorithm. IEICE Trans Fundam Electron Commun Comput Sci. 2015;E98–A(12): 2723–7.
3. Gao Z, Dai L, Han S, Chih-Lin I, Wang Z, Hanzo L. Compressive sensing techniques for next-generation wireless communications. IEEE Wirel Commun. 2018; 2–11.
4. Shi Z, Zhou C, Gu Y, Goodman NA, Qu F. Source estimation using coprime array: a sparse reconstruction perspective. IEEE Sens J. 2017;17(3):755–65.
5. Hayashi K, Nagahara M, Tanaka T. A user's guide to compressed sensing for communications systems. IEICE Trans Commun. 2013;E96–B(3):685–712.
6. Donoho DL. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. Discret Comput Geom. 2006;35(January):617–52.
7. Yilmaz O. Stable sparse approximations via nonconvex optimization. In IEEE international conference on acoustics, speech and signal processing, 2007; 2008. p. 3885–8.
8. Chartrand R. Nonconvex compressed sensing and error correction. In: IEEE international conference on acoustics, speech and signal processing 2007; 2007, no. 3. p. 889–92.
9. Chartrand R. Exact reconstruction of sparse signals via nonconvex minimization. IEEE Sig Process Lett. 2007;14(10):707–10.
10. Xu Z, Chang X, Xu F, Zhang H. $L_{1/2}$ regularization: a thresholding representation theory and a fast solver. IEEE Trans Neural Netw Learn Syst. 2012;23(7):1013–27.
11. Ahmad R, Schniter P. Iteratively reweighted $L_1$ approaches to sparse composite regularization. IEEE Trans Comput Imaging. 2015;1(4):220–35.

# Study on the Influence of Image Motion on Image Radiation Quality

Fan Yang[1(✉)], Zhaocong Wu[2], Jisheng Zeng[1], and Zheng Wei[1]

[1] South China Sea Institute of Planning and Environmental Research, SOA, Guangzhou, China
yangfan_whu@sina.cn
[2] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

**Abstract.** This paper conducted depth and systematic research on the influence of image motion on radiation quality. The exploration of the regular image radiation quality could be a great significance for improving the image quality. The source of abnormal image motion was analyzed in terms of platform motion and flutter, attitude control error, and image motion compensation error. The model of point spread function under abnormal image motion was established based on the analysis of superposition model between adjacent pixels and PSF variation model. The radiation information and MTF caused by PSF changes were analyzed. The experiment shows that the authenticity of the radiation information expressing the imaging target is weakened. The larger the abnormal image motion, the faster the MTF curve declines, indicates the overall image quality decreases, the spatial resolution decreases, and the sharpness decreases.

**Keywords:** Abnormal image motion · Point spread function · Radiation quality

## 1 Introduction

The technology of high-resolution remote sensing in China has been continuously improved and matured to meet the growing demand for image data for users in various industries. However, there is still a certain gap between China and foreign advanced high-resolution remote sensing commercial satellite imagery in information extraction and application. The key issue in the further development of remote sensing and the application of image data in China is image quality [1].

During imaging of remote sensing satellite, the radiation quality of images will inevitably be affected by the imaging conditions and environment [2], including solar lighting, atmospheric conditions, imaging targets, satellite platforms, optical cameras, electronic circuits, and data processing.

The article starts from the source of abnormal image motion and studies its influence on the radiation quality of images. The paper establishes the relationship between the image motion and imaging radiation quality. It analyzes the principle of the superposition and change of imaging point spread function (PSF) under the condition of abnormal image motion, establishes the model of PSF in the case of image motion,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_23

and analyzes the influence principle and regular pattern from the aspects of the authenticity of radiation information and MTF.

By analyzing the effect of image motion on imaging radiation quality and its influence regular pattern, it helps to analyze the key influencing factors from the source. It can effectively control the quality of remote sensing imaging from the perspective of the design of remote sensing satellite systems and improve the image quality of high-resolution remote sensing satellites in China.

## 2 The Source of Abnormal Image Motion in Imaging

In the imaging process of remote sensing sensors, many factors cause the abnormal image point motion [3], including the motion and chatter of platform, attitude stability, pointing accuracy and other control errors, speed-to-height ratio errors, image motion compensation, and time synchronization accuracy errors, etc. The position of the image point deviating from the ideal decreases geometric quality, such as image geometric distortion, spatial misalignment, and reduced spatial resolution, and also decreases the radiation quality such as image blur, changes in radiation information and system MTF.

The sources of abnormal image motion mainly include:

(1) Platform motion and flutter. When the remote sensing sensor is in orbit, the satellite platform will be affected by factors such as unavoidable external space environment interference and platform itself works. Vibration will be transmitted to the sensor through the platform, affecting the attitude stability of the sensor, resulting in abnormal image motion.

(2) Attitude control error. The satellite platform will cause real-time changes in the roll angle, pitch angle, and yaw angle during movement. The change of the attitude angle causes an angle error between the imaging light and ideal conditions. The image produces horizontal or vertical length distortion or angular rotation, and the position of the image point in the image deviates from the ideal position [4].

Through the analysis of attitude and attitude angular velocity error, imaging distance error, and other factors affecting the magnitude of image motion, the yaw angle error in attitude error is the most important influence factor of image motion.

(3) Image motion compensation error. Satellite orbital motion, earth rotation, and attitude angle change cause the relative motion of the image point on the focal plane. The onboard image motion compensation device compensates the relative movement of the image point according to the image shift speed. Due to the variation of the attitude angle and the attitude angular velocity caused by the platform factor, the deviation of the image-moving speed causes the deviation of the drift angle compensation and the integration time, resulting in abnormal image point motion.

## 3 The Model of Point Spread Function Under Image Motion

### 3.1 The PSF of Imaging System

PSF describes the spread of a point light source after passing through an optical imaging system. In the imaging process, the input signal of each pixel in the TDICCD camera is the electromagnetic wave information corresponding to the ground imaging target of the pixel. The center position of the pixel is determined by the imaging geometric model, and the pixel signal is integrated by PSF during the integration time [5].

R. Richter proposed that the PSF can be represented by a two-dimensional Gaussian function [6]. As shown in Eq. 1, where $\sigma$ is the Gaussian width of the PSF, $(x, y)$ is the image coordinate value of a pixel, and $d$ is the distance between the pixel and the target pixel, the integral of the PSF function is 1.

$$\mathrm{PSF}(x, y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{d^2}{2\sigma^2}\right) \tag{1}$$

The paper investigates the phenomenon of point spread of pixels and the effect of image motion on PSF based on this model. Each pixel produces a two-dimensional Gaussian distribution centered on the pixel in the image. Due to the point spread phenomenon of the pixel, the electromagnetic wave signal collected by each pixel is an intensity distribution that is a PSF model, weakening toward the surroundings. The entire image is formed by the superposition of Gaussian distributions generated by each pixel on all CCD arrays. The two-dimensional and three-dimensional Gaussian distribution are shown in Fig. 1.

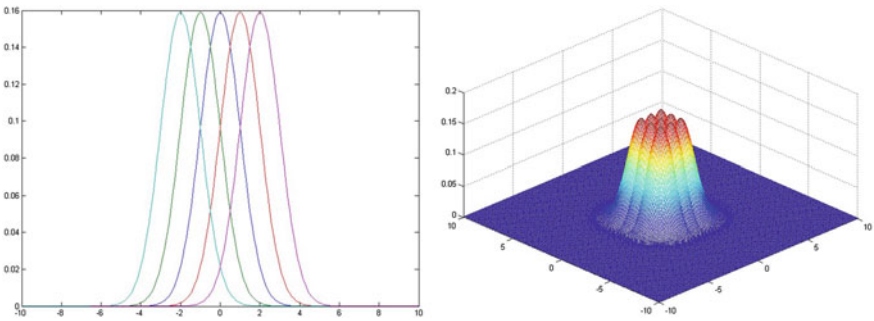


**Fig. 1.** Gaussian distribution of PSF in two-dimension and three-dimension

### 3.2 The Overlay of PSF in Adjacent Pixel

Due to the existence of point spread phenomenon, the pixel will influence the neighboring pixels while being affected by the PSF of neighboring pixel points. The PSF is superimposed to form a point spread effect surface [7]. According to the characteristics of the Gaussian function, the Gaussian distribution of a point drops below 1% of the

peak from the center of a Gaussian width, and the effect of more than two Gaussian-width distances is negligible. The phenomenon of point spread of a single pixel mainly affects the range of the distance between two pixels around it.

The information of each pixel in the image is the integral of the point spread effect surface in this pixel area. After being discretized, an image is formed. The paper builds an overlay model of the neighboring pixels in the direction of the point spread along TDI and vertical TDI. The model is expressed as Eq. 2.

$$\begin{aligned} \mathrm{PSF}(x) = & \int_{-d/2}^{d/2} \mathrm{PSF}(0)\mathrm{d}x \\ & + \int_{-3d/2}^{-d/2} \mathrm{PSF}(-d)\mathrm{d}x + \int_{-2d}^{-3d/2} \mathrm{PSF}(-2d)\mathrm{d}x \\ & + \int_{d/2}^{3d/2} \mathrm{PSF}(d)\mathrm{d}x + \int_{3d/2}^{2d} \mathrm{PSF}(2d)\mathrm{d}x \end{aligned} \tag{2}$$

Due to the existence of the pixel PSF effect, the plane where the original luminance alternates become flattened, the gray value of the pixel is affected by the neighboring pixels. The dark pixels are brightened by the diffusion phenomenon of bright pixels in the surrounding area. The bright pixels are darkened by the diffusion of surrounding dark pixels. In the image, the edges are blurred, and the contrast is reduced.

### 3.3 The Analysis of PSF Changes Under Abnormal Image Motion

In the presence of image motion in the TDICCD, the imaging position of the ground-based target changes on the detector. The center position of the signal changes with respect to the center of the pixel. The image motion causes the image pixel size changes, and the Gaussian width becomes larger. The effect of the point spread phenomenon on the surrounding pixels becomes larger [8]. Since the energy of the electromagnetic radiation signal received by a single pixel is fixed, the peak energy of the PSF decreases and the light energy distribution of the camera changes.

Under the influence of the image shift, the PSF of the pixel can be expressed as Eq. 3. In the formula, $\Delta x$ and $\Delta y$, respectively, represent the image point motion of the corresponding pixel of the imaging target along and vertical TDICCD direction.

$$\mathrm{PSF}(x+\Delta x, y+\Delta y) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(x+\Delta x)^2+(y+\Delta y)^2}{2\sigma^2}\right] \tag{3}$$

The PSF changes in the case of image motion are shown in Fig. 2 in the two-dimensional view and the three-dimensional view. In the presence of image motion, the

central position of the point spread changes with the displacement of the image point, and the peak signal of the point spread weakens, but the diffusion range increases [9].

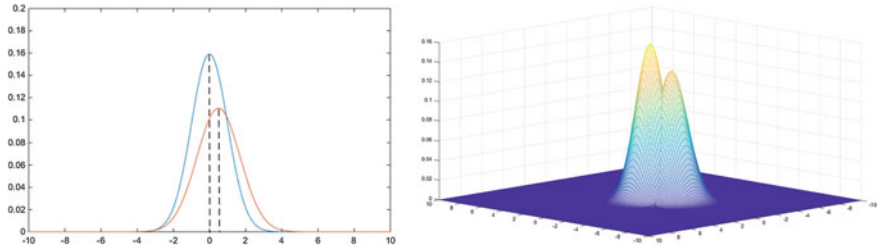


**Fig. 2.** Gaussian distribution of PSF changes in the case of image motion

The effects caused by the PSF change under the abnormal image motion include:

(1) The weakening of the pixel center signal causes a reduction in the energy of the pixel information;
(2) The increase in the spread range of the pixel causes an increase in the extent of the surrounding pixels. The influence distance is related to the size of the image motion;
(3) The amplitude of the pixel spread becomes smaller, and the influence on the surrounding pixels increases;
(4) The variation of the diffusion center of pixel point causes signal confusion between pixels.

### 3.4 The Model of PSF Under Abnormal Image Motion

Through the above-mentioned PSF change in the case of image motion and the superposition principle of adjacent pixels PSF, the paper establishes the PSF model under the condition of abnormal image motion. The model for a single pixel PSF is modeling as follows:

(1) A single pixel is affected by pixels whose peripheral distance is in the range of $[\Delta x \in (-2d, 2d), \Delta y \in (-2d, 2d)]$, and the influence of the point spread is related to the distance. The point spread function in pixels can be expressed as Eq. 4. In the equation, $m$, $n$ is used to identify the pixels around the center pixel that are within the influence of the $[-2d, 2d]$ point spread.

$$\mathrm{PSF}_{IM}(x, y) = \mathrm{PSF}_{(0,0)}(x, y) \otimes \sum_{m=-2d}^{2d} \sum_{n=-2d}^{2d} \mathrm{PSF}_{(m,n)}(m - x, n - y) \tag{4}$$

(2) Only when the center pixel has an image motion, the PSF influence varies with the direction and the amount of the image motion, and the point spread of the surrounding pixels to the center pixel also varies with the image motion. The PSF for the center pixel under image motion can be expressed as Eq. 5.

$$\begin{aligned} \mathrm{PSF}_{IM}(x,y) &= PSF_{(0,0)}(x+\Delta x_0, y+\Delta y_0) \\ &\otimes \sum_{m=-2d}^{2d}\sum_{n=-2d}^{2d} \mathrm{PSF}_{(m,n)}(m-x+\Delta x_0, n-y+\Delta y_0) \end{aligned} \tag{5}$$

(3) In the case of image motion in the pixels within the surrounding influence distance, the PSF model of the center pixel can be expressed as Eq. 6. This model takes into account both the PSF variation in the case of a single pixel image motion and the change in the superposition effect of the surrounding pixel PSF.

$$\begin{aligned} \mathrm{PSF}_{IM}(x,y) &= \mathrm{PSF}_{(0,0)}(x+\Delta x_0, y+\Delta y_0) \\ &\otimes \sum_{m=-2d}^{2d}\sum_{n=-2d}^{2d} \mathrm{PSF}_{(m,n)}(m-x+\Delta x_m, n-y+\Delta y_n) \end{aligned} \tag{6}$$

Based on the PSF change model, this paper studies the effect of image motion on image radiation quality.

## 4 The Influence on Radiation Quality of Abnormal Image Motion

Based on the PSF model under image motion, the paper investigates the variation of radiation information in the pixel caused by PSF changes and analyzes the effect of image motion on the authenticity of image radiation information and MTF.

### 4.1 The Influence on the Validation of Radiation Information

In an ideal situation, the gradation information of the pixels in the image is determined by the PSF of pixel and the surface radiation information received by the camera. The radiation information $g(x, y)$ of each pixel can be expressed as Eq. 7. In the equation, $f(x, y)$ is the surface radiation information received by the pixel, and $n(x, y)$ is the noise of the imaging system.

$$g(x,y) = \mathrm{PSF}(x,y) \otimes f(x,y) + n(x,y) \tag{7}$$

In the image, the radiation information of each pixel is the integration and quantization of the radiation information within the pixel [10], and the model can be expressed as Eq. 8.

$$G(x,y) = \int_{-0.5}^{0.5}\int_{-0.5}^{0.5} g(x,y)\mathrm{d}x\mathrm{d}y \tag{8}$$

The gray value $\mathrm{DN}_{(0,0)}$ of the center cell in the case of image motion is expressed as Eq. 9.

$$\begin{aligned}\mathrm{DN}_{(0,0)} = &\int_{-0.5}^{0.5}\int_{-0.5}^{0.5} \mathrm{PSF}_{IM(0,0)}(x,y) \otimes f_{(0,0)}(x,y)\mathrm{d}x\mathrm{d}y \\ &+ \sum_{p=-2d}^{2d}\sum_{q=-2d}^{2d}\int_{-0.5}^{0.5}\int_{-0.5}^{0.5} \mathrm{PSF}_{(p,q)}(p-x,q-y) \otimes f_{(p,q)}(p-x,q-y)\mathrm{d}x\mathrm{d}y\end{aligned} \tag{9}$$

Similarly, the gray values of other neighboring pixels can be derived based on the mutual influence between the PSFs of the pixels.

In the presence of abnormal image motion, the change of radiation information caused by image motion is reflected in the authenticity of image radiation information. In addition to the surface radiation information, the radiation information of the pixel is also affected by the change of the PSF. The extent and range of diffusion of neighboring pixels in the center pixel increase. The radiation information is confused with surrounding pixel radiation and radiation information of the pixel changes [11].

Therefore, the radiation information of the pixel increases due to the influence of the surrounding radiation information, and the authenticity of the radiation information expressing the imaging target is weakened [12, 13].

### 4.2 The Influence on the MTF

The change of the PSF under the image motion results in the change of the image MTF. The MTF is the two-dimensional Fourier transform of the PSF [14]. The model is expressed as Eq. 10.

$$\mathrm{MTF}(u,v) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{PSF}_{IM}(x,y)\mathrm{e}^{-j2\pi(ux+vy)}\mathrm{d}x\mathrm{d}y \tag{10}$$

The yaw angle error is the most important influence factor of image motion. Therefore, the influence of abnormal image motion on the MTF is studied, mainly through the influence of the yaw angle attitude error on the PSF along and vertical TDICCD direction [4].

The influence regular of yaw angle error on MTF is shown in Fig. 3. The thesis compares the MTF curve under the condition of ideal attitude and MTF curve and the condition of different yaw angle error. As the yaw angle error increases, the larger

the influence range of the PSF effect, the faster the MTF curve declines. The experimental results show that under the influence of yaw angle error, the overall image quality decreases, the spatial resolution decreases, and the sharpness decreases.

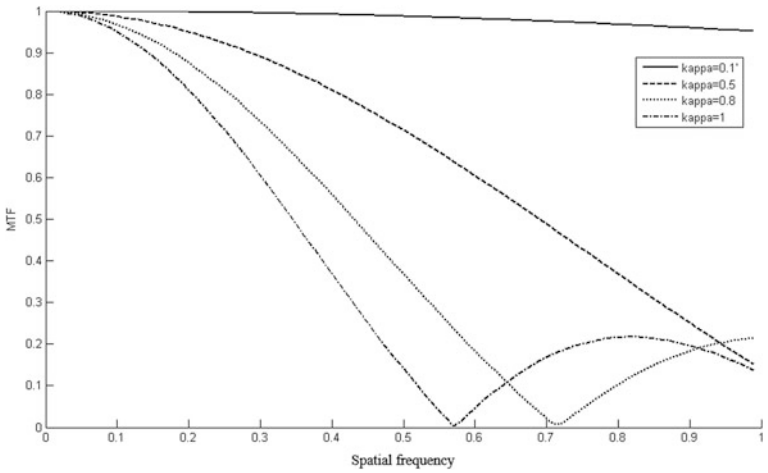



**Fig. 3.** Influence of different yaw angle errors on MTF

## 5 Summary

The paper analyzed the main source of abnormal image motion and established the PSF model under abnormal image motion, then studied the effect of image motion on the authenticity of image radiation information and MTF. The results showed that under the abnormal image motion, the authenticity of the radiation information expressing the imaging target is weakened, and the MTF curve declines.

**Acknowledgments.** This work was supported in part by the Key Laboratory of Technology for Safeguarding of Marine Rights and Interests and Application, SOA, and supported by the Director Foundation of Marine Science and Technology in South China Sea Branch, SOA, NO. 1713. Heartfelt thanks are also given for the comments and contributions of reviewers and members of the editorial team.

## References

1. Shiping C. Some issues about the remote sensing image quality. Spacecraft Recov Remote Sens. 2009;30(2):10–7.
2. Jun H, Dong W, Tianyu S. Application and development of recent space optical imaging remote sensors. Chin J Opt Appl Opt. 2010;3(6):519–33.
3. Wang T, Zhang Y, Zhang Y, et al. Construction and validation of rigorous imaging model for high-resolution remote sensing satellites. J Remote Sens. 2013;17(5):1087–102.

4. Wu Z, Yang F, Wu Y, et al. Impact of satellite platform attitudes' error on imaging geometric quality of TDICCD. J Nat Univ Defense Technol. 2017;39(2):101–6.
5. Huang C, Townshend JRG, Liang S, et al. Impact of sensor's point spread function on land cover characterization: assessment and deconvolution. Remote Sens Environ. 2002;80 (2):203–12.
6. Richter R, Bachmann M, Dorigo W, et al. Influence of the adjacency effect on ground reflectance measurements. Geosci Remote Sens Lett IEEE. 2006;3(4):565–9.
7. Jun W, Jiaqi W, Lu E, et al. Calculation of optical transfer function with two-dimensional image motion. Acta Opt Sin. 2001;21(5):581–5.
8. Feng X, Lelin C, Wei Z. Research on effect of PSF on point target detection performance. Infrared Laser Eng. 2007;36:177–81.
9. Bing X, Binliang J. The PSF estimation method and image restoration algorithm based on image-motion analysis of spaceborne TDICCD Camera. J Astronaut. 2010;31(3):937–40.
10. Lihong Y, Bianhong Z, Xingxiang Z, et al. Gaussian fitted estimation of point spread function and remote sensing image restoration. Chin Opt. 2012;5(2):181–8.
11. Costello TP, Mikhael WB. Restoration of digital images with known space-variant blurs from conventional optical systems. In: AeroSense'99 international society for optics and photonics; 1999:71–9.
12. Ruiz CP, Lopez FJA. Restoring SPOT images using PSF-derived deconvolution filters. Int J Remote Sens. 2002;23(12):2379–91.
13. Hochman G, Yitzhaky Y, Kopeika NS, et al. Restoration of images captured by a staggered time delay and integration camera in the presence of mechanical vibrations. Appl Opt. 2004;43(22):4345–54.
14. Wei J, Shiping C, Qiaolin H, et al. A simulation study of MTF compensation on optical sampled imaging system. Spacecraft Recov Remote Sens. 2007;28(4):28–33.

# Flame Image Segmentation Algorithm Based on Motion and Color Saliency

Yuyan Li and Lin Wang(✉)

Academy of Information Science and Technology, Dalian Maritime University, Dalian 116026, China
wanglin_dmu@163.com

**Abstract.** This paper proposed a flame segmentation algorithm based on the saliency of motion and color. First, feature point detection is performed on the video image using the scale-invariant feature transform (SIFT) algorithm, and the optical flow field of moving object in the adjacent frame is acquired by the optical flow method. According to the optical flow vector difference between the target pixel point and the surrounding neighborhood pixels, the motion saliency map is obtained based on the Munsell color system. Then, the *LSI* flame color statistical model based on the Lab and HSI space is used to extract color saliency map of video images. Finally, under the Bayes framework, the motion saliency map and the color saliency map are fused in an interactive manner to obtain the final flame segmentation map. Experimental results show that the proposed algorithm can effectively segment the flame image in different scenarios.

**Keywords:** Visual saliency · Flame segmentation · Optical flow · *LSI* color statistics model

## 1 Introduction

Image segmentation refers to segmenting the image into meaningful regions based on the features such as grayscale, texture, gradient, and shape of the image, so as to obtain an image representation that is easier to understand and analyze. In recent years, scholars have studied various image segmentation methods [1]. Some of the hybrid methods are closely related to the visual attention mechanism [2], which is mainly based on the saliency or prior knowledge of the image itself to select the visual information process. Image saliency-based image target analysis not only does not require a large amount of training data, but also can quickly focus on significant targets in the image. Therefore, many scholars try to introduce the concept of visual saliency into video image segmentation.

At present, there are many methods to generate saliency maps based on visual saliency. Among them, the classic Itti model [3] uses Gauss Pyramid to generate multi-scale images based on the human visual characteristics. The underlying features such as color, brightness, direction are calculated, and the attention maps at each scale are calculated according to the difference in the central periphery. Finally, the attention maps are fused to get the saliency map. But the model only takes into account the local features of the image when extracting the salient features of the image and does not


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_24

consider the image salient features globally, so the detected areas are relatively rough and contain more background regions. Hou and Zhang [4] proposed a spectral residual method. The method considers that the information in an image can be divided into two parts: redundant information and saliency information. By using the algorithm operation on the amplitude spectrum of the image, the redundant information is obtained by convolution between the average filter and the amplitude spectrum. The saliency information is characterized by the difference between the image amplitude spectrum and its convolution, and the significant region is obtained by inverse transformation to the space domain. The processing speed of this method is faster, and the boundaries of the significant areas detected are relatively clear.

In this paper, saliency theory is introduced into flame image segmentation, and a flame image segmentation algorithm is proposed which combines motion saliency and color saliency. First, we combine the SIFT algorithm with the optical flow method and use the motion vector difference between pixels to get the motion saliency map. Then, the *LSI* color statistical model is obtained by combining the Lab space and the HSI space, and the color saliency map is obtained by using the difference of the *L*, *S*, and *I* components of each pixel. Finally, in the Bayesian framework, the motion saliency map and color saliency map are fused to get the flame segmentation.

## 2 Motion Saliency Feature Extraction

Motion saliency is an inherent characteristic of the flame, but it will be affected by the shape, structure, size, ambient light, and other interference conditions of the combustion. In this paper, based on the SIFT algorithm to obtain more accurate feature points, and the Munsell color system is used to visualize the moving objects, optical flow method is used to extract movement regions. The motion saliency map is obtained through the difference between the pixel points. The motion feature extraction algorithm in this paper can further improve the accuracy of the motion saliency map.

### 2.1 Optical Flow Field Calculation Based on SIFT Operator

The optical flow is defined as the moving speed of the pixel points in the video frame. If the optical flow vector is calculated for each pixel in the image, the resulting result is called dense optical flow, and the amount of calculation is large. If the pixels of certain characteristics are selected first, the optical flow vector is only calculated for these few pixels, and the sparse optical flow is obtained, which can greatly reduce the amount of calculation.

In this paper, the SIFT algorithm is used to detect the $p_{t-1}$ of feature point in the $t-1$ frame, and the Lucas-Kanade optical flow algorithm is used to obtain the corresponding position of the $t$ frame feature point $p_t$. The arrow light flow diagram can show the motion direction of each pixel, and the arrow size indicates the speed of the pixel motion. As can be seen from Fig. 1, compared with Harris corner detection, the feature point detection algorithm in this paper can obtain more stable feature points, which makes the detection of moving objects more accurate.

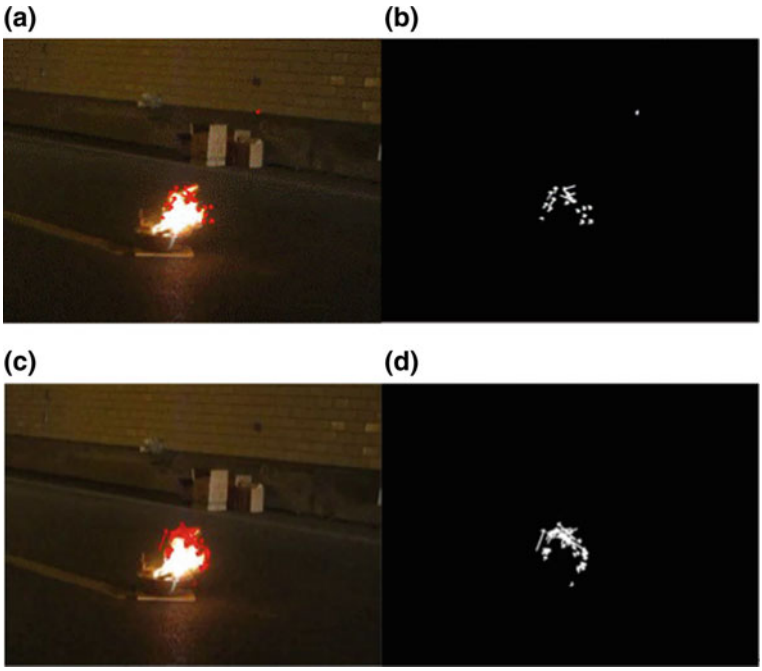



**Fig. 1.** **a** Harris operator optical flow vector map and **b** shows the corresponding optical flow vector binary map, **c** SIFT operator optical flow vector and **d** shows the corresponding optical flow vector binary map

## 2.2 Motion Saliency Feature Extraction

The video frame image uses the SIFT-based optical flow field algorithm to calculate the motion characteristics. For the motion saliency of the target, the motion velocity of each effective tracking point is extracted as a motion feature, and the movement difference of the optical flow vectors of the target pixel point and the surrounding neighborhood pixel points is calculated. The motion saliency map $S_1$ is obtained by Formula (1):

$$S_1(f_i) = \sum_{j=1}^{n} w_{ij} D_m(f_i, f_j) \tag{1}$$

where, $n$ denotes the set of pixel points adjacent to pixel $i$, the weight $w_{ij} = w'_{ij} / \sum_j w'_{ij}$, $w'_{ij} = 1/L_{ij}$ is normalized by the inverse ratio of the Euclidean distance of two pixel positions. $L_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ denotes Euclidean distance between pixels, that is, the closer the pixel distance is, the larger the weight value and the farther the

distance, the smaller the weight. The difference in motion characteristics is defined by Eq. (2) as the Euclidean distance of the motion velocity vector:

$$D_m(f_i,f_j) = \sqrt{(vx_i - vx_j)^2 + (vy_i - vy_j)^2} \tag{2}$$

where, $vx_i$, $vx_j$, $vy_i$ and $vy_j$ are motion velocity vectors of feature points, respectively.

In this paper, we use the Munsell color system to visualize the significance of the obtained motion. The extracted motion flame and moving vehicle significance map are shown in Fig. 2. As shown in Fig. 2, this algorithm can completely extract the moving object in the picture, where different colors indicate different directions of motion, and the depth of color indicates the speed of the movement.

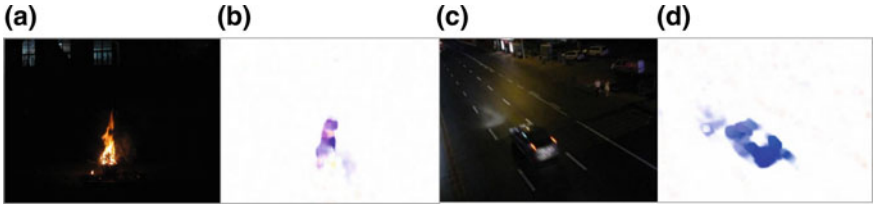



**Fig. 2.** Saliency map of the flame and the moving vehicle. **a** Flame image, **b** flame saliency map, **c** non-flame image, **d** Mobile vehicle saliency map

## 3 Color Saliency Feature Extraction

Color saliency is an important characteristic of flame. Usually, the components of RGB color space are too correlated, which is easy to be affected by illumination and brightness and does not match the human eye recognition process. This paper converts the RGB color space into the Lab and HSI color spaces that are similar to the human visual system and obtains the *LSI* color model. The Gaussian filtering of the image is performed using a DoG filter, and the color saliency map of the image is extracted by the size of the image pixel difference. The color feature extraction algorithm in this paper can further improve the accuracy of the color saliency map.

### 3.1 LSI Color Statistics Model

The lightness value $L$ component in the Lab space closely matches the perception of human vision, which represents the darkest black at $L = 0$, and the brightest white at $L = 0$. In the HSI space, $S$ component denote the degree of saturation, and the depth of the color is described, that is, the extent to which the solid color is diluted with white, and the influence of the ground reflection on the flame combustion can be effectively suppressed. $I$ component denote the brightness, which means that the brightness of light is perceived by the human eye, which corresponds to image gray scale and image brightness and cannot be actually measured.

In this paper, we use the *LSI* color statistical model to extract flame candidate regions. The experimental results are shown in Formula (3), the values of each *L*, *S*, *I* pixels of the flame image are:

$$\begin{cases} 0 \le L \le 100 \\ 0 \le S \le 0.2 \\ 127 \le I \le 255 \end{cases} \tag{3}$$

### 3.2 Color Saliency Feature Extraction

We use Gaussian blur in the *LSI* color model to eliminate subtle texture details and the effects of noise and coding artifacts, the saliency map $S_2$ for an image $I$ can be formulated as:

$$S_2(x, y) = \left\| I_\mu - I_{LSI}(x, y) \right\|_2 \tag{4}$$

where $I_\mu = [L_\mu \quad S_\mu \quad I_\mu]^{\mathrm{T}}$ is the mean feature vector of the $t$-th image, $I_{LSI}(x, y) = [L_{LSI} \quad S_{LSI} \quad I_{LSI}]^{\mathrm{T}}$ is the corresponding image pixel vector value in the Gaussian-blurred version (using a $3 \times 3$ separable binomial kernel) of the $t$-th image, and $\|\|_2$ is the $L_2$ norm.

In Fig. 3, column (a) shows the input images, column (b) shows the saliency maps extracted from the HSI color space, column (c) shows the saliency maps extracted from the Lab color space, and column (d) shows the saliency maps extracted from the *LSI* color space. Comparing the experimental results of column (b) and column (c), the *LSI* color statistical model presented in this paper is used to segment the flame, the detection result is more accurate, and the influence of the reflective part generated on the ground during flame burning can be effectively suppressed.

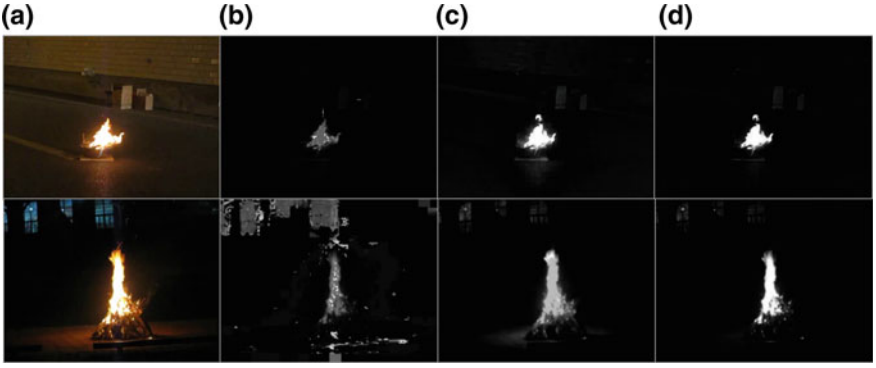



**Fig. 3.** **a** Input images, **b** saliency maps of HSI color space extraction, **c** saliency maps of Lab color space extraction, **d** saliency maps of *LSI* flame color statistical model extraction

## 4 Flame Segmentation Algorithm Based on Bayesian Model

The saliency detection algorithm based on the Bayesian model is mainly used to describe the saliency by the posterior probability in [5–7].

In this paper, one of the motion saliency maps $S_1$ and color saliency maps $S_2$ obtained from Eqs. (1) and (4) is considered as a priori $S(m)(m = \{1, 2\})$, and use the other one $S(n)(n \neq m, n = \{1, 2\})$ to compute the likelihood. The foreground and background regions of $S(m)$ are denoted by $F_m$ and $B_m$, and the number of pixel points in the region is denoted by $N_{F_m}$ and $N_{B_m}$, respectively. The distribution characteristics of the saliency values of $S(n)$ in the two regions $F_m$ and $B_m$ were statistically analyzed, and the histogram of the statistical distribution of $S(n)$ in the foreground and background regions is obtained and normalized. In each region, we compute the likelihoods by comparing $S(n)$ and $S(m)$ in terms of the foreground and background bins at pixel $z$:

$$p(S_n(z)|F_m) = \frac{N_{b_{Fm}}(s_n(z))}{N_{F_m}} \tag{5}$$

$$p(S_n(z)|B_m) = \frac{N_{b_{Bm}}(s_n(z))}{N_{B_m}} \tag{6}$$

Therefore, the posterior probability of $S(m)$ as a priori is shown in Eq. (7):

$$p(F_m|S_n(z)) = \frac{S_m(z)p(S_n(z)|F_m)}{S_m(z)p(S_n(z)|F_m) + (1 - S_m(z))p(S_n(z)|B_m)} \tag{7}$$

We use these two posterior probabilities to compute an integrated saliency map, $\mathrm{Sal}(S_1(z), S_2(z))$, based on Bayesian integration:

$$\mathrm{Sal}(S_1(z), S_2(z)) = p(F_1|S_2(z)) + p(F_2|S_1(z)) \tag{8}$$

## 5 Experimental Results and Analysis

At present, there is no standard test library for flame samples. In order to verify the effectiveness of the algorithm, the self-photographed flame video and Internet sharing flame video are used to test the algorithm of this paper. We compare the algorithm of this paper with the saliency model of [4, 8–10]. The segmentation results are shown in Fig. 4. In Fig. 4, column (a) shows original flame video images for five different scenes. From top to bottom, there are Nocturnal outdoor snow stack flame video images, Nocturnal campus flame video images, Daytime Field Flame video image, Road tunnel flame and driving vehicle lights interference video images and the traffic video images captured by night crossing. Column (b) shows the segmentation result graph obtained by the SR saliency calculation method [4], the algorithm is fast in calculation speed, but the extracted saliency detection effect map cannot segment the moving flame region from the background. Column (c) shows the segmentation result

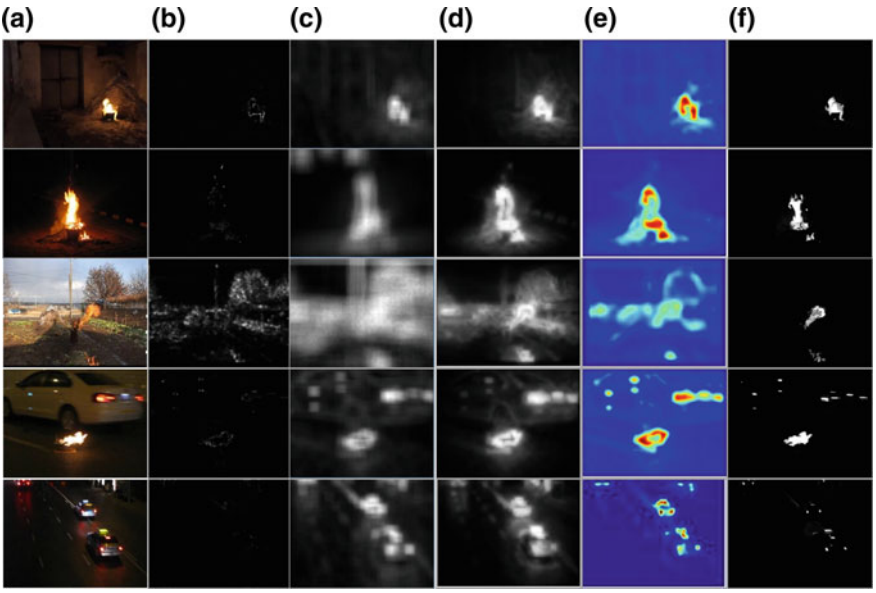



**Fig. 4.** **a** Input images, **b** saliency maps obtained by SR saliency calculation method, **c** saliency maps obtained by SUN-saliency calculation method, **d** saliency maps obtained by context-aware calculation method, **e** saliency maps obtained by AIM model calculation method, **f** saliency maps obtained by this algorithm

graph obtained by the SUN-saliency calculation method [8], the algorithm overemphasizes the edges of the flame rather than the entire target. Column (d) shows the segmentation result graph obtained by the Context-Aware saliency method [9], the accuracy of the algorithm detection is improved, but since only the K closest cell contrasts are superimposed, the detection result sometimes overemphasizes the edges, and the effect of the global contrast cannot be achieved. Column (e) shows the segmentation result graph obtained by the AIM model [10], the algorithm is based on the principle of maximizing the sampling information in the scene to predict the saliency region, but the detection results can only get part of the flame and cannot get the whole area of the flame segmentation. Column (f) shows the experimental result graph obtained by the algorithm, It can be seen that, compared to the above methods, the top three non-interfering flame image segmentation maps from top to bottom are complete and clear in the boundary, and can effectively eliminate the influence of the ground reflective area. However, in the fourth segmented figure, the moving lights and tunnel lights with flame colors of the moving body are also completely segmented in the reflective area. In the fifth segmented figure, the taillights and ceiling lights of vehicles with flame colors are also completely segmented. Pseudo-flame interference can be further eliminated using multi-feature identification of flames.

## 6 Conclusion

This paper proposed a flame segmentation algorithm based on the saliency of motion and color. The algorithm extracts the saliency feature of motion and color from the image sequence. In order to improve the accuracy of the flame segmentation, the SIFT algorithm is used to detect the key points and calculate the motion vector of each frame image combined with the LK optical flow method and the motion saliency map is obtained from the motion vector difference of the pixels. Then, the color saliency map is obtained in the *LSI* color statistical model based on the color and luminance characteristics of the flame. Finally, in the Bayesian framework, the saliency map of motion and color are fused to get the flame segmentation. Experimental results show that the algorithm can segment the flame area accurately and effectively eliminate the influence of the ground reflective area during flame burning. But we cannot exclude the interference of moving vehicle taillights, and we need to further use pattern recognition and other methods to carry out multi-feature comprehensive identification of the flame.

**Acknowledgments.** This paper was supported by the Project of Applied Basic Research Project of Ministry of Transport of the People's Republic of China (Grant No. 2015 319 225 210).

## References

1. Estrada FJ, Jepson AD. Benchmarking image segmentation algorithms. IJCV. 2009;85: 167–81.
2. Tsotsos JK, Culhane SM, Wai WYK, Lai Y, Davis N, Nuflo F. Modelling visual attention via selective tuning. ArtifIntell. 1995;78(1–2):507–45.
3. Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell. 1998;20(11):1254–9.
4. Hou X, Zhang L. Saliency detection: A spectral residual approach. In: CVPR; 2007.
5. Rahtu E, Kannala J, Salo M, Heikkilä J. Segmenting salient objects from images and videos. In: ECCV; 2010. p. 366–79.
6. Xie Y, Lu H, Yang MH. Bayesian saliency via low and mid level cues. IEEE Transa Image Process. 2013;22(5):1689–98.
7. Xie Y, Lu H. Visual saliency detection based on bayesian model. In: ICIP; 2011. p. 645–8.
8. Li X, Lu H, Zhang L, Ruan X, Yang M-H. Saliency detection via dense and sparse reconstruction. In: ICCV; 2013.
9. Zhang L, Tong MH, Marks TK, Shan H, Cottrell GW. SUN: a Bayesian framework for saliency using natural statistics. J Vis. 2008;8(7):1–20, 12.
10. Bruce N, Tsotsos J. Saliency based on information maximization. In: NIPS; 2006.

# A Study on Ring Array Imaging Method for Concealing Weapon Detection

Shuliang Gui(✉) and Jin Li

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, Sichuan, China
guishuliang@l26.com

**Abstract.** In this paper, a ring array imaging method for near-filed security screening is proposed for densely populated areas security imaging. With the fixed transceivers, this method can avoid the motion phase error, which is ineluctable in traditional synthetic aperture radar (SAR). Besides, the number of transceivers is less than planar array radar with the same quality of imaging results. The convolution back propagate (CBP) algorithm is adopted to reconstruct image to avoid the phase error, which is created by plane-wave hypothesis for near-field imaging in high-frequency band. Moreover, the simulation results of CBP and other imaging algorithm are described and compared. And the experimental results are shown to verify the performance of this method.

**Keywords:** Ring array · Near-filed imaging · Security screening

## 1 Introduction

In recent years, terrorism has been increased globally. Most of them occurred in densely populated areas, such as gymnasium, railway station, and music festival. Nevertheless, the traditional security measures, such as X-ray security apparatus and metal detection apparatus, cannot detect and recognize conceal weapons rapidly and effectively [1]. With the development of radar imaging technology and the improvement of computer hardware performance, radar imaging has been new measures in the security inspection field. Especially for the millimeter-wave (MMW) and Terahertz (THz) imaging technology, they can recognize different material items with the radar cross section (RCS) in the high-resolution images. Besides, it has been possible for real-time security imaging with the fast imaging algorithm [2–5].

The main system model of security imaging radar can be divided into two types: mechanical scanning radar and electronic scanning radar. The former one uses few antennas moving to get the echo with a big aperture, and its principle is similar to CSAR. As a typical MMW security imaging radar of the former type, SafeView(L-3) has been widely applied in airports. It can provide a high-resolution whole-body MMW image of one person with its vertical antennas moving along a fixed track. However, it cannot satisfy the needs of densely populated areas security imaging with its cooperative imaging method [6, 7]. Besides, the motion phase error will be a great challenge due to the ultrashort wavelength in MMW and THz band. On the contrary, the electronic scanning radar uses a lot of fixed antennas as a real aperture. The planar array


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_25

developed by Rohde & Schwarz is one of it with 768 TX elements and 768 RX elements. The sequence of transmiting and receiving is controlled by electronic switches, which means that the echo can be acquired in much less than one second without any motion phase error. Nevertheless, it costs a lot for the huge amount of elements [8, 9]. Here, we provide a ring array to avoid the cooperative imaging, motion phase error, and huge costs.

As a near-field imaging, the security imaging has to consider about near-field spherical wavefront. Especially for the MMW and THz imaging, the plane-wave hypothesis will bring a massive phase error, which leads to the severe defocus and deformation at the edge of image [10]. A common way is to reconstruct image via back-projection (BP) algorithm. However, it costs a huge amount of computation [11]. Besides, the Taylor expansion is adopted to decompose the phase error caused by plane-wave hypothesis [12]. Here, the CBP algorithm is adopted to figure it out with less computation than BP and more accuracy than Taylor expansion method.

The remainder of this paper is organized as follows. In Sect. 2, the system geometric model and signal mathematics model are introduced. Section 3 presents micro-drone detection and parameters estimation using micro-motion signatures. The simulation and experimental results are demonstrated and analyzed in Sects. 4 and 5, respectively. And the conclusion of this study is given in Sect. 6.

## 2 Signal Model

Figure 1 shows the geometry of the ring array and a point A of target. As the figure shown, the ring array is formed with N transceivers, which are located in a vertical circle at the plane $z = 0$ with the radius $r_0$ and the fixed angle $\theta$ (for $i$th transceiver can be expressed as $\theta_i$). Assuming that the signal of this ring array radar is step frequency signal, the signal transmitting and receiving steps can be summarized as: (1) set a signal frequency, (2) all the transceivers transmit and receive signal in turns, (3) set the next signal frequency, and repeat the step 2 until all the frequency point signals are transmitted and received.

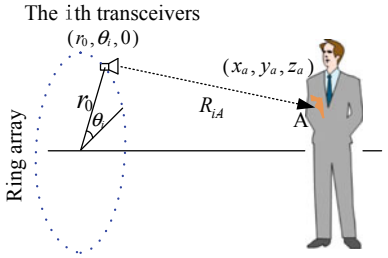



**Fig. 1.** Geometry of the ring array and a point $A$ of target

For a target located at $(x_a, y_a, z_a)$, the distance of propagation between $i$th transceiver and point $A$ can be given as:

$$R_{i-A} = \sqrt{(r_0 \cos\theta_i - x_a)^2 + (r_0 \sin\theta_i - y_a)^2 + z_a^2} \tag{1}$$

The $i$th transceiver echo which reflected from target $A$ can be expressed as:

$$s(\theta_i, f) = \delta(x_a, y_a, z_a) \exp\left(-j4\pi \frac{f}{c} R_{i-A}\right) \tag{2}$$

where $f$ is the frequency of signal, $\delta(x_a, y_a, z_a)$ is the scattering coefficient of point $A$, and $c$ is the speed of light. Hence, all the transceivers echo data reflected from all the targets in imaging scene $O$ can be collected as:

$$s(\theta, f) = \iiint_O \delta(x, y, z) \exp\left(-j4\pi \frac{f}{c} R\right) \mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{3}$$

where $(x, y, z) \in O$ and $R$ is given as:

$$R = \sqrt{(r_0 - x)^2 + (r_0 - y)^2 + z^2} \tag{4}$$

## 3 Imaging Reconstruction Algorithm

As the echo data expressed in (3), for BP algorithm, the scattering coefficient of scene $O$ can be given as:

$$\delta(x, y, z) = \iint_{\theta f} s(\theta, f) \exp\left(j4\pi \frac{f}{c} R\right) \mathrm{d}f\mathrm{d}\theta \tag{5}$$

For a fixed plane $z = h$, with the polar format, the coordinate $(x, y, z)$ can be assumed as:

$$x = \rho \cos\phi, y = \rho \sin\phi, z = h \tag{6}$$

where $\rho = \sqrt{x^2 + y^2}$, $\phi = \mathrm{angle}(x + jy)$. By inserting (6) into (4), the distance $R$ can be rewritten as:

$$\begin{aligned} R &= \sqrt{(r_0 \cos\theta - \rho \cos\phi)^2 + (r_0 \sin\theta - \rho \sin\phi)^2 + h^2} \\ &= \sqrt{r_0^2 + \rho^2 + h^2 - 2r_0\rho \cos(\theta - \phi)} \end{aligned} \tag{7}$$

Assuming that:

$$F(\theta-\phi,f)=\exp\left(j4\pi\frac{f}{c}\sqrt{r_0^2+\rho^2+h^2-2r_0\rho\,\cos(\theta-\phi)}\right) \tag{8}$$

Hence, (5) can be rewritten as:

$$\delta(\rho,\phi)=\iint\limits_{\theta f} s(\theta,f)F(\theta-\phi,f)\mathrm{d}f\mathrm{d}\theta \tag{9}$$

As the transceivers fixed in a complete which means that $\theta\in[-\pi,\pi]$, we can get that:

$$\int\limits_{-\pi}^{\pi} s(\theta,f)F(\theta-\phi,f)\mathrm{d}\theta=s(\phi,f)\odot F(\phi,f) \tag{10}$$

where $\odot$ means the symbol of convolution in angle domain. With the theory of Fourier transform, (10) can be converted into:

$$s(\phi,f)\odot F(\phi,f)=\mathrm{IFFT}_{k_\phi}\left[S\left(k_\phi,f\right)*F\left(k_\phi,f\right)\right] \tag{11}$$

where

$$S\left(k_\phi,f\right)=\mathrm{FFT}_\phi[s(\phi,f)] \tag{12}$$

$$\mathrm{F}\left(k_\phi,f\right)=\mathrm{FFT}_\phi[F(\phi,f)] \tag{13}$$

Hence, (9) can be rewritten as:

$$\delta(\rho,\phi)=\mathrm{IFFT}_{k_\phi}\left[\int\limits_f S\left(k_\phi,f\right)\mathrm{F}\left(k_\phi,f\right)\mathrm{d}f\right] \tag{14}$$

Here, we use twice fast Fourier transform (FFT) and once fast inverse Fourier transform (IFFT) to take the place of the integral operation in angle dimension, which has a large number of samples to satisfy Nyquist sampling theorem in high-frequency band. At last, by performing a two-dimension (2D) interpolation operation, the image of scene $O$ can be reconstructed.

The flowchart of the above image reconstruction algorithm is shown in Fig. 2.

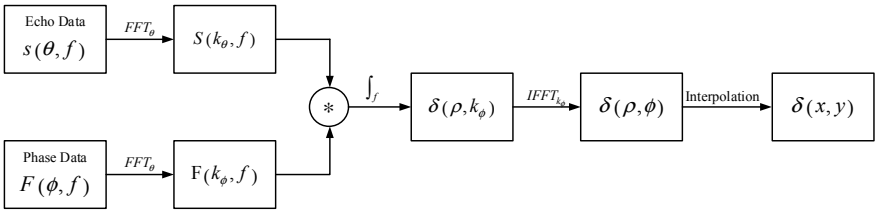



**Fig. 2.** Flowchart of the image reconstruction algorithm

## 4 Simulation Results

The simulation parameter is shown in Table 1. Here, $f_0$ and $B$ are the center frequency and bandwidth of signal, respectively; $N_f$ and $N_\theta$ are the number of frequency samples and transceivers, respectively.

**Table 1.** Simulation parameters

| $f_0$ | $B$ | $r_0$ | $h$ | $N_f$ | $N_\theta$ |
|---|---|---|---|---|---|
| 330 GHz | 28.8 GHz | 1 m | 1 m | 256 | 1024 |

The simulation imaging targets are points, and the location of these points is shown in Fig. 3. And the results of simulation via different imaging algorithms are shown in Fig. 4 with normalized operation.

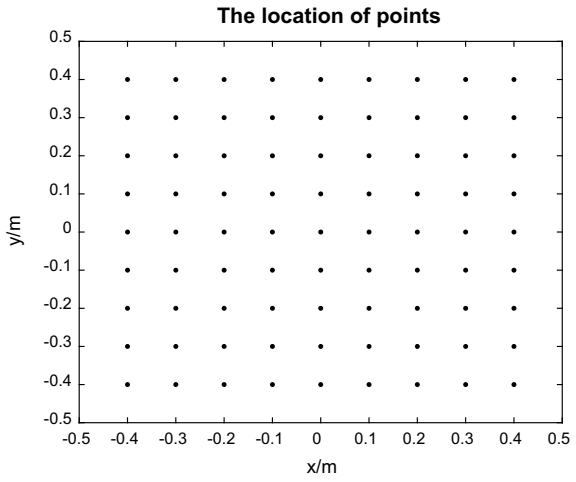



**Fig. 3.** Location of simulation imaging point targets

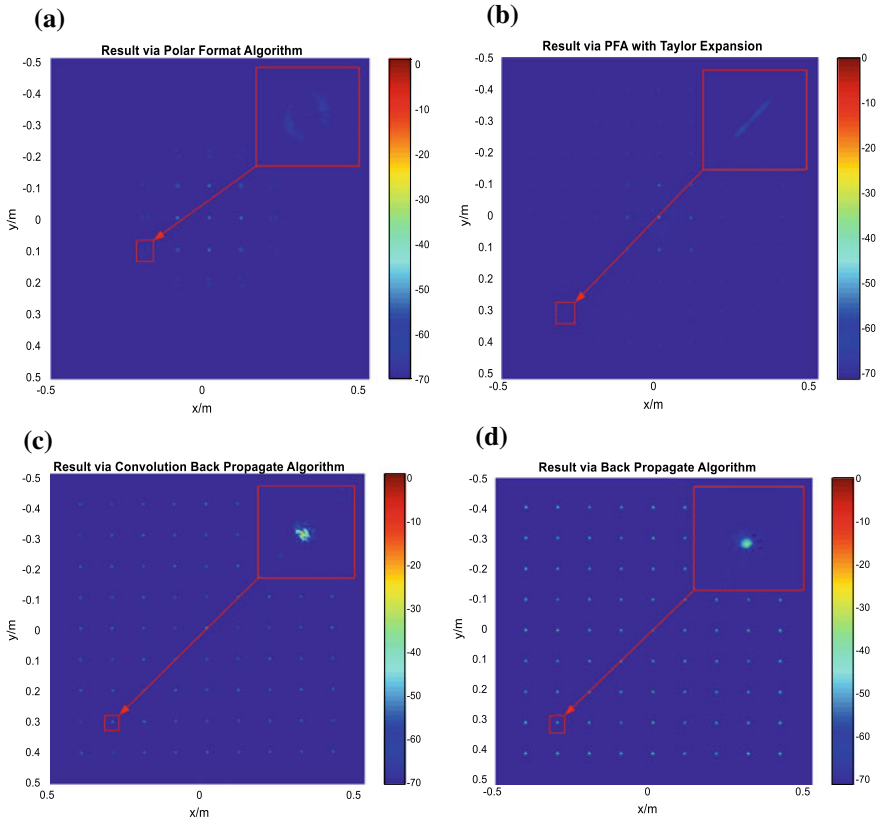



**Fig. 4.** Results of simulation via different imaging algorithms: **a** polar format algorithm (PFA); **b** PFA with Taylor expansion correction (PFA-TEC); **c** back propagate algorithm; **d** convolution back propagate algorithm

By comparing with Fig. 4a, b, it is obvious that the scene range of effective imaging has been expanded with the Taylor expansion correction. But the defocus and deformation phenomenon still appears in the edge of images. Besides, as the results shown in Fig. 4c, d, the image can be reconstructed via BP algorithm and CBP algorithm without defocus and deformation both. Nevertheless, due to the affection of interpolation precision, Fig. 4d is not as evenly as Fig. 4c. The run times via different algorithm are shown in Table 2 with the CPU (Inter Core i5-4590) and 8 GB RAM. As Table 2 told, the run time of CBP imaging algorithm is the shortest except polar format algorithm. Hence, the CBP is the most suitable imaging algorithm for the proposed ring array radar security imaging.

**Table 2.** Run times with different algorithms

| Algorithms | Run time (s) |
|---|---|
| PFA | 2.18 |
| PFA-TEC | 103.46 |
| BP | 229.73 |
| CBP | 17.53 |

## 5 Experimental Result

Here, we use a THz transceiver with inverse synthetic aperture radar (ISAR) imaging method to simulate the proposed ring array imaging method. The parameters of this experiment are same as Table 1 expect that $r_0 = 2m$.

Figure 5a shows the picture of target, and Fig. 5b shows the imaging result of targets via CBP algorithm. As Fig. 5 shown, the proposed imaging method has reconstructed the image of pistol successfully via CBP algorithm.

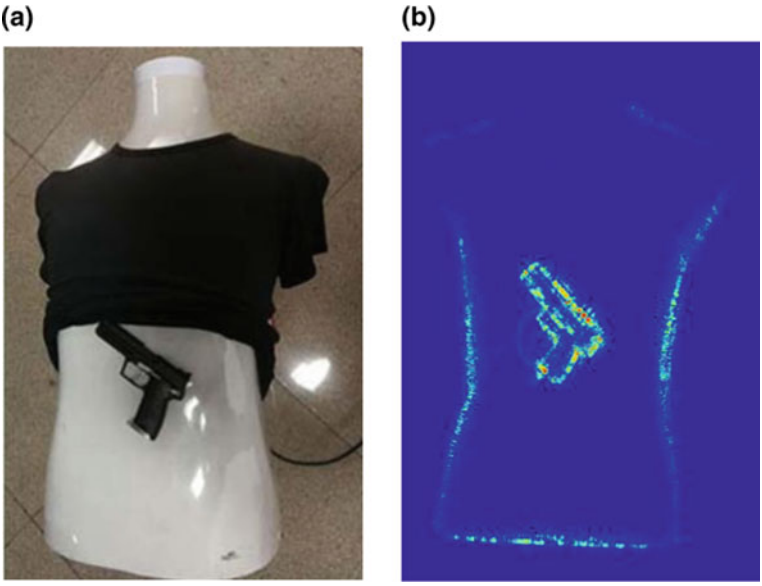



**Fig. 5.** **a** Experimental target. **b** Imaging result via convolution back propagate algorithm

## 6 Conclusion

A ring array imaging method for concealing weapon detection is proposed in this paper. The main contribution of this paper is as follows. First, the imaging method can be used as security screening for densely populated area with less antennas. Second, the

fast imaging algorithm without plane-wave approximation phase error is addressed. In the end, the simulation and experimental results have been comparatively analyzed to verify the performance of the proposed imaging method and algorithm.

**Acknowledgments.** This work was supported by the 2011 Project for Collaborative Innovation Center of Information Sensing and Understanding.

## References

1. Zhang B, Pi Y, Yang X. Terahertz imaging radar with aperture synthetic techniques for object detection. In: IEEE international conference on communications workshops; 2013; IEEE. p. 921–5.
2. Sheen DM et al. Standoff concealed weapon detection using a 350-GHz radar imaging system. In: Proceedings of SPIE, vol 7670.1; The International Society for Optical Engineering; 2010. p. 115–8.
3. Sheen DM, McMakin DL, Hall TE. Three-dimensional millimeter-wave imaging for concealed weapon detection. IEEE Trans Microw Theory Techn. 2001;49(9):1581–92.
4. Demirci S, Cetinkaya H, Yigit E, Ozdemir C, Vertiy A. A study on millimeter-wave imaging of concealed objected: application using back-projection algorithm. Prog Electromagn Res. 2012;128:457–77.
5. Chattopadhyay G. Terahertz science, technology, and communication. In: International conference on computers and devices for communication; 2012. p. 1–4.
6. Sheen DM, Mcmakin DL. Combined illumination cylindrical millimeter-wave imaging technique for concealed weapon detection. ProcSpie. 2000;4032(2000):52–60.
7. Sheen DM, Mcmakin DL. Circularly polarized millimeter-wave imaging for personnel screening. In: Proceedings of SPIE, vol 5789; The International Society for Optical Engineering; 2005. p. 117–26.
8. Ahmed SS, Schiessl A, Schmidt LP. A novel fully electronic active real time imager based on a planar multistatic sparse array. IEEE Trans Microw Theory Tech. 2011;59(12):3567–76.
9. Schiessl A, Ahmed S S, Genghammer A et al. A technology demonstrator for a 0.5 m × 0.5 m fully electronic digital beamforming mm-wave imaging system. In: Proceedings of the 5th European conference on Antennas and propagation (EUCAP); IEEE; 2011. p. 2606–9.
10. Guo Q, Chang T, Geng G, et al. A high precision terahertz wave image reconstruction algorithm. Sensors. 2016;16(7):1139.
11. Demirci S, Cetinkaya H, Tekbas M et al. Back-projection algorithm for ISAR imaging of near-field concealed objects. In: General assembly and scientific symposium; XXXth URSI. IEEE, 2011. p. 1–4.
12. Garber WL. Extensions to polar formatting with spatially variant postfiltering. In: Proceedings of SPIE, vol 8051, No. 5; The International Society for Optical Engineering, 2011. p. 361–372.

# Image Parameters Evaluation for Road Lighting Based on Clustering Analysis

Yi Xiong, Ning Lv, Xufen Xie(✉), and Yingying Shang

Research Institute of Photonics, Dalian Polytechnic University, Dalian, China
xiexufenll@foxmail.com

**Abstract.** Road lighting is a main factor which impacts on traffic accident rate. The valuable lighting evaluations are the fundament of road lighting design. We propose five classes parameters which come from road lighting images to evaluate the quality of road lighting in this paper. We first calculate 10 image parameters from road lighting images. It includes mean value of gray level, variance of gray level, radiation precision steepness, gray level entropy, second moment of angle, contrast, autocorrelation, inverse difference moment, detail energy, and edge energy. Then, we divide the above 10 parameters into five categories using cluster analysis. These categories are mean value class, variance class, contrast class, detail energy class, and information-related class. Finally, combined with the physical meaning of the parameters, the evaluation index of the traditional road lighting and the characteristics of the human eye, we connect these five categories with the average brightness of pavement, the uniformity of road surface brightness, glare, road sign inducibility, and psychological factors. The experimental results show that the road lighting image parameters have good clustering properties, and the clustered image parameters can reflect the quality of road lighting.

**Keywords:** Road lighting evaluation · Image characteristic parameters · Systematic cluster analysis

## 1 Introduction

Glare of road lighting and unclear sight of driver are important factors of traffic accidents at night. The investigation of the international lighting committee (CIE) manifests that good nighttime road lighting environment can effectively reduce the urban traffic accident rate by 30–40%. Therefore, the improving quality of road lighting can reduce traffic efficiency and ensure traffic safety. To improve the quality of road lighting, we must evaluate road lighting more effectively. Road lighting measurement can get quantitative information of road lighting and provide important reference for road lighting construction. For a long time, road lighting design has been carried out with reference to the "luminance standards" of pavement. The brightness information of the road surface and the light curtain caused by the stray light can affect the visibility of the road area and the objects around the road. Through imaging for road lighting scenes, we get digital images including light sources [1], roads, and scenery. Combined with imaging technology and digital image processing technology,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_26

the calculation and estimation of the brightness, glare, and many statistical parameters in the scene can be used to evaluate the road lighting in an all-round way.

This paper proposed a method to evaluate road lighting quality based on road lighting image parameters. The clustering analysis technology is used to classify gray level mean, gray level variance, radiation precision steepness, gray entropy, second moment of angle, contrast, autocorrelation, inverse difference moment, detail energy and edge energy into gray level mean class, gray level variance class, contrast class, detail energy class, and information-related class. Combined with the physical meaning of the parameters and the evaluation index of traditional road lighting and the characteristics of human eyes, the parameters are analyzed.

## 2 Extraction of Image Parameters

The road lighting scene includes the characteristics of the road lighting. And the image of road lighting scene sampled typical characteristics of road lighting [2]. Therefore, the road lighting can be analyzed and evaluated by the road lighting images. A lot of image parameters describe the characteristics of image scene. These characteristics can be used to evaluate scene characteristics and lighting quality. The common image characteristic parameters are shown in Table 1. These parameters can be categorized as three major parts from gray level parameters, texture parameters [4], and edge parameters.

**Table 1.** Parameter of image quality evaluation and its significance

| Type | No. | Name |
|---|---|---|
| Gray level parameters | 1 | Gray level mean |
| | 2 | Gray level variance |
| | 3 | Radiation precision steepness |
| | 4 | Gray entropy |
| Edge parameters | 5 | Edge energy |
| Texture parameters | 6 | Second moment of angle |
| | 7 | Contrast |
| | 8 | Autocorrelation |
| | 9 | Inverse difference moment |
| | 10 | Detail energy |

## 3 Cluster Analysis of Road Lighting Image Parameters

### 3.1 System Clustering Analysis

There are a lot of parameters in images. The relativity of the parameters needs to be analyzed. Cluster analysis is to aggregate samples according to the similarity of quality

characteristics [5]. According to the comprehensive nature of the category, multiple varieties are aggregated using cluster analysis. We used the deviation square sum method (Ward method) to analyze the 10 image parameters in case (Q) system cluster analysis. One road lighting image is shown in Fig. 1, and the cluster analysis results are

**Fig. 1.** Original gray level image

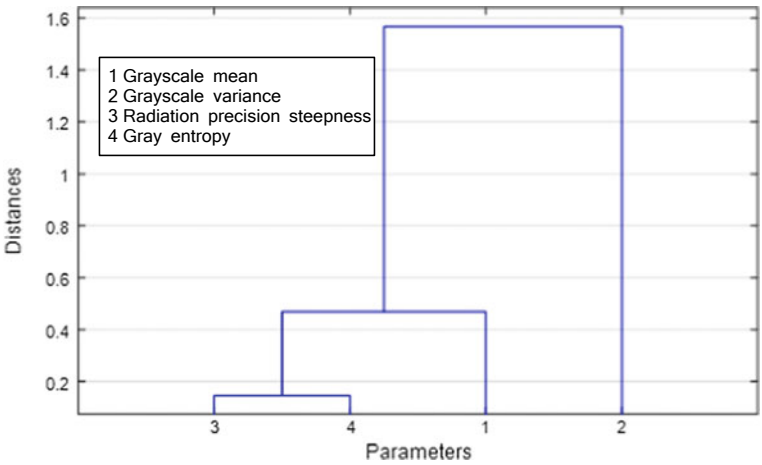



**Fig. 2.** Gray parameter cluster analysis dendrogram

shown in Figs. 2, 3, and 4.

### 3.2 Preliminary Cluster Screening of All Kinds of Image Parameters

According to the dendrogram of cluster analysis (Fig. 2), when the distance between classes is = 0.4, the four gradation parameters can be divided into three categories. The first group aggregates two gray scale parameters, namely radiation precision steepness and gray entropy. The second is the gray level mean and the third is the gray level variance.

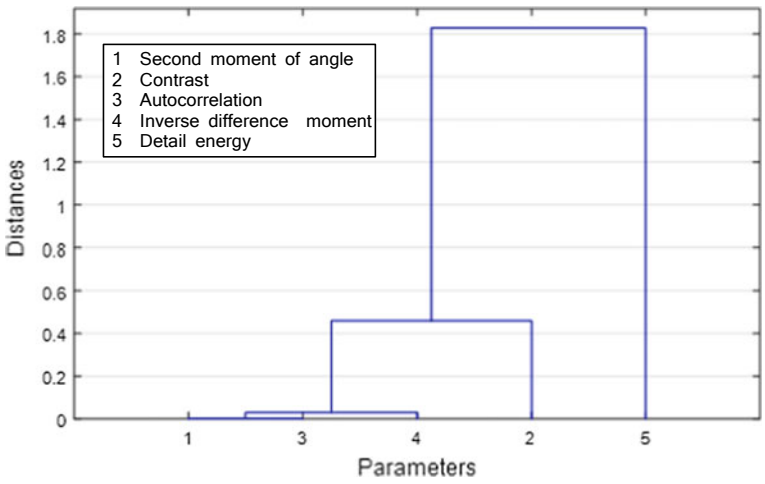



**Fig. 3.** Texture parameter cluster analysis dendrogram

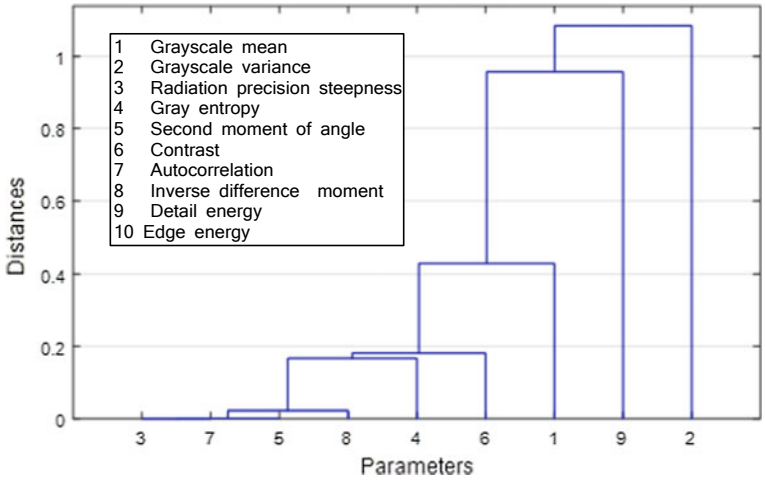



**Fig. 4.** Two screening parameter cluster analysis dendrogram

From the cluster analysis dendrogram (Fig. 3), we can see that five texture parameters are divided into three classes when the distance between classes is = 0.4. The first class aggregates three-edge parameters; they are second moment of angle, autocorrelation, and inverse difference moment. The second category is contrast, and the third is the detail energy.

By clustering analysis, three categories are separated from gray and texture parameters. A random selection of parameters from each class is carried out again for cluster analysis. The parameters selected in this paper are gray level mean, gray level

variance and gray entropy in gray parameters, contrast, inverse difference moment and detail energy in texture parameters, and edge energy in edge parameters. According to the dendrogram of cluster analysis (Fig. 4), seven texture parameters can be divided into five categories: gray level mean class, gray level variance class, contrast class, detail energy class, and information-related class. The first class has gray level mean, the second class has gray level variance, the third kind has contrast, the fourth kind has the detail energy, and the fifth class has edge energy, the gray entropy, inverse difference moment, second moment of angle, autocorrelation, and radiation precision steepness. Among them, the fifth category also contains parameters that are not selected for re-clustering because of repeated clustering (Table 2).

**Table 2.** Classification of image parameters

| No. | Class name | Including original image parameters |
|---|---|---|
| 1 | Gray level mean class | Gray level mean |
| 2 | Gray level variance class | Gray level variance |
| 3 | Contrast class | Contrast |
| 4 | Detail energy class | Detail energy |
| 5 | Information-related class | Edge energy, gray entropy, inverse difference moment, second moment of angle, autocorrelation, radiation precision steepness |

## 4 The Significance of Image Parameters Evaluation

### 4.1 Evaluation Indicators for Road Lighting

The Urban Road Lighting Design Standard (CJJ 45-2015, it implemented by the Ministry of Housing and Urban-Rural Development on June 1, 2016) states that the average road surface brightness or average road surface illumination, road surface brightness uniformity, and longitudinal uniformity or evenness of road surface

**Table 3.** Design index and significance of urban road lighting

| No. | Name |
|---|---|
| 1 | Average road surface brightness |
| 2 | The total brightness of road surface |
| 3 | Road brightness longitudinal uniformity |
| 4 | Glare limitation of lighting |
| 5 | Road lighting inducibility |
| 6 | Environmental ratio SR |
| 7 | Average road surface illumination |
| 8 | Road illumination total uniformity |

illumination, glare limitation, environmental ratio, and inducibility are valid evaluation indexes [9].

According to CJJ 45-2015, Table 3 shows the design index and significance of urban road lighting. The average brightness of the road is the average value of the brightness of each point obtained at the point set on the road in accordance with the relevant provisions of the international lighting committee (CIE).

The brightness of the road pavement manifests the overall uniformity of brightness and brightness of the average ratio of the minimum. Longitudinal uniformity of road surface luminance refers to the ratio of minimum and maximum luminance brightness on the same lane center. Glare is a visual phenomenon that causes discomfort or reduces the ability to observe objects or details due to the discomfort of the brightness distribution or luminance range in the field of vision, or the existence of extreme contrast. In order to prevent glare phenomenon, it needs to calculate the road lighting glare limit. The proper installation of lamps and lanterns along the road can provide drivers with visual information on the direction, line type, and slope of the road, which is called the inducibility of road lighting. SR refers to the environment than the average level of illumination and the adjacent roadway outside 5 m wide area like 5 m wide roadway illumination than the average level. Road surface illuminance on average represents the mean value according to CIE relevant provisions set in advance on the road on the measured or calculated the illuminance. The total surface illumination uniformity is on the road and the average illuminance ratio of minimum illumination.

### 4.2 Significance of Clustering Image Parameters

The human visual system (HVS) [6] model is used to simulate the visual perception process. In the actual visual environment, objects in turn undergo visual nonlinearity, visual sensitivity bandpass, visual multi-channel, and masking effects [10]. At the same time, they are affected by human visual physiological functions in the process, and ultimately produce images in the human eye, then output the result. Because of some indicators in road lighting, it is impossible to describe the road lighting quantitatively, and there are indicators which are influenced by the individual's intuitive feelings [3].

Considering all these considerations, this paper only discusses some indicators for evaluating road lighting. To imaging the actual environment of road lighting, we can obtain the evaluation parameters related to the average brightness of road surface, the total uniformity of road brightness, the limit of glare, the environment ratio, and the inducibility. The correspondence relationship is shown in Fig. 5.

(1) Glare is the visual condition of discomfort and visibility of objects, due to the incorrect distribution or range of luminance in the field of vision, or the existence of extreme brightness contrast in space or time. The contrast in image parameters is the ratio of the brightest part and the darkest part of the image, or the logarithmic difference. Glare is closely related to brightness and shade ratio in the environment. Image contrast is also closely related to brightness and shade ratio in the environment. Therefore, there is a certain correlation between contrast in road lighting images and glare in actual road lighting.

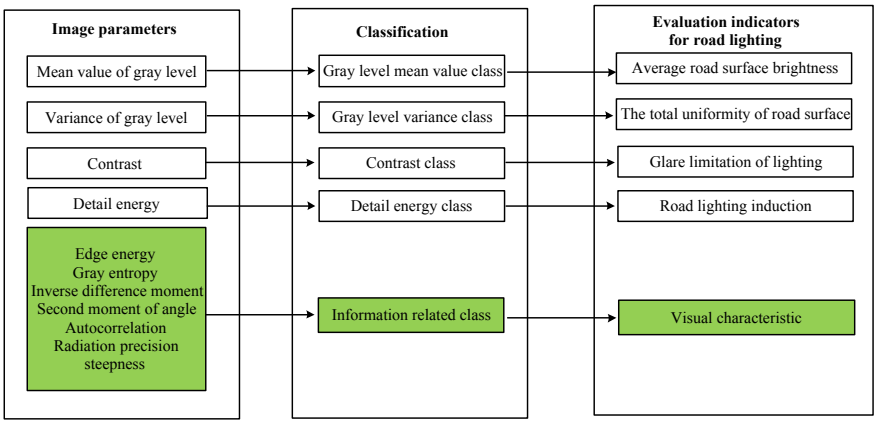



**Fig. 5.** Significance analysis of image parameter clustering

(2) The total uniformity refers to the ratio of the minimum brightness to the average luminance on the road surface. The lower the total uniformity under the same illumination, the harder it is to detect obstacles. The variance of the gray level reflects the richness of the gradation level of the image, which reflects the size of the high-frequency part of the image. The larger the gray scale variance, the greater the difference between the high-frequency and low-frequency components. The high-frequency part of the image and the road obstacles are the parts that need attention in image and environment. The smaller the image gray variance, the more prominent the high-frequency part of the image is. The lower the total brightness of the road surface is, the less obvious the road obstacles are. Therefore, there is a certain correlation between the variance of image gray and the total uniformity of road surface brightness.

(3) Mean brightness of road surface refers to the luminance average of all test points on the road surface. The mean value of the image reflects the brightness of the image as a whole, and it is a quantitative reflection of the actual radiation energy of the scene objects. The average gray level mean reflects the average brightness of the image. Mean brightness and image mean value are the mean values of environmental brightness. Therefore, we can conclude that the average brightness of pavement can be reflected to some extent by extracting and analyzing the gray level mean of road lighting images [9].

(4) Road lighting facilities should be able to provide good inducibility. It plays an important role in traffic safety and comfortable driving. The detail energy describes the richness of the detail edges of the image from the local part of the image. The greater the value of detail energy, the more abundant the information provided by the image. If road lighting provides a lot of useless information to the human eye, it will lead to wrong visual induction. This will lead to distraction of drivers and cause safety hazards. It can be seen that by extracting the detail energy in the road lighting image, the inducement of road lighting can be analyzed to a certain extent [8].

(5) The system cluster analysis includes the edge energy, the gray entropy, and the inverse distance. According to the physical properties, it is classified as an information-related class, which represents the relationship between gray values.

## 5 Conclusions

A new method for evaluating the quality of road lighting is presented in this paper. The method is based on the images parameters of road lighting,gray level mean value, gray level variance, radiation precision steepness, gray entropy, second moment of angle, contrast, autocorrelation, inverse difference moment, detail energy, and edge energy are clustered, and it is classified as gray level mean class, gray level variance class, contrast class, detail energy class, and information-related class. By analyzing the physical meaning of the parameters, a correspondence relationship between the evaluation indexes of traditional road lighting and the characteristics of human eyes is established. 20 frame road lighting images are selected and analyzed in experiments. The results show that road image parameters have good clustering properties. To some extent, clustered image parameters can evaluate the quality of road lighting. The method has good operability, and it can make up for parameter indicators and fills the shortage of lacking indicators in road evaluation. It has great application prospects in road lighting evaluations.

## References

1. Yingkui H, Zhonglin C, Yingpiao L. Light effects of common road light sources under intermediate vision conditions. J Chongqing Univ. 2007;30(1):139–41.
2. Kang W. Research on road lighting detection based on luminance imaging technology. Doctoral dissertation. Zhejiang University; 2016.
3. Yiying W, Doudou C, Liang Z, Jun M, Wenhui N. Image quality evaluation method based on spatial similarity of masking effect. J Hefei Univ Technol: Nat Sci Edn. 2015;10: 1339–41.
4. Xiaobing X, Lei C, Jianping W. Research and application of road lighting characteristics based on intermediate vision. J Hefei Univ Technol (Nat Sci). 2013;36(6):704–8.
5. Liyan G, Xianjun M, Naiqiao L, Jinfeng B. Evaluation of apple processing quality based on principal component and cluster analysis. J Agric Eng. 2014;30(13):276–85.
6. Chunhua P, Tonglin Z, Hao L. HVS evaluation method for image quality. Comput Eng Appl. 2010;46(4):149–51.
7. Chen X, Zheng X, Wu C. Portable instrument to measure the average luminance coefficient of a road surface. Meas Sci Technol. 2014;25(3):35203–9.
8. Cattini S, Rovati L. Low-cost imaging photometer and calibration method for road tunnel lighting. IEEE Trans Instrum Meas. 2012;61(5):1181–92.
9. China Academy of Building Research. Urban road lighting design standards CJJ45-2015. China Building Industry Press; 2016.
10. Shuqin L, Lifang Y, Gong Y, Xingsheng L. Review of image quality assessment. Chin Sci Technol Pap. 2011;06(7):501–6.

# A Novel Method for Detecting the Circle on Motion-Blurred Image

Fengjing Liu[1], Xing Zhou[2], Ju Huo[2(✉)], Yunhe Liu[1], Ming Yang[3], and Shuai Liu[1]

[1] Beijing Institute of Spacecraft System Engineering, Beijing 100094, China
[2] School of Electrical Engineering, Harbin Institute of Technology, Harbin 150001, China
torch@hit.edu.cn
[3] Control and Simulation Center, School of Astronautics, Harbin Institute of Technology, Harbin 150001, China

**Abstract.** As a typical feature point with the distinct advantage of being detected easily, the circle has been widely used for camera calibration and motion measurement. However, motion blur may cause a negative effect on the accuracy of the center location. In this paper, the developed method for the circle detection on motion blur image is proposed, which consists of two procedures. Wiener filtering is used to restore a degraded image in the first step. Zernike moment is utilized to subpixel central location in the second step. Image restoring simulation and center detection experiments are provided to verify the performance of the method. Results show that the clarity of the images restored by Weiner filtering is high and the circles on the restored image can be detected successfully and located accurately.

**Keywords:** Stereo vision · Code recognition · Circle detection · Motion blur

## 1 Introduction

As a non-contact metrology, the visual measurement has been developed rapidly and commonly applied in the field of measurement for its high accuracy and speed [1, 2]. In the visual measurement, checkerboard corners, crossing lines and circles have always been collocated on the surface of measured objectives. Since circular markers have the advantage of being easily detected, they are widely used in machine vision field. Due to the rotation invariance of the circle, encoding circle is usually adopted to accelerate matching procedure and diminish the complexity of measuring procedure [3]. Based on the above mentioned, the coded target is utilized in this paper. The code of this target is prone to be recognized, so this target can accelerate the stereo matching process while calibrating the binocular camera.

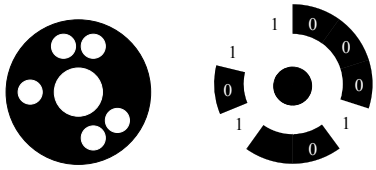
In recent years, various studies have been carried out on the circle detection and location since the high accurate location can develop the accuracy of image matching and measuring results [4, 5]. Due to the uses of its result in camera calibration, 3D construction and other follow-up processes, matching image is considered as a key step of measuring process by using the coordinates of feature points in different images.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_27

Besides, in measuring process, especially in motion measurement, the coordinates of the circles on the objectives are utilized to obtain motion trails, deformation parameters and other measuring information. Therefore, in measurement process, high-precision central location can improve the precision of measurement. However, in shoot process, the tremble of hand and the motion of objects can result in the motion blur [6]. Suppose that detect the circle in blurred image directly, the precision of central location could not meet the standard of measurement. To solve the problem, this paper proposes a method of detecting the circle on motion-blurred image. The method includes two simple but effective approaches. Firstly, estimate the motion parameters and restore image by Weiner filtering. Secondly, detect the edge of the circle by using Zernike moment and locate central. This method is carefully tested on simulation and real experiment to validate its performance. Results reveal that detecting the circle on motion-blurred image by this method is more effective.

## 2 Coded Target

Two commonly used coded targets, shown in Fig. 1, are circular coded target and Schneider's coded target [7], respectively. With a view of the count of codes, Schneider's coded target is adopted in this paper.




(a) Circular coded target (b) Schneider's coded target

**Fig. 1.** Coded target

The white pattern is associated with digit 0, while the black pattern is associated with digit 1; this association is shown in Fig. 1b. While decoding the pattern, each pattern is been read anticlockwise or clockwise. Regard the minimum of these k binary numbers as the code of each target to guarantee the uniqueness of the decoded result. Table 1 lists the number of codes with different code digits. The code target proposed in this paper is associated with a ten-digit binary number, and that means this coded target includes ten code patterns. The size of the coded target is illustrated in Fig. 2; $\phi_{\min}$ is the diameter of circular target, while $\phi_{\text{mid}}$ is the internal diameter and $\phi_{\max}$ is the outside diameter of coded band.

**Table 1.** Number of codes with different code digits

| Coded digits | 4 | 8 | 10 | 12 | 15 |
|---|---|---|---|---|---|
| The number of codes | 5 | 35 | 107 | 351 | 21,291 |

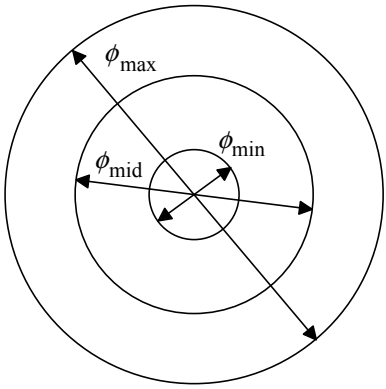



**Fig. 2.** Coded target size

Decoding process is as follows: After ellipse fitting, the circle which is the inverse affine transformation of the ellipse with diameter $\frac{\phi_{max}+\phi_{mid}}{2}$ is set as the decoded target. The gray value of each $360^\circ/k$ degree is extracted firstly, and the corresponding binary digit is obtained secondly. The minimum of all the code numbers is set as the decoded result. This process is illustrated in Fig. 3.

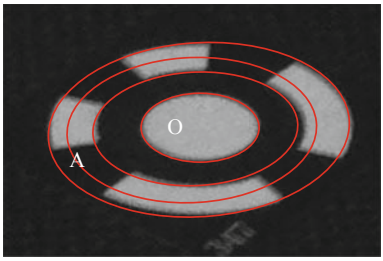



**Fig. 3.** Schematic diagram of decoding

## 3 The Method of Restoring Image

### 3.1 The Degeneration and Restoration Model of Motion-Blurred Image

As a common distortion, motion-blurred image can be resulted due to the relative movement between the object and the camera. In general, the model of degenerated image in time domain can be described by

$$g(x,y) = h(x,y) * f(x,y) + n(x,y) \tag{1}$$

where $h(x, y)$ denotes point spread function (PSF), $f(x, y)$ denotes clear image, $*$ is convolution operator and $n(x, y)$ is noise. The expression of point spread function is [8]

$$h(x,y) = \begin{cases} 1/L, & \sqrt{x^2+y^2} \le L, x/y = -\tan(\theta) \\ 0, & \text{else} \end{cases} \tag{2}$$

where $L$ denotes motion scale and $\theta$ is motion direction. The expression indicates obviously that the estimates of motion scale and motion direction are the key factors of restoring image. The Fourier transform of motion-blurred image without noise is written as

$$G(u,v) = H(u,v)F(u,v) = F(u,v) \cdot \frac{\sin(\pi uL)}{\pi uL} e^{-j\pi uL} \tag{3}$$

Thus, it can be seen that the black stripes would occur in the spectrum image of blurred image because of the zeros of sine [9]. The orientation of black strips is perpendicular to the motion direction; hence, the motion direction can be found by extracting orientation of spectrum image of blurred image. After the obtainment of both motion direction and scale, blurred image can be restored by Wiener filtering. Making the mean square error (MES) of restored image and original image is the fundamental principle of Wiener filtering. However, the power spectrum of original image and noisy image, which is used in Wiener filtering, is hard to obtain. Therefore, approximation of Wiener filtering is adopted in this work

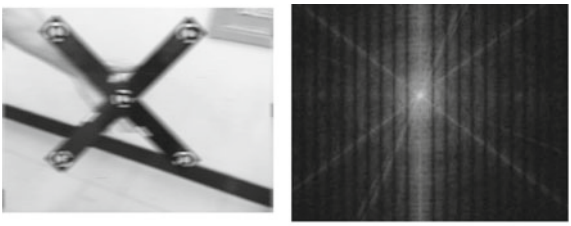
$$\widehat{F}(u,v) = \left[ \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K} \right] G(u,v) \tag{4}$$

where

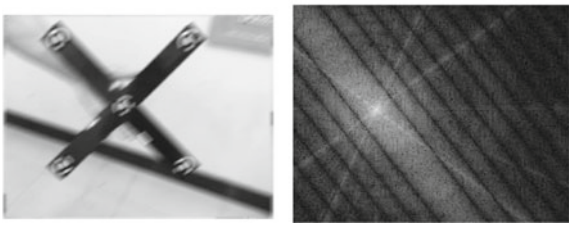
$$K \approx \frac{\text{Average noise power spectrum}}{\text{Average image power spectrum}}.$$

### 3.2 The Estimation of Motion Direction and Scale

As Fig. 4 shows, the light and dark strips occur in the spectrum of blurred image, and these stripes are perpendicular to motion direction, so the motion direction can be determined by spectrum analysis. The orientation can be extracted by Radon transform. The corresponding angle of the maximum value is vertical to strips, namely motion direction. Nevertheless, noise exists in actual degeneration model and it has a negative effect on the accuracy of direction evaluation. Therefore, the median filter algorithm is adopted to preprocess the image firstly, and Fourier transform is applied secondly. After Radon transform, fit a curve with the maximum value of each angle. The corresponding angle of the peak of the curve is the direction of the motion.



(a) Horizontal direction



(b) 30 degree

**Fig. 4.** Spectrum of motion-blurred image

After the motion direction is determined, the motion scale can be obtained by differential autocorrelation function [10]. First of all, rotate the blurred image at the motion angle and obtain the horizontal differential of the rotated image. The distance between two minimums is the motion scale of blurred image. Once motion scale and direction are obtained, image can be restored by Wiener filtering afterward. Considering the effect of noise, this paper uses morphological filter and thresholding method before central location to denoise.

## 4 High-Precision Circle Detection Method

### 4.1 Subpixel Edge Detection Method

S. Ghosal and R. Mehrotra proposed a subpixel edge detection method based on Zernike moment [11]. The gray value of image is mapped to a unit circle, and the edge feature is shown in Fig. 5. The part within the circle of line $L$ can be regarded as ideal edge.

The Zernike moment of 2D image is described as

$$Z_{nm} = \frac{n+1}{\pi} \iint\limits_{x^2+y^2 \le 1} f(x,y) V_{nm}^{*}(\rho,\theta) \mathrm{d}x\mathrm{d}y \tag{5}$$

* represents complex conjugate. $V_{nm}(\rho,\theta)$ is integral kernel function, and it can be written in a form that corresponds with the polar geometry

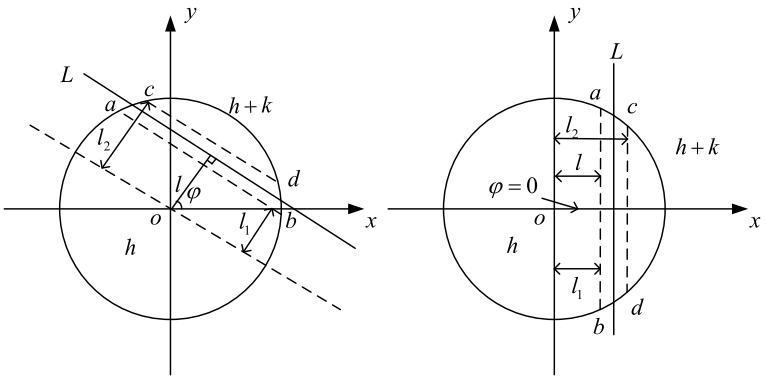



**Fig. 5.** Ideal edge detection model of the Zernike moment

$$V_{nm}(\rho,\theta) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!\rho^{n-2s}}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n+|m|}{2}-s\right)!} \mathrm{e}^{-im\theta} \tag{6}$$

According to the rotation invariance of Zernike moment, the edge can be detected by the Zernike moment which is been turned $\varphi$ degree. Obtain the Zernike moment $\{Z_{00}, Z_{11}, Z_{20}, Z_{31}, Z_{40}\}$ of image firstly, and $l$, $k$, $h$, $\varphi$ can be obtained secondly. The edge can be expressed by these parameters. Suppose the size of the moment template is $N \times N$, the error caused by Zernike moment template can be eliminated by

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \frac{Nl}{2} \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix} \tag{7}$$

where $(x, y)$ is the edge detected by Sobel operator. $(x_s, y_s)$ is the subpixel edge detected by Zernike moment.

However, in actual imaging process, the gray value near the edge follows Gaussian distribution, so the ideal edge detection is inaccurate. Suppose the edge and $y$-axis are parallel, and the actual edge model is shown in Fig. 6. The actual distance between the edge and the center of model is $l$. The subpixel edge position error is determined by

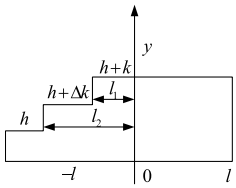



**Fig. 6.** Actual edge model

$$E = \frac{\delta(\delta-1)(1-l_2)\left(\sqrt{(1-l_1^2)^3}-\sqrt{(1-l_2^2)^3}\right)}{\delta\sqrt{(1-l_1^2)^3}+(1-\delta)\sqrt{(1-l_2^2)^3}} \tag{8}$$

where

$$\delta = \Delta k/k \tag{9}$$

If $\delta \neq 1$ and $\delta \neq 0$, it holds that $E \neq 0$. In addition, position error $E$ increases with the distance between $l_1$ and $l_2$. The error of actual subpixel edge can be compensated by

$$l = l_r - E \tag{10}$$

The precision can be improved after the error compensation. And the edge pixel can be evaluated for higher accuracy by

$$k \geq k_t \cap |l_2 - l_1| \leq l_t \tag{11}$$

where $k_t$ and $l_t$ are the thresholds of $k$ and $l$, respectively.

### 4.2 Algorithm of Central Location

While the spatial circle plane is perpendicular to the optical camera axis, the coordinate fitting center of ellipse projection in spatial plane is equal to the projection of the spatial circle center. Often, the spatial circle plane is not perpendicular to the optical camera axis. In this circumstance, a line passing through the two points is in the image plane. The slope $k$ of this line is the independence of the radius of the spatial circle. And after camera imaging process, the actual circle center projection, the geometrical center of ellipse and fitting center of ellipse are collinear.

Unqualified edge should be removed according to the circle criterion after the circular edge detection. Deal with regular edge and obtain two points that have the longest distance, and the midpoint of them is the center of the ellipse projection. Afterward, judge the distance of these points and fitting center. If the distance is closer than a given threshold, the midpoint of these two centers is regarded as the actual projection of the spatial circle center. If not, make a line which through these two centers intersects image plane at $e_{\text{image}}(u_e, v_e)$ and $f_{\text{image}}(u_f, v_f)$. Suppose the coordinate of the actual projection of the spatial circle center is $O_{\text{image}}(u_o, v_o)$. According to the invariance of straight line and the cross ratio while affine transformation, a coordinate of ellipse center can be determined by

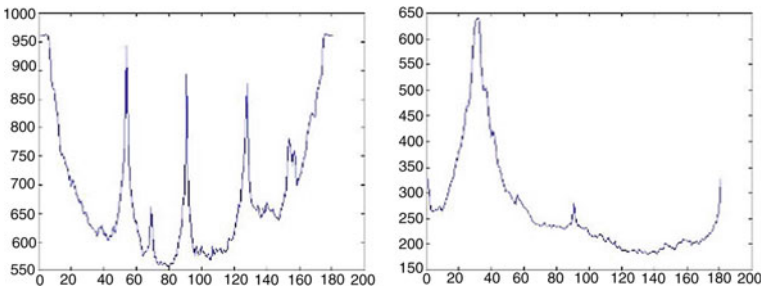
$$\begin{cases} \frac{u_o-u_e}{u_f-u_e} = \frac{R}{2R} \\ \frac{v_o-v_f}{v_f-v_e} = \frac{R}{2R} \end{cases} \Rightarrow \begin{cases} u_o = \frac{u_e+u_f}{2} \\ v_o = \frac{v_e+v_f}{2} \end{cases} \tag{12}$$

where $R$ is the radius of the spatial circle center. The midpoint of these two centers of ellipse is set as the actual projection of the circle center.

## 5 Experiment and Analysis

### 5.1 Simulation Experiment

The simulation experiment is used for verifying the proposed restore method. Add motion blur to the image by computer firstly, and secondly, get the motion direction and motion scale by the method described in this paper. The motion direction acquired by the Radon transform in frequency domain is displayed in Fig. 7. It is easy to see that the corresponding angle of the maximum value of Radon transform is equal to the motion angle. Use Wiener filtering to restore image later, and the result of the restored process is shown in Fig. 8. The definition of the restored image is fine.




(a) Horizontal (b) 30 degree

**Fig. 7.** Result of identification of motion direction

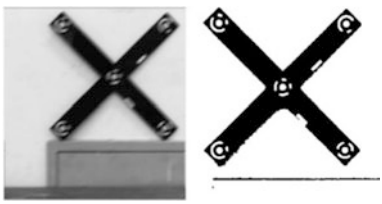


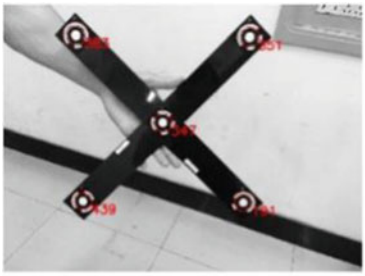
**Fig. 8.** Restored image

### 5.2 Measuring Experiment

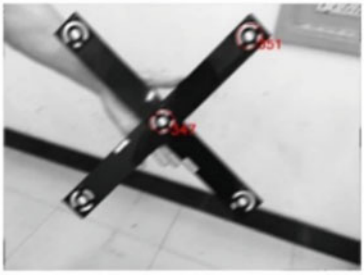
The experiment of measurement did with EoSens® 3CL series high-speed and accuracy camera is made by Mikrotron. The model of camera is MC3010. The resolution of this kind of camera is 1280 × 1024 pixel, the size of pixel is 0.008 mm/pixel, and the model of the lens is Zoom-Nikkor 24–85 mm/1:2.8-4D.

Add the motion blur with 21 pixels and 11 degrees firstly. Restore blurred image by Weiner filtering secondly. Next, decode the coded targets in the original image, blurred image and restored image by algorithm mentioned in this paper. The decoding results

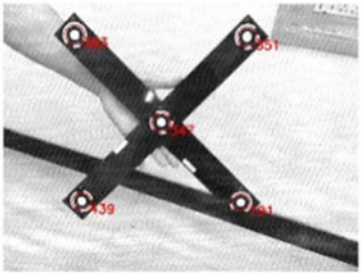
are shown in Fig. 9. The coded targets in blurred image cannot be decoded successfully, and all the coded targets in original image and restored image are recognized by decoding process. Detect circles in original and restored image and locate centrals; the coordinates of centrals are listed in Table 2. The result shows that the difference of central coordinates between original image and restored image is tiny, and these results can verify the accuracy of the circle detection in restored image meets the requirement.



(a) Original image



(b) Motion-blurred image



(c) Restore image

**Fig. 9.** Result of decoding

**Table 2.** Result of the circle detection

| Code number | 191 | 439 | 347 | 351 | 383 |
|---|---|---|---|---|---|
| Original image | 326.649 | 135.723 | 231.105 | 334.179 | 127.189 |
| | 252.545 | 250.802 | 156.576 | 54.1206 | 51.8835 |
| Restored image | 326.651 | 135.721 | 231.083 | 334.186 | 127.241 |
| | 252.523 | 250.822 | 156.581 | 54.0857 | 51.8926 |

## 6 Conclusion

A novel method based on Weiner filtering and Zernike moment is proposed in this paper for detecting the circle on motion-blurred image. The coded target is used to accelerate the matching procedure. To improve the accuracy of central location, restore image by Weiner filtering firstly, and subpixel central location based on high accuracy edge detection by Zernike moment is adopted secondly. Simulation and real experiments are carried out to verify the proposed method. Results show that the accuracy and stability of this method are satisfied. This method is very useful for visual measuring process that extracts information by detecting the circle.

## References

1. Pinto T, Kohler C, Albertazzi A. Regular mesh measurement of large free form surfaces using stereo vision and fringe projection. Opt Lasers Eng. 2012;50(7):910–6.
2. Aguilar JJ, Torres F, Lope MA. Stereo vision for 3D measurement: accuracy analysis, calibration and industrial applications. Measurement. 1996;18(4):193–200.
3. Zhou XG, Nai guang L, Deng WY et al. Image point correspondence using coded targets. J Beijing Inst Mach. 2002.
4. Guo SY, Zhang XF, Zhang F. Adaptive randomized hough transform for circle detection using moving window. In: International conference on machine learning and cybernetics. IEEE; 2006. p. 3880–5.
5. Kim HS, Kim JH, et al. A two-step circle detection algorithm from the intersecting chords. Pattern Recogn Lett. 2001;22(6–7):787–98.
6. Deng ZF, Xiong YL. Identification of motion-blur direction based on frequency-domain algorithm. Opto-Electron Eng. 2007;34(10):98–101.
7. Schneider C-T. 3D Vermessung von Oberflächen und Bauteilen durch Photogrammetrie und Bildverarbeitung. In: Proceedings of the IDENT/VISION'91, May 14–17; Stuttgart, Germany; 1991. p. 90–3.
8. Moghaddam ME, Jamzad M. Finding point spread function of motion blur using Radon transform and modeling the motion length. In: IEEE international symposium on signal processing and information technology. IEEE; 2004. p. 314–17.
9. Zhou Y. Study on Wiener filtering for restoration of motion blurred image. Comput Eng Appl. 2009;45(19):181–3.
10. Chen Q. Identification the scale of the point spread function from the motion blurred image. Comput Eng Appl. 2004;40(23):15–7.
11. Ghosal S, Mehrotra R. Orthogonal moment operators for subpixel edge detection. Pattern Recogn. 1993;26(2):295–306.

# Image Enhancement of Finger Vein Patterns Based on the Guided Filter

Tao Zhan, Hui Ma(✉), and Na Hu

College of Electronic Engineering, Heilongjiang University,
Harbin 150080, China
{574762975,735366516,1965551595}@qq.com

**Abstract.** To solve the problem that image enhancement of finger vein patterns based on traditional filtering methods fails to intuitively highlight the feature of edge protection, the experimental study model based on the guided filter is proposed. Through adding the comparison experiment between guided filter and bilateral filter, and doing the binary processing to the finger vein image after the process of the guided filtering and bilateral filtering, it can be found that some noises exist around the vein texture. In order to reduce or eliminate the interference, a traditional average filtering method is applied for denoising, which not only highlights the vein texture details but eliminates the interference in the post-processing, and at the same time, adjusting the filter parameters will cause a significant impact on the enhancement of finger vein image. A comparison experiment in false recognition rate between two filtering algorithms is conducted, and visual and numerical evaluations are performed on finger vein image after the process of enhancement and binarization; the result indicates that the guided filter has better edge protection feature and lower false recognition rate than the bilateral filter.

**Keywords:** Processing · Guided filter · Finger vein image enhancement · Bilateral filter · Binary processing · Median filter

## 1 Introduction

With the continuous development of information technology and biomedicine, biometrics is being integrated into everyday life such as access control, attendance checking, payment and so on [1], and as we all know, the most common biometrics include face recognition, hand vein recognition, fingerprint identification. However, face recognition and fingerprint identification have some shortcomings. On the one hand, face will change with age, on the other hand, fingerprint can be easily copied; therefore, compared with them, hand vein recognition technology has good advantages, which mainly reflect in its higher universality, independence and high stability. But as a new member of biometric family for finger vein recognition technology, a lot of research results have been made with in-depth research in recent years; in addition, some problems and limitations still exist. First of all, although finger vein recognition has the advantages of high safety and precision, the finger vein cannot be collected in a covert manner which is similar to face; secondly, the quality of finger vein image needs


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_28

to be improved. Low-quality and deformed finger vein images are the main contributors to reduce recognition performance [2]; the main reason is that these low-quality images will affect the extraction and recognition of finger vein features, so enhancing texture details of vein image is one of the important research contents in finger vein recognition technology [3, 4].

Filters are widely used in image processing, pattern recognition, computer vision and other related fields [5], and filter processing is an important process in many image processing algorithms. Among them, the image smoothing margin filter has drawn great attention and has become the basic step of some image processing applications.

At present, the finger vein image enhancement algorithms mainly include the enhancement algorithm in the space and frequency domain, in which the algorithms of contrast enhancement and image texture feature enhancement are frequently used. The former relies mainly on improved histogram equalization technique, while the later on Retinex technique, and the purpose is to enhance the contrast of image, thus making a distinction between region of interest (ROI) and background area. Retaining more texture details of finger vein image is the advantage of this kind of algorithm; however, enhancing the faint texture details of finger vein image is the disadvantage, so the quality of the finger vein image used in the experiment is relatively high, and this kind of algorithm is mainly used to enhance what texture details are clear.

In fact, the methods above for image enhancement only have specific functions, so the universality, robustness and high efficiency of the algorithm await innovation. In order to meet the needs of efficient processing in the era of big data. More scholars shift their research direction to the composite image enhancement technology that can keep the main features of the image and restrain the noise; among them, striking the balance between highlighting the edges and suppressing image noise has become a key issue for research. The guided filter which has wide applications in the field of image processing such as edge enhancement, image fogging and smoothing [6–8] was first proposed by He et al. [9]. At Computer Vision Conference 2010. Tomasi et al. [10] proposed a bilateral filter [11, 12] that takes into account both the neighboring pixel values and the geometrical proximity. Ye et al. [13] proposed a facial expression analysis preprocessing algorithm based on the adaptive guided filter, and it is outstanding in eliminating noise and other aspects. How to highlight the edge-preserving characteristics of the guided filter becomes the problem to be discussed in this paper; through adding the comparison experiment of bilateral filter, the conclusion can be drawn from the visual evaluation and the numerical evaluation. Compared with the bilateral filter algorithm, the experimental results show that the guided filter has a better characteristic of edge protection.

## 2 Image Enhancement of Finger Vein Patterns Based on the Guided Filter

### 2.1 Basic Theory of the Guided Filter

The guided filter is one kind of filter technique that just appeared in recent years, and it is widely applied in the field of image restoration for its characteristic of edge

protection as well as smoothing background. The characteristic of edge protection between guided filter and bilateral filter is their greatest similarity; however, compared with bilateral filter, the guided filter has the advantage of overcoming the problem of gradient inversion and low complexity. In the definition of the guided filter, a local linear model is used, which defines the output image as a local linear model related to the leading image, and uses the leading image to guide the filter process of the image. The filter expression is:

$$Y = \text{guidefilter}(p, I, r, \lambda), \tag{1}$$

In this expression, $r$ is the size of the filter window, $\lambda$ is the regularization parameter, $I$ is the guided image, $p$ is the input image to be filtered, $Y$ is the output image, and both $I$ and $p$ are given. The linear relationship of the guided filter assumed is as follows:

$$Y_i = a_k I_i + b_k, \quad \forall_{i \in \omega_k} \tag{2}$$

$Y$ is the value of the output pixel, $I$ is the value of the input image, $i$ and $k$ are the indices of the pixels, and $a$ and $b$ are the coefficients of the linear function when the center of gravity of the window is located at $k$. In reality, the input image is not necessarily the image to be filtered, which can be also other images such as the guided filter. The following equation can be got by doing the gradient operation on both sides of Eq. (2):

$$\Delta y = a \Delta I, \tag{3}$$

Equation (3) shows that when the input image $I$ has a gradient, the output $y$ has a similar gradient, which explains the characteristics of edge protection of the guided filter.

In the algorithm of the guided filter, the core content is to calculate the optimal solution of the linear coefficient $(a_k, b_k)$, and what is hoped is to make the difference minimum between the output value and the real value, which means the following expression is minimum:

$$E(a_k, b_k) = \sum_{i \in \omega_k} \left( (a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2 \right), \tag{4}$$

In this expression, $p$ must be the image to be filtered, the coefficient before $a$ not only prevent $a$ from being too large but a important parameter for adjusting the filter. Through the least square method, the following expression can be got:

$$a_k = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \varepsilon}, \quad b_k = \bar{p}_k - a_k \mu_k, \tag{5}$$

In the expression, $\mu_k$ is the average of $I$ in the window $\omega_k$, $\sigma_k^2$ is variance, $|\omega|$ is the number of pixels, and $\bar{p}_k$ is the average of the image $p$ to be filtered. It can be found that

a pixel is contained by multiple windows when calculating the linear coefficient of each window; that is to say, each pixel is described by multiple linear functions. Therefore, the output value of a certain point can be obtained by averaging all the linear function values that contain the point, and the output image $y$ can be written as follows:

$$y_i = \frac{1}{|\omega|} \sum_{k:i \in \omega_k} (a_k I_i + b_k) = \bar{a}_i I_i + \bar{b}_i, \tag{6}$$

Due to the local linear characteristic of the guided filter and the limitation of the cost function, the output image gets the details of the leading image while preserving the overall characteristics of the original image. The guided filter is similar to the bilateral filter on the effect of filtering, but the greatest advantage of the guided filter in handing some details is algorithms that the time complexity has nothing to do with the size of the window can be written, so it is much more efficient when handling the image with large window.

### 2.2 Image Enhancement of Finger Vein

For the contrast of the input, the leading image shown in Fig. 1a is relatively low; after the process of the self-guided filter, the result is shown in Fig. 1b, then the image of finger vein is first enhanced by a linear model, and the enhanced image is shown in Fig. 1c. As can be seen from the figure, although the enhanced image achieves the purpose of denoising to some extent, the detail of vein texture has not been well highlighted. At the same time, doing the same operation to the enhanced image second time, the result is shown in Fig. 1d, e; a conclusion that the contrast of the image is clearly improved as well as the texture details are better highlighted can be drawn from the figures.

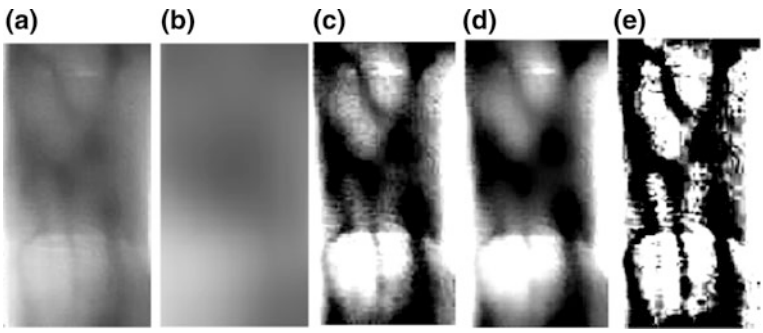



**Fig. 1.** **a** Original image, **b** filtered image, **c** enhanced image, **d** twice filtering image, **e** twice enhancement image

### 2.3 Step of Finger Vein Enhancement Algorithm

Step 1: Take the appropriate filter radius $r$ and regularization parameters $\lambda$ of the input finger vein image, and do the processing of the self-guided filter, whose purpose is to eliminate the noises and get the filtered image $y_1$; the lost texture details of the filtered image are relatively less.

Step 2: For the result in the first step, use the linear enhancement model to handle it, and the model can be written as: $I' = (I - y_1) * 8 + y_1$.

Step 3: Taking the result of step 2 as input image and leading image $I$, then the filtered image $y_2$ can be got by doing the same processing second time; as a result, more texture details are lost.

Step 4: Repeat step 2 with the result of step 3 to obtain the second enhanced image of finger vein.

### 2.4 Image Binarization of Finger Vein

Because the finger vein image still exists in some noises after a series of process, the works of binarization can play a role of distinguishing the texture details from the noise. A method called OTSU is applied to do the work of binarization. The algorithm is derived by using the theory of the least squares based on the gray histogram, and its basic theory is to divide the grayscale value of the image into two parts by the optimal threshold so as to maximize the variance between the two parts, which means the maximum separation.

In general, the process of gray normalization should be done before the binarization processing aiming to distinguish the texture details from the noises, thus leading a better effect of enhancing. The result is shown in Fig. 2a after doing the binarization processing to the enhanced image of finger vein. For some noises still exist around the texture of finger vein, median filter can be used to eliminate the noises, and the result can be seen in Fig. 2b.

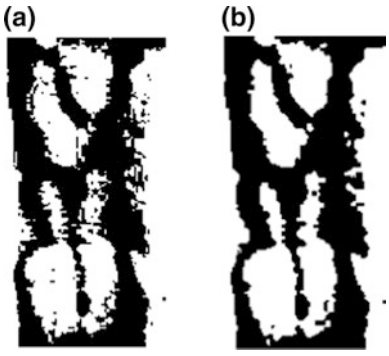



**Fig. 2.** **a** Binarization image and **b** third filtering image

## 3 Comparison Experiment of Finger Vein Image Enhancement Based on Bilateral Filter

The bilateral filter also possesses edge protection feature, which is different from the guided filter and belongs to nonlinear filtering method. In general, in order to extract the edge of the image, the combination of bilateral filter and image edge detection or segmentation is the focus of the research. Therefore, the bilateral filter not only has better edge protection but reduces noises in the aspect of image processing. In this paper, the same method is applied to handle the finger vein image with bilateral filter, and the results of the comparison experiment are shown in Fig. 3.

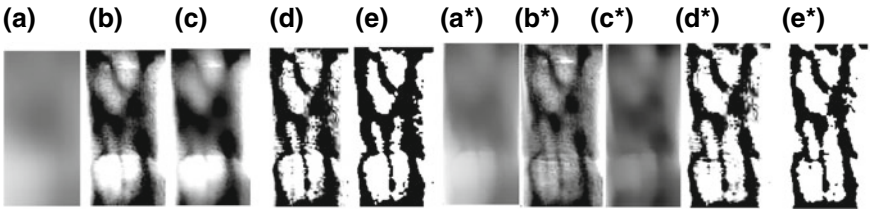



**Fig. 3.** **a**–**e** Represent the results of the guided filtering processing, while **a***–**e*** represent the results of the comparison experiment, which is described as follows: **a**, **a*** filtered image, **b**, **b*** enhanced image, **c**, **c*** twice filtering image, **d**, **d*** binary image, **e**, **e*** thrice filtering image

## 4 Experiment Analysis of Result

### 4.1 Database Selection

In this paper, the database of finger vein image (FV database) selected from Tianjin Key Laboratory of Intelligent Signal and Image Processing includes 64 samples, and 15 finger vein images are contained in every sample; at the same time, all the finger images have been accurately processed in the extraction of region of interest (ROI).The size of the image is set to 170 pixel $\times$ 91 pixel.

### 4.2 Filter Parameter Selection

The parameter selection of the filter has the decisive influence on the image enhancement. Therefore, for the finger vein database used in this experiment, different guided filter parameters are chosen for the comparison experiment and chose the optimal parameter through the numerical evaluation.

The full name of SSIM is structural similarity index. It is a measure of evaluating the similarity of two images. It was first proposed by the Laboratory for Image and Video Engineering at the University of Texas located in Austin. If one image is pre-compressed and compressed, the SSIM algorithm can be used to evaluate the quality of the compressed image (Fig. 4).

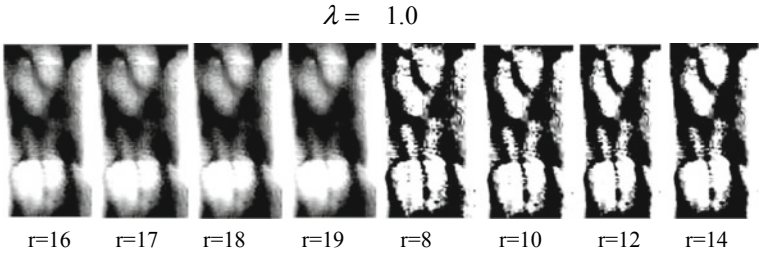



**Fig. 4.** Enhanced image at a different radius and twice enhancement image at a different radius

$$\lambda = 0.1$$

A conclusion that filter $r = 16$ is the first filter optimal parameter while $r = 8$ is the second can be drawn from Tables 1 and 2, and regularization parameter $\lambda = 0.1$ is the best in this experiment.

**Table 1.** SSIM value of the first enhanced image at different radius

| $r$ | SSIM |
|---|---|
| $r$ = 16 | 0.9942 |
| $r$ = 17 | 0.9937 |
| $r$ = 18 | 0.9933 |
| $r$ = 19 | 0.9928 |

**Table 2.** SSIM value of the twice enhanced image at different radii

| $r$ | SSIM |
|---|---|
| $r$ = 8 | 0.9457 |
| $r$ = 10 | 0.9389 |
| $r$ = 12 | 0.9345 |
| $r$ = 14 | 0.9316 |

### 4.3 Image Quality Evaluation

Contrast ($C$): Calculating the standard deviation as the contrast parameter.
Definition ($D$): Calculating the average gradient of all the pixels as the definition parameter.

A conclusion that the guided filter algorithm is better than the bilateral filter algorithm can be drawn from Table 3.

**Table 3.** Comparison between two filtering algorithms in C, D and SSIM

| Evaluation index | Guided filter | Evaluation index | Bilateral filter |
|---|---|---|---|
| *C* | 78.4997 | *C* | 78.4882 |
| *D* | 49.0138 | *D* | 49.0079 |
| SSIM | 0.9921 | SSIM | 0.9912 |

### 4.4 False Recognition Rate Evaluation

It can be seen from Table 4 that the finger vein image processed by guide filtering in the case of the same template radius has a significantly lower false recognition rate than the finger vein image processed by bilateral filtering (Table 4).

**Table 4.** Comparison between two filtering algorithms in false recognition experiment

| Guided filter | False recognition rate (%) | Bilateral filter | False recognition rate (%) |
|---|---|---|---|
| $r = 16$ | 0.4687 | $r = 16$ | 0.7812 |
| $r = 17$ | 0.6250 | $r = 17$ | 0.7812 |
| $r = 18$ | 0.6250 | $r = 18$ | 0.9375 |
| $r = 19$ | 0.7812 | $r = 19$ | 1.0937 |

## 5 Conclusion

To solve the problem that image enhancement of finger vein patterns based on traditional filtering methods fails to intuitively highlight the features of protecting the edge, the experimental study model based on the guided filter is proposed without changing its linear complexity. Filter parameters are adjusted several times in the experiment in order to obtain the best experimental results. Through adding the comparison experiment of bilateral filter, visual and numerical evaluations are performed on finger vein image after the process of enhancement and binarization and a comparison between two filtering algorithms in false recognition experiment is conducted to evaluate which one is even better; the result indicates that the guided filter has better edge protection feature and lower false recognition rate than the bilateral filter. The next work is to optimize the parameters in the filtering algorithm so as to effectively improve the effect of the filtering algorithm on finger vein image enhancement.

## References

1. Cao W, Wang HB, Shi J, Yu R, Tao L. Enhancement algorithm of finger vein image based on weighted guided filter with edge detection. Laser Optoelectron Prog. 2017;54(02): 172–80.
2. Qiu JH, Xu W, Wang YF. Study of finger vein recognition application. Study Inf Secur. 2016;2(01):86–92.

3. You L, Li H, Wang JW. Finger-vein recognition algorithm based on potential energy theory. In: Proceedings of the 2015 IEEE 16th international conference on industrial technology; 2016. p. 742–5.
4. Yang WM, Qin C, Wang XJ et al. Cross section binary coding for fusion of finger vein and finger dorsal texture. In: Proceedings of 2016 IEEE international conference on industrial technology; 2016. p. 742–5.
5. Wu K, Han GL, Wang YQ, Wu XT. Multi-scale guided filter and its application in image dehazing. Opt Precis Eng. 2017;25:2182–94.
6. Kou F, Chen W, Li Z, et al. Content adaptive image detail enhancement. IEEE Sig Process Lett. 2015;22(2):211–5.
7. Liu W, Cui YF, Wu XL. Image defogging algorithm based on constrained evolutionary time-frequency weighted filtering. Comput Sci. 2014;41(09):311–4.
8. He KM, Sun J, Tang XO. Single image haze removal using dark channel prior. IEEE Trans Pattern Anal Mach Intell. 2009;33(12):2341–53.
9. He KM, Sun J, Tang XO. Guided image filtering. IEEE Trans Pattern Anal Mach Intell. 2013;35(6):1397–409.
10. Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Proceedings of the 6th international conference on computer vision; 1998. p. 839–46.
11. Kapik B, Yannoulis NC, Shams N, Unoprostone Monotherapy Study Group—eu. Fast "0(1)" bilateral filtering. Miami: CVPR 2009 IEEE conference in computer vision and pattern recognition, 2009.
12. Sun L et al. Adaptive bilateral filter considering local characteristics. In: Hefei: 2011 six international conference in image and graphics (ICIG); 2011. p. 187–92.
13. Ye C, Yong JZ, Kamen Lvanov et al. Pre-processing for muscle motion analysis: adaptive guided image filtering for speckle reduction of ultrasound images. In: Osaka Japan: 35th annual international conference of the IEEE EMBS; 2011. p. 4026–4029.

# Image Segmentation of LBF Model with Variable Coefficient Regularized Area Term

Liyan Wang(✉), Jing Liu, and Yulei Jiang

Department of Mathematics, Dalian Maritime University, Dalian 116026, China
wangliyan@dlmu.edu.cn

**Abstract.** In this paper, an improved LBF model based on local regional information is proposed for image segmentation. The basic idea is to add the regularized area term to the energy function of the LBF model and establish a variable coefficient with adaptive capability composed of image regional information. Compared with the LBF model, the proposed model increases the driving force of the evolution curve, making the result better when dealing with the images with weak boundaries and intensity inhomogeneity. At the same time, it effectively solves the problem that the LBF model is sensitive to the initial position and size of the evolution curve. This model is used to segment medical images with complex topological structure and intensity inhomogeneity. Experimental results show that regardless of the initial curve of any position or any size, it has little influence on the segmentation result; moreover, the localization of deep-depressed image boundaries is more accurate, so we get the conclusion that the new model has corresponding improvements in segmentation accuracy and robustness.

**Keywords:** LBF model · Medical image segmentation · Intensity inhomogeneity · Level-set method · Model adaptability

## 1 Introduction

Image segmentation is a very important part of image processing, and its purpose is to extract the desired parts from the image for subsequent research. Medical image segmentation is an essential step in diagnosing tumors, guiding operations and visualizing medical data, which provides a reliable basis for clinical medical diagnosis and pathological research. However, medical images have the characteristics of intensity inhomogeneity, deep depression and topological complexity, which put forward higher requirements for image segmentation technology.

The geometric active contour models based on partial differential equation [1–11], which use curve evolution theory and level-set method to segment images, have been widely used in medical image segmentation, they have the advantages of dealing with the splitting or merging of contour flexibly and the stability of numerical solution, and


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_29

they are usually divided into two categories: edge-based and region-based models. The edge-based models [1, 4, 7–10] mainly use the edge information to drive the contour toward object boundaries, for example, GAC model [1]. These models have good segmentation effect for strong boundary and low-noise images, but an ideal effect cannot be obtained when a medical image with intensity inhomogeneity is segmented. The region-based models [5, 6, 11] utilize the regional statistical information to detect objects, and they have good segmentation results for the images with weak edges and intensity inhomogeneity, for example, the well-known CV [11] model and LBF [6] model belong to this kind. The CV model uses the global information to segment images and has certain robustness to noise, but it cannot achieve the best effect when it is used to segment medical images. In order to make up for the defects of the CV model, Li et al. put forward the LBF model. The LBF model has better performance on dealing with intensity inhomogeneous images, which is based on the local area information of the image to drive curve evolution, and it does not affect the determination of the target boundary even if the image boundary is blurred or broken because it doesn't have to make use of the image edge information can complete segmentation. However, the LBF model also has some shortcomings; for example, the evolution curve depends on the selection of the initial position and the size of the initial curve.

When the LBF model is used to segment medical images with intensity inhomogeneity and deep depression, due to the LBF model is sensitive to the initial contour, so the initialization process is extremely important, sometimes only relying on the local fitting term to segment the image is prone to boundary leakage or over-segmentation, failing to achieve an ideal segmentation effect. Besides on, we consider adding the area term to the energy function of the LBF model and defining the coefficient of the area term as a function that integrates the image regional information. Compared with the LBF, new model increases the driving force during evolution, which makes the model better when dealing with medical images of intensity inhomogeneities and deep depression. At the same time, the variable coefficient $S_n(I)$ can automatically change the size and symbol according to the image information and then change the direction and speed of the curve, and the role of the area item also changes with it. The inward or outward motion is automatically selected at different positions so that the initial curve can be flexibly placed, greatly reducing the restriction on the initialization of the model.

## 2 LBF Model

LBF model is a kind of variable region geometric active contour model, which is mainly composed of fitting energy items, and it is shown in literature [6]. Specifically, for each point in the image, the energy function that represented by level-set formation is defined in the following way

$$E^{\mathrm{LBF}}(\phi, f_1(x), f_2(x)) = \lambda_1 \int \left[ \int K_\sigma(x-y)|I(y) - f_1(x)|^2 H(\phi(y))\mathrm{d}y \right] \mathrm{d}x + \lambda_2 \int \left[ \int K_\sigma(x-y)|I(y) - f_2(x)|^2 (1 - H(\phi(y)))\mathrm{d}y \right] \mathrm{d}x + \mu \int_\Omega \delta_\varepsilon(\phi)|\nabla\phi|\mathrm{d}x\mathrm{d}y + v \int_\Omega \frac{1}{2}(|\nabla\phi| - 1)^2 \mathrm{d}x\mathrm{d}y \quad (1)$$

where $I$ is an input image, $\phi$ is a level-set function, $\lambda_1, \lambda_2, \mu, v$ are positive constants, $K_\sigma(x)$ is a Gaussian kernel function, $f_1(x)$ and $f_2(x)$ are two intensity fitting values in the local small neighborhood of pixel $x$ (the size of the neighborhood depends on the scale $\sigma$ of the Gaussian function), and $H(x)$ and $\delta(x)$ represent the Heaviside function and the Dirac function, respectively.

For a fixed level-set function $\phi$, we minimize the function $E^{\mathrm{LBF}}(\phi, f_1, f_2)$ in (1) with respect to the functions $f_1$ and $f_2$ to get the expression as follows

$$f_1(x) = \frac{K_\sigma(x) * [H_\varepsilon(\phi(x))I(x)]}{K_\sigma(x) * H_\varepsilon(\phi(x))}, f_2(x) = \frac{K_\sigma(x) * [(1 - H_\varepsilon(\phi(x)))I(x)]}{K_\sigma(x) * (1 - H_\varepsilon(\phi(x)))} \quad (2)$$

The first two terms in the above Eq. (1) represent local fitting terms, which enable the evolution curve to reach the target edge better; the third term is the length term, which ensures the smoothness of the curve and the occurrence of isolated area items as little as possible; and the fourth term is the distance regularization term to avoid the reinitialization steps of the level-set function during evolution. The LBF model introduces the Gaussian function $K_\sigma(x)$ as the weight of the fitting term on the basis of the CV model, to control the range of the fitting function. Therefore, the selection of the initial curve position and size has a great influence on the segmentation result and the segmentation time.

## 3 Elaboration of the Model

Due to there is no regularized area term in the LBF model, when the image edge is blurred or discontinuous, the LBF model only depends on the length term and the local fitting term to segment the image, so that the evolution curve may not stay at the real boundary of the object, resulting in boundary leakage or excessive segmentation.

We consider adding the area item to the LBF model, and it can be represented by the level-set function as follows

$$A(\phi) = \int_\Omega H(\phi(x, y))\mathrm{d}x\mathrm{d}y \quad (3)$$

In order to enhance the adaptability of the model, avoid the constant area factor defines the role of the area term and affects the final segmentation result, while improving the sensitivity of the model to the initial position, we define a new area term coefficient as follows

$$S_n(I) = \begin{cases} \frac{m}{I_0^p + 1}, & \nabla g \cdot N < 0 \\ \frac{-m}{I_0^p + 1}, & \nabla g \cdot N > 0 \end{cases} \tag{4}$$

where $m, p > 0$ are fixed constants, $\nabla g$ is the gradient of the stop function $g(x)$, $N$ represents inner normal vector and it always points to the interior of the zero-level-set curve, and $I_0 = |\nabla(G_\sigma * I)|$ is the Gaussian filtered gradient of the image. In order to reduce the effect of noise on the segmentation result, we chose the following stop function as follows

$$g = \frac{\alpha}{1 + |\nabla(G_\delta * I)|^\beta} \tag{5}$$

where $\alpha, \beta$ are positive constants. $\alpha$ is used to control the evolution speed and $\beta$ is used to control the sensitivity of the evolution curve to noise, the bigger $\alpha$ is, the smaller $\beta$ is, and the faster the movement speed is. Generally, $\alpha = 1, \beta = 1$ is used, and different values can be selected according to the characteristics of the image.

We know that the direction of the gradient is always the direction where the function changes the fastest, so the direction of $\nabla g$ always points to the direction of $g$ increasing and then $\nabla g$ always points to the direction of deviation from the edge regardless of whether the initial curve is inside or outside the target object. It is assumed that when the contour curve moves outside the target boundary, $N$ and $\nabla g$ are in the opposite direction, i.e., $\nabla g \cdot N$ is negative, so $S_n(I) > 0$; if the active contour is inside the boundary, then the direction of $N$ and $\nabla g$ is the same, i.e., $\nabla g \cdot N$ is positive, so $S_n(I) < 0$. Therefore, in our proposed method, the dependence of the curve on the initial position is overcome, the initial curve can be defined at any position, and the image with deep sag or discontinuity can also be segmented. In the area where the image gray value changes little, the gradient modulus $I_0$ of the image is approximately zero, and then the value of $S_n(I)$ is $m$ or $-m$, and the evolution curve can continue to move until the real boundary of the target area stops evolving. As can be seen from the above definition of (4), there is an inverse relationship between $I_0$ and $S_n(I)$, in the part where image gray value is inhomogeneous, $I_0$ is small and $S_n(I)$ is large, which can accelerate the evolution of the zero level assembly to the target boundary; when the zero-level set is near the target boundary, the value of $S_n(I)$ is smaller and evolution is slower, so the zero-level set will slowly move to the target boundary without creating a boundary leak.

In summary, the improved energy function of LBF model can be written as

$$\begin{aligned} E_{\text{all}}^{\text{LBF}}(\phi, f_1, f_2) &= \lambda_1 \int \left[ \int K_\sigma(x-y)|I(y)-f_1(x)|^2 H(\phi(y))\mathrm{d}y \right] \mathrm{d}x \\ &+ \lambda_2 \int \left[ \int K_\sigma(x-y)|I(y)-f_2(x)|^2 (1-H(\phi(y)))\mathrm{d}y \right] \mathrm{d}x \\ &+ \mu \int_\Omega \delta_\varepsilon(\phi)|\nabla\phi|\mathrm{d}x\mathrm{d}y + \nu \int_\Omega \frac{1}{2}(|\nabla\phi|-1)^2 \mathrm{d}x\mathrm{d}y \\ &+ S_n(I) \int_\Omega H(\phi(x,y))\mathrm{d}x\mathrm{d}y \end{aligned} \tag{6}$$

Using the variational method and gradient descent method to minimize the energy function $E_{\text{all}}^{\text{LBF}}$, we obtain the partial differential equation satisfying the final level-set evolution as follows

$$\begin{aligned} \frac{\partial\phi}{\partial t} &= -\delta_\varepsilon(\phi)(\lambda_1 e_1 - \lambda_2 e_2) + \mu\delta_\varepsilon(\phi)\mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) \\ &+ \nu\left(\nabla^2\phi - \mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right) - S_n(I)\delta_\varepsilon(\phi) \end{aligned} \tag{7}$$

Heaviside function is approximated by $H_\varepsilon(x)$ in the form of smooth regularization, and the Dirac function is replaced by the derivative $\delta_\varepsilon(x)$ of $H_\varepsilon(x)$, the parameter $\varepsilon$ controls the rate of change of the function, and their functional expressions are as follows [6]

$$H_\varepsilon(x) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\left(\frac{x}{\varepsilon}\right)\right), \ \delta_\varepsilon(x) = H'_\varepsilon(x) = \frac{1}{\pi}\frac{\varepsilon}{\varepsilon^2 + x^2} \tag{8}$$

$e_1$ and $e_2$ are the functions as below

$$e_1 = \int_\Omega K_\sigma(y-x)|I(x)-f_1(y)|^2\mathrm{d}y, \ e_2 = \int_\Omega K_\sigma(y-x)|I(x)-f_2(y)|^2\mathrm{d}y \tag{9}$$

## 4 Comparative Experiment and Result Analysis

The new model, which we call it the ALBF model, has been applied on the following set of experiments on medical images to fully verify its feasibility. In our experiment, the images we selected were all from the clinical images of patients in a hospital. The experimental parameters were as follows

The proposed model: $\lambda_1 = 1$, $\lambda_2 = 2$, $\mu = 0.003 \times 255 \times 255$, $v = 1$, $\sigma = 3$, $\varepsilon = 1$, time step $\tau = 1$, $\alpha = 1$, $\beta = 2.8$, $m = 200$, $p = 2$, and iteration number: fifty times.

LBF model: $\lambda_1 = 1$, $\lambda_2 = 2$, $\mu = 0.003 \times 255 \times 255$, $v = 1$, $\sigma = 3$, $\varepsilon = 1$, time steps $\tau = 1$, and iterations: fifty times.

Figure 1 is the segmentation results of two models with different initial curve sizes. We can see the results of the 50 iterations of the proposed model obtained with either the large or small contours are better than those of the LBF model, and the results obtained by the proposed model for the three contours are approximately the same. It can be seen that the proposed model is less sensitive to the initial contour size.

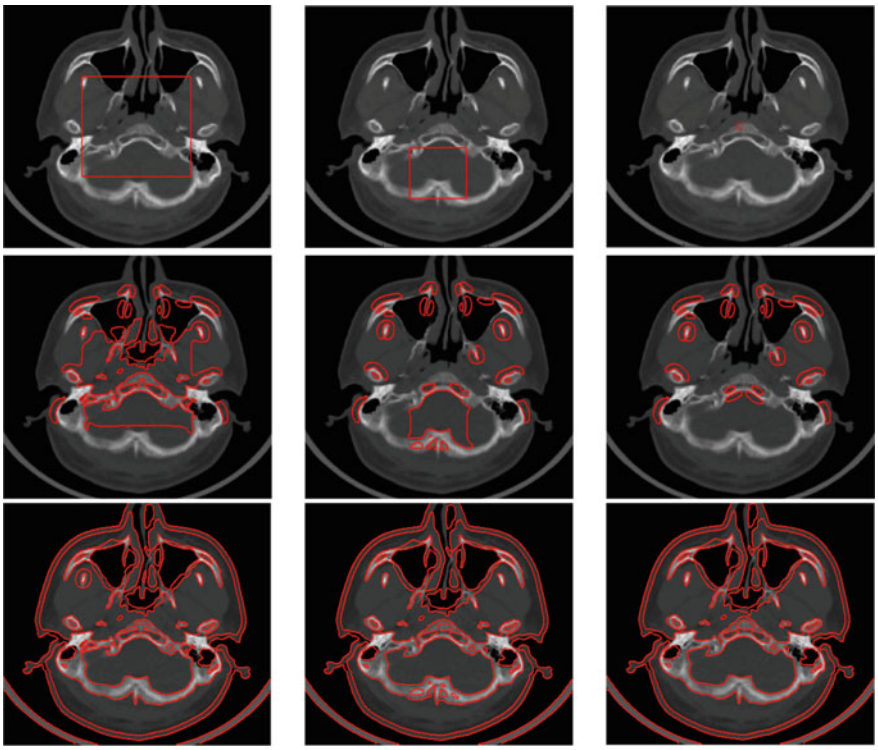


**Fig. 1.** Comparisons of the segmentation results between the proposed model and the LBF model for head CT of $512 \times 512$. The first row: initial contour. The second row: the segmentation results of the LBF model. The third row: the segmentation results of the proposed model

Figure 2 shows the segmentation results of the two models when selecting different positions of the initial curve. It can be seen from Fig. 2 that the 50 iteration results obtained by using the contours of different positions in proposed model are better than those obtained by the LBF model. It can be concluded that the position of the initial curve has little effect on proposed model.

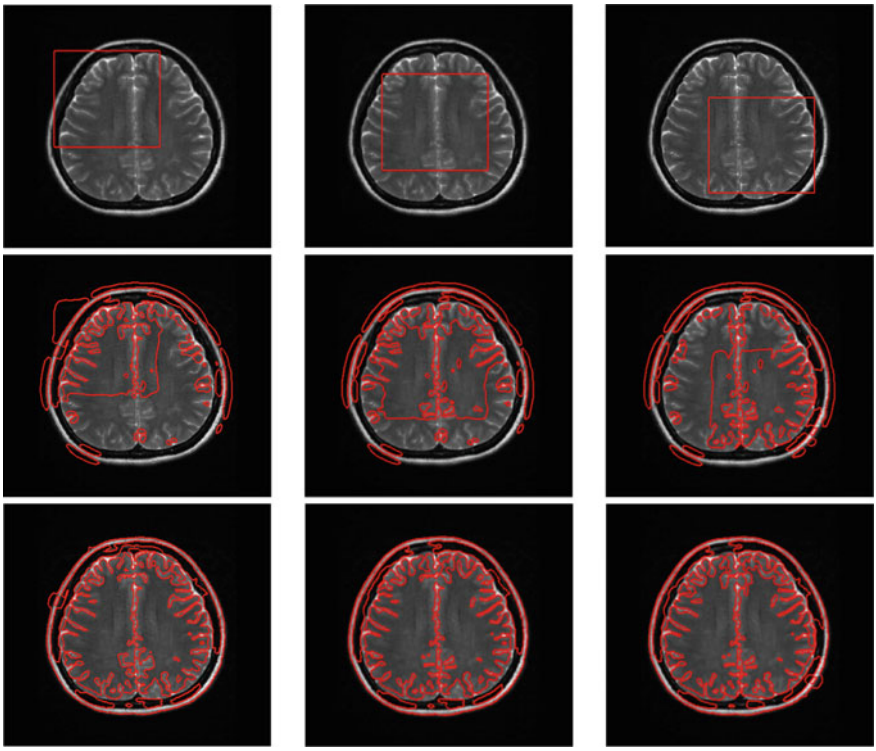


**Fig. 2.** Comparisons of the segmentation results between the proposed model and the LBF model for head MRI of 512 × 512. The first row: initial contour. The second row: the segmentation results of the LBF model. The third row: the segmentation results of the proposed model

Figure 3 shows that the proposed model and LBF model for intensity inhomogeneity and depression of medical image processing results of 50 iterations. We can get the conclusion that the results of the new model in this paper are better than those of the LBF model for areas with uneven grayscale and pitted, so the model is effective.

## 5 Summary

In this paper, the energy function of the LBF model is improved. The area term is added to the original LBF energy function, and the area coefficient associated with the image information is defined at the same time, which effectively solves the problem that the LBF model is sensitive to the position and size of the initial contour and cannot adaptively determine the direction of movement. This allows the model to define an initial contour of arbitrary size and shape at any position, which enables better

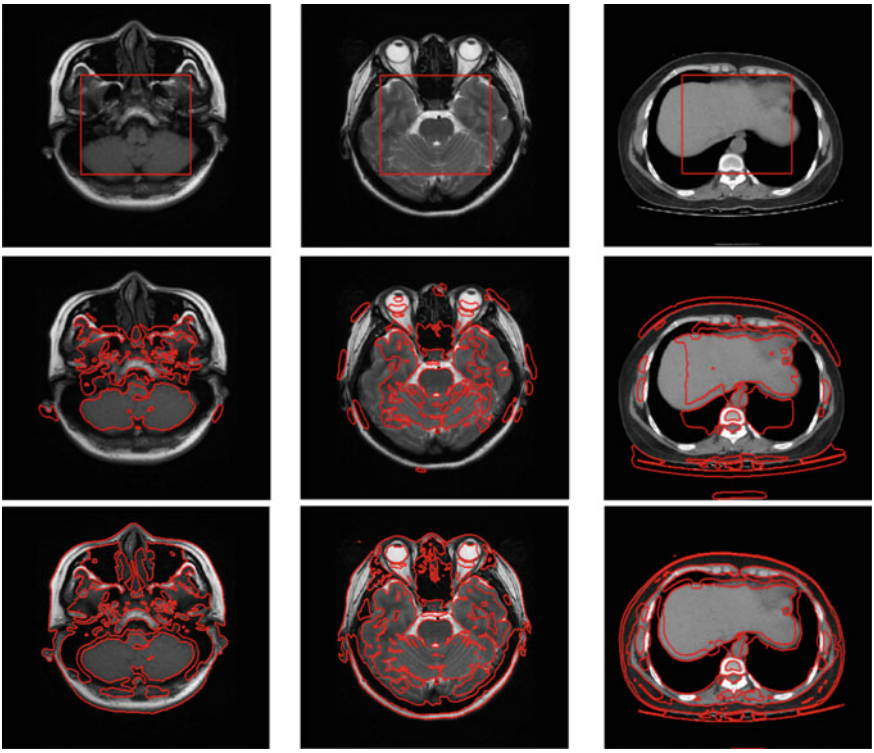


**Fig. 3.** Results of different models for medical image segmentation. The first row: initial contour. The second row: the segmentation results of the LBF model. The third row: the segmentation results of the proposed model

segmentation of depth-depressed and intensity inhomogeneous images, such as more complex medical images, while improving segmentation accuracy and facilitating practical application.

## References

1. Caselles V, Kimmel R, Sapiro G. Geodesic active contours. Int J Comput Vis. 1997;22 (1):61–79.
2. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. IEEE Trans Image Process. 1998;7(3):359–69.
3. Li C, Xu C, Gui C, Fox MD. Level set evolution without re-initialization: a new variational formulation. In: IEEE computer society conference on computer vision and pattern recognition, vol 1; 2005. p. 430–6.
4. Li C, Xu C, Gui C, Fox MD. Distance regularized level set evolution and its application to image segmentation. IEEE Trans Image Process. 2010;19(12):3243–54.
5. Li C, Kao CY, Gore JC, Ding Z. Minimization of region-scalable fitting energy for image segmentation. IEEE Trans Image Process. 2008;17(10):1940–9.

6. Li C, Kao CY, Gore JC, Ding Z. Implicit active contours driven by local binary fitting energy. In: IEEE conference on computer vision and pattern recognition; 2007. p. 1–7.
7. Kass M, Witkin A, Terzopolos D. Snakes: active contour models. Int J Comput Vis. 1988;1 (4):321–31.
8. Malladi R, Sethian JA, Vemuri BC. Shape modeling with front propagation: a level set approach. IEEE Trans Pattern Anal Mach Intell. 1995;17(2):158–75.
9. Gao S, Bui TD. Image segmentation and selective smoothing by using mumford-shah model. IEEE Trans Image Process. 2005;14(10):1537–49.
10. Sedlák J, Morávek J. Optimal approximations by piecewise smooth functions and associated variational problems. Commun Pure Appl Math. 2010;42(5):577–685.
11. Chan TF, Vese LA. Active contours without edges. IEEE Trans Image Process. 2001;10 (2):266–77.

# Research on Image Super-Resolution Reconstruction of Optical Image

Aiping Jiang(✉), Xinwei Li, and Han Gao

Heilongjiang University, Harbin, China
jap62@126.com, 863301127@qq.com, ghl9951230@163.com

**Abstract.** Currently, the image super-resolution reconstruction method based on sparse representation has limited ability to process the details of the edge. Therefore, based on the dictionary learning, the local variance feature edge gradient estimation image fast super-resolution reconstruction is improved and optimized based on dictionary training. The dictionary training process includes cluster analysis of high-resolution images, local variance extraction, and sparse filtering. The reconstruction process includes local variance detection of the low-resolution image and threshold judgment, and then the image is reconstructed according to the gradient value.

**Keywords:** Sparse representation · Super-resolution reconstruction · Dictionary construction · Gradient estimation

## 1 Introduction

Digital images, as the medium for the transmission of information on the Internet, are always around us and play an important role. Digital image processing technology is generally used in military reconnaissance, medical devices, satellite remote sensing, face recognition, and monitoring and plays an important role [1, 2]. Since images are often disturbed by the device during acquisition and transmission, high-resolution images are converted to low-resolution images. Therefore, if you want to obtain high-resolution information in low-resolution images, you must think of ways to improve the resolution of the image through certain technical means.

The document reconstructs high-resolution images using prior constraints of input low-resolution image blocks (LR) and high-resolution (HR) image blocks that have the same sparse representation coefficients under a specific dictionary pair get a better reconstruction effect [3]. A paper proposes using a Gaussian mixture model clustering algorithm to cluster training image blocks and then uses improved K-singular value decomposition (K-SVD) algorithm to quickly obtain high- and low-resolution dictionary pairs, thus improving the problem of poor super-resolution reconstruction of images based on dictionary learning [4]. The article introduced non-local self-similarity and super-Laplacian prior as regularization constraints on sparse dictionary reconstruction models, thereby enhancing the effectiveness of the sparse model [5]. In this paper, double sparse dictionary learning and local variance edge gradient estimation are used to rapidly reconstruct the image.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_30

## 2 Algorithm

### 2.1 Dictionary Optimization

#### 2.1.1 K-SVD Optimization

Taking into account that the original Yang dictionary algorithm does not study the high-resolution image will make unwanted noise during the presentation process and does not classify the image, so start from this aspect and further dictionary optimization. K-SVD algorithm is a dictionary construction algorithm. It uses multiple images to train high-resolution dictionary and low-resolution dictionary. From the perspective of the matrix, the purpose of dictionary learning is to decompose the input $Y$ matrix into the dictionary $D$ and sparse matrix $X$; the formula is:

$$Y \approx D * X \tag{1}$$

$X$ is called a coefficient matrix and needs to be sparse as much as possible. $D$ is the dictionary that is trained, and a normalized vector is a matrix.

Obtain a high-resolution combination $Y^{\mathrm{h}} = \{y_1, y_2 \ldots y_n\}$ of satellite optical images, where $k$ samples are chosen as the atoms of $D$, and the sparse coefficient matrix is set as the 0 matrices. Fix the dictionary, find the sparse coding of each sample, assume the dictionary $D = [a_1, a_2, a_3, a_4]$, and calculate the sparse coding $x$ of $Y$.

First, find the vector $a$ that has the largest distance from the vector $y$ from $a_1$, $a_2$, $a_3$, and $a_4$, that is, $a_1 * y$, $a_2 * y$, $a_3 * y$, and $a_4 * y$. Assuming $a_2 * y$ is maximal, we choose $a_2$ as the first atom, and then the first encoding vector is $x_1 = (0, b, 0, 0)$, where $b$ is the unknown number. The solution coefficient, $y - b * a_2 = 0$.

Then use $x_1$ and $a_2$ as multiplication to reconstruct the data and calculate the residual value, $y' = y - b * a_2$. If the residual is less than a given $\varepsilon$, it will end the iterative loop; otherwise, it will continue to iterate. Then continue to perform a residual approximation of the atomic vectors in the rest of the dictionaries. Update the residual value. Fix sparse coefficients and update the dictionary. The essence of dictionary learning is to find the optimal solution from the dictionary. First of all, from the mathematical model, the sparse decomposition algorithm refers to the selection $k$ columns of atom in the trained dictionary, and the $k$ column atom is used to solve the decomposed image signal $y$ by linear combination:

$$y = A\theta \tag{2}$$

In the sparse representation of Eq. (2), $y$ denotes the sparse decomposition signal, $\theta$ denotes the column vector of the sparse $N * 1$ to be decomposed, $A$ denotes a redundant dictionary, and $M * N$ denotes a vector.

The dictionary construction process in this paper first performs Gaussian filtering on multiple high-resolution images and then divides them into two types of images. One type of image is a high-resolution image related to the input, and the other type is an image that has nothing to do with the high resolution. Then denoise the image and save it to the image library. The HR image is processed into an LR image of a degraded

model. After the HR image is filtered, the LR and HR images are segmented. Finally, the high and low dictionary pairs are generated by the optimized K-SVD algorithm.

### 2.1.2 Feature Extraction

For natural images, high-frequency image information has a large impact on human vision, and humans easily obtain high-frequency information. Therefore, it is of great research significance to use low-resolution high-frequency information generated by an algorithm. Several one-dimensional filters, namely

$$\begin{aligned} f_1 &= [-1, 0, 1]\ f_2 = f_1^T \\ f_3 &= [1, 0, -2, 0, 1]\ f_4 = f_3^T \end{aligned} \tag{3}$$

In Eq. (3), $f_1, f_2, f_3$, and $f_4$ represent four different filters that are used to extract the four feature vectors of the image so that the vectors are in close contact with each other. The K-SVD dictionary optimization algorithm training based on feature dictionary is as follows:

(1) Acquiring a high-resolution combination of satellite optical images $Y^{\mathrm{h}} = \{y_1, y_2 \ldots y_n\}$, $f_1$, $f_2$, $f_3$, and $f_4$ as feature extraction vectors, the goal is to output high- and low-resolution training dictionaries $D_{\mathrm{h}}$ and $D_{\mathrm{l}}$.
(2) Randomize the matrix, initialize the Gaussian matrix with high and low dictionaries, and then downsample the high-resolution set $Y^{\mathrm{h}}$ by $k$ times to make the set of low-resolution images, $M = 0$.
(3) $k$-magnification of the generated low-resolution image produces a new training set and performs low-resolution to high-resolution reconstruction of the set. The new set is labeled $X^{\mathrm{l}} = \{x_1, x_2 \ldots x_n\}$.
(4) To update the dictionary, extract the low-resolution information and use the obtained vector to do a convolution calculation with the new set $X^{\mathrm{l}}$ to obtain a low-resolution image feature block.
(5) Fix the dictionary $D$ and update the sparse coefficient $\alpha$ by iterating. In the iterative process, the final result is mainly made to approach the defined variable $\delta$, every time $M = M + 1$, and the condition is satisfied.

## 3 Cluster-Based Dictionary Learning Method

In this paper, the variance is used to perform feature estimation when the variance is greater than the threshold value. If the condition is not satisfied, Yang's method is used for reconstruction. In the image training process, the image blocks need to be averaged.

$$K = \sum_{i=1}^{k} \sum_{j \in L} (x_j - \mu_i)^2 \tag{4}$$

In formula (4), $\mu_i$ represents the selected pixel, $x_j$ represents all the pixels in the sample, $K$ represents the new value after iteration, and the value of the pixel in the

image is randomly selected from the $N$ samples as the centroid. Continue loop iterations until the new centroid is equal to the original centroid or less than the given value to stop the loop.

Since the variance of each image block is calculated, each image block has a large error. Therefore, it is necessary to extract the pixels that represent the most global information in each image block and perform local variance calculation. The variance is set to 10 in this paper. When the threshold value is greater than the variance, the image will be compressed and sensed by the gradient estimation reconstruction when it is less than the variance. At the threshold, bicubic interpolation is performed on the image block. The local variance of the image is calculated as follows.

$$m_x(i,j) = \frac{1}{(2n+1)^2} \sum_{k=i-n}^{i+n} \sum_{j-n}^{j+n} x(k,l) \tag{5}$$

$$\delta_2(i,j) = \frac{1}{(2n+1)^2} \sum_{k=i-n}^{i+n} \sum_{l=j-n}^{j+n} [x(k,l) - m_x(i,j)]^2 \tag{6}$$

In Eqs. (5) and (6), $x$ represents a high-resolution block, $m$ and $n$ represent the window size of the image block, $\delta$ represents a local variance, and the known image block information can solve the local variance, calculate, and continue to do the next step.

The method used in this paper is a morphological gradient method. First, input a low-resolution image and calculate the variance after cubic interpolation, and calculate which group of image blocks the variance is similar to. Then find the image block and calculate the input low-resolution image through the sparse filter matrix sparse features, namely

$$K = WY \tag{7}$$

$W$ is the sparsely filtered matrix of the corresponding image block, and $Y$ represents the mean value of the image. If there are $N$ similar sample blocks in the image block, the selected sample image block is returned by the coefficients calculated by the formula (6). As follows:

$$a = \min \|K - F_{\mathrm{M}} a\|_2^2 \tag{8}$$

From Eqn. (8), we can see that $F_{\mathrm{M}} = [K_1, K_2, \ldots K_N]$ represents a feature matrix composed of multiple samples and $a$ represents a regression coefficient. Then, the resolution of the input corresponding to the high-resolution image block gradient will be high resolution according to the image block subset in the training library. The coefficient regression and gradient of the image block are used to perform the estimation calculation.

In the sample, the final estimated value is a linear superposition of the gradient field samples and a morphological estimate of the low-resolution, high-resolution gradient. Calculate the reconstruction formula as follows:

$$X^{1+1} = X^t - \varepsilon \frac{\partial X}{\partial t} = X^t + \varepsilon \cdot \frac{\partial E(X|Y, \nabla X)}{\partial X} \tag{9}$$

$$\frac{\partial X}{\partial t} = -\frac{\partial E(X|Y, \nabla X)}{\partial X} = -\frac{\partial E(X|Y)}{\partial X} \tag{10}$$

From Eqs. (9) and (10), $X$ indicates a high-resolution image, $Y$ indicates a low-resolution image, $\nabla X$ indicates a gradient field of an estimated high-resolution image, and $E(X|Y)$ indicates a local variance estimation based on dictionary learning.

## 4 Simulation Experiment Analysis

This article compares the PSNR value of the Yang algorithm with the PSNR value of the algorithm. It is proved that the PSNR values of the reconstructed image of this algorithm are higher than those of the original Yang algorithm.

Tables 1 and 2 can be intuitively drawn. Compared with the algorithm of this paper, the algorithms in bilinear interpolation, cubic difference, and Yang are similar in the peak signal-to-noise ratio and structure. Compared with the other three methods, the peak signal-to-noise ratio has improved. The higher peak signal-to-noise ratio indicates that the better the recovery effect and the less noise, the algorithm of this paper is obviously better than the other three super-resolution reconstruction algorithms, and the edge is also improved to some extent, although the naked eye is not obvious, from the data analysis can clearly draw conclusions.

**Table 1.** PSNR value of the 512 dictionary reconstruction

| Image name | Bilinear interpolation | Bicubic interpolation | Yang | Algorithm |
|---|---|---|---|---|
| Summer resort | 29.740025 | 31.882536 | 33.453914 | 34.035114 |
| Kunming Airport | 28.675323 | 31.746233 | 33.763832 | 34.274221 |
| Nest | 28.427301 | 32.742532 | 34.284642 | 35.714225 |
| Water Cube | 27.294122 | 31.421641 | 33.187225 | 35.827135 |
| Capital Airport | 29.735262 | 31.728113 | 33.931018 | 35.716269 |

The image size in the training library is different. In this experiment, the image block size uses 5 * 5 image blocks, and the number of overlapping pixels is six. In this paper, the super-resolution reconstruction of the image of the Capital International Airport is performed. Two different dictionaries are selected to reconstruct the same image, so as to compare the degree of dictionary training, respectively using different dictionaries for reconstruction as shown in Fig. 1.

At this point, the dictionary is switched again. When the dictionary size is 1024, other parameters remain unchanged, as shown in Fig. 2.

**Table 2.** PSNR value of the 1024 dictionary reconstruction

| Image name | Bilinear interpolation | Bicubic interpolation | Yang | Algorithm |
|---|---|---|---|---|
| Summer resort | 29.740025 | 31.882536 | 33.453914 | 35.132208 |
| Kunming Airport | 28.675323 | 31.746233 | 33.763832 | 36.300179 |
| Nest | 28.427301 | 32.742532 | 34.284642 | 36.901352 |
| Water Cube | 27.294122 | 31.421641 | 33.187225 | 37.125249 |
| Capital Airport | 29.735262 | 31.728113 | 33.931018 | 36.937919 |

（a）LR test chart

（b）Bilinear interpolation reconstruction image

（c）Bicubic interpolation reconstruction image

（d） HR image

（e）Yang algorithm reconstruction image

（f） This article algorithm reconstructs images

**Fig. 1.** Capital International Airport low-resolution image reconstruction

（a）LR test chart

（b） Bilinear interpolation reconstruction image

（c）Bicubic interpolation reconstruction image

（d）HR image

（e）Yang algorithm reconstruction image

（f） This article algorithm reconstructs images

**Fig. 2.** Capital International Airport low-resolution image reconstruction

Figures 1 and 2 show the comparative results of the low-resolution image reconstruction of Capital International Airport in the results obtained by the algorithm. Among them, Figs. 1a and 2a are the LR test images, Figs. 1b and 2b are the reconstruction result of the bilinear interpolation algorithm, Figs. 1c and 2c are the reconstruction result of the bicubic interpolation algorithm, Figs. 1d and 2d are the original HR images, Figs. 1e and 2e are the result of the Yang algorithm reconstruction, and Figs. 1f and 2f in this article reconstruct the results.

## 5 Conclusions

Two different sparse dictionaries are used to compress the super-resolution reconstruction of the same image. The training library images are mainly divided into two categories. The first one is the image information related to the input low resolution, and the second one is the image information independent of the input low resolution. Reconstructed by two types of images. The PSNR values of the same image reconstructed with different dictionary images were analyzed and compared. The results show that the reconstructed image of this algorithm has a certain improvement over the peak signal-to-noise ratio of Yang algorithm.

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation China (61601174), in part by the Postdoctoral Research Foundation of Heilongjiang Province (LBH-Q17150), in part by the Science and Technology Innovative Research Team in Higher Educational Institutions of Heilongjiang Province (No. 2012TD007),in part by the Fundamental Research Funds for the Heilongjiang Provincial Universities (KJCXZD201703), and in part by the Science Foundation of Heilongjiang Province of China (F2018026).

## References

1. Yang X, Wu W, Liu K, et al. Fast multisensor infrared image super-resolution scheme with multiple regression models. J Syst Archit Euromicro J 2016;64(C):11–25.
2. Weiss N, Rueckert D, Rao A. Multiple sclerosis lesion segmentation using dictionary learning and sparse coding. In: Medical image computing and computer-assisted intervention—MICCAI 2013. Berlin, Heidelberg: Springer; 2013. p. 735–42.
3. Yang J, Wright J, Huang TS, et al. Image super-resolution reconstruction via sparse representation. IEEE Trans Image Process. 2010;19(11):2861–73.
4. Zhan S, Fang Q, Yang F, et al. Image super-resolution reconstruction via improved dictionary learning based on coupled feature space. Acta Electron Sin. 2016;44(5):1189–95.
5. Li J, Wu J, Chen Z, et al. Self-learning image super-resolution method based on sparse representation. Chin J Sci Instrum. 2015;36(1):194–200.

# Jitter Detection for Gaofen-1 02/03/04 Satellites During the Early in-Flight Period

Ying Zhu(✉), Mi Wang, Yufeng Cheng, Lin Xue, and Quansheng Zhu

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, People's Republic of China
yzhul003@whu.edu.cn, yzhu_l003@foxmail.com

**Abstract.** Satellite jitter, the fluctuation of satellite point, has negative influence on the geometric quality of high-resolution optical satellite imagery. Gaofen-1 02/03/04 satellites are Chinese operational remote sensing satellites, which have Earth observation capability with 2/8 (panchromatic/multispectral image) meter resolution. This paper presented a jitter detection method based on multispectral imagery for Gaofen-1 02/03/04 satellites. Three short strip datasets captured in the early in-flight period were used to conduct the experiments. The results indicate that satellite jitter with a frequency of 1.1–1.2 Hz, and amplitude of 1–2 pixel exists during the early in-flight period.

**Keywords:** Jitter detection · Multispectral imagery · Gaofen-1

## 1 Introduction

Satellite jitter, which is the fluctuation of satellite point, has become one of the important factors that affect the geometric quality of high-resolution optical satellite imagery products [3, 7]. Satellite jitter has been observed in many satellites and sensors, such as QuickBird, ASTER, ALOS, ZiYuan-3, Tianhui-1 [1, 4–6, 8, 9].

Jitter detection is one of the most important steps in the ground processing system to guarantee the quality of the image products. The method of jitter detection can be categorized into two classes. One is based on high-accuracy attitude sensor with high frequency. The other one is based on imagery with parallax observation.

Gaofen-1 02/03/04 satellites launched on March 13, 2018, are Chinese operational remote sensing satellites. These three satellites adopt same imaging design. Two 2/8 m optical cameras with a panchromatic and four multispectral bands CCDs parallel installed on focal plane, which constructs parallax observation. So the jitter detection based on multispectral images can be applied on the Gaofen-1 02/03/04 satellites.

This paper presents a jitter detection method for Gaofen-1 02/03/04 satellites, and some results are found by using the data captured during the early in-flight period, which can provide an efficient information for in-flight test.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_31

## 2 Methodology

### 2.1 Jitter Detection Based on Multispectral Imagery

Multispectral image of Gaofen-1 satellite has four bands including blue (B1), green (B2), red (B3) and infrared band (B4). Four bands of CCDs are placed in parallel on focal plane with certain distance, shown in Fig. 1. Different bands capture the same scene with a certain time lag when satellite is steady and the coordinate difference of the corresponding points on two different bands image is almost fixed. However, the coordinate difference will vary with time when satellite jitter exists. So satellite jitter can be detected by analyzing band-to-band registration error of multispectral imagery.

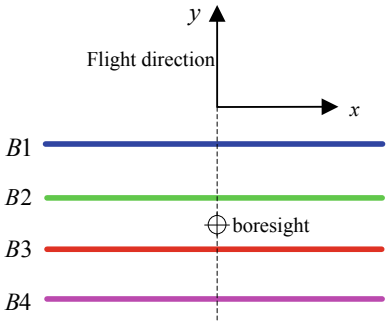



**Fig. 1.** Design of multispectral camera

### 2.2 Workflow of Jitter Detection

There are five main steps in the whole workflow of jitter detection using multispectral imagery.

- **Dense Points Matching**

To ensure the sample points distribute evenly and sufficiently, every pixel in the sample direction is taken as detection candidate point. Correlation matching and least squared matching [2] are successively applied to obtain corresponding points with sub-pixel accuracy between two bands. Corresponding points whose coordinate differences are beyond the threshold value will be detected and removed.

- **Parallax image generation**

After dense points matching, the coordinate differences of every point can be calculated. $\Delta x_{u,k}$, $\Delta y_{u,k}$ are the point difference in sample direction and in line direction indexed with $u$, $k$. where $u$ is line number, $k$ is point number. And two parallax images which present coordinate difference in sample, and line directions are generated by $\Delta x_{u,k}$ and $\Delta y_{u,k}$, respectively.

- **Registration error curve calculation**

Considering the limitation of matching accuracy, the average of coordinate differences of corresponding points in each line is calculated by the formula (1).

$$\begin{aligned} \Delta x(t) &= \sum_{k=1}^{k=V_u} \Delta x_{u,k}(t) \Big/ V_u \\ \Delta y(t) &= \sum_{k=1}^{k=V_u} \Delta y_{u,k}(t) \Big/ V_u \end{aligned} \tag{1}$$

where, $\Delta x(t)$ $\Delta y(t)$ are the registration error across track and along track, $V_u$ is the total number of points in line $u$; $t$ is the imaging time of each line.

- **Spectrum analysis**

To reconstruct the jitter, the periodic parameters of registration error in both directions, including frequency, amplitude and phase, should be determined. First, the main frequency and amplitude of registration error are analyzed by Fourier transform. Then, the Sinusoidal function is used to fit the registration error by taking the spectrum analysis result as the initial value, so the accurate frequency $f$, amplitude $A_\mathrm{r}$ and phase $\varphi'$ can be estimated. And the registration error caused by satellite jitter can be expressed as follows:

$$g(t) = A_\mathrm{r} \cdot \sin(2\pi f t + \varphi') \tag{2}$$

- **Jitter reconstruction**

The frequency of satellite jitter is the same as the registration error. The amplitude and phase of jitter can be calculated using formula (3) and (4).

$$A_d = A_\mathrm{r}/(2\sin(\pi f \Delta t)) \tag{3}$$

$$\varphi = \varphi' - \pi/2 - \pi f \Delta t \tag{4}$$

where, $\Delta t$ is the time interval of two different bands.

Thus, the jitter can be determined by using formula (5)

$$d(t) = \frac{A_\mathrm{r}}{2\sin(\pi f \Delta t)} \cdot \sin(2\pi f t + \varphi' - \pi/2 - \pi f \Delta t) \tag{5}$$

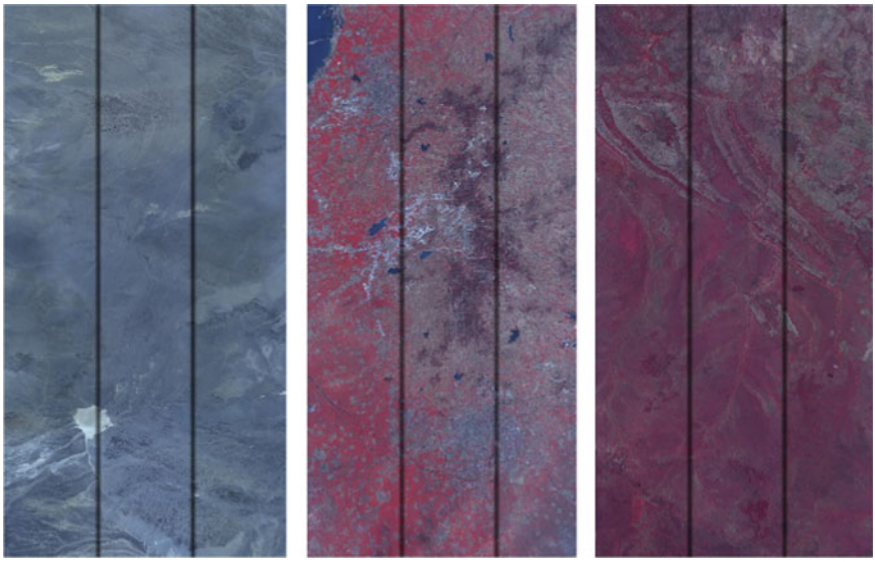
# 3 Experiments and Discussions

## 3.1 Data Description

Three datasets were used to detect the jitter. The basic information about the data is shown in Table 1. The raw multispectral images contain three sub-CCD images. The browse diagrams of three images are shown in Fig. 2.

**Table 1.** Basic information of the datasets

| Image ID | Satellite ID | Center location | Image size |
|---|---|---|---|
| Scene A | 02 | E100.0 N43.0 | 9167 × 1536 × 3 |
| Scene B | 03 | E120.1 N36.9 | 9167 × 1536 × 3 |
| Scene C | 04 | E134.9 S16.3 | 9167 × 1536 × 3 |



(a) Scene A (b) Scene B (c) Scene C

**Fig. 2.** The browse diagrams of three images

### 3.2 Results and Discussions

#### Jitter detection results of Scene A

*Parallax image generation results*

Considering that the radiometric character of infrared band image is much different with that of visible band images, only images of three visible bands were used in the experiments. The parallax images of sub-CCD images between B1 and B2 and parallax images between B2 and B3 were generated as shown in Figs. 3 and 4.

It is obvious that the coordinate difference across the track is periodically varied with imaging line, but there is no obvious change along the track.

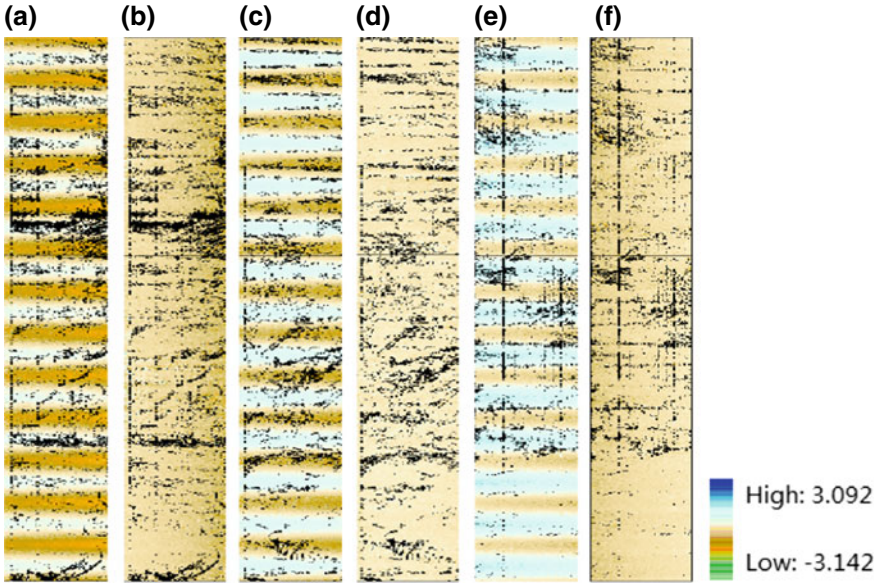



**Fig. 3.** Parallax image generation between B1 and B2 of Scene A. **a** CCD 1 across the track; **b** CCD1 along the track; **c** CCD 2 across the track; **d** CCD 2 along the track; **e** CCD 3 across the track; **f** CCD 3 along the track

*Registration error curves calculation*

The registration curves between two different bands of three sub-CCD images were generated by the method presented in Sect. 2. Figures 5 and 6 show the registration error curves of combination B1–B2 and B2–B3, respectively. The results show that the registration error curves in across track direction have obvious periodicity. But registration error along the track is very small and has a random distribution.

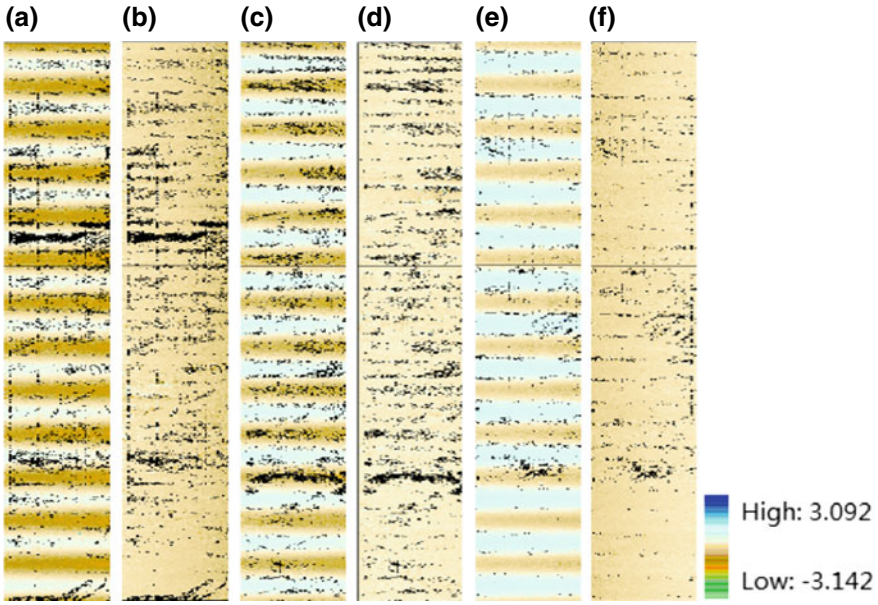



**Fig. 4.** Parallax image generation between B2 and B3 of Scene A. **a** CCD 1 across the track; **b** CCD1 along the track; **c** CCD 2 across the track; **d** CCD 2 along the track; **e** CCD 3 across the track; **f** CCD 3 along the track

*Spectrum analysis*

Table 2 shows the fitting results of registration error in across track direction by Sinusoidal function. The estimated frequency from different band combinations of different sub-CCD images is almost same. As the time interval of different bands is not identical, the amplitudes and phases are different. But the amplitudes and phases are almost the same with same band combination of different sub-CCD images.

*Jitter reconstruction*

Jitter reconstruction results were shown in Table 3. The frequency of satellite jitter is equal to the registration error curve. The averaged frequency is 1.2038 Hz. And the averaged amplitude is 1.3021 pixels in image space, which equals to 10.4168 m in object space. Due to the beginning time between combination B1–B2 and B2–B3 has a certain time interval, the estimated phases by two band combinations are different, but the results of different sub-CCD images with same band combination are almost the same.

## Jitter detection results of Scene B and C

Jitter detection was conducted with Scene B and C and satellite jitter in the across track direction was also found. Jitter reconstruction results of Scene B and C are shown

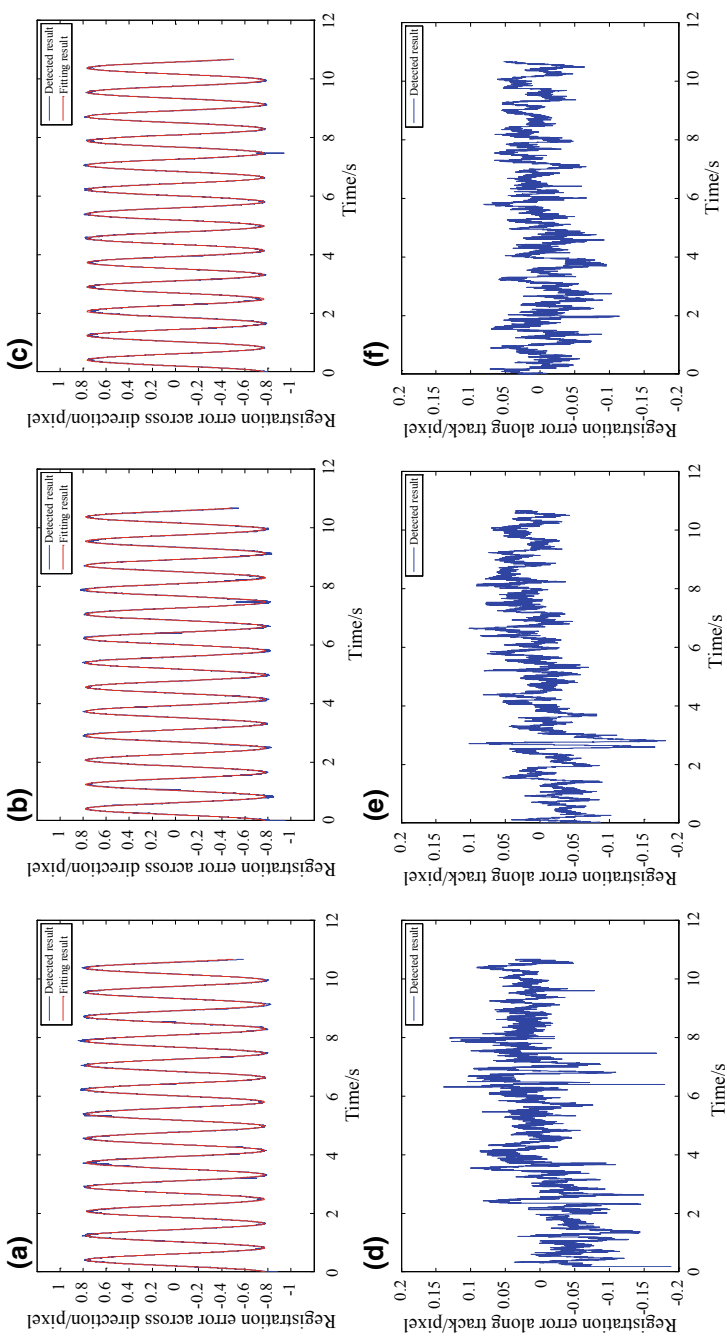



**Fig. 5.** Registration error curves between B1 and B2 of Scene A. **a–c** Registration error of CCD 1–3 across the track; **d–f** registration error of CCD 1–3 along the track

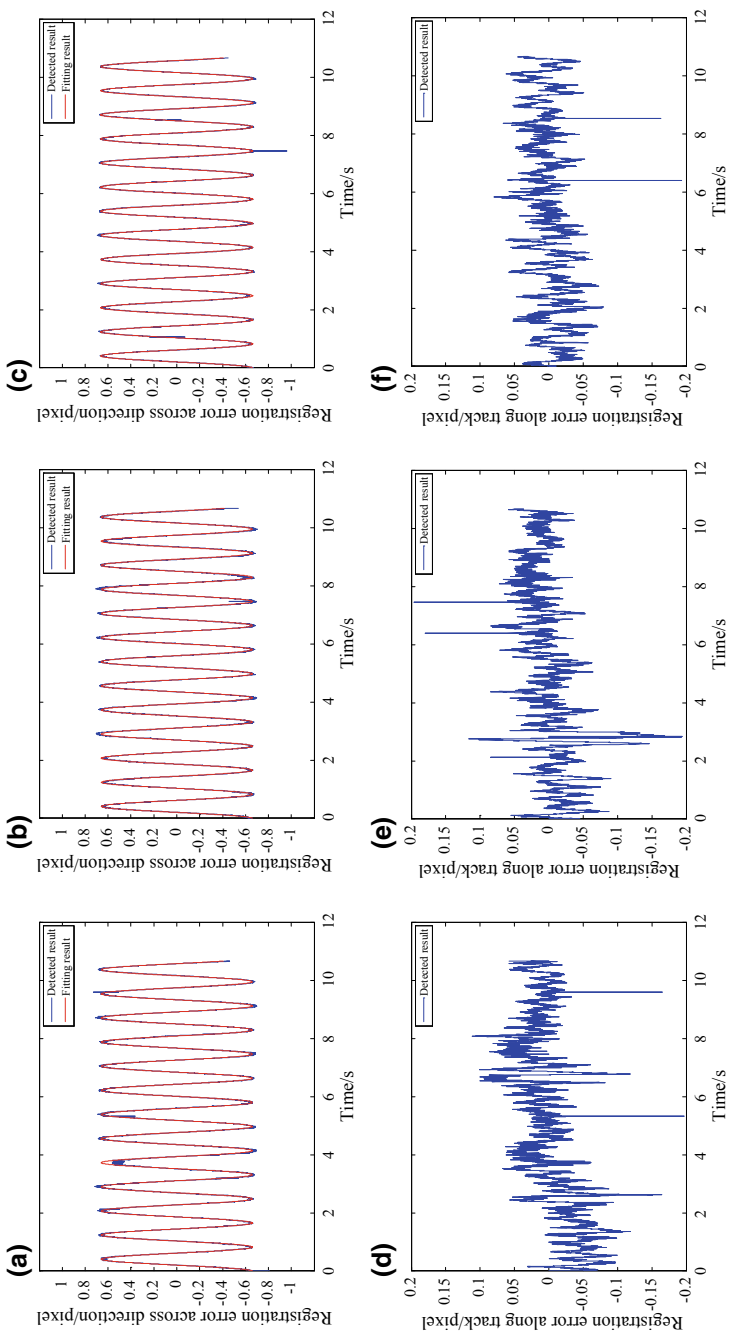



**Fig. 6.** Registration error curves between B2 and B3 of Scene A. **a–c** Registration error of CCD 1–3 across the track; **d–f** registration error of CCD 1–3 along the track

**Table 2.** The fitting results of registration error across the track of Scene A

| Image ID | Band combination | CCD ID | Frequency/Hz | Amplitude/pixels | Phase/rad |
|---|---|---|---|---|---|
| Scene A | B1–B2 | CCD1 | 1.2038 | 0.7717 | 4.7733 |
| | | CCD2 | 1.2039 | 0.7772 | 4.7712 |
| | | CCD3 | 1.2039 | 0.7656 | 4.7692 |
| | B2–B3 | CCD1 | 1.2036 | 0.6604 | 4.7659 |
| | | CCD2 | 1.2040 | 0.6591 | 4.7521 |
| | | CCD3 | 1.2038 | 0.6632 | 4.7574 |

**Table 3.** Jitter reconstruction result of Scene A

| Image ID | Band combination | CCD ID | Frequency/Hz | Amplitude/pixels | Phase/rad |
|---|---|---|---|---|---|
| Scene A | B1–B2 | CCD1 | 1.2038 | 1.2912 | 2.8990 |
| | | CCD2 | 1.2039 | 1.3058 | 2.8982 |
| | | CCD3 | 1.2039 | 1.2827 | 2.8953 |
| | B2–B3 | CCD1 | 1.2036 | 1.3075 | 2.9398 |
| | | CCD2 | 1.2040 | 1.3106 | 2.9271 |
| | | CCD3 | 1.2038 | 1.3149 | 2.9317 |
| | Average value | | 1.2038 | 1.3021 | – |

together in Table 4. During the imaging period of Scene B, the satellite suffered a jitter with frequency of 1.1009 Hz and amplitude of 1.1844 pixels. And the satellite suffered a jitter with frequency of 1.2220 Hz and amplitude of 2.0890 pixels during the imaging period of Scene C.

**Table 4.** Jitter reconstruction result of Scene B and C

| Image ID | Band combination | CCD ID | Frequency/Hz | Amplitude/pixels | Phase/rad |
|---|---|---|---|---|---|
| Scene B | B1–B2 | CCD1 | 1.1010 | 1.1655 | 0.5646 |
| | | CCD2 | 1.1009 | 1.2018 | 0.5686 |
| | | CCD3 | 1.1010 | 1.1685 | 0.5698 |
| | B2–B3 | CCD1 | 1.1009 | 1.1924 | 0.5064 |
| | | CCD2 | 1.1007 | 1.1871 | 0.5091 |
| | | CCD3 | 1.1010 | 1.1913 | 0.5011 |
| | Average value | | 1.1009 | 1.1844 | – |
| Scene C | B1–B2 | CCD1 | 1.2218 | 2.0471 | 0.5250 |
| | | CCD2 | 1.2221 | 2.0322 | 0.5208 |
| | | CCD3 | 1.2220 | 2.1317 | 0.5211 |
| | B2–B3 | CCD1 | 1.2221 | 2.1135 | 0.4571 |
| | | CCD2 | 1.2220 | 2.1007 | 0.4574 |
| | | CCD3 | 1.2220 | 2.1090 | 0.4576 |
| | Average value | | 1.2220 | 2.0890 | – |

## 4 Conclusions and Future Work

In this paper, a jitter detection method based on multispectral imagery for Gaofen-1 02/03/04 satellites was presented. Three short strip datasets captured in the early in-flight period were used to conduct the experiments. The results show that satellite jitter with a frequency of 1.1–1.2 Hz, and amplitude of 1–2 pixel was detected in the across track direction. In the future, more datasets such as a whole strip data will be used to further investigate the spatiotemporal characteristics of the satellite jitter to determine the jitter source and compensate the negative influence on geometric accuracy.

**Acknowledgments.** The authors would like to thank the editors and anonymous reviewers for their valuable comments, which helped improve this paper. This work was supported by the National Natural Science Foundation of China with Project No. 91738302 and. 91438203. The experimental data are provided by CRESDA. These supports are valuable.

## References

1. Ayoub F, Leprince S, Binety R, et al. Influence of camera distortions on satellite image registration and change detection applications. In: 2008 IEEE international geoscience and remote sensing symposium, vol 2, No II; 2008. p. 1072–5.
2. Gruen A. Development and status of image matching in photogrammetry. Photogram. Rec. 2012;27(137):36–57.
3. Iwasaki A. Detection and estimation satellite attitude jitter using remote sensing imagery. Adv Spacecraft Technol. 2011;13:257–72.
4. Sun T, Long H, Liu B, et al. Application of attitude jitter detection based on short-time asynchronous images and compensation methods for Chinese mapping satellite-1. Opt Express. 2015;23(2):1395–410.
5. Takaku J, Tadono T. High resolution DSM generation from ALOS prism-processing status and influence of attitude fluctuation. In: 2010 IEEE international geoscience and remote sensing symposium, Honolulu, July 25–30; 2010.
6. Teshima Y, Iwasaki A. Correction of attitude fluctuation of Terra spacecraft using ASTER/SWIR imagery with parallax observation. IEEE Trans Geosci Remote Sens. 2008;46(1):222–7.
7. Tong X, Ye Z, Xu Y, et al. Framework of jitter detection and compensation for high resolution satellites. Remote Sens. 2014;6(5):3944–64.
8. Tong X, Xu Y, Ye Z, et al. Attitude oscillation detection of the ZY-3 satellite by using multispectral parallax images. IEEE Trans Geosci Remote Sens. 2015;53:3522–34.
9. Wang M, Zhu Y, Jin S, et al. Correction of zy-3 image distortion caused by satellite jitter via virtual steady reimaging using attitude data. ISPRS J Photogram Remote Sens. 2016;119:108–23.

# Hybrid Fractal Image Coding

Guohua Jin, Qiang Wang, and Sheng Bi(✉)

College of Information Science and Technology, Dalian Maritime University, Dalian, China
dmusipi@hotmail.com

**Abstract.** In order to improve the performance of fractal coding methods, a new method is proposed in this paper. Firstly, we find that the range blocks with large variances play a more important role in causing the degradation of decoded images, and the effect of the remaining range blocks can be ignored. Secondly, the range blocks with larger variances will be encoded in an extended domain block pool, and the remaining ones will be encoded with the no-search fractal encoding method. Finally, two fractal coding methods are used to assess the performance of the proposed method. Experiments show that compared with the previous methods, the proposed method can provide shorter encoding time, better quality of decoded images and fewer bits per pixel.

**Keywords:** Fractal image coding · Collage error · No-search · Quality of decoded images

## 1 Introduction

Different from conventional image compression techniques, fractal image coding is characterized by its novel idea, potential high compression ratio, fast decoding, resolution independence and predictable quality of decoded images [1–4]. Thus, it has attracted much attention from the researchers worldwide. However, it suffers from high computational complexity in the encoding process. In order to overcome this problem, some researchers perform block matching in the feature space [5–7]. Due to the low dimension of the features selected, the encoding process can be finished in a short time. In order to accelerate the encoding process furthermore, no-search fractal image encoding is also proposed at the expense of the quality of decoded images [8].

In our research, we find that the range blocks with large variances play a more important role in determining the quality of decoded images. On the contrary, the effect of the remaining range blocks can be ignored. Thus, in our research, we design an extended domain block pool (EDBP) for the range blocks with large variances. EDBP can reduce the collage errors effectively, and then, the quality of decoded images can be expected to improve. But it leads to larger computational complexity in the encoding process, and the bits per pixel (Bpp) increase as well. In order to accelerate the encoding process and reduce Bpp, we adopt the no-search fractal coding method to

Guohua Jin and Qiang Wang—These authors contributed equally to this work and should be considered co-first authors.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_32

encode the remaining range blocks with small variances. Finally, two fractal encoding methods, Jacquin's and He's methods, are used to assess the performance of the proposed method [1, 6]. Experiments show that compared with the previous methods, the proposed method has shorter encoding time, fewer Bpp and better quality of decoded images simultaneously.

This paper is organized as follows: An introduction about conventional fractal image coding will be reviewed in Sect. 2. In Sect. 3, a hybrid fractal encoding method is proposed. By analyzing the importance of variances, we encode the range blocks with larger variances in an extended domain block pool and the remaining range blocks with no-search fractal image coding. In Sect. 4, the experimental procedures and performance of the proposed method will be presented and discussed. Finally, the conclusions are given in Sect. 5.

## 2 Conventional Fractal Image Coding

Fractal image coding is to establish an iterated function system (IFS) whose fixed point can approximate the input image well. It mainly consists of a series of block matching operations between domain blocks and range blocks. The range blocks come from dividing the input image uniformly into nonoverlapping $B \times B$ blocks. The domain blocks can be obtained by sliding a $2B \times 2B$ window over the input image from the left to the right and from the top to the bottom. Generally, the sliding step is set to be $2B$. In order to make the domain blocks match range blocks well, the domain blocks are firstly contracted to the same size of range blocks and then extended with eight isometric transformations. After performing affine transformations on contracted domain blocks, for arbitrary range block, its best-matched domain block can be found by minimizing the following function:

$$\mathrm{CE}(\boldsymbol{R}) = \min_{\alpha,\beta} \|\boldsymbol{R} - \alpha\boldsymbol{D} - \beta\boldsymbol{I}\|^2 \tag{1}$$

where $\boldsymbol{I}$ denotes a matrix whose elements are all ones. $\alpha$ and $\beta$ denote the scaling coefficient and offset coefficient for the affine transformation, respectively. CE($\boldsymbol{R}$) is the collage error for the range block $\boldsymbol{R}$, and $\boldsymbol{D}$ denotes the associated best-matched domain block. In order to minimize Eq. (1), we set the derivatives of Eq. (1) with respect to $\alpha$ and $\beta$ to zeros and the optimal $\alpha$ and $\beta$ can be computed as

$$\alpha = \left\langle \boldsymbol{R} - \bar{r}\boldsymbol{I}, \quad \boldsymbol{D} - \bar{d}\boldsymbol{I} \right\rangle \Big/ \left\|\boldsymbol{D} - \bar{d}\boldsymbol{I}\right\|^2, \ \beta = \bar{r} - \alpha\bar{d} \tag{2}$$

where $\langle \bullet, \bullet \rangle$ denotes the inner product. $\bar{r}$ and $\bar{d}$ are the mean values of $\boldsymbol{R}$ and $\boldsymbol{D}$, respectively. Figure 1 illustrates the above fractal encoding process.

At the decoding phase, arbitrary image can be selected as the initial image. The same transformations as those in the encoding process are applied to the initial image recursively. After about ten iterations, the decoding process converges to the final fixed

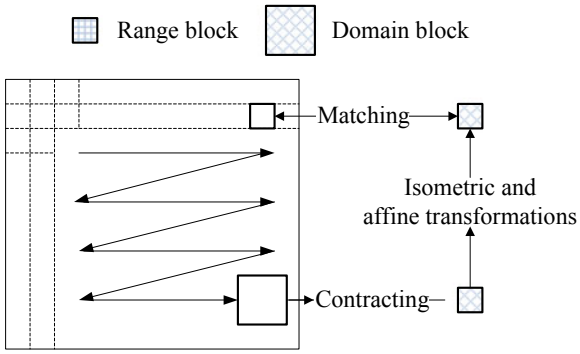



**Fig. 1.** Illustration of the fractal encoding process.

point image. We take the 256 × 256 Lena image, for example, as the input image. If the 256 × 256 Goldhill image is selected as the initial image which is shown in Fig. 2a, the associated first five iteration images in the decoding process are illustrated in Fig. 2b–f, respectively.


**Fig. 2.** **a** Initial image. **b–f** First five iteration images in the decoding process.

## 3 Proposed Method

From Sect. 2, we know that in the encoding process, we spend the same encoding time and memory space for each range block, but the quality of decoded range blocks differs greatly. In our research, we find that the range blocks with larger variances play a more important role in leading to the degradation of decoded range blocks. Thus, a larger domain block pool is designed for them, and then, we can get better quality of decoded images at the expense of longer encoding time and less compression ratio. For the range blocks with small variances, by adopting the no-search fractal coding method, we can get higher compression ratio and shorter encoding time at the expense of slightly worse quality of decoded images. Finally, for the whole input image, the encoding time, compression ratio and quality of decoded images can be all expected to be improved.

In the fractal encoding process, the range blocks are encoded one by one. We define the accumulated collage error (ACE) and accumulated variance (AVAR) as follows:

$$\begin{cases} \mathrm{ACE} = \sum_{i=1}^{\mathrm{Num1}} \mathrm{CE}(\boldsymbol{R}_i) \\ \mathrm{AVAR} = \sum_{i=1}^{\mathrm{Num1}} \mathrm{VAR}(\boldsymbol{R}_i) \end{cases} \tag{3}$$

where Num1 denotes the number of the coded range blocks. By substituting Eq. (2) back to Eq. (1), we can get

$$\mathrm{CE}(\boldsymbol{R}) = \|\boldsymbol{R} - \alpha\boldsymbol{D} - \beta\boldsymbol{I}\|^2 = \|\boldsymbol{R} - \bar{r}I\|^2 - \alpha^2\left\|\boldsymbol{D} - \bar{d}I\right\|^2 \le \|\boldsymbol{R} - \bar{r}I\|^2 \tag{4}$$

From Eq. (3), we can see that for arbitrary range block $\boldsymbol{R}$, its variance provides an upper limit for its collage error, and the range blocks with large collage errors must have corresponding large variances. Thus, only the range blocks with large variances have the possibility to lead to large collage errors. Thus, from Eqs. (3) and (4), we have

$$\mathrm{ACE} \le \mathrm{AVAR} \text{ especially } \mathrm{ACE}^{\mathrm{All}} \le \mathrm{AVAR}^{\mathrm{All}} \tag{5}$$

where $\mathrm{ACE}^{\mathrm{All}}$ and $\mathrm{AVAR}^{\mathrm{All}}$ denotes the ACE and AVAR of all range blocks. Furthermore, we can calculate the average collage error (ACER) as follows:

$$\mathrm{ACER} = \mathrm{ACE}^{\mathrm{All}} / \mathrm{NumR} = \sum_{i=1}^{\mathrm{NumR}} \mathrm{CE}(\boldsymbol{R}_i) \Big/ \mathrm{NumR} \tag{6}$$

In our research, we adopt the peak signal to noise ratio (PSNR) to measure the quality of decoded images as follows:

$$\mathrm{PSNR} = \log_{10}\left(255^2 \Big/ \frac{1}{M \times N}\sum_{j=1}^{N}\sum_{i=1}^{M}\left(\boldsymbol{f}^{\mathrm{Original}}(i,j) - \boldsymbol{f}^{\mathrm{Decoded}}(i,j)\right)^2\right) \tag{7}$$

where $\boldsymbol{f}^{\mathrm{Original}}$ and $\boldsymbol{f}^{\mathrm{Decoded}}$ denote the original image and the decoded image, respectively. $M$ and $N$ are the height and width of the images. Moreover, the logarithmic relationship between the average collage error (ACER) and the peak signal-to-noise ratio (PSNR) of the decoded image can be formulated as

$$f(x) = \beta_1 + \beta_2 \log_{10}(x) \tag{8}$$

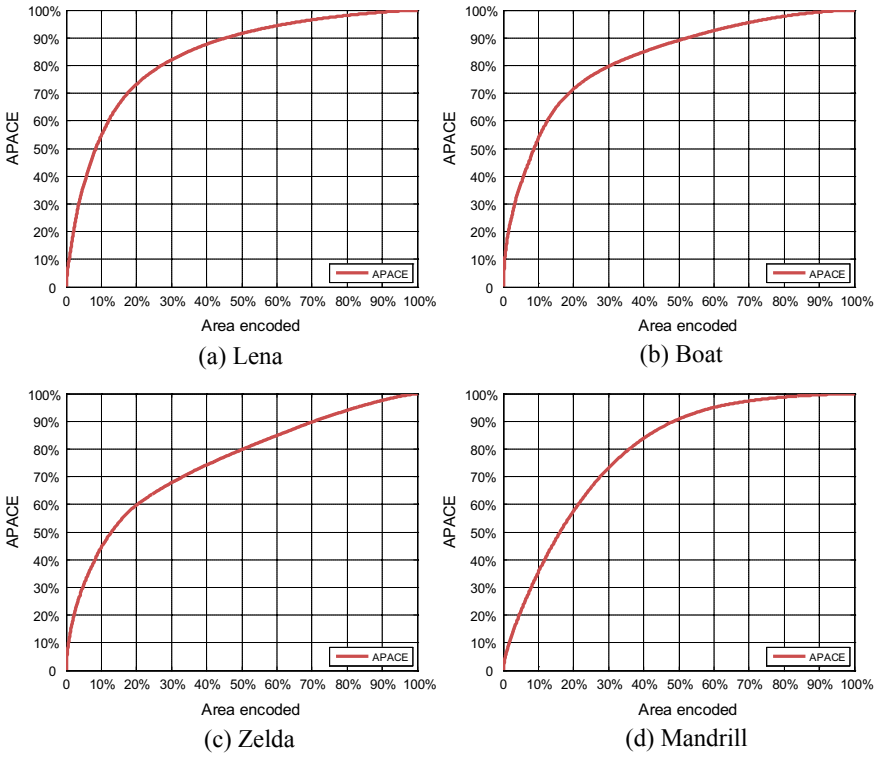
where $\beta_1$ and $\beta_2$ are constant coefficients. $x$ and $f(x)$ denote the ACER and PSNR of the decoded image, respectively.

We take 512 × 512 Lena, Boat, Zelda and Mandrill, for example, to illustrate the importance of the variances of range blocks quantitatively. In the fractal encoding process, we sort the range blocks by their variances from largest to smallest and then encode the range block in order one by one. Then, the actual percentage of accumulated collage error (APACE) can be denoted as

$$\mathrm{APACE} = \mathrm{ACE}/\mathrm{ACE}^{\mathrm{All}} = \sum_{i=1}^{\mathrm{Num1}} \mathrm{CE}(\boldsymbol{R}_i) \Big/ \sum_{i=1}^{\mathrm{NumR}} \mathrm{CE}(\boldsymbol{R}_i) \tag{9}$$

where NumR denotes the total number of range blocks.

From Fig. 3, we can see that in the sorted range block sequence, the first 50% range blocks contribute 91.57, 89.06, 79.72 and 90.74% collage errors for Lena, Boat, Zelda and Mandrill, respectively. The importance of the range blocks with large variances can be verified. Thus, if we can reduce the collage errors of the range blocks with large variances, $\mathrm{AVAR}^{\mathrm{All}}$ will decrease effectively and then, $\mathrm{ACE}^{\mathrm{All}}$ and ACER can be expected to decrease as well. According to Eqs. (7) and (8), the PSNR quality of decoded images can be improved.




(a) Lena (b) Boat

(c) Zelda (d) Mandrill

**Fig. 3.** APACE versus area for Lena, Boat, Zelda and Mandrill, respectively.

We set a threshold $\sigma$ for the variances of range blocks, and all the range blocks can be divided into two different categories: the range blocks with large variances and the range blocks with small variances. For the former ones, their respective best-matched domain blocks will be all searched in an extended domain block pool. Besides the conventional DBP as in Sect. 2, we also design a supplementary domain block pool

(SDBP) which can be constructed with two steps: (1) slide a $4B \times 4B$ window over the input image from the left to the right and from the top to the bottom, and the sliding step is set to be $4B$. (2) Contract all the domain blocks to the size of $B \times B$ by averaging every $4 \times 4$ pixels, and perform eight isometric transformations on them. As stated in Sect. 2, for arbitrary range block with large variances, its best-matched domain block will be searched in both DBP and SDBP. In this case, compared with the conventional method in Sect. 2, the collage errors can be expected to decrease effectively at the expense of more encoding time and memory space. For the latter ones, as illustrated in Fig. 4, appoint the domain block which has the same center with the range block as the best-matched domain block directly. Although the no-search method will lead to larger collage errors, according to Eq. (4), we know that smaller variances will provide an upper limit for their respective collage errors. This implies that small collage errors can be maintained. Moreover, the no-search method can provide shorter encoding time and less memory space. In addition, according to Eq. (8), since $\beta_2$ is a negative value, we know that smaller collage errors imply better decoded image quality. Finally, for the whole image, we can expect to obtain shorter encoding time, less memory space and better quality of decoded images.

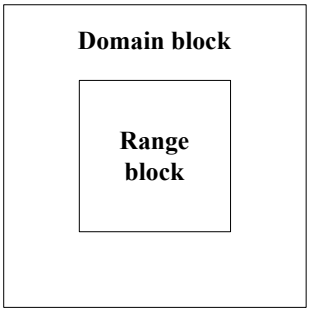



**Fig. 4.** Range block and its associated best-matched domain block.

## 4 Experiments

In this section, four $256 \times 256$ images, Lena, Camera, Couple and Zelda are used. The size of range blocks and the threshold $\sigma$ for the variances of range blocks are set to be 4 and 8, respectively. The size of domain blocks and search step $\delta$ are both set to be 8. The coefficients, $s$ and $o$, are quantized by 5 and 7 bits, respectively. The experiments are carried out on a Pentium Dual-Core 2.93 GHz PC and programmed using MATLAB software. We will compare Jacquin's and He's methods with our proposed method by bits per pixel (Bpp), encoding time and quality of decoded images, respectively [1, 6]. Based on the discussions in Sect. 2, we design the encoding procedures as follows:

Step 1: Given an input image, divide it uniformly into $B \times B$ range blocks. Calculate the variances of range blocks and sort them by their variances from largest to smallest.

Step 2: Slide a $2B \times 2B$ window over the input image from the left to the right and from the top to the bottom. The sliding step is set to be $\delta$. After contracting and perform eight isometric transformations on the domain blocks, a domain block pool (DBP) can be obtained.

Step 3: Slide a $4B \times 4B$ window over the input image as in Step 2. The sliding step is set to be $4B$. Contract the domain blocks into the same size of range blocks and then perform eight isometric transformations on them. A supplementary domain block pool (SDBP) is obtained.

Step 4: For the range blocks with large variances, their respective best-matched domain block is searched within DBP and SDBP. For arbitrary range block, we will use one bit to label which domain block pool, DBP or SDBP, is used, and thus an extra column vector $V1$ is needed. The remaining range blocks with small variances are encoded with the no-search fractal encoding method. Another $1 \times$ NumR column vector $V2$ is needed to label the above two kinds of range blocks.

Table 1 illustrates the comparison results between Jacquin's method and our proposed method. Similar comparison results between He's method and our proposed method are listed in Table 2 as well. From the second and third rows of Tables 1 and 2, we can see that the proposed method can provide better quality of decoded images than Jacquin's method. Compared with Jacquin's method, since the range blocks with large variances can find their respective best-matched domain blocks in a larger domain block pool, EDBP, this will result in smaller collage errors, and then, better quality of decoded images can be obtained. From the fourth to seventh rows of Table 1, we can

**Table 1.** Performance comparison between Jacquin's method and the proposed method [1].

| Images | | Lena | Camera | Couple | Zelda |
|---|---|---|---|---|---|
| PSNR (dB) | Jacquin's [1] | 31.96 | 29.37 | 34.74 | 35.02 |
| | Proposed | **32.12** | **29.65** | **35.10** | **35.23** |
| Bpp | Jacquin's [1] | 1.56 | 1.56 | 1.56 | 1.56 |
| | Proposed | **1.45** | **1.17** | **1.34** | **1.36** |
| Time (s) | Jacquin's [1] | 5.23 | 5.23 | 5.20 | 5.22 |
| | Proposed | **5.06** | **3.26** | **4.34** | **4.42** |

**Table 2.** Performance comparison between He's method and the proposed method [6].

| Images | | Lena | Camera | Couple | Zelda |
|---|---|---|---|---|---|
| PSNR (dB) | He's [6] | 31.78 | 29.15 | 34.31 | 34.71 |
| | Proposed | **31.93** | **29.26** | **34.52** | **34.94** |
| Bpp | He's [6] | 1.56 | 1.56 | 1.56 | 1.56 |
| | Proposed | **1.45** | **1.17** | **1.34** | **1.36** |
| Time (s) | He's [6] | 1.45 | 1.44 | 1.48 | 1.45 |
| | Proposed | **1.17** | **0.80** | **1.02** | **1.03** |

see that compared with Jacquin's method, the proposed method can provide less encoding time and need less memory space. For the range blocks with smaller variances, since their best-matched domain blocks are directly designated without the searching process, and there is no need to store the position information of the best-matched domain blocks, shorter encoding time and less memory space can be achieved.

## 5 Conclusions

In our research, we propose a method to improve the performance of fractal coding methods. By analyzing the importance of the variances of range blocks, we find that the range blocks with large variances play a more important role in determining the quality of decoded images. Thus, we design an extended domain block pool for the range blocks with large variances. For the remaining range blocks with small variances, the no-search fractal coding method is adopted. Experiments show that the proposed method can improve the performance of fractal coding methods effectively and make them more useful in practical applications.

## References

1. Jacquin AE. Image coding based on a fractal theory of iterated contractive image transformations. IEEE Trans Image Process. 1992;1(1):18–30.
2. Fisher Y. Fractal image compression: theory and Application. New York: Springer; 1994.
3. Wohlberg B, Jager G. A review of the fractal image coding literature. IEEE Trans Image Process. 1999;8(12):1716–29.
4. Wang Q, Bi S. Improved method for predicting the peak signal-to-noise ratio quality of decoded images in fractal image coding. J Electron Imaging. 2017;26(1):013024.
5. Lai CM, Lam KM, Siu WC. Improved searching scheme for fractal image coding. Electron Lett. 2002;38(25):1653–4.
6. He CJ, Xu XZ, Yang J. Fast fractal image encoding using one-norm of normalized block. Chaos, Solitons Fractals. 2006;27(5):1178–86.
7. Wang Q, Liang D, Bi S. Fast fractal image encoding based on correlation information feature. In: Proceedings of the 3th international congress on image and signal processing, Yantai, China; 2010. p. 540–3.
8. Furao S, Hasegawa O. A fast no search fractal image coding method. Sig Process Image Commun. 2004;19(5):393–404.

# Image Encryption Algorithm Based on Chaotic Sequence

Xiaodi Chen and Hong Wu(✉)

College of Electronic Engineering, Heilongjiang University, Harbin 150080, China
2002060@hlju.edu.cn

**Abstract.** With the rapid development of communication technology and the Internet, the security of image information has also received more attention. The information characteristics of an image can be described by pixel position and pixel value. Therefore, the encryption algorithm for the image is generally based on these two aspects (Commun Nonlinear Sci Number Simulat 17(7):2969–2977, [1]; Int J Bi-furcation Chaos 7(7):1579–1997, [2]). Using a method alone can easily be cracked by an attacker, so image encryption often combines pixel position scrambling with gray value diffusion. In this paper, chaotic sequences are applied to image data encryption, and pixel position scrambling and gray value diffusion are combined to obtain ciphertext images. And the feedback mechanism is introduced in the diffusion link, so that the ciphertext of each pixel is not only related to the current plaintext, but also relevant to the ciphertext of all the previous pixels, thereby increasing the complexity of the ciphertext.

**Keywords:** Pixel position · Pixel value · Gray value diffusion · Feedback mechanism · Chaotic sequence

## 1 Ciphertext Feedback Image Gray Value Diffusion

Image gray value diffusion is to regard the image as a matrix, and the purpose of changing the gray value of the pixel in the image is achieved by a certain algorithm [3]. The more commonly used algorithm is to XOR the image matrix and the chaotic matrix in the form of stream encryption. Nevertheless, the security of the system is completely determined by the chaotic sequence [4, 5]. In order to overcome this issue, this paper introduces a feedback mechanism in the diffusion link, so that the ciphertext of each pixel is not only relevant to the current plaintext, but also relevant to the ciphertext of all the previous pixels, which increases the complexity of the ciphertext.

The pixel diffusion method of ciphertext feedback is as shown in Eqs. (1.1) and (1.2). After the current pixel which XOR with the chaotic sequence, which obtains an intermediate quantity $a(i)$, and afterward the ciphertext of the previous pixel which XOR with the $a(i)$, and its consequence is used as the ciphertext of the current pixel, and subsequently a constant $c(0)$ is set to fulfill the diffusion of the first pixel.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_33

$$a(i) = \text{bitxor}(d(i), I(i)) \tag{1.1}$$

$$c(i) = \text{bitxor}(a(i), c(i-1)) \tag{1.2}$$

$d(i)$ is the chaotic sequence which used for diffusion, $I(i)$ is plaintext pixel which after scrambling, $c(i)$ for the currently processed ciphertext pixel, $c(i-1)$ which is the previous ciphertext pixel.

In order to make the pixel pixels behind the position also affect the previous pixel encryption, this paper uses two rounds of diffusion encryption, that is, one round of positive sequence encryption and one round of reverse order encryption. The ciphertext is more secure.

## 2 Chaotic Sequence Generation Algorithm

(1) Given the initial values and parameters of the Logistic map, and the Logistic map iteratively generates the first layer of chaotic sequences.

$$\{x_{k0} | k = 0, 1, 2\ldots, M\} \tag{2.1}$$

(2) Each of $x_{k0}$ is used as the initial value of the Cubic map, and a sequence of length $N$ is iterated; $M$ $x_{k0}$ produces $M$ chaotic sequences, and they are connected in series to obtain a second sequence.

$$\{x_{ki} | k = 0, 1, 2\ldots, M; i = 0, 1, 2, \ldots, N\} \tag{2.2}$$

(3) According to the value of $x_{k0}$ and $x_{kN}$ which iterated by the initial value $x_{k0}$, after the parameter $\mu$ of the Logistic map and the parameter a of Cubic map are dynamically changed according to the following method to form a chaotic system whose parameters are mutually controlled.

$$\mu = \begin{cases} 3.9 & x_{kN} \geq 0 \\ 4 & x_{kN} < 0 \end{cases} \tag{2.3}$$

$$a = \begin{cases} 1 & x_{k0} \geq 0.5 \\ 4 & x_{k0} < 0.5 \end{cases} \tag{2.4}$$

(4) Discard the front-end sequence of each subsystem iteration to ensure that the system enters a chaotic state. Give the system a different initial value $y_{00}$, after repeat the above steps to generate a composite sequence $\{y_{ki}\}$.

(5) Select 0 as the threshold and quantize the chaotic sequence $\{x_{ki}\}$ and $\{y_{ki}\}$ according to the binary decision method of Eq. (2.3) to obtain the binary 0 and 1 sequence $\{X_{ki}\}$ and $\{Y_{ki}\}$. $\{X_{ki}\}$ XOR $\{Y_{ki}\}$ with to obtain random sequence $\{Z_{ki}\}$ which aims to encryption.

## 3 Image Encryption Algorithm Based on Improved Chaotic Sequence

The improved chaotic sequence which is applied to image data encryption and the ciphertext image obtains by combining pixel position scrambling and gray value diffusion.

### 3.1 Pixel Position Scrambling

Instead of simply listing headings of different levels, we recommend that every heading is followed by at least a short passage of text.

(1) Enter the plaintext image and read the image size as $M \times N$.
(2) Set the initial value $x_{00} = 0.1$, $y_{00} = 0.1001$, initial parameter $\mu = 4$, and generate a chaotic sequence $\{x_1, x_2, x_3, \ldots, x_{M\times N\times 8}\}$ of length $M \times N \times 8$ by the improved algorithm proposed in Sect. 2. Convert every 8 bits in a chaotic sequence into a decimal number. At this time, the sequence is grayscale sequence $\{d_1, d_2, d_3, \ldots, d_{M\times N}\}$ whose the value range is [0, 256] and the length is $M \times N$.
(3) Sorting the sequences in order from minor to large, which obtains ordered sequence $\{d_i, d_k, \ldots, d_p\}$ and position sequence $\{w\}$, and reconfiguring the position sequence into a position matrix $L$ whose the size is $M \times N$; the plaintext image matrix is rearranged according to the position matrix $L$ to generate an image $I$ whose pixel position is scrambled.

### 3.2 Pixel Gray Value Diffusion

The pixel gray value is diffused to eliminate the statistical features such as the histogram of the plaintext image, which serves to confuse the ciphertext image with the plaintext image.

(1) Convert the scrambled image $I$ into the $I_2$ sequence, whose the length is $1 \times (M \times N)$.
(2) The $i$th pixel in the one-dimensional pixel sequence $I_2$ and the element of the chaotic sequence $\{d\}$ is sequentially subjected to one round of positive sequence encryption in accordance with Eqs. (1.1) and (1.2) $(i = 1, 2, \ldots, M \times N)$.
(3) Subsequently, the elements in the pixel sequence $I_2$ are subjected to the second round of reverse order encryption in the order of $(i = M \times N, M \times N - 1, \ldots, 2, 1)$ and the chaotic sequence $\{d\}$ according to the formula (1.1) and (1.2), and the ciphertext sequence $\{c\}$ is obtained, and the diffusion of all the pixel points is completed. Reshape the sequence $\{c\}$ into $M \times N$ matrix to obtain a ciphertext image.

The decryption process is the inverse of the encryption process. The gray sequence and the position sequence are obtained by the key and the parameter. Decrypting the two rounds of diffusion encryption according to the inverse operations of the Eqs. (1.1) and (1.2). Reconstruct the scrambled image according to the position matrix. And the plaintext image is obtained.

## 4 Summary

This article describes in detail how to apply chaos theory to the field of image encryption. And the sequence key which generates by the compound chaotic sequence proposed in Sect. 2. After the pixel spatial position and pixel gray value of the image which are scrambled and diffused. Finally, it introduces a feedback mechanism to greatly improve the security of the algorithm.

## References

1. Xue KP, Hong PL. Security improvement on an anonymous key agreement protocol based on chaotic maps. Commun Nonlinear Sci Number Simulat. 2012;17(7):2969–77.
2. Short KM. Signal extraction from chaotic communication. Int J Bifurcat Chaos. 2016;7 (7):1579–997.
3. Wei L (2013) Research on image encryption algorithm based on four-dimensional hyperchaotic map. Harbin, China.
4. Vattulainen I, Ala-Nissila T, Kankaala K. Physical models as tests of randomness. Phys Rev E. 2015;52(3):3205–14.
5. Li TY, Yorke JA. Entropy and chaos. Adv Math. 2013;3:122–8.

# Weighted Nuclear Norm Minimization Image Denoising Method Based on Noise Variance Estimation

Shujuan Wang[1(✉)], Ying Liu[1], Hong Liang[2], and Yanwei Wang[3]

[1] College of Science, Harbin Engineering University, Harbin, China
wangshujuan@hrbeu.edu.cn
[2] College of Automation, Harbin Engineering University, Harbin, China
[3] College of Mechanical Engineering, Harbin Institute of Petroleum, Harbin 150001, China

**Abstract.** Weighted nuclear norm minimization (WNNM) uses image non-local similarity to deal with image denoising; this method not only maintains the detailed texture edge structure but also reduces the impact on distortion of the image after denoising. However, WNNM method assumes that the noise variance of the image is known, where the parameter is set by subjective experience that will result in incompleteness in theory. To handle this issue, it is proposed to pre-estimate noise variance based on discrete wavelet transformation (DWT). The simulation result shows that compared with original WNNM method, pre-estimate noise variance in image denoising has a faster algorithm running speed and a higher image signal-to-noise ratio after denoising.

**Keywords:** WNNM algorithm · Discrete wavelet transformation · Singular value decomposition · Image denoising

## 1 Introduction

The essential problem of image denoising is matrix recovery which can only be recovered under certain conditions. On one side, the original matrix is low-rank. On the other side, pollution such as damage and noise is sparse, when some elements of the matrix are severely damaged sparse and low-rank matrix decomposition can be used to restore the original matrix. In this way, if the original matrix is low-rank of arbitrary size and with existence of small amount of sparse noise, the restoration of the matrix can be described as the optimization problem represented by the following formula (3).

Let the size $m \times n$ of the original image be $f(x, y)$, $\eta(x, y)$ is noise, $\xi(x, y)$ is image that contaminated by noise, the following equation can be derived.

$$\begin{aligned} &\xi(x, y) = f(x, y) + \eta(x, y), \\ &x = 1, 2, \ldots, m, y = 1, 2, \ldots, n \end{aligned} \tag{1}$$

The purpose of image denoising is to recover the original image that is not contaminated by noise $f(x, y)$ and to remove noise interference $\eta(x, y)$ from noisy image


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_34

$\xi(x, y)$. For noisy image $\xi(x, y)$, it discretizes into a $m \times n$ size matrix after sampling and quantization process. So, we can rewrite (1) as a matrix as:

$$Y = X + E \tag{2}$$

from which matrix $X$ represents the original image $f(x, y)$, matrix $E$ stands for noise $\eta(x, y)$.

When data matrix has a low-rank structure and only some small portion elements are heavily polluted by noise, which indicates noise matrix is sparse, then matrix noise reduction problem could be described as the following model:

$$\begin{aligned} \min \quad & \operatorname{rank}(X) + \gamma \|E\|_0 \\ \text{s.t.} \quad & X + E = Y \end{aligned} \tag{3}$$

In this equation, the objective function is the rank of clean image matrix $X$ and zero norm of noise matrix $E$, $\gamma$ represents the weights of noise, which is used to adjust the relation between the rank of matrix $X$ and sparseness of noise matrix $E$. It is claimed that, the difficulty of solving Eq. (3) is *NP*, so, the objective function should be loosen [2].

**Lemma 1** *$S = \left\{X \in R^{m \times n}, \|X\|_2 \leq 1\right\}$, then the convex envelope of $\phi(X) = \operatorname{Rank}(X)$ can be expressed as $\phi_{\text{env}} = \|X\|_* = \sum_i^{\min\{m,n\}} \sqrt{\lambda_i(XX^T)}$.*

It can be seen from the lemma that the solution to the matrix rank minimization problem can be transformed into the solution of the nuclear norm minimization problem [1, 2, 3]. Model expresses as follows:

$$\begin{aligned} \min \quad & \|X\|_* + \gamma \|E\|_1 \\ \text{s.t.} \quad & X + E = Y \end{aligned} \tag{4}$$

where $\gamma$ represents the weights of noise, model (4) converts the non-convex optimization problem of model (3) into approximate solution to convex optimization problem [4].

## 2 Weighted Minimum Nuclear Norm Denoising

The weighted kernel norm minimization problem is expressed as follows:

$$\hat{X} = \arg\min_X \gamma \|Y - X\|_F^2 + \|X\|_{\varpi,*} \tag{5}$$

where $\gamma$ is noise weight, for local blocks $y_i$ in image $Y$. We can find its non-local similarity block in the image by some method [5, 6] (such as block matching, that is, finding the block with similar texture and structure in the whole image), by stacking these non-local similar blocks into a matrix $Y_i$. Finally, by aggregating all the denoising blocks, the entire image will be estimated. Have the following expression:

$$\hat{X}_i = \arg\min_{X_i} \gamma \|Y_i - X_i\|_F^2 + \|X_i\|_{\varpi_i,*} \tag{6}$$

where $\gamma$ is noise weight, which is used to balance the contradiction between sparse matrix and low-rank matrix, $\text{rank}(X_i) = r_i$, $\|X_i\|_{\varpi,*} = \sum_{j=1}^{r} \left|\varpi_j \sqrt{\lambda_j(X_i X_i^T)}\right|$, $\sqrt{\lambda_j}$ represents the $j$th singular value of block matrix $X_i$, $\varpi_j$ represents the $j$th weight of singular value of block matrix $X_i$, $\varpi = [\varpi_1, \varpi_2, \ldots, \varpi_{r_i}]$, $\varpi_j \geq 0$.

We assume

$$\varpi_j = \frac{c\sqrt{n}}{\left(\sqrt{\lambda_j(X_i X_i^T)} + \varepsilon\right)}$$

where $c > 0$ is a constant, $n$ is the number of similar patches of $Y_i$, $\varepsilon = 10^{-16}$ in order to avoid the error term with zero denominator.

## 3 Noise Variance Estimation

With the weights defined above, the WNNM algorithm can be used directly to solve the model in (6). However, there is still a problem that the singular value $\sqrt{\lambda_j(X_i X_i^T)}$ is unknown. We assume that the noise energy follows uniform distribution over each subspace spanned by the base pairs of $U$ and $V$ [2], then initialize singular value. $\sqrt{\lambda_j(X_i X_i^T)}$ can be estimated as

$$\sqrt{\hat{\lambda}_j(X_i X_i^T)} = \sqrt{\max\left\{\left[\sqrt{\lambda_j^2(Y_i Y_i^T)} - n\sigma_n^2\right], 0\right\}}$$

where $\sqrt{\lambda_j(Y_i Y_i^T)}$ is the $j$th singular value of $Y_i$, $n$ is the number of similar patches to $Y_i$, $\sigma_n^2$ is the noise variance. By applying the above process to each block and reaggregating all the blocks together, a noise-free image $X$ can be reconstructed.

The above solution $\varpi$ needs to estimate the singular value of the original image, and the calculation process needs to accurately calculate the noise variance parameter of the image. Thus, the original algorithm has two disadvantages:

(1) In the process of denoising, the noise concentration is often selected in a subjective manner by setting a fixed noise estimation parameter within a certain noise range and finally selecting the results experiments.
(2) The algorithm can get better denoising result, but the program runs longer, and the time cost is high.

$L$ stands for the low frequency of the signal while $H$ represents the high frequency of the signal. On the other words, $L$ is the smooth part of the signal, $H$ is the detail part. The secondary wavelet decomposition of two-dimensional image can be shown as follows (Fig. 1).

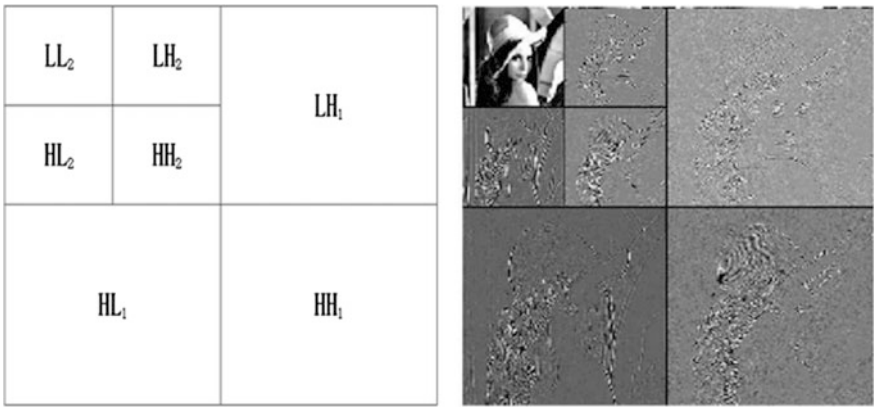



**Fig. 1.** Secondary wavelet decomposition

Selecting the appropriate wavelet and transform the given signal then get the transformed wavelet coefficients $W$. Let wavelet of the noisy image be

$$\omega_x(i) = \omega_n(i) + \omega_s(i),$$

where $\omega_n(i)$ represents the noise wavelet coefficient, $\omega_s(i)$ is the wavelet coefficient for image edge information. In our experiments, we use this method to compare with Donoho and Johnstone's estimation methods. The noise variance is estimated as follows:

$$\sigma = \frac{1}{0.6745MN} \sum_{i=1}^{M} \sum_{j=1}^{N} |\omega_n(i,j)| \tag{7}$$

The noise estimation process is as follows:

1. Firstly, implement wavelet transform on the noisy image, and the high-frequency sub-band *HH* obtained after the second-order wavelet decomposition is obtained;
2. Then, applying Sobel operator to perform edge extraction on the sub-band, which can obtain high-frequency edge information;
3. After that, subtracting the sub-band edge information from the sub-band obtained by decomposing the second wavelet of the noisy image to obtain noise information;
4. Finally, use formula (7) to estimate noise variance.

An algorithm based on noise variance estimation is proposed, as follows:

1. Input noise image $f(x, y)$, sampling quantization into matrix $Y$, initialize $\widehat{X}^0 = Y$, $Y^0 = Y$;
2. for $k = 1 : k$ do iterative regularization

$$Y^k = \widehat{X}^{k-1} + \delta\left(Y - \widehat{Y}^{k-1}\right);$$

3. Estimate noise variance using discrete wavelet transformation $\sigma_n^2$;
4. Choose an arbitrary $y_i$ from $Y^k$;
5. Find block matrix group $Y_i$ like $y_i$;
6. Estimate weight vector $\varpi$;
7. Singular value decomposition $[U, \Sigma, V] = \mathrm{SDV}(Y_i)$;
8. Derive estimation: $\widehat{X}_i = \mathrm{US}_{\varpi}(\Sigma)V^T$;;
9. Aggregate $X_i$ to form a clear image $\widehat{X}^k$.

## 4 Simulation

Due to space limitations, we take the Monarch image as an example. Firstly, a Gaussian white noise with a standard deviation of 20 is added to the Monarch diagram. Figure 2b shows the Monarch image after adding $\sigma = 20$ Gaussian noise image. We conduct the second-order discrete wavelet decomposition on noisy image to obtain the high-frequency sub-band *HH*. For the wavelet denoising image, the edge of the image is extracted by the Sobel operator. Finally, the variance is calculated by subtracting the image edge information from the high-frequency sub-band *HH*. The process is as follows:

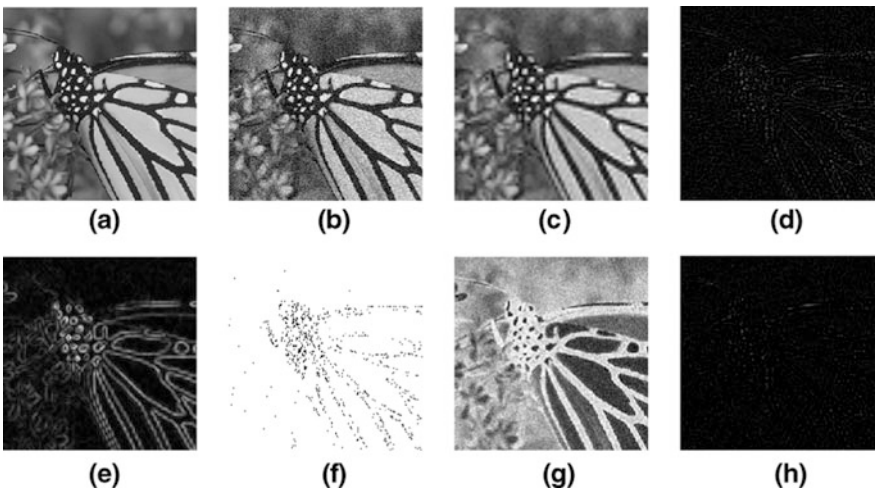



**Fig. 2.** Noise estimation. **a** Original Monarch image, **b** noisy image, **c** wavelet first-order decomposition denoising image, **d** wavelet first-order decomposition high-frequency sub-band HH, **e** the edge of image extracted by Sobel operator, **f** reduce the edge information, **g** subtract the image of the edge, **h** final noise image

In order to verify the superiority of the proposed method, each image is compared with the method by *DJ* method, and the results are as follows (Table 1).

**Table 1.** Standard deviation estimation comparison table

| Image | St.dev. | *DJ* method | Improved method |
|---|---|---|---|
| Monarch | 20 | 28.5860 | 25.9439 |
| | 40 | 44.0566 | 42.5434 |
| | 60 | 61.9536 | 61.0306 |
| | 80 | 80.8701 | 80.2661 |
| | 100 | 100.2945 | 99.8717 |

Obviously, the method of this paper estimates the noise variance more accurately. The pre-estimated noise parameters are then used for algorithm image denoising. The simulation results are as follows.

Adding Gaussian noise with standard deviation of 20, 30, 40, 50, 60, 70, 80, 90, 100, respectively, compare median filtering, Wiener filtering, wavelet threshold denoising, WNNM, and improved WNNM. The peak signal noise of the post image is as shown in Fig. 3.

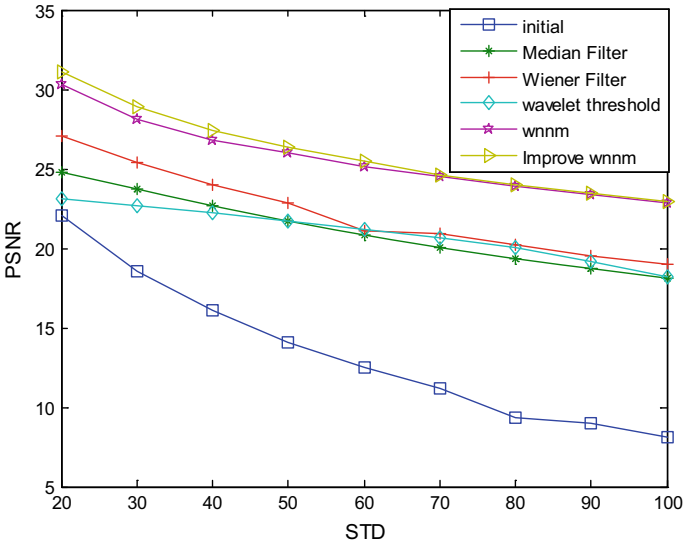



**Fig. 3.** Comparison of peak signal-to-noise ratios of five denoising algorithms

Compared with other algorithms, the image distortion obtained by the improved WNNM algorithm is smaller and the similarity with the original image is higher, which fully demonstrates the superiority of the improved WNNM algorithm.

## 5 Conclusion

In this paper, the wavelet transform is used to estimate the noise signal variance, and weighted nuclear norm minimization algorithm is introduced to establish a new image denoising method. The specific application uses Sobel operator to extracts the edge information of the high-frequency sub-band of the second-order wavelet decomposition in noisy image and combines the edge information of the image to estimate the variance of the noise signal. The simulation results show that the proposed method not only reduces the time for selecting the optimal noise parameters by comparing the experimental results, but also estimates the noise concentration in advance to improve the signal-to-noise ratio. These demonstrate feasibility and efficiency of the improved algorithm on image denoising.

**Acknowledgments.** This work was supported by the Fundamental Research Funds for the Central Universities under Grant No. HEUCFP201802.

## References

1. Dong W, Shi G, Li X. Nonlocal image restoration with bilateral variance estimation: a low-rank approach. IEEE Trans Image Process. 2013;22(2):700–11.
2. Gu S, Zhang L, Zuo W et al. Weighted nuclear norm minimization with application to image denoising. In: IEEE conference on computer vision and pattern recognition, Columbus: IEEE; 2014. p. 2862–9.
3. Gu S, Xie Q, Meng D, et al. Weighted nuclear norm minimization and its applications to low level vision. Int J Comput Vis. 2016;121:1–26.
4. Canh TN, Dinh KQ, Jeon B. Multi-scale/multi-resolution Kronecker compressive imaging. In: 2015 IEEE international conference on image processing (ICIP), IEEE; 2015. p. 2700–04.
5. Papyan V, Elad M. Multi-scale patch-based image restoration. IEEE Trans Image Process. 2016;25(1):249–61.
6. Dong W, Shi G, Li X. Nonlocal image restoration with bila-teral variance estimation: a lowrank approach. IEEE Tran Image Process. 2013;22(2):700–11.

# The Design and Implementation of Display System for Campus Internet Student Behavior

Xinxin Huang, Mei Nian, Haifang Li, and Bingcai Chen(✉)

School of Computer Science and Technology, Xinjiang Normal University, Urumqi, China
{345566810,cbc9}@qq.com

**Abstract.** The campus online behavior analysis needs to graphically display the distribution of the network traffic and the students' application status in the campus network at present or in a period of time. After capturing and analyzing the campus network traffic data, statistics and analysis are performed on the acquired data. Based on the result, a visual display can be obtained. It helps the network administrator understand whether the use of campus network traffic is reasonable and can facilitate student management personnel to understand whether the students' Internet behavior is healthy. In this paper, the B/S framework is used to capture and analyze the data obtained by the campus network using Web pages.

**Keywords:** Internet behavior · Web traffic · Graphical

## 1 Introduction

With the rapid development of the network, more and more people begin to contact the network. The network brings convenience to students in the aspect of life and study. In colleges and universities, the online information of all students' must pass through the gateway. It is feasible to use the student's online data to analyze the online behavior of students [1].

Statistical analysis and visualization of students' online data can help school administrators understand the daily Internet behavior of students and can also help them understand how long students spend in various applications which can provide the factual basis for managers to manage students well. It can also help network managers understand the use of broadband in schools and the bandwidth that schools spend on various applications which can provide reliable data support for the network construction and optimization of universities [2].


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_35

## 2 The Design of Campus Network Information Display System

### 2.1 The Overall Framework and Function of the Campus Network Information Display System

The campus network information display system needs to store the student online data which are captured and parsed in the MySQL database and then search through the database in the foreground using the JFreeChart mapping component. The data that meets the conditions is visually displayed. The design of the system is mainly to enable school administrators to better understand the students' daily online behavior and let network center managers understand the use condition of bandwidth in school's network so as to optimize the network more rationally. According to the different requirements of the school administrators (student administrators pay attention to the length of time students spend on various applications while the network administrator focuses on the traffic that the students spend on various applications), The system mainly includes the following functional modules student administrator module and network administrator module [3].

### 2.2 Graphic Display Design

The online time is the most important behavior of the students' online behavior. The student administrator module classifies and integrates the students' online information through the statistics and analysis of the students' online information. Internet traffic is the most important feature of online behavior for network administrators. Since the length of the Internet using time is similar to that of the online traffic graph, the following is an example of the Internet time.

1. Internet time chart display. After entering the system, the student administrator can select the date to be viewed to view the chart of the online time on each application in the whole school, including the online time chart of the day, week, and month.
2. Personal online time chart display. Due to the large number of college students, the system provides two ways for student administrators to view the length of their personal online time. The student administrator can search for the online behavior of the designated student based on the student number or student name.
3. College online time chart display. The student administrator can perform a targeted analysis on the statistical analysis of the online time for a specific college student. After entering the system, the student administrator selects the college first and then clicks the user analysis to select the date to be viewed to view the online duration information of the college students in the current day, week, or month.
4. Class online time chart display. The student administrator can get a targeted analysis of the statistical analysis of the online time of a specific class of students. After entering the system, the student administrator selects the class first and then clicks the user analysis to select the date to be viewed to view the online time duration information of the class for the current week or the current month.

### 2.3 Database Design

The structural design of the database will directly affect the system's operating efficiency and display effect. According to the overall functional structure of the system, this paper has designed four data tables user information table, data table, id to sort table, and user table.

## 3 Campus Network Information Display System Implementation

### 3.1 JFreeChart Introduction

JFreeChart is a powerful graphics component. Through its own development package, it can draw various statistical graphs, such as column charts, pie charts, and fold lines it can also produce diagrams, such as area diagrams, and 2D or 3D diagrams. After the data is collected at the server, the dataset required by the chart is constructed. The server generates a temporary image file, returns the address of the image to the client, and displays it as a picture [4].

### 3.2 The Realization of the Chart Display Function

The requirements of the student administrator module and the network administrator module are different, one focuses on the surfing time while the other focuses on the Internet traffic. However, in terms of achieving technical consistency, this paper uses the student administrator module as an example to illustrate the implementation of the chart display function.

1. The line graph is implemented in the Student Administrator module, which provides a cartographic dataset via Java. The author uses the data structure language to operate the back-end database through the JDBC data interface, then extracts and organizes the data that meets the logical business requirements. Since JFreeChart's own data collection method cannot meet the requirements of the system, we adjust the dataset according to the needs of the system which the information that should be included in the line chart covers the time, the type of application, and the duration of the application. According to the JDBC data interface query time, application use time, and the application type of group statistics, we initialize XYSeries as the application name and then add each application's usage time to the XYSeries. Finally, add XYSeries to the XYSeriesCollection dataset. Create a line chart by using the CreateXYLineChart method in the line graph component ChartFactory provided by JFreeChart, set the background color of the picture by using the SetBackgroundPaint method in the XYPlot object, and set the thickness of the line in the line graph by using the XYLineAndShaperenderer object. Draw the temporary picture on the server side. Return the drawn picture address to the client browser for presentation to the student administrator [5].
2. The realization of the pie chart is based on the needs of the student administrator. The pie chart needs to show the proportion of various applications and the length of time

spent on the Internet. The JDBC data interface is used to query the type of application and the corresponding using duration of the application and group statistics corresponding to the type of application. We pass the query's access dataset to the DefaultPieDataset default dataset object in the JFreeChart component to complete the dataset addition. Then we use the createPieChart3D method in the JFreeChart component ChartFactory to create a pie chart and pass the created dataset as a parameter.

3. Bar chart implementation is based on the needs of the student administrator, and the bar chart needs to show the duration of various applications used in each time period. Compared to line and pie charts, the dataset design of the column chart is more complicated. First, we query the time and duration of the application through the JDBC data interface and group statistics by time and application category. Then we pass the query data according to the corresponding time through the addValue method of the DefaultCategoryDataset objectAdd to the dataset. Finally, we create a column chart using the createBarChart method in the JFreeChart component ChartFactory [6].

## 4 Conclusion

Statistics and analysis of students' online behavioral data mainly aim at the statistics and analysis of college students' Internet surfing behavior at ordinary times. They count the length of time students spent on various applications and the amount of Internet traffic. Students' online duration on applications allows school administrators to effectively understand students' daily information. Online habits and Internet traffic allow network managers to effectively understand the school network bandwidth usage, which can help to provide a reliable database for campus network optimization.

**Acknowledgments.** This work was supported by CERNET Innovation Project (NGII20160604), Tendering Project of the key disciplines of the Xinjiang Normal University under Grant 17SDKD1201 and NSFC (61771089).

## References

1. Zuhui H, Shi Y. Research on college students' online behavior analysis and data mining. Chin Distance Educ. 2017;02:26–32.
2. Bo L. Current status of college students online and countermeasures. Chin Foreign Entrepreneurs. 2014;15:152–3.
3. Lili Y, Mei N, Tingting Y. Personalized network teaching resource system. Comput Syst. 2012;21(06):174–7.
4. Wei H, Wanjun L. Application of JFreeChart in java web project. Sci Technol Eng. 2008;10:2699–701.
5. Ruifa Y, et al. Drawing web statistics chart based on JFreeChart the design and implementation of. Netw Secur Technol Appl. 2010;12:64–6.
6. Hongbo Wu. Web data diagram based on JFreeChart technology performance. Electron Sci Technol. 2010;23(01):97–9.

# Part II
# Digital Signal Processing

# Research on Data Flow Partitioning Based on Dynamic Feature Extraction

Wei Wang(✉) and Min Zhang

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
weiwang@tjnu.edu.cn

**Abstract.** With the rapid development of the Internet of things, social networks, and e-commerce, the era of big data has arrived. Although big data has great potential for many areas such as industry, education, and health care, getting valuable knowledge from big data can be a daunting task. Big data has the characteristics of high-speed change, and its content and distribution characteristics are in dynamic changes. Most current models are static learning models that do not support online updating, making it difficult to learn dynamically changing big data features in real time. In order to solve this problem, this paper proposed a method to support incremental recursive least squares (IRLS) regression parameter estimation and variable sliding window algorithm to analyze and judge the trends of dynamic characteristics of data streams, which can provide early warning, status assessment, and decision support for monitoring objects and improve the accuracy and adaptability of data flow classification. The real-time computational and analysis accuracy are obviously improved than the traditional algorithm, and the simulation results verify the effectiveness of the proposed algorithm.

**Keywords:** Trend analysis · Dynamic data mining · Incremental recursive least squares method · Variable sliding window

## 1 Introduction

Dynamic data flow refers to an ordered sequence of data consisting of a large number of continuously arriving, potentially infinitely long, fast-changing data with transient, real-time, and infinite characteristics. In recent years, data flow has been widely used in many fields such as stock exchange, telephone communication record, network traffic monitoring, sensor network, and so on. Data stream contains a lot of information, which can be used as the basis for intelligent decision-making. The purpose of feature extraction and description of dynamic data stream is to extract the change information and analyze the data stream signal to achieve decision support, such as stock market curve analysis, audio waveform analysis, ECG waveform recognition analysis, and industrial production curve identification. At present, however, most of the data streams are generated by variety of dynamic systems in real time, which is a high-speed incoming data, so the data stream data transmission, calculation, and storage become very difficult. Due to the different application fields, the characteristics of the dynamic


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_36

data flow curve vary widely, and most of them exist in a process system of nonlinear, multivariable coupling, events and large time delays. Therefore, the data flow trend analysis method requires strong adaptability and high analysis accuracy. The established data model can be applied to dynamic data flow analysis and prediction in various fields [1].

The traditional data flow curve trend analysis algorithm has sliding window (SW) algorithm [2–4], Extrapolation for Online Segmentation of Data (OSD) [5], bottom-up algorithm [4], and top-down algorithm [4], but they have obvious deficiencies. In the following, we give some shortcomings of SW and OSD algorithms in data segmentation:

A. There is no limitation to the maximum length of sliding window in SW algorithm. When the threshold of detection point is relatively large, the length of window may be too long and the fitting error becomes large. And the arriving data needs a complete regression modeling and data segment segmentation point judgment. As the data increases, the computational efficiency is extreme.
B. The OSD algorithm sets the minimum sliding window on the basis of the SW algorithm and only substitutes the newly arrived data into the established model to analyze the extrapolated cumulative error. Although the efficiency is improved, the transition point in the sliding window cannot be detected.
C. The SW and OSD algorithms have no limitation on the maximum length of the sliding window. When the detection threshold is relatively large, the length of the window may be so long that the trend analysis error becomes larger.
D. Both SW and OSD use the conventional least squares method for curve fitting. Compared to the conventional least squares method, incremental recursive least squares method is more efficient and has higher fitting accuracy.

In order to improve the precision of trend analysis and make the data stream feature more accurate, this paper improves the traditional data flow trend analysis algorithm and proposes incremental recursive least squares (IRLS) [6] linear regression modeling to get the model fitting parameters. Using variable sliding window algorithm [1] according to the trend of eigenvalues to split the data stream. The algorithm limits the lower limit of the length of a data stream. And the algorithm only recursively calculates the parameters of the regression model before the new data stream segmentation starts until the minimum segmentation length is reached, and does not perform segmentation point detection. The experimental results show that the proposed algorithm can get better fitting accuracy, can update fitting parameters in real time, can segment data segments online, and has high computational efficiency and small fitting error. It is suitable for the dynamic model of data. Moreover, based on the real-time trend analysis of data flow, the segment data segment can get better classification accuracy and recognition efficiency for the later target recognition processing. It is widely used in agriculture, commerce, health care, and other fields.

In this paper, a real-time algorithm that can adjust the trend parameters online according to the elements of the new arrival data stream is proposed. In addition, the variable sliding window algorithm is used to quickly detect the data stream segmentation points and determine the segmentation. The algorithm has high precision and low computational time complexity. The remainder of the paper is organized as follows. In

Sect. 2, the data flow problem will be described. Then, data stream linear regression modeling and variable sliding window algorithm will be introduced in Sect. 3. Experimental results for blocking the data flow will be shown in Sect. 4. The conclusion and discussion are in Sect. 5.

## 2 Data Flow Problem Description

The real-time trend analysis of data flow is based on real-time segmentation according to some statistical property indices (such as mean square error, cumulative error, and generalized likelihood ratio statistics), so that the data in the divided data segments is subject to the same statistical model. Segment obeys different statistical models.

In order to facilitate the study, we define the one-dimensional continuous data flow:

$$Y = \{\upsilon_{t_1}, \ldots, \upsilon_{t_i}, \ldots, \upsilon_{t_c}, \ldots\} \tag{1}$$

where $t_c$ is the current moment.

Data stream segmentation (this paper uses the mean square error comparison) to divide $Y$ into a series of continuous non-empty data segments (i.e., sliding windows): $\{Y_1, \ldots, Y_j, \ldots, Y_s, \ldots\}$.

The $j$th data segment is:

$$Y_j = \{\upsilon_{t_{j,1}}, \ldots, \upsilon_{t_{j,\lambda}}, \ldots, \upsilon_{t_{j,n_j}}\} \tag{2}$$

Corresponding data arrival time:

$$t_{j,\lambda} \in \{t_1, \ldots, t_i, \ldots, t_c, \ldots | j \in N, 1 \leq j \leq s; \lambda \in N, 1 \leq \lambda \leq n_j\} \tag{3}$$

In Eqs. (2) and (3), the length of the data segment $Y_j$ is denoted by $n_j$, that is, the length of the sliding window $Y_j$, and $t_{1,1} = t_1$. Note that the data segment $Y_s$ includes the data segment of the current data $\upsilon_{t_c}$.

Suppose the data in $Y_j$ can be fitted by linear regression model, that is:

$$\upsilon(t) = f(t, \theta_j) + \varepsilon_j(t), t \in \{t_{j,1}, \ldots, t_{j,n_j}\} \tag{4}$$

where $f(t, \theta_j) = a_j t + b_j$ is the linear regression model of data segment $Y_j$, $\theta_j = [a_j, b_j]^{\mathrm{T}}$ is the model parameter vector, parameter $a_j$ is the trend characteristic value of data segment $Y_j$, and $\varepsilon_j(t)$ is the independent and identically distributed zero-mean white noise. For the purpose of algorithm description, let the first data element $\upsilon_{t_{j+1,1}}$ of data segment $Y_{j+1}$ be the dividing point of $Y_j$.

The basic task of the real-time trend analysis of data stream is to perform the following calculation on the newly arrived data stream element $\upsilon_{t_c}$ based on the currently accepted data sequence $Y_{s,n} = \{\upsilon_{t_{s,1}}, \ldots, \upsilon_{t_{s,n}}\}$: (1) split point detection (such as detecting whether $\upsilon_{t_c}$ is used as the division point of data segment $Y_{s,n}$; (2) establish a regression model for the current data segment to calculate the current fitted model parameter values $a_j$ and $b_j$.

## 3 Data Flow Real-Time Trend Analysis Method

In order to overcome the shortcomings of existing trend analysis algorithms, this paper presents a new real-time trend analysis algorithm for data streams to overcome the drawbacks of high computational complexity of SW algorithm and low accuracy of OSD algorithm in traditional methods.

### 3.1 Incremental Recursive Least Squares Algorithm (IRLS)

In order to identify the time-varying system in real time, Zhou et al. proposed an incremental recursive least squares (IRLS) algorithm [7]. The algorithm can use the newly arrived data to correct the original model parameters to obtain new model parameters.

Let the current linear regression model constructed for the data sequence $Y_{s,n} = \{\upsilon(t_{s,1}), \ldots, \upsilon(t_{s,n})\}$ be $f(t, \theta_{s,n})$, and the data sequence be expressed as $Y_{s,n} = [\upsilon(t_{s,1}), \ldots, \upsilon(t_{s,n})]^{\mathrm{T}}$ in vector form, then $Y_{s,n} = U_{s,n} + \varepsilon_{s,n}$. Among them:

$$U_{s,n} = \begin{bmatrix} t_{s,1} & \cdots & t_{s,n} \\ 1 & \cdots & 1 \end{bmatrix}^{\mathrm{T}}, \mathrm{rank}(U_{s,n}) = 2 < n \tag{5}$$

$\varepsilon_{s,n}$ is the expected random error vector of zero. The parameter $\theta_{s,n}$ is estimated by the least squares method, even if $\theta_{s,n}$ is satisfied

$$\min \mu_{s,n} \left\| \varepsilon_{s,n} \right\|^2 = \left\| Y_{s,n} - U_{s,n} \theta_{s,n} \right\|^2 \tag{6}$$

Let $p_{s,n} = U_{s,n}^T U_{s,n}$, and $q_{s,n} = U_{s,n}^T Y_{s,n}$. Derivative of $\mu_{s,n}$ and make its initial derivative of 0, can be obtained $\theta_{s,n} = p_{s,n}^{-1} q_{s,n}$. When the data stream element $\upsilon(t_c)$ is reached, it is marked as $\upsilon(t_{s,n+1})$, and the current data sequence is expanded to $Y_{s,n+1}$, then there is:

$$p_{s,n+1} = U_{s,n+1}^T U_{s,n+1} = U_{s,n}^T U_{s,n} + U_{n+1}^T U_{n+1} = p_{s,n} + U_{n+1}^T U_{n+1} \tag{7}$$

where $U_{n+1} = [t_{s,n+1}1]$. Similar to Eq. (7), for $q_{s,n+1}$ there is:

$$q_{s,n+1} = U_{s,n+1}^T Y_{s,n+1} = q_{s,n} + U_{n+1}^T \upsilon(t_{s,n+1}) \tag{8}$$

In this case, the parameter vector of the regression model satisfies the recurrence equation

$$\theta_{s,n+1} = p_{s,n+1}^{-1} q_{s,n+1} = \theta_{s,n} + p_{s,n+1}^{-1} U_{n+1}^T (\upsilon(t_{s,n+1}) - U_{n+1}\theta_{s,n}) \tag{9}$$

In order to avoid the inversion in Eq. (9), a matrix inversion lemma is introduced

$$[A + BCD]^{-1} = A^{-1} - A^{-1}B[DA^{-1}B + C^{-1}]^{-1}DA^{-1} \tag{10}$$

Let $A = p_{s,n}, B^T = D = U_{n+1}, C = 1, \beta_{s,n} = p_{s,n}^{-1}$, available

$$\beta_{s,n+1} = p_{s,n+1}^{-1} = \beta_{s,n} - (1 + U_{n+1}\beta_{s,n}U_{n+1}^T)^{-1}\beta_{s,n}U_{n+1}^T U_{n+1}\beta_{s,n} \tag{11}$$

The initial value $\beta_{s,0} = \varsigma I$, $\varsigma$ is a positive number, and $I$ is $2 \times 2$ unit matrix. When the data sequence noise is larger, $\varsigma$ takes a smaller value; on the contrary, $\varsigma$ takes a larger value. Substituting (11) into (9) yields a recursive formula for regression model parameters: $\theta_{s,n+1} = \theta_{s,n} + \Delta\theta_{s,n}$

$$\text{Where} \quad \Delta\theta_{s,n} = \beta_{s,n}U_{n+1}^T(1 + U_{n+1}\beta_{s,n}U_{n+1}^T)^{-1}(\upsilon(t_{s,n+1}) - U_{n+1}\theta_{s,n}) \tag{12}$$

where $(1 + U_{n+1}\beta_{s,n}U_{n+1}^T)$ is a scalar, thus avoiding matrix inversion. The initial value $\theta_{s,0}$ of Eq. (12) is generally taken as zero vector.

### 3.2 Variable Sliding Window Algorithm

The maximum length of the sliding window in the SW algorithm is not limited. When the detection threshold is relatively large, the length of the window may be so long that the trend analysis error becomes larger. However, the OSD algorithm defines the minimum sliding window length so that the mutation in the minimum sliding window point cannot be detected.

Aiming at the shortcomings of sliding window in SW and OSD algorithms, this paper presents a variable sliding window algorithm which dynamically changes the setting window length to segment the data sequence reasonably. The algorithm first sets

the length of reference window and the length of the longest data window. Starting from the starting point of the current data segment, regression modeling is re-established for each newly arrived data stream element to improve accuracy. When the current data segment is smaller than the length of the reference window, the fitted standard deviation of the model is compared with the return value of the noise function $G$ to detect whether there is an abnormal point in the reference window. If the length of the current data segment is greater than or equal to the reference window length, when the mean square error is greater than the standard segmentation threshold set in advance, the newly arrived data is considered as the segmentation point of the current data segment. If the length of the current data segment is greater than the length of the longest data window, the data point that has the closest fitting square variances to the standard segmentation point threshold is searched as the dividing point of the current data segment from the beginning of the current data segment. The variable sliding window algorithm solves the fixed window problem of the SW algorithm and the OSD algorithm and realizes the reasonable segmentation of the data stream, so the accuracy of the trend analysis is improved.

## 4 Experiments Results

In order to verify the validity of the proposed algorithm, the monitoring data collected in 60 days from 54 mica2 sensor nodes in Intel Berkeley Research Laboratory's sensor network laboratory is used (mica2dot nodes: processor is Atmegal28L, RF chip is cc1000, flash is 128 kb, RAM is 4 kb, and the transmission rate is 76.8 kbps). The data set is sampled every 31 s by temperature, humidity, light intensity, and node voltage. The data is collected in the TinyDB network query processing system, and the system is built on the TinyOS platform. Sensor distribution was shown in Fig. 1, and gives the original data plot of individual sensors (Fig. 2) and overlay sensors (Fig. 3).

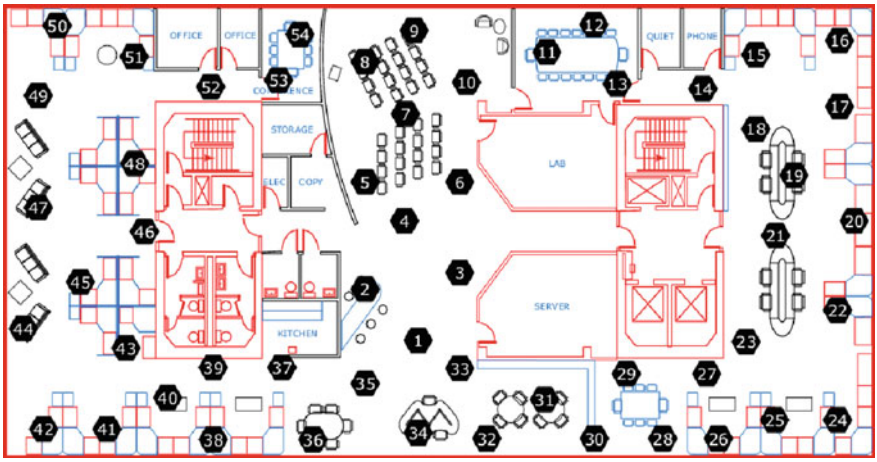



**Fig. 1.** Sensor layout

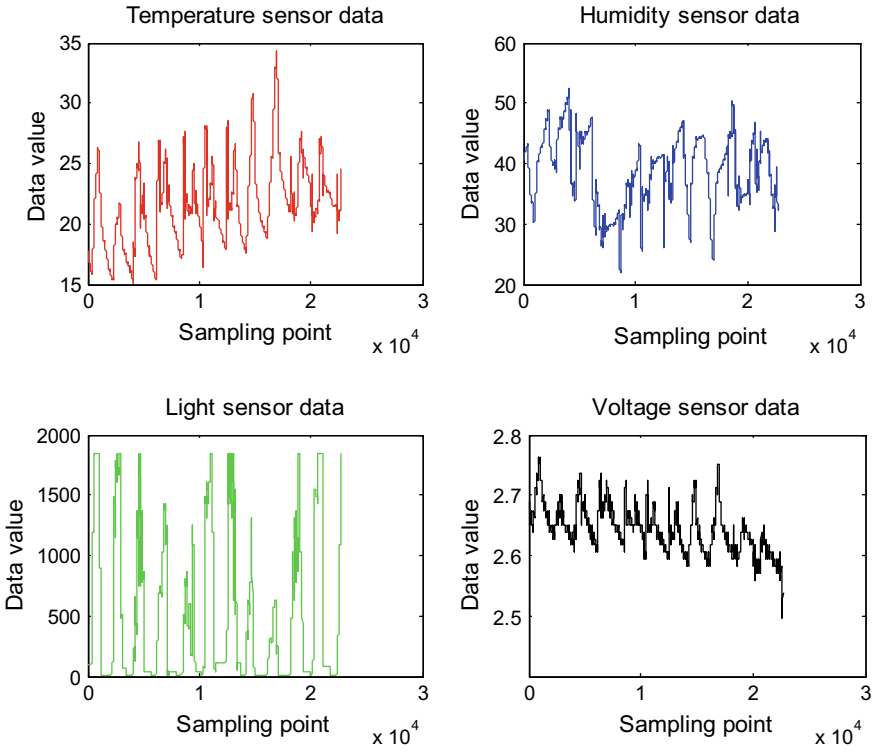



**Fig. 2.** Original data plot of individual sensors

The hardware configuration of the emulation computer is the main frequency p42.5GHZ, and the RAM is 512M. The software environment is matlab r2014a under the 64-bit win10 operating system. The sampling frequency of the four sensors is the same and we combine the sensor data obtained by the block to improve the accuracy of target recognition and accurately determine the environment suitable for spring, summer, autumn and winter.

Therefore, we use the temperature and humidity sensor to determine the segment length. The number of data points in the segment is mainly, and other sensors reasonably set the threshold according to the number of data points of the temperature and humidity sensor. The four sensor data at the same time are respectively modeled by incremental recursive least squares regression, and the appropriate reference window length, maximum data window length, split point threshold and noise function are set in the field, and the split point is detected by the dynamic sliding window. The four sensors have 22,673 data, respectively. For the sake of clarity, only the first 800 data points are taken.

Experiments and SW algorithm are used to analyze the trend of the same data stream, and the two algorithms are compared from the fitting accuracy, fitting error and algorithm running time. This experiment only gives the fitting effect of temperature, humidity and illumination sensor. Since the voltage value changes with temperature, it is not mentioned here as an auxiliary information in the later target recognition.

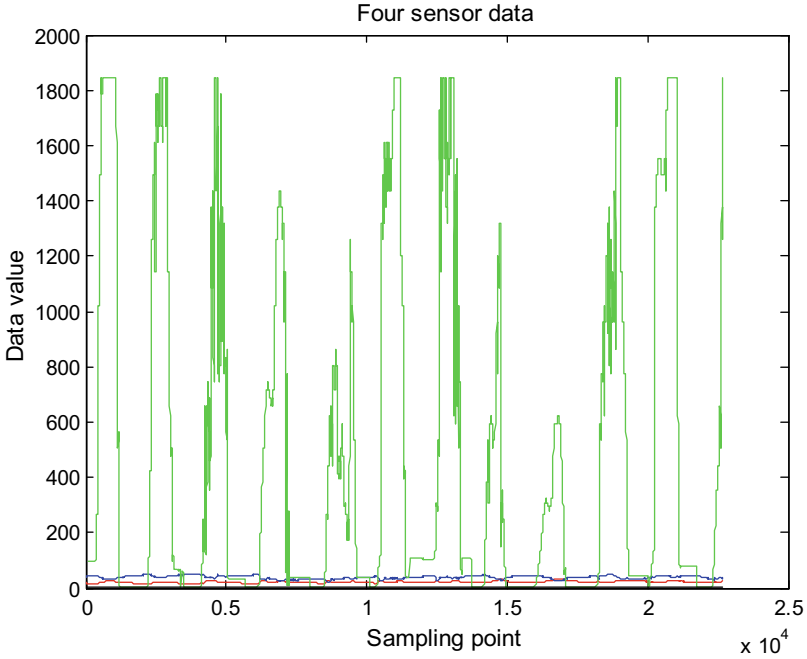



**Fig. 3.** Original data plot of overlay sensors

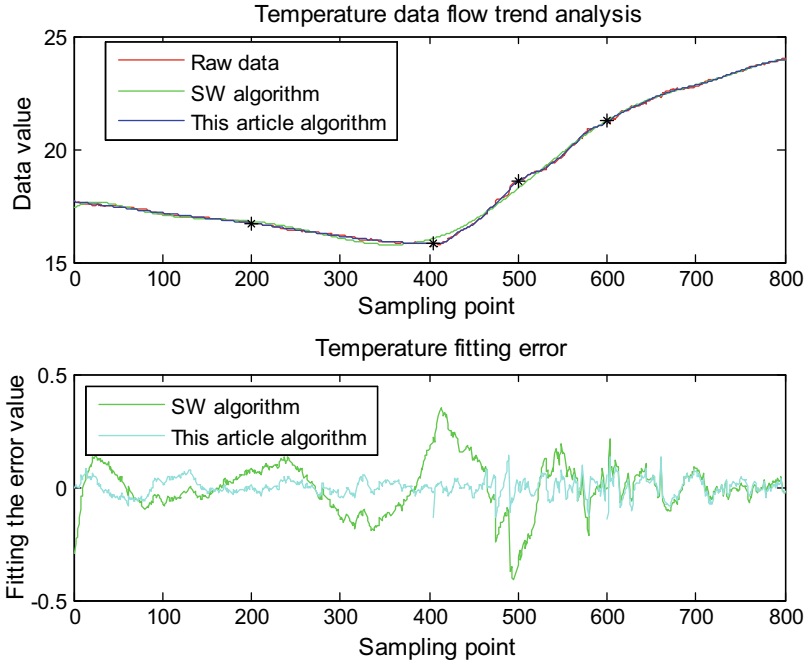



**Fig. 4.** Temperature data flow trend analysis

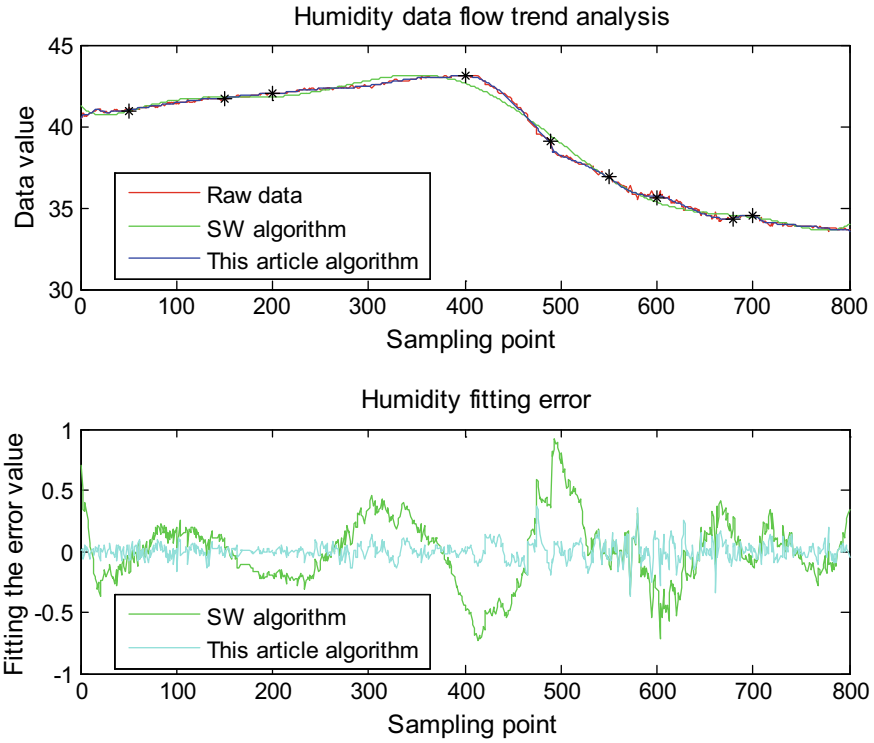



**Fig. 5.** Humidity data flow trend analysis

Figures 4, 5 and 6 respectively show the fitting effect of temperature, humidity and illumination. It can be seen that compared with SW algorithm, the algorithm can better approximate the original data, and the fitting error value can be clearly obtained. The fitting effect of the algorithm is better, and the obtained fitting parameters are more accurate. This shows that the segmentation points determined by the algorithm are reasonable and the trend analysis accuracy is the highest. Table 1 shows the running time of the algorithm. It can be seen that the algorithm has low computational complexity and high efficiency.

## 5 Conclusions

In this paper, incremental recursive least squares method and variable sliding window algorithm are combined to propose a data stream real-time trend extraction algorithm. In order to solve the data flow elements that arrive constantly, this algorithm uses incremental mechanism to determine the data sequence regression model parameters and dynamically segment data segments, and extract the trend characteristics of data flow in real time. The algorithm not only calculated faster, but also higher accuracy. After this method is used to segment the data stream, the data can be extracted later to

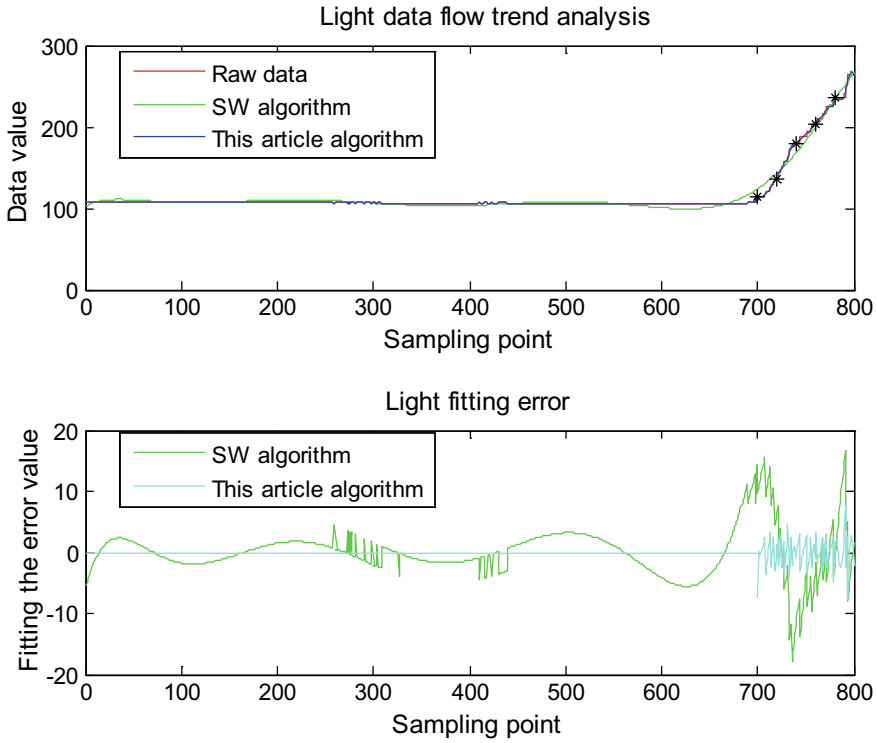



**Fig. 6.** Light data flow trend analysis

**Table 1.** Three sensors' linear modeling runtime based on two algorithms

| Data | This article algorithm (s) | Sliding window (SW) (s) |
|---|---|---|
| Temperature sensor data | 0.1719 | 1.6296 |
| Humidity sensor data | 0.1652 | 1.6699 |
| Light sensor data | 0.0788 | 1.9244 |

analyze and process a piece of data, which provides a good solution for some algorithms that cannot handle dynamic data nowadays, and lays a good data foundation for later target recognition.

**Acknowledgments.** This paper is supported by Natural Youth Science Foundation of China (61501326, 61401310) and Natural Science Foundation of China (61271411). It also supported by Tianjin Research Program of Application Foundation and Advanced Technology (15JCZDJC31500) and Tianjin Science Foundation (16JCYBJC16500).

## References

1. Wang C, Pang X, Lu Z, et al. Research on data flow classification based on dynamic feature extraction and neural network. Comput Syst Appl. 2010;30(6):1539–42.
2. Koski A, Juhola M, Meriste M. Syntactic recognition of ECG signals by attributed finite automata. Pattern Recogn. 1995;28(12):1927–40.
3. Shatkay H, Zdonik S. Approximate queries and representations for large data sequences. In: Proceedings of 12th IEEE international conference on data engineering. Washington: IEEE Computer Society; 1996. p. 546–53.
4. Keogh E, Chu S, Hart D, et al. Segmenting time series: a survey and novel approach. In: Proceedings of IEEE international conference on data mining. Los Jose: IEEE Computer Society; 2001. p. 289–96.
5. Sylvie C, Carlos GB, Cathering C, et al. Trends extraction and analysis for complex system monitoring and decision support. Eng Appl Artif Intell. 2005;18(1):21–36.
6. Zhou Q, Wu T. Research and application of a data flow trend analysis method. Control Decis Mak. 2008;23(10):1182–5.
7. Zhou Q, Cluett W. Recursive identification of time-varying systems via incremental estimation. Automatica. 1996;32(10):1427–31.

# Digital Signal Processing Technology of DVOR Navigation Signal

Zhengbo Yang(✉), Jiaquan Ye, Jing Liu, Ping Yang, and Fei Liang

The Second Research Institute of CAAC, Chengdu, China
yangzhengbo@caacsri.com, l96l02849@qq.com

**Abstract.** In order to make the civil aviation navigation system digital and miniaturization, a digital signal processing technology of applying the Doppler VHF Omnidirectional Range (DVOR) is proposed to improving the processing speed and the performance of the receiver data. This paper introduces the basic principle of DVOR signal and the basic principle and realization method of digital signal processing technology. Finally, we completed the design of digital signal demodulation, filtering, and comparison phase by LabVIEW 2016, so that the DVOR measure system can display information such as bearing.

**Keywords:** DVOR · Digital signal processing · Phase comparison

## 1 Introduction

Civil aviation navigation equipment is the most important infrastructure of civil aviation and is closely related to flight safety. With the continuous growth of flight flow and the implementation of large-scale airport Cat II and Cat III instrument landing system carry out, the airport operation has become more and more dependent on the navigation facilities. The operation detection of navigation equipment is an important means to ensure the aviation safety. The traditional test system uses to be built by the separate instruments. But this kind of system is costly, low automatically, and poor scalability. With the development of civil air transport, modern navigation receivers generally combine the functions of receivers such as DVOR, LOC, GS, and MB. Most airborne flight electronic devices are highly digitized and integrated, which cannot be manually checked by human anymore. So, the developed countries in the world have been using the automatic test equipment to complete such works. The digital signal processing technology uses LabVIEW and signal receiver to achieve the analysis of navigation signals. The basic idea of software radio is based on a common, standard, and modularized hardware platform to achieve the various functions of the radio station through the programming of software. And it will be produced by the design method of hardware-based, application-oriented. The automatic measure system puts forward higher requirements for the flexibility and comprehensiveness of the signal source, which is hard to be meeting by the traditional one. By contrast, the signal receiver based on the software radio has a more outstanding advantage and can suit the need more thoroughly because of its flexibility, openness, and other characteristics.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_37

## 2 DVOR Signal

The basic function of the Doppler VHF Omnidirectional Range is to provide a complex radio signal consisting of a 30-Hz reference phase signal and a 30-Hz variable phase signal and the corresponding RF signals. The complex radio signals are received by the airborne DVOR receiver and demodulated by demodulation module, and the measured azimuth angle of the aircraft relative to the ground VHF omnidirectional range is the DVOR azimuth. DVOR operating frequency range is 108–117.95 MHz, and the distance between channels is 0.05 MHz.

### 2.1 DVOR Reference Phase Signal

DVOR reference phase signal, which is the 30-Hz low-frequency signal amplitude modulation directly on the carrier frequency $f_0$, was transmitted by the high-frequency signal through the central antenna in the center of the rotating antenna to the space generated by the radiation, including only simple of the RF carrier signal, and the expression is:

$$U_R = U_{RM}[1 + m_{\mathrm{R}} \cos(2\pi Ft)] \cos 2\pi f_0 t \tag{1}$$

where $m_{\mathrm{R}}$ is the 30-Hz reference signal on the RF carrier amplitude modulation factor, $F$ is the baseband signal frequency, and $f_0$ is the RF frequency.

The intensity of the radiation signal and the phase of the 30-Hz signal are unchanged in the radial azimuths of the reference phase signal in the space 0°–360°, and the horizontal direction of the radiation field is a circle.

### 2.2 DVOR Variable Phase Signal

DVOR variable phase signal includes RF carrier signal and 9960-Hz subcarrier signal. 9960 Hz subcarrier is 30 Hz baseband signal frequency modulation, modulation index is 16, two pairs of orthogonal sideband antenna respectively radiation sinusoidal modulation sideband wave and cosine modulation sideband wave, the field strength changes according to the 30 Hz law, so it is generate a 30 Hz sine law changes in the amplitude modulation wave in space, the expression is:

$$U_V = U_{VM}[1 + m_V \cos(2\pi f_{\mathrm{s}} t + m_f \cos(2\pi Ft + \alpha))] \cos 2\pi f_0 t \tag{2}$$

where $m_V$ is the amplitude modulation index of the subcarrier to the RF carrier, $f_{\mathrm{s}}$ is the subcarrier frequency, $m_f$ is the modulation index of the subcarrier for the 30-Hz baseband signal, $F$ is the baseband signal frequency, $\alpha$ is the current VOR radial azimuth, and $f_0$ is the RF frequency.

### 2.3 Synthetic Signals

The composite signal received by the receiver in space includes the superposition of the reference phase signal and the variable phase signal, as shown in Fig. 1, and the expression is:

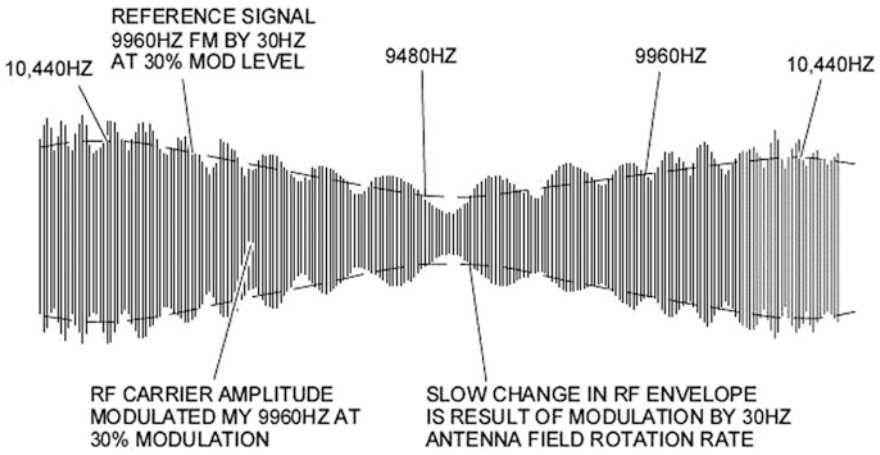



**Fig. 1.** Spatial synthesis signal

$$\left[1 + m_{\mathrm{R}} \cos(2\pi Ft) + m_V \cos(2\pi f_s t + m_f \cos(2\pi Ft + \alpha))\right] \cos 2\pi f_0 t \tag{3}$$

The receiver obtains the direction information by demodulating and comparing the phase difference between the two.

## 3 Systems

The digital signal processing function of digital filtering, digital extraction, digital down-conversion, digital demodulation, and data processing, which is based on software radio peripheral equipment, completed with large capacity FPGA to directly digital signal processing. For this kind of software radio system, a real station cannot be observed from the entity, and it solves the radio communication problems from the angle of software. The universal software radio is the closest to the ideal software radio and is more suitable for testing, teaching, and research because of its openness, flexibility, programmability, and man–machine interface.

### 3.1 Hardware Part

Software radio peripheral equipment which simulates the instrument landing system in the airborne signal receiver is used to receive the navigation signal from the ground-based navigation aid, through high-speed continuous extraction and processing to achieve digital signal processing in order to facilitate the modular processing on the computer and data analysis. The hardware part mainly includes crystal oscillator, amplifier, quadrature demodulator, low-pass filter, analog-to-digital converter, digital down converter, and other modules to ensure the collected signal which is not distorted. And the hardware also needs to guarantee the matching degree of automatic work and the realization of continuous signal acquisition. It has resolved the traditional problems of instability and in continuity effectively.

### 3.2 Software Part

LabVIEW, which is a graphical programming language, is a measure system software development platform. It uses the program method of data stream; the data streams between the nodes in program block diagram determine the order of the program. Signal generation, signal acquisition, and digital signal processing are performed in the LabVIEW 2016 environment.

Digital signal processing in the LabVIEW program connects according to the form of pipeline, and it is consistent with the signal flow direction, so it has a very high efficiency, short delay, and high processing speed. To a certain extent, it has made up for the low speed of the general processor processing signal. The flow chart is shown in Fig. 2.

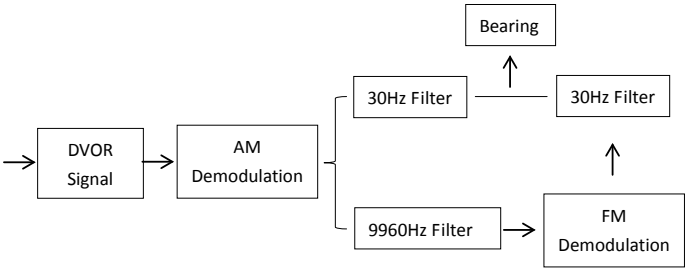



**Fig. 2.** System block diagram

### 3.3 Data Processing

Compared with the traditional digital signal processing method, the program based on the LabVIEW development by using the pipeline form connection has a high efficiency and a higher processing speed. The navigation signal uses software radio peripherals to simulate DVOR signal transmission. In the waveform generation program, the reference phase signal and the variable phase signal are simultaneously loaded on the RF

carrier, and the spatial synthesis of the DVOR signal is simulated and transmitted through the transmitting antenna. DVOR receiver part mainly consists of receiving antenna, RF receiving front end, digital down-conversion, amplitude demodulation, audio signal demodulation, and azimuth signal demodulation and other modules. The received DVOR signal turns into zero IF signal through the amplification of high frequency and the down-conversion which achieves the analysis of azimuth through the down sampling, band-pass filter, amplitude demodulation and frequency demodulation, and other processing methods.

## 4 Results

According to the design of the hardware structure to connect the hardware, set the waveform information in the program interface. The baseband signal is 30-Hz sine wave, FM subcarrier is 9960 Hz sine wave, frequency offset is 480 Hz, and modulation index is 0.3. Considering the machine performance and running time, the carrier frequency of the DVOR analog signal is 1 MHz as an example. In the waveform generator, it needs to initialize the device and adjust the parameters firstly, then, set the IP address, the signal amplitude is 1 V, and the center frequency is 1 MHz. The receiver write data and the generator transmit data should select the same sampling rate, in order to ensure the spectrum which the generation of the signal is correct.

The test signal is generated by software radio simulation and program programming. In the program, we need to set the virtual channel Tx1, and the sampling frequency must meet the Nyquist theorem; so here, we select 2 MHz and the sampling clock is set to "continuous samples," which will continue 1 s. The sampled data can be displayed in real time in the waveform chart and transmitted continuously via a software radio in order to receive the antenna reception.

In order to realize the measurement and verification of the signal, it is necessary to demodulate the acquired waveform and restore the corresponding signal. The collected DVOR space synthesis signal is coherently demodulated to obtain the outer envelope of the spatial synthesis signal, including the reference phase signal of 30 Hz, the 9960 Hz FM subcarrier, and the 1020-Hz audio identification signal. The 30-Hz reference phase signal is directly obtained through the 30 Hz filter. The front panel of the program can be seen in its time domain waveform and frequency domain waveform and show frequency, phase, and amplitude of the 30-Hz reference phase signal. 9960 Hz subcarrier is filtered and demodulated to obtain 30-Hz variable phase signal. Its corresponding parameters can be seen in the front panel. The phase difference between the reference phase signal and the variable phase signal indicates the current azimuth information, and the reading azimuth is the DVOR bearing. After running, the front panel is shown in Fig. 3. When the program is running, the reference phase signal and variable phase signal can display the corresponding frequency and amplitude information. BEARING displays bearing information, and press the right corner of the "STOP" button to lock the current tab page.

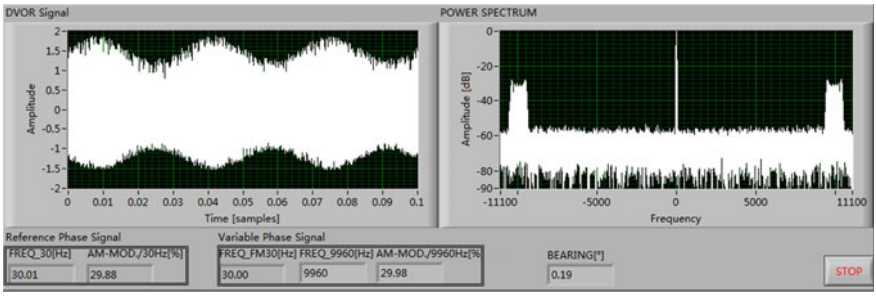



**Fig. 3.** Test the front panel

## 5 Conclusions

At present, software radio technology has been widely used in the field of military and civilian communications. Compared with the traditional equipment, this kind of test equipment has a broader application prospect because of its openness and flexibility. While the digital signal processing is the key to the transition from the analog signal to digital signal. Through the practical test and verification of the navigation equipment, the program based on the LabVIEW can achieve the test and analysis of each parameter by combining the software radio equipment, which is convenient, efficient, and handy. The test system not only can replace the original traditional test system, but also can always record and view the test data, and it is easy to call and match at any time.

**Acknowledgments.** This paper is supported by the National Key R&D Plan (2017 YFB0503402). The authors would like to thank the anonymous reviewers and editors for their helpful comments and suggestions.

## References

1. DVOR4000 technical manual description, operating, maintenance. Italy: Thales ATM S.p.A.; 2000
2. Ni Y, Lu D, Wang Y, Cui M. Navigation principles and systems; 2015.
3. Doppler VOR VRB-52D, Interscan Navigation Systems Pty. Ltd.; 2016.
4. Aeronautical radio navigation aids Part 2: technical requirement for very high frequency omnidirectional range (VOR), MH/T 4006.2; 1998.
5. AWA Defence Industries Pty Ltd. Doppler VOR VRB-51D, vol 9. Australia; 1990.

# A Vibration Signal Preprocessing Method for Hydraulic Pump Based on ILCD Fusion

Han Dong[1(✉)] and Sun Jian[2]

[1] Shijiazhuang Campus of Army Engineering University, Shijiazhuang, China
han58228@163.com
[2] Luoyang Electronic Equipment Experiment Center, Luoyang, China

**Abstract.** Hydraulic pump vibration signal preprocessing is the basis for failure prediction. The vibration signal preprocessing method based on ILCD fusion is proposed to solve the problem that the vibration signal is nonlinear and the feature information is weak. Firstly, the high-frequency harmonics are combined with ILCD to decompose multi-channel vibration signals, and ISC components can be achieved. Secondly, the sensitive factors are defined as evaluation indexes, and the sensitive components containing the fault feature information are screened and weighted fusion. The reconstructed signal is obtained to reduce the noise and other interference components and effectively extract the fault feature information. Finally, the effectiveness of this method is verified by measuring the vibration signal of the hydraulic pump.

**Keywords:** Signal preprocessing · ILCD · Information fusion · Hydraulic pump

## 1 Introduction

The vibration signal of the hydraulic pump presents strong nonlinearity and non-stationary in the event of a fault. In order to solve the distortions of some of the ISC components in local characteristic-scale decomposition (LCD), Jinde et al. [1] proposed ILCD method, which uses the piecewise polynomial instead of the original linear LCD connection, closer to the envelope, effectively reducing the appearance of distortion, the decomposition component, and the actual component closer. However, during the ILCD, some noise and interference components are still randomly distributed in some ISC components, affecting the extraction of feature information. In this paper, the high-frequency harmonic algorithm is introduced into ILCD. Firstly, the high-frequency harmonics and ILCD are combined to decompose the multi-channel vibration signals to obtain the corresponding ISC components. Secondly, the sensitive factors are defined, the ISC components are screened, and the sensitive components which contained feature information are obtained. ILCD fusion algorithm is proposed, which is to be fused to obtain the reconstructed signal. Finally, the validity of the proposed method is verified by measuring the vibration signal of the hydraulic pump.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_38

## 2 ILCD on Vibration Signals

Considering the characteristics of signal decomposition by ILCD, this paper combines the high-frequency harmonic [2] algorithm with the ILCD algorithm in Ref. [1] to obtain ISC components and focus the anomalous components in the first ISC component as much as possible, so that the sensitive components can be easily selected. Specific methods are as follows:

Step 1 Construction of high-frequency harmonic signals:

$$H(t) = U\cos(2\pi f) \tag{1}$$

$$x'(t) = x(t) + H(t) \tag{2}$$

$x(t)$ is the original signal, $H(t)$ is the high-frequency harmonic component, $U$ is the amplitude, usually take $U = \max(x(t))$; $f$ is the frequency, usually take $f = f_s/2$.

Step 2 Determine $x'(t)$ extreme point $X_k$ and the corresponding moment $\tau_k$, $k = 1, 2, \ldots, M$, where $M$ is the extreme point.

Step 3 Use segmented triad to connect the adjacent extreme, and calculate $L_k$:

$$L_k = a_{k-1}X_{k-1} + a_kX_k + a_{k+1}X_{k+1}, k = 2, 3, \ldots, M-1 \tag{3}$$

$a_k = 0.5$,

$$a_{k-1} = \frac{-2\tau_k^3 + 3\tau_{k-1}\tau_{k+1}^2 - 6\tau_{k-1}\tau_k\tau_{k+1} + 3\tau_k^2(\tau_{k-1} + \tau_{k+1}) - \tau_{k+1}^3}{2(\tau_{k-1} - \tau_{k+1})^3},$$

$a_{k+1} = -a_{k-1}$. By extending the left and right endpoint values, calculate the $L_1$ and $L_M$ using Eq. (3).

Step 4 Fitting $L_1 \sim L_M$ by cubic spline function, get the baseline $BL_1(t)$;

Step 5 Separate baseline:

$$h_1(t) = x(t) - BL_1(t) \tag{4}$$

Step 6 If $h_1(t)$ meets the ISC component condition [3], output $\mathrm{ISC}_1 = h_1(t)$; otherwise, $h_1(t)$ is taken as the original data, repeat Step 1–4, and separate the baseline:

$$h_{11}(t) = h_1(t) - BL_{11}(t) \tag{5}$$

Step 7 If $h_{11}(t)$ still does not meet the ISC conditions, then continue to repeat Steps 1–4, cycle $k$ times, until the conditions are met, output $\mathrm{ISC}_1 = h_{1k}(t)$;

Step 8 Separating the $\mathrm{ISC}_1$ component from the original data, achieve the new signal $u_1(t)$:

$$u_1(t) = x(t) - \mathrm{ISC}_1 \tag{6}$$

Step 9 Take $u_1(t)$ as the original data, repeat Steps 2–8, and get $\mathrm{ISC}_2$. Loop $n$ times until it is monotonous or constant for $u_n$, resulting in $\mathrm{ISC}_1, \mathrm{ISC}_2, \ldots, \mathrm{ISC}_n$ and trend $u_n(t)$.

By decomposing the signal using LCD, the characteristic information of the signal can be refined into each ISC component to facilitate the capture of sensitive information.

## 3 Vibration Signal Preprocessing Based on ILCD Fusion

### 3.1 Sensitive Component Selection

In general, the correlation coefficient is often used to reflect the linear relationship between variables, and mutual information can reveal the nonlinear relationship between variables [4, 5]. There are many ways to calculate it. In this paper, the equations based on Shannon entropy are used:

$$\begin{cases} I(y_1, \ldots, y_n) = \sum_{i=1}^{n} H(y_i) - H(y_1, \ldots, y_n) \\ H(y_i) = -\int p(y_i) \lg p(y_i) \mathrm{d}y_i \end{cases} \tag{7}$$

$H(y_i)$ is the entropy of the variable $y_i$, and the $p(y_i)$ is the probability density function of the variable $y_i$. For solving the joint probability density function, the probability of the sample point falling in the multi-dimensional space can be used to approximate the estimation.

Based on mutual information, the sensitivity factor $\lambda$ as a screening indicator is constructed. With the same sampling parameters' setting, the normal signal is $y_{\mathrm{nor}}$, the fault signal is $y_f$, and the $\mathrm{ISC}_i$ is the No.$i$ ISC components through the fault signal ILCD, and then the sensitivity of the $\mathrm{ISC}_i$ can be defined as:

$$\lambda_i = I_{1i} - I_{2i} \tag{8}$$

$I_{1i}$ is mutual information of $\mathrm{ISC}_i$ and $y_f$, and $I_{2i}$ is mutual information of $\mathrm{ISC}_i$ and $y_{\mathrm{nor}}$. Through analysis, $\lambda_i$ can quantitatively reflect the amount of sensitive feature information contained in $\mathrm{ISC}_i$. The higher the value of $\lambda_i$ is, the bigger difference of the $I_{1i}$ and $I_{2i}$ are, and the more fault sensitive features the $\mathrm{ISC}_i$ contains. Therefore, the ISC component corresponding to the maximum value of the sensitivity factor $\lambda$ is the sensitive component.

In practical application, it is impossible to know whether the signal under test is normal or faulty in advance. However, for the same batch of samples of the same type, the vibration signal data under normal conditions can be measured and stored; that is, the $y_{\mathrm{nor}}$ in the Eq. (8) can be obtained. Therefore, for a signal measured in the current state of a certain sample, if it is a normal signal, after ILCD, the $I_{1i}$ and $I_{2i}$ are completely identical according to Eq. (8); that is, the sensitivity factors of all ISC

components are 0, which means there is no ISC component that contains fault sensitive information. Correspondingly, the energy is evenly distributed in each frequency band for the normal signal, and the ISC components obtained by the ILCD do not contain the fault sensitive features. For the fault signal, its energy is concentrated in a few fault characteristic frequency bands; some ISC components can acquire fault sensitive information after ILCD. Therefore, the mutual information method used to select the ISC components can effectively remove the noise component and interference component and capture the sensitive information better, so as to improve the descriptive ability of the degraded state.

### 3.2 Vibration Signal Fusion-Based ILCD

In this paper, the three-direction vibration signal is taken as an example to illustrate the algorithm. According to the sensitivity factor $\lambda$ defined in Sect. 3.1, the higher the $\lambda$ is, the stronger its ability to express the characteristic information and the greater the contribution to the characteristic. Therefore, the fusion weight of the selected ISC component can be calculated on the basis of $\lambda$. Suppose the $\mathrm{ISC}_{i-j}$ is the sensitive component selected by the No.$i$ signal, $\lambda_i$ is the corresponding sensitivity factor, $i = 1, 2,$ and 3, so the fusion weight $k_i$ can be got.

$$k_i = \lambda_i / \mathrm{sum}(\lambda_1 + \lambda_2 + \lambda_3) \tag{9}$$

The weighted fusion reconstructed signal $x_{\mathrm{Final}}$ can be described as:

$$x_{\mathrm{Final}} = k_1 x_1 + k_2 x_2 + k_3 x_3 \tag{10}$$

According to the previous analysis, the specific steps of vibration signal fusion method-based ILCD are as follows:

① Acquisition of vibration signals $X_{x\mathrm{nor}}$, $X_{y\mathrm{nor}}$, $X_{z\mathrm{nor}}$ and $X_x$, $X_y$, $X_z$ of the hydraulic pump under normal conditions and fault conditions, using the ILCD algorithm with high-frequency harmonics. According to Eqs. (1)–(6), resolve $X_x$, $X_{y,}$ and $X_z$ separately to get the corresponding ISC component set;

② Using Eq. (8), the sensitivity factor $\lambda$ of each ISC component of $X_x$ signal is calculated, and the maximum $\lambda$ is selected as the sensitive component of $X_x$ signal. The sensitivity factor of each ISC component of $X_y$ and $X_z$ is calculated separately to obtain the corresponding sensitive component ISC and the corresponding $\lambda$;

③ Using $\lambda$ corresponding to each sensitive component, calculate the corresponding weight $k_i$ according to Eq. (9);

④ According to Eq. (10), reconstructed signal $x_{\mathrm{Final}}$ is carried out by using $k_i$, which is the result of the preprocessing of the hydraulic pump vibration signal.

## 4 Experimental Data Verification

In order to verify the effectiveness and practicability of the proposed method, this method is applied to the vibration signal preprocessing of hydraulic pump loose slipper fault (shown in Fig. 1). Hydraulic pump type is SY-10MCY14-1EL, a total of seven plungers, drive motor model Y132M-4, rated speed of 1480 rp/min, and the cycle of 0.041 s. The vibration accelerometer is installed at the pump cover (shown in Fig. 2), sampling frequency is 12 kHz, and time is 1 s. The DH-5920 vibration signal test and analysis system are used to store the vibration signal into the computer, as shown in Fig. 3.


**Fig. 1.** Fault of loose slipper

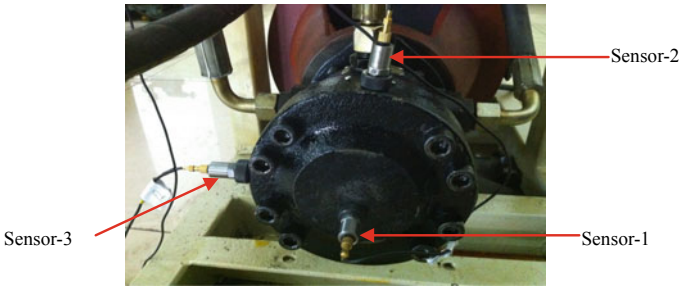



**Fig. 2.** Installation of vibration sensors on hydraulic pump

The vibration signal collected by the sensors Sensor1–Sensor3 is recorded as X1–X3. Figure 3 shows the time domain diagram of X1–X3, whose frequency domain is shown in Fig. 4.

Through analysis of the results in the time domain, the seven shock vibration signals in each period, which is mainly decided by the structure of the hydraulic pump. In the frequency domain, there are obvious modulation phenomena of vibration signal. Since the actual speed of the hydraulic pump shaft is 1480 rpm/min, the shaft

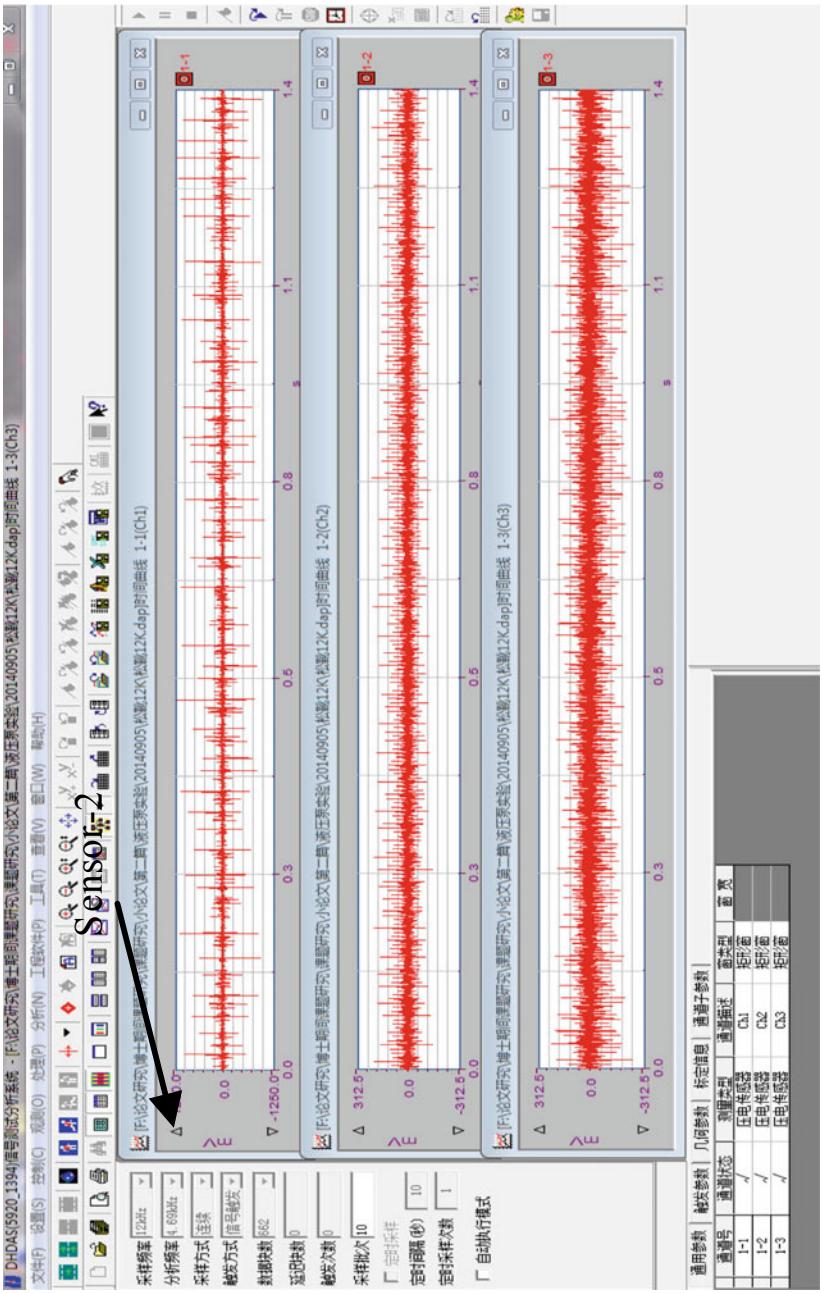
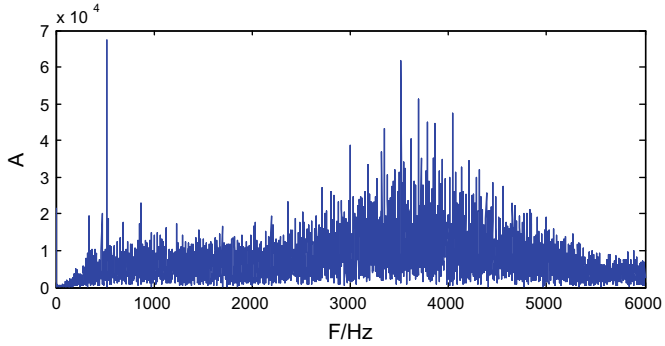
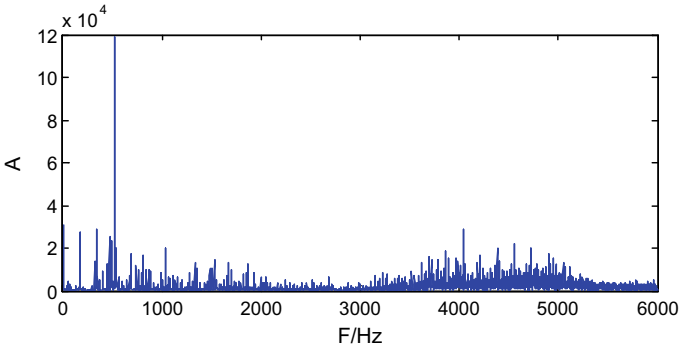


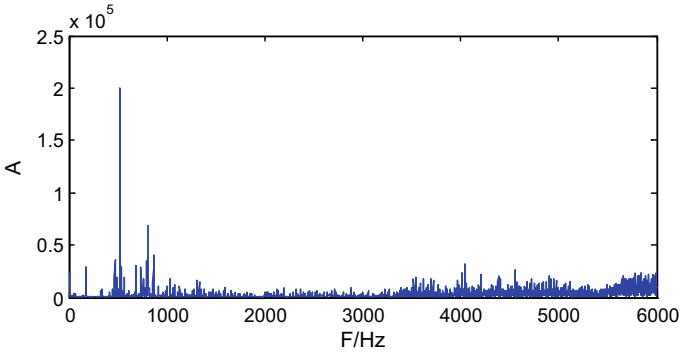

**Fig. 3.** Acquisition of multi-channel signals by DH-5920




(a) Frequency domain of signal X1




(b) Frequency domain of signal X2




(c) Frequency domain of signal X3

**Fig. 4.** Frequency domain of multi-channel signals of hydraulic pump

frequency is 1480/60 = 24.6 Hz, the plunger pump has seven plungers, and the inherent impact frequency of the hydraulic pump vibration signal is 24.60 × 17 = 172.2 Hz; The characteristic frequency of the loose slipper fault should be equal to or close to the shaft frequency 24.6Hz in theory. However, both inherent

frequency and fault characteristic frequency are annihilated in noise and interference. Therefore, the proposed ILCD fusion method can be used for hydraulic pump signal processing.

First, the signal X1 is processed using the ILCD method with high-frequency harmonics. According to Eqs. (1)–(6), ILCD yields eight ISC components and one trend item $u(t)$. Sensor1 is used to collect the vibration signal $X1_{nor}$ of the hydraulic pump under normal conditions according to the same sampling parameters. The sensitivity factor $\lambda$ of each ISC component is calculated, respectively, by Eq. (8), and the results are −51.26, 0.315, 0.963, 2.621, 4.836, 2.138, and 0.792. Therefore, the $ISC_{1–5}$ is taken as a sensitive component of X1, and the corresponding maximum sensitivity factor is $\lambda_1 = 4.836$. For comparative analysis, $ISC_{1–1}$, $ISC_{1–4}$, and $ISC_{1–5}$ frequency components of the three components are shown in Fig. 5.

By analyzing Fig. 5a, it can be found that the noise and interference components of the original signal X1 concentrate into the first component $ISC_{1–1}$ mostly, which contains almost no fault-related information, Therefore, the sensitivity factor is far less than other components; from Fig. 5b, $ISC_{1–4}$ has been able to reflect part of the characteristic information, that is, the inherent frequency information 172.2 Hz and its multiplier, but the loose slipper fault information is still relatively weak; it can be clearly seen from Fig. 5c that both the loose slipper information of 24.6 Hz and the inherent frequency of 172.2 Hz are reflected to a certain degree, so the sensitivity factor of $ISC_{1–5}$ is the highest, indicating that its contribution to characteristic information is the largest also, which is consistent with the previous analysis.

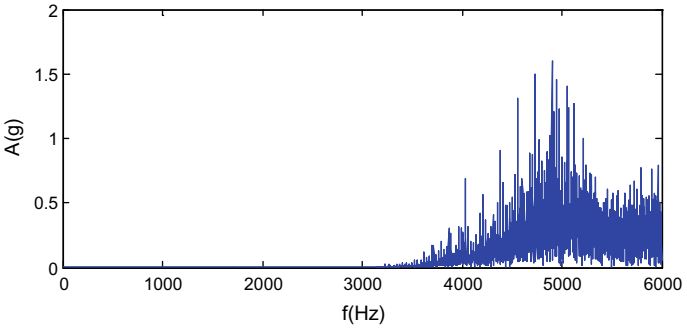
Using the same method, the signals X2 and X3 are processed, respectively, and the sensitive factors screened are $ISC_{2–4}$ and $ISC_{3–5}$, respectively. The frequency domain is shown in Fig. 6, and the corresponding sensitive factors are $\lambda_2 = 3.021$ and $\lambda_3 = 2.892$, respectively.

It can be seen from Fig. 6, the sensitive components $ISC_{2–4}$ and $ISC_{3–5}$ screened by X2 and X3 signals also contain relatively clear loose slipper information and inherent information. Compared with $ISC_{1–5}$ selected by X1 signal, the characteristic information proportion is relatively low, and the values of $\lambda_2$ and $\lambda_3$ are less than $\lambda_1$. According to Eq. (9), the fusion weights of three sensitive components are: $k_1 = 0.450$, $k_2 = 0.281$, $k_3 = 0.269$. According to Eq. (9), the reconstructed signal is:

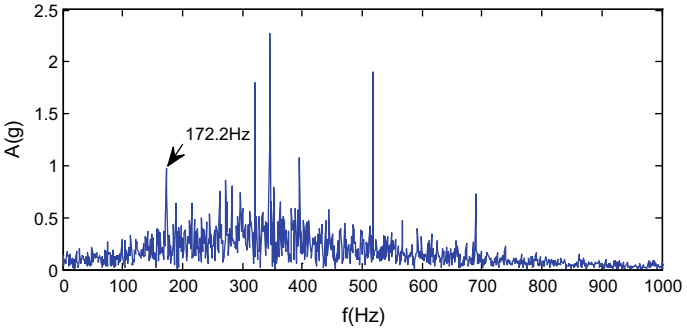
$$x_{\text{Final}} = 0.450\text{ISC}_{1-5} + 0.281\text{ISC}_{2-4} + 0.269\text{ISC}_{3-5}$$

The spectrum is shown in Fig. 7.

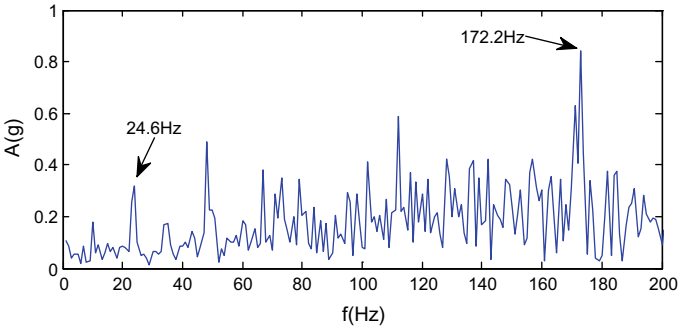
It can be seen from Fig. 7 that the fault feature information has been preliminarily extracted through the fusion pretreatment of the fault signal. It can be clearly seen that the fault characteristic frequency 24.6 Hz and its multiplier of loose slipper, also the inherent frequency 172.2 Hz and its multiplier of the hydraulic pump. The noise interference in original signal and modulation phenomena are effectively suppressed.




(a) Frequency domain of $ISC_{1\text{-}1}$

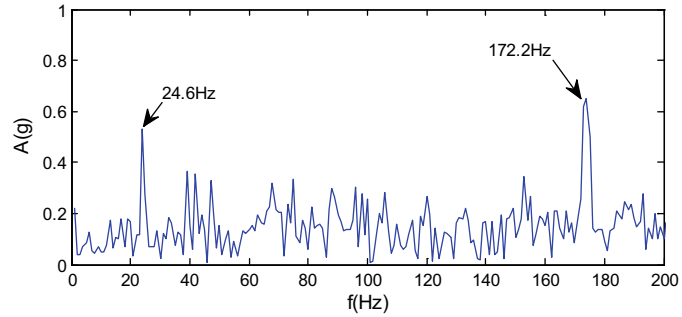



(b) Frequency domain of $ISC_{1\text{-}4}$

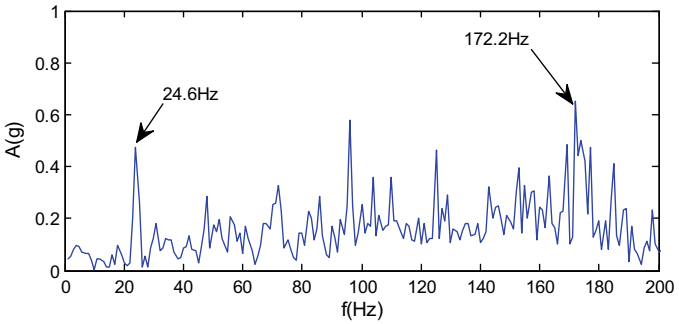



(c) Frequency domain of $ISC_{1\text{-}5}$

**Fig. 5.** Frequency domain of ISC




(a) Frequency domain of $ISC_{2\text{-}4}$




(b) Frequency domain of signal $ISC_{3\text{-}5}$

**Fig. 6.** Frequency domain of sensitive components

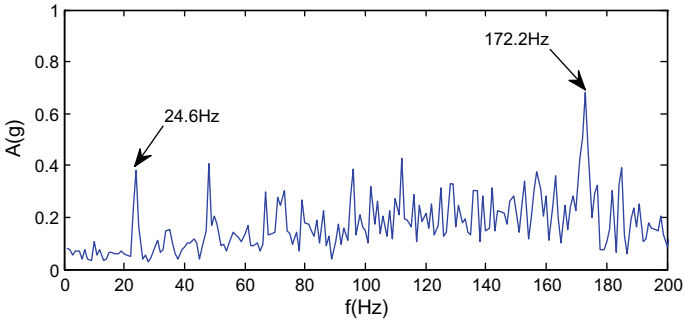



**Fig. 7.** Effect of preprocessing for signals of hydraulic pump based on ILCD fusion algorithm

## 5 Conclusion

The hydraulic pump vibration signal preprocessing method based on ILCD fusion is proposed, and its effect is verified using hydraulic pump measured vibration signal. The main conclusions are as follows:

(1) By combining high-frequency harmonics with ILCD, the ISC components are screened to reduce the influence of noise and interference components, and the sensitive feature information is effectively extracted.
(2) A fusion method based on ILCD is proposed, the sensitive components of multi-channel vibration signals obtained are fused, the reconstructed signals are obtained, the integrity of feature information is improved, and the preprocessing effect is improved.
(3) Experimental results show that the method proposed in this paper has a better preprocessing effect on multi-channel vibration signals and can extract the required feature information efficiently.

## References

1. Jinde Z, Junsheng C, Ming Z. A new time-frequency analysis method based on improved local characteristic-scale decomposition and normalized quadrature. Acta Electron Sin. 2015;43(7):1419–23.
2. Petrea L, Demian C, Brudny JF. High-frequency harmonic effects on low frequency iron losses. IEEE Trans Magn. 2014;50(11):6301204.
3. Cheng J, Zheng J, Yang Y. A nonstationary signal analysis approach-the local characteristic-scale decomposition method. J Vibr Engi. 2012;25(2):215–20.
4. Urgaonkar R, Neely MJ. Optimal routing with mutual information accumulation in wireless networks. IEEE J Sel Areas Commun. 2012;30(9):1730–7.
5. Zhao X, Patel TH, Zuo MJ. Multivariate EMD and full spectrum based condition monitoring for rotating machinery. Mech Syst Sig Process. 2012;27:712–28.

# GhostMobileCloud: A Cloud-Based Mobile Terminal User Accessing Remote 3D Scene Real-Time Communication Method

Ronghe Wang(✉), Bo Zhang, Jianning Bi, Xinhai Zhang, Xiaolei Guo, and Dong Jiao

National Engineering Laboratory for Public Security Risk Perception and Control by Big Data (PSRPC), Academy of Electronics and Information Technology of China Electronics Technology Group Corporation, No. 11 Shuangyuan Rd., Shijingshan District, Beijing 100041, People's Republic of China
wangronghe@buaa.edu.cn

**Abstract.** In this paper, we propose a real-time interactive method for mobile terminal users to access remote three-dimensional scene based on cloud services. We realize a dynamic and fluent method to access remote three-dimensional scene for mobile terminal. The method achieves a good user experience. We distribute mobile phone user task request and pad user task request from PC terminal user task request by the load balancing server and transfer these tasks to cloud server. So that mobile users transmit the interactive access operation in the scene to cloud server equivalent in the way of instruction set. We complete user command related work and high-quality rendering process by cloud server, which makes mobile terminal user access remote 3D scene immersive interaction. The mobile terminal user only needs to send a command request and receive high-quality pictures. The cloud servers receive instructions from user terminal and complete all related interactive work and transmit results in the form of high-quality pictures to mobile terminal user. This paper solves the technical challenge of mobile device users to access remote real-time interactive 3D scene in a manner of cloud services and continuous interval frame image for the first time.

**Keywords:** Cloud rendering · Large-scale three-dimensional scene · Mobile terminal · Remote interaction

## 1 Introduction

With the development of computer vision, virtual reality, augmented reality, enhancement virtual environment, etc., the various scale 3D scene has become a part of people's lives. There is no doubt that the current mainstream PC can directly access remote scene and relate interactive operation, but it is limited by many factors that mobile terminal users access remote 3D scene [1], such as hardware limitations,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_39

reliance on third-party plug-ins, the limitations of browser technology and network transmission speed limit.

**Hardware limitations**: In order to solve the problem of current mobile terminal rendering performance [2–8], mobile terminal has converted rendering work from original CPU to special GPU chip. However, mobile devices do not have a unified standard and different manufacturers have different product realization principles, so the rendering effects of different mobile devices are different [9–14].

Reliance on third-party plug-ins: Most of mobile terminal users need to install third-party plug-ins software to access remote 3D scene. It is difficult for some project requirements. Some users may refuse to install these third-party plug-ins, which hinders the promotion of related products.

**The limitations of browser technology**: The current mainstream mobile browser in the 3D scene rendering is widely used WebGL technology, but the technology in the mobile browser is not mature, and in the mainstream browser support for the technology is different. In this case, same program in different mobile devices on rendering effect is not only different, but also will be inefficient rendering and rendering errors in the process of rendering.

**Network transmission speed limit**: When mobile terminal needs to render 3D scenes locally, it needs to download material-related files from cloud server. The mobile terminal network transmission speed in response to these materials has some difficulty. If users download files by Wi-Fi or 4G network, it will spend a lot of time when the size of a single material file reaches 10 M magnitude. This will seriously affect user experience. If users store render material file in mobile device, it is difficult to avoid post material file update problem. Taking up too much storage space may also cause users to refuse to use.

## 2 Related Work

Remote rendering technology refers to server to bear most of the graphics computing work and store entire 3D scene model data. This can effectively reduce the burden of mobile devices. The existing remote rendering technology mainly includes the remote rendering method based on image, the method of remote drawing based on depth image and the remote rendering technology based on graphics.

The remote rendering method based on image refers to server to compress and send image view and view image stream to client according to user's viewpoint information of 3D scene rendering, and the client only need to decode compressed image sequence can be directly to display a 3D rendering results. At present, the main research teams include Mobile Systems Design laboratory of California University, GamePipe laboratory of University of Southern California, ARTEMIS (Advanced Research & Techniques for Multidimensional Imaging Systems) of Paris Institute of Telecommunications laboratory, System and Network Security laboratory of Taiwan national

University and NMSL (Networking and Multimedia Systems Lab) laboratory of Taiwan national Tsinghua University.

The method of remote drawing based on depth image introduces depth information and uses depth image to represent structure information of 3D scene. The server generates the color image and the depth image of the reference frame and sends them to the client according to user viewpoint information. When the viewpoint changes within a certain range, the client transforms 3D image into a new viewpoint image according to reference frame of the color image and the corresponding depth image.

The remote rendering technology based on graphics refers to graph calculation work by the server and the client. At present, there is less research on the distance drawing based on graphics on mobile devices. Remote rendering technology based on 3D graphics is that the server and the client assume certain graphics calculation work respectively. Most of the complex 3D graphics computing works are done by the server, and the client only needs to perform a small amount of 3D graphics calculation to get the final results.

## 3 The Core Method

### 3.1 Redistribution of Operations

The rendering work is transferred to cloud rendering cluster that has large-scale rendering computing ability, and mobile terminal realizes the remote 3D scene real-time interactive access only by sending instruction set and accepting picture.

### 3.2 Improvement of Interaction Mode

The mobile device communicates with the cloud rendering cluster through the instructions and displays the 3D scene by receiving the rendered image.

### 3.3 Improvement of Transmission Mode

Using a picture to show the 3D scene helps to get rid of the dependence of the third-party plug-ins, because the mobile device itself browser can finish this work. The resource needed by the rendering scene is loaded by the heterogeneous scene material resource manager, and the transmission process of the user's direct contact with these complex and large resource files is shielded.

The rendering architecture achieved by above several key design ideas realizes that the interactive experience of the remote 3D virtual scene similar to PC ends through the browser as the access platform on the mobile terminal device. In general, we implement a method for real-time interactive access to a remote 3D scenario by a mobile terminal user.

# 4 System Implementation

## 4.1 Characteristics of Three-Tier Architecture

The hierarchical structure of this paper is shown in Fig. 1.

This paper is the three-tier architecture based on socket communication connection and loose coupling. The top layer is the communication layer. The main task of this layer is to balance the distribution of the user's access request via the load balancing server and send the task to different servers mirror area. Then, inside of area, the task scheduling server further distributes the task and sends it to the rendering cluster for processing. Different interaction strategies are provided for different users by the rendering layer. The rendering layer distributes the user's access request to the less load renderer. In idle rendering preset area, a user base scene, as well as camera and rendering machine module, is generated. The command receiver receives the user's interactive operation instructions and calls other modules to complete all the operations contained in the user's instruction, and the operation process is transmitted to the end user in a continuous rendering frame image. The bottom layer is a data management layer. This layer manages the data resources of the scene. The renderer can update the resource file from this layer and write user interact results back to the configuration file.

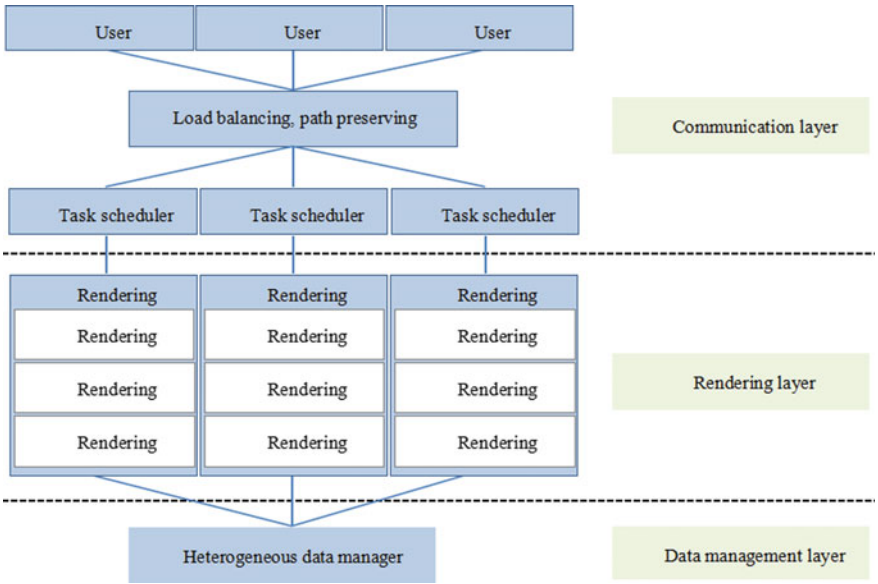



**Fig. 1.** Hierarchical structure of this paper

## 4.2 Feature of the Cloud Rendering Server

The cloud rendering server completes mobile terminal user's access request. U3D.exe executable program runs on cloud server. Each U3D.exe can start a thread by external

user commands, and generate a separate user virtual scene unit, and generate a camera, instruction transceiver and other function module in user virtual scene unit at the same time. User virtual scene unit communicates with the middle layer server through the socket. When users need to access to remote scene, the server will start an U3D.exe program through the socket and transfer the user instructions and began to work. The user can see all interactions operation in the virtual scene by the cameras. After receiving the instruction from the user, instructions transceiver triggers other modules to complete the interactive operation. The interaction operation is sent to the mobile terminal user in the form of a continuous interval rendering frame. So the large-scale computing focus on high performance distributed cloud server to complete, and the mobile terminal only sends commands and receives pictures. This way maximizes the server's high computational performance and minimizes the computational load of the mobile terminal and greatly improves the efficiency of mobile terminal users and remote 3D scene interaction.

### 4.3 Feature of Shunt and Mobile Terminal, and Architecture of Communication Layer

The access request of the mobile terminal users is allocated to the different server mirror areas to ensure the pressure balance of the cloud server. On the other hand, an independent cloud rendering server cluster is maintained inside each mirror area. The cluster is unified management by the only task scheduling server in the mirror area. When the user operation requests are received, task scheduling server will transform them into cluster identifiable operating instructions and then decide the render target machine according to the occupancy rate of the rendering machine in rendering cluster and user access cache. After receiving the user's instruction stream, the cloud server completes all user interaction operations and makes high-quality real-time rendering of interactive process and sends the result to the mobile terminal user in the form of pictures.

### 4.4 Rendering Layer, Scene and Scene Unit

Rendering layer contains many rendering servers. They are evenly divided ownership according to the number of mirror server area. They are unified management by task scheduling server. The main function of rendering layer is to dynamically manage the user's virtual scene and analyze the user instructions. These instructions include inter-operation, rendering the scene and other functions. The configuration information of the virtual scene and the scene material are provided by heterogeneous scene material resource management layer. In this way, the massive computations are concentrated on the high-performance cloud rendering cluster to complete. The mobile terminal sends commands and receives pictures only.

### 4.5 Heterogeneous Scene Resource Management Layer

It is very important for the whole system that the scene materials are unified management, because they are universal. The independence of scene material facilitates the

change of the system and reduces the storage pressure of rendering layer. The rendering layer can obtain the materials through the interface provided by the heterogeneous scene material resource manager. Considering the data exchange process may lead to system performance degradation, the layer designs the asynchronous download, asynchronous decompression and delegate callback mechanism, which ensures the rendering layer resource request process and the main loop of the rendering layer run in parallel. User virtual scene unit module is shown in Fig. 2.

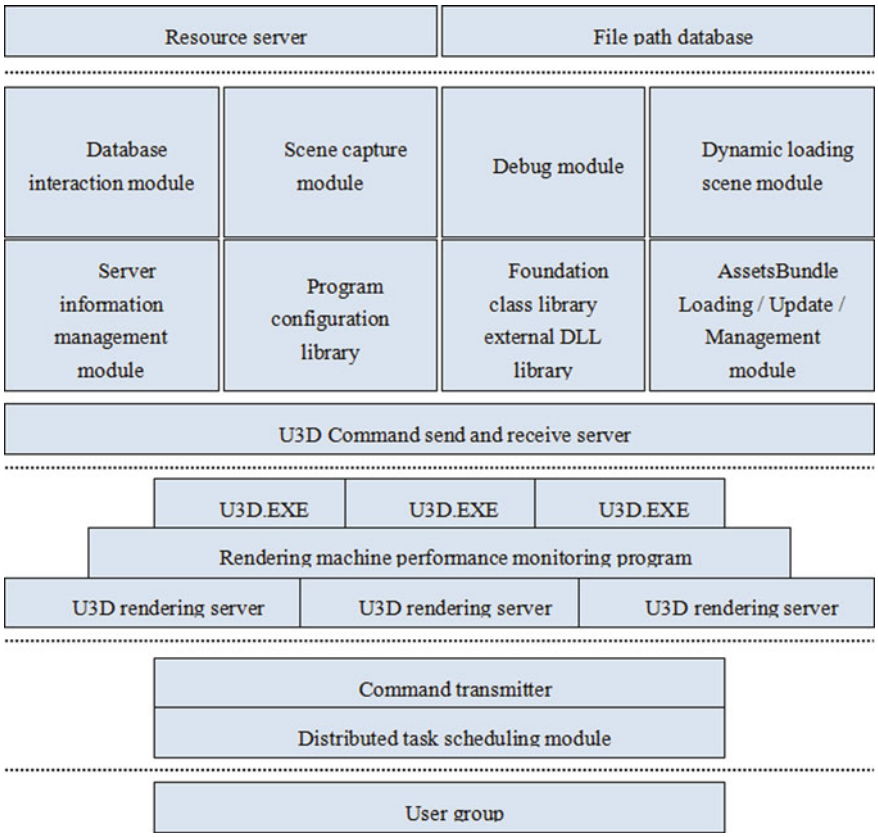



**Fig. 2.** User virtual scene unit module

As shown in the figure, the user group through the load server shunts. The user group assigns the access request of the mobile terminal user to the rendering server through the renderer performance-monitoring program. A thread is started in the rendering server's U3D.exe to complete the user remote real-time interactive access to their own user virtual scene. The lens capture module is used to realize the roaming function and realize the user on the scene in the roaming process self-timer function. The debug module is used to deal with the exception of the scene, get system status and

dynamically modify the internal parameters. The dynamic loading scene module is used to achieve asynchronous loading of the scene. The program configuration library is used to achieve the correct configuration of the program. The foundation class library external DLL library is used to achieve the basic library and external DLL call function. The AssetsBundle loading, update, management module are used to achieve synchronous download function.

## 5 Experiments and Comparison

The effects of our method on different devices are shown in Fig. 3.

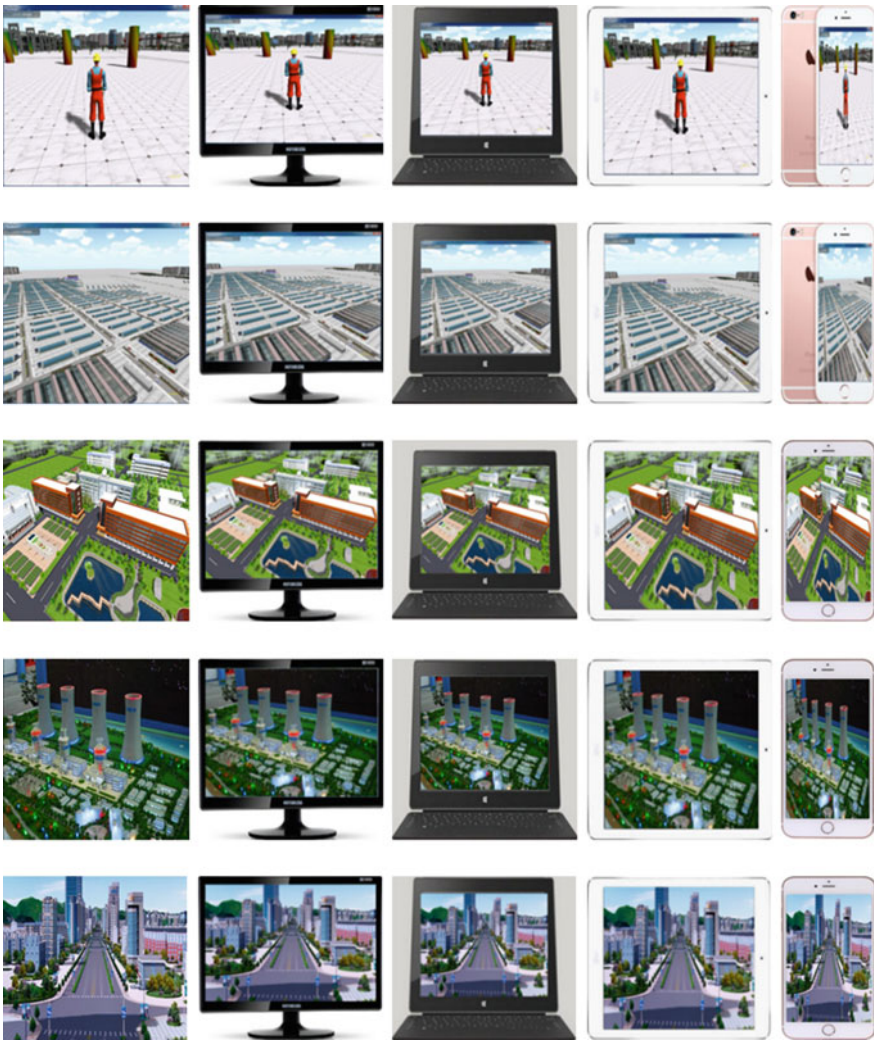


**Fig. 3.** Effect of our method on different devices

The function and performance of the 3D scene rendering are tested by experiments. The corresponding experimental results are given. When user navigates in virtual park scene, power plant scene and the virtual city scene, the client rendering results are shown in Fig. 4.

The data transmission when users view 3D scene is shown in Table 1.

(a) Virtual park scene (b) Power plant scene (c) Virtual city scene

**Fig. 4.** 3D scene test data set

**Table 1.** Data transmission when users view 3d scene

| Data of 3D scene | Total processing | Total visible | Proportion of transmission (%) |
|---|---|---|---|
| Patch | 3,319,336 | 309,722,592 | 0.81 |
| Vertex | 1,707,673 | 159,329,108 | 0.27 |
| Frame | 12,035 | 1215 | 1.23 |

As shown in the table, using cloud services, the total number of frame transmission is only about 1% of the total frame drawing data. Therefore, our method can effectively reduce the repeat data transmission, improve the transmission efficiency.

(1) The reuse of efficient rendering frame

We realize the efficient pictures reuse, and the comparison of average computation time with other algorithms is shown in Table 2.

**Table 2.** Comparison of average computation time with other algorithms

| 3D scene | Average drawing time (ms) | | | |
|---|---|---|---|---|
| | VFC | CHC | CHC++ | Our method |
| Virtual scene | 14.5 | 10.24 | 9.63 | 8.23 |
| Power plant | 20.42 | 10.61 | 9.81 | 9.57 |
| Virtual city | 15.76 | 10.89 | 9.92 | 7.79 |
| Average | 15.65 | 10.38 | 9.47 | 8.18 |

**Table 3.** Average processing time of remote rendering server

| Time performance (each frame) | Average processing time (ms) | | |
|---|---|---|---|
| | Virtual park scene | Power plant scene | Virtual city scene |
| Average server response time | 20.8 | 22.1 | 18.8 |
| Average drawing time of client | 8.17 | 8.23 | 8.07 |
| Average transmission time of data | 0.12 | 0.13 | 0.09 |
| Total time | 29.09 | 30.46 | 26.96 |

(2) The use frequency of cloud server

The parameters of the total number of paths are 100, 500 and 1000, as shown in table. The complexity of the algorithm is $O(n2)$, where $n$ is the total number of paths. We place a large rendering scene on the server to reduce the user's waiting time and improve the user's real experience.

Average processing time of remote rendering server is shown in Table 3.

## 6 Conclusions

This paper provides a method based on cloud rendering for mobile terminal users real-time interactive access to remote 3D scene by a distributed dynamic loading scene. The method mainly solves the problem of limited computing ability and core technology bottleneck constraints in mobile terminal users real-time interactive access to remote 3D scene.

**Acknowledgments.** We sincerely thank each one of the reviewers and editors' work to the paper. This paper is supported by National Natural Science Foundation Projects of China (Grant No. 61300007). The National Key Research and Development Program of China:Research and development of intelligent security card port monitoring and warning platform (Grant No. 2016YFC0800507), Innovation Foundation Program of China Electronics Technology Group Corporation:Research on holographic and abnormal behavior intelligent warning technology for social security risk targets.

## References

1. Deri L, Fusco F. Realtime MicroCloud-based flow aggregation for fixed and mobile networks. In: Wireless communications and mobile computing conference; 2013. p. 96–101.
2. Wang S, Dey S. Adaptive mobile cloud computing to enable rich mobile multimedia applications. IEEE Trans Multimedia. 2013;15(4):870–83.
3. Wang S, Dey S. Cloud mobile gaming: modeling and measuring user experience in mobile wireless networks. ACM SIGMOBILE Mobile Comput Commun Rev. 2012;16(1):10–21.

4. Zhao Z, Hwang K, Villeta J. Game cloud design with virtualized CPU/GPU servers and initial performance results. In: The workshop on scientific cloud computing DATE. ACM, 2012; p. 23–30.
5. Tizon N, Moreno C, Cernea M, et al. MPEG-4-based adaptive remote rendering for video games. In: Proceedings of the 16th ACM international conference on 3d web technology; 2011. p. 45–50.
6. Huang CY, Hsu CH, Chang YC, et al. GamingAnywhere: an open cloud gaming system. In: Proceedings of the 4th ACM multimedia systems conference 2013; p. 36–47.
7. Wu W, Arefin A, Kurillo G, et al. CZLoD: a psychophysical approach for 3D tele-immersive video. ACM Trans Multimedia Comput Commun Appl. 2012;8(3s):39.
8. Shi S, Nahrstedt K, Campbell R. A real-time remote rendering system for interactive mobile graphics. ACM Trans Multimedia Comput Commun Appl (TOMCCAP) 2012;8(3s):46.
9. Wu W, Arefin A, Kurillo G, et al. Color-plus-depth level-of-detail in 3D tele-immersive video: a psychophysical approach. In: ACM international conference on multimedia, ACM; 2011. p. 13–22.
10. Koppel M, Wang X, Doshkov D, et al. Depth image-based rendering with spatio-temporally consistent texture synthesis for 3-D video with global motion. In: Proceedings of the 19th IEEE international conference on image processing (ICIP 12); 2012. p. 2713–16.
11. Zhu M, Mondet S, Morin G, et al. Towards peer-assisted rendering in networked virtual environments. In: Proceedings of the 19th ACM international conference on multimedia; 2011. p. 183–92.
12. Zhang K, Liang X, Shen X, et al. Exploiting multimedia services in mobile social networks from security and privacy perspectives. IEEE Commun Mag. 2014;52(3):58–65.
13. Laurila JK, Gatica-Perez D, Aad I, Blom J, Bornet O, Do T-M-T, Dousse O, Eberle J, Miettinen M. The mobile data challenge: big data for mobile computing research, 2012.
14. Fernando N, Loke SW, Rahayu W. Mobile cloud computing: a survey. Future Gener Comput Syst. 2013;29:84–106.

# Research on Control Method of Electric Proportional Canard for Two-Dimensional Trajectory Correction Fuze of Movable Canard

Dan Fang(✉) and Yi Wang

Army Engineering University, Shijiazhuang, China
daoxy001@126.com

**Abstract.** The two-dimensional trajectory correction fuze of movable canard is the hot spot of research currently with the advantages of low cost and strong correction ability. For the two-dimensional trajectory correction fuze of movable rudder, continuous sine control, interval sinusoidal control, and constant control angle control are used to control the electric proportional canard. By analyzing the influence of different control methods on the ballistic characteristics and the correction ability, the canard control methods are evaluated and an optimal control method is proposed. It is greatly significant for the development of two-dimensional trajectory correction fuze for movable canard in theory and engineering.

**Keywords:** Two-dimensional ballistic correction fuze · Electric proportional canard · Control method

## 1 Introduction

Grenades, important member of the ammunition family, are important offensive fire suppression ammunition. However, they are widely dispersed and cannot meet the requirement of high strike accuracy of ammunition in modern warfare. The information transformation of grenade to improve strike accuracy is the current research hot spot.

The two-dimensional correction control mechanism commonly used in conventional ammunition information transformation can be roughly divided into three categories: correction control based on aerodynamic force, correction control based on direct force, and correction control based on inertial force. The aerodynamic-based correction control changes the aerodynamic force of the projectile by changing the aerodynamic shape of the projectile. The commonly used aerodynamic layout includes duck canard, normally arranged aerodynamic rudder, and wedge-shaped block mounted on the tail of a grenade [1–7]. The duck canard arrangement has different classifications of a pair of canard pieces, multiple pairs of canard pieces, fixed canard deviation, and adjustable canard deviation. The aerodynamic layout of the tail-mounted wedge block has a continuous correction feature with the greater ability to modify the grenades compared with the duck canard layout [8, 9]. The correction based on direct


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_40

force is mainly to achieve ballistic correction through pulse [10, 11]. Inertial force-based correction control achieves ballistic correction by mounting movable masses or rotating gyros inside the projectile [12, 13], but the application of that modified control method is only demonstrated on mortar shells.

In the correction mechanism based on the direct force, the effective correction control force provided is small due to the high speed of the grenades and the large pulse coverage area. After adding the rotation isolator, the correction control mechanism-based aerodynamic force can be effectively applied to grenades, such as the US 'Excalibur'-guided projectile, PGK, and 'silver bullet' correction fuze. The two-dimensional trajectory correction fuze is a new concept proposed in recent years. The replacement of the original ordinary fuze can realize the information transformation of the grenade. The two-dimensional trajectory correction fuze based on movable canard has the characteristics of strong correction ability and low cost, which is the key point in the research of the current grenade trajectory correction control.

In this paper, the canard control method for two-dimensional trajectory correction fuze of movable canard is studied. Based on the analysis of the affection of three control methods, namely continuous sinusoidal control, interval sine control, and constant control angle control, the correction ability of different control methods is studied and the canard control methods are eventually evaluated.

## 2 Electric Proportional Canard of Movable Canard Two-Dimensional Trajectory Correction Fuze

The two-dimensional trajectory correction fuze of movable canard has added satellite antenna, movable canard, missile-borne computer, and bearings on the basis of the original fuze, which leads to major changes in internal structure and aerodynamic shape as shown in Fig. 1.

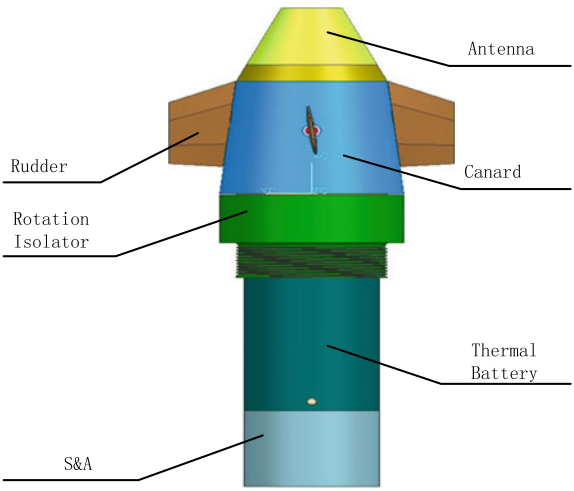



**Fig. 1.** Composition of movable canard two-dimensional trajectory correction fuze

The shape of movable canard two-dimensional trajectory correction fuze is shown in Fig. 2. The rudder 1 and the rudder 3, called the differential rudder, have the same rudder angle while the rudder direction deviation is different. The rudder 2 and the rudder 4, called control rudder, have the same rudder deflection angle and the rudder direction deviation. During the flight of the projectile, a rotating torque is formed on the differential rudder under the effect of air so that the canard rotates in reverse with respect to the projectile, and the fuze tail rotates clockwise (looking from the bullet tail to the warhead) to form the isolation from canard to the roll angular velocity of the projectile.

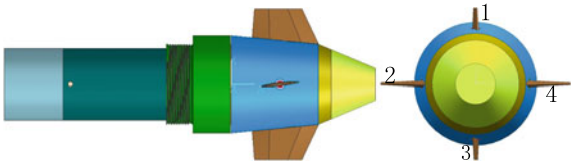



**Fig. 2.** Shape of movable canard 2D trajectory correction fuze

During the flight of the projectile, when it is in the uncontrolled state, the canard overcomes the rolling damping moment, the friction torque, and the interaction torque between the canard and the projectile under the action of the steering torque, and maintains the stable rotation speed less than 10 r/s. When the control rudder is in an uncontrolled state, the rudder angle is 0. When it is in the controlled state, the rudder blade performs two rudder partial motions during the one rotation of the canard, and the air acting on the control rudder forms the control force and the control moment and changes the attitude of the projectile to change the force of the projectile so as to achieve trajectory correction.

## 3 Canard Control Method

There are three canard control methods: continuous sine control, interval sine control, and constant control angle control. The control states with one period of the three methods are shown in Fig. 3.

### 3.1 Influence of Canard Control Methods on Trajectory Characteristics

Grasping the influence of different canard control methods on ballistic characteristics is the basis of selecting control methods. This section focuses on analyzing the effects of three canard control methods on ballistic characteristics. In the simulation analysis, the initial simulation conditions are as follows: projectile initial velocity is 879 m/s, firing angle is 40.5°, and standard meteorological condition is adopted.

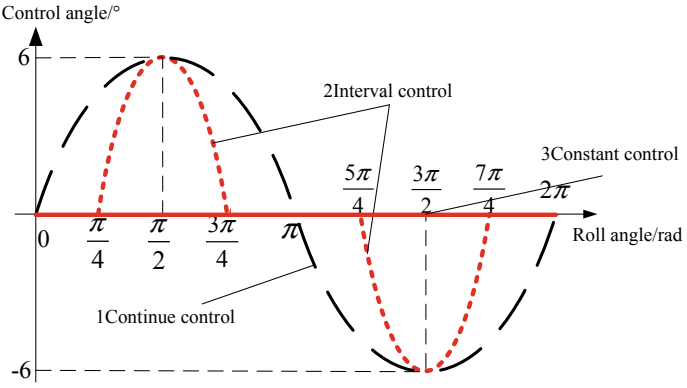
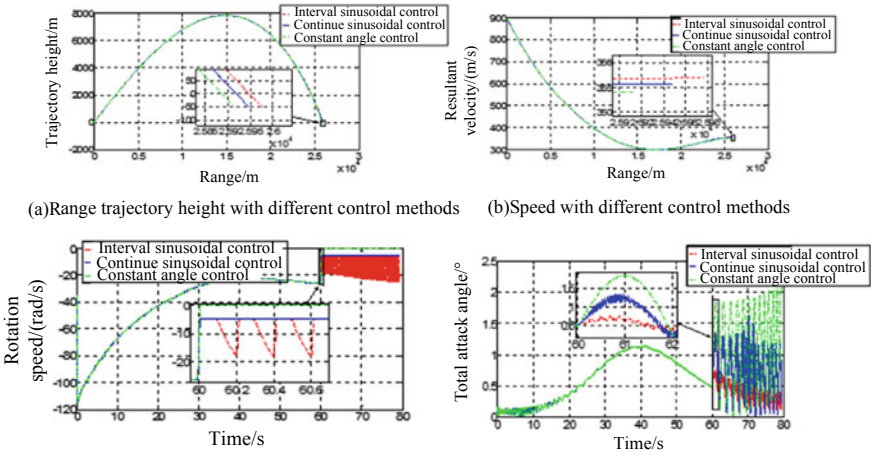



**Fig. 3.** Three control algorithms of control rudder

## 1. Influence of Different Control Methods on Trajectory Characteristics

According to Fig. 4, different control methods have different loss in range, so the three control methods have different set points. The ranges are 25,975, 25,950, and 25,935 m (shown in Fig. 4a). The effect of different control methods on the engaged speed can be ignored. Under that condition, the numerical values are 357.0, 356.0, and 354.0 m/s (shown in Fig. 4b). The constant angle control rotation speed remains 0 (rad/s). Continuous sinusoidal control remains constant at 5 (rad/s) after controlled. Interval sinusoidal control increases to 17 (rad/s) when uncontrolled and returns back to 5 (rad/s) after control (shown in Fig. 4c). In terms of total attack angle, we can conclude that the total attack angle changes significantly after the start of control in the three control methods, but the peak value does not exceed 3°, which meets the ballistic stability requirements (shown in Fig. 4d).



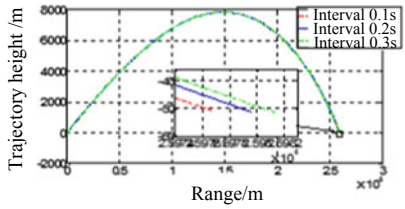

(a)Range trajectory height with different control methods (b)Speed with different control methods

(c)Rotation speed with different control methods (d)Total attack angle with different control methods

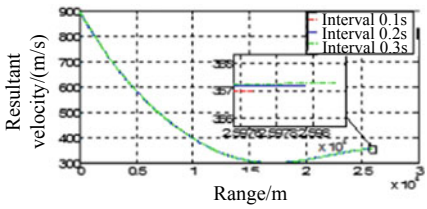
**Fig. 4.** Ballistic simulation with three control methods

2. **Influence of Different Interval time of Interval Sine Control on Trajectory Characteristics**

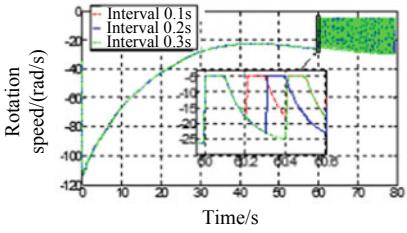
According to Fig. 5, the interval sinusoidal control has different loss in range for different interval time, so the set points are different and range is 2598.3, 2597.8, and 2598.6 m (shown in Fig. 5a). The resultant velocity in end is proportional to the interval time. Under those conditions, the values are 357.0, 357.2, and 357.3 m/s (shown in Fig. 5b). The effect of the interval time on the rotation speed is that the peaks are different when the speed increases during control. It is 17, 23, and 25 rad/s (shown in Fig. 5c). In terms of total attack angle, we can conclude that there is a change in the total attack angle after starting control at different intervals, but the peak value does not exceed 3° to meet the ballistic stability requirements (shown in Fig. 5d).



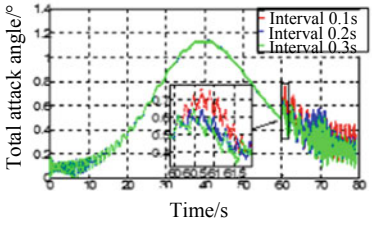

(a)Trajectory height with different interval time




(b)Resultant velocity with different interval time




(c)Rotation speed with different interval time

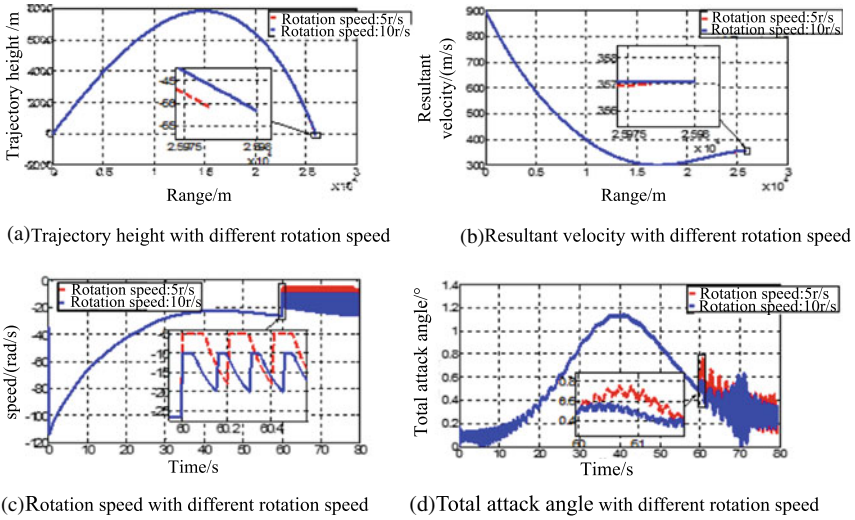



(d)Total attack angle with different interval time

**Fig. 5.** Trajectory simulation of interval control for different starting time

3. **Influence of Different Rotation Speed on Trajectory Characteristics of Movable Wing**

According to Fig. 6, the range loss at different speeds is small. There is basically no change in the landing point in the order of 2597.7 and 2598.0 m (shown in Fig. 6a). The change in combined speed is basically the same, and the deviation is 0.2 m/s (shown in Fig. 6b). The initial speed is different after starting control, so when the controlled state is maintained, the initial speed is kept and the impact is small (shown in Fig. 6c). The change range of the total attack angle is less than 3°, which coincides with the ballistic stability performance indicators (shown in Fig. 6d).




(a)Trajectory height with different rotation speed

(b)Resultant velocity with different rotation speed

(c)Rotation speed with different rotation speed

(d)Total attack angle with different rotation speed

**Fig. 6.** Trajectory simulation with different rotation speeds

In summary, the different control methods, different speeds, and different intervals have different disturbances on the total attack angle, and there are also differences in the amplitude of up and down disturbances. However, the amplitudes do not exceed 3°, which meets the stability index of the ballistics. The influence in other aspects will not cause obvious deviation on the trajectory.

### 3.2 Correction Ability

The correction ability is the maximum correction distance that can be achieved by the bullets, which can be divided into longitudinal correction ability and lateral correction ability.

The projectile landing point and correction ability under different controls are as follows.

According to Fig. 7, it can be seen that the influence of different starting time under constant angle control on the landing point of the projectile is also different. The range of the lateral deviation obtained from the control at 20 s is the largest, and the range gradually decreases as the time increases. The sphere of correction capability is inversely proportional to the starting time. The maximum longitudinal far modification is 282.0 m. The maximum longitudinal short modification is 274.0 m. The maximum lateral left modification is 373.0 m. The maximum lateral right modification is 369.0 m (shown in Fig. 8). Under the same condition, the length and distance of the obtained lateral deviations are different for the projectiles in the controlled state.

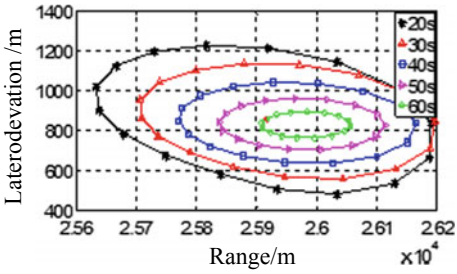



**Fig. 7.** Constant angle control of the landing point of different starting time

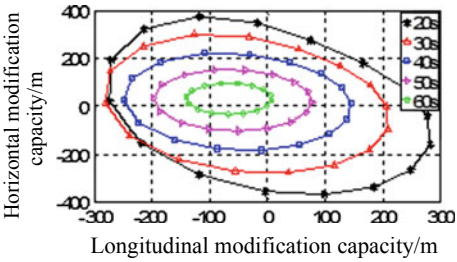



**Fig. 8.** Constant angle control correction capability

According to Fig. 9, it can be seen that the different starting time under interval control has different effects on the dropping point of the projectile. The same range of sideways deviation is obtained with the same 20 s control. The range gradually decreases as the time increases. The size of the range is inversely proportional to the start-up time, with the maximum longitudinal far modification of 199.3 m, the maximum longitudinal short modification of 113.5 m, the maximum lateral left modification of 224.0 m, and the maximum lateral right modification of 231.9 m (shown in Fig. 10). Under the same condition, the length and distance of the obtained lateral deviations are different for the projectiles in the controlled state.

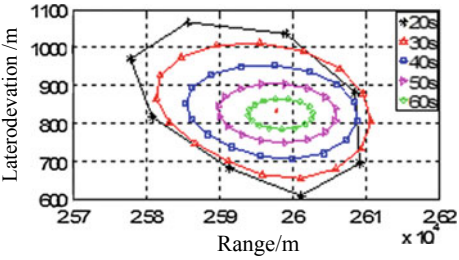



**Fig. 9.** Interval control the landing point of different starting time

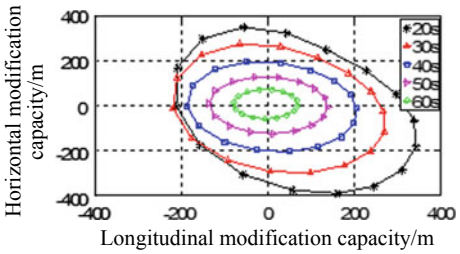



**Fig. 10.** Interval control correction capability

According to Fig. 11, it can be seen that the different starting time under the interval control has different influence on the landing point of the projectile. The range of the side deflection obtained from 20 s control is the largest, and the range gradually decreases as the time increases. The range of corrective ability is inversely proportional to the starting time. The maximum longitudinal far modification is 169.6 m. The maximum longitudinal short modification is 143.2 m. The maximum lateral left modification is 238.9 m. The maximum lateral right modification is 217.1 m (shown in Fig. 12).

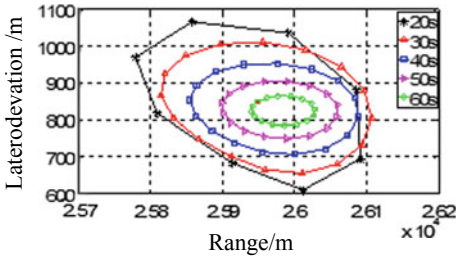



**Fig. 11.** Continuous control of projectile landing points for different starting times

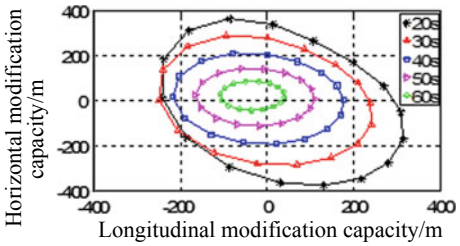



**Fig. 12.** Continuously controlled correction capability

The correction capabilities of different control methods are not the same. A large range means that the range of correction with the control trajectory is wide, indicating that the corrective ability of the correction component is strong. The correction range of different control methods is inversely proportional to the start-up time and decreases with the increase of the start-up time. The three control methods have the largest correction ability at a starting time of 20 s. The constant angle control is larger than other correction ranges and has stronger correction ability.

Similar to the method of calculating the correction ability of the control rudder of different structures, the correction ability and range loss of different canard control methods are calculated using the maximum rudder angle of 6° as an example. The results are shown in Tables 1 and 2.

From the two tables, the range loss of the constant angle control scheme is 8.6%, but within the acceptable range (less than 10%). The constant angle control is more robust than the other two control modes.

**Table 1.** Comparison of range loss with different canard control schemes

| Canard control scheme | Range (%) | Range loss (m) | Range loss percent (%) |
|---|---|---|---|
| Without rudder | 28374.1 | – | – |
| Constant control (starting time: 60 s) | 25915.7 | 2458.4 | 8.6 |
| Continuous control (rotation speed: −5 r/s) | 25948.4 | 2425.7 | 8.5 |
| Interval control (interval time: 0.1 s) | 25978.1 | 2396.0 | 8.4 |

**Table 2.** Correction ability of projectiles with different canard control schemes

| Canard control scheme | Correction ability (m) | | | |
|---|---|---|---|---|
| | Longitudinal far modification | Longitudinal short modification | Horizontal left modification | Horizontal right modification |
| Continuous control | 169.6 | 143.2 | 238.9 | 217.1 |
| Interval control | 199.3 | 113.5 | 224.0 | 231.9 |
| Constant control | 282.0 | 274.0 | 373.0 | 369.0 |

### 3.3 Canard Control Method Evaluation

The optimization of movable canard structure parameters should follow the following principles

(1) In the flight process, the stability requirements should be satisfied. The variation ranges of the total attack angle with the three control methods, different interval times, and different speeds are all less than 3° to satisfy the stability requirements.
(2) The loss on the firing is as small as possible to meet the large range. The constant angle control has a percentage of the range loss of 8.6%. The continuous sine control has a percentage of the range of 8.5%. The interval sine control has a percentage of the range of 8.4%. However, the range loss of all three control schemes is within the acceptable range (less than 10%).
(3) The correction capability is enough to ensure that the CEP meets the tactical technical indicators. From Tables 1 and 2, the longitudinal correction ability and lateral correction ability of the constant angle control are the largest compared with the other two control methods.

The distribution of the landing point of the uncontrolled projectile with different movable canard structural parameters is basically equal. Therefore, the choice of the structural parameters of the movable canard is to ensure that the range loss is minimized based on the correction ability.

## 4 Conclusion

Different control methods, rotation speeds, and interval times have different disturbances on the total attack angle. There are also differences in the amplitude of up and down disturbances while the amplitude does not exceed 3° to meet the requirements of trajectory stability. The impact on other aspects will not cause obvious deviation on trajectory. The constant angle control scheme has a larger range loss, but within an acceptable range (less than 10%). Moreover, the longitudinal correction capability and lateral correction capability are the largest compared with the other two control methods. Therefore, the constant angle control is adopted.

## References

1. Clancy J, Bybee T, Friedrih W. Fixed canard 2-D guidance of artillery projectile. US: 6981672B2. 2006.
2. Costello M. Modeling and simulation of a differential roll projectile. In: Proceedings of the 1998 AIAA modeling and simulation technologies conference; 1998. p. 490–9.
3. Wernert P, Theodoulis S. Modeling and stability analysis for a class of 155 mm spin-stabilized projectiles with course correction fuse (CCF). In: Proceedings of the 2011 AIAA atmospheric flight mechanics conference and exhibit; 2011. p. 1–13.
4. Theodoulis S, Gassmann V, Wernert P. Guidance and control design for a class of spin-stabilized fin-controlled projectiles. J Guid Control Dyn. 2013;36(2):517–31.
5. Gagnon E, Lauzon M. Course correction fuse concept analysis for in-service 155 mm spin-stabilized gunnery projectiles. In: Proceedings of the 2008 AIAA guidance, navigation and control conference and exhibit; 2008. p. 1–20.
6. Fresconi F, Cooper G, Celmins I et al. Flight mechanics of a novel guided spin-stabilized projectile concept. Aerosp Eng. 2013;226:327–340.

7. Gross M, Costello M, Fresconi F. Impact point model Predictive control of a spin-stabilized projectile with instability protection. In: AIAA atmospheric flight mechanics conference. Boston; 2013. p. 1–21.
8. Fresconi F, Plostins P. Control mechanism strategies for spin-stabilized projectiles. In: Proceedings of 47th AIAA aerospace sciences meeting including the new horizon forum and aerospace exposition, Orlando, Florida, January 2009. AIAA Paper; p. 1–23.
9. Ollerenshaw D, Costello M. On the swerve response of projectiles to control input. J Guidance Control Dyn. 2008;31(5):1259–65.
10. Corriveau D, Berner C, Fleck V. Trajectory correction using impulse thrusters for conventional artillery projectiles. In: 23th International symposium on ballistics, Tarragona, Spain; 2007. p. 639–46.
11. Corriveau D, Wey P, Berner C. Thrusters pairing guidelines for trajectory corrections of projectiles. J Guidance Control Dyn. 2011;34(4):1120–8.
12. Frost G, Costello M. Control authority of a projectile equipped with a internal unbalanced part. J Dyn Syst Meas Control. 2006;128(4):1005–12.
13. Rogers J, Costello M. Control authority of a projectile equipped with a controllable internal translating mass. J Guidance Control Dyn. 2008;31(5):1323–33.

# Normalization of Statistical Properties of Sea Clutter Based on Non-coherent Accumulation

Yi Liu(✉), Shufang Zhang, and Jidong Suo

College of Information Science and Technology,
Dalian Maritime University, Dalian, China
1042031044@qq.com, yunxiaoyuji1987@163.com

**Abstract.** For clutter with long tailing characteristics, such as Weibull distribution, Ruili distribution, lognormality, and K-distribution, CFAR processors corresponding to various distributions need to be used when CFAR processing technology is used. Otherwise, it is difficult to obtain constant false alarm characteristics when the statistical characteristics of clutter change. From the viewpoint of improving the robustness of CFAR, according to the central limit theorem and the logarithmic compression principle of the signal, this paper attempts to accumulate and average the radar signal before CFAR processing according to the pulse accumulation characteristics and median limit theorem of radar signals. The clutter of the post-PDF is close to a normal distribution, which effectively eliminates the trailing effect of the clutter characteristics, i.e., it effectively suppresses the sharp peak interference and the distribution characteristics of the normalized clutter.

**Keywords:** CFAR · Radar pulse accumulation · Normalization · Smearing · Radar clutter

## 1 Introduction

Constant false alarm rate (CFAR) detector is an efficient way for controlling the false alarm rate (FAR) of radar to avoid the receiver fault caused by high FAR [1]. Among them, the most common one is called Mean Level CFAR (ML CFAR) detection processor which is based on the Cell Averaging CFAR (CA-CFAR) detector [2]. The hypothesis of CFAR detection processor is that, the clutter obeys certain statistical distribution characteristics, and all or parts of the parameters are known. Due to the improvement of radar resolution and the resolution performance of the same radar is various under different operating parameters, the statistical characteristics of radar clutter will change, and the amplitude of sea clutter will not obey the Rayleigh distribution in low-resolution condition. It is obvious that the processing effect of various CFAR processors based on CA-CFAR detection processor is reduced. Many studies show that the logarithmic normal (lognormal) distribution, Weibull distribution, and K-distribution model can better describe the statistical distribution characteristics of sea clutter amplitude in high-resolution [3] condition. Literature [4, 5] prove that the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_41

ordered statistic CA-CFAR (OSCA-CFAR) class processors (OSGO-CFAR, OSSO-CFAR) have constant false alarm performance under the condition when the PDF of its K-distribution is known. However, in practical applications, it is difficult to establish a variety of CFAR processors to adapt to the clutter background of various statistical characteristics.

The literature [6] proposes several PDF models for the CFAR detection of most sea clutter based on low-grain incidence angle high-resolution radar with different degrees of tailing effect, which include lognormal distribution, Weibull distribution, and K-distribution. Statistical studies have shown that the K-distribution is the most suitable model for sea clutter in high-resolution and/or low grazing incidence radar applications. The literature [7] analyzes the deconvolution error caused by the asymmetry caused by the clutter with tailing effect and gives a deconvolution iteration-based tailing suppression and noise with improved distribution symmetry. The suppression filter design method effectively reduces the deconvolution error, but the deconvolution iterative algorithm used by this method is somewhat complicated for radar clutter processing. The literature [8] pointed out that the commonly used parameter estimation method is more sensitive to these spike data, leading to larger estimation error. For radar target scattering center parameter estimation and feature extraction based on decay index (DE) and model, the noise background is often non-Gaussian, and the distribution density function exhibits a long tailing property which uses M-estimation to implement long-delay dragging. The literature [9] pointed out that the K-distribution is the most suitable model for the long-tailed sea clutter effect, although it has high-resolution and/or low grazing angle radar clutter statistical properties. This paper discusses the "ideal fixed" and "ideal CFAR" threshold detection criteria for ideal high-resolution radar detection performance. However, there is no discussion on how to reduce or overcome the effect of smearing on the robustness of target detection performance.

It can be seen from the above that the main reason affecting the target parameter estimation and detection effect is the tailing effect and statistical parameter inconsistency of the statistical distribution characteristics of different clutters. Therefore, how to effectively treat the clutter prior to detection makes it possible to effectively reduce the tailing effect of the clutter after the pre-detection treatment, and it is desirable to be able to tend to a normal distribution. This article is based on this purpose to start research.

This paper first analyzes the statistical characteristics of the signal model and discusses the basic method of weakening the tailing effect of clutter statistical distribution characteristics. The Weibull Clutter, the Ruili Clutter, the lognormal clutter and the K-distributed clutter are accumulated five times, and the statical histogram of the accumulated clutter is obtained, as well as the mean and variance estimates after the clutter accumulation. The normal distribution of the theory is fitted and compared to verify the validity of the proposed method. In addition, the *W*-test method [10] was used to test the normality of the clutter after 5 times accumulation of K-distribution clutter, which further validated the effectiveness of the proposed method. The research method of this paper is to use the CA-CFAR processor based on normal distribution to achieve the robust and constant false alarm handling effect in the complicated and changeable clutter environment.

## 2 Analysis of Signal Models and Statistical Characteristics

Let $M$ denote the number of pulse accumulation of radar, at $i$th ($1 \geq i \geq M$) pulse repetition period, $j$th range quantification cell, received signal is expressed as

$$x(i,j) = \begin{cases} s(i,j) + c(i,j) + n(i,j), H_1 \\ c(i,j) + n(i,j), \qquad H_0 \end{cases} \tag{1}$$

where $s(i, j)$ is the target echo component and $c(i, j)$ is the echo component of the sea clutter in the received signal, and $n(i, j)$ is the noise component generated by the receiver. And $H_1$ and $H_0$ denote the existence of the target and the absence of the target, respectively.

For ship navigation radar and coastal surveillance radar, the sea clutter component is much larger than the noise component. In order to simplify the problem, the noise component in the formula (1) can be ignored, and (1) becomes:

$$x(i,j) = \begin{cases} s(i,j) + c(i,j), H_1 \\ c(i,j), \qquad H_0 \end{cases} \tag{2}$$

If the radar pulse signal is non-coherently accumulated, the $j$th distance unit performs $M$ pulse repetition period of slow time sampling and accumulates the average, the echo signal is

$$y(j) = \frac{1}{M} \sum_{i=1}^{M} x(i,j) \tag{3}$$

The non-coherent accumulation of the following radar signals is accomplished in a range unit, therefore, the distance unit $j$ in the (3) can be omitted, and (3) can be written as below:

$$y = \frac{1}{M} \sum_{i=1}^{M} x(i) = \begin{cases} \frac{1}{M} \sum_{i=1}^{M} (s_i + c_i), H_1 \\ \frac{1}{M} \sum_{i=1}^{M} c_i, \qquad H_0 \end{cases} \tag{4}$$

In a certain range cell of the radar, during the accumulation of $M$ pulses, the sea surface reflected clutter $c_i$ ($i$ = 1, 2, …, $M$) generated by each radar transmitting pulse should be independent and identically distributed. According to the central limit theorem, it is assumed that each quantity in the random variable sequence $\{c_1, c_2, \ldots, c_M\}$ has the same mathematical expectation and variance, ie.

$$E[c_i] = \mu, D[c_i] = \sigma^2 \tag{5}$$

and let

$$c_\Sigma = \sum_{i=1}^{M} c_i \tag{6}$$

When $M \to \infty$, $\xi = \frac{1}{\sqrt{M}\sigma}(c_\Sigma - M\mu) \sim N(0, 1)$, i.e.

$$p(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2}\right) \tag{7}$$

Thus, the statistical distribution density function of the clutter $\bar{c} = \frac{1}{M} c_\Sigma$ with $M$ times incoherent accumulation is

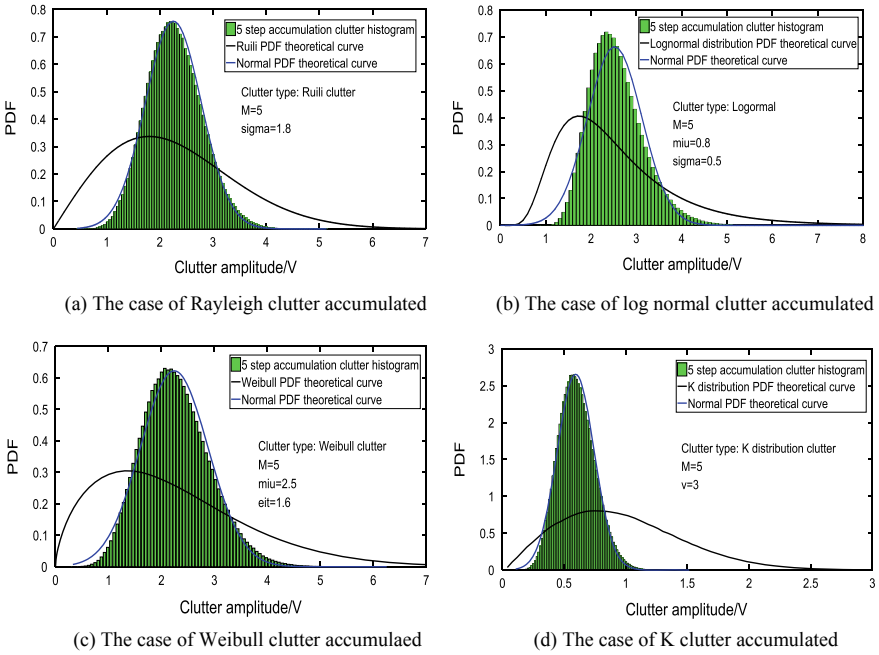
$$p(\bar{c}) = \frac{1}{\sqrt{2\pi}\left(\frac{\sigma}{\sqrt{M}}\right)} \exp\left(-\frac{(c-\mu)^2}{2\left(\frac{\sigma}{\sqrt{M}}\right)^2}\right) \tag{8}$$

It can be seen that the mean value of the clutter $\bar{c}$ accumulated is unchanged, and the variance is $1/\sqrt{M}$ of original one before accumulated, the amplitude of spikes decreases, so, the trailing effect of the statistical distribution characteristics of clutter is weakened greatly also. This is undoubtedly favorable for improving the performance of CFAR processor based on clutter mean and improving the applicability of the processor.

## 3 Experimental Verification

In practice, $M$ is difficult to be very large, but experimental analysis shows that even if $M$ is not very large, the tailing of the clutter after non-coherent accumulation will be effectively compressed, close to the normal distribution. Fig. 1 illustrates the comparison between normal PDF, histogram accumulated for 5 times, and PDF of the original clutter not accumulated, where (a) is Rayleigh clutter, (b) is lognormal clutter, (c) is Weibull clutter, and (d) is K clutter respectively. The statistical histogram is compared with the theoretical curve generated from the normal distribution of the mean and variance estimates of the five accumulated average clutters. From the figure, it can be seen that the clutters of various distribution characteristics are non-coherent for 5 times. After the accumulation, the tailing effect is greatly reduced and is close to the normal distribution.

For the shore ship navigation and monitoring radar, the antenna scanning cycle is 20 r/min, and the 3 dB antenna has a horizontal beam width of 0.45°. When the radar pulse repetition rate is 2000 Hz, the realized pulse accumulation number is $M > 7$.




(a) The case of Rayleigh clutter accumulated

(b) The case of log normal clutter accumulated

(c) The case of Weibull clutter accumulaed

(d) The case of K clutter accumulated

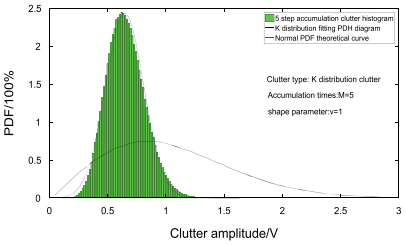
**Fig. 1.** Comparison of trailing improvement for 4 clutters accumulated with 5 times

For a ship navigation radar, the antenna rotates 24 times per minute, and the 3 dB antenna has a horizontal beam width of 1.1°. When the radar pulse repetition rate is 2000 Hz, the realized pulse accumulation number is $M > 15$. It can be seen that for marine radars used in ships and for shore-based and monitoring, it can achieve a large number of pulse accumulations.

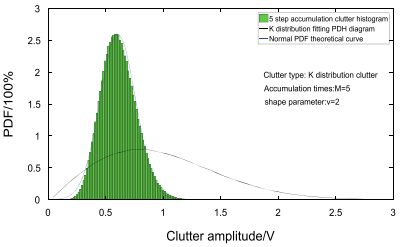
The distribution parameters of different clutters are different, after finite accumulation, it is different from the pseudo synthetic graph of the normal distribution theory PDF characteristic, the histogram of the 5 accumulation of the K-distribution clutter with different parameters and the fitting of the theoretical PDF of the normal distribution are shown in Fig. 2.

According to the *W*-test method, the normality of the accumulated clutter is tested experimentally and the experimental result is shown in Fig. 3.

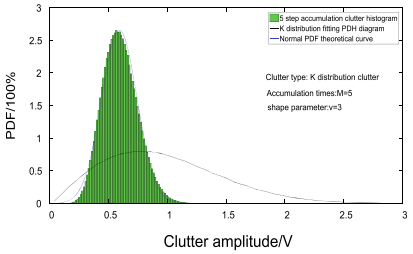
From Figs. 1 and 2, it can be seen that although the clutter cannot be completely fitted with the theoretical PDF curve of normal distribution after a finite number of accumulation processes, it has been very close. The symmetry has been improved, and the trailing effect has been effectively reduced. That is to effectively suppress the sharp spike interference, which is undoubtedly very beneficial for the realization of the CFAR processing. According to the "*W* inspection threshold table," Fig. 3 shows that



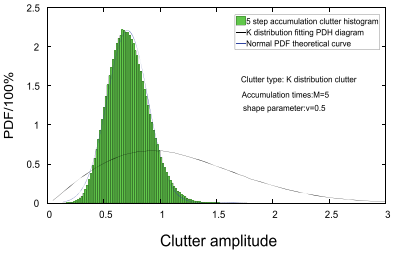

(a) The curve fitting between histogram of 5 times accumulation of K clutter in v=1 and normal PDF theory




(b) The curve fitting between histogram of 5 times accumulation of K clutter in v=2 and normal PDF theory



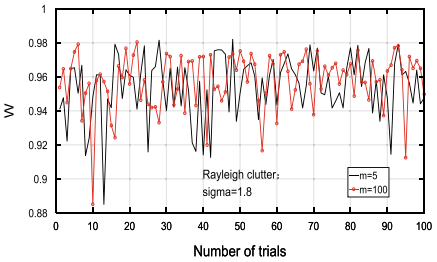

(c) The curve fitting between histogram of 5 times accumulation of K clutter in v=2 and normal PDF theory



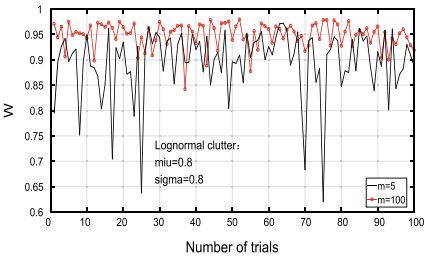

(d) The curve fitting between histogram of 5 times accumulation of K clutter in v=2 and normal PDF theory

**Fig. 2.** K clutter after five cumulative histograms and normal PDF theoretical curve fitting and trailing improvement comparison

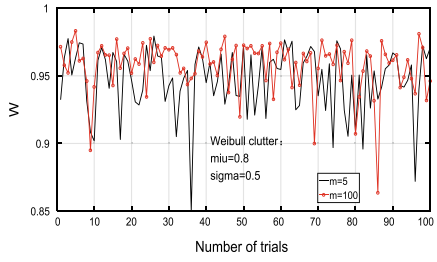
the clutter, which is accumulated over 5 times, is subject to normal distribution on the whole. At the same time, in Fig. 3, we can also analyze the K-distribution of the clutter, with the increase of the accumulation number, the *W* value will be gradually reduced. This is mainly due to the influence of the correlation of the K-distribution clutter, which leads to the accumulation of over the number of times, which will lead to the deviation from the normality. But the accumulation of radar pulses is usually around 10, so this effect is not serious.



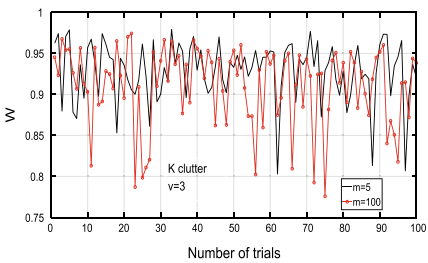

(a) The curve of W-test for Rayleigh clutter.




(b) The curve of W-test for logarithmic normal clutter.




(c) The curve of W-test for Weibull clutter.




(d) The curve of W-test for K clutter.

**Fig. 3.** Test results of *W*-test

## 4 Conclusions

In this paper, based on the radar pulse accumulation characteristics and the median limit theorem, the non-coherent accumulation of radar signals is averaged before the CFAR processing, so that the PDF of the accumulated clutter is close to the normal distribution, which simplifies the clutter distribution characteristics. Eliminate the trailing effect of clutter characteristics. This method undoubtedly uses the CA-CFAR processor based on normal distribution to achieve the CFAR processing effect in a complicated and changeable clutter environment and provides a clutter preprocessing method before the CFAR processing. This is of great significance for improving the robustness of the CFAR processor.

## References

1. Raghavan RS. CFAR detection in clutter with a Kronecker covariance structure. IEEE Trans Aerosp Electron Syst. 2017;53(2):619–29.
2. Finn HM, Johnson RS. Adaptive detection mode with threshold control as a function of spatially sampled clutter-level estimates. RCA Rev. 1968;29:414–64.
3. Youzhong H, Yingning P et al. Radar automatic detection and processing. Beijing: Tsinghua University Press; 1999.
4. Chengpeng H, Chaohuan H, Jin R, Jianping Y. Performance analysis of OSGO- and PSSO-CFAR in K clutter background. J Electron Inf Technol. 2005;27(7):1061–4.

5. Chengpeng H, Chaohuan H, Jianping Y. Performance analysis of OSCA-CFAR in K clutter background. Electron Inf Countermeas Technol. 2006;21(2).
6. Abdel-Nabi MA, Seddik KG, El-Badawy E-A. Spiky Sea clutter and constant false alarm rate processing in high-resolution maritime radar systems. In: International conference on computer and communication engineering (ICCCE 2012), 3–5 July 2012, Kuala Lumpur, Malaysia.
7. Yamauchi H, Somha W.A noise suppressing filter design for reducing deconvolution error of both-directions downward sloped asymmeric RTN long-tail distributions. In: 2015 International workshop on CMOS variability (VARI); 2015. p. 51–6.
8. Shi Z, Zhou J, Zhao H, Fu Q. Robust estimation of scattering center parameters in long-tailed K-distribution clutter. J Electron Inf Technol. 2007;29(12).
9. Ritcey JA. Radar Detection in K-distributed clutter plus noise using L-statistics. 978-1-5386-1823-3/17/$31.00 ©2017 IEEE
10. Yi Z, Xiubin Z, Xiaohua T, Hongsen X. Knowledge-aided signal detection algorithm in non-Gaussian clutter. Sig Process. 2012;28(1):60–6.

# Improved Max-Log-MAP Turbo Decoding by Extrinsic Information Scaling and Combining

Lei Sun and Hua Wang(✉)

School of Information and Electronics,
Beijing Institute of Technology, Beijing, China
wanghua@bit.edu.cn

**Abstract.** Turbo codes are among the best error-correcting codes, but trade-offs between performance and complexity in decoding are required for hardware implementation. In this paper, a novel extrinsic information scaling scheme for max-log-MAP decoder is proposed. It scales and combines extrinsic information generated at successive iteration round. The proposed method is evaluated for 3GPP LTE turbo codes in terms of decoding performance, complexity, and convergence. The simulation results show it has decoding gain near to log-MAP while keeps almost the same computation complexity as max-log-MAP with slight increment in memory resource. Moreover, it maintains insensitivity to SNR estimation error of max-log-MAP algorithm. Compared with conventional scaling scheme, it accelerates extrinsic information exchange between two constituent decoders to get better convergence and decoding performance.

**Keywords:** Turbo codes · Extrinsic information · Scaling factor

## 1 Introduction

Turbo codes have been widely adopted in many wireless communication standards such as CCSDS and LTE due to their excellent error-correcting performance. The constituent soft-input, soft-output (SISO) decoders exchange their extrinsic information by employing maximum a posteriori (MAP) algorithm iteratively. To reduce computational complexity, it is usually performed in logarithmic domain, turning to log-MAP (LM) algorithm. For further simplicity of hardware implementation, sub-optimal algorithms such as max-log-MAP (MLM) were proposed. MLM algorithm has the least computational complexity and owns the advantage of insensitivity to SNR mismatch [7,12]. However, it has about 10% performance degradation relative to the optimal LM [8].

To reduce the performance loss of MLM, several trade-off modification methods were proposed. Substantially, there are two kinds. The first kind aims at


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_42

substituting the correction function in LM with its approximation variants [2,6]. Although their performance can get close to the optimal LM, they are susceptible to SNR mismatch [4], especially in high-order modulation cases, which limits its application in high-throughput systems [12]. The other kind is the extrinsic information (EI) scaling scheme [1,11,13]. It makes compensations by multiplying a constant value, which is called the scaling factor (SF), on the EI from output of the SISO decoder. It is insensitive to SNR mismatch error as MLM algorithm [3,4]. Moreover, compared with MLM, it provides an additional 0.2–0.4 dB decoding gain with almost the same computational complexity [11]. EI scaling prompts the mutual information exchange between constituent decoders to accelerate the convergence of turbo decoding process, which can be monitored by extrinsic information transfer (EXIT) chart [5,10].

In this paper, a novel EI scaling and combining scheme is proposed. It features the scaling combination of EI at both current and previous iteration rounds. It retains the properties of EI scaling algorithm and further improves the decoding performance as well. EXIT analysis indicates that it accelerates the convergence of the decoding process compared with conventional scaling scheme. The paper is arranged as follows. In Sect. 2, we review conventional MAP family turbo decoding algorithms. The scaling combination scheme is proposed in Sect. 3. In Sect. 4, we give the simulation results in three aspects of performance, complexity, and convergence. Finally, conclusions are drawn in Sect. 5.

## 2 Review of MAP Family Turbo Decoding Algorithms

In this section, we briefly review some typical MAP family decoding algorithms in turbo codes, including the optimal algorithm and its sub-optimal variants, which contain various notations and guidance that are necessary to our following statement.

### 2.1 The Log-MAP Algorithm

Define log-likelihood ratio (LLR) for reliability estimation of $k$th information bit $u_k$ as:

$$L(u_k) := \ln \frac{\Pr(u_k = 0|\mathbf{Y})}{\Pr(u_k = 1|\mathbf{Y})} \tag{1}$$

where $\mathbf{Y}$ is the received sequence. Then, the a posteriori probability of $u_k$ can be computed as follows:

$$\begin{aligned} L_{apo}(u_k) = &\max^*_{(s',s),u_k=0} [A_{k-1}(s') + M_k(s',s) + B_k(s)] \\ &- \max^*_{(s',s),u_k=1} [A_{k-1}(s') + M_k(s',s) + B_k(s)] \end{aligned} \tag{2}$$

where $A_k(s)$, $B_k(s)$, and $M_k(s',s)$ are forward, backward, and branch metrics in logarithmic domain. Recursive relations for first two metrics are:

$$A_k(s) = \max^*_{s' \in (s',s)} (A_{k-1}(s') + M_k(s',s)) \tag{3}$$

$$B_k(s) = \max^*_{s\in(s',s)}(M_{k+1}(s',s) + B_{k+1}(s)) \tag{4}$$

For bivariable cases, max* operation is defined as:

$$\max^*(a,b) := \ln(e^a + e^b) \tag{5}$$

For multivariable cases, the computation can be operated recursively. Using Jacobian logarithmic simplification [6], it can be written as:

$$\begin{aligned}\max^*(a,b) &= \max(a,b) + \ln(1 + e^{-|a-b|}) \\ &= \max(a,b) + f_c(a,b)\end{aligned} \tag{6}$$

The term $f_c(a,b)$ is called the correction function. LM algorithm is optimal as it takes all the trellis paths into consideration and calculates the maximum a posteriori probability values for each information bit. However, it is not suitable for hardware implementation because of its nonlinear computation in correction function. Moreover, it is susceptible to SNR mismatch.

### 2.2 The Max-Log-MAP Algorithm

Replacing max* function with max approximation in (2)–(4), we attain MLM algorithm. It neglects the correction term, so it has the least computation complexity but becomes sub-optimal consequently. It has performance degradation of 0.3–0.5 dB relative to optimal LM. However, it is a basic decoding algorithm in most hardware implementations. There are two main reasons for this. Firstly, MLM merely contains addition (subtraction) and maximum comparison selection operations to make the add-compare-select (ACS) unit qualified to its computation tasks [9]. Secondly, MLM is insensitive to the estimation of channel noise as the susceptible correction function part is discarded. Therefore, the decoder can work without accurate channel estimation.

## 3 Scaling Max-Log-MAP Algorithm

After SISO decoder generates the a posteriori information $L_{apo}(u_k)$, EI is extracted as follows:

$$L_{ext}(u_k) = L_{apo}(u_k) - L_{apr}(u_k) - L_{int}(u_k) \tag{7}$$

where $L_{apo}(u_k)$, $L_{ext}(u_k)$, and $L_{int}(u_k)$ on right-hand side indicate a posteriori information, a priori information (AI), and the channel intrinsic information of $u_k$, respectively. In turbo decoding process, EI output from one constituent decoder becomes AI input of the other. In this way, EI is exchanged between two SISO constituent decoders alternately. Finally, after a fixed number of iterations, the estimation of $u_k$ is attained from $L_{apo}(u_k)$.

EI scaling scheme is a modification to MLM indirectly. It is independent of a posteriori information generation. Therefore, its computational complexity is nearly the same as MLM. For conventional scaling scheme of MLM (S-MLM) [1,11], its additional burden is merely a multiplier. In comparison, the decoding performance is improved by 0.2–0.4 dB under optimum scaling factor in case of 3GPP standard.

As information entropy of EI is strongly related to its variance, increasing it properly will help faster converge during iteration [5]. Inspired by Chase's combining method in HARQ mechanism, an EI scaling and combining scheme of MLM (SC-MLM) is proposed:

$$L_{ext}(u_k) = s[L_{apo}(u_k) - L_{apr}(u_k) - L_{int}(u_k)] + (1-s)L_{ext}^{last}(u_k) \tag{8}$$

where $s$ is called scaling factor (SF) and $L_{ext}^{last}(u_k)$ is the EI generated at the last iteration. For the first iteration round, it is initialized to 0. As shown in (8), both scaling and combining operations are used for the evolution of EI.

The structure of the SC-MLM decoder is depicted in Fig. 1. The subscripts $a$ and $e$ denote the AI and EI, respectively, while the superscript *last* denotes the EI generated at the last iteration. Compared with conventional turbo decoder, the EI should be saved (dotted line) temporarily for scaling and combining at the next iteration; therefore, additional memory is required.

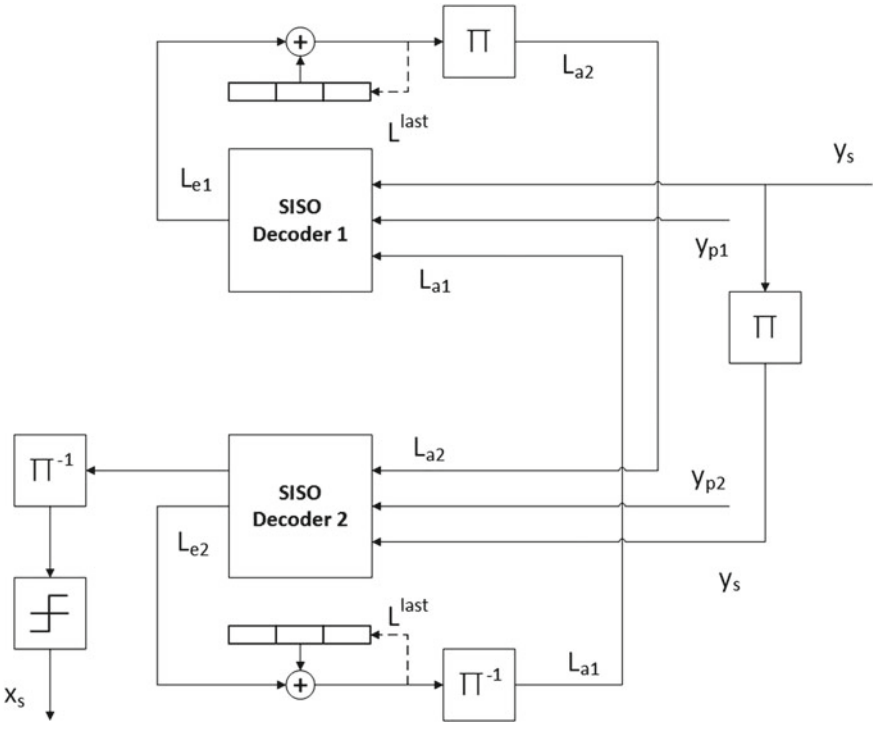



**Fig. 1.** The structure of SC-MLM Turbo decoder

## 4 Simulation Results

Unless otherwise mentioned, our proposed method is evaluated for generator polynomial $(13,15)_{\text{oct}}$ Turbo codes in 3GPP LTE standard. The code length is 2048 with rate 1/3. They are transmitted over AWGN channel with QPSK modulation. All the BER results are obtained through over $10^5$ transmission frames. For a decoder, the maximum number of iterations is 4 and no SNR mismatch is assumed. The simulation results comparison will be conducted in three aspects including decoding performance, computational complexity, and iteration convergence.

### 4.1 Performance Analysis

Figure 2a shows the BER performance between LM, MLM, linear LM (LLM) [2], and two scaling MLM schemes versus $E_b/N_0$ ranging from 0 to 2dB. SF is set to 0.7. In practice, considering about throughput and latency, SISO decoder usually works in a parallel manner, whose results are shown in Fig. 2b with a parallel degree of 8. It should be noted that LM is no longer optimal under parallel decoding. The simulation results show that the improvement under parallel decoding is quite more evident.

SF is a crucial parameter in scaling scheme. The BER performance versus a range of SF from 0.5 to 0.95 for SC-MLM is shown in Fig. 3 under $E_b/N_0$ 1.2, 1.5, and 1.8 dB conditions. SF value corresponding to the lowest BER is considered to be optimum. The simulation results imply irregular relevance between optimum SF and SNR.

Sensitivity to SNR mismatch between MLM, SC-MLM, and LLM is shown in Fig. 4 under 1.2 dB $E_b/N_0$. The label of $x$-axis $\Delta E_b/N_0$ denotes SNR mismatch, where negative values represent underestimation, while positive values stand for overestimation. The simulation results show LLM is sensitive to SNR mismatch as LM, especially under underestimation condition. Comparatively, SC-MLM maintains the insensitive characteristic of MLM.

### 4.2 Complexity Analysis

The computational complexity of both recursive metrics and EI is shown in Table 1. Here, $N$ is the code length, $M$ is the number of encoder shift registers, and $X$ is defined as $X = N * 2^M$. For simplicity, logarithmic and exponential operations are considered as identical, the same with subtractions and additions. As the scaling schemes only have an impact on EI computation, their operation numbers are nearly equal.

### 4.3 Convergence Analysis

The EXIT chart and iterative decoding trajectories are given in Fig. 5, where $I_A$ and $I_E$ represent the mutual information of AI and EI respectively. Vertical lines

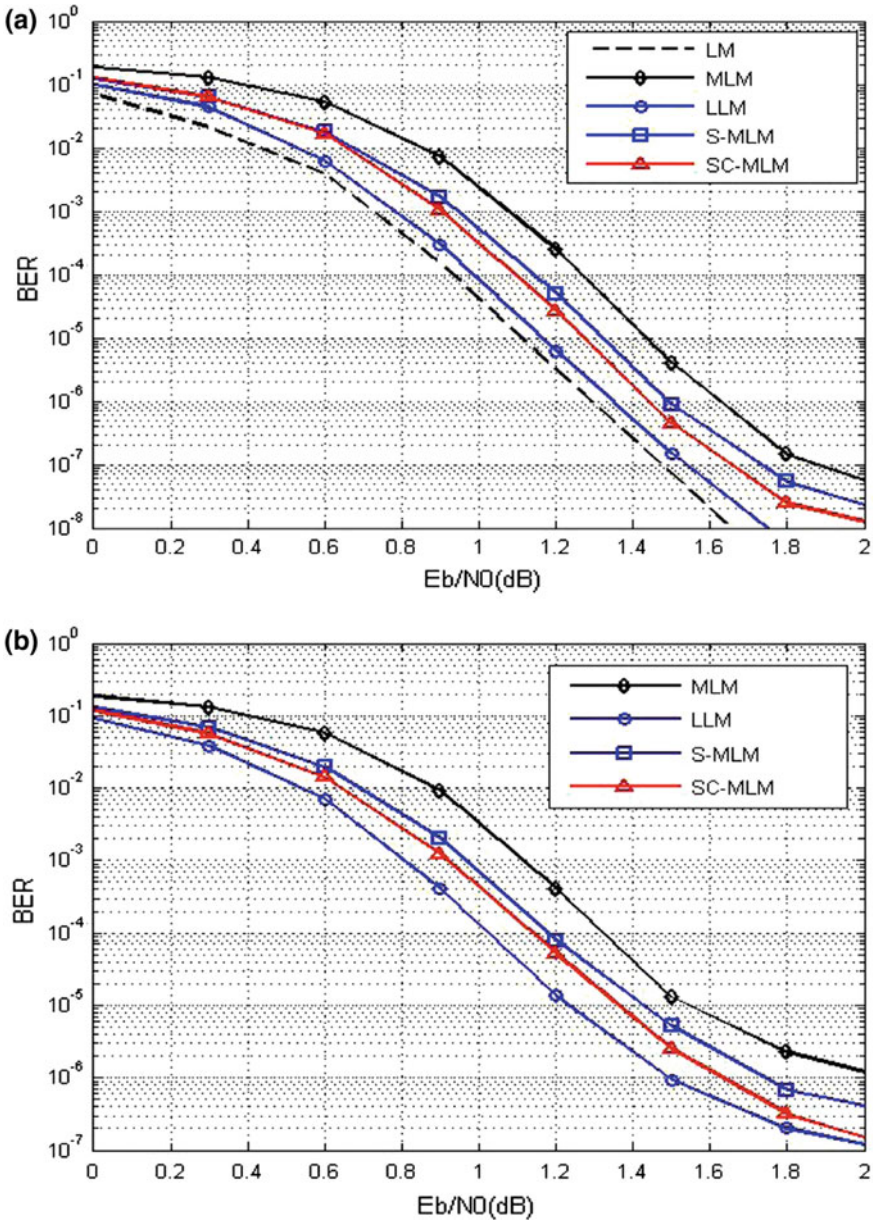



**Fig. 2.** BER performance comparison. **a** serial decoding, **b** parallel decoding

in trajectories represent half-iteration in SISO decoder 1, while horizontal lines denote that in the other SISO decoders. EXIT curve is the envelope of optimal LM decoding trajectories. Usually, decoding procedure starts at the original

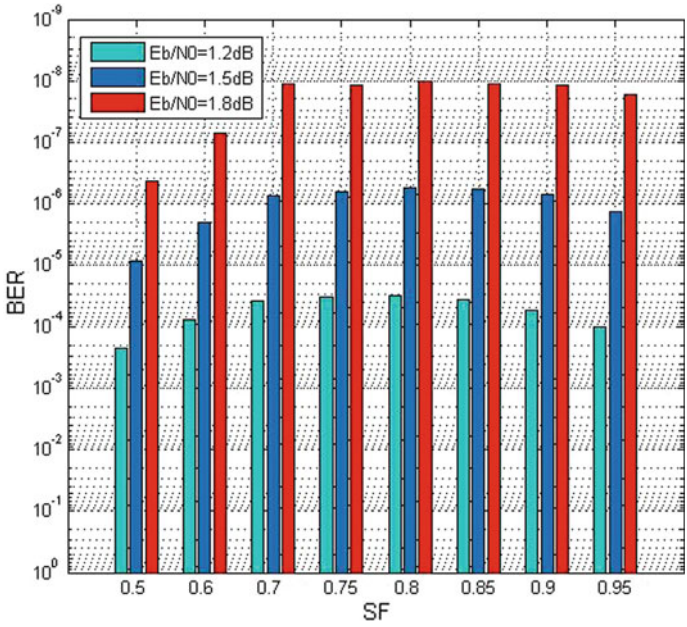



**Fig. 3.** BER performance versus different scaling factor

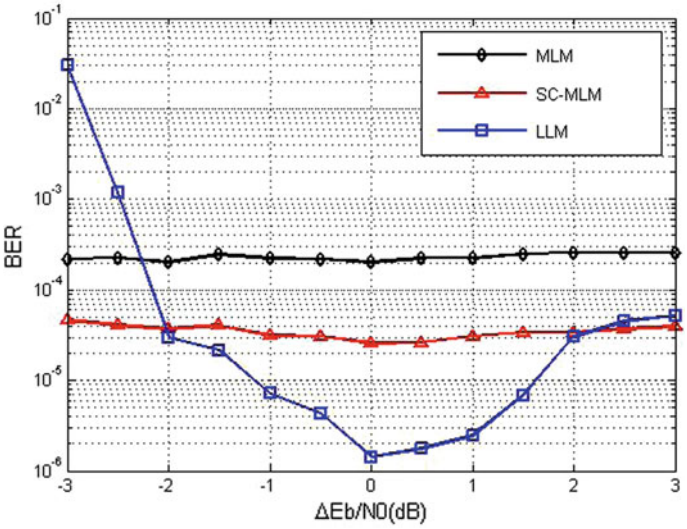



**Fig. 4.** BER performance comparison under SNR mismatch

**Table 1.** Operation numbers in both recursive metrics and EI computation

| | Exponents | Comparisons | Additions | Multiplications |
|---|---|---|---|---|
| LM | $7X + 4N$ | 0 | $9X + 2N$ | 0 |
| MLM | 0 | $3X - 4N$ | $6X + 6N$ | 0 |
| LLM | 0 | $3X - 4N$ | $18X - 2N$ | $5X - 4N$ |
| S-MLM | 0 | $3X - 4N$ | $6X + 6N$ | $2N$ |
| SC-MLM | 0 | $3X - 4N$ | $6X + 8N$ | $4N$ |

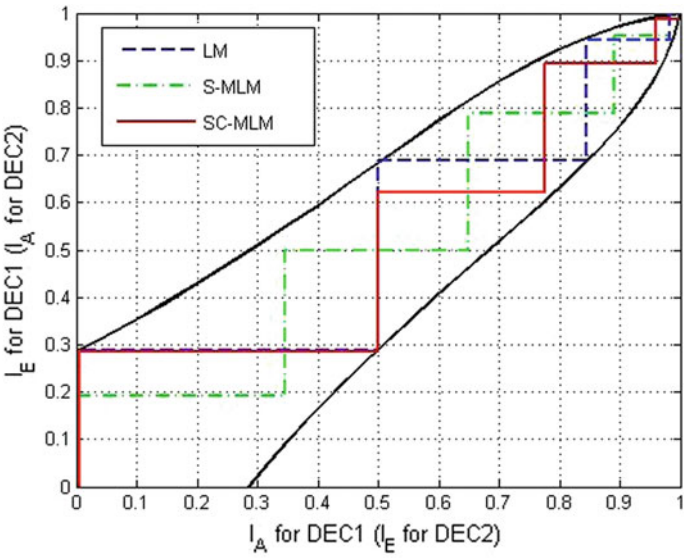



**Fig. 5.** EXIT chart and iterative decoding trajectories

point and terminates at the right top after several half-iterations. Assume either mutual information larger than 0.9 for sufficient EI exchange to terminate the iteration; LM and SC-MLM need three rounds, while S-MLM requires 4. It can be observed from trajectories that SC-MLM scheme accelerates EI exchange to make iterative decoding converge faster. The simulation is conducted under BPSK modulation with 1dB $E_b/N_0$.

## 5 Conclusion

In this paper, a novel EI scaling and combining scheme for MLM algorithm is proposed to improve turbo decoding performance. It prompts EI exchange to converge faster by EI scaling and combining. Compared with conventional MLM, it provides additional 0.2 dB decoding gain, which is more obvious under parallel decoding, while maintains almost the same computational complexity.

Compared with LM, it inherits the insensitivity to SNR mismatch of MLM. Compared with S-MLM, it accelerates the EI exchange to make earlier iteration termination. The simulation is also made to find the optimum SF. In a hardware implementation, additional memory should be needed for saving EI generated at each iteration round. In conclusion, it is a well trade-off scheme between performance and complexity in turbo decoding.

## References

1. Alvarado A, Nunez V, Szczecinski L, Agrell E. Correcting suboptimal metrics in iterative decoders. In: IEEE international conference on communications; 2009. p. 1–6.
2. Cheng JF, Ottosson T. Linearly approximated log-map algorithms for turbo decoding. In: Vehicular technology conference proceedings, 2000. Vtc 2000-Spring Tokyo, vol 3. 2000 IEEE; 2000. p. 2252–56.
3. El-Khamy M, Wu J, Lee J, Roh H, Kang I Near-optimal turbo decoding in presence of SNR estimation error. In: Global communications conference; 2013. p. 3737–42.
4. El-Khamy M, Wu J, Lee J, Kang I. Online log-likelihood ratio scaling for robust turbo decoding. IET Commun. 2013;8(2):217–26.
5. Hagenauer J. The exit chart—introduction to extrinsic information transfer in iterative processing. In: 2004 European signal processing conference; 2004. p. 1541–48.
6. Liu Z, Wu B, Ye T. Improved turbo decoding with multivariable taylor series expansion. IEEE Commun Lett. 2017;99:1–1.
7. Papaharalabos S, Mathiopoulos PT, Masera G, Martina M. Non-recursive max* operator with reduced implementation complexity for turbo decoding. IET Commun. 2012;6(7):702–7.
8. Robertson P, Villebrun E, Hoeher P. A comparison of optimal and sub-optimal map decoding algorithms operating in the log domain. In: IEEE international conference on communications, 1995. ICC'95 Seattle, gateway to globalization, vol 2; 1993. p. 1009–13.
9. Roth C, Belfanti S, Benkeser C, Huang Q. Efficient parallel turbo-decoding for high-throughput wireless systems. IEEE Trans Circ Syst I Regul Pap 2014;61(6):1824–35 (2014)
10. Sapra R, Jagannatham AK. Exit chart based ber expressions for turbo decoding in fading MIMO wireless systems. IEEE Commun Lett. 2015;19(1):10–3.
11. Vogt J, Finger A. Improving the max-log-map turbo decoder. Electron Lett. 2000;36(23):1937–9.
12. Worm A, Hoeher P, Wehn N. Turbo-decoding without snr estimation. IEEE Commun Lett. 2000;4(6):193–5.
13. Yue DW, Nguyen HH. Unified scaling factor approach for turbo decoding algorithms. IET Commun. 2010;4(8):905–14.

# An Efficient Classification Method of Uncertain Data with Sampling

Jinchao Huang[1(✉)], Yulin Li[2], Kaiyue Qi[3], and Fangqi Li[3]

[1] School of Cyber Security, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China
hjc2015@sjtu.edu.cn

[2] School of Computer Engineering, University of Illinois at Urbana-Champaign, 508 E University Ave., Champaign, IL 61820, USA
yulinli2@illinois.edu

[3] School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China
{tommy-qi,solour_lfq}@sjtu.edu.cn

**Abstract.** Current research on the classification for uncertain data mainly focuses on the structural changes of the classification algorithms. Existing methods have achieved encouraging results; however, they do not take an effective trade-off between accuracy and running time, and they do not have good portability. This paper proposed a new framework to solve the classification problem of uncertain data from data processing point. The proposed algorithm represents the distribution of raw data by a sampling method, which means that the uncertain data are converted into determined data. The proposed framework is suitable for all classifiers, and then, XGBoost is adopted as a specific classifier in this paper. The experimental results show that the proposed method is an effective way of handling the classification problem for uncertain data.

**Keywords:** Classification · Uncertain data · Sampling · XGBoost

## 1 Introduction

In recent years, with rapid developments of the Internet and intelligent, the efficiency of social has been unprecedentedly improved. The effect, that machine learning and artificial intelligence can achieve, is far more than the effect of the manual operation. In logical thinking, e-commerce, medical, financial, and some other fields, business decisions become ever more reliant on machine learning and artificial intelligence. After the optimization, artificial intelligence systems can help to deliver goods and information to destinations at a faster speed and a lower price. As an important part of artificial intelligence, the performance


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_43

of the classification algorithm has receiving increasing attention. More accurate classification means a better sense of user experience, which also brings a lot of benefits to businesses.

Most of the existing classification methods are applied only to determined data, such as decision tree classification [10], K-nearest neighbor (KNN) classification [7], naive Bayes classification [5], artificial neural network classification, support vector machine (SVM) classification [13], and so on. However, in practical applications, data often have characteristics of randomness, fuzziness, and multiple uncertainty. Specifically, in many cases, we cannot get accurate data, but obtain the probability distribution of data, where traditional classification methods no longer apply.

There has been a growing attention on uncertain data mining. Zhang and Bi first proposed classification algorithms for uncertain data [1], and they extended SVM to process the data affected by noise. Qin et al. [8] and Smith [12] both improved C4.5 and presented uncertain decision tree algorithms DTU and UDT separately. There are also some scholars who modified naive Bayes classifiers to deal with the uncertain data [6,9,11].

The above methods mainly pay attention to improvements of model itself. Most of them need complicated computing and therefore cannot satisfy the requirement of real time in practical applications. In this paper, we propose a sampling-based classification algorithm for uncertain data. The new algorithm focuses on the transformation of data form, in particular to taking advantage of sampling to convert probability distribution form to determined data. When sample times are suitable, the data obtained can represent the distribution of raw data perfectly, and then, the original problem can be transformed into a traditional classification problem. Empirical comparisons conducted on a set of datasets demonstrate the effectiveness of our proposed approach.

The rest of this paper is organized as follows. In Sect. 2, we give a brief introduction about the presence of data uncertainty. The new method is described in Sect. 3. Descriptions of datasets and experiment results are presented in Sect. 4. Section 5 concludes the paper and discusses the prospect of further research.

## 2 Modeling in the Presence of Uncertainty

Like traditional methods, a dataset of the classification algorithm for uncertain data consists of $d$ training tuples, represented as $\{t_1, t_2, \ldots, t_d\}$, and each tuple is associate with a feature vector $V_i = (f_{i,1}, f_{i,2}, \ldots, f_{i,k})$ and a class label $c_i \in C$. The difference is that each member $f_{i,j}$ of the feature vector is represented by probability distribution rather than deterministic value, as in Table 1. Table 1 shows a sequence of uncertain samples about credit card transaction, and each row represents a record.

The common used method is to build a model that maps each feature vector composed of pdf to a probability distribution $P$ on $C$. Specifically, when given a tuple $t_0 = (f_{0,1}, f_{0,2}..., f_{0,k}, c_0)$, the trained model will predict the class label $c_0$ based on $P_0 = M(f_{0,1}, ..., f_{0,k})$ with high accuracy. And it is defined that $P_0$

**Table 1.** A sequence of uncertain samples

| ID | Age | Income (million yuan) | Maximum purchase (million yuan) | Purchase time | Label |
|---|---|---|---|---|---|
| s1 | 18–23 | 9.3–11.4 | 0.2–0.3 | 19:00–23:00 | 0 |
| s2 | 27–34 | 13.9–17.0 | 0.6–0.7 | 16:00–20:00 | 1 |
| s3 | 50–61 | 13.1–16.1 | 0.1–0.2 | 16:00–19:00 | 0 |
| s4 | 54–67 | 10.0–12.2 | 0.2–0.3 | 13:00–16:00 | 0 |
| s5 | 27–33 | 17.0–21.6 | 0.6–0.7 | 10:00–13:00 | 1 |

predicts $c_0$ if $c_0 = \arg\max_{c \in C}\{P_0(c)\}$. Especially in [12], before starting training model, a pdf would be stored as a set of s sample points by approximating the associated value $f_{i,j}$ with a discrete distribution, and this algorithm has achieved quite good results on accuracy. However, existing methods do not take an effective trade-off between accuracy and running time. And there is still plenty of room for development.

## 3 The New Method

As discussed above, most of the existing methods change the classifier structure to make it applicable for uncertain data, where the feature form of input data is a probability distribution. These approaches are applied only to some specific classifiers, which means poor portability. Hence, in this paper, we propose a new method that is suitable for all classifiers.

### 3.1 Framework of the New Method

The framework of the new method is shown in Fig. 1. As we can see in Fig. 1, the new method mainly contains two parts: the training phase and the testing phase. And the training part can also be divided into data processing section and model training section.

In data processing, we propose a sampling method to convert uncertain data to determined data. When facing the feature form of input data is a probability distribution, the sampling method samples each feature of every tuple according to its distribution, and then, we can obtain a determined tuple composed of sampling values. By taking repeated sampling method, the raw data are expanded by a series of sampling tuples, which also means the distributions of raw data are represented by these sampled values. After that, the determined data are used as input to train the classifier. Then, the same operation is performed on the test data. It is easy to see that the original problem is transformed into a traditional classification problem, and this framework is suitable for all classifiers.

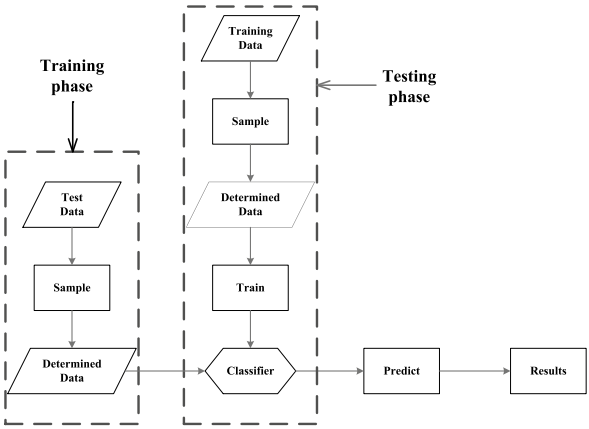



**Fig. 1.** Framework of the new method

### 3.2 Classifier Fitting

Gradient boosting algorithm is a powerful machine learning technique and has produced some state-of-the-art results in data mining. It learns an ensemble of trees which makes a trade-off between classification accuracy and model complexity. In this paper, we adopt XGBoost [2], a boosting algorithm proposed recently, as a classifier. When facing a large-scale dataset, XGBoost is able to obtain the trained model with good accuracy in a short time. The above advantage of XGBoost ensures that our method always achieves a good effect.

## 4 Experiment and Results

### 4.1 Datasets

To evaluate the performance of the new method, we conduct experiments on several standard benchmark datasets. The standard datasets are obtained from the UCI machine learning repository [3], including ionosphere, satellite, wavefore-21, and wavefore-40. The details are shown in Table 2. And because there is no benchmark uncertain dataset that can be found, we model the data uncertainty with a Gaussian distribution and a controllable parameter $w$, given in [12]. Supposing that the value of features is $x_i$, and $|X_i|$ is the threshold of original feature $X_i$, then we define that $a_{i,t} = x_{i,t} - |X_i|$ and $b_{i,t} = x_{i,t} + |X_i|$, and they are regarded as mean and variance of the Gaussian distribution $f_{i,j}$, separately. From these, we can see that parameter $w$ plays the role of determining the degree of injected uncertainty.

### 4.2 Evaluation Metrics and Comparison Baselines

To measure the performance of various methods, we use accuracy and running time to analyze the results.

**Table 2.** Statistics of datasets

| Datasets | Feature numbers | Number of training set | Number of test set | Class numbers |
|---|---|---|---|---|
| Ionosphere | 34 | 200 | 151 | 2 |
| Satellite | 36 | 4435 | 2000 | 8 |
| Wavefore-21 | 21 | 4000 | 1000 | 3 |
| Wavefore-40 | 40 | 4000 | 1000 | 3 |

We compare the new method with two existing algorithms: VFDT [4] and UDT [12]. VFDT is a decision tree algorithm with fast learning speed when values of experimental data are determined. And UDT is a decision tree algorithm for data with static uncertainty, which has achieved a high accuracy. Note that, in the experiments, except for the VFDT algorithm, all the other algorithms use the data injected uncertainty as input.

### 4.3 Results and Analysis

We first examine the accuracy of the algorithms, and the results are shown in Figs. 2, 3, 4, and 5, which correspond to four datasets, respectively. For each figure, the horizontal axis is the value of parameter $w$, and the vertical axis is accuracy. As mentioned above, VFDT is a special case, which runs on datasets without uncertainty, so the performance curve of VFDT is always a straight line. From these figures, it is easy to see that the new method performs much better than the others. In Fig. 2, all accuracies of the new method achieve more than 94% under four different $w$ conditions, and accuracies of the other two algorithms only attain 90% or so. As for Waveform-40 dataset, the promotion is more obvious, where the precision is improved more than 10% when $w = 0.01$. The results of all datasets have fully verified that our method has a great effect on accuracy rate, which also indicate that our sampling strategy can describe the distribution of raw data well.

Then, we study the execution time of the algorithms, and the results are represented in Table 3. Because the running time of UDT is too long, we only give the results of VFDT and our proposed method. From Table 3, we can see that the running times of the two algorithms are about the same. Experiments on ionosphere, waveform-21, and waveform-40 demonstrate that our method is faster, but result on satellite is reversed. In addition, the running time of our new method increases slowly with the increase of uncertainty, meaning that it is more difficult to construct a proper model when there is more uncertainty. On the whole, our method has good performance on the execution time.

Such good results prove that the new method is an effective way of handling the classification problem for uncertain data.

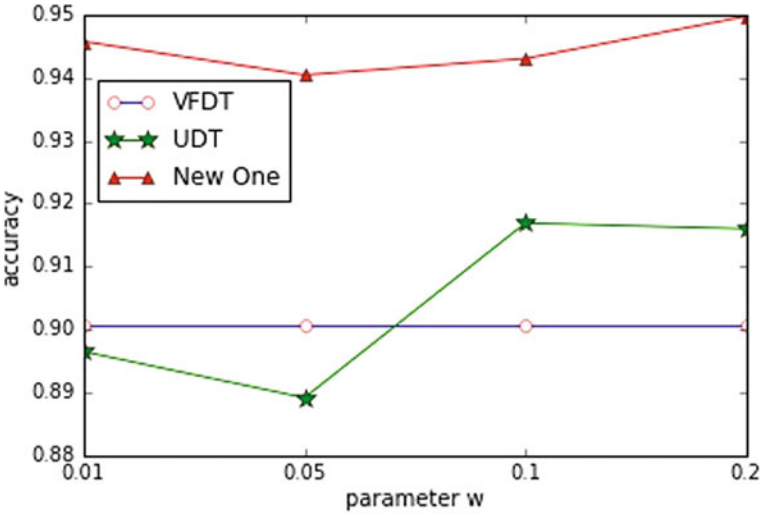



**Fig. 2.** Comparisons of accuracy (ionosphere)

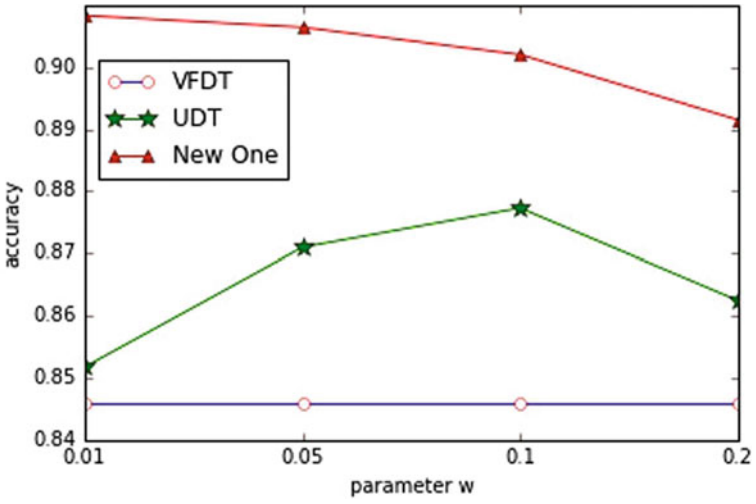



**Fig. 3.** Comparisons of accuracy (satellite)

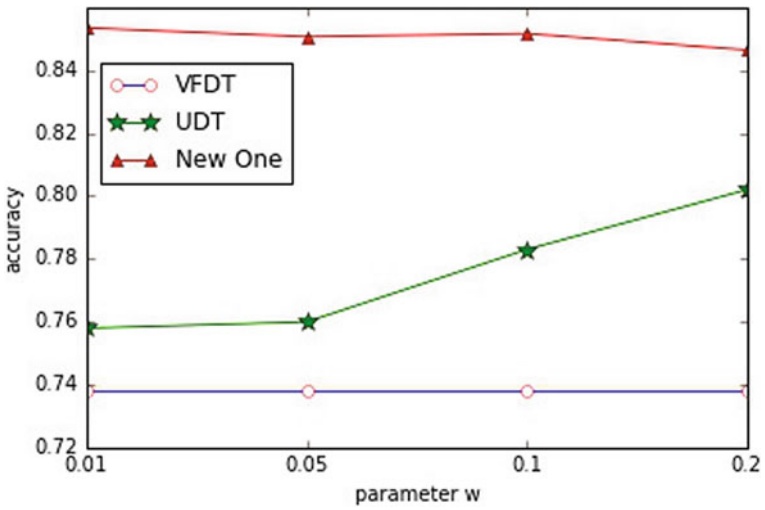



**Fig. 4.** Comparisons of accuracy (wavefore-21)

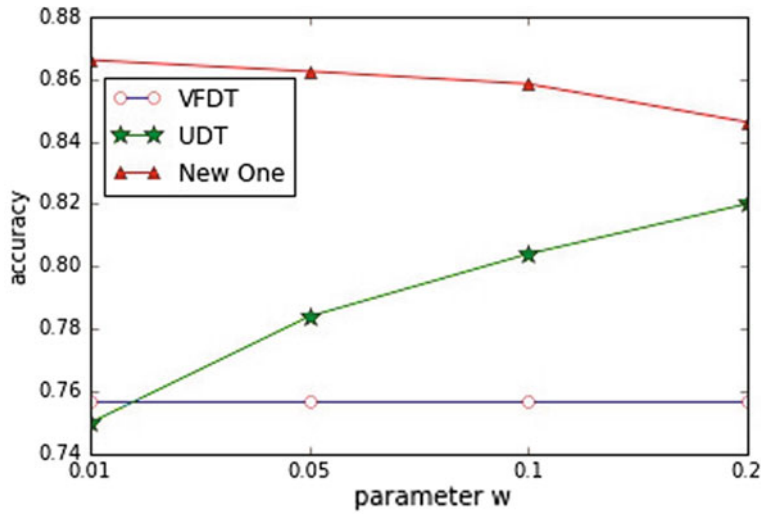



**Fig. 5.** Comparisons of accuracy (wavefore-40)

**Table 3.** Comparisons of running time (unit: second)

| Datasets | VFDT | New one ($w = 0.01$) | New one ($w = 0.05$) | New one ($w = 0.1$) | New one ($w = 0.2$) |
|---|---|---|---|---|---|
| Ionosphere | 0.46 | 0.107 | 0.105 | 0.114 | 0.108 |
| Satellite | 4.12 | 14.11 | 14.65 | 16.62 | 18.04 |
| Wavefore-21 | 5.62 | 4.96 | 5.19 | 5.5 | 6.12 |
| Wavefore-40 | 10.62 | 7.24 | 7.71 | 8.07 | 8.72 |

## 5 Conclusion

In this paper, we propose a new framework to solve the classification problem of uncertain data, from data processing point. The new framework is suitable for all classifiers, and we adopt XGBoost as a classifier, specifically. Then, we conduct a series of experiments on several standard benchmark datasets. The results show that the new method obtains a pretty good performance on accuracy and execution time, which fully proves that the proposed method is an effective way of handling the classification problem for uncertain data.

**Acknowledgments.** This research work is funded by the National Key Research and Development Project of China (2016YFB0801003).

## References

1. Bi J, Zhang T. Support vector classification with input data uncertainty. In: Proceedings of neural information processing systems, vol 17; 2004. p. 161–8.
2. Chen T, Guestrin C. Xgboost: a scalable tree boosting system; 2016. p. 785–94. (2016)
3. Dheeru D, Karra Taniskidou E. UCI machine learning repository; 2017. URL http://archive.ics.uci.edu/ml
4. Domingos P, Hulten G. Mining high-speed data streams. In: Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining; 2000. p. 71–80.
5. Duda RO, Hart PE. Pattern classification and scene analysis. Hoboken: Wiley; 1973.
6. He J, Zhang Y, Li X, Wang Y. Bayesian classifiers for positive unlabeled learning. Berlin, Heidelberg: Springer; 2011.
7. Peterson L. K-nearest neighbor. Scholarpedia. 2009;4(2):1883.
8. Qin B, Xia Y, Li F. DTU: a decision tree for uncertain data. In: Advances in knowledge discovery and data mining, Pacific-Asia conference, PAKDD 2009, Bangkok, Thailand, April 27–30, 2009, Proceedings; 2009. p. 4–15.
9. Qin B, Xia Y, Wang S, Du X. A novel bayesian classification for uncertain data. Knowl-Based Syst. 2011;24(8):1151–8.
10. Quinlan JR. Induction on decision tree. Mach Learn. 1986;1(1):81–106.

11. Ren J, Lee, SD, Chen X, Kao B, Cheng R, Cheung D. Naive bayes classification of uncertain data. In: IEEE international conference on data mining; p. 944–9.
12. Tsang S, Kao B, Yip KY, Ho WS, Lee SD. Decision trees for uncertain data. IEEE Trans Knowl Data Eng. 2011;23(1):64–78.
13. Vapnik VN. The nature of statistical learning theory. Technometrics. 1997;8(6):1564.

# An Improved Robust Kalman Filter for Real-Time Detection of Cycle Slips in the Single-Frequency Carrier Phase Measurements Validated with BDS Data

Ye Tian(✉) and Yizhe Jia

Academy of Space Electrical Information Technology, Xi'an 710100, China
tianye_504@163.com

**Abstract.** The detection of cycle slips and outliers in single-frequency carrier phase data, or any other type of un-expected changes in the single-frequency carrier phase measurements of the GNSS, is one of the major data preprocessing problems that needs to be addressed, especially single-frequency receivers account for most of the market share and GNSS carrier phase data are used for real-time applications that require reliable position results. In this contribution the improved RKF (Robust Kalman Filter) is designed to detect cycle slips or unexpected changes when single-frequency carrier phase is interfered by small outliers, in order to improve the cycle slip detection success rates. Real BDS single-frequency data have been used to test and evaluate the algorithm, where the simulation results indicate that the improved RKF has a higher cycle slip detection success rate than the RKF when observations are interfered by small outliers, proving the efficiency of the algorithm proposed in this paper.

**Keywords:** Improved RKF · Cycle slip detection · Small outlier · BDS

## 1 Introduction

Carrier phase measurements, which are applied in high-precision positioning applications, often suffer from cycle slips and outliers because of signal interruptions, low signal-to-noise ratios, and poor receiver device accuracy [1]. Such unexpected changes should be detected and repaired [2], if possible, before carrier phase measurements are used as high-precision measurements in real time or near real time. Clearly, this is an onerous requirement, but one which ensures that all other data-processing configurations, such as conventional data post-processing, can be catered for using the same algorithm, though perhaps with more relaxed testing criteria.

Currently, single-frequency receivers account for most of the market share, and detection and repair algorithms in real time based on single-frequency measurements are extremely valuable. At present the Kalman filter is mainly used to detect and repair cycle slips and outliers [3]. A new cycle slip detection method is proposed in the Ref. [4] on the basis of the grey theory. In the Ref. [5], the outlier test is performed on the residual of the polynomial model to detect the cycle slips. The above mentioned methods can't guarantee detection success rate when measurements contain outliers


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_44

and positioning applications have high real-time requirements. Therefore, the cycle slip detection method for single-frequency carrier phase measurements should have a level of performance of robustness to ensure the accuracy of the detection in real time applications.

In designing a real time robust detection algorithm, one task has to be performed [6, 7]. Generate residuals that can robustly reflect abrupt changes in the single-frequency measurements in real time, which can be considered as detection of cycle slips [8]. In this paper the improved RKF is proposed and analysed based on a simulation study and numerical experiments. In addition, the performance, in which the basic RKF and the method proposed in this paper is applied for the cycle slips repair procedure, is assessed based on the real BDS data. Last, the methodology described in this paper is summarized, and recommendations regarding quality control of single frequency carrier phase measurements are presented.

## 2 Methodology

### 2.1 The Robust Kalman Filter

First of all, we briefly introduce the RKF is used as the basis for the cycle slip detection algorithms. We assume that the phase measurements at time $t_{k+1}$ can be represented by a constant acceleration model:

$$\varphi_{k+1} = \varphi_k + \dot{\varphi}_k \Delta t + \ddot{\varphi}_k \left( \frac{\Delta t^2}{2!} \right) + \text{error}_k \tag{1}$$

where $\Delta t$ is the time difference $t_{k+1} - t_k$, $\varphi_k$ is the phase measurements in cycles at time $t_k$, $\dot{\varphi}_k$ is the phase rate and $\ddot{\varphi}_k$ is the phase acceleration. The error term in Eq. 1 represents a random process of zero mean. A system of dynamics in discrete formulation can be developed from Eq. 1 as follows:

$$\mathbf{x}_{k+1} = \mathbf{\Phi} \mathbf{x}_k + \mathbf{w}_k, \quad k = 1, 2, 3 \ldots \tag{2}$$

where $\mathbf{\Phi}$ is the transition matrix, the process error $\mathbf{w}_k$ is assumed to be a stationary white-noise sequence with covariance matrix $\mathbf{Q}_k$. The GNSS measurements $\mathbf{L}_k$ would be the phase and phase rate, accordingly, the measurements equations may be written in the form:

$$\mathbf{L}_k = \mathbf{A}_k \mathbf{x}_k + \mathbf{v}_k, \quad k = 1, 2, 3 \ldots \tag{3}$$

The error $\mathbf{v}_k$ is assumed to be a stationary white-noise sequence with covariance matrix $\mathbf{R}_k$. The extrapolated state vector and error covariance matrices are given as the discrete Kalman filter solutions:

$$\bar{\mathbf{x}}_k = \mathbf{\Phi}_{k,k-1} \hat{\mathbf{x}}_{k-1} \quad \overline{\mathbf{P}}_k = \mathbf{\Phi}_{k,k-1} \widehat{\mathbf{P}}_{k-1} \mathbf{\Phi}_{k,k-1}^T + \mathbf{Q}_k \tag{4}$$

The innovations and its covariance can be written in form:

$$\overline{\mathbf{V}}_k = \mathbf{A}_k\bar{\mathbf{x}}_k - \mathbf{L}_k \quad \mathbf{Q}_{\overline{\mathbf{V}}_k} = \mathbf{A}_k\overline{\mathbf{P}}_k\mathbf{A}_k^T + \mathbf{R}_k \tag{5}$$

In order to overcome the outlier influence on the filtering results, the RKF uses the equivalent weight matrix $\widetilde{\mathbf{K}}_k$ based on IGGIII to control filter results. The RKF equivalent weight matrix is defined as:

$$\widetilde{\mathbf{K}}_k = \begin{cases} \mathbf{K}_k & (f_k \le k_0) \\ \mathbf{K}_k \frac{k_0}{f_k}\left[\frac{k_1 - f_k}{k_1 - k_0}\right]^2 & (k_0 < f_k \le k_1) \\ \mathbf{0} & (f_k > k_1) \end{cases} \tag{6}$$

where $\mathbf{K}_k$, which is defined as $\mathbf{K}_k = \mathbf{\Phi}_{k,k-1}\mathbf{A}_k\mathbf{Q}_{\overline{\mathbf{V}}_k}^{-1}$, is the standard Kalman filter equivalent weight matrix, $k_0$ and $k_1$ are two thresholds, $f_k$ is the standardized residual, and the updated state vector and its covariance matrix at the sampling position $k$ as:

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + \widetilde{\mathbf{K}}_k\overline{\mathbf{V}}_k \quad \widehat{\mathbf{P}}_k = (\mathbf{I} - \widetilde{\mathbf{K}}_k\mathbf{A}_k)\overline{\mathbf{P}}_k(\mathbf{I} - \mathbf{A}_k^T\widetilde{\mathbf{K}}_k^T) + \widetilde{\mathbf{K}}_k\mathbf{R}_k\widetilde{\mathbf{K}}_k^T \tag{7}$$

## 2.2 The Improved Robust Kalman Filter

When the RKF is used to detect the cycle slips or unexpected changes, the detection results are susceptible to the small outliers. So it is necessary to make the innovations overcome the effects of small outliers to improve the detection rate. In order to reduce the amount of variation in the innovations it may be assumed that the GNSS observations $\mathbf{L}_k$ contain errors which are serially and exponentially correlated. Therefore, the measurements errors $\mathbf{v}_k$ are described as the output of a known linear system driven by white noise $\boldsymbol{\varepsilon}_k$:

$$\mathbf{v}_{k+1} = \boldsymbol{\psi}\mathbf{v}_k + \boldsymbol{\varepsilon}_k \tag{8}$$

where $\boldsymbol{\psi}$ is a transition matrix and the process error $\boldsymbol{\varepsilon}_k$ is a random which follows a normal distribution with zero mean and covariance matrix $\mathbf{Q}_\varepsilon$, uncorrelated with the system noise $\mathbf{w}_k$. Define a weighted difference $\boldsymbol{\zeta}_k$ between the next measurements $\mathbf{z}_{k+1}$ and the current measurements $\mathbf{z}_k$, which does not contain any the sequentially correlated measurements noise of $\mathbf{v}_k$, as follows:

$$\begin{aligned} \boldsymbol{\zeta}_k &= \mathbf{z}_{k+1} - \boldsymbol{\psi}\mathbf{z}_k = \{\mathbf{A}\mathbf{\Phi} - \boldsymbol{\psi}\mathbf{A}\}\mathbf{x}_k + \{\mathbf{A}w_k + \varepsilon_k\} \\ &= \widetilde{\mathbf{A}}x_k + \tilde{\mathbf{v}}_k \quad k = 1, 2, 3\ldots \end{aligned} \tag{9}$$

where the uncorrelated random errors and design matrix as follows:

$$\tilde{\mathbf{v}}_k = \mathbf{A}\mathbf{w}_k + \varepsilon_k \quad \widetilde{\mathbf{A}} = \mathbf{A}\boldsymbol{\Phi} - \boldsymbol{\psi}\mathbf{A} \tag{10}$$

The covariance matrix of the error vector $\tilde{\mathbf{v}}_k$, is, omitting subscripts for the sake of simplicity:

$$\begin{aligned} \widetilde{\mathbf{R}}_k &= E\left\{\tilde{\mathbf{v}}\tilde{\mathbf{v}}^T\right\} = E\left\{[\mathbf{A}\mathbf{w}+\boldsymbol{\varepsilon}][\mathbf{A}\mathbf{w}+\boldsymbol{\varepsilon}]^T\right\} \\ &= \mathbf{A}E\left\{\mathbf{w}\mathbf{w}^T\right\}\mathbf{A} + E\left\{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathbf{T}}\right\} = \mathbf{A}\mathbf{Q}\mathbf{A}^T + \mathbf{R} \end{aligned} \tag{11}$$

In order to reduce the impact of small outliers, we can use a sequential and recursive form for determination of a linear-minimum-variance estimate of the states $E[\mathbf{x}_{k+1}|\mathbf{z}_{k+1}]$ at the $k+1$ sampling position. The filter estimate of state $E[\mathbf{x}_{k+1}|\mathbf{z}_{k+1}]$ and its covariance matrix $\mathbf{M}_{k+1}$ are written in the form:

$$\begin{aligned} E[\mathbf{x}_{k+1}|\mathbf{z}_{k+1}] &= \boldsymbol{\Phi}_k E[\mathbf{x}_k|\mathbf{z}_{k+1}] + \mathbf{Q}_k\mathbf{A}_k^T(\widetilde{\mathbf{R}}_k)^{-1}\left\{\boldsymbol{\zeta}_k - \widetilde{\mathbf{A}}_k E[\mathbf{x}_k|\mathbf{z}_{k+1}]\right\} \\ \mathbf{M}_{k+1} &= \tilde{\boldsymbol{\Phi}}_k\mathbf{P}_k\left(\widetilde{\boldsymbol{\Phi}}_k\right) + \widetilde{\mathbf{Q}}_k \end{aligned} \tag{12}$$

where the matrix involved are defined by the expressions:

$$\widetilde{\boldsymbol{\Phi}}_k = \boldsymbol{\Phi}_k - \mathbf{Q}_k\mathbf{A}_k^T(\widetilde{\mathbf{R}}_k)^{-1}\mathbf{A}_k \quad \widetilde{\mathbf{Q}}_k = \mathbf{Q}_k - \mathbf{Q}_k\mathbf{A}_k^T(\widetilde{\mathbf{R}}_k)^{-1}\mathbf{A}_k\mathbf{Q}_k^T \tag{13}$$

For the first measurement $\mathbf{z}_1$ there is insufficient information to form the vector of measurement differences. So we calculate the measurement difference $\boldsymbol{\zeta}_1$ from the second measurements, and we also calculate the covariance matrix as initial conditions from second measurements. The measurement innovations $\widetilde{\mathbf{V}}_{k+1}$, in which $E[\mathbf{x}_{k+1}|\mathbf{z}_{k+1}]$ is used to mitigate the small outliers, can be defined as:

$$\widetilde{\mathbf{V}}_{k+1} = \mathbf{z}_{k+1} - \mathbf{A}E[\mathbf{x}_{k+1}|\mathbf{z}_{k+1}] \tag{14}$$

with covariance matrix $\mathbf{Q}_{\widetilde{\mathbf{V}}_{k+1}}$ is defined as:

$$\mathbf{Q}_{\widetilde{\mathbf{V}}_{k+1}} = E\left\{\widetilde{\mathbf{V}}_{k+1}\widetilde{\mathbf{V}}_{k+1}^T\right\} = \mathbf{A}\mathbf{M}_{K+1}\mathbf{A}^T \tag{15}$$

Herein $\widetilde{\mathbf{V}}_k$ and $\mathbf{Q}_{\widetilde{\mathbf{V}}_k}$ are used instead of $\overline{\mathbf{V}}_k$ and $\mathbf{Q}_{\overline{\mathbf{V}}_k}$. We also use the equivalent weight matrix $\widetilde{\mathbf{K}}_k$ based on IGGIII as Eq. 6 defined to control filter results. But the robust residual $f_k$ is redefined as:

$$f_k = \frac{\left\|\widetilde{\mathbf{V}}_k\right\|_2}{\sqrt{\left\|\mathbf{Q}_{\widetilde{\mathbf{V}}_k}\right\|_2}} \tag{16}$$

$\hat{\mathbf{x}}_k$ and $\widehat{\mathbf{P}}_k$ can be calculated by the Eq. 7. The improved RKF, using measurements differences $\zeta_k$ and state $E[\mathbf{x}_{k+1}|\mathbf{z}_{k+1}]$, can reduce the amount of variation in the innovations and mitigate the small outliers impact on innovations. These improved steps and methods can effectively improve the cycle slips or unexpected change detection success rate regardless of small outliers, which is validated with BDS data in the next subsection.

## 3 Test Description

The RKF and the improved RKF are separately applied to real single-frequency BDS data which the B1 frequency data is selected under normal ionospheric conditions. Both real and simulated cycle slips and unexpected changes are processed together. The experiment adopts the data, which sampling interval is 30 s, in JFNG observation station on January 6, 2017, using measurement data of C03 stood for GEO satellite and C13 stood for MEO satellite.

In the simulation, the small outlier of 0.12 cycle is added on C03 phase measurements at a random epoch and the cycle slip of 1 cycle is added on the same measurements after the small outlier epoch. In Fig. 1, the RKF is used to detect the unexpected change of carrier phase measurements, showing the innovation which is defined in Eq. 5 and the innovation which is defined in Eq. 14, with the same measurements from the improved RKF. By comparison, it can be seen that RKF can't overcome the impact of small outliers, resulting in failure detection of cycle slips, but the improved RKF can successfully detect the cycle slips, which shows better robustness.

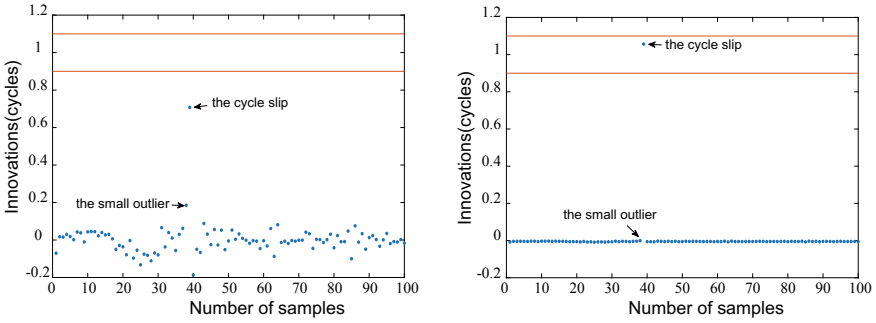



**Fig. 1.** The innovations of C03 from the RKF and improved RKF

In the another simulation, artificial cycle slips and outliers as shown in Table 1, in which the integer represents the cycle slips and the decimal represents the outliers, are added on the C13 carrier phase measurements. From Fig. 2, we can obviously see that RKF can't accurately detect the cycle slips when carrier phase measurements before the cycle slips is affected by the small outliers, while the improved RKF can detect the

**Table 1.** Artificial cycle slips and outliers

| Epoch | $k-4$ | $k-3$ | $k-2$ | $k-1$ | $k$ | $k+1$ | $k+2$ |
|---|---|---|---|---|---|---|---|
| Value | +0.11 | −1 | +0.15 | +1 | 0 | +0.4 | −1 |

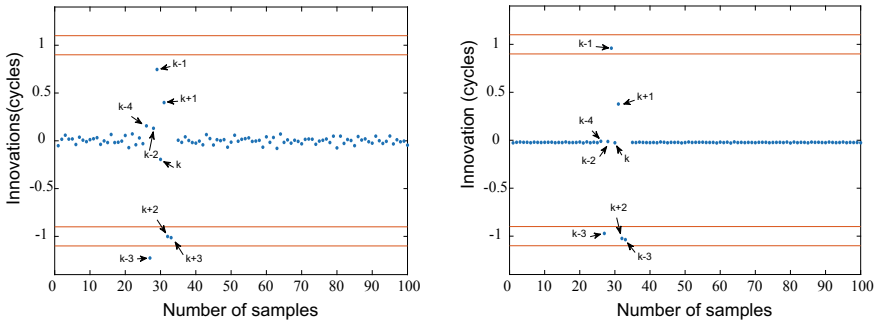



**Fig. 2.** The innovations of C13 from the RKF and improved RKF

cycle slips more accurately in the same situation. And we can also see that when carrier phase measurements before the cycle slips is affected by the larger outliers, both methods have the same ability to detect the cycle slips.

When the outliers with different standard deviations (SD) are added to the epoch before the cycle slips, Table 2 shows the cycle slip detection rates for the two methods. From Table 2, we can see that when the standard deviation of outlier is small than 0.3, the RKF has a lower detection rates. It is worth mentioning that when the outlier is small, in which standard deviation from 0.1 to 0.2, the residual of RKF $f_k$ is less than $k_0$ leading to the $\widetilde{\mathbf{K}}_k$ and $\mathbf{K}_k$ are equal, and the best estimate $\hat{\mathbf{x}}_k$ is affected by the small outlier which lead the $\bar{\mathbf{x}}_{k+1}$ isn't optimally estimated. So the RKF can't overcome the impact of small outliers. However, for the improved RKF, it can control the standard deviation of the innovations and reduce the outliers impact on the next epoch, which can effectively improve the detection success rate.

**Table 2.** Artificial cycle slips and outliers

| SD of outliers (cycle) | 0.1 | 0.2 | 0.3 | 0.4 |
|---|---|---|---|---|
| RKF (%) | 75.36 | 73.67 | 72.30 | 70.56 |
| Improved RKF (%) | 85.07 | 77.82 | 72.39 | 70.53 |

## 4 Conclusions and Remarks

In this paper the improved RKF has been investigated and applied to the single-frequency carrier phase measurements. The improved RKF takes advantage of weighted differences and the linear-minimum-variance state to reduce the small outliers

effect on cycle slip detection. This algorithm validated with BDS Data is also capable of fast and real-time feedback, identifying when any cycle slips or abrupt changes occur, as well as controlling the standard deviation of the innovations to reduce the impact of the outliers. According to the experimental results show that the improved RKF has a higher cycle slip detection rate than the RKF when the previous epoch when the epoch before the cycle slip is disturbed by small outliers. All in all, the improved RKF could easily be extended to the other data or other systems except BDS, and this algorithm appears to be a good choice for the single-frequency carrier phase measurements quality control. Further analysis on the accuracy of single-frequency cycle slips detection under the case of more complex outlier interference will be investigated in future.

## References

1. Ju B, Gu D, Chang X et al. Enhanced cycle slip detection method for dual-frequency BeiDou GEO carrier phase observations. GPS Solutions, 2017; p. 1–12.
2. Zhang X, Zeng Q, He J, Kang C. Improving TurboEdit real-time cycle slip detection by the construction of threshold model. Geomatics Inf Sci Wuhan Univ. 2017;42(3):285–92.
3. Zhou W. Research and realization on theories and methods of precise positioning base on BeiDou navigation satellite system. PLA Information Engineering University, 2013.
4. Wang C, Wang J. Cycle slip detection and correction of single frequency undifferenced phase observation. J Tongji Univ (Nat Sci). 2012;40(9):1393–8.
5. Li X, Wu M, Zhang K. Applying outlier test theory to detect and correct cycle-slip. Chin J Sci Instrum. 2012;33(10):2315–21.
6. Yao Y, Gao J, Li Z, Tan X. Research on robust kalman filter of observations with unequal precision in precise point positioning. Geomatics Inf Sci Wuhan Univ. 2017;42(3):314–20.
7. Zhang X, Pan Y, Zuo X, Wang J. An improved robust kalman filtering and its application in PPP. Geomatics Inf Sci Wuhan Univ. 2015;40(7):858–64.
8. Yang Y, He H. Adaptive robust filtering for kinematic GPS positioning. Acta Geodaetica Et Cartographica Sinica. 2001;30(4):293–8.

# A Low-Complexity Shorten Regenerating Code with Optimal Repair Bandwidth

Ke Li(✉), Shushi Gu, Ye Wang, Jian Jiao, and Qinyu Zhang

Communication Engineering Research Center, Harbin Institute of Technology, Shenzhen, Guangdong, China
like@stu.hit.edu.cn
{gushushi,wangye,jiaojian,zqy}@hit.edu.cn

**Abstract.** Regenerating codes (RGCs) are considered as optimal trade-off between repair bandwidth and storage amount per node in a distributed storage system (DSS). Actually, due to the limited nodes' amount and bandwidth resources in networks, redundant bandwidth cost and assisted nodes connections will involve as employing the traditional RGC. Specific to these problems, we propose a new code, named shorten minimum storage regenerating code (sMSR) with two novel targets, unit storage cost (USC) and unit repair bandwidth (URB), and construct it by removing some information bits of the mother code generated by product matrix in encoding process. Additionally, in order to improve the availability of sMSR, we implement Binary Addition and Shift Implementable Convolution (BASIC) to decrease the computation complexity. The simulation results demonstrate that our code improves repair efficiency of MSR codes in practical DSS.

**Keywords:** Distributed storage system · Shorten codes · Regenerating codes · BASIC · Repair bandwidth

## 1 Introduction

In a distributed storage system (DSS) by a large number of inexpensive and unreliable devices, it is crucial to make some efforts to avoid the data loss caused by frequent node failures. A typically feasible mechanism is to introduce redundancy, which can be achieved by two common strategies: *replication* and *erasure codes* [1]. Replication is the simplest scheme, but increases storage overhead since storing multiple copies. Erasure codes enable tolerance to failures with the lower storage overhead at the expense of excessive bandwidth. For repairing one node requires, maximum distance separable (MDS) codes have to download the whole file to reconstruct only a small fraction.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_45

By reviewing the repair problem as a single-source, multi-broadcast problem, Dimakis et al. proposed *Regenerating codes* (RGCs) reaching optimal trade-off between storage amount per node and repair bandwidth required reconstructing one failed node [2]. $(n, k, d)$ Regenerating codes consist of $n$ storage nodes that $\alpha$ data symbols each, and download $\beta(\leq\alpha)$ symbols from each of $d(\geq k)$ arbitrary surviving nodes to generate a failed node. The parameter set of a regenerating code that aims to store the maximized file of size $B$ reliably must satisfy the following condition.

$$B \leq \sum_{i=0}^{k-1} \max\{\alpha, (d-i)\beta\} \tag{1}$$

For these codes, if the bandwidth is constrained to be minimum, a trade-off between the repair bandwidth $d\beta = \gamma$ and the storage per node $\alpha$ can be set with the two extreme points being achieved for exact-regenerating codes. The obtained extreme points are referred to as the minimum bandwidth regenerating (MBR) and minimum storage regenerating (MSR) [2]. Rashmi researched further on the trade-off given in [2] and presented approaches to construct MSR and MBR codes in a product matrix framework (PM-MSR and PM-MBR) [3].

Conventionally, the more assisted nodes connected when repairing, the less bandwidth occupied. Due to nodes' amount and bandwidth resources in DSS are limited, it is worth researching how to decrease the connected nodes and repair bandwidth. Specific to these problems, Oggier et al. put forward the concept of repairable locality denoting the number of assisted nodes [4] and proposed locally repairable regenerating codes [5] based on the polynomial operation. In our paper, we combine the shorten codes [6] with regenerating codes, and then propose a construction method of shorten MSR codes based on product matrix framework. Moreover, to enhance the practicability of regenerating codes, we also implement the Binary Addition and Shift Implementable Convolution (BASIC) [7] in our code to decrease the computation complexity. The simulation results show that our coding scheme reduces the repair bandwidth and computing consumption compared with the traditional MSR codes for practical DDS.

## 2 Shorten Regenerating Codes

### 2.1 Unit Storage Cost and Unit Repair Bandwidth

Conventionally, the performance of regenerating code with specific parameters $(n, k, d)$ can be evaluated by two targets $\alpha$ and $\gamma$, which, respectively, denotes storage and repair bandwidth cost per node. However, under the circumstance that bandwidth resource of network is limited, such evaluations may be inaccurate. For example, consider PM-MSR codes (product matrix minimum storage RGCs), with three different $(n, k, d)$ as in Table 1, and prescribe that the parameter constraints are $d = 2k - 2, \beta = \gamma/d = 1$. In this case ,we can easily get the

**Table 1.** Performances of PM-MSR codes with different $(n, k, d)$

| $n$ | $k$ | $d$ | $\alpha$ | $\gamma$ | $B$ | $\delta$ | $n\alpha\delta$ | $\gamma\delta$ |
|---|---|---|---|---|---|---|---|---|
| 4 | 2 | 2 | 1 | 2 | 2 | 3 | 12 | 6 |
| 5 | 2 | 2 | 1 | 2 | 2 | 3 | 15 | 6 |
| 6 | 3 | 4 | 2 | 4 | 6 | 1 | 12 | 6 |

storage cost per node $\alpha = d - k + 1 = k - 1$ and the affordable size of data block each coding process. In Table 1, we hypothesis the data size $M = 6$ and hence we can work out the stripe pieces $\delta = M/B$. Finally, the actual storage and repair bandwidth overhead in the whole process are $n\alpha\delta$ and $\gamma\delta$. From Table 1, we can observe that when we just take $\alpha$ and $\gamma$ as the criteria, the $(4, 2, 2)$ and $(5, 2, 2)$ PM-MSR coding schemes ought to perform more excellently, but due to the bandwidth limitation $B$, we have to consider the stripe pieces $\delta$, and in this case, the $(6, 3, 4)$ PM-MSR code becomes the best. Hence, to validly evaluate the actual performance of specific RGC, we propose two novel targets *unit storage cost* (*USC*) and *unit repair bandwidth* (*URB*), which, respectively, expresses hard disk space occupied each unit stored data block and hard disk space occupied each unit stored data block, and their computational formula are shown in (2) and (3).

$$USC_{MSR} = n \cdot \frac{k-1}{k(k-1)} = \frac{n}{k} \tag{2}$$

$$URB_{MSR} = \frac{d}{k(k-1)} \tag{3}$$

### 2.2 Analysis of Shorten Regenerating Codes

In last subsection, we propose two new targets to choose optimal coding parameters, while the problem that the required nodes are overmuch persists. To solve the problem, this subsection presents a scheme—shorten MSR code (sMSR), which is mainly achieved by removing $t$ information bits of PM-MSR [3]. Assuming the original parameter set of regenerating codes is $(n, k, d)[\alpha, \gamma, B]$, after being shortened, it turns to $(n_s, k_s, d_s)[\alpha_s, \gamma_s, B_s]$.

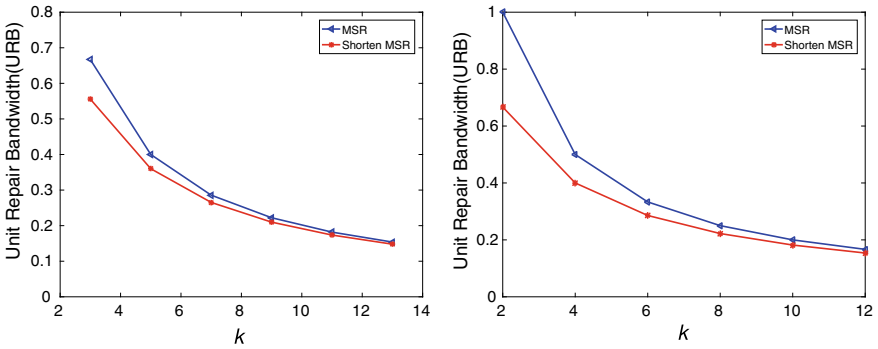
$$\begin{aligned} n_s &= n - t \\ k_s &= k - t \\ d_s &= d - t \\ \alpha_s &= \alpha \\ \gamma_s &= \gamma - t \\ B_s &= B - t \cdot \alpha \end{aligned} \tag{4}$$

Shorten MSR codes have the same parameter constraint $d \geq 2k-2$ as PM-MSR codes. For convenience, we unify the constraint, that is, $d = 2k-2$. Through (2)–(4), we obtain the $USC$ and $URB$ of sMSR with $(n_s = n, k_s = k, d_s = 2k+t-2)$.

$$USC_{sMSR} = \frac{(n)(k+t-1)}{(k)(k+t-1)} = \frac{n}{k} \tag{5}$$

$$URB_{sMSR} = \frac{2k+t-2}{(k+t-1)(k)} \tag{6}$$

Fairly, compare the formulas in the same parameter set $(n, k, 2k+t-2)$; we discover that the (USC) is unchanged, but $URB$ makes a difference displayed in Fig. 1. Figure 1 illustrates shortening indeed reduce the $URB$, indicating that it needs less nodes connected but repair bandwidth unchanged, and the more bits shortened, the smaller $URB$ is.




(a) MSR code and sMSR code with $(t = 1)$

(b) MSR code and sMSR code with $(t = 2)$

**Fig. 1.** Unit repair bandwidth of MSR code and sMSR code.

### 2.3 Construction of Shorten MSR Regenerating Codes

Our coding scheme is established on product matrix framework [3], and the construction of mother code can be obtained from [3]. As already stated, the encoder output is a $n \times \alpha$ codematrix $\boldsymbol{C}$. The central idea in shortened MSR codes is to built a $d \times \alpha$ message matrix $\boldsymbol{M}$ which has redundant information in it enabling $t$ rows of output codematrix all 0. (1) sMSR Encoder: A shortened MSR code has the same coefficient matrix $\boldsymbol{\Psi}$ as the PM-MSR mother code, which can be chosen as an $n \times d$ Vandermonde matrix as (7) shows. The message matrix $\boldsymbol{M}$ is a $d \times \alpha$ matrix, whose elements stem from a encoder input

vector $(u_1, u_2, \ldots, \widetilde{u_{d+1}}), \widetilde{u_{d+2}}, \ldots, \widetilde{u_B})$ of length $B$, where $\widetilde{u_{d+1}}), \widetilde{u_{d+2}}, \ldots, \widetilde{u_B}$ are redundant symbols whose value from the consideration that the original $t$ rows of codematrix. For example, assume the original row of codematrix is $\underline{c}_6 = (u_1 + 6u_2 + 10u_4 + 8\widetilde{u_5}u_2 + 6u_3 + 10\widetilde{u_5} + 8\widetilde{u_6})$. To turn it to zero vector, we can set $\widetilde{u_5} = -(u_1 + 6u_2 + 10u_4)/8$ and $\widetilde{u_6} = -(u_2 + 6u_3 + 10u_5)/8$.

$$\boldsymbol{\Psi} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & \Psi_1 & \Psi_2 & \ldots & \Psi_{d-1} \\ 1 & \Psi_1^2 & \Psi_2^2 & \ldots & \Psi_{d-1}^2 \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & \Psi_1^{n-1} & \Psi_2^{n-1} & \ldots & \Psi_{d-1}^{n-1} \end{bmatrix} \tag{7}$$

(2) sMSR Decoder: The data collector contacts any $k_s (k_s = k - t)$ nodes to construct data matrix $\boldsymbol{\Psi}_{S-DC}\boldsymbol{M}$. $\boldsymbol{\Psi}_{S-DC}$ is polished with $t$ row vectors to obtain a $k \times d$ new coefficient matrix $\boldsymbol{\Psi}_{DC}$. Fill the code matrix $\boldsymbol{\Psi}_{S-DC}\boldsymbol{M}$ with $t$ rows' zero vectors, attaining a new data matrix $\boldsymbol{\Psi}_{DC}\boldsymbol{M}$. Decode it as PM-MSR does.

(3) sMSR Repair: The newcomer connects $d_s$ assisted nodes denoted as $h_1, h_2, \ldots, h_{d-1}$, and attains a $d_s \times d$ coefficient matrix written as $\boldsymbol{\Psi}_{s-repair}$. Fill $\boldsymbol{\Psi}_{s-repair}$ with $t$ shortened nodes' coefficient vectors to compose a novel $d \times d$ coefficient matrix $\boldsymbol{\Psi}_{repair}$. Download $\beta$ data from each assisted node to form a $d_s \times \beta$ data matrix, and fill it with $t$ rows' zero vectors, getting a $d \times \beta$ data matrix to repair the disabled node.

## 3 Complexity Optimization of Shorten MSR Codes

### 3.1 Complexity Analysis of Shorten MSR Codes

Primarily, we state the complexity of finite field operations. Suppose a primitive polynomial $g(z)$ with the highest power series $\omega$ in finite field $GF(2^\omega)$, polynomial addition needs $\omega$ times XOR, and multiplication needs $\mu\omega$ times XOR, $\mu = (\omega - 1) + \|g(z)\|_0$. Now we turn to research complexity of shorten MSR. Solving matrix equation involves frequently in decoding and repair process, so we regard it as a mathematical model to investigate the complexity. Suppose a matrix equation $\boldsymbol{Ax} = \boldsymbol{B}$, $\boldsymbol{A}$ is a $n \times n$ invertible matrix, both $\boldsymbol{x}$ and $\boldsymbol{B}$ are $n \times 1$ matrix, while $\boldsymbol{x}$ unknown. Solving $\boldsymbol{A}^{-1}$ needs $\mu\omega \cdot \frac{n^2(n+1)(n-2)+(n+2)(n-1)}{2} + n(2n-1)\sum_{i=2}^{n-1} i(i-1)\omega$ times XOR. $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{B}$ needs $n^2\mu\omega + n(n-1)\omega$ times, simply approximately, the total XOR times is $\frac{n^4}{2}\mu\omega$.

Assume the original file totally has $Bm$ symbols divided into $B$ original code blocks with $m$ symbols each. The finite field will be $GF(2^\omega)$, and $\omega = m$. In encoding process, first, work out the $t\alpha$ complementary redundancy by the $t\alpha$ sets of equations, the XOR times is $\frac{t^4\alpha^4}{2}\mu\omega$. Then, we multiple the $(n + t) \times (2k + 2t - 2)$ coefficient matrix and the $(2k + 2t - 2) \times \alpha$ message matrix, which total needs $\alpha(n + t)(2k + 2t - 2)$ times multiplications and $\alpha(n + t)(2k + 2t - 3)$ times additions. Converting to XOR, it totally operates $\alpha(n + t)\omega[(2k +$

$2t-2)\mu+(2k+2t-3)]$ times, and averaging to every original symbol, it is $2(n+t)\mu+\frac{t^4k^3\mu}{2(k+t-1)}$. Similarly, we can determine the XOR times in decoding process and repair process average to every symbol, which, respectively, are $\frac{\mu(k+t)[2(k+t)^2+t+1]}{k}$ and $\frac{3(2k+t+2)\beta\mu}{k}$.

We can observe that the complexity of every process is deeply related to $\mu$, while $\mu=(\omega-1)+\|g(z)\|_0$; thus, we can recognize the complexity is closely bound up to the size of finite field and weight of primitive polynomial. Table II shows the weight of primitive polynomial in finite field.

### 3.2 Shorten MSR Codes Based on BASIC

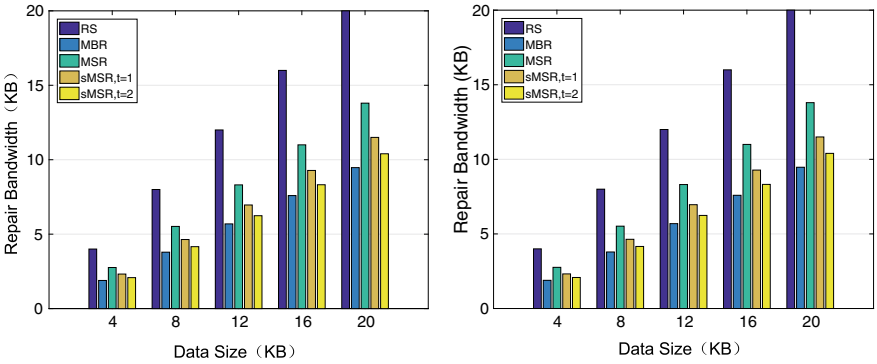
The central idea of Binary Addition and Shift Implementable Convolution (BASIC) is to convert the complicated polynomial multiplication in finite field $GF(2^w)$ to binary addition and shift by convolution, which can be adaptive for computer processing. The shift of BASIC can be realized by simply pointer move in physical layer; therefore, we can ignore the complexity of multiplication times. Meanwhile, the addition in BASIC is bitwise XOR, which depends on the polynomial term with the highest order defined as the degree of polynomial. Hou et al. determined the numbers of XOR in BASIC MSR average to every symbol each process [7]. Because BASIC sMSR cannot guarantee data being shifted of any number of nodes is 0, we cannot shorten bits arbitrarily. We make the constraint to the number of shorten node is 1.

The coefficient matrix $\boldsymbol{\Psi}$ of BASIC sMSR can be chosen as Vandermonde matrix based on shift operation $z$ like (8), and the other processes keep the same. The first row of code matrix $\boldsymbol{C}$ is all 0 denoting the shorten node. Due to the shift in BASIC sMSR which could be $(n-1)(2k-1)$ when the parameter is $(n,k)$, BASIC sMSR may occupy a little more repair bandwidth than sMSR. However, in actual DSS, the magnitude of stored data is GB at least, so in this case, the bandwidth overhead brought by shift can be ignored, which means BASIC sMSR can maintain high performance of sMSR when used in actual DSS.

$$\boldsymbol{\Psi}=\begin{bmatrix}1 & 1 & 1 & 1 & 1\\ 1 & z & z^2 & \dots & z^{d-1}\\ 1 & z^2 & z^4 & \dots & z^{2(d-1)}\\ 1 & \vdots & \vdots & \ddots & \vdots\\ 1 & z^{n-1} & z^{2(n-1)} & \dots & \Psi^{(d-1)(n-1)}\end{bmatrix} \tag{8}$$

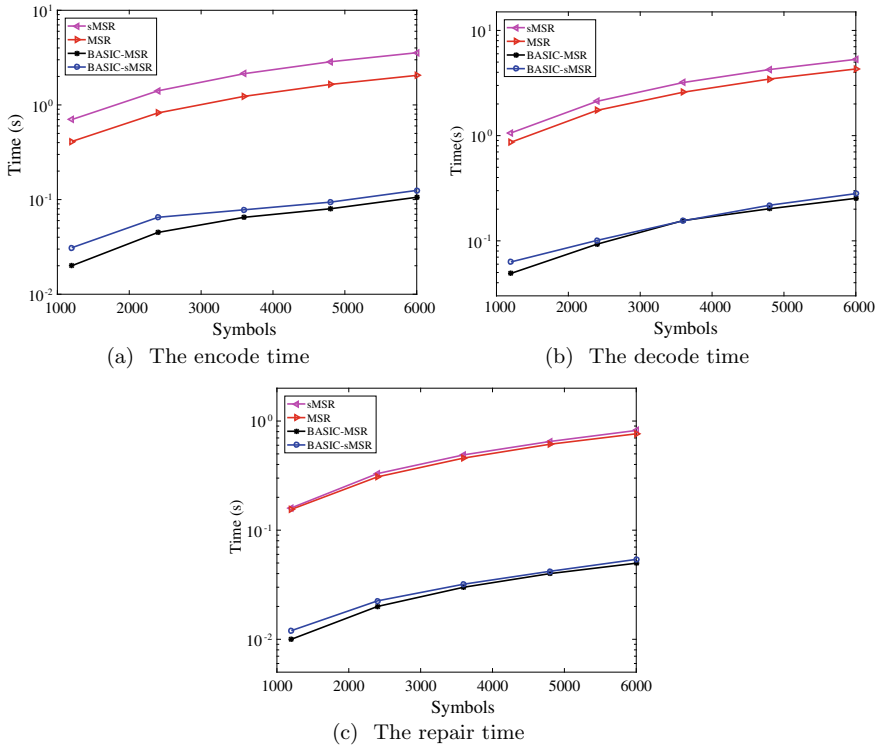
## 4 Simulation and Comparison

In this section, we first simulate the performance of RS, MBR. MSR. sMSR ($t=1$ and $t=2$) with two different coding parameters, (a) $n=8,k=3$; (b) $n=16,k=6$, in processing five original files with size of 4, 8, 12, 16, 20 KB. The comparison diagrams are displayed in Fig. 2. From Fig. 2, we can observe




(a) n=8,k=3 (b) n=16,k=6

**Fig. 2.** Repair bandwidth comparison of RS, MBR, MSR, sMSR ($t = 1$), sMSR ($t = 2$)




(a) The encode time (b) The decode time

(c) The repair time

**Fig. 3.** Comparison of MSR, sMSR, BASIC MSR, and BASIC sMSR codes

that through the shortening, the performance of MSR codes can be improved and the more bits shortened, the less bandwidth cost. The results are evidences that our shorten codes can efficiently decrease the bandwidth cost.

Furthermore, we simulate four schemes, MSR, sMSR, BASIC MSR, BASIC sMSR, and compare the elapsed time of them in each process to illustrate intuitively the complexity difference. The time axis in Fig. 3 take the logarithm to make the tiny difference, caused by shortening, between sMSR and MSR easier to observe. Figure 3 manifests the complexity of BASIC sMSR approximately equals that of MSR, while is much less than MSR and sMSR in every process. Therefore, we can deem that our further work can substantially reduce the computation complexity, maintaining the same bandwidth cost.

## 5 Conclusions

In our paper, we propose a new code integrated with shorten codes, BASIC operations, and regenerating codes, which is called BASIC shorten MSR codes (BASIC sMSR). We also provide the general construction of BASIC sMSR based on the product matrix framework. From the comparison results with the traditional regenerating codes, we can observe that our code gives better performance regarding the repair bandwidth, the assisted nodes connected when repairing and execution time while maintaining the storage cost unchanged; thus, we can certify that our code, BASIC shorten MSR codes, with advantages of low repair bandwidth and computation complexity can be more practical in DSS for big data storage.

**Acknowledgments.** This work was supported by the National Natural Sciences Foundation of China under Grant 61701136, Grant 61771158, Grant 61501140, and Grant 61525103, and Project funded by China Postdoctoral Science Foundation 2018M630357, and Shenzhen Basic Research Program under Grant JCYJ2017081114233370, Grant JCYJ20170811154309920 and Grant ZDSYS20170728090330586.

## References

1. Ma Y, Nandagopal T, Puttaswamy KPN, Banerjee S. An ensemble of replication and erasure codes for cloud file systems. In: Proceedings IEEE. INFOCOM; 2013. p. 1276–84.
2. Dimakis AG, Brighten Godfrey P, Wu Y, Wainwright MJ, Ramchandran K. Network coding for distributed storage systems. IEEE Trans Inf Theory. 2010;56(9):4539–51.
3. Rashmi KV, Shah NB, Kumar PV. Optimal exact-regenerating codes for distributed storage at the MSR and MBR points via a product-matrix construction. IEEE Trans Inf Theory. 2011;57(8):5227–39.
4. Oggier F, Datta A. Self-repairing homomorphic codes for distributed storage systems. In: Proceedings IEEE. INFOCOM; 2011. p. 1215–23.
5. Oggier F, Datta A. Self-repairing codes for distributed storage a projective geometric construction. In: Information theory workshop; 2011. p. 30–4.

6. Goldwasser JL. Shortened and punctured codes and the macwilliams identities. Linear Algebra Appl. 1997;253(253):1–13.
7. Hou H, Shum KW, Chen M, Li H. Basic codes: low-complexity regenerating codes for distributed storage systems. IEEE Trans Inf Theory. 2016;62(6):3053–69.

# An Adaptive Thresholding Method for Background Subtraction Based on Model Variation

ShaoHu Peng[1], MingJie Deng[1], YuanXin Zhu[1], ChangHong Liu[1], Zhao Yang[1], Xiao Hu[1], Yuan Wu[2], and HyunDo Nam[3(✉)]

[1] School of Mechanical and Electrical Engineering, Guangzhou Higher Education Mega Center, GuangZhou University, GuangZhou 510006, Guangdong Province, China

[2] B3-4A1 Merchants Guangming Science Park, No. 3009 Guanguang Road, Guangming District, Shenzhen City, Guangdong Province 518107, People's Republic of China

[3] Department of Electronics and Electronical Engineering, Dankook University, Yongin-si, Gyeonggi-do 448-701, South Korea

hdnam@dankook.ac.kr

**Abstract.** Background subtraction is an important task in computer vision. Pixel-based methods have a high processing speed and low complexity. But when the video frame with camouflage problem is processed, this kind of methods usually output incomplete foreground. In addition, the parameters of many algorithms are invariable. These methods cannot tackle non-static background. In this paper, we present an adaptive background subtraction algorithm derived from ViBe. Gaussian Kernel template is used to model initialization and update. Standard deviation is used to measure background dynamics. We test our algorithm on a public dataset, named changedetection.net. The results show that we can handle most of scenarios. Compared to ViBe, we achieve better result generally, especially in dynamic background and camera jitter categories.

**Keywords:** Background subtraction · GMM · Adaptive thresholding · Visual surveillance · Computer vision

## 1 Introduction

Background subtraction is an essential step in many computer vision tasks, such as object tracking, motion recognition, and pedestrian detection, etc. This technique is widely used in visual surveillance and motion sensing game. In practical application scenarios, there are many challenges that can make background subtraction algorithms generate false detection, like shadows, light change, and dynamic background, etc.

Background subtraction algorithms can be mainly categorized into pixel-based method and region-based method. According to [1], pixel-based methods assume that the time series of observations are independent at each pixel (e.g., [2–4]). Such algorithms have low computational complexity and run quickly, but they ignore the relationship between the incoming pixel and its neighboring pixels, resulting in poor


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_46

background segmentation in some scenes. One of the classic pixel-based approaches is the Gaussian Mixture Model (GMM) [2]. In this approach, the background model of each pixel is represented by a mixture of Gaussians and updated through an online approximation. Pixel values which do not conform to the background distributions will be considered as foreground pixels. The algorithm can handle with slow lighting changes and recovers quickly if background reappears. Region-based approaches take inter-pixel relationship into consideration, which means more pixels are considered in pixel classification (e.g., [5–7]). The foreground mask can be more complete. But these approaches have higher computational complexity. The approach proposed in [7] is one of the region-based approaches. The authors of [7] subdivide the video frames into regions. By choosing the continuation of the background, a tessellation is grown. Only few parameters are needed in this algorithm. The algorithm is able to cope with heavy clutter.

The pixel-based algorithms and region-based algorithms have shown good performance in background subtraction. However, these kinds of approaches have a poor performance under the non-static background. The motivation of this work is to develop a robust background subtraction algorithm that can work well in the scenarios of non-static background. To do this, we present a new background subtraction algorithm based on pixel region. The proposed method utilizes the standard deviation of the model samples to measure the dynamics of the background. Due to the fact that the pixel values of the non-static background have higher variations than those of the static background, a dynamic distance threshold is defined according to the standard deviation. It increases or decreases adaptively based on the standard deviation, which enables the algorithm adapt to dynamic change of scenes. Furthermore, Gaussian Kernel template is used to assign Gaussian coefficients for the incoming pixel and its neighboring pixels, which considers the inter-pixel spatial relationship. Since these pixels are used to initialize the model according to their probability, a reliable background model is built in the first frame. Due to the utilization of the dynamic thresholding strategy, the proposed method is able to achieve good performance in background subtraction.

The changedetection.net dataset [8] is used to evaluate the proposed algorithm. Experimental results show that overall results of the proposed method are better than ViBe [9], GMM [10] and local self-similarity [11]. Due to the use of pixel region in pixel classification, we obtain more complete foreground masks in thermal category that contains the camouflage problem. Since the proposed adaptive model has a better control in the scenes containing camera jitter and dynamic background, the proposed approach can subtract the background correctly under disturbance of dynamic scenes and has a good performance in baseline and shadow categories.

The paper is organized as follows. In Sect. 2, we present the related background subtraction algorithms. The proposed method is described in detail in Sect. 3. Experimental results are presented in Sect. 4. Section 5 concludes this paper.

## 2 Related Works

ViBe [9] is an efficient non-parametric algorithm. The background models are initialized from a single frame in this method. Model samples are selected randomly from eight-connected neighbors of the incoming pixel in the first frame. To determine if a pixel belongs to the background, ViBe [9] compares it with the model samples. If there are at least two model samples whose Manhattan distance is less than a given threshold, the pixel is classified as background. The algorithm utilizes the conservative update scheme, which means foreground pixels are never included into the background model. In the model update process, a background pixel is used to replace a model sample selected randomly, which is different from many other algorithms which replace the oldest model sample. ViBe [9] also uses the background pixel to randomly update the model from one of the pixels in the neighborhood. This strategy can help remove "ghost" problems. In addition, time subsampling can extend lifespan of the background samples and reduce the computational complexity to obtain higher frames per second.

Zivkovic [10] proposed an improvement of GMM algorithm. In this algorithm, a weight is used to indicate how much of a data sample belongs to a Gaussian component. If there is a sample whose Mahalanobis distance from a component is less than a specific threshold (the sample is close to the component), the component is then updated. If there are no close components, a new component is created. The oldest one is abandoned if there are enough components. Moreover, the number of components per pixel can adaptively change according to scenes.

In [11], the method distinguishes the foreground and background by comparing the local self-similarities between the model and the sequence frames. Instead of using the histogram or color intensity, it uses the local self-similarity descriptor to detect the foreground. Furthermore, the color intensity Euclidean distance and morphological operation are used for post-processing.

## 3 Description of the Algorithm

In this section, we explain the principle of the proposed algorithm. Compared to ViBe [9], we add Gaussian Kernel template to initialize and update the model. Besides, a $3 \times 3$ pixels region is applied to determine if a pixel belongs to background. The distance threshold can dynamically change based on an estimation of the background dynamics.

### 3.1 Background Model Initialization and Update

The authors of ViBe [9] fill the model of an incoming pixel with the values randomly taken in its eight-connected neighborhood in the first frame. Spatial relationship is applied in model initialization of ViBe [9], but it is not enough. If some neighboring pixels are closer to the incoming pixel than the others, higher pick-up probability of initialization should be occupied by these pixels. Furthermore, the incoming pixel also should be used for initialization. This is due to the fact that by considering the spatial relationship, the incoming pixel is more related to the model than its neighbors. Instead

of randomly taking the eight-connected neighborhood to fill the model, the proposed method fills the model samples from the 3-by-3 pixels by assigning a pick-up probability to each pixel. As shown in Fig. 1, for a 3-by-3 pixel region, the pick-up probabilities are determined based on the Gaussian Kernel template. Therefore, the use of Gaussian Kernel allows the proposed method to take more spatial relationships between the incoming pixel and its surrounding into consideration. Every model sample is an RGB vector. After $N$ iterations, there are $N$ samples in the model.

| | | |
|---|---|---|
| 0.0625 | 0.125 | 0.0625 |
| 0.125 | 0.25 | 0.125 |
| 0.0625 | 0.125 | 0.0625 |

| | | |
|---|---|---|
| 0.0833 | 0.1667 | 0.0833 |
| 0.1667 | 1.00 | 0.1667 |
| 0.0833 | 0.1667 | 0.0833 |

**Fig. 1.** Probability of initialization procedure (left), probability of update procedure (right)

In the update procedure, if a pixel is classified as a background pixel, the method not only uses it to update its corresponding background model, but also updates one of its eight-neighboring pixel models, as ViBe [9] did. However, different from ViBe, we follow the specific possibilities to choose the adjacent pixel, instead of choosing randomly. In this process, we also use the Gaussian Kernel template. The model is not updated if the pixel belongs to foreground. Specific possibilities are assigned as shown in Fig. 1.

The model is not updated every time. In the first 150 frames, it is updated every $T_{\text{initial}}$ frame. After that, we update the model every $T_{\text{update}}$ frames.

### 3.2 Adaptive Euclidean Distance Threshold

In a video, there are some static areas (i.e., house) and areas with high dynamics (i.e., fountain). In this case, some (high) dynamic backgrounds with great changes cannot be detected and will be classified as foreground incorrectly if an invariable distance threshold is used. To solve this problem, this paper proposes to change the distance threshold adaptively (increases for dynamic background areas, and decrease for static areas). To do this, the standard deviation $\sigma$ is employed to be a measurement of background dynamics. Hence, an array to store the modulus of model samples (RGB vector) is created and standard deviation $\sigma$ of this modulus is calculated.

For the case that there are small differences among samples of pixels in static region, the standard deviation is small. When the background models of pixels in

high-dynamic areas are updated, samples can be replaced with the values which have a great difference from the other samples, resulting in a high standard deviation.

After calculating the standard deviation, the distance threshold can change adaptively as follows:

$$R(x_i) = \begin{cases} R(x_i)\left(1 + F_{\text{updown}}\right), \text{ if } \sigma F_{\text{control}} > R(x_i) \\ R(x_i)\left(1 - F_{\text{updown}}\right), \text{ else} \end{cases} \quad (1)$$

where $F_{\text{updown}}$ and $F_{\text{control}}$ are fixed parameters, $x_i$ stands for pixels in different location. In high dynamic areas, the product of $\sigma F_{\text{control}}$ is usually greater than $R(x_i)$, which means $R(x_i)$ should increase. In this situation, the distances between the pixel in dynamic background and its background samples can be smaller than $R(x_i)$, and the pixel is classified as background correctly. Pixel regions with assigned grayscale value of 85 in ground truth picture of Fig. 2 are Non-Roi.

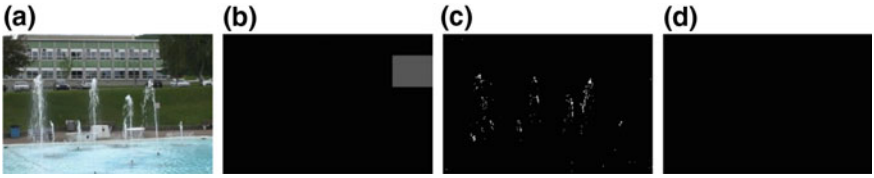



**Fig. 2.** Frame 626 of sequence fountain 01 (dynamic background category): input image (**a**). Ground truth (**b**). Foreground mask of ViBe [9] (**c**). Foreground mask of proposed method after using adaptive threshold (**d**)

### 3.3 Pixel Classification

We use a pixel region to judge if an incoming pixel belongs to the background. If there are $Q_{\min}$ model samples in the background model whose Euclidean distances from the pixel are smaller than a specific threshold, the pixel is determined to be similar with its background model. All pixels in the pixel region (including the incoming pixel and its eight-connected neighboring pixels) are then traversed. An incoming pixel is classified as background if there are at least $K$ pixels which are similar with their corresponding models. Otherwise, it is considered as foreground.

The classification of a pixel involves the computation of $9 \times N$ Euclidean distances between all nine pixel and their own model samples, and of $9 \times N$ comparison with their related thresholds $R(x_i)$. Note that, as we only need to find a few matches, the judgement of similarity process of a pixel will stop when $Q_{\min}$ matched samples have been found and the classification process of a pixel breaks when there are $K$ pixels similar with their own models.

## 4 Experimental Results

We compare our method with ViBe [9], GMM [10] and local self-similarity [11] using the public dataset offered on the http://www.changedetection.net Web site. The dataset includes 31 real-world video sequences, grouped in six categories: baseline, camera jitter, dynamic background, intermittent object motion, shadows, and thermal, which stands for different kinds of challenges. We use median blur as a post-processing operation.

These methods are compared based on the following metrics: True Negatives (TN), True Positives (TP), False Positives (FP), False Negatives (FN), Recall (Re), Specificity (Sp), False Positive Rate (FPR), False Negative Rate (FNR), Percentage of Wrong Classifications (PWC), Precision (Pr), *F*-Measure ($F1$) more details of these metrics can be referred to [8].

### 4.1 Parameter Settings

The parameters and their values are given hereafter:

- $N = 35$: the number of background model samples.
- $mR_{\text{initial}} = 16.5$: initial threshold.
- $mR_{\text{low}} = 16.5$: lower bound of the distance threshold.
- $mR_{\text{high}} = 50$: upper bound of the distance threshold.
- $F_{\text{control}} = 5.5$: the factor in the controller for the distance threshold which controls the value $\sigma F_{\text{control}}$.
- $F_{\text{updown}} = 0.05$: the rate which makes the distance threshold increase or decrease.
- $Q_{\text{min}} = 3$: required number of matching samples to determine that the pixel is similar with its background model.
- $K = 6$: the number of pixels which are similar with their own model in the region to label an incoming pixel as background.
- $T_{\text{initial}} = 1$: update factor in the first 150 frames.
- $T_{\text{update}} = 2$: update factor after the first 150 frames.

### 4.2 Result and Discussion

In Table 1, we present the complete results of the proposed method and the overall results of ViBe [9] (labeled as "*V*" in Table 1), GMM [10] (labeled as "*G*") and local self-similarity [11] (labeled as "*L*"). The data for the algorithm comparison is obtained from the changedetection.net dataset [8]. The best overall result is mentioned in bold. We can find that there are five metrics obtained from the proposed method are better than those of other methods, including *F*-measure which is closed correlated with the method rankings and the most important metric used for comparison. Compared to other approaches, our approach outputs a more complete foreground mask, which means the True Positive is higher. The method can also handle the area which can easily cause False Positive, such as dynamic background. So we get more True Negative. Therefore, we get higher score in five metrics.

**Table 1.** Complete results of the proposed method and the overall results of other methods

| Category | Re | Sp | FPR | FNR | PBC | Pr | *F*1 |
|---|---|---|---|---|---|---|---|
| Baseline | 0.8251 | 0.9966 | 0.0034 | 0.1749 | 1.0357 | 0.8826 | 0.8510 |
| CameraJ | 0.7483 | 0.9824 | 0.0176 | 0.2517 | 2.5781 | 0.7216 | 0.7096 |
| Dynamic | 0.6080 | 0.9982 | 0.0018 | 0.3920 | 0.5388 | 0.7900 | 0.6388 |
| Intermit | 0.4178 | 0.9591 | 0.0409 | 0.5822 | 7.7019 | 0.6979 | 0.4302 |
| Shadow | 0.7797 | 0.9901 | 0.0099 | 0.2203 | 1.8353 | 0.8279 | 0.7972 |
| Thermal | 0.5565 | 0.9968 | 0.0032 | 0.4435 | 2.6845 | 0.9513 | 0.6678 |
| Overall | 0.6559 | **0.9872** | **0.0128** | 0.3441 | **2.7291** | **0.8119** | **0.6824** |
| Overall (*V*) | 0.6821 | 0.9830 | 0.0170 | 0.3179 | 3.1178 | 0.7357 | 0.6683 |
| Overall (*G*) | 0.6964 | 0.9845 | 0.0155 | 0.3036 | 3.1504 | 0.7079 | 0.6596 |
| Overall (*L*) | **0.9354** | 0.8512 | 0.1488 | **0.0646** | 14.2954 | 0.4139 | 0.5016 |

We present in Table 2 the *F*-measure and Precision in all six categories obtained by the proposed method and the other methods. The best result is mentioned in bold. Table 2 shows that in camera jitter, dynamic background, and thermal, our method ranks first. Especially in camera jitter and dynamic background categories, the proposed method achieves 18.37 and 13.02% improvement in *F*-measure and 36.43 and 47.77% improvement in Precision compared to ViBe [9], respectively. In these two categories, the background is not static and changes all the time, which can easily cause False Positive. In our algorithm, standard deviation in background model samples is used to measure background dynamics, and the threshold can be adaptively changed. Therefore, our algorithm has a better control of these two categories. Camouflage is a big problem in thermal category. We use a pixel region to determine if a pixel belongs to background. Compared to pixel-based methods, the condition for pixels to be included into the background is stricter in our method, which means we can detect foreground objects similar with the background. Although our *F*-measure and Precision

**Table 2.** *F*-measure and Precision in all categories

| Category | Metrics | Proposed | ViBe [9] | GMM [10] | Local [11] |
|---|---|---|---|---|---|
| Baseline | *F*-measure | 0.8510 | **0.8700** | 0.8382 | 0.8494 |
| | Precision | 0.8826 | **0.9288** | 0.8993 | 0.7564 |
| CameraJit | *F*-measure | **0.7096** | 0.5995 | 0.5670 | 0.2074 |
| | Precision | **0.7216** | 0.5289 | 0.4872 | 0.1202 |
| DynamicBg | *F*-measure | **0.6388** | 0.5652 | 0.6328 | 0.0949 |
| | Precision | **0.7900** | 0.5346 | 0.6213 | 0.0518 |
| Intermittent | *F*-measure | 0.4302 | 0.5074 | 0.5325 | **0.5329** |
| | Precision | **0.6979** | 0.6515 | 0.6458 | 0.4445 |
| Shadow | *F*-measure | 0.7972 | **0.8032** | 0.7319 | 0.5951 |
| | Precision | 0.8279 | **0.8342** | 0.7232 | 0.4673 |
| Thermal | *F*-measure | 0.6778 | 0.6647 | 0.6548 | **0.7297** |
| | Precision | **0.9513** | 0.9363 | 0.8706 | 0.6433 |

in baseline and shadow categories is lower than ViBe [9], they are almost all higher than 0.8, which means the method has a good behavior in these categories. We get the worst result in the intermittent object motion. In this category, the foreground objects usually stop for a while, resulting in False Negative. When the object which has been static for a while starts to move, there is a big difference between the real background left by the object and the model initialized with the foreground pixels, which can cause False Positive.

## 5 Conclusion

In this paper, we present a non-parametric region-level algorithm derived from ViBe [9]. The background model of an incoming pixel is initialized with a $3 \times 3$ pixel region. Gaussian Kernel template is used for distributing probability. Initialization and update take inter-pixel relationship into consideration. We also use a $3 \times 3$ pixel region to classify a pixel, and the foreground mask is more complete. Adaptive distance threshold is used for non-static scenarios, so that our method can handle with dynamic background and camera jitter. The experimental results demonstrate that our method is more robust than ViBe [9]. For future work, we will study how to solve the intermittent object motion problem and detect foreground more precisely in thermal scenes.

**Acknowledgments.** This work was supported by Guangdong provincial scientific and technological project (ID: 2017B020210005), Student's Platform for Innovation and Entrepreneurship Training Program (ID: 201711078010), and the National Natural Science of Foundation of China (No. 61501177).

## References

1. Bouwmans T, Maddalena L, Petrosino A. Scene background initialization: a taxonomy. Pattern Recogn Lett. 2017;96:3–11.
2. Stauffer C, Grimson W. Adaptive background mixture models for real-time tracking. In: IEEE international conference on computer vision and pattern recognition; 1999. p. 246–52.
3. Chen CC, Aggarwal J. An adaptive background model initialization algorithm with objects moving at different depths. In: IEEE international conference on image processing; 2008; p. 2664–7.
4. Bevilacqua A, Stefano LD, Lanza A. An effective multi-stage background generation algorithm. In: IEEE international conference on video and signal based surveillance; 2005. p. 388–93.
5. Farin D, De With PHN, Effelsberg W. Robust background estimation for complex video sequences. In: IEEE international conference on image processing, vol. 1; 2003. p. I-145–8.
6. Lin HH, Liu TL, Chuang JH. Learning a scene background model via classification. IEEE Trans Sig Process. 2009;57(5):1641–54.
7. Colombrai A, Fusiello A. Patch-based background initialization in heavy cluttered video. IEEE Trans Image Process. 2010;19(4):926–33.

8. Goyette N et al. Changedetection.net: a new change detection benchmark dataset. In: IEEE international conference on computer vision and pattern recognition workshops (CVPRW); 2012. p. 1–8.
9. Barnich O, Van Droogenbroeck M. ViBe: a universal background subtraction algorithm for video sequences. IEEE Trans Image Process. 2011;20(6):1709–24.
10. Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction. In: IEEE international conference on pattern recognition; 2004. p. 28–31.
11. Jodoin J-P, Bilodeau G, Saunier N. Background subtraction based on local shape. Technical report; 2012. arXiv:1204.6326v1.

# Performance Analysis of Whale Optimization Algorithm

Xin Zhang[1,2], Dongxue Wang[1,2], and Xiu Zhang[1,2(✉)]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
[2] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China
ecemark@tjnu.edu.cn, 1187713920@qq.com, zhang210@126.com

**Abstract.** Through the research and analysis of a relatively novel natural heuristic, meta-heuristic swarm intelligence optimization algorithm, this swarm intelligence algorithm is defined as a whale optimization algorithm. The algorithm builds a mathematical model by simulating a social behavior of humpback whales. This optimization algorithm was inspired by the bubble-like net hunting phenomenon that humpback whales prey on. By analyzing the four benchmark optimization problems with or without offset and rotation, the convergence performance of the whale optimization algorithm and the ability to solve the optimization problem are proved. The performance of the whale optimization algorithm is based on the computer simulation technology. Through the convergence curve obtained from the experiment, we can see that the whale optimization algorithm performs best for the five benchmark optimization problems without rotation.

**Keywords:** Swarm intelligence · Whale optimization algorithm · Convergence · Numerical optimization

## 1 Introduction

The swarm intelligence optimization algorithm has the characteristics of simple and easy-to-understand concept, easy-to-implement programming, no need of gradient information, and avoidance of local optimal solution. Therefore, swarm intelligence algorithm has been well applied in various fields, such as particle swarm optimization algorithm and genetic algorithm [1,2]; artificial bee colony algorithm, ant colony algorithm, and whale optimization algorithm [3]. There are two main points of the optimization problem. One is to find the global optimal solution. The second is to have a high convergence speed. In the actual programming code, if there is no suitable function, then there will be a local optimal situation. Therefore, it is not possible to reach the global optimum


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_47

result, and it is necessary to find the most suitable optimization method. The method can adjust the search direction in the search space in real time, so as to achieve the global optimal effect [4,5].

The whale optimization algorithm (WOA) was a novel swarm intelligence optimization search method proposed by Mirjalili and Lewis in 2016 [3]. The algorithm is inspired by the simulation of hunting behavior of humpback whales and the global search of groups through whales. And then surround the prey, and attack the prey on the process of hunting to achieve the purpose of optimizing the search. Whale optimization algorithm has the advantages of simple principle, few parameter setting, easy programming, and so on. In function optimization, it is obviously superior to differential evolution algorithm [6], gray wolf optimization algorithm, and gravity in terms of convergence speed performance and accuracy in solving. The search algorithm has been proved that the whale optimization algorithm has been well applied in reservoir optimization scheduling. Therefore, this paper will conduct a deep research on the whale optimization algorithm, study its main performance, and conduct computer simulation tests.

## 2 Whale Optimization Algorithm

The WOA algorithm consists of three stages. They are the surrounding prey, bubble net attack method, and search for prey.

### 2.1 The Surrounding Prey

Humpback whales can accurately identify prey positions and surround them in time because the optimal position of the prey in the search space is unknown. Therefore, the whale optimization algorithm temporarily assumes that the best candidate is the target prey or near-optimal hunting position. After the best search agent is defined, other search agents will try to update their location to find the best search agent. This search behavior is mainly represented by the following equation:

$$\boldsymbol{D} = |\boldsymbol{C} \cdot \boldsymbol{X}^*(t) - \boldsymbol{X}(t)| \tag{1}$$

$$\boldsymbol{X}(t+1) = \boldsymbol{X}^*(t) - \boldsymbol{A} \cdot \boldsymbol{D} \tag{2}$$

where $t$ represents the current iteration, $\boldsymbol{A}$ and $\boldsymbol{C}$ are coefficient vectors, $\boldsymbol{X}^*$ is the position vector of the best solution obtained so far, $\boldsymbol{X}$ is the position vector, and $|\cdot|$ is an absolute value. It should be noted here that if there is a better solution, $\boldsymbol{X}^*$ should be updated in each iteration.

### 2.2 Bubble Net Attack Method

In order to mathematically model the bubble net behavior of humpback whales, the following two methods have been designed: (1) shrinkage bracketing mechanism and (2) spiral update position. Shrinkage bracketing mechanism is achieved

by reducing the value of $\boldsymbol{A}$ in Eq. (1). Spiral update position depicts a spiral texture resembling a snail shell in a two-dimensional plane. First, we should calculate the distance between the whales and the prey and then establish a spiral equation between the whales and the target prey to simulate the humpback whales spiral.

### 2.3 Search for Prey

The same method can also be used to search for prey based on changes in vector $\boldsymbol{A}$. However, in reality, humpback whales can conduct random searches based on the location of other whales. Therefore, we can use a $\boldsymbol{A}$ with a random value range greater than 1 or less than $-1$ to force the search agent and stay away from the referenced whales. The exploration phase is the opposite of the development phase. We can update the search agent's location in the exploration phase in real time based on a randomly selected search agent, instead of the best search agent location currently found.

## 3 Measurement of Convergence

There are many ways to solve complex problems in various optimization algorithms [7,8]. For various types of optimization problems [9,10], the algorithms used are also different, and the optimized performance shown varies. Some of these optimization algorithms can find local optimal solutions in a short time [11]. Thus, it is very necessary to study the properties of the WOA algorithm. Moreover, the analysis of WOA might discover the advantages and disadvantages of the algorithm for different problems.

**Table 1.** High-dimensional reference function names

| Function | Name |
|---|---|
| $f_1(x)$ | Sphere function |
| $f_2(x)$ | Schwefel's problem 2.22 |
| $f_3(x)$ | Schwefel's problem 1.2 |
| $f_4(x)$ | Schwefel's problem 2.21 |
| $f_5(x)$ | Rosenbrock function |

In order to verify the convergence performance of the WOA algorithm and solve the performance of the optimization problem, the experiment uses five benchmark functions and conducts an experimental test with four combinations of shifting and rotating, which is $f_1(x)$–$f_5(x)$. The high-dimensional reference function is shown in Table 1.

Functions of Table 1 are classical ones without any displacement and flipping. Problem models of real-world applications are usually more complex than

these classical functions. To obtain a more meaningful study, shifting and rotating operations are attached with these functions, i.e., function without shifting without rotating, function with shifting without rotating, function without shifting with rotating, and function with shifting with rotating. The convergence of WOA is analyzed based on the four cases in the next section.

## 4 Simulation Experiment

As shown in Fig. 1, for $f_1(x)$ without shifting and without rotation, the optimal function value obtained by WOA is less than $10^{-200}$, which is much smaller than the commonly used threshold $10^{-8}$, so it can be considered that WOA algorithm finds the optimal solution.

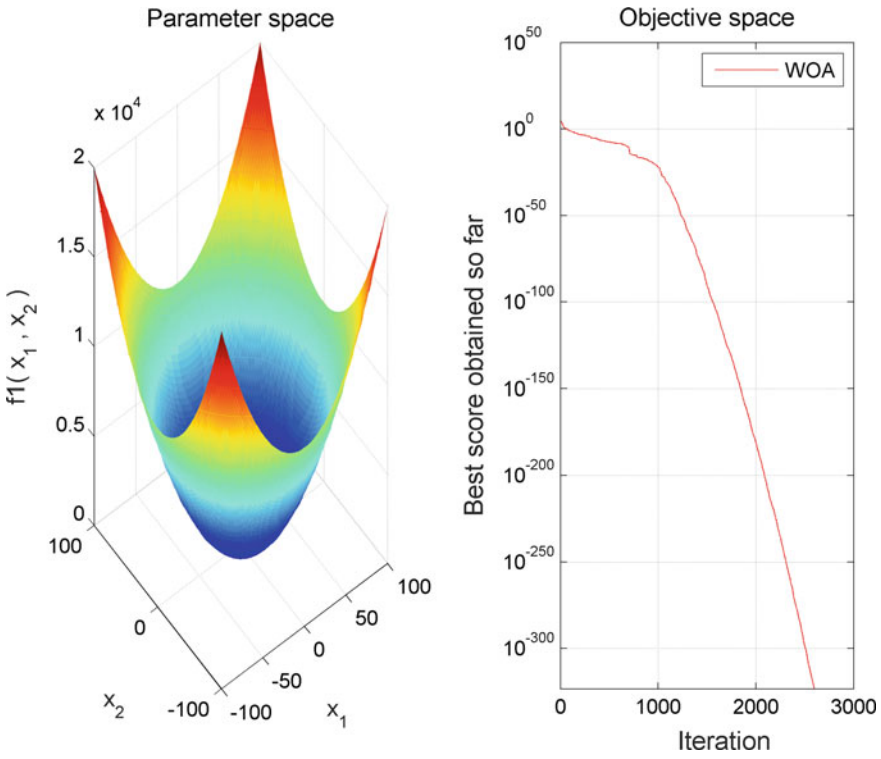



**Fig. 1.** WOA algorithm solves $f_1(x)$ without shifting without rotating.

As shown in Fig. 2, for the function $f_1(x)$ with shifting without rotation, the optimal function value obtained by the WOA is less than $10^{-200}$, which is much smaller than the commonly used one. The threshold value is $10^{-8}$, so the WOA can be considered to find the optimal solution.

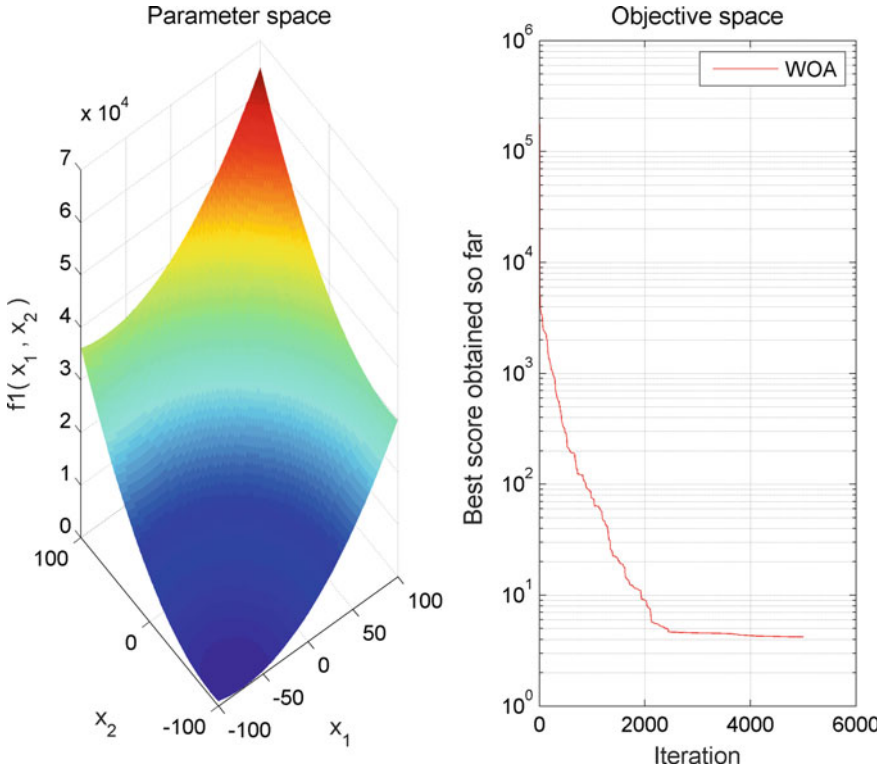



**Fig. 2.** WOA algorithm solves $f_1(x)$ with shifting without rotating.

As shown in Fig. 3, for the function $f_1(x)$ without shifting with rotation, the optimal function value is greater than 3.855, which is much larger than the commonly used one. The threshold value is $10^{-8}$, so it can be considered that the whale optimization algorithm does not find the optimal solution.

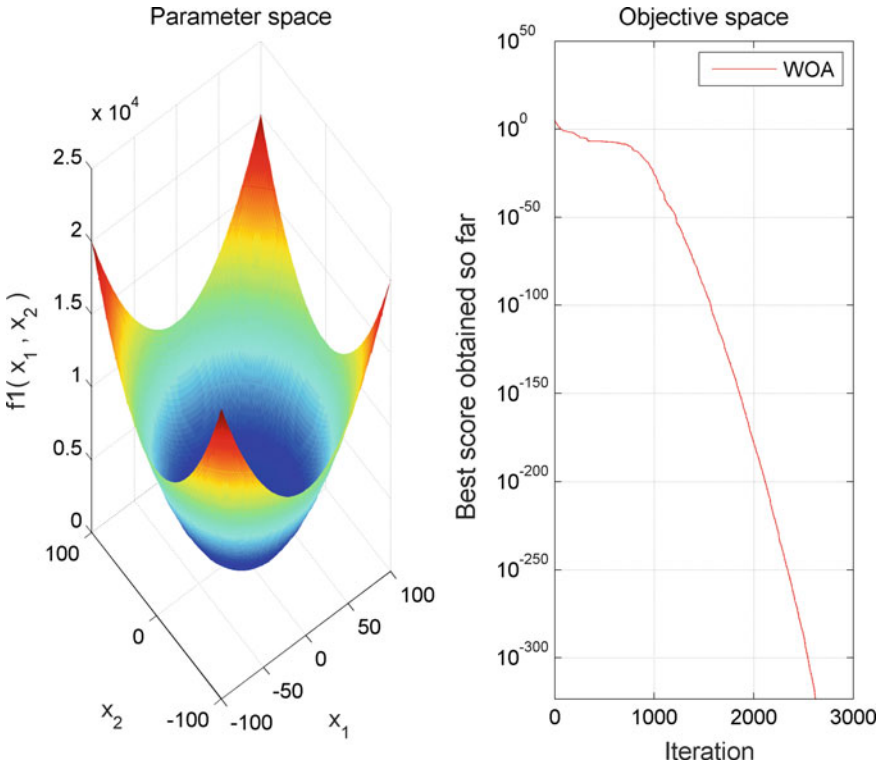



**Fig. 3.** WOA algorithm solves $f_1(x)$ without shifting with rotating.

As shown in Fig. 4, for the function $f_1(x)$ with shifting with rotation, the optimal function value obtained by the whale optimization algorithm is greater than 3.417, which is much larger than the commonly used one. The threshold is $10^{-8}$, so the WOA algorithm can be considered as not finding the optimal solution.

For the saving of space, the convergence curve for $f_2(x)$, $f_3(x)$, $f_4(x)$, and $f_5(x)$ are not presented. The results are summarized in Table 2. In this table, "Yes" denotes if WOA finds global optimum of the related function; "No" denotes WOA cannot find global optimum. It can be seen from Table 2 that the convergence performance of WOA is the best for solving functions with shifting without rotation.

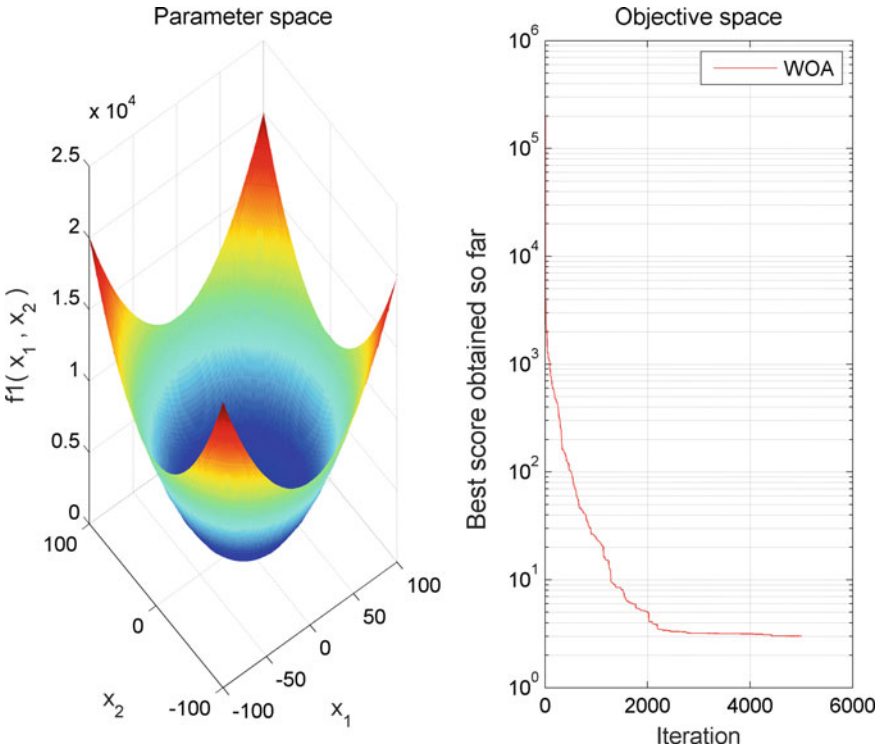



**Fig. 4.** WOA algorithm solves $f_1(x)$ with shifting with rotating.

**Table 2.** Summary of the results of the WOA algorithm for $f_1(x)$ $f_5(x)$ without shifting without rotating, with shifting without rotating, without shifting with rotating, and with shifting with rotating. Yes for finding global optimum; otherwise No

| Operation | $f_1(x)$ | $f_2(x)$ | $f_3(x)$ | $f_4(x)$ | $f_5(x)$ |
|---|---|---|---|---|---|
| Without shifting without rotating | Yes | Yes | No | No | No |
| With shifting without rotating | Yes | Yes | Yes | Yes | No |
| Without shifting with rotating | No | No | No | No | No |
| With shifting with rotating | No | No | No | No | No |

## 5 Conclusion

In summary, we can clearly see the convergence curve obtained by the WOA for five optimization problems. From the comparison of these convergence curves and the summary table, it can be seen that the convergence performance has no rotation in the case of offset. The effect is best. The effect is the worst when there exist both shifting and rotating.

The WOA has the shortcomings of local optimal solution and slow convergence speed. With the emergence of various swarm intelligence algorithms, other optimization algorithms can be introduced to improve the practical problems and then fill in the deficiencies of the whale optimization algorithm.

**Acknowledgments.** This research was supported in part by the National Natural Science Foundation of China (Project No. 61601329, 61603275), the Tianjin Higher Education Creative Team Funds Program, the Applied Basic Research Program of Tianjin (Project No. 15JCYBJC51500, 15JCYBJC52300), and the Doctoral Fund Project of Tianjin Normal University (Project No. 043-135202XB1602).

## References

1. Gotmare A, Bhattacharjee SS, Patidar R. Swarm and evolutionary computing algorithms for system identification and filter design: A comprehensive review. Swarm Evol Comput. 2017;32:68–84.
2. Yuen SY, Chow CK. A genetic algorithm that adaptively mutates and never revisits. IEEE Trans Evol Comput. 2009;13:454–72.
3. Mirjalili S, Lewis A. The whale optimization algorithm. Adv Eng Softw. 2016;95:51–67.
4. Zhang, X, Zhang, X, Fu, WN. Fast numerical method for computing resonant characteristics of electromagnetic devices based on finite-element method. IEEE Trans Magn. 2017;53, Article No. 7401004.
5. Wu Z, Xia X. Optimal switching renewable energy system for demand side management. Solar Energy. 2015;114:278–88.
6. Zhang X, Zhang X. Improving differential evolution by differential vector archive and hybrid repair method for global optimization. Soft Comput. 2017;21:7107–16.
7. Oliva D, El Aziz MA, Hassanien AE. Parameter estimation of photovoltaic cells using an improved chaotic whale optimization algorithm. Appl Energy. 2017;200:141–54.
8. Zhang S, Zhou Y, Li Z, Pan W. Grey wolf optimizer for unmanned combat aerial vehicle path planning. Adv Eng Softw. 2016;99:121–36.
9. Cheng MY, Prayogo D. Fuzzy adaptive teaching clearing-based optimization for global numerical optimization. Neural Comput Appl. 2018;29:309–27.
10. Gaudioso M, Giallombardo G, Mukhametzhanov M. Numerical infinitesimals in a variable metric method for convex nonsmooth optimization. Appl Math Comput. 2018;318:312–20.
11. Li G, Cui L, Fu X, Wen Z, Lu N, Lu J. Artificial bee colony algorithm with gene recombination for numerical function optimization. Appl Soft Comput. 2017;52:146–59.

# Face Template Protection Algorithm Based on DNA Encoding Encryption

Futong He and Jiaqi Zhen(✉)

College of Electronic Engineering, Heilongjiang University,
Harbin 150080, China
zhenjiaqi2011@163.com

**Abstract.** With the rapid development of information technology, biometric technology has been widely used. However, biological information is limited to everyone, once it is leaked, it can no longer be safely used, which will cause lifelong damage to users. Therefore, the study of biological template protection technology is of great significance. This paper introduces the deoxyribonucleic acid (DNA) coding theory into biometric template protection and proposes a face template protection algorithm based on DNA coding encryption. Experiments and analyses are carried out on Olivetti Research Laboratory (ORL) face database. The results show that the proposed algorithm not only enhances the security of the original template, but also ensures the recognition performance of the system.

**Keywords:** Biometrics · Biometric template DNA coding · Face template protection

## 1 Introduction

With the rapid development of social information, people have been using keys, certificates, user names, passwords and bank cards for identification in the past. Since certificates, keys, and bank cards are easily stolen and lost, user names and passwords are easily forgotten, so biometric identification technology emerges, which utilizes human biometric features for identification.

For biometric systems, biometric templates are used to verify and match the biometric features which are stored in the database and represent most of the biometric information. Biometrics are unique [1]. Once lost, they are permanently lost, unlike traditional passwords and passwords that can be updated or reset. Biological template leaks can seriously threaten the security and privacy of individuals. Therefore, template protection is an important research direction of biometric system security, and it is also a research hotspot of biometric recognition.

In this paper, a face template protection algorithm based on DNA encoding encryption is proposed [2]. The other parts of the paper are as follows: the second part introduces the related contents, the third part is the algorithm implementation, the fourth part is the experimental results and analysis. The last part is the conclusion.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_48

## 2 Related Works

### 2.1 Biological Template Protection Method

Generally speaking, a good biometric template protection algorithm [3] should have the following four characteristics:

(1) Diversity: using this scheme, different security templates can be generated for different databases, even if one database template is lost, it can not be used to attack other databases, thus avoiding the impact of cross-matching.
(2) Revocability: once the template is compromised, the system can easily undo the original template and regenerate a new template so that it does not expose the original biological information.
(3) Security: even if a biological template is lost or stolen, it is difficult for an attacker to recover the original biometric information by using the lost template.
(4) Authentication performance: a good biometric template does not degrade the performance of original biometric recognition in the encrypted domain.

The current biometric template protection algorithms can be divided into feature transformation method and biometric encryption method.

The feature transformation method is to transform the biological template $T$ into $F(T, K)$ and store $F(T, K)$ in the database through the transformation function $F$. Usually, the key or random number is used as the parameter of the transformation function. The transformed feature template $F(Q, K)$ is obtained by applying the transformation function $F$ to the tested biometric $Q$ and compared with $F(T, K)$. According to the characteristics of the transformation function, the characteristic transformation method can be divided into salinization method and irreversible transformation method.

Biometric encryption is also known as the auxiliary data method. It is an encryption key extracted from registered users. The comparison is done by verifying the validity of the extracted encryption key. Encryption method can be divided into key binding method and key generation method.

### 2.2 DNA Coding Methods and Preprocessing

In the double spiral structure of DNA, the DNA sequence contains four bases: A, C, G, T. In these four bases, A and T are complementary to each other, C and G are paired with each other. The four bases are coded with the 0, 1 in the computer, and each base corresponds to two binary numbers [4]. Therefore, the bases in the DNA sequence can be coded with 11 and encoded after the pairing rules. Be paired with 00 and 11, and 01 and 10. There are eight DNA coding methods that meet the complementary pairing criteria, as shown in Table 1.

DNA complementary replacement: Assume $x_i$ is the $i$-DNA encoded user template. $x_i^j$ is the $j$ dimension of $x_i$. Thus, $x_i$ is a sequence consisting of four bases, and then, in accordance with the principle of complementary pairs of bases, the nucleotides in the nucleotide $x_i$ encoded by DNA are specified as follows:

**Table 1.** Rules of DNA encoding

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| A | 00 | 00 | 01 | 01 | 10 | 10 | 11 | 11 |
| T | 11 | 11 | 10 | 10 | 01 | 01 | 00 | 00 |
| C | 01 | 10 | 00 | 11 | 00 | 11 | 01 | 10 |
| G | 10 | 01 | 01 | 00 | 11 | 00 | 10 | 01 |

$$\begin{cases} x_i^j \neq D(x_i^j) \neq D\big(D(x_i^j)\big) \neq D\big(D\big(D(x_i^j)\big)\big) \\ x_i^j = D\big(D\big(D\big(D(x_i^j)\big)\big)\big) \end{cases} \tag{2.1}$$

where $D(x_i^j)$ and $x_i^j$ are a pair of base pairs, they are complementary, and the base pair satisfies the emittance condition. From Formula (2.1), six reasonable base pair complementary combinations can be obtained, as shown in Table 2.

**Table 2.** Type of complementary substitutions of base pairs

| Type | Type representation |
|---|---|
| 1 | (AT)(TC)(CG)(GA) |
| 2 | (AT)(TG)(GC)(CA) |
| 3 | (AG)(GT)(TC)(CA) |
| 4 | (AG)(GT)(CT)(TA) |
| 5 | (AC)(CT)(TG)(GA) |
| 6 | (AC)(CG)(GT)(TA) |

## 3 Proposed Algorithm

### 3.1 Logistic Map

Logistic mapping [5] is defined as follows:

$$x_{n+1} = \mu x_n(1 - x_n) \tag{3.1}$$

while $\mu$ belongs to (0, 4]. From the bifurcation diagram, we can see that the sequences generated by logistic map are chaotic sequences when $3.5699 < \mu \leq 4$ (Fig. 1).

### 3.2 Algorithm Steps

(1) Let $U$ be the face feature matrix extracted by PCA [6]. For the matrix $U$ of $m \times n$, we use the Formula (3.2) to normalize it to the value $U_1$ of 0 to 1, and then perform the operation of Formula (3.3) to get the integer $U_2$ between 0 and 255, and then convert $U_2$ into a binary matrix $U_3$ of $(m \times n) \times 8$ size [7]. Where map min max()

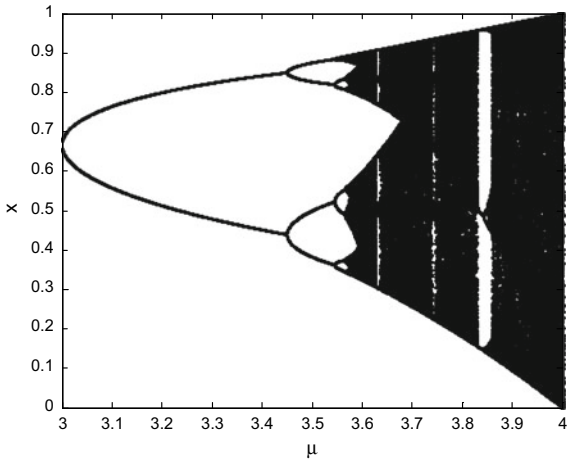



**Fig. 1.** Bifurcation diagram of the logistic map

is the normalized function in MATLAB and round( ) is the whole function in MATLAB.

$$[U_1,\ PS] = \text{mapminmax}(U,\ 0,\ 1) \tag{3.2}$$

$$U_2 = \text{round}(U_1 \times 255) \tag{3.3}$$

(2) Random generation of an integer $r_1$ between 1 and 8 corresponding to the $r_1$ encoding rule in Table 1 is generated at random. According to this kind of DNA coding rule, $U_2$ is converted into a DNA coding sequence $U_3$ of $(m \times n) \times 4$ size.
(3) Given the initial value of logistic chaotic [8] sequence $x_0$ and parameter $\mu$ and using the logistic sequence generated by Formula (3.1) to generate the iterative times of DNA complementary substitution. Let $C = \{c_1,\ c_2,\ \ldots,\ c_m\}$ represent the final iteration number and the $i$ dimension $c_i$ of $C$ is calculated by Formula (3.4).

$$c_i = \text{fix}(\text{mod}(l_i \times 10,\ 4)) \tag{3.4}$$

Among them, fix($x$) is a function that gets the nearest integer to the element $x$ in the null direction, and mod() is the operation of finding modules. Then, $r_2$ is generated randomly to decide which complementary base pairs are used in Table 2. Finally, according to $c_i$, DNA sequence $U_3$ corresponds to the complementary substitution sequence $u_4^i$ in $U_4$ can be obtained by $u_3^i$, the formula is as follows:

$$\begin{cases} u_4^i = u_3^i, & c_i = 0; \\ u_4^i = D(u_3^i), & c_i = 1; \\ u_4^i = D(D(u_3^i)), & c_i = 2; \\ u_4^i = D(D(D(u_3^i))), & c_i = 3. \end{cases} \tag{3.5}$$

(4) Randomly generate a 1–8 integer $r_3$, using the $r_3$ DNA encoding rule in Table 1, converting DNA sequence $U_4$ into M's one-dimensional binary sequence, and then converting the binary sequence into $m$ row $n$ column decimal matrix $U_5$, $U_5$ for the feature template matrix after DNA encoding.

Decryption steps:

(1) Using the known random number $r_3$, using the $r_3$ DNA encoding rule in Table 1, the encrypted biometric template matrix $U_5$ is converted into a one-dimensional DNA-coded sequence $U_4'$ with a $(m \times n) \times 4$ size.
(2) Transform $U_4'$ into $U_3'$ according to the complementary DNA substitution method [9] and the $r_2$ iterative sequence $C$ in Table 2.
(3) According to the $r_1$ DNA encoding in Table 1, the $U_3'$ is converted to a decimal matrix $U_2'$ with a size of $m \times n$, and $U_2'$ is divided by 255 to get the inverse normalization operation of the $U_1'$, $U_1'$ re-execution (3.6) gets $U'$.

$$U' = \text{mapminmax}(\text{'}reverse\text{'}, U_1') \tag{3.6}$$

## 4 Experimental Results and Analysis

In order to test the proposed algorithm, we conducted experiments on the ORL face database and compared with the experimental results of the feature faces. It contains a lot of comparative results. In this experiment, five samples of each person were used as training sets and others for testing. The original template uses PCA algorithm to extract facial feature vectors [10].

In this algorithm, we use the false match rate (FMR) and false no match rate (FNMR) to evaluate the performance of the algorithm. FMR and FNMR are the two parameters of the performance evaluation of the main recognition algorithm. In addition, the equal error rate (EER) can be used to evaluate the overall performance of the algorithm. It unifies FMR and FNMR, FMR increases with the increase of threshold, FNMR decreases with the increase of threshold, and the intersection of both is the value of EER. In the detection error tradeoff (DET) curve, the $X$-axis and $Y$-axis represent FMR and FNMR, respectively, and the lower DET curve implies higher performance.

Figure 2 compares the DET curves between the original feature face and the face template protection algorithm based on logistic mapping encryption. Table 3 lists the EER of the original face PCA and the face template protection algorithm based on logistic map encryption, which is 3.9%. From Fig. 2, we can see that the two curves

coincide together, that is to say, using the DNA encoding encryption algorithm does not reduce the recognition performance of the original system. The algorithm satisfies the requirements of authentication performance, diversity, revocation, and security.

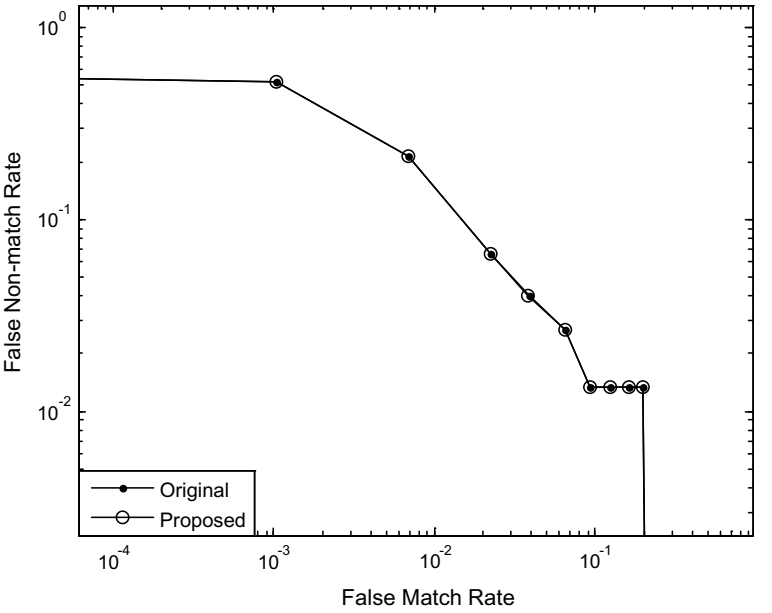



**Fig. 2.** Comparison of DET curve

**Table 3.** Comparison table of EER

| Algorithm | EER(%) |
|---|---|
| Original | 3.9 |
| Proposed | 3.9 |

## 5 Conclusion

This paper proposes a face template protection algorithm based on DNA encoding encryption, first converts human face features to integers 0–255, then converts them to binary, and then converts the binary sequences into DNA sequences using one of the DNA encoding methods. Using the one-dimensional logistic chaotic system to generate iterations of DNA complementary substitution, the DNA sequence is iterated to generate a new DNA sequence. Experimental results show that the algorithm not only enhances the security of the original template, but also guarantees the recognition performance of the system.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China under Grant No. 61501176, Heilongjiang Province Natural Science Foundation (F2018025), University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (UNPYSCT-2016017).

## References

1. Kevenaar TAM, Lemma AN. System and method for verifying the identity of an individual by employing biometric data features associated with the individual. US, US9160522; 2015.
2. Zhang XC, Han F, Niu Y. Chaotic image encryption algorithm based on bit permutation and dynamic DNA encoding. Comput Intell Neurosci. 2017:1–11.
3. Shao X, Xu H, Veldhuis RNJ et al. A concatenated coding scheme for biometric template protection. IEEE international conference on acoustics, speech and signal processing; 2012. p. 1865–8.
4. Zhang Q , Wei X. A novel couple images encryption algorithm based on DNA subsequence operation and chaotic system [J]. Optik Int J Light and Electron Opt; 2013;124(23): 6276–6281.
5. Clark F, Brook BW, Delean S, et al. The theta-logistic is unreliable for modelling most census data. Methods Ecol Evol. 2010;1:253–62.
6. Bari AT, Reaz MR, Islam AK, et al. Effective encoding for DNA sequence visualization based on nucleotide's ring structure. Evol Bioinform. 2013;9:251–61.
7. Zhu Q. Multi-directional two-dimensional PCA with matching score level fusion for face recognition. Neural Comput Appl. 2013;23:169–74.
8. Zhang X, Cao Y. A novel chaotic map and an improved chaos-based image encryption scheme. Sci World J. 2014;2014:1–8.
9. Awad A, Miri A. A new image encryption algorithm based on a chaotic DNA substitution method. In: IEEE international conference on communications; 2012. p. 963–71.
10. Cui H, Liu J, Chen P, et al. Facial feature extraction based on MB_LBP and improved Fast PCA algorithm. Microcomput Appl. 2015;15:29–32.

# Arousal Recognition Using EEG Signals

Xiang Ji, Xiaomin Tong, Xinhai Zhang, Yunxiang Yang(✉), Jing Guo, Bo Zhang, and Jing Cheng

China Academy of Electronics and Information Technology, Beijing 100041, China
yyxsdu@126.com

**Abstract.** As an indicator of emotion intensity, arousal plays an important role in emotion recognition. However, the accuracy rate-based EEG signals have been far away from human's satisfactory due to the lack of effective methods. In this paper, we propose a novel framework for recognizing arousal levels by using EEG signals. Instead of using time domain feature and frequency domain feature of EEG, we select the EEG feature directly from a large number of EEG signals by using the feature selection method after data standardization. Based on our method, feature with most distinguished ability has been found. The experimental results on the open data set DEAP show that the arousal accuracy has been significantly improved by using our method.

**Keywords:** Arousal recognition · Affective computing · EEG feature · Feature selection

## 1 Introduction

Affective computing aims to recognize the type and intensity of the user's emotion while watching video streams or accepting other stimulations. As an important factor of emotion tagging, arousal measures the users' attention levels [1]. Video with higher arousal means this video will attract more attention. Arousal recognition plays important roles in the video recommendation, video classification and advertising design.

Researches expect to use the electroencephalogram (EEG) responses to recognize the emotion when the user is watching videos due to human is actually the source of the emotion, which is called implicated video tagging [2, 3]. There were some researches that try to extract distinguished feature from EEG signals for recognizing the video arousal. For example, Koelstra et al. [4] recorded EEG signals of 32 participants as each watched 40 one-minute long excerpts of music videos. These signals were used to find a database named DEAP for researchers to use. Ningmeng [5] extracted time domain feature and frequency domain feature from EEG signals for emotion recognition.

However, the accuracy of recognition-based EEG signals is still far away from satisfactory. Most useful information has not been extracted from the EEG signals by using time domain feature and frequency domain feature extracted methods [5]. To solve this problem, instead of extracting time domain feature and frequency domain


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_49

feature, this paper extracted distinguished features from EEG signals directly by using data normalized and feature selection algorithms. The experimental result has indicated that our proposed framework outperforms the traditional methods.

This paper proposes a novel framework for the video arousal recognition based on EEG signals of the DEAP database. At first, in order to make the data with consistent range, the EEG signals of each participant have been normalized according to the column. Second, the most powerful feature is extracted directly from EEG signals by using the mRMR feature selection method [6–8]. Finally, the video arousal has been classified by SVM [9] based on the selected features. The experimental result demonstrated that, contrasting the traditional time domain feature and frequency domain feature, our proposed framework can extract more distinguished features directly and the accuracy of arousal recognition has been significantly improved.

## 2 DEAP Database

In the DEAP database, EEG signals of 32 participants were recorded when each participant watched 40 one-minute long excerpts of music videos. Thirty-two participants contain 16 females and 16 males. The health of all participants met the test requirement. There was no mental illness or cranial nerve injury. All participants had a good mental state, and normal hearing and eyesight. Before the experiment, each one has been told the attention matters. The experimental process kept quiet and not disturbed, so that the subjects could concentrate on the experiment. Each music video is 60 s long. Forty music videos come from emotion tagging and evaluation Web site. Thirty-two-channel EEG signals were recorded while watching music videos. Thirty-two-channel EEG was in accordance with the international standard. The dimensionality of EEG signals obtained in the experiment is $40 \times 32 \times 8064$, which means each participant watched 40 videos $\times$ 32-channel EEG signals $\times$ each channel record 8064 data in one minute. The experimental scene is shown in Fig. 1.

Each participant annotated arousal level when watching each video. The level is from 1 to 9. Low arousal level means the video is inactive (e.g., uninterested, bored). High arousal level means the video is active (e.g., alert, excited). Finally, for each participant, each video was annotated with arousal score. All the arousal scores are divided into two classes: One class is from level 1 to 5, which means low arousal, and the label of the video is defined as $-1$. The other class is from 6 to 9, which means high arousal, and the label of the video is defined as 1. This paper will research the video arousal two-classification problem and improve the accuracy significantly.

## 3 Feature Selection of EEG Signals

A lot of EEG signals will collect from the acquisition experiment (i.e., Fig. 1). The distinguished feature has been involved in these signals. Existing time and frequency domain feature cannot select all distinguished information. Hence, this paper proposed an effective method to select these distinguished information directly from the EEG signals, based on which the accuracy of classification has been significantly improved.

**Fig. 1.** The experiment of EEG acquisition [4]

In order to unify the data range, EEG signals corresponding 40 music videos from each channel of each participant were normalized. The data dimensionality is $40 \times 8064$, in which 40 means each participant watched 40 music videos and 8064 means data length collected from each channel within 1 min. The data is normalized according to the columns as follows:

$$Data(j) = (Data(j) - \mu(Data(j)))/\sigma(Data(j)) \tag{1}$$

in which $j$ means the $j$th column of $Data$ and $\mu(Data(j))$ means the mean value of the $j$th column of $Data$. $\sigma(Data(j))$ means the standard deviation of $j$th column of $Data$.

For EEG signals of each channel of each participant, the most distinguished feature could be selected by mRMR method. mRMR algorithm aims to select the most correlated feature for classification, and meanwhile the correlation between features is lowest, which fully reaches the request of the EEG feature selection. mRMR deified the following two functions for the largest correlation and smallest redundancy [6–8]:

$$\max D(S,c),\ D = \frac{1}{|S|}\sum_{x_i \in S} I(x_i;c) \tag{2}$$

$$\min R(S),\ R = \frac{1}{|S|^2}\sum_{x_i,x \in S} I(x_i,x_j) \tag{3}$$

in which $S$ means feature set. $c$ means the class. $I(x_i;c)$ means mutual information between feature and class. $I(x_i,x_j)$ means mutual information between feature and feature. The mutual information is defined as follows:

$$I(x;y) = \int\int p(x,y)\log\frac{p(x,y)}{p(x)p(y)}\mathrm{d}x\mathrm{d}y \tag{4}$$

in which $x$, $y$ mean tow random values. $p(x)$, $p(y)$, $p(x,y)$ mean the probability density functions.

The largest correlation between feature and class and the smallest redundancy between features are achieved through the following functions:

$$\max \Phi(D,R), \quad \Phi = D - R \tag{5}$$

Assume a data set $S_m$. The best feature set selects the feature from rest data $\{S - S_m\}$ which meet the largest value of the function (5) as follows:

$$\max_{x_j \in X - S_{m-1}} \left[ I(x_j;c) - \frac{1}{m-1}\sum_{x_j \in S_{m-1}} I(x_j;x_i) \right] \tag{6}$$

## 4 Arousal Recognition of Music Video

Algorithm flow of this paper is shown in Table 1. First, EEG signals corresponding 40 music videos from each channel of each participant were normalized. Second, mRMR algorithm is used to select feature from each EEG channel of each participant. Then, the selected feature is used to classify the arousal levels of 40 videos and obtain the accuracy of classification. Finally, the mean accuracy of 40 EEG channels is defined as the final accuracy of the 40 music videos.

As the same with the literature [5], four cross-validations were used in the experiment. Concretely, the EEG data of 40 videos was divided into four groups, with three servings as training sets each time and the remaining 1 as a test set. SVM classifier was repeated 4 times to ensure that each data was tested, and the result of the 4 test results was averaged. As a result, the classification accuracy of one channel for 40 videos is obtained.

## 5 The Proposed Algorithm of This Paper

**Algorithm 1: The algorithm of arousal recognition-based EEG signals**
**Input:** EEG signals of 32 objects watching 40 music videos.
**Output:** The accuracy of arousal classification of music video.
**Steps**: **1**. According to formula (1), the EEG signals of each subject were standardized according to the column data.
**2**. The EEG signals of each participant were selected by the mRMR feature selection method, and the EEG features with high classification accuracy were obtained.

**Table 1.** Arousal classification result of music video-based EEG signals

| Participant | Traditional time domain feature classification (%) | Traditional frequency domain feature classification (%) | Chi-square feature selection method | This paper (%) |
|---|---|---|---|---|
| S1 | 77.5 | 52.5 | 86.5 | 88.3 |
| S2 | 62.5 | 67.5 | 83.4 | 87.4 |
| S3 | 80 | 80 | 80.7 | 91.9 |
| S4 | 70 | 72.5 | 79.1 | 88.3 |
| S5 | 60 | 62.5 | 85.5 | 86.9 |
| S6 | 55 | 67.5 | 86.7 | 90.2 |
| S7 | 65 | 80 | 79 | 86.3 |
| S8 | 52.5 | 57.5 | 80. | 83.9 |
| S9 | 72.5 | 62.5 | 85 | 89.6 |
| S10 | 55 | 65 | 69.5 | 80.7 |
| S11 | 77.5 | 70 | 76.1 | 85.9 |
| S12 | 82.5 | 82.5 | 82.5 | 89.2 |
| S13 | 85 | 85 | 85.2 | 95.1 |
| S14 | 67.5 | 72.5 | 73.9 | 88.7 |
| S15 | 60 | 67.5 | 86.8 | 88 |
| S16 | 65 | 77.5 | 89.2 | 90.3 |
| S17 | 72.5 | 70 | 71.7 | 82.7 |
| S18 | 70 | 77.5 | 77.6 | 87.3 |
| S19 | 67.5 | 67.5 | 71.3 | 80.9 |
| S20 | 77.5 | 82.5 | 81.0 | 93 |
| S21 | 80 | 80 | 81.3 | 92.3 |
| S22 | 55 | 65 | 78.9 | 87.7 |
| S23 | 77.5 | 80 | 78.1 | 92 |
| S24 | 92.5 | 85 | 83.5 | 92.1 |
| S25 | 80 | 80 | 80.4 | 90.6 |
| S26 | 62.5 | 50 | 67.9 | 77.5 |
| S27 | 77.5 | 72.5 | 79.9 | 90.5 |
| S28 | 50 | 62.5 | 75.9 | 83.6 |
| S29 | 82.5 | 70 | 80.6 | 88 |
| S30 | 50 | 72.5 | 79.8 | 85.6 |
| S31 | 75 | 70 | 79.9 | 86.7 |
| S32 | 70 | 80 | 81.9 | 90.9 |
| Average | 69.6 | 71.5 | 80% | 87.9 |

**3**. The arousal of the music video is classified by the EEG feature of each channel, and the classification accuracy of 40 music videos per channel is obtained. Then, the classification accuracy of the 32 channels is averaged to get the accuracy rate of the pleasure classification of the 40 music videos.

**Output:** The classification accuracy rate, as shown in Table 1.

## 6 Experimental Result

In the DEAP database, there were 32 subjects and each participant watched 40 music videos. The number of EEG channels was 32. Each participant watched each video and produced 32 × 8064 EEG data. Thirty-two indicated 32 EEG channels per person. The length of the data collected by each channel in 1 min is 8064. EEG signals are standardized. Then, the mRMR algorithm is used to select the normalized data, and thus the arousal feature of each EEG channel for 40 videos is obtained. Based on the arousal feature of each channel, 40 videos were classified with arousal level. We calculate the average classification accuracy of 32 channels as the arousal accuracy of 40 videos for each participant. The results of the experiment are shown in Table 1, in which S1–S32 represent first subjects to thirty-second participant.

The traditional method extracts the EEG feature with six kinds of time domain characteristics (mean, standard deviation, the mean absolute value of the first order of the original signal, the mean absolute value of the second order of the original signal, first order of standardized signal and second-order difference mean absolute value) from the DEAP database. The Fourier transform of the five frequency bands is used for each frequency band, and the average energy of each EEG is taken as a frequency domain feature.

In order to demonstrate the effectiveness of our feature selection algorithm, we add the experiment on the validity of the feature selection method. Concretely, we use chi-square feature selection method to carry out the experiment. The chi-square feature selection algorithm is based on the chi-square statistics method, which through continuous iteration finds the effective feature of the data. The experimental results are shown in Table 1. From the results, we can see that the feature selection algorithm adopted in this paper can get higher classification accuracy than the chi-square feature selection.

The classification accuracy of time domain (i.e., the traditional method) and frequency domain feature (i.e., the traditional method) in 32 participants is shown in Table 1. The average classification accuracy of time domain feature is 69.6%, and the highest classification accuracy of time domain feature is 92.5%. The average classification accuracy of frequency domain feature is 71.5%, and the highest classification accuracy of frequency domain feature is 85%. The average accuracy of the chi-square feature selection algorithm is 80%, and the highest classification accuracy of the chi-square feature selection algorithm is 89.2%. Compared with the traditional method and the chi-square feature selection algorithm, the average accuracy of this paper is 87.9%, which is 18.3% higher than the traditional time domain feature, 16.4% higher than the traditional frequency domain feature and 7.9% higher than the chi-square feature

selection algorithm. Meanwhile, the highest accuracy of this paper among 32 participants is 95.1%, which is 2.6% higher than the traditional time domain feature, 10.1% higher than the traditional frequency domain feature and 5.9% higher than chi-square feature selection method. All these experimental results demonstrated our algorithm outperforms traditional algorithms.

The advantage of this paper is that original frequency and time domain features can only extract a limited number of features from the massive brain response EEG data. A large number of effective features will be filtered out, and as a result it is difficult to obtain all the distinguishing features (i.e., only six time domain features and five frequency domain features). Our proposed algorithm uses the normalize algorithm and selects feature with strong classification ability from the EEG data directly, so as to obtain higher classification accuracy.

## 7 Conclusion

It has been a hot research point for the computer to have human emotion recognition ability. Recently, many researchers have explored brain response to recognize the video emotion when human watching video. In order to improve the recognized ability based on the EEG signal, a novel framework is proposed in this paper. Compared with the traditional time and frequency domain feature, this paper uses data standardization and mRMR feature selection method to obtain more distinguished features and much higher arousal classified accuracy. The experimental results (i.e., Table 1) demonstrate that EEG signals involve much useful emotion information and this paper provides a very effective solution to the video emotion recognition.

**Acknowledgments.** This paper is supported by Beijing Nova Program (Z181100006218041) and National Key R&D Program of China (2017YFC0820106).

## References

1. Wu D, Courtney C, Lance B, Narayanan S, Dawson M, Oie K, Parsons T. Optimal arousal identification and classification for affective computing using physiological signals: virtual reality Stroop task. IEEE Trans Affect Comput. 2010;1(2):109–18.
2. Pantic M, Vinciarelli A. Implicit human-centered tagging. IEEE Signal Process Mag. 2009;26 (6):173–80.
3. Hanjalic A, Xu L. Affective video content representation and modeling. IEEE Trans Multimedia. 2005;7(1):143–54.
4. Koelstra S, Muhl C, Soleymani M, et al. Deap: a database for emotion analysis; using physiological signals. IEEE Trans Affect Comput. 2012;3(1):18–31.
5. Ningmeng H. Emotion recognition based on EEG. Guangzhou, China: South China University of Technology; 2016.
6. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max- relevance, and min-redundancy. IEEE Trans Pattern Anal Machine Intell. 2005;27(8):1226–38.

7. Li B, Hu L, Chen L, Feng K, Cai Y, Chou K. Prediction of protein domain with mrmr feature selection and analysis. PLoS ONE. 2012;7(6):e39308.
8. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expres. J Bioinform Comput Biol. 2005;3(2):185–205.
9. Chang C, Lin C. LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. 2011;2(3):27.

# Systematic Raptor Codes with UEP Property for Image Media Transmission

Guoqing Chen, Shushi Gu(✉), Ye Wang, Jian Jiao, and Qinyu Zhang

Communication Engineering Research Center, Harbin Institute of Technology, Shenzhen, Guangdong, China
chenguoqing@stu.hit.edu.cn,
{gushushi, wangye83, jiaojian, zqy}@hit.edu.cn

**Abstract.** In order to improve unequal error protection (UEP) for compressed images, a novel rateless transmitting scheme is proposed in this paper. Based on 3GPP Raptor codes, this scheme can accurately match the image with different priority to ensure the efficiency and reliability. These Raptor codes have the linear complexity of encoding and decoding. In addition, a new framework called Short Interweaving Systematic Raptor (SISR) codes is proposed. Based on Markov process, the theoretical analysis of degree distribution design is given. Simulation results show that, compared to the traditional Raptor code, SISR can prominently improve the PSNR to obtain better image quality with different packet loss rate.

**Keywords:** 3GPP-systematic Raptor code · Unequal error protection · Image media · Peak signal-to-noise ratio

## 1 Introduction

With the rapid development of networking and multimedia technology, streaming media business has become an important part of network services. According to the *The 41st China Statistical Report on Internet Development* [1], as of December 2017, Chinese Internet users have reached 772 million, the number of mobile Internet users reached 753 million, respectively. As of December 2017, the number of live webcast users reached 422 million, and streaming media services represented by short videos have achieved significant growth. There is no more important issue than ensuring the validity and stability of streaming media files.

Presently, forward error correction (FEC) codes on application layer attract much attention. Rateless codes, such as LT [2] and Raptor codes [3], have the characteristics about low complexity, flexible coding rate and non-feedback channel, which are very suitable for broadcast. The 3rd Generation Partnership Project (3GPP) has proposed 3GPP-systematic Raptor codes based on Multimedia Broadcast Multicast Service (MBMS) [4]. Systematic Raptor codes not


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_50

only improve the efficiency of decoding, but also enhance the quality of images, videos and other multimedia business. Besides, systematic Raptor codes are also adopt in studies that value high-quality video and huge file downloading in 5G for proper system configuration in application layer [5].

In some streaming media system, different parts of the data require different protections. Unequal error protection (UEP) [6] is of great significance to image recovery probability. In [7], an optimized UEP framework for single-layered video is proposed to solve different priorities of video packets. They develop an efficient algorithm to make use of the good characteristics codes at large block length. In [8], the authors put forward to that a guaranteed quality of service problem can be transformed into a convex optimization problem, and optimize rateless codes for UEP for streaming media. How to provide a kind of codes with low complexity and having the advantages of systematic codes is a critical problem.

In this paper, we present a kind of new UEP-systematic Raptor codes, termed Short Interlocking Systematic Raptor (SISR) codes. This code can flexibly adjust the redundancy of important encapsulation, so that the generator matrix is constructed easily. Through Markov analysis, we design a very short interweave distribution and obtain the optimal proportion of redundancy. Simulation results show that SISR can quickly approach a certain threshold with low redundancy, to ensure the high efficiency and reliability of image transmission.

## 2 System Model and Systematic Raptor Codes

Since systematic symbols and middle symbols are both introduced in 3GPP-systematic Raptor codes, decoding success rate and decoding complexity have improved significantly. Raptor coding bits consist of two parts: systematic and redundant symbols. The length of redundancy symbols can be changed according to channel conditions. The coding steps of 3GPP-systematic Raptor codes used in this paper are as follows.

Firstly, it is initialization: The length of source symbol $A$ is assumed as $K$: $a_1$, $a_2$, ..., $a_K$. Then calculate the number of the LDPC code $S$, the number of the Half code $H$ and the number of the intermediate code $L = K + S + H$ according to the length of source symbol $K$.

Secondly, as shown in Fig. 1, we construct the reversible generated matrix $\mathbf{G}$ using LDPC matrix $\mathbf{G}_{LDPC}$, Half matrix $\mathbf{G}_{Half}$ and LT matrix $\mathbf{G}_{LT}$ according to the relationship of pre-coding.

Thirdly, we calculate the matrix $\mathbf{G}^{-1}$, the inverse matrix of $\mathbf{G}$. Get the intermediate symbols $C$: $c_1$, $c_2$, ..., $c_L$, by (1). $A'$ is get from putting 0 with $(S+H)$ times at the beginning of the symbol $A$.

$$\boldsymbol{G}^{-1} \times A' = C \tag{1}$$

Fourthly, the original symbols which have $K$ bits are sent directly. Then encoding the intermediate symbols by LT code and getting the redundancy symbols with the length of $P$. The length of encoding symbols is $N = K + P$. The

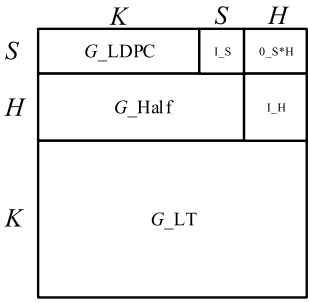



**Fig. 1.** Structure of encoding matrix $\boldsymbol{G}$.

encoding symbols are made by systematic symbols and redundancy symbols: $a_1$, $a_2$, ..., $a_K$, $p_1$, $p_2$, ... and $p_P$.

Encoding Symbol ID (ESI) is added to the head of packet. After erased on BEC, the decoding steps are as follows.

The first step is receiving first $K$ symbols which are original. Checking out that if packets lost by ESI. If there is no packet loss, we can get all source symbols directly. If lost, then jump to the second step;

The second step is all of the code symbols received are regarded as the LT code of intermediate symbol $C$ and structure decoding equations. To solve intermediate symbol $C$: $c_1$, $c_2$, ..., $c_L$, Gaussian decoding algorithm is used;

The third step is obtaining encoding symbols $A'$ by multiply matrix $\boldsymbol{G}$ and matrix $C$ by (2), where $\boldsymbol{G}$ is the generated matrix and $C$ is intermediate symbol. $K$ behind bits are the source symbols for accomplishing the decoding process.

$$\boldsymbol{G} \times C = A' \tag{2}$$

## 3 Design of UEP-Systematic Raptor Code

### 3.1 Short Intertwined Systematic Raptor Code

A scheme is proposed that is named Short Intertwined Systematic Raptor (SISR) codes. A short LT code which with the length of $P'$ is added between systematic symbols and redundant symbols. The short LT codes protect the important symbols to improve success rate of decoding process and has little influences on the performance of whole systematic decoding. The receiver uses Gaussian decoding algorithm to recover the important packets with the length of $K_1$ and higher priority. The algorithm of SISR is shown as in Table 1.

To illustrate clearly, we explain the algorithm of SISR for a compression image after wavelet transformation, with the code length of original symbols $K$=128. In three levels of image wavelet decompression, the whole image can be divided into level 1, level 2 and level 3 frequency components, and produces seven belts. Level 1 is the low-frequency area. SISR increases the rate of high-frequency area and decreases the rate of low-frequency area with a constant total

**Table 1.** Algorithm of SISR's encoding and decoding

| Algorithm of SISR's Encoding and Decoding |
|---|
| **Encoding**: |
| (1) $K$ bits length source symbols $A$ add with 0 and then multiplied with inverse matrix $\mathbf{G}^{-1}$, getting intermediate symbols $C$. Then we send the systematic bits $A$; |
| (2) Encode the important $K_1$ bits pockets $A_1$ by short LT code. Multiplied $A_1$ with $\mathbf{G}_{SISR}$ which is short LT coding generated matrix, getting interweaved symbols which with the length of $P$' and outputting it; |
| (3)Encode the intermediate $C$ symbols by LT code. Multiplied $C$ with LT code matrix $\mathbf{G}_{LT}$ so that we get the redundancy symbols. Then output it |
| **Decoding**: |
| (1) Check if the front $K$ symbols received successfully. If they are received completely, we get the source symbols without decoding. But if they are not received completely, then jump to the second step; |
| (2) Gaussian decoding the important symbols and the intertwined symbol, restore the original symbols $A_1$; |
| (3) Gaussian decoding the whole symbols received include the important symbols $A_1$ and the rest encoding code, restore the intermediated symbols $C$. Getting original $A$ symbols by multiplied $C$ with the generated matrix $\mathbf{G}$. Then decoding successfully |

bit rate $R_{tot}$. SISR coding scheme has excellent ability of image restoration. Short LT interweaved code plays an important role which makes SISR has UEP characteristics to protect low-frequency area.

### 3.2 Degree Distribution Design of Short LT Code

Important packets are usually more than the rest packets. In order to improve the recovery property of important packets, we construct short LT algorithm to encode them, making good use of their character of short length. It needs to design the degree distribution polynomial of short LT code, i.e., $T(x) = \sum_i t_i x$. The process of decoding can be described by Markov chain. A data packet decoded can be expressed as a state node. Its transition probability depends on the receiving probability of particular code package. The Markov process ends up with all nodes being decoded successfully.

By the transition matrix of Markov chain, the least code length of completely decoding is obtained as shown in (3):

$$E[T] = \frac{1}{t_1} + \frac{6t_1}{t_1 - 3} + \frac{18t_1}{(3 - t_2)(3 - 2t_1 - t_2)} + \frac{9t_1}{2(t_1 + t_2)(3t_1 + 2t_2)}. \tag{3}$$

In 3GPP-systematic Raptor codes, the systematic bits with the length of $K_1$ are regard as $K_1$ packets of degree-1. The solution is assigning degree-1 coefficient to the rest degrees according to the proportion until encoding degree distribution satisfies the coefficients of minimum $E[T]$. For $K = 3, 4, 5, 6$, the coefficients of degree distribution are shown in Table 2.

**Table 2.** Optimal weights for degree distribution in cases $K = 3, 4, 5, 6$

| $K$ | 3 | 4 | 5 | 6 |
|---|---|---|---|---|
| $t_1$ | 0.517 | 0.429 | 0.370 | 0.327 |
| $t_2$ | 0.397 | 0.430 | 0.451 | 0.467 |
| $t_3$ | 0.086 | 0.100 | 0.102 | 0.099 |
| $t_4$ | | 0.041 | 0.055 | 0.068 |
| $t_5$ | | | 0.021 | 0.024 |
| $t_6$ | | | | 0.014 |

Gaussian decoding can make a full use of received code symbols. We use Gaussian decoding in short code and LT code in SISR. The failure probability of each packet is $P_{SISR}$. Then we have

$$\varepsilon \geq P_{SISR} \geq \varepsilon(1 - \frac{\alpha}{K_1})^{N_1(1-\varepsilon)-k_1} \tag{4}$$

Among them, $\varepsilon$ is the loss rate of a single packet, $\alpha$ is the average degree of encoding packets, $K_1$ is the code length of original important packet, $N_1$ is the length of important encoding packet. As we known from the experience formula, the 3GPP-systematic Raptor code's decoding failure probability $P_F$ should be:

$$P_F(x) = \begin{cases} 1 & \text{if } N(1-\varepsilon) < K \\ 0.85 \times 0.567^{N(1-\varepsilon)-K} & \text{if } N(1-\varepsilon) \geq K \end{cases} \tag{5}$$

In SISR scheme, the failure probability of less important packets $P_{FL}$ equals to the whole failure probability of 3GPP-systematic Raptor codes, i.e., $P_{FL}(x) = P_F(x)$. By Gaussian decoding, the failure probability of short interweaved code will be $P_{FG}$=$1 - (1 - P_{SISR})^{k_1}$. Therefore, the failure probability of important packets in SISR scheme is $P_{FM}$:

$$P_{FM} = P_{FG} \cdot P_{FL} \tag{6}$$

### 3.3 Image Restore Performance Analysis

Image compression technology is a typical anti-interference measure in image transmission. CCSDS recommends specific standard compression algorithm of

discrete wavelet transform (DWT) for gray image compression. We use 9/7 discrete wavelet transform, which is under the CCSDS standard. The $j_{th}$ fundamental frequency component of original image is $LL_{j-1}(f)$:

$$LL_{j-1}(f) = LL_j(f) + HL_j(f) + LH_j(f) + HH_j(f) \tag{7}$$

The high-level restoration of image wavelet transform is the result of the multiple inverse transforming. The failure probability of EEP image transmission through BEC channel is $P_F$. We can get the restored image:

$$\begin{aligned} c_{LL}^{EEP}(n_1, n_2) =& (1 - P_F)(LLc_{LL}^1 + HLc_{HL}^1 \\ &+ LHc_{LH}^1 + HHc_{HH}^1) \end{aligned} \tag{8}$$

The SISR scheme protects the low-frequency component of the image $LL_{cLL^j}$. We can get received image:

$$\begin{aligned} c_{LL}^{SISR}(n_1, n_2) =& (1 - P_{FM})LLc_{LL}^1 \\ &+ (1 - P_{FL})(HLc_{HL}^1 + LHc_{LH}^1 + HHc_{HH}^1) \end{aligned} \tag{9}$$

Peak signal-to-noise ratio (PSNR) is a way to measure the quality of image transmission, and it is a international standard to measure the quality of images. As shown in (10), where MSE is the mean square error between original image preprocessing:

$$PSNR = 10 \cdot log_{10}(\frac{(2^n - 1)^2}{MSE}) \tag{10}$$

Mean square error of pixels MSE is shown as follows:

$$MSE = \frac{\sum_i \sum_j (s_{i,j} - r_{i,j})^2}{i \cdot j}, \tag{11}$$

where $s$ is the gray value of original image pixel, $r$ is the gray value of restored image pixel, $i, j$ are image rows and columns, respectively. Compared with the method of EEP, low-pass functions and high-pass functions have the characteristics of the integral $\sum h(n) = 1, \sum g(n) = 0$, that is to say low-frequency filters save energy while high-frequency filters lose energy. When $N \cdot (1 - \varepsilon) - k$ is small, $P_{FM} < P_F$ and $P_{FL}$ is little less than $P_F$. The SISR scheme protects more energy of image from losing. Area with a higher energy has a closer value between $r_{i,j}$ and $s_{i,j}$.

## 4 Simulation Results

First of all, we check the range of SISR image protection. Secondly, the best important package redundancy is obtained by a amount of simulation experiments. Thirdly, the comparisons with the other UEP-systematic Raptor codes are presented.

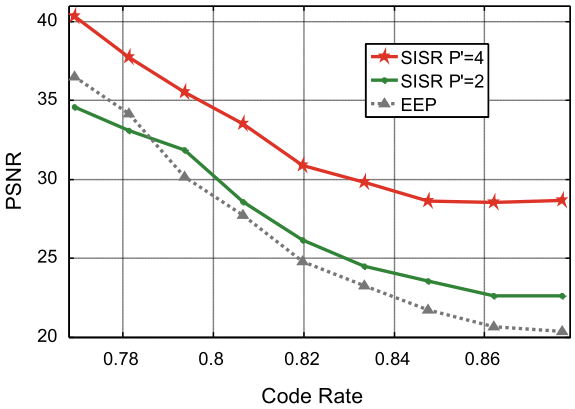
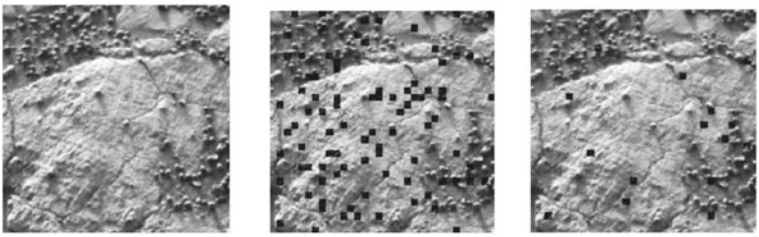



**Fig. 2.** Image protection of UEP and EEP with $K = 128$, $K_1 = 8$, $\varepsilon = 0.15$.

Compared the SISR scheme of different important redundant, we get the trend of image protection along with the change of coding rate. Simulation conditions are as follows. The original packet length is $K = 128$, important packet length is $K_1 = 8$, packet loss rate of channel is $\varepsilon = 0.15$, and the simulation times are 10,000. The simulation results are shown in Fig. 2. The star line, solid line and dotted line refer to the scheme of $P' = 4, P' = 2$ and EEP. With rate decreasing, PSNR of EEP scheme is higher than that of $P' = 2$ SISR scheme at the rate point of 0.79. PSNR of EEP scheme and $P' = 2$ SISR scheme achieves the standard value at the rate points of 0.79 and 0.80 respectively. However, PSNR of SISR scheme that is $P' = 4$ achieves the standard value at a better rate point 0.83. By using SISR, it approaches the standard value 30–40 dB more rapidly, which proves SISR has an excellent protection on images protection.



(a) Original image (b)EEP scheme (c) SISR scheme

**Fig. 3.** SISR-UEP versus EEP in restore with $\varepsilon = 0.15$.

Under the conditions that the erasure probability is $\varepsilon = 0.15$, original image, image of EEP-systematic Raptor codes and image of SISR scheme of the photographs are shown in Fig. 3a–c. Obviously, we can find that the SISR scheme significantly improves images transmission.

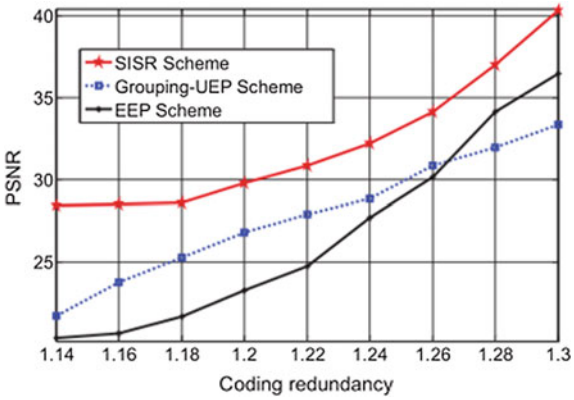



**Fig. 4.** Trend of SISR, grouping UEP and EEP with a increasing coding redundancy.

As shown in Fig. 4, these two UEP schemes make PSNR of image approach to 30dB faster than that of EEP scheme. PSNR of EEP scheme will be beyond the grouping UEP scheme after 1.26 redundancy rate. The proposed SISR scheme achieves 2–3 dB better than the grouping scheme and EEP scheme within 1.14–1.3 redundancy rate. Moreover, the EEP scheme would excess the SISR scheme when the redundancy rate is more than 1.3. In a word, SISR scheme can effectively protect the image at a normal redundancy rate.

## 5 Conclusions

In this paper, an UEP 3GPP-systematic Raptor codes scheme is studied, named SISR, for transmitting high-quality images on application layer. The scheme combines the original systematic Raptor codes and short LT codes, maintaining the good performance of systematic codes structure. In a constant code rate, we find the optimal redundancy of SISR short code $P'_{max}$ by Monte Carlo simulation. The results prove that SISR scheme has an outstanding performance of PSNR, which is much better than EEP scheme. Compared with other systematic Raptor codes, SISR is more efficient on image recovery and suitable for high-speed data transmission of streaming media.

**Acknowledgments.** This work was supported in part by the National Natural Sciences Foundation of China under Grant 61701136, Grant 61501140, Grant 61771158 and Grant 61525103, and China Postdoctoral Science Foundation Grant 2018M630357, and Shenzhen Basic Research Program under Grant JCYJ20170811154233370, Grant JCYJ20170811154309920 and Grant ZDSYS201707280903305.

## References

1. China Internet Network Information Center.: The 41st China Statistical Report on Internet Development, 2018; p. 21–5.
2. Luby M, Mitzenmacher M, Shokrollahi M, et al. Efficient erasure correcting codes. IEEE Trans J Inf Theory. 2001;47(2):569–84.
3. Shokrollahi A. Raptor codes. IEEE Trans J Inf Theory. 2006;52(6):2551–67.
4. 3GPP, Multimedia Broadcast/Multicast Service (MBMS); Protocols and Codecs, 3GPP TS 26.346 v10.0.0, Mar. 2011.
5. Xiong K, Zhang Y, Fan P, Yang HC. Evaluation framework for user experience in 5G systems on systematic rateless-coded transmissions. IEEE Access. 2016;4:9108–18.
6. Hsiao H-F, Ciou Y-J. Layer-aligned multipriority rateless codes for layered video streaming. IEEE Trans. 2014;24(8):2–5.
7. Luo Z, Song L, Zheng S, Ling N. Raptor codes based unequal protection for compressed video according to packet priority. IEEE Trans. 2013;15(8):2208–12.
8. Cao Y, Blostein SD, Chan W-Y. Optimization of unequal error protection rateless codes for multimedia multicasting. J Commun Networks. 2015;17(3):221–9.

# Novel Combination Policy for Diffusion Adaptive Networks

Qiang Fan[1], Wang Luo[1(✉)], Wenzhen Li[2], Gaofeng Zhao[1], Qiwei Peng[1], Xiaolong Hao[1], Peng Wang[1], Zhiguo Li[1], Qilei Zhong[1], Min Feng[1], Lei Yu[1], Tingliang Yan[1], Shaowei Liu[1], Yuan Xia[1], Bin Han[1], Qibin Dai[1], Jie Wang[3(✉)], and Guan Gui[3]

[1] NARI Group Corporation/State Grid Electric Power Research Institute, Nanjing, China
luowang@sgepri.sgcc.com.cn
[2] East Inner Mongolia Electric Power Company Limited, Hohhot, China
[3] Nanjing University of Posts and Telecommunications, Nanjing, China
2018010223@njupt.edu.cn

**Abstract.** Diffusion adaptive networks are received attractive applications in various fields such as wireless communications. Selections of combination policies greatly influence the performance of diffusion adaptive networks. Many diffusion combination policies have been developed for the diffusion adaptive networks. However, these methods are focused either on steady-state mean square performance or on convergence speed. This paper proposes an effective combination policy, which is named as relative-deviation combination policy and uses the Euclidean norm of instantaneous deviation between the intermediate estimation vector of alone agent and the fused estimation weight to determine the combination weights of each neighbor. Computer simulations verify that the proposed combination policy outperforms the existing combination rules either in steady-state error or in convergence rate under various signal-to-noise ratio (SNR) environments.

**Keywords:** Combination policy · Diffusion adaptive networks · ATC · CTA · Relative-deviation combination policy

## 1 Introduction

Diffusion adaptive networks consist of a collection of nodes with the ability of adaptation, learning, and information exchange. Adaptive networks have many applications ranging from research in biological networks [1, 2] to environmental monitoring and smart city [3]. There are many strategies for these problems in adaptive networks, such as consensus [4–6], incremental [7, 8], and diffusion strategies [9–13]. Compared to incremental strategies and consensus strategies, diffusion strategies are more stable, robust, and scalable [14]. Based on the above reasons, this paper will focus on diffusion implements.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_51

The implement of diffusion strategies has two phases. One stage named combination stage is that each sink collects and fuses information collected from its neighbors. Another stage named adaptation stage is that every node in the network updates its own estimation via least mean square algorithms. According to the implementation order of the two stages, two different diffusion algorithms named adapt-then-combine (ATC) diffusion least mean square (LMS) algorithms and combine-then-adapt (CTA) diffusion LMS algorithm have been described in [10]. Whether in ATC diffusion LMS algorithm or in CTA diffusion LMS algorithm, combination weights of every neighbor play an important role in the performance of adaptive networks.

Many combination policies have been proposed in previous research, such as averaging rule [15], Laplacian rule [16–18], metropolis rule [16], relative-degree rule [19], and relative-variance combination rule [20]. These combination rules either have a high steady-state error or have a low convergence rate. Inspired by the previous construction of combination policy, this paper proposes a new more effective combination rule. The proposed combination rule bases on Euclidean norm of instantaneous deviation between the intermediate estimation vector of alone agent and the fused estimation weight to determine combination matrix. Computer simulation results demonstrate that the new combination policy outperforms the existing combination rules in both mean square steady-state error and convergence rate.

**Notation**: $\mathbb{E}(\cdot)$ denotes the statistic expectation of a vector. $(\cdot)^{\mathrm{T}}$ denotes matrix transpose. $(\cdot)^{-1}$ denotes matrix inverse. The operator $\|\cdot\|$ denotes the $\ell_2$-norm of a vector. The operator $|.|$ denotes an absolute value function. And normal font letters denote scalars, boldface lowercase letters denote column vectors, and boldface uppercase letters denote matrices.

The rest of this paper is organized as follows. Section 2 introduces the system model and problem formulation. Traditional combination policies are reviewed, and a new more effective combination policy is proposed in Sect. 3. Computer simulations are conducted in Sect. 4. Section 5 concludes the paper.

## 2 System Model and Problem Formulation

Consider a collection of $N$ nodes distributed over a space region. They are going to estimate a common parameter $\mathbf{w}^o$ collaboratively which is a column vector of size M. Two nodes that can share information with each other are defined as neighbors. The number of neighbors of node $k$ is called the degree of node $k$ denoted by $n_k$. And neighbors of node $k$ are denoted by $\mathcal{N}_k$. At every time instant $i$, each agent $k$ senses a column signal vector $\mathbf{u}_{k,i}$ and obtains a scalar measurement $d_k(i)$. The covariance matrix of $\mathbf{u}_{k,i}$ is defined as $\mathbf{R}_{u,k} = \mathbb{E}\left\{\mathbf{u}_{k,i}\mathbf{u}_{k,i}^T\right\}$. The relation between the known information $\left\{d_k(i),\ \mathbf{u}_{k,i}\right\}$ and the unknown parameter $\mathbf{w}^o$ is constructed as follows:

$$d_k(i) = \mathbf{u}_{k,i}^T\mathbf{w}^o + v_k(i) \tag{1}$$

where $v_k(i)$ is additive white Gaussian noise (AWGN) in the environments. $v_k(i)$ is a stationary sequence of independent zero-mean random variables with a finite variance $\sigma^2_{v,\,k}$.

Each node in the network estimates the parameter $\mathbf{w}^o$ by minimizing local cost function defined as follows:

$$J_k(\mathbf{w}_k) = \mathbb{E}\left|d_k(i) - \mathbf{u}_{k,i}^T \mathbf{w}_k\right|^2 \tag{2}$$

The network seeks the optimal estimation of $\mathbf{w}^o$ cooperatively by minimizing the global cost function constructed as follows:

$$J^{\text{global}}(w) = \sum_{k=1}^{N} \mathbb{E}\left|d_k(i) - \mathbf{u}_{k,i}^T \mathbf{w}\right|^2 \tag{3}$$

One diffusion strategy for solving the above problem adaptively in a distributed manner is ATC diffusion LMS algorithm [21]. The implement of the diffusion strategy is described in detail below. Firstly, we should design an $N \times N$ matrix $\mathbf{A}$ with non-negative entries $\{a_{\ell,k}\}$ satisfying:

$$\mathbb{I}_{N\times 1}^{\mathrm{T}} \mathbf{A} = \mathbb{I}_{N\times 1}^{T}, \text{ and } a_{\ell,k} = 0 \quad \text{if } \ell \notin \mathcal{N}_k \tag{4}$$

Here, $\mathbf{A}$ is a combination matrix and $a_{\ell,k}$ denotes the combination weight of node $\ell$ when it shares information with node $k$ as shown in Fig. 1a, and $\mathbb{I}_{N\times 1}$ is a column vector of size $N$ with all its entries as one. The update equation of ATC diffusion LMS algorithm is defined as follows:

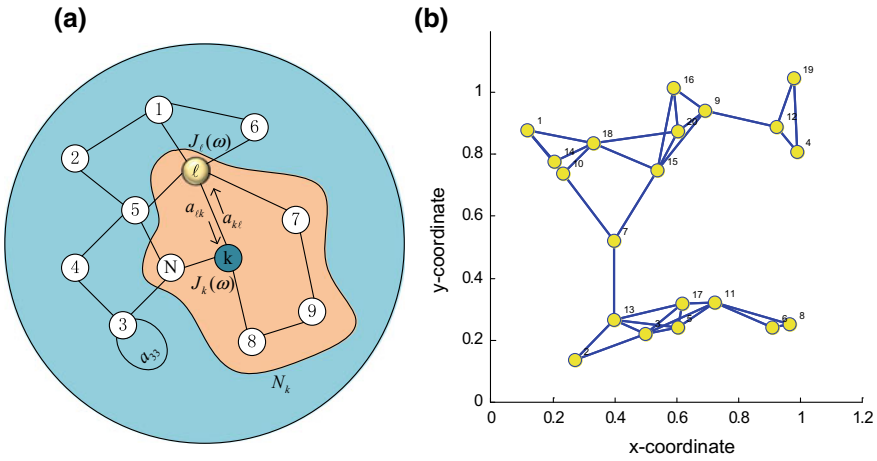



**Fig. 1.** **a** A diffusion network and the neighborhood of agent $k$ are denoted by the brownfield. **b** Topology of the adaptive network

$$\begin{cases} \boldsymbol{\psi}_{k,i} = \mathbf{w}_{k,i-1} + \mu_k \mathbf{u}_{k,i}^{*}\left[d_k(i) - \mathbf{u}_{k,i}\mathbf{w}_{k,i-1}\right] \\ \mathbf{w}_{k,i} = \sum_{\ell \in \mathcal{N}_k} a_{\ell,k}\boldsymbol{\psi}_{\ell,i} \end{cases} \tag{5}$$

where $\mu_k$ is the step size of node $k$. The update equation has two steps including incremental update and spatial update as Eq. (5) shows. In incremental update phase, the node $k$ uses its own known data $\left\{d_k(i), \mathbf{u}_{k,i}\right\}$ to update its estimated value by using stochastic gradient decent method from $\mathbf{w}_{k,i-1}$ to an intermediate estimation $\boldsymbol{\psi}_{k,i}$. In the spatial update phase, sink $k$ fuses the intermediate estimation $\left\{\boldsymbol{\psi}_{k,i}\right\}$ collected from its neighborhood through combination weights $\left\{a_{\ell,k}\right\}$ to obtain the updated weight estimation $\mathbf{w}_{k,i}$. Naturally, the selection of combination coefficients $\left\{a_{\ell,k}\right\}$ influences the performance of the network.

## 3 Review the Previous Combination Rules and Propose a Novel Combination Policy

Many combination rules have been designed in the literatures. Three represents of the previous combination policies are described in detail in Table 1. The first two combination policies mainly depend on the degree of the nodes to allocate combination weights. These selections may degrade the performance of adaptive networks because of ignoring the noise profile across a network. Since the nodes accessing to more neighbors may be nosier than the other nodes, it is insufficient to construct the combination weights to their neighbors by only relying on the degree of the nodes. Therefore, it is an important task to design the combination policy by taking into account the noise profile existing in the nodes as well.

**Table 1.** Three represents of the previous combination policies

| Combination rules | Reference |
|---|---|
| $a_{\ell k} = \begin{cases} \frac{1}{n_k}, & \text{if } k \neq \ell \text{ are neighbors or } k = \ell \\ 0, & \text{otherwise} \end{cases}$ | Uniform policy [15] |
| $a_{\ell k} = \begin{cases} \frac{n_\ell}{\sum_{m \in N_k} n_m}, & \text{if } k \text{ and } \ell \text{ are neighbors or } k = \ell \\ 0, & \text{otherwise} \end{cases}$ | Relative-degree policy [19] |
| $a_{\ell,k}(i) = \begin{cases} \frac{\sigma_{\ell,k}^{-2}(i)}{\sum_{j \in \mathcal{N}_k} \sigma_{j,k}^{-2}(i)}, & \text{if } \ell \in \mathcal{N}_k \\ 0, & \text{otherwise} \end{cases}$ | Relative-variance policy [20] |

The third combination rule in Table 1 named relative-variance policy bases on the noise quality of each node to determine their combination weights. The noise variance of agents is obtained by relying on instantaneous data approximations, and its recursion formula is derived specifically in Table 2. But the relative-variance policy achieves the better performance in terms of steady-state error at a large cost of low convergence rate.

Inspired by the relative-variance combination rule, this paper proposes a more simple and effective combination policy. Instead of the square of Euclidean norm of the instantaneous deviation between $\boldsymbol{\psi}_{k,i}$ and $\mathbf{w}_{k,i-1}$ to estimate noise degree across the network, we let the Euclidean norm of instantaneous deviation between $\boldsymbol{\psi}_{k,i}$ and $\mathbf{w}_{k,i-1}$ determine the combination weights of each neighbor. The proposed new combination policy named relative-deviation combination policy is constructed as follows:

$$a_{\ell,k} = \begin{cases} \dfrac{\left\|\boldsymbol{\psi}_{\ell,i}-\mathbf{w}_{\ell,i-1}\right\|^{-1}}{\sum_{m\in\mathcal{N}_k}\left\|\boldsymbol{\psi}_{m,i}-\mathbf{w}_{m,i-1}\right\|^{-1}}, & \text{for} \quad \ell\in\mathcal{N}_k \\ 0, & \text{others} \end{cases} \tag{6}$$

In this selection of combination policy, sink $k$ fuses the intermediate information from its neighbors in proportion to the inverses of the Euclidean norm of the instantaneous deviation between $\boldsymbol{\psi}_{k,i}$ and $\mathbf{w}_{k,i-1}$. The combination rule is practically meaningful. $\left\|\boldsymbol{\psi}_{\ell,i}-\mathbf{w}_{\ell,i-1}\right\|^{-1}$ plays the same role as $\left\|\boldsymbol{\psi}_{\ell,i}-\mathbf{w}_{\ell,i-1}\right\|^{-2}$ and is more direct than the latter.

**Table 2.** Derivation procedure for recursion formula of $\sigma_{\ell,k}^{-2}(i)$

The ATC algorithm is described as follows:

$$\boldsymbol{\psi}_{k,i} = \mathbf{w}_{k,i-1} + \mu_k \mathbf{u}_{k,i}^{*}[d_k(i) - \mathbf{u}_{k,i}\mathbf{w}_{k,i-1}] \tag{7}$$

$$\mathbf{w}_{\mathrm{k,i}} = \textstyle\sum_{\ell\in\mathcal{N}_k} a_{\ell,k}\boldsymbol{\psi}_{\ell,i} \tag{8}$$

For sufficiently small step-sizes, the estimation $\mathbf{w}_{k,i-1}$ approaches $\mathbf{w}^{o}$, by using model equation (1), an equation can be written as the follows:

$$\boldsymbol{\psi}_{k,i} \approx \mathbf{w}^{o} + \mu\mathbf{u}_{k,i}^{*}v_k(i) \tag{9}$$

And then:

$$\mathrm{E}\left\|\boldsymbol{\psi}_{k,i} - \mathbf{w}^{o}\right\|^2 \approx \mu^2\sigma_{v,k}^2\mathrm{Tr}(\mathbf{R}_{u,k}) \tag{10}$$

Using instantaneous approximation:

$$\mathrm{E}\left\|\boldsymbol{\psi}_{k,i} - \mathbf{w}^{o}\right\|^2 \approx \left\|\boldsymbol{\psi}_{k,i} - \mathbf{w}_{k,i-1}\right\|^2 \tag{11}$$

Combine (9) and (10) and the recursion formula is derived as the follows:

$$\sigma_{\ell,k}^2(i) = (1-v_k)\sigma_{\ell,k}^2(i-1) + v_k \times \left\|\boldsymbol{\psi}_{\ell,i} - \mathbf{w}_{\ell,i-1}\right\|^2 \tag{12}$$

where $v_k$ is a positive value close to 1.

## 4 Simulation Results

In this section, we conduct computer simulations to compare the proposed combination policy to previous combination rules including uniform, relative-degree, and relative-variance combination rules. We conduct two experiments in various SNR environments. Among these simulations, the parameter values of the system model and algorithms are uniform except SNR. Parameter values are set as in Table 3.

**Table 3.** Parameter values of the simulations

| Parameters | Values |
|---|---|
| The number of agents | $N = 20$ |
| Input signal | Pseudorandom binary sequences |
| The length of the estimated parameter | $M = 64$ |
| Distribution of the estimated parameter | $\mathrm{CN}(0, 1)$ |
| Gaussian noise distribution | $\mathrm{CN}\left(0, \delta_n^2\right)$ |
| Step size of all agents | $\mu = 0.005$ |

In all experiments, we set uniform step sizes across all agents and all $\mathbf{R}_{u,k}$ equal to each other. The network in this work consists of 20 nodes, and the topology of the network is designed that follows a rule that if the distance of two nodes is less than or equal to a threshold value, the two nodes are defined as neighbors. The topology is generated as shown in Fig. 1b.

In a real setting, each agent in the adaptive network expending a large geographic region has a distinct separation in the quality of SNR. Taking the practical conditions into consideration, this paper conduct the experiments with each node has different SNR environments. We design two experiments under different noise degree environments to confirm the effectiveness of the new combination policy. Two different noise degree environments are defined as bad noise degree environments (the value of SNR from 5 to 10 dB) and general noise degree environments (the value of SNR from 10 to 20 dB), respectively. The computer simulation results are shown as follows.

Figure 2a shows the average performance of all nodes under different combination policies in bad noise degree environments, and Fig. 2b shows the steady performance of every node under the same conditions. As the results in Fig. 2a, b demonstrate, whether regarding the average performance of all nodes or the steady performance at each node, relative-variance combination policy outperforms uniform and relative-degree combination rules regarding steady-state error at cost convergence rate. Compared to relative-variance combination policy, our proposed policy named relative-deviation combination policy obtains superior performance in terms of steady-state error and convergence rate under the bad noise degree environments.

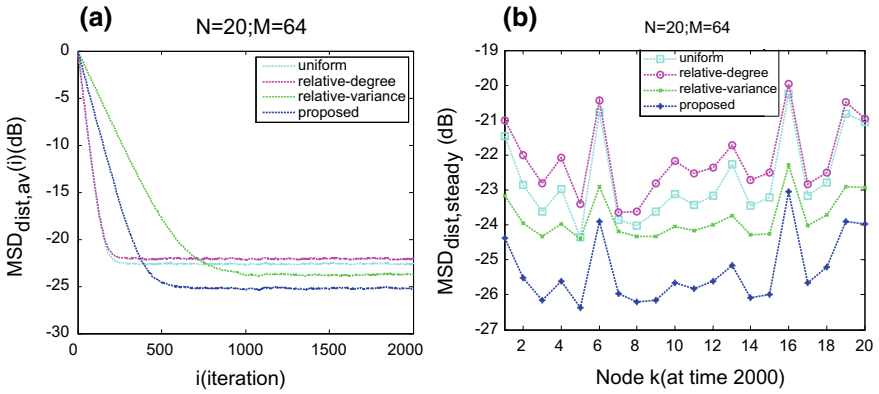



**Fig. 2.** **a** Average performance of all nodes under different combination policies in bad noise degree environments. **b** Steady performance of every node under bad noise degree environments

Figure 3a shows the average performance of all nodes under different combination policies, in general noise degree environments, and Fig. 3b shows the steady performance of every node under the same conditions. In Fig. 3a, b, along with the quality of SNR enhancing, the superiority of combination policies based on noise profile across the network including relative-variance and the proposed relative-deviation combination policy significantly improved regarding steady-state error. Compared to relative-variance combination policy, the proposed relative-deviation combination rule keeps considerable merits in both steady-state error and convergence rate under these conditions.

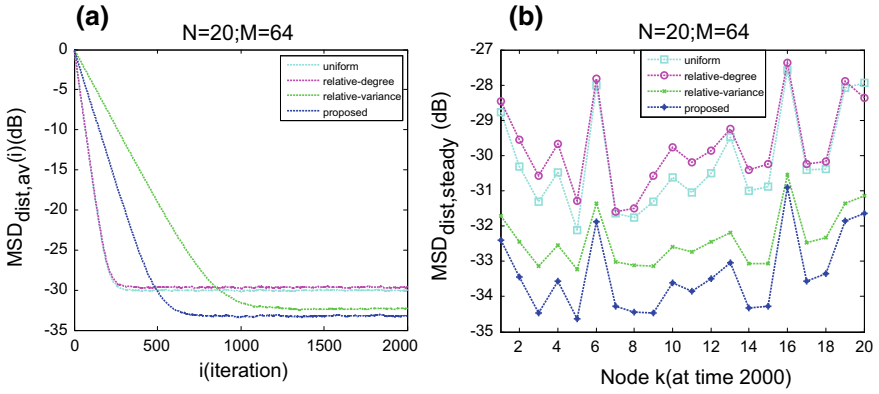



**Fig. 3.** **a** Average performance of all nodes under general noise degree environments. **b** Steady performance of every node under general noise degree environments

## 5 Conclusion

This paper proposed an effective combination policy named relative-deviation combination policy. We conducted two experiments under two different noise degree environments. The computer simulation results demonstrated that the proposed new combination policy outperforms the existing combination rules in terms of both steady-state error and convergence rate in various SNR environments. The next work is to seek more essential factors in combination policies and construct more effective combination rules.

**Acknowledgments.** This research was funded by State Grid Corporation Science and Technology Project (named "research on intelligent preprocessing and visual perception for transmission and transformation equipment").

## References

1. Tu SY, Sayed AH. Mobile adaptive networks. IEEE J Sel Top Signal Process. 2011;5(4):649–64.
2. Cattivelli FS, Sayed AH. Modeling bird flight formations using diffusion adaptation. IEEE Trans Signal Process. 2011;59(5):2038–51.
3. Chen L, IEEE MS, Ho Y, Lee H, Wu H, Liu H. An open framework for participatory PM2.5 monitoring in smart cities. IEEE Access. 2017;5:14441–54.
4. Nedic A, Ozdaglar A. Distributed subgradient methods for multi-agent optimization. IEEE Trans Automat Contr. 2009;54(1):48–61.
5. Kar S, Moura JMF. Distributed consensus algorithms in sensor networks with imperfect communication: link failures and channel noise. IEEE Trans Signal Process. 2009;57(1):355–69.
6. Srivastava K, Nedic A. Distributed asynchronous constrained stochastic optimization. IEEE J Sel Top Signal Process. 2011;5(4):772–90.
7. Rabbat MG, Nowak RD. Quantized incremental algorithms for distributed optimization. IEEE J Sel Areas Commun. 2005;23(4):798–808.
8. Lopes CG, Sayed AH. Incremental adaptive strategies over distributed networks. IEEE Trans Signal Process. 2007;55(8):4064–77.
9. Chen J, Richard C, Hero AO, Sayed AH. Diffusion LMS for multitask problems with overlapping hypothesis subspaces. In: IEEE international workshop on machine learning for signal processing; 2014. p. 1–6.
10. Sayed AH. Adaptive networks. Proc IEEE. 2014;102(4):460–97.
11. Sayed AH, Tu SY, Chen J, Zhao X, Towfic Z. Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior. IEEE Signal Process Mag. 2013;30:155–71.
12. Sayed AH. Diffusion adaptation over networks. E-Ref Sig Process. 2013;61:1419–33.
13. Chen J, Sayed AH. Diffusion adaptation strategies for distributed optimization and learning over networks. IEEE Trans Signal Process. 2012;60(8):4289–305.
14. Sayed AH. Adaptation, learning, and optimization over networks. Found Trends Mach Learn. 2014;7(4–5):1–501.

15. Tu S, Member S, Sayed AH. Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks. IEEE Trans Signal Process. 2012;60 (12):6217–34.
16. Blondel VD, Hendrickx JM, Olshevsky A, Tsitsiklis JN. Convergence in multiagent coordination, consensus, and flocking. In: Proceedings of the 44th IEEE conference on decision and control, and the European control conference; 2005. p. 2996–3000.
17. Xiao L, Boyd S. Fast linear iterations for distributed averaging. In: Pmedingr of the 42nd LEEK conference on UecMon and Conlrd; 2003. p. 65–78.
18. Scherber DS, Papadopoulos HC. Locally constructed algorithms for distributed computations in Ad-Hoc networks. In: Information processing in sensor networks (IPSN); 2004. p. 11–9.
19. Xiao L, Boyd S, Lall S. A scheme for robust distributed sensor fusion based on average consensus. In: Information processing in sensor networks; 2005. p. 63–70.
20. Cattivelli FS, Lopes CG, Sayed AH. Diffusion recursive least-squares for distributed estimation over adaptive networks. IEEE Trans Signal Process. 2008;56(5):1865–77.
21. Tu S, A. Sayed AH. Optimal combination rules for adaptation and learning over networks. In: IEEE international workshop on computational advances in multi-sensor adaptive processing (CAMSAP); 2011. p. 317–20.

# A Study of Transmission Characteristic on OFDM Signals Over Random and Time-Varying Meta-surface of the Ocean

Yu-Han Dai[1], Hui Li[1,2(✉)], Yu-Cong Duan[1], and Yan-Jun Liang[1]

[1] School of Information Science and Technology, Hainan University, Haikou 570228, China
hitlihuilll2@163.com
[2] School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China

**Abstract.** With the promotion of "the Belt and Road" and twenty-first-century "Maritime Silk Road" and the establishment of an information network in Hainan Province, the construction of network covering Southern China Sea is particularly important, which should combine with cable, wireless and satellite transmissions. The complexity of the impact of the marine environment on radio wave propagation and the particularity of communication between vessels are important factors to be considered when designing marine radio communication systems. This paper studies the models of sea channel for sea surface diffuse reflection. In this model, the performance of OFDM transmission was analyzed.

**Keywords:** Marine communication · OFDM · Diffuse reflection · Channel modeling

## 1 Introduction

China is a country with a long coastline, which has millions of square kilometers of ocean area. At the same time, China's social economy is developing rapidly and the international trade is increasing. It is urgent to build a strong modern navy because of the security of international trade personnel. Secondly, due to the particularity of the maritime communication business in the geographical conditions and development time, there is still a large gap between the terrestrial wireless data transmission technology and the maritime communication technology.

At present, offshore of wireless data transmission system has many shortcomings in the transmission reliability, transmission rate, transmission distance and anti-interference ability in the domestic. The problem of the contradiction between the limited frequency resources and the development demand for the wireless communication restricts the development of high-speed data communication in the modern port and coastal ships is becoming more and more serious. Therefore, there is a wide application market and development prospects in the wireless transmission technology


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_52

in the communication in ship shore. In recent years, with the deepening of the informatization and modernization information construction and unceasing enhancement of the information degree, the wireless communication system will turn into the high-speed broadband wireless communication from the narrowband and low speed.

The application of orthogonal frequency division multiplexing (OFDM) is analyzed and studied in this paper. OFDM has high spectrum efficiency and strong ability to resist the frequency-selective fading. It can be effective against the interference in the signal waveform and is used in a multipath environment and high-speed data transmission of the fading channel [1]. The literature [2, 3] adopted an improved algorithm of the maximum likelihood estimation of suboptimal to realize the signal detection of OFDM. The simulation results show that the improved maximum likelihood estimation algorithm is much smaller than the maximum likelihood estimation algorithm. The literature [4] used the sixth-, fourth- and second-order cumulants to detect OFDM signals. This algorithm deals with an intermediate frequency; therefore, it is not necessary to synchronize with the carrier of the signal. However, the same algorithm is based on OFDM signal obeying normal distribution under the Gaussian channel, while the single carrier does not have a foundation, so the algorithm is only applicable to the Gaussian channel. The literature [5] proposed a strong self-correlation to carry out OFDM detection, which has low computational complexity and good detection performance under the Gaussian channel. The literature [6] studied the signal detection and parameter estimation of OFDM in Rayleigh fading channel. The literature [7, 8] mainly studied the role of OFDM in 4G mobile communication system. The literature [9, 10] studied the future OFDM wireless communication outlook and OFDM simulation on IEEE 802.16 system. The literature [11, 12] studied the signal optimization design in OFDM communication system and the influence of residual time error on channel estimation.

At present, OFDM technology is mainly used for cellular systems on the ground. A large number of documents studied OFDM transmission channel and the related technology of ground in the cellular, but OFDM technology in the application of sea still lacks a certain research. The literatures [13, 14] proposed the future maritime mobile communication air interface and the application analysis of OFDM technology in a maritime wireless communication system. Therefore, the application of OFDM in a maritime wireless communication system is studied in this paper.

## 2 System Model

### 2.1 Characteristics of Marine Microwave Transmission

Compared with terrestrial communication, the microwave communication has its own environmental characteristics. First of all, in the topography, there are a few obstacles in the sea; the direct result of this is the large gap between radio waves. So when the wave travels at sea, the diffraction loss is smaller than on land. At the same time, the propagation gap increases, increasing the wave reflection. In addition, if the grazing angle is small, the reflection coefficient in the microwave band is relatively large, and the effect of the reflected wave is greater than that on land [15]. In the communication

process, the microwave can be regarded as the propagation in free space, then consider the influence of topography and atmosphere on the blocking, reflection and atmospheric absorption of the microwave, and then combine the two. In the process of sea microwave transmission, in addition to the free space loss, there is mainly the surface reflection of the microwave. Secondly, as a carrier for antenna buoy highly generally is low, the surface of the surge of the microwave transmission can be equivalent to the influence of the obstacle influence on the microwave shade at sea. In addition, the atmospheric absorption of the sea and the attenuation of rain and fog also have an impact on the microwave transmission.

## 2.2 OFDM Technology and Performance Analysis

OFDM is a special kind of multicarrier modulation technology, it makes the channel divide into several orthogonal subchannels, and high-speed data signal is converted into parallel low-speed data streams and then modulated on each channel to transmit. The orthogonal signals can be separated by using the relevant technology at the receiving terminal, which can reduce the intersymbol interference (ISI) between the subchannels. Each channel on the signal bandwidth is less than the channel bandwidth, so every subchannel can be looked on as a flat fading channel, in order to eliminate the intersymbol interference. And because the bandwidth of each subchannel is only a small part of the original channel bandwidth, channel equalization becomes relatively easy to get [16]. OFDM can make full use of channel bandwidth without a band-pass filter to select the waveform that is orthogonal even if aliasing. Schematic diagram of OFDM is plotted in Fig. 1.

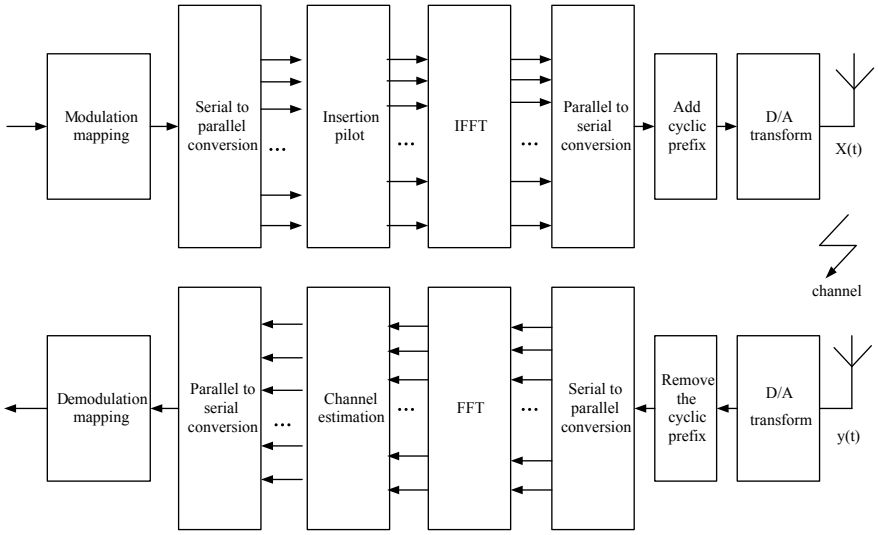



**Fig. 1.** Basic schematic diagram of OFDM

## 3 Fading Channel Model

When the radio waves transport in the sea, marine radio wave propagation and channel modeling studied the diffuse channel multipath component in the received signal at the receiving terminal are not only caused by adjacent vessels, but the rough sea surface also can produce diffuse reflection caused by multipath components [17]. The random phase and amplitude in different multipath components can cause the fluctuation of signal intensity, resulting in small-scale decline and multipath interference effect. The multipath effect of the sea surface is shown in Fig. 2.

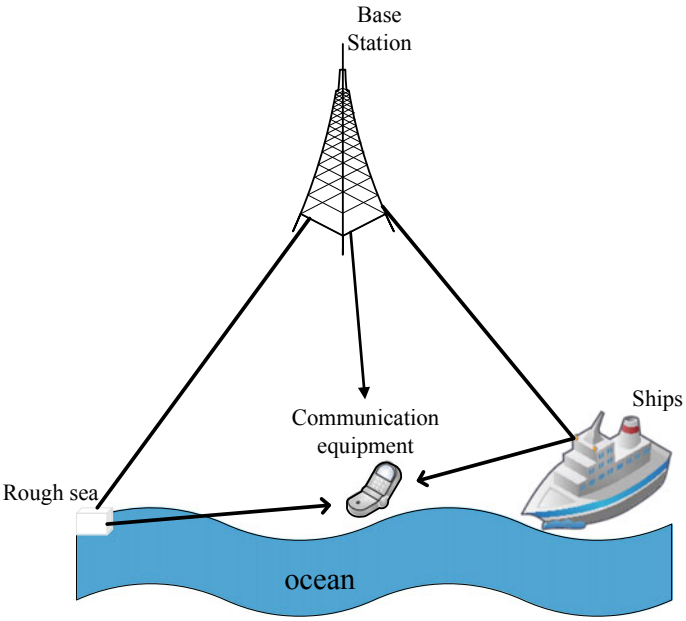



**Fig. 2.** Multipath effect of the sea surface

The propagation path of radio waves in the rough sea surface consists of three parts, as shown in Fig. 3, namely direct path, specular reflection path and diffuse reflection path. The direct path refers to the visual path between the radio frequency transmitter and the receiver. A specular reflection path is the path of a specular reflection. Diffuse path refers to the path through rough surface diffuse reflection [18]. Therefore, the model that can approximate the signal propagation under the rough sea surface is equivalent to the Rician channel model.

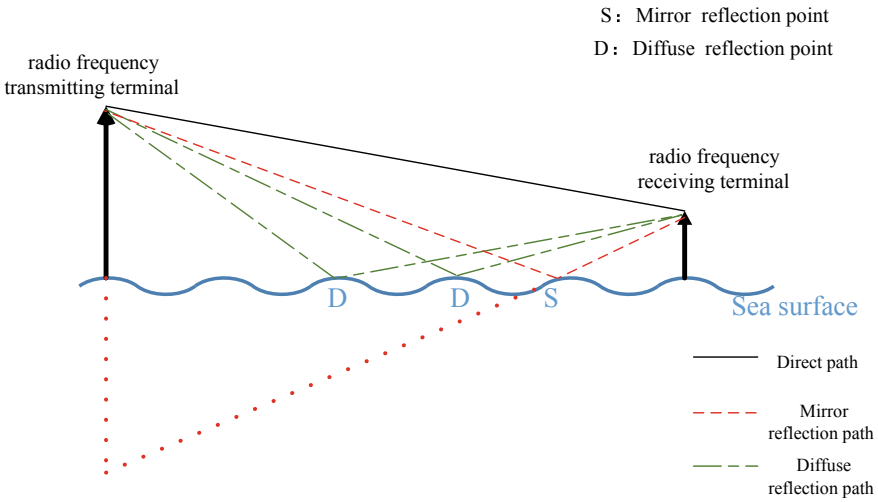



**Fig. 3.** Multipath model considering the diffuse reflection of the sea surface

### 3.1 Rician Channel Model

When there is a static (non-fading) signal in the receiver, such as the direct wave (LOS, line of sight) component (there is a direct wave on the sea), the envelope of the signal received by the receiver obeys the Rician distribution. In this case, the multipath component randomly arriving from different angles superimposed a DC component [19]. When the main signal component is weakened, the Rician distribution will change to Rayleigh distribution. The probability density function of Rician distribution is

$$p(r) = \begin{cases} \frac{r}{\sigma^2}\exp(-\frac{r^2+c^2}{2\sigma^2})I_0(\frac{Cr}{\sigma^2}), & C \geq 0,\ r \geq 0 \\ 0, & r < 0 \end{cases} \tag{1}$$

Among them, C represents the peak value of the main signal components, R means the received signal envelope, and $I_0(\cdot)$ represents the first modified Beckinsale function. In order to better analyze the Rician distribution, we define the ratio of the power of the main signal to the variance of multipath components as the Rician factor $K$, and the expression of $K$ can be written as

$$K = \frac{C^2}{2\sigma^2} \tag{2}$$

The Rician distribution is completely determined by the Rician factor $K$. The common dB is expressed as

$$K = 10\log\frac{C^2}{2\sigma^2} \tag{3}$$

When $C = 0$ and $K = -\infty$ dB, it shows that the multipath fading is the largest and the Rician distribution becomes Rayleigh distribution. When $K = +\infty$, there is no fading phenomenon. Figure 4 shows the probability density function of the Rician distribution.

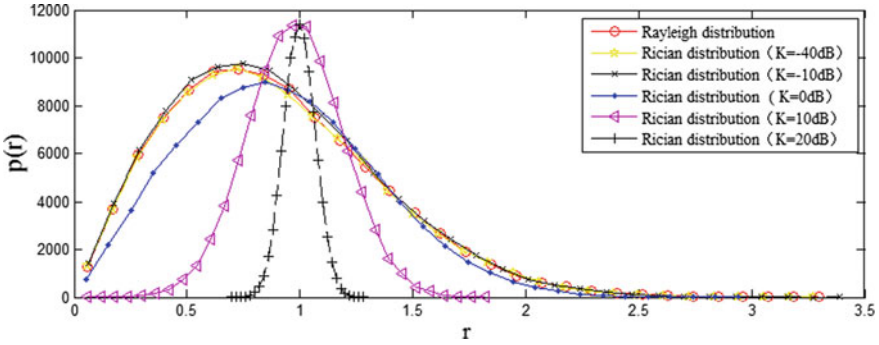



**Fig. 4.** Probability density function of different Rician factor distributions

### 3.2 OFDM System Model

Due to the reflection of direct and reflex of various obstacles between the mobile station and the base station in the rough sea, the received signals are usually composed of multipath signals. Due to the random variation of the phase, amplitude and arrival time of the multipath signal, the rapid fluctuation of the envelope of the received signal is caused, which is usually called fast fading. The fast fading variation of the envelope amplitude of the received signal obeys the Rician distribution. At the same time, it can be proved theoretically that the envelope amplitude of the received signal obeys the Rician distribution by using the Fresnel principle of the equivalent wave propagation [20].

OFDM equivalent baseband model: Set $N$ to represent the number of subcarriers, $T$ to express the OFDM symbol period, $d_i$ as the data symbol assigned to the sub-channel of the $i$, and $f_i$ as the carrier frequency of the subcarrier of the $i$, and the OFDM symbol can be expressed as

$$s(t) = \sum_{i=0}^{N-1} d_i e^{j2\pi f_i t}, \quad 0 \le t \le T \tag{4}$$

$f_i = f_c + i/T$, and $S(T)$ is sampled at the rate of $N/T$, that is, $t = kT/N$, and $0 \le k \le N - 1$; then, the equivalent baseband signal of OFDM symbol can be expressed as

$$S_k = S\left(\frac{kT}{N}\right) = \sum_{I=0}^{N-1} d_i e^{j2\pi ik/N}, \quad 0 \le k \le N-1 \tag{5}$$

It can be seen that $S_k$ is equivalent to inverse discrete Fourier transform (IDFT) operation for $d_i$. Similarly, at the receiving end, discrete Fourier transform (DFT) operation on $S_k$ can restore the transmitted data symbol $d_i$. Therefore, the modulation and demodulation in the OFDM system can be implemented by IDFT/DFT operations, respectively. In actual application, the further use of fast Fourier transformation (FFT) operation can significantly reduce the complexity of the system operation.

In the OFDM system, the lower bit rate makes the OFDM system effectively resist the intersymbol interference (ISI) caused by multipath propagation. In order to eliminate intersymbol interference to the maximum extent, the protection interval is inserted between each symbol of OFDM, and the length of the protection interval is generally greater than the maximum delay extension in the wireless channel and can completely overcome the impact of ISI. However, due to the influence of multipath propagation, the orthogonality between subcarriers is also destroyed, resulting in interchannel interference (ICI). In order to further eliminate ICI, a prefix is formed in the protection interval guard interval (GI), which is called cyclic prefix. Because of the cyclic prefix, the difference between the periods of each subcarrier in the symbol period of OFDM is an integer. So ensuring that the orthogonality between the subcarriers will not destroyed by the influence of the multipath propagation.

The signal of binary data after phase-shift keying (PSK) or quadrature amplitude modulation (QAM) and serial/parallel transformation is $X(k)$, and OFDM modulation is carried out after the insertion of the pilot, that is, inverse fast Fourier transform (IFFT) operation; then, the signal in the time domain is $x(n)$

$$X(n) = \mathrm{IFFT}\{X(K)\} = \sum_{K=0}^{N-1} X(K)\mathrm{e}^{-j2\pi nk/N}, \quad 0 \le n \le N-1 \tag{10}$$

After the protection interval GI is inserted, the signal can be expressed as

$$x_g(n) = \begin{cases} x(N+n), n = -N_g, & -N_g+1, \ldots, -1 \\ x(n), & n = 0,\ 1,\ 2,\ \ldots,\ N-1 \end{cases} \tag{11}$$

In this, $N_g$ is the length of GI. The signal is filtered into the multipath fading channel after the baseband shaping and filtering, and there is additive noise interference. The signal $y_g(n)$ of the receiver can be expressed as

$$y_g(n) = x_g(n) * h(n) + w(n) \tag{12}$$

Among them, $h(n)$ is the impulse response of the channel, and $w(n)$ is additive white Gaussian noise. The signal obtained after removing the protection interval is expressed as $y(n)$, OFDM demodulation and FFT operation, and the obtained frequency domain signal can be expressed as $Y(k)$

$$Y(k) = \mathrm{FFT}\{y(n)\}\mathrm{e}^{-\frac{j2\pi kn}{N}}, \quad k = 0,\ 1,\ \ldots,\ N-1 \tag{13}$$

If the protection interval length is longer than the channel maximum multipath delay length, the intersymbol interference of OFDM cannot be considered. $Y(k)$ can be expressed as

$$Y(k) = X(k) * H(k) + W(k), \quad k = 0,\ 1,\ \ldots,\ N-1 \tag{14}$$

$H(k)$ is the corresponding frequency response of the channel impulse response function on the K subcarrier, and $W(k)$ is the Fourier transform of additive white Gaussian noise $\omega(n)$. The principle of channel estimation based on the pilot is to estimate the channel by the partial sequence in the transmitter and the receiver, and then the channel information of all subcarriers $H(k)$ is predicted by this part of the channel information through interpolation or the correlation of the time–frequency domain.

The commonly used channel estimation algorithms are the least square (LS) algorithm and the linear minimum mean square error (LMMSE). The complexity of LS algorithm is relatively low under a certain bit error rate. Its expression is

$$\hat{H}_{\text{ls}} = X^{-1}Y = \left[\frac{Y(0)}{X(0)}\frac{Y(1)}{X(1)}\cdots\frac{Y(N-1)}{X(N-1)}\right] \tag{15}$$

For the construction of the simulation system, the length of the cyclic prefix is determined directly by the size of the bandwidth, the bit rate and the maximum delay extension. The length of the general cyclic prefix is 2–4 times of the maximum delay extension, and the cyclic prefix is too long to increase the loss of signal-to-noise ratio (SNR). The ring prefix is too short and cannot effectively resist the intersymbol interference (ISI) caused by multipath. After determining the length of the cyclic prefix, the OFDM symbol length is generally 5 times the length of the cyclic prefix in the actual application, so that the loss of the SNR caused by the insertion of the cyclic prefix is only about 1 dB.

After determining the length of the OFDM symbol and the length of the prefix, the effective data part length obtained by the length of the OFDM symbol subtracted from the length of the cyclic prefix, its reciprocal is the subcarrier interval. And then the number of subcarriers is obtained by dividing the total bandwidth by the subcarrier interval, and the required bit rate is divided by the OFDM symbol rate, which is the number of bits to be transmitted for each OFDM symbol. Finally, the modulation mode of the subcarrier is determined according to the number of subcarriers and the number of bits required for each OFDM symbol [21].

## 4 Simulation and Analysis

### 4.1 Error Rate Analysis

Compared with terrestrial wireless communication, the environment of maritime wireless communication is different. The main factors that influence the maritime channel include multipath fading effect and Doppler effect. Communications at sea usually work in a complex and changeable channel environment. When the radio wave

is transmitted over the sea, the transmission path is composed of direct wave propagating in free space, the reflected waves reflected on the sea surface, scattered diffraction waves formed by interference from other obstacles and complex waves superposed and overlapped by multiple waveforms, so that time delay and power loss will appear when reaching the mobile terminal. Finally, the signals at the receiver side need to be linearly combined, and the total signal is actually the convolution of signal and channel impact response. The frequency response of the frequency domain has different response patterns at different frequencies, so there will be deeper frequency-selective fading.

Figure 5 depicts the performance curve of the signal under white Gaussian noise, Rayleigh channel and Rician channel. The signal is modulated by QPSK. The Rayleigh channel adopts the Jakes model, the scattering signal component of the multipath Rician channel is modeled by Jakes, and the fading factor $K$ is 10 dB. As we can see from Fig. 5, the bit error rate and symbol error rate of QPSK after the Rayleigh fading channel and the Rician fading channel are much higher than the bit error rate and symbol error rate under the additive white Gaussian noise (AGWN) channel. The bit error rate and symbol error rate of the signal passing through the Rician channel are less than that of the Rayleigh channel. In the rough channel, the channel model is similar to the Rician channel model.

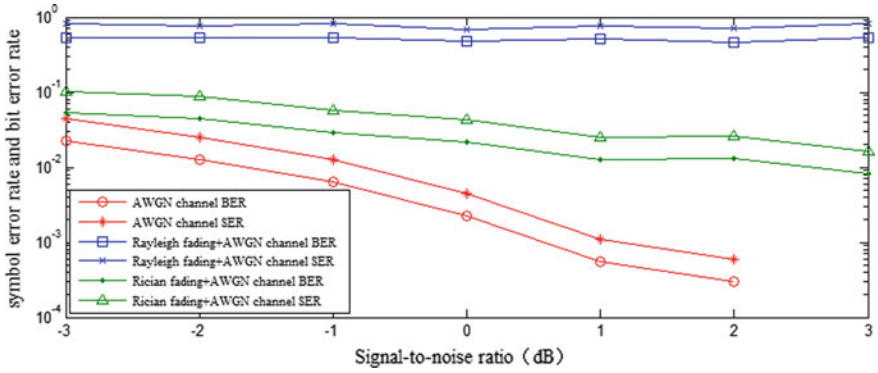



**Fig. 5.** Performance simulation of different fading channels

## 4.2 Comparison of OFDM Performance Under Different Modulation Modes

This part describes the simulation of OFDM transmission on the sea. First, the related parameters are presented for further analysis, as shown in Table 1.

Figure 6 reflects the simulation of OFDM on the Rician channel, through different modulation modes: QPSK and 16-QAM. For computer simulation, the number of subcarriers of an OFDM symbol is 52 sampling points, and the cyclic prefix length is 16 sampling points. The scattering signal component of multipath Rician channel is

modeled by Jakes, and the fading factor K is set to 10 dB. It can be seen from the graph that OFDM has different modulation modes under the Rician channel, and the bit error rate of single channel and multipath channel will also be different. When the SNR of the system is 25 dB, the bit error rate of QPSK is lower than that of QAM. When the channel is a multipath channel, the bit error rate will be larger than that of single-path channel. The error code performance of QPSK is obviously better than that of QAM. Because of the same information rate, the system bandwidth occupied by QPSK is 2 times that of 16-QAM. In this respect, the good bit error performance of QPSK is at the expense of system bandwidth. Therefore, when designing the system, we should take the best balance between the frequency utilization rate and the bit error rate as the principle.

**Table 1.** OFDM system parameters

| Name | Parameter |
|---|---|
| Bit rate | 20 Mbps |
| FFT Sample point number | 64 |
| Subcarrier number | 64 |
| Subcarrier modulation | 16-QAM/QPSK |
| Cyclic prefix length | 16 |
| Maximum Doppler shift | 100 |
| Channel | Rician channel |
| OFDM symbol interval | 4 μs |
| Sampling interval | 50 ns |

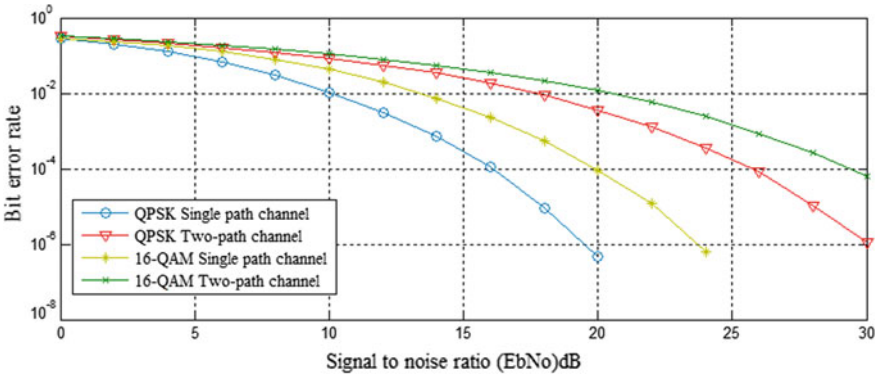



**Fig. 6.** OFDM under different modulation modes in Rician channel

### 4.3 Comparison of OFDM Performance Under Different Channel Estimation Methods

From Fig. 7, we can see that the performance of channel estimation using training symbols is less than that of ideal channel estimation. The bit error rate of OFDM system under ideal channel is less than that of training symbol. When bit error ratio (BER) is $10^{-2}$, the performance of QPSK and 16-QAM training symbol channel estimation is about 3 dB compared with the performance loss under ideal channel estimation. This is because in this frame, the channel is not fixed, and when the channel estimation is used by the training symbol, the data symbol is compensated according to the value of the channel estimated by the training symbol. Therefore, the error is inevitably produced.

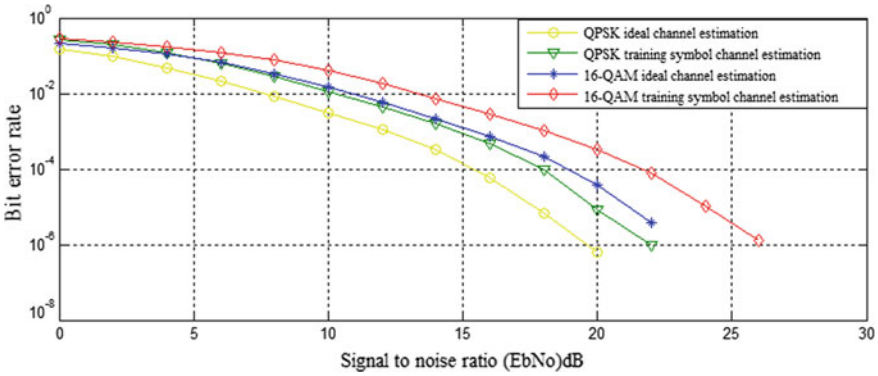



**Fig. 7.** OFDM under different channel estimation methods in Rician channel

## 5 Conclusions

This paper is mainly about channel modeling and modeling simulation for random time-varying rough sea communication environment. The performance of OFDM system under the rough sea surface is studied. Under the rough sea surface, the channel model can be simulated as the Rician channel model. Then, we further study the role of OFDM in maritime communications by modeling OFDM on Rician channel.

In the future, we will further study the role of OFDM in the rough sea surface communication. In subsequent research, some improved algorithms can be considered for better communication between OFDM in the rough sea surface. At the same time, we will also carry out the research of peak-to-average power ratio suppression algorithm in OFDM on meta-surface ocean.

**Acknowledgments.** This article is supported by the National Natural Science Foundation of China (No. 61661018), the Zhejiang Provincial Natural Science Foundation of China (Grant No. LY15F010003), the International Scientific and Technological Cooperation Project (No. 2015DFR10510) of National Science and Technology Ministry and Hainan Provincial Key R. & D. Projects of China (No. ZDYF2016010 and ZDYF2018012). Hui Li is the corresponding author of this article.

## References

1. Zhen HL. Design and simulation of OFDM communication system. Northwestern Polytechnical University; 2003.
2. Reddy SB, Yucek T, Arslan H. An efficient blind modulation detection algorithm for adaptive OFDM systems. In: IEEE vehicular technology conference, vol. 1; 2003. p. 1895–9.
3. Yucek T, Arslan H. A novel sub-optimum maximum-likelihood modulation classification algorithm for adaptive OFDM systems. In: Wireless communications and networking conference, vol. 3; 2004. p. 739–44.
4. Wang B, Ge LD. A novel algorithm for identification of OFDM Signal. Wireless Commun Networking Mobile Comput. 2005;1:261–4.
5. Cao P, Peng H, Dong YK, Wang B. An OFDM blind detection and parameter estimation algorithm based on cyclic prefix. J Univ Inform Eng. 2010;11(2):196–200.
6. Yuan ZL. Research on OFDM signal detection and parameter estimation in Rayleigh fading channel. Southwest Jiao Tong University; 2011.
7. Liu QP, Dong JT. Application of OFDM technology in 4G mobile communication system. Electronic Test. 2014;05:102–4.
8. Liu GS. Application of 4G communication network based on OFDM Technology. Netw Security Technol Appl. 2016;07:85–7.
9. Arraño HF, Azurdia-Meza CA. OFDM: today and in the future of next generation wireless communications. In: 35th IEEE convencion de Estudiantes de Centroamerica y Panama; 2016. p. 12–6.
10. Koffman I, Roman V. Broadband wireless access solutions based on OFDM access in IEEE 802.16. IEEE Commun Mag. 2013;40(4):96–103.
11. Dan LL, Xiao Y, Lei X. Signal optimization design in OFDM communication system. National Defense Industry Press; 2015.
12. Kao YA. The influence of residual timing errors on channel estimation in OFDM system. In: International conference on computer, electronics and communication engineering; 2017:5–6.
13. Zhu GN. Research on future air mobile communication air interface and OFDM Technology. Dalian Maritime University; 2012.
14. Wang J. Application and analysis of OFDM technology in maritime wireless communication system. Wirel Commun Technol. 2011;20(02):36–9.
15. Zhao HL, Tong LL. Transmission characteristics of microwave channel at sea. J Harbin Univ Sci Technol. 2006;11(2):16–20.
16. Wang X. OFDM signal detection and modulation recognition. University of Science Technology China; 2009.
17. Huang F. Research on characteristics of maritime wireless radio propagation and channel modeling. Hainan University; 2015.
18. Liao XF. The study of maritime wireless transmission based on SC-FDE. Hainan University; 2014.

19. Meng ZQ. Research on transmission rate of large scale MIMO system under Rician channel. Dalian Maritime University; 2017.
20. Heiskala J, Terry J. OFDM wireless local area networks. Electronic Industry Press; 2003.
21. Wu K. Simulation research and implementation of wireless fading channel simulation. University of Electronic Science and Technology; 2014.

# Parallel Implementation of SUMPLE Algorithm in Large-Scale Antenna Array

Yan Di[1(✉)], Shuai Weiyi[1], Liu Peijie[1], Sun Ke[2], and Li Xiaoyu[2]

[1] Department of Electronics and Optical Engineering, Space Engineering University, Beijing 101416, China
yandimail@126.com
[2] Beijing Aerospace Control Center, Beijing 102206, China

**Abstract.** Digital signal processing on high-performance computing platform based on general-purpose processor is one of the development directions of software-defined radio, which provides a new idea for the signal combining of large-scale antenna array. In this paper, two parallel implementations of SUMPLE combined algorithm are proposed, which are multi-antenna-group method (MAGM) based on task decomposition principle and multi-signal-block method (MSBM) based on data decomposition principle. The processing capability and real-time performance of the two methods are analyzed in detail. Results show that MAGM is not suitable for the high-performance computing platform because of large and frequent data transmissions from central processing unit (CPU) to graphics processing unit (GPU). The traditional iterative method of SUMPLE cannot be applied to MSBM due to the dependencies among signal blocks.

**Keywords:** Antenna array · High-performance computing platform · Task decomposition · Data decomposition

## 1 Introduction

Large-scale antenna arrays are being used for radio astronomy, positioning as well as communication applications. The size of arrays being built today is approaching hundreds of antennas, such as Allen Telescope Array (ATA) and Atacama Large Millimeter Array (ALMA) [1, 2]. SUMPLE algorithm is one of the most important algorithms for signal combining [3–5]. The beamformer needs to receive all the signals of the antenna array and calculate weights of each channel at the same time. When the antenna array is large enough, two issues will be more prominent [6]: First, it is difficult to transmit all the signals to the beamformer simultaneously. Second, it is difficult to calculate the weights simultaneously. Therefore, the application of SUMPLE algorithm in large-scale antenna array is limited.

Signal processing on high-performance computing platforms has been applied [7], which provides a new idea for the beamforming of the large-scale antenna array. The typical high-performance computing platform is central processing unit (CPU) and graphics processing unit (GPU) heterogeneous system [8]. In this paper, the beamformer implemented on CPU-GPU heterogeneous system is called soft beamformer


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_53

(Softformer). In order to make full use of the parallel processing resources of GPU and improve the Softformer's combined capability, we need to study the parallel implementation of SUMPLE algorithm, which is also the main content of this paper.

In Sect. 2, we describe the basic principle of SUMPLE algorithm. In Sect. 3, we describe two kinds of parallel implementations, which are multi-antenna-group method (MAGM) and multi-signal-block method (MSBM). Then, we analyze the real-time performance and combined capability. The conclusions are discussed in Sect. 4.

## 2 SUMPLE Algorithm

Phase correction is the key technology of antenna array, that is what we call signal combined algorithm, mainly including SIMPLE, SUMPLE, Eigen, and so on. SUMPLE is one of the most excellent algorithms. Consider signal received by the antenna can be expressed as:

$$S_{i,k} = s_{i,k} + n_{i,k}, \quad i = 1,\ 2,\ \ldots,\ N_{\text{ant}} \tag{1}$$

where $N_{\text{ant}}$ is the number of antennas, subscript $k$ is the time serial number, subscript $i$ is the antenna serial number, $S_{i,k}$ is the noisy signal, $s_{i,k}$ is the non-noise signal, and $n_{i,k}$ is the noise. Weight of the signal can be expressed as:

$$W_{i,K} = w_{i,K} + \eta_{i,K}, \quad i = 1,\ 2,\ \ldots,\ N_{\text{ant}} \tag{2}$$

where subscript $K$ is the serial number of integral period, $w_{i,K}$ is the ideal weight, and $\eta_{i,K}$ is the weight error caused by noise.

SUMPLE correlates the received signals of all antennas with the weighted sum of all the other antennas to obtain the complex weights and then weights all the signals. Weights are calculated as follows:

$$W_{i,K+1} = R_{K+1} W_{i,K} \left\{ \frac{1}{L} \sum_{t=KL}^{(K+1)L-1} \left[ \left( S_{i,t} W_{i,K}^{*} \right) \sum_{j=0, j\neq i}^{N_{\text{ant}}-1} \left( S_{j,t}^{*} W_{j,K} \right) \right] \right\} \tag{3}$$

where $L$ is the number of sampling points in an integral period, $W_{i,K}^{*}$ is the complex conjugate of $W_{i,K}$, and $R_{K+1}$ is the normalized amplitude factor and meets the following condition:

$$\sum_{i=1}^{N} W_{i,K+1} W_{i,K+1}^{*} = N_{\text{ant}} \tag{4}$$

In (3), weights of the $K + 1$th integral period is derived from weights of $K$th integral period. This makes weights of different integral period which must be calculated according to the order, constraining the parallelism of SUMPLE.

## 3 Parallel Implementation of SUMPLE

In this section, the parallel implementations are classified as MAGM and MSBM according to the parallel decomposition principle. This section analyzes the feasibility of these two methods and uses the following parameters: The number of antennas is 500, the sampling rate of the received signal is 100 MHz and 8 bit, the length of the integral period is 100,000, and the network bandwidth of Softformer is 20 Gbps.

### 3.1 MAGM

All the antennas are divided into several antenna groups, and each antenna group uses a Softformer. Softformer can be classified as level-1 Softformer, level-2 Softformer, and so on. Level-1 Softformer combines the antenna signals in the group. Level-2 Softformer combines the signals which are the combined results of Level-1 Softformer. The basic principle of MAGM is shown in Fig. 1.

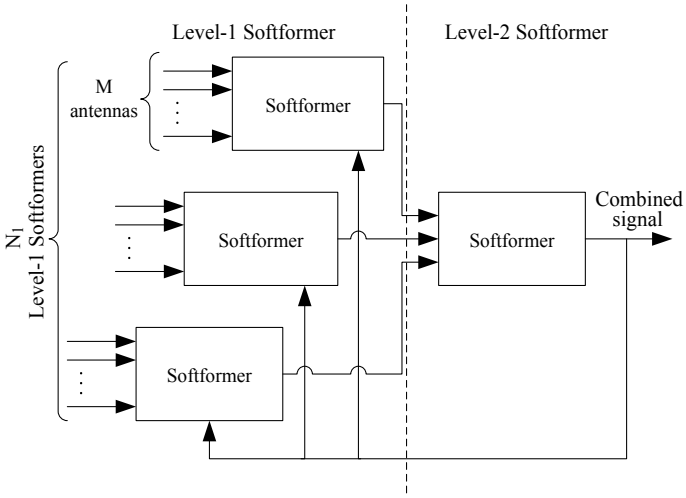



**Fig. 1.** Basic principle of MAGM

The detailed steps of MAGM are as follows:

**Step 1**: Initialize weights of each Softformer.
**Step 2**: Each level-1 Softformer performs combined according to weights of the previous integral period.
**Step 3**: Each level-2 Softformer performs combined according to weights of the previous integral period.
**Step 4**: The total combined signal is fed back to Softformers at all levels.
**Step 5**: Each level-1 Softformer performs weight calculation and combined according to the total combined signal.

**Step 6**: Each level-2 Softformer performs weight calculation and combined according to the total combined signal.

Repeat steps 2–6 to continue signal combining. The computational process of the CPU-GPU heterogeneous system is shown in Fig. 2.

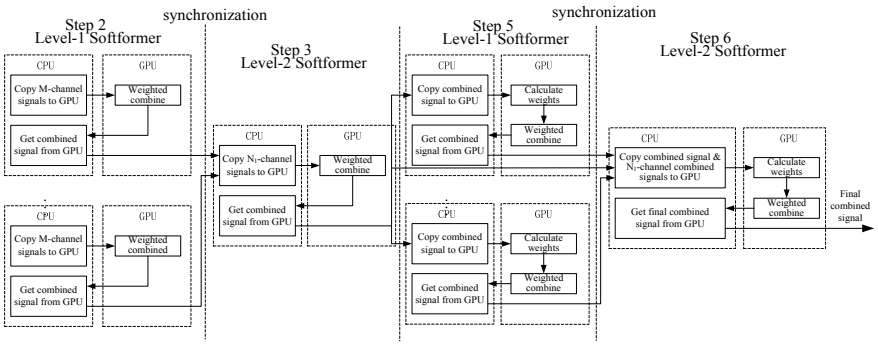



**Fig. 2.** MAGM's computational process on the CPU-GPU heterogeneous system

### 3.2 Real-Time Performance and Combined Capability of MAGM

In Fig. 2, if each Softformer is capable of processing $M$ channels of signal and the number of level-1 Softformers is $N_1$, then up to $M \times N_1$ channels of signal can be processed. If 1 level-3 Softformer is added and the number of level-2 Softformers is $N_2$, then up to $M^2 \times N_2$ channels of signal can be processed. For convenience, only level-1 and level-2 Softformer are analyzed in this paper. Take the hypotheses in this section as an example, $M = 25$, and the combined task of 500 antennas can be achieved by using 20 level-1 Softformers and 1 level-2 Softformer.

In order to ensure real-time processing, the following condition must be met:

$$T_{\text{total}} < \frac{L}{f_s} \tag{5}$$

where $L$ is the length of integral period, $f_s$ is the sampling frequency, and $T_{\text{total}}$ is the total processing time from step 2–6:

$$T_{\text{total}} = \max\{T_2(1), \ldots, T_2(N_1)\} + T_3 + \max\{T_5(1), \ldots, T_5(N_1)\} + T_6 \tag{6}$$

where subscripts of $T$ indicate the step numbers, and for example, $T_2(N_1)$ is the processing time of the $N_1$th Softformer in step 2. For convenience, we assume that:

$$T_2(1) = \cdots = T_2(N_1), T_5(1) = \cdots = T_5(N_1) \tag{7}$$

So, (6) can be simplified to:

$$T_{\text{total}} = T_2 + T_3 + T_5 + T_6 \tag{8}$$

Each $T$ consists of two parts: the data copy time from CPU to GPU memory which is expressed as $T_{\text{tran}}$ and GPU computing time which is expressed as $T_{\text{pro}}$. Then, $T_{\text{total}}$ can be expressed as:

$$T = T_{\text{tran}} + T_{\text{pro}}, T_{\text{total}} = T_{\text{total, tran}} + T_{\text{total, pro}} \tag{9}$$

From steps 2 to 6, the total amount of data copied from CPU to GPU is $ML + N_1L + L + (N_1 + 1)L = 6.7 \times 10^6$. So in one second, a total of 6.7 GB data are copied from CPU to GPU. For PCI-e 3.0 with 16 GB/s unidirectional bandwidth, it also takes 0.42 s. In fact, the 6.7 GB data are divided into 1000 times, and each time is divided into five steps. The time spent on frequent copies of small data will increase exponentially or even exceed 1 s.

We can see that, in MAGM, there are lots of data transmission from CPU to GPU, and these data transfers have obvious logical relations with kernel functions. Obviously, MAGM is not suitable for CPU-GPU heterogeneous systems.

### 3.3 MSBM

In MSBM, each Softformer combines all the signals in one integral period, and different Softformers process signals in different integral period. The basic principle of MSBM is shown in Fig. 3, $A(i,j)$ represents the $i$th integral period of the $j$th antenna. Because each Softformer in the MSBM receives signals from all antennas, there is no need for feedback aiding of the final combined signals. The detailed steps of MSBM are as follows:

**Step 1**: Initialize weights of each Softformer.
**Step 2**: The Softformer receives all the signals in one integral period.
**Step 3**: The Softformer performs weight calculation and signal combining.

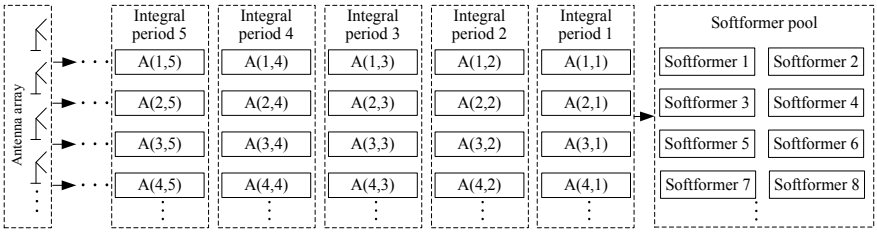



**Fig. 3.** Basic principle of MSBM

### 3.4 Real-Time Performance and Combined Capability of MSBM

The processing time of an integral period consists of two parts. One is the network transmission time, the time it takes to transmit all the signals in one integral period from AD sampling units to the Softformer which is expressed as $T_{\text{tran}}$, and the other is the computation time which is expressed as $T_{\text{pro}}$.

$$T_{\text{total}} = T_{\text{tran}} + T_{\text{pro}} \tag{10}$$

For any $T_{\text{total}}$, when the number of Softformers which is expressed as $N_0$ satisfies (11), the real-time performance can be guaranteed.

$$N_0 \geq \left\lceil \frac{T_{\text{total}}}{LT_{\text{s}}} \right\rceil \tag{11}$$

where $T_{\text{s}}$ is the sampling period. The physical meaning of (11) is shown in Fig. 4. In Fig. 4, two Softformers take turns to receive signals and perform processing to ensure real-time performance. The greatest advantage is that we can keep the real-time performance by increasing the number of Softformers, while reducing the bandwidth and computing speed of a single Softformer.

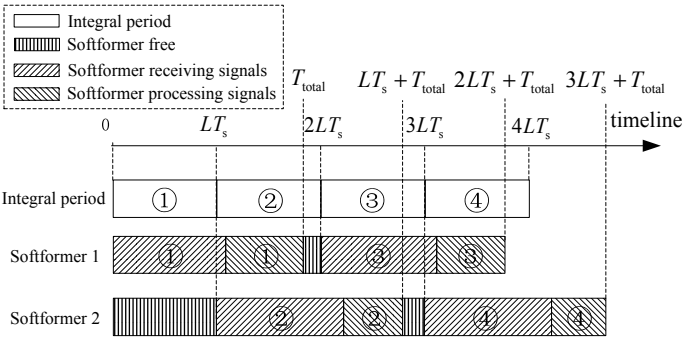



**Fig. 4.** Physical meaning of formula 11

Take the hypotheses in this section as example, in one integral period, the total amount of signals from 500 antennas is $4 \times 10^8$ bit, so $T_{\text{tran}}$ is 20 ms. The transmission time from CPU to GPU memory is about 3.1 ms which can be hidden in the network transmission. The combining of 500 antennas probably requires $2.5 \times 10^{10}$ times of multiplication and addition operations. If the Softformer is based on the NVIDA TESLA K80 and the floating-point computing power is fully utilized, $T_{\text{pro}}$ will be 3 ms. Therefore, tens of ms is required for one integral period, so tens of Softformers can achieve real-time combining.

## 4 Conclusions

In this paper, two kinds of parallel implementations of SUMPLE algorithm in large-scale antenna array based on high-performance computing platform were studied. In MSBM, each Softformer combines all the signals in one integral period, and different Softformers process signals in different integral period. By increasing the number of Softformers, MSBM's combined capability can be enhanced. At the same time, the network bandwidth pressure and calculation pressure of a single Softformer are reduced. This paper provides some guidance for the parallel implementation of SUMPLE algorithm on CPU-GPU heterogeneous system, and future test validation will be carried out on a specific CPU-GPU platform.

## References

1. Harp GR. Allen telescope array digital processing requirements driven by radio frequency interference concerns. Radio Sci. 2005;40:1–9.
2. Dickman RL, Millimeter radio astronomy and the genesis of ALMA. In: 2015 IEEE MTT-S international microwave symposium, Phoenix, AZ; 2015. p. 1–4.
3. Shi X, Wang Y, Dang H. Study on array signal processing for deep space communications. In: 2008 second international symposium on intelligent information technology application, Shanghai; 2008, p. 611–5.
4. Foutz J, Spanias A, Bellofiore S, Balanis CA. Adaptive eigen-projection beamforming algorithms for 1D and 2D antenna arrays. IEEE Antennas Wirel Propag Lett. 2003;2:62–5.
5. Rogstad DH. The SUMPLE algorithm for aligning arrays of receiving radio antennas: coherence achieved with less hardware and lower combining loss, IPN Progress Report; 2005. p. 42–162, August 2005[Online]. Available: https://tda.jpl.nasa.gov/progress_report/42-162/162B.pdf.
6. Barott WC et al. Real-time beamforming using high-speed FPGAs at the allen telescope array. In: Radio science, vol. 46; 2011. p. 1–17.
7. Islam MS, Kim CH, Kim JM. Computationally efficient implementation of a hamming code decoder using graphics processing unit. J Commun Netw. 2015;17:198–202.
8. Yang Y, Xiang P, Mantor M, Zhou H. CPU-assisted GPGPU on fused CPU-GPU architectures. In: IEEE international symposium on high-performance comp architecture, New Orleans, LA; 2012. p. 1–12.

# An Improved Square Timing Error Detection Algorithm

Pengfei Guo(✉), Celun Liu, Mucheng Li, Bizheng Liang, and Jianguo Li

School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
{guopengfei,liucelun,limucheng,liangbizheng,jianguoli}@bit.edu.cn

**Abstract.** Timing recovery is one of the most important issues that need to be solved for a receiver. The key of timing recovery is the extraction of timing errors. Currently, the timing recovery in the satellite communication systems cannot meet the requirement of the real-time passive intermodulation (PIM) interference suppression. It calls for the study on improved algorithm that can be applied in the satellite communication system. And the improved algorithm should have small computation and a high running speed. This paper first introduces the overall structure of the traditional square timing recovery loop, and then proposes an improved square timing error detection algorithm. By using the Goertezl algorithm, this improved algorithm optimizes the step–extracting spectral components at the symbol rate 1/T. Finally, this paper simulates and analyzes the improved algorithm, as well as parameters' influence on the performance of the algorithm. It is proved that the improved square timing error detection algorithm can greatly reduce the amount of computation under the premise of guaranteeing performance.

**Keywords:** Timing recovery · Square timing error detection algorithm · Reduce computation

## 1 Introduction

Timing recovery is essential for the receiver in digital communication system. It means to extract timing information from the received signal to make the receiver sample at the best position of the symbol [1], thereby reducing the bit error rate. In the all-digital receiver, the sampling clock with high stability and high accuracy is independent of the signal clock [2]. When such a sampling clock is used to sample the received signal directly, there will be some timing error because the sampling clock frequency is not synchronized with the symbol rate of the received symbol. And the timing recovery algorithm is to eliminate this timing error.

In general, the timing recovery loop first needs to extract the timing information from the received signal [3], estimate the timing error in combination


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_54

with the local sampling clock, and then use the estimated value of the timing error to control the interpolator in real time, including updating the filter's coefficient and determining the specific interpolation time. The output sequence of the interpolator is exactly synchronized with the received signal. That is the process of timing recovery [4]. The expected advantages of the timing recovery algorithm are as follows: fast synchronization, low computation, and easy for digital realization [5]. The current timing recovery algorithm applied in the receiver of satellite communication system has a great number of computation to do, thus requiring much time. It fails to satisfy the needs of the real-time PIM interference suppression. Based on the actual demand, this paper proposes an improved square timing error detection algorithm, simulates, and analyzes the algorithm. The algorithm can largely reduce computation under the premise of good performance and is advantageous for digital realization.

## 2 Square Timing Recovery Loop

The square timing recovery loop is an efficient, stable, and non-data-aided (NDA) timing recovery loop [6]. This loop does not require feedback of the signal and directly extracts the timing information from the received signal, so that the timing error can be quickly captured. And square timing error detection algorithm is quite independent of the carrier. The timing recovery would not be influenced by the frequency offset and phase offset of the carrier, so it could be prior to carrier synchronization. Thus, timing recovery has been widely used in all-digital receivers [7]. The square timing recovery loop mainly consists of match filter, sampling unit, squaring timing error detector, timing control unit, and interpolating filter. Passing through the match filter and the sampling unit, the timing error in the received signal is extracted by the squaring timing error detector. Together with the sampling sequence, the timing error decides the best position for sampling in the interpolation filter to realize timing recovery [8].

Figure 1 shows the whole structure of the square timing recovery loop.

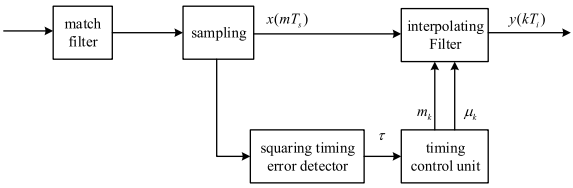



**Fig. 1.** The square timing recovery loop

In this loop, the squaring timing error detector plays a key role. And it applies square timing error detection algorithm, whose principle is that the received baseband signal can generate baseband sampling sequence after passing match filter and sampling unit, then by modular squaring the sampling sequence, the

frequency spectrum must include the spectral component at the symbol rate 1/T. The timing error of the sampling point is expressed in the frequency domain as phase rotation of the spectral component. Martin Oerder proved that the normalized phase of that spectral component is an unbiased estimate of the normalized timing error in his paper [4]. The following part explains the specific process of the loop. The signal received can be written as:

$$r(t) = \sum_{n=-\infty}^{\infty} a_n g_T(t - nT - \varepsilon(t)T) + n(t) = u(t) + n(t) \tag{1}$$

where $a_n$ is the transmitted symbol sequence, $g_T(t)$ is the transmission signal pulse, $g_T(t)$ is the duration of the symbol, $n(t)$ is the channel noise which is assumed to be white and Gaussian distribution with power density $N_0$, and $\varepsilon(t)$ is an unknown, slowly varying timing error. Timing recovery means, to some extent, an accurate estimate of $\varepsilon(t)$. Since $\varepsilon(t)$ varies very slowly, we can assume that $\varepsilon(t)$ will remain constant for a period of time, so in the digital realization, we can group the data and then process the received signals section by section. The analysis followed assumes that the received signals are divided into groups and every group contains L symbols. After passing the matched filter (impulse response $g_R(t)$), $r(t)$ is sampled at the rate $T/N$, then we have samples

$$\tilde{r}_k = \tilde{r}(kT/N) = \sum_{n=-\infty}^{\infty} a_n g(\frac{kT}{N} - nT - \varepsilon T) + \tilde{n}(\frac{kT}{N}) \tag{2}$$

In the equation, $g(t) = g_T(t) * g_R(t)$.

The sequence is then sent to the square timing error detector, which applies square timing error detection algorithm. Figure 2 shows how the algorithm works.

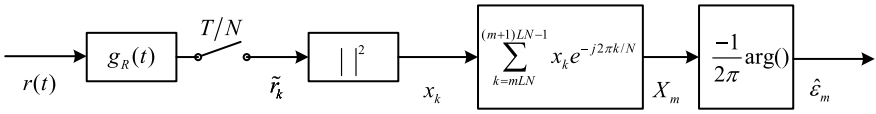



**Fig. 2.** Square timing error detection algorithm

First, the sampling sequence is modular squared, then we have

$$x_k = \left| \sum_{n=-\infty}^{\infty} a_n g(\frac{kT}{N} - nT - \varepsilon T) + n(\frac{kT}{N}) \right|^2 \tag{3}$$

In order to extract the spectral components at the symbol rate 1/T, the DFT of $x_k$ should be computed to get the complex Fourier coefficient at the symbol rate. With each group's length as L, and oversampling rate as N, the

L-th sample in the LN-point DFT of $x_k$ represents the symbol rate. Then, the spectral component can be expressed as

$$X_m = \sum_{k=mLN}^{(m+1)LN-1} x_k e^{-j2\pi k/N} \tag{4}$$

To calculate the unbiased estimate of the normalized timing offset is to figure out the normalized phase of $X_m$, namely

$$\hat{\varepsilon}_m = -1/2\pi \arg(X_m) \tag{5}$$

The estimate of timing error $\hat{\varepsilon}_m$, along with the sampling sequence, will lead to timing recovery after passing the timing control unit and interpolation filter [9].

## 3 Square Timing Error Detection Algorithm

As mentioned above, the normalized phase of the spectral component at the symbol rate is an unbiased estimate of the normalized timing error. It is the L-th sample in the LN-point DFT of $x_k$ that needs to be calculated. If FFT algorithm was used here like the traditional way, a lot of computation results were invalid. So, this paper proposes an improved square timing error detection algorithm which uses the Goertzel algorithm to detect the information at a certain frequency in the signal. That how Goertzel algorithm improves DFT will be analyzed in detail in this part [10]. Suppose that $x_n$ is a sample sequence with a length of N, $n \in [0, N-1]$, and n is an integer. Then, the DFT of $x[n]$ is defined as

$$X_k = \sum_{n=0}^{N-1} x_n w_N^{kn}, \mathrm{k} = 0, 1, 2, \cdots, N-1 \tag{6}$$

wherein $W_N = e^{\frac{-j2\pi}{N}}$. The DFT summation can be written in another way as

$$X_k = \sum_{n=0}^{N-1} \left(w^{-k}\right)^{N-n} \cdot x_n \tag{7}$$

It can be found that the powers of W in the DFT summation can be treated as the power terms in a polynomial, with $x[n]$ as the multiplier coefficients. Then, this polynomial can be evaluated in the nested form:

$$X_k = (\cdots(((W^{-k} \cdot x_0 + x_1) \cdot W^{-k} + x_2) \cdot W^{-k} + x_3) \cdots + x_{N-1}) \cdot W^{-k} \tag{8}$$

As the nested form is evaluated term by term, it produces a sequence of intermediate results y:

$$
\begin{aligned}
&y_{-1} = 0 \\
&y_0 = W^{-k} y_{-1} + x_0 \\
&y_1 = W^{-k} y_0 + x_1 \\
&\cdots \\
&y_{N-1} = W^{-k} y_{N-2} + x_{N-1} \\
&y_N = W^{-k} y_{N-1}
\end{aligned}
$$

The final result $y_N$ equals the desired DFT sum. The following iteration is a summary of the above process:

$$
\begin{aligned}
&y_n = W^{-k} y_{n-1} + x_n, n = 0, ..., N-1 \\
&y_N = W^{-k} y_{N-1}
\end{aligned}
$$

The iteration equation in the discrete domain is equivalent to the following transfer function:

$$\frac{Y}{X} = \frac{1}{1 - W^{-k} z^{-1}} \tag{9}$$

So the transfer function of the Goertezl filter is:

$$H_k[z] = \frac{1}{1 - W_N^{-k} z^{-1}} = \frac{1 - W_N^k z^{-1}}{1 - 2\cos[2\pi k/N] z^{-1} + z^{-2}} \tag{10}$$

Multiply $W_N^{-k}$ will lead to the final result. The block diagram of the Goertezl algorithm is as Fig. 3 The realization of the filter can be expressed as

$$s_k[n] = x[n] + 2\cos[2\pi k/N] s_k[n-1] - s_k[n-2] \tag{11}$$

$$X[k] = W_N^{-k} \cdot y_k[N-1] = W_N^{-k} \cdot (s_k[N-1] - W_N^k s_k[N-2]) \tag{12}$$

And $s_k[-1] = s_k[-2] = 0$. The filter is composed of two parts: the formula (11) means the first part will be recursively calculated for N times, and the formula (12) means the latter part only calculates once in the end. $X[k]$ will be figured out.

Using Goertezl algorithm, the improved square timing error detection algorithm is advantageous in the following aspects:

1. Reducing computation. It uses a recursive real arithmetic structure to avoid intricate complex operations;
2. Saving storage space. The required storage space is small, and only the corresponding coefficients and some intermediate results are stored;
3. Improving real-time performance. The Goertezl algorithm can immediately process every sampling result, improving the calculation efficiency while the FFT needs all the values of the data block before processing;
4. There is no limit on the number of sampling points N, which can be arbitrarily selected.

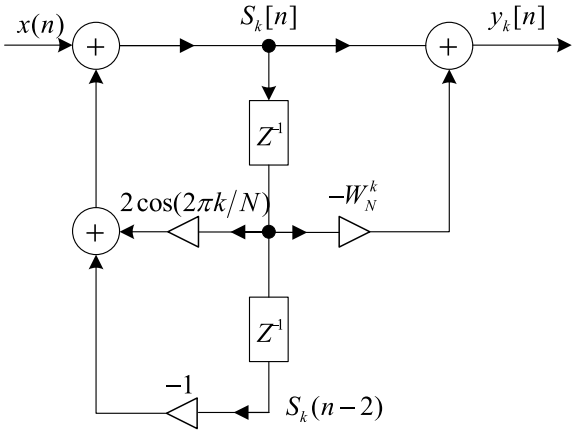



**Fig. 3.** The block diagram of Goertezl algorithm

The key advantage of the improved square timing error detection algorithm lies in the reduction of computation. Here is a detailed analysis of the computation. From the realization of the filter, we can see that for each sample value, the formula (11) calculates once recursively, and calculates for N times all together. And the formula (12) only calculates once in the end after recursive calculation. Every operation of the formula (11) needs two real additions and one real multiplication, and one of the multipliers in real multiplication is constant $2\cos[{}^{2\pi k}/_{N}]$ (this process also saves storage space). Every operation of the formula (12) needs one real addition and two real multiplications. Therefore, the Goertezl algorithm requires a total of 2N+1 real additions and N+2 real multiplications to finish one operation. The computation of Goertezl algorithm is far less than that of FFT since it requires a large amount of unnecessary operations. So, we only compare the computation of Goertezl algorithm with that of DFT. When it comes to DFT, one operation needs real multiplying complex for N times and complex addition for N-1 times, namely 2 N real multiplications and 2N-2 real additions. Table 1 shows the reduction of computation.

**Table 1.** Comparison between the computation of Goertezl algorithm and DFT

| | Real addition | Real multiplication |
|---|---|---|
| Goertezl algorithm | 2N+1 | N+2 |
| DFT | 2N-2 | 2N |

Table 1 shows that Goertezl algorithm reduces real multiplication by almost a half and greatly improved the performance of square timing error detection algorithm.

## 4 Simulation Results and Analysis

First, the BER performance of the improved square timing error detection algorithm is simulated and analyzed. The simulation conditions are set as follows: timing error is 0.25T; every group's length L is 64; all the received symbols are divided into 2000 groups; QPSK modulation; cubic interpolation is used in interpolation filter; oversampling rate is 4. And comparing theoretical BER curve with improved square timing recovery loop BER curve when Eb/No is from −2 to 7 leads to the simulation results seen in Fig. 4.

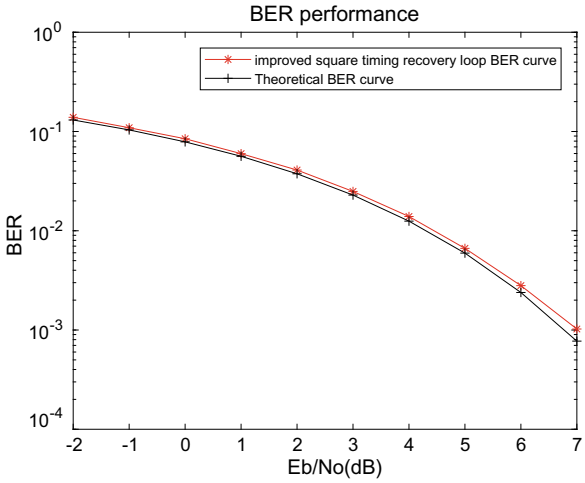



**Fig. 4.** BER performance of the improved square timing recovery loop

It can be drawn from the Fig. 4 that after timing recovery performed by the improved square timing error detection algorithm, the BER performance is extremely close to the theoretical BER performance, proving that improved square timing error detection algorithm can realize timing recovery. Second, L is an important parameter in the improved square timing error detection algorithm. Based on the analysis above, the value of L does not influence the overall computation. Therefore, the following analysis focuses on the impact on the variance of timing estimate caused by different value of L. Suppose that L is 4, 8, 16, 32, 64, 128, 256, and 512 while other conditions remain the same, and explore the connection between L and variance of timing estimate.

Figure 5 shows that as L increases, the variance of timing estimate reduces significantly. When L is big enough, the timing error stemmed from the extracted spectral components at the symbol rate 1/T will be accurate and stable. But when L is too large, too many symbols are needed before processing, resulting in long-time delay. So, L should be in a reasonable range, for example, 64, to meet the requirements of stability and real-time performance.

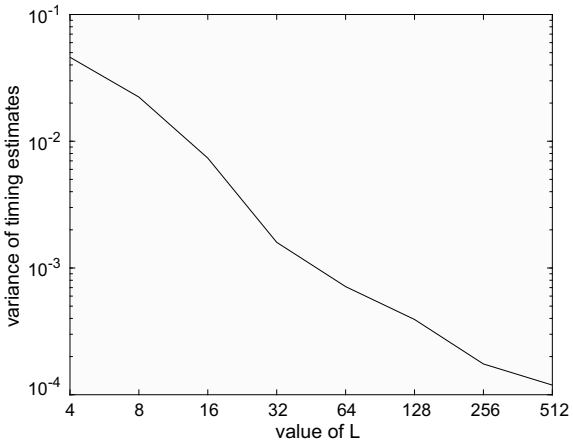



**Fig. 5.** Connection between L and variance of timing estimate

## 5 Conclusion

The timing recovery algorithm adopted by all-digital receivers is expected to be fast and have small computation burden. After analyzing the square timing recovery loop, this paper proposes an improved square timing error detection algorithm. The algorithm can accurately realize timing recovery while greatly reducing computation at the same time. With all those advantages, the improved algorithm can be useful in the real-time PIM interference suppression and meet the needs of satellite communication system.

**Acknowledgment.** This work was supported by the National Natural Science Foundation of China under Grant 61601027.

## References

1. D'Andrea AN, Luise M. Optimization of symbol timing recovery for QAM data demodulators. IEEE Trans Commun. 1996;44:399–406.
2. Xia J, Li Y, Xu T, Chen J, Song Y. Timing recovery algorithm selection for duobinary signal used in next generation 25G-EPON. IEEE Photonics J. 2017;9:1–7.
3. Gardner FM. A BPSK/QPSK timing-error detector for sampled receivers. IEEE Trans Commun. 1986;34:423–9.
4. Oerder M, Meyr H. Digital filter and square timing recovery. IEEE Trans Commun. 1988;36:605–12.
5. Scardoni F, Magarini M, Spalvieri A. Impact of self noise on tracking performance of non-data-aided digital timing. J Lightwave Technol. 2015;33:3755–62.
6. Ayat M, Mirzakuchaki S, Beheshti-Shirazi AA. Design of new digital blind feedforward timing recovery algorithm for linear modulations in digital communication receivers. IET Commun. 2016;10:1736–43.

7. Gardner FM. Interpolation in digital modems I. Fundamentals. IEEE Trans Commun. 1993;41:501–7.
8. Johansson H, Gustafsson O. Linear-phase FIR interpolation, decimation, and mth-band filters utilizing the farrow structure. IEEE T CIRCUIT-I. 2005;52:2197–207.
9. Lamb D, Chamon FO, Nascimento VH. Efficient filtering structure for spline interpolation and decimation. Electron Lett. 2016;52:39–41.
10. Ghouse MA. Modular realizations of the DFT from the Goertzel DFT modules (of smaller dimension). IEEE Trans Signal Process. 1993;41:427.

# Software-Defined Space-Based Integration Network Architecture

Peng Qin[1], Haijiao Liu[1], Xiaonan Zhao[2(✉)], Yingyuan Gao[1], Zhou Lu[1], and Bin Zhou[1]

[1] China Academy of Electronics and Information Technology, Beijing 100041, China
[2] Tianjin Normal University, Tianjin 300387, China
xiaonan5875@163.com

**Abstract.** In the future, information service calls for higher requirements for space-based information network infrastructure. However, there are problems of independent development and redundant construction among existing space-based information systems. To solve the problems, a software-defined space-based integration network architecture is proposed in this article, which could provide information transmission, networked awareness, position navigation and timing (PNT) service on demand.

**Keywords:** Software-defined network · Space-based integration network architecture · System architecture · Technical architecture · Key technologies

## 1 Introduction

Space-based systems are developing in directions such as integration of heterogeneous networks, intelligent and low-granularity resource allocations, as well as real-time and accurate information services. However, existing space-based resources are mainly allocated to multi-satellites and multi-systems, and space-based network in a single type cannot meet the needs of diverse users, e.g., global coverage, high reliability and low delay. Multi-network integration is an important trend of the development of space-based network techniques, but existing networking techniques still lack in complete and unified system architecture and the standards of air interfaces are miscellaneous which hinder the management of multi-type networks and the scheduling of ubiquitous resources and services. Thus, the new architecture of network integration should be built, and a standard network integration system should be proposed. Further, we should break through some enabling techniques such as intelligent operation and maintenance control, resource scheduling, space-time datum and spatial awareness so as to promote full-coverage, time-sensitive and high-precision space-based resource services in the background of ubiquitous network integration.

Facing the needs of global coverage, high reliability and low delay in space-based information services, researches on space-based network techniques and deployments of infrastructures are in full swing at home and abroad. With respect to space-based networking, hybrid and multilayer networking schemes in high, medium, low orbits and near space are surging at abroad. Some projects such as Wideband Global Satellite


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_55

(WGS) system, Other 3 Billion (O3b) system and OneWeb adopt the structure of 'ground-based space network' to realize multilayer and three-dimensional coverage of space, air and ground, and serve various types of users. Europe proposes the Integrated Space Infrastructure for Global Communications (ISICOM) to guarantee global communication and vigorously promotes researches on the integration of satellite communication systems and terrestrial mobile communication networks (e.g., 4G/5G).

China keeps advancing the construction of space-based information system in the past years and has achieved preliminary results. In the field of space networks, 'technical innovation 2030 major project' space-integrated ground network [1–3] deploys global-coverage strategic information infrastructures based on concepts of 'space-based networking, space–ground interconnection and global services.' In the meantime, China also arranges some research projects such as 'key techniques and demonstrations in the future integration networks' and 'fundamental theories and key techniques in space information networks,' and conducts some researches on key techniques, e.g., the network architecture of space information network and multi-protocol optimization and simulation verification.

The development of new technologies such as microelectronics, software-defined network (SDN) [4], network functions visualization (NFV) [5, 6], cloud computing and big data not only contributes to the decoupling of data plane and control plane and the visualization of network functions, but also promotes the level of integration and softwarization for the satellite payloads and enhances the capability of network management control, providing foundation for the evolution of space-based network integration.

## 2 Related Works

### 2.1 State-of-the-Art of Space-Based Network Integration

With the goal of multilayer networking and multi-functional integration, many countries all over the world make relevant developing plans and launch systematical constructions, and some representative systems are as follows.

International Maritime Satellite Organization started up the network construction of Inmarsats in 1979, which has evolved to the fifth generation now. The project builds a space–ground integrated network consisting of eight high-orbiting Inmarsats and 22 ground stations, and realizes the interconnection among Inmarsats, mobile communication networks and terrestrial networks.

O3b is a global satellite Internet service provider, and it aims at providing the Internet service for those 'other three billion' population who have no access to the Internet. Up to now, O3b has deployed 16 satellites, which adopted transparent forwarding strategy at an orbit altitude of 8062 km. In addition, the satellites use the Ka frequency band and each satellite has 16 wave beams in order to provide users with a communication solution that can rival to 4G to the 'last mile' problem.

Iridium was proposed by Motorola in 1987, and the next generation (Iridium Next) started up construction in 2007. Iridium Next updates to a broadband information network, and each satellite carries various kinds of payloads for services such as

communication, climate change surveillance, multispectral imaging of the earth and so on. Iridium Next is equipped with cross-links and onboard Internet protocol (IP) switching functions. The downstream rate in L frequency band can be up to 1.5 Mbps and in Ka frequency band can be up to 30 Mbps.

OneWeb system plans to deploy and networking 720 low-orbiting satellites. The satellites are distributed across 18 orbit planes at an orbit altitude of 1200 km, and Ku frequency band is used, achieving 6 Gbps transmission rate for a single satellite, 50 Mbps for user access rate and a total system capacity of 6 Tbps. A single satellite weights about 150 kg and is produced by the devices of the civil grade in assembly lines, costing less than 1 million dollars.

SpaceX proposed a construction plan of satellite networks, and at least 4257 small satellites are planned to be deployed across 43 mixed orbits using Ku or Ka frequency band to provide space broadband wireless access services. The company plans to launch two experimental satellites, and more satellites will come into service in 2020. A single satellite weights about 386 kg and can provide broadband services with a system capacity up to 1 Gbps for a single user.

NASA proposed Space Communication and Navigation (SCaN) in 2007 to develop an integrated network infrastructure for space communication and navigation and provide services in the surface of Moon and Mars.

Europe proposed ISICOM, which not only integrates heterogeneous network nodes distributed across the high, medium, low orbits, near space or ground into a communication network with massive system capacity based on IP and mixed microwave/laser, but also provides users with value-added information services through integrating Galileo and GMES.

China is establishing BeiDou Navigation Satellite System (BDS), which consists of five geostationary earth orbit (GEO) satellites, three inclined geosynchronous satellite orbit (IGSO) satellites and 27 medium earth orbit (MEO) satellites. BDS can provide PNT services and provide services such as short message and data acquisition by means of integrating communication payloads in the GEO satellites. BDS preliminarily realizes the integration of navigation and communication.

### 2.2 Software-Defined Networks

Researchers in Stanford University proposed the concept of SDN [4]. The basic idea of SDN is to separate the control function such as routing decisions in routers from hardware devices, and the central controller not only decouples control plane and forwarding plane to achieve more flexible control, but also makes devices simpler.

SDN has drawn extensive attention since it came out, and many research institutes have carried out work on interface standardization, key techniques and related experiments. Some manufacturers put forward SDN solutions that accelerate the pace of SDN industrialization in terrestrial networks. China Unicom Research Institute indicates that SDN will become the fundamental network architecture and will experience a long-term evolution and transition process (Fig. 1).

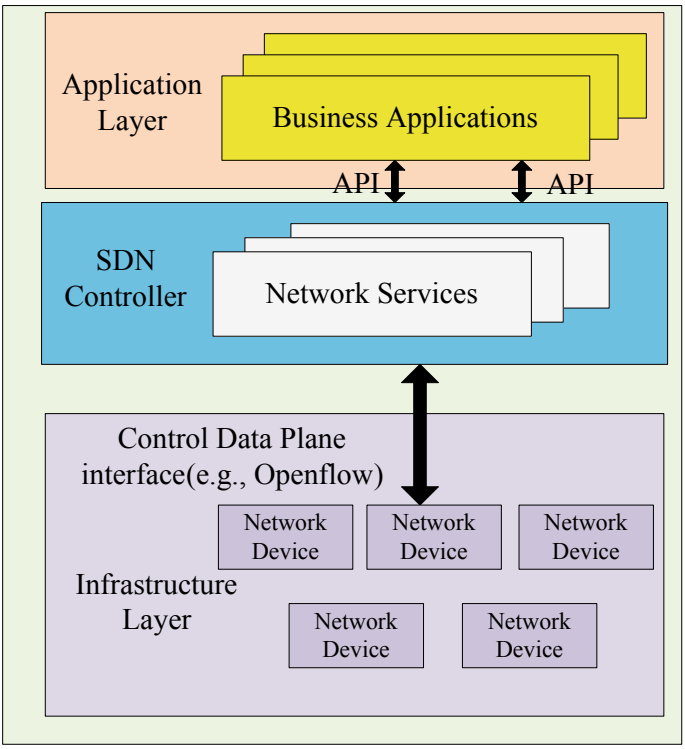



**Fig. 1.** SDN architecture

### 2.3 Network Visualization

Network visualization (NV) [5] is to abstract, accumulate and isolate resources on the same physical network infrastructure to allow heterogeneous visual networks sharing the same substrate network. This kind of network architecture is convenient to deploy new businesses, thus boosting the development of new innovation networks. All social circles pay much attention to NV and initiate a series of research projects. The European Union starts up FP7 research project—4WARD (Architecture and Design for the Future Internet)—and puts forward a network system architecture design scheme based on NV. Canada launches the Nouveau project, which builds a flexible, controllable and safe network based on the idea of NV.

### 2.4 Network Functions Visualization

NFV [6] uses the visualization technique to run all kinds of network services in standard servers, switches and storage systems. Network services are no longer provided in the form of exclusive hardware, but in the form of software encapsulated in virtual machines that can run in data centers, network nodes or even terminal user devices. NFV can optimize network configurations and topologies in real time

according to the variation of real network traffic and can realize agile deployment of network services, thus adapting to the business development fast and flexibly. As for manufacturers, NFV can help to reduce the development circle and accelerate the product launch. Network operators can reduce their investment on hardware devices and lower energy consumption of data centers by using NFV techniques. On October 22, 2012, European Telecommunications Standards Institute (ETSI) published the NFV white paper, which propelled the researches and applications on NFV.

## 3 Space-Based Integration Network Architecture

According to the future application demand, based on the concept of SDN, the network architecture of 'equipment entering network, information entering the cloud, service to terminal' is adopted, which is able to provide on-demand information transmission and distribution, network awareness, PNT service.

### 3.1 System Architecture

As shown in Fig. 2, the software-defined space-based integration network architecture consists of heterogeneous software-defined space-based nodes and ground support facilities [7, 8]. The software-defined space-based nodes have functions of transmission and switching, information awareness, navigation and positioning. The space-based integration network architecture is based on unified control of software definition, virtualization of network resources and network slicing.

Physical network: The space-based nodes use the general function module to deploy according to the different needs, and each node is equipped with the general hardware platform, in which the function of nodes is configured by software definition technology. By adjusting the coverage area, power and frequency bandwidth, the transmission/sensing/PNT function nodes are reconstructed, and the open interconnection with the mobile communication network and terrestrial Internet is achieved. The space-based transmission function nodes make use of space-based network programming and online allocation of network resources to dynamically respond to demand for access to network and information transmission and enhance the reliable access capability of mobile nodes. The space-based sensing function nodes use the sensing resources distributed in the middle and low orbit to carry out multi-dimensional cooperative intelligent detection, forming a global full-time coverage of cross-domain intelligent perceptual information network. The space-time PNT function nodes synthetically utilize many kinds of navigation information enhancement means to realize the navigation positioning accuracy to the specific user in the service area.

Information cloud: According to the demand of users, space-based transmission resources, sensing resources and PNT resources are dynamically organized to form a multi-dimensional spatiotemporal global logic model. On this basis, a unified cloud platform framework is constructed to achieve the aggregation of space-based distributed information resources and to provide network and communication, remote sensing and geographic information, PNT service in multicenter form. At the same time, the powerful data processing capability of the cloud platform is used to provide

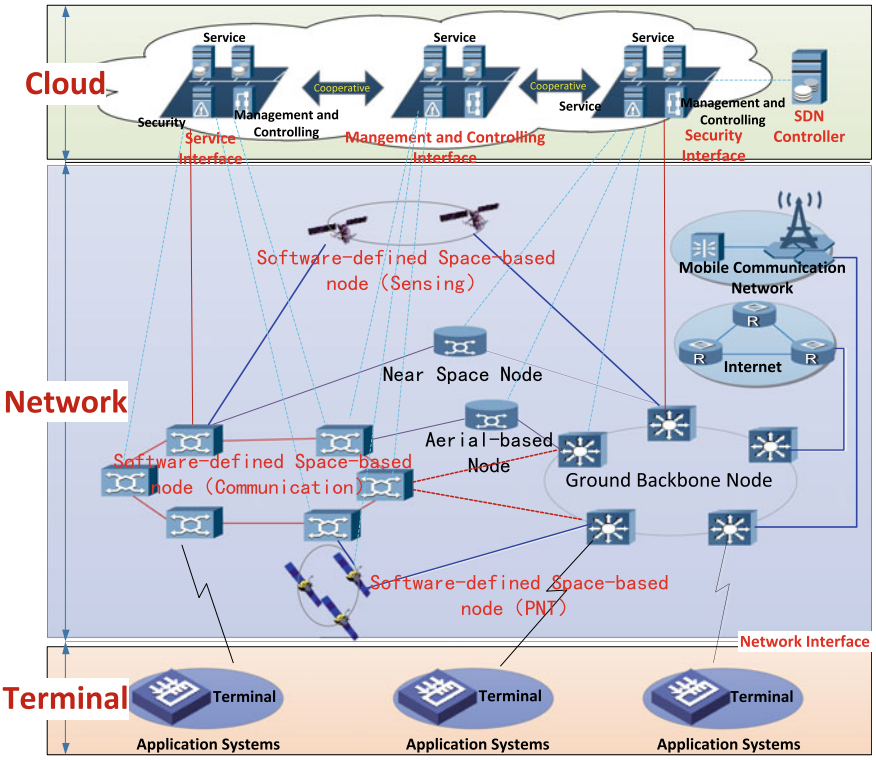



**Fig. 2.** Software-defined space-based integration network architecture

operation and maintenance control and security protection for space-based integration network. In the aspect of operation and maintenance control, the SDN controller is constructed on the cloud platform, which undertakes the tasks of computing, scheduling and managing space-based integration network. In the aspect of security protection, the cloud platform undertakes the tasks of link security encryption, network domain isolation, identity authentication and so on.

Service terminal: The functions of transmission, awareness and PNT are extended to the client and integrated with local applications to construct application systems that meet different requirements. Combined with the application scenario of space-based integration network, it provides three kinds of services, such as the information enhancement of hot spots, the space-based time-sensitive service and the collection of massive environmental information. The information enhancement of hot spots is oriented to the application area, aggregates the space resources on demand and provides enhanced communication, PNT and other information services to meet the needs of users at high speed and real time. The space-based time-sensitive service can meet the demand of large capacity, high reliability, low delay and accurate synchronization of time through the intelligent and rapid processing of the task. The collection of

massive environmental information makes use of large-scale sensors to obtain all elements of environmental information for social governance, to provide multi-source information support.

### 3.2 Technical Architecture

As shown in Fig. 3, the technical architecture of software-defined space-based integration network is divided into resource layer, SDN control layer, service layer, application layer and the security and maintenance domain across the previous four layers.

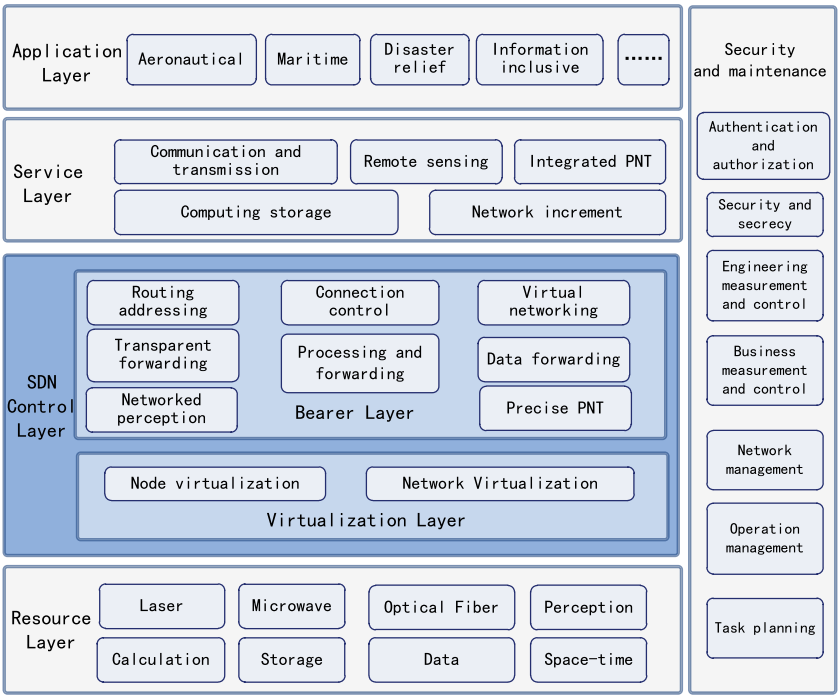



**Fig. 3.** Technical architecture of software-defined space-based integration network

Resource layer: It is the physical entity of the operation of the space-based integration network, which is the foundation to support the other three layers and called by the SDN control layer upward. It mainly includes laser, microwave, fiber and other transmission resources, computing, storage, data resources, perception and time and space reference resources, as well as internal links, external links and other links. It is realized by the antenna, transceiver equipment, computing storage devices and so on.

SDN control layer: It is the core of the operation of the space-based integration network, including the virtualization layer and the carrying layer. The virtualization layer is composed of two parts: node virtualization and network virtualization, which

virtualizes the physical resources, such as all kinds of space-based nodes and network links according to the unified description and is called on demand by the SDN controller. Because the SDN controller has a unique global perspective, the resource scheduling is always optimal, which is especially suitable for the space-based integration network under the condition of limited space resources. Under the global scheduling of SDN controller, the carrying layer mainly realizes the functions of service adaptation, protocol conversion, routing addressing, controllable security exchange, networked perception, precision PNT, data forwarding, flow balance, network destruction reconfiguration and so on.

Service layer: It provides communication transmission services, remote sensing information services, integrated PNT services, computing storage services, network value-added services and so on. Among them, for computing storage, it supports unified management and scheduling of networked computing storage resources, and provides flexible and scalable elastic computing service capabilities and massive data storage service capabilities. For communication and transmission, it provides various communication services, such as voice, data, video, status presentation, messages and so on. For remote sensing and integrated PNT services, it provides users with hierarchical remote sensing, integrated PNT service products, which are task-oriented and on demand. For service generation and provisioning, it provides network value-added services such as status presentation, network resource status and network resource control for application systems, which is realized by service encapsulation of part of the functions through service interfaces or components.

Application layer: It is implemented by user center, user terminal, software and data. Through the space-integrated ground network, it provides quality services in scenarios such as space, maritime, overseas, remote areas and emergency situations. It also studies application capabilities in different scenarios, including technical indicators, terminal forms and operating models.

Security and maintenance: It is implemented by various types of security protection equipment, software, control and control centers, and auxiliary facilities, including certification and authorization, security and confidentiality, engineering monitoring and control, business monitoring and control, network management, operation management and network mission planning. It investigates its classification, capability and system, and studies the deployment of this function between the nodes of the space and the ground, highlighting the integration of military and civilian networks.

## 4 Key Technologies

(1) Space-based network topology design and optimization technology
Topology optimization design is carried out from the aspects of frequency orbital resources, network coverage, communication capacity and rate requirements, navigation and positioning accuracy requirements, etc. Topological structure performance evaluation factors and evaluation models are constructed, and topological structure optimization design algorithms are designed, achieving heterogeneous integration of network structure and efficient management and operation of the network.

(2) Software-defined intelligent network node technology
From antenna, transmission, switching and processing, software-defined technologies can be used to design space-based nodes that can be reconfigured on orbit and function-based, enabling the decoupling of the satellite hardware platform and network functions, configuring communication, awareness and navigation according to the task requirements and finally realizing task aggregation and timely response.

(3) Distributed cooperative transmission technology in space-based integration network
For multi-star application scenarios such as formation flying, concentric constellation and mixed orbit constellation in space-based fusion networks, under the conditions of different satellite channel models, it breaks through distributed cooperative transmission technologies and achieves integrated multi-star communications.

(4) Layered cross-domain edge computing technology
Focusing on the differences between different satellite computing resources and computing methods in space-based integration networks, the hierarchical distributed computing resource coordination scheduling mechanism is analyzed, which provides the cross-domain collaborative multi-satellite computing capabilities, multi-level storage computing capabilities and distributed processing capabilities.

(5) The task-oriented cognitive model building technology of supply and demand of space-based resources
Through the unified modeling of mission requirements and space-based resource characteristics, the requirements for space-based resources are accurately expressed and parametrically described, enabling space-based information resource convergence and precise services,which provides model support for task-oriented application services of space-based integration network.

(6) Networked cooperative awareness technology
Based on satellite sensing system development planning and application task requirements, an optimized scheduling model for satellite-aware resources is established to realize resource services based on online interactions, intelligent demand coordination and analysis, and awareness of demand coordination and management, to enhance space-based awareness of resource scheduling capabilities.

(7) Space-time reference unified technology for heterogeneous orbital inter-satellite links
Based on the inter-satellite link network of the space-based integration network, a reliable and unified space-time reference is obtained during the sensing and networking of large-scale heterogeneous orbit satellites, to realize the maintenance and fault tolerance of space-time reference and the unification of multi-source integration on the space, and to provide a unified and accurate space-time reference.

## 5 Conclusions

In this article, the software-defined space-based integration network architecture is proposed. The system architecture of the proposed space-based integration network consists of the physical network, information cloud and service terminal, and the technical architecture consists of resource layer, SDN control layer, service layer, application layer, security and maintenance. Simultaneously, the key technology system to support space-based integration network is proposed, which provides a theoretical and technical basis for the construction of software-defined space-based integration network.

## References

1. Manqing W, et al. The Overall idea of space integrated ground network. Satell Netw. 2016;3:30–6.
2. Zhou L, et.al. The overall idea of space integrated ground network. In: The 12th annual conference of satellite communications; 2016. p. 1–11.
3. Peng Q, et al. The SoS capability reference model of space-ground integrated network. J Chin Acad Electron Inf Technol. 2016;6:629–35.
4. Open Networking Foundation. Software-defined networking: the newnorm for networks. White Paper, vol. 4; 2012.
5. Chowdhury N, Boutaba R. Network virtualization: state of the art and research challenges. IEEE Commun Mag. 2009;47(7):20–6.
6. Virtualisation NF. Introductory white paper. In: SDN and openflow world congress, Darmstadt, Germany; 2012.
7. Peng Q, et al. Design of ground systems for SCN driven by concurrent multitasking. J China Acad Electron Inf Technol. 2015;5:492–6.
8. Peng Q, et al. Consideration on the construction of integrated management and control system of space-based information network. J China Acad Electron Inf Technol. 2017;10:443–7.

# An Improved Interacting Multiple Model Algorithm Based on Switching of AR Model Set

Wenqiang Wei, Jidong Suo(✉), and Xiaoming Liu

Dalian Maritime University, Dalian, China
sjddmu@dlmu.edu.cn

**Abstract.** To solve the problem of complex maneuvering target tracking, an improved interacting multiple model (IS-AR-IMM) algorithm based on AR model set switching is proposed in this paper. The algorithm is proposed on the basis of the AR-IMM algorithm based on AR model and S-IMM algorithm based on model set switching. Simulation shows that IS-AR-IMM algorithm not only solves the problem of increasing computational complexity and model selection caused by the increase of models in AR-IMM and S-IMM algorithms, but also has better tracking performance.

**Keywords:** Radar target tracking · AR model · Model set switching · Interacting multiple model

## 1 Introduction

Maneuvering target tracking is based on modeling, estimation, tracking, and prediction of the moving state of the target, so as to achieve real-time calculation of target states [1]. It is an important research topic in vessel traffic service (VTS) system. Because of the rapid development of the maritime industry and the increasing complexity of maritime transportation, it is necessary to constantly update the tracking technology to improve the tracking accuracy of the target.

Two key parts of maneuvering target tracking are maneuvering target tracking model and adaptive filtering algorithm. They are the important research direction in the field of target tracking; that is, the target tracking model can be more satisfied with the actual movement, and the new filtering algorithm is constantly improved or found to adapt to the changeable moving target. Therefore, the selection of target state quantities, the establishment of models, and the rationality of filtering algorithm selection will directly affect the accuracy of target tracking [2].

In this paper, we first study and analyze an IMM algorithm based on AR model. It applies the AR model with different orders to the IMM framework to track maneuvering targets. Among them, the first- and second-order AR model can describe uniform motion and uniform acceleration motion. Next, we will refer to the interacting multiple model algorithm based on model set switching proposed in document [3]. It executes the IMM algorithm in parallel with two different model sets and then uses the probability of different models to compare the motion state of the target; thus, the result


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_56

of the better model set is used as the output of the time. Although the algorithm effectively improves the probability of the model matching and tracking accuracy, the computation is greatly increased.

The IS-AR-IMM algorithm defines two different model sets to track the two states of the target's non-maneuverability and maneuver, respectively. If the target motion state changes, the model sets will automatically switch between the model sets, and only one model set is used for target tracking at the most time. The complexity of the improved algorithm is obviously smaller than that of the S-IMM algorithm, and it also has superior tracking performance. Finally, the feasibility of the proposed model set and algorithm is verified by simulation.

## 2 AR-IMM Algorithm Steps

Because of the unique operation mechanism of IMM algorithm, it has been widely used in the field of target tracking. In addition, combined with the optimal solution of the coefficient of AR model in document [4], the tracking advantage is more prominent when the number of AR coefficients is $N = 4$. So, the next section will introduce the specific steps of IMM algorithm based on AR model.

Step 1: the initialization of the filter
Step 2: the calculation of the state transition matrix
Step 3: interactive input of the model
Step 4: model filtering
Step 5: model probability update
Step 6: state estimation fusion
Step 7: let $k = k + 1$, proceed with step 2.

## 3 Improved Interacting Multiple Model Algorithm Based on AR Model Set Switching

In the AR-IMM algorithm, the model probability μdetermines the matching degree between the state of the target and the model. If the probability of a model is larger, it is considered that it matches the state of the target at this time. In addition, for tracking complex and changeable maneuvering targets, if the number of models in the model set is too few, then it cannot completely describe the target motion state of each maneuvering time, and the tracking effect will have a certain influence. If the number of models is too many, the amount of calculation will become larger, and there are more mismatched models at each time of the target movement, and they also have a certain probability, so the tracking effect will also be affected.

Considering the shortcomings of the large amount of operation of the S-IMM algorithm, and combining the methods applied to the AR model in the previous section, an improved interactive multiple model algorithm (IS-AR-IMM) based on AR model set switching is proposed.

(1) The composition and switching of AR model set

The IS-AR-IMM algorithm sets two different models set to track the two states of non-maneuver (non-M) and maneuver, respectively. This method will compare the sum of probability of the matching models in the current model set and the probability of the mismatch model to determine the motion state of the target. If the sum of the probability of the matching model is greater than the probability of the mismatched model, the motion state of the target has not been changed, and the current model set can match the moving state of the target. Conversely, it can be explained that the motion state of the target has changed, and the current model set cannot match the motion state of the target very well. At this point, the system will automatically transform into another set of models, and then track the target in real time. Therefore, only when the system changes the target motion state, the model set can be transformed, so there is only one set of model sets for target tracking at the most time, and the simulation later confirms this.

(2) IS-AR-IMM algorithm steps

Suppose that $M^{\,i}$ is a model set at $k = 0$ moments and run AR-IMM algorithm for it. The specific implementation steps of the IS-AR-IMM algorithm are as follows:

Step 1: The output state estimation of AR-IMM is $\boldsymbol{x}_{k+1/k+1}$; the corresponding covariance $\boldsymbol{P}_{k+1/k+1}$; the probability of model set $M^{i(j)}$ of the j model $\mu_{k+1}^{i(j)}$.

Step 2: Determines whether the model set matches the current target motion state.

If the current model set is $M^{\,1}$, then the target's non-maneuvering state is tracked at this time. If $\mu^{1}_{(\text{non}-\text{m})k+1} < \mu^{1}_{(\text{m})k+1}$, the target is judged to be maneuvering, indicating that the model set $M^{\,1}$ does not apply to the current target motion state. If $\mu^{1}_{(\text{non}-\text{m})k+1} > \mu^{1}_{(\text{m})k+1}$, the target is judged to be maneuvering, indicating that the model set $M^{\,1}$ is suitable for the current target motion state. The case of model set $M^{\,2}$ is homogeneous.

Step 3: Select the matching model set according to the result of step 2.

If the result of judgment shows that the model set $M^i$ is a matching model set, then execute the AR-IMM algorithm corresponding to $M^i$ at the next moment, then go back to step 1, and continue the iteration operation. If the result shows that the model set $M^i$ is not a matching model set, then proceed to the next step.

Step 4: Initialize another model set.

Specific rules: If the current model set is $M^1$, then select the corresponding state of the model with larger model probability in the two first-order AR model at the last time to initialize the first-order AR model in the model set $M^2$, and the corresponding state of the second-order AR model in the model set $M^1$ to initialize two second-order AR modes in the $M^2$. The case of model set $M^2$ is homogeneous.

After the initialization is completed, the current model set is defined empty, and then the AR-IMM algorithm corresponding to another model set $M^j$ is executed.

In step 4, the same model in the other model set $M^j$ is initialized using the state of the model with larger model probability in the model set $M^i$. It is because the model with larger model probability in the two same models in the model set $M^i$ can match the current motion state of the target more. While the maneuver changes at the next moment, it will become a mismatched model, its model probability will be relatively smaller, and so its fusion weight will be relatively smaller, which can further improve the tracking accuracy.

Step 5: Go back to step 1 to enter the iteration calculation for the next sampling.

The single-step flowchart of the IS-AR-IMM algorithm is shown in Fig. 1.

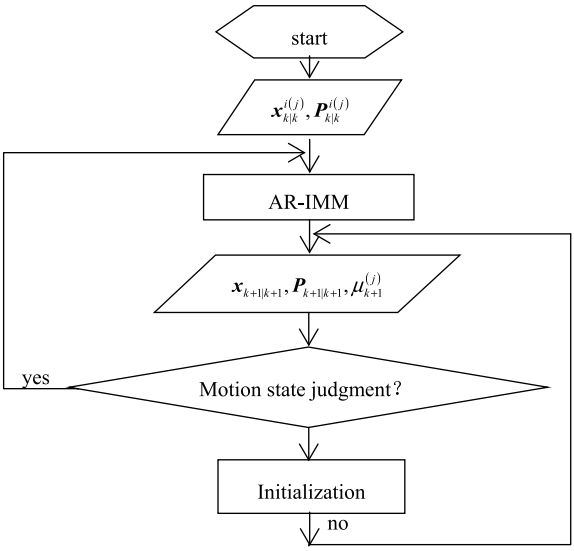



**Fig. 1.** Single-step flowchart of IS-AR-IMM algorithm

The single-step flowchart of the initialization of the IS-AR-IMM algorithm is shown in Fig. 2.

## 4 Simulations

In this section, the tracking effect of IS-AR-IMM and AR-IMM1 and AR-IMM2 will be compared by simulations.

This simulation of a two-dimensional turning scene of a target ship in a certain water area consists of three parts: uniform speed, fast turning, and slow turning. Assuming that the ship initially performs uniform motion. Between 41 and 52 s, the ship turns fast, and the acceleration on $X$-axis and $Y$-axis is $x_a = 1$ m/s$^2$ and $y_a = 1$ m/s$^2$, respectively. Then, the ship is moving at a uniform speed, from 71 to 80 s, the ship is turning slowly, and the acceleration on $X$-axis and $Y$-axis is $x_a = 0.1$ m/s$^2$ and

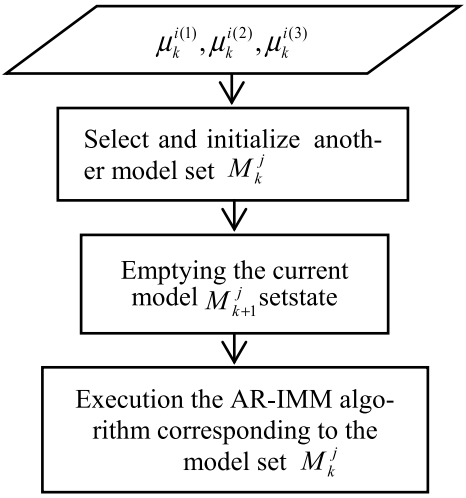



**Fig. 2.** Initialization single-step flowchart of IS-AR-IMM algorithm

$y_a$ = 0.2 m/s$^2$, respectively. It is assumed that only the location information of the ship can be obtained. The variance of the observed noise is 100 m$^2$, the sampling interval is $\Delta t$ = 1 s, and the total tracking time is 120 s. The experimental results were obtained on average through 200 Monte Carlo experiments.

Here, the model set $M^1$ is first set up, which consists of two first-order AR models and one second-order AR model, of which $q_{r1}^1 = 0.01$, $q_{r2}^1 = 0.1$, $q_{r3}^1 = 0.5$. Then, the model set $M^2$ is set up, which consists of one first-order and two second-order AR models, of which $q_{r1}^2 = 0.01$, $q_{r2}^2 = 0.1$, and $q_{r3}^2 = 0.5$. The number of the coefficients of the AR models are all $N = 4$. The state noise covariance of each model is all $\boldsymbol{Q}_{r3}^{i(j)} = q_{rj}\Delta t \cdot \boldsymbol{I}$, and here, $i$ and $j$ are used to represent the $i$ and $j$ models in the model set, respectively. The transfer probability matrix of the model is also defined as

$$\boldsymbol{\pi}_{ij} = \begin{bmatrix} 0.90 & 0.05 & 0.05 \\ 0.05 & 0.90 & 0.05 \\ 0.05 & 0.05 & 0.90 \end{bmatrix} \tag{1}$$

Through the data in Table 1, the operation time of IS-AR-IMM is a little longer than that of AR-IMM, and about 7% more. This is because the IS-AR-IMM algorithm determines whether the model set needs to switch at every moment, and if necessary, a AR-IMM algorithm needs to be executed at that time. It can also be seen that the S-IMM algorithm in [3] is more than two times as much time as IMM and AR-IMM,

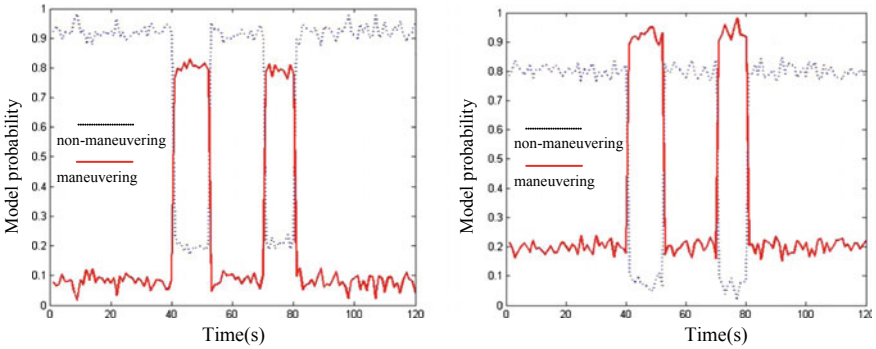
**Table 1.** Operation time of different algorithms for maneuvering target tracking

| Algorithms | AR-IMM1 | AR-IMM2 | S-IMM | IS-AR-IMM |
|---|---|---|---|---|
| Operation time (s) | 0.89 | 0.88 | 1.83 | 0.95 |

because IMM1 and IMM2 must be calculated at every sampling time in S-IMM, and at the same time, the state of the target needs to be detected to select the best tracking result. Therefore, the operation time of S-IMM is long and the complexity is high.

Further comparative analyses of the four algorithms are made below from the two aspects: model probability and tracking accuracy.

The probability of non-maneuvering and maneuvering models corresponding to the respective models set in the two AR-IMM algorithms is shown in Fig. 3a, b, respectively. From the graph, the probability of the matching model is always greater than the probability of the mismatching model in the whole process of the target, whether it is the AR-IMM1 algorithm or the AR-IMM2 algorithm.



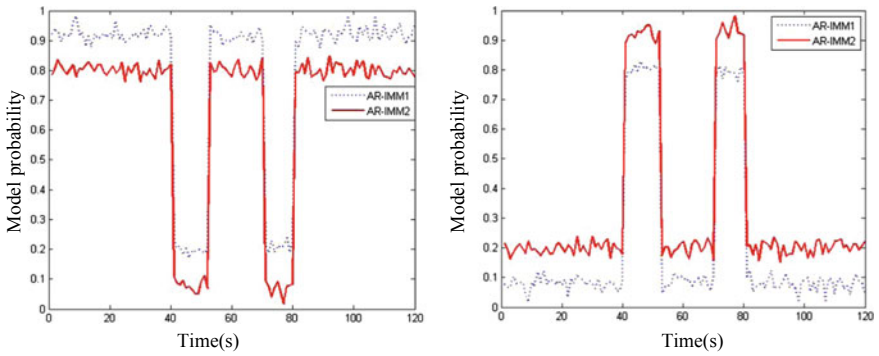

(a) AR-IMM1 (b) AR-IMM2

**Fig. 3.** Comparison of model probabilities between AR-IMM1 and AR-IMM2 algorithm

In Fig. 4a, b, the probabilities of the non-maneuvering model and the maneuvering model of the model sets corresponding to the two AR-IMM algorithms are compared, respectively. For example, if the ship is in a non-maneuverable state (0–40 s), the sum of probabilities of the first-order AR models of the model set in the AR-IMM1 algorithm is significantly greater than the probability of the first-order AR model of the model set corresponding to the AR-IMM2 algorithm. When the ship is in a maneuvering state (41–52 s), the sum of the probabilities of the second-order models of the model set by the AR-IMM2 algorithm is significantly greater than the probability of the second-order model in the model set corresponding to the AR-IMM1 algorithm.

Combining Fig. 4a, b and the steps of the IS-AR-IMM algorithm in this paper, we can see that when the target is in a non-maneuverable state, the IS-AR-IMM algorithm will select a more matched model set $M^1$ and execute the AR-IMM1 algorithm M; when in a maneuvering state, IS-AR-IMM will select a more matching model set $M^2$ and execute AR-IMM2 algorithm. At the same time, the state of the current model set will be emptied after initializing another model set at each maneuvering change time.

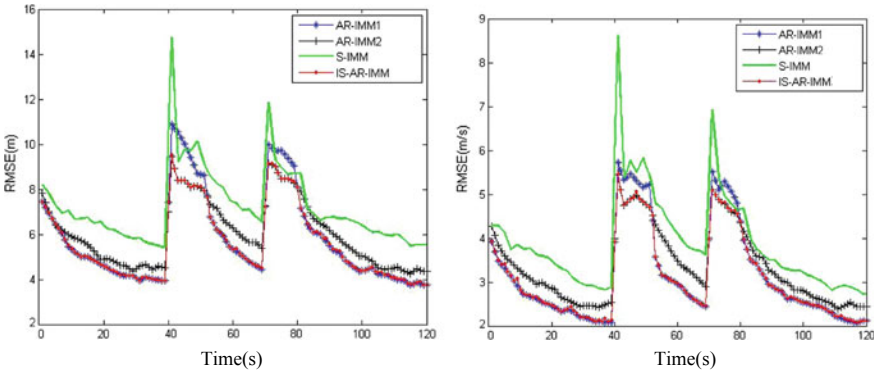
Next, we will make detailed tracking performance comparison analysis through the RMSE estimated by these four algorithms.




(a) non-maneuvering model (b) maneuvering model

**Fig. 4.** Comparison of model probabilities between AR-IMM1 and AR-IMM2 algorithm

Figure 5a, b, respectively, describes the RMSE curves of the four algorithms of AR-IMM1, AR-IMM2, S-IMM, and IS-AR-IMM for estimating the position and velocity of the target ship. It is found that when the ship is moving at uniform speed, the RMSE curve of the IS-AR-IMM algorithm is basically the same as that of the AR-IMM1 algorithm. There are the following reasons:




(a) position (b) velocity

**Fig. 5.** Comparison of the estimation error between four algorithms

(1) The number of the coefficients of the two AR models is the same, all of which are $N = 4$.
(2) The number of non-maneuverable and maneuvering models of the two is the same. The non-maneuverable model contains two first-order AR models, and the maneuvering model contains two second-order AR models.
(3) The corresponding state noise intensity of each model is the same.

Compared with AR-IMM2 algorithm, the position and velocity tracking performance of IS-AR-IMM algorithm is about 8 and 15%, respectively, at a uniform speed between 0 and 40 s. Compared with the AR-IMM1 algorithm, the location and speed tracking performance of the IS-AR-IMM algorithm is about 5 and 10%, respectively, in the fast turning state from 41 to 52 s. Compared with the AR-IMM2 algorithm, the position and velocity tracking performance of the IS-AR-IMM algorithm is about 12 and 20%, respectively, at the uniform speed of 53–70 s. Compared with the AR-IMM1 algorithm, the location and speed tracking performance of the IS-AR-IMM algorithm is about 6 and 14%, respectively, in the state of slow turning between 71 and 80 s. Compared with the AR-IMM2 algorithm, the position and velocity tracking performance of the IS-AR-IMM algorithm is about 6% and 15%, respectively, at the uniform speed between 81 and 120 s.

In addition, from Fig. 5a, b, we can see that compared with the S-IMM algorithm proposed in [3], the AR-IMM and IS-AR-IMM algorithms mentioned in this paper show obvious advantages in tracking accuracy. In combination with Table 1, although the IS-AR-IMM algorithm is a little longer than the AR-IMM algorithm in computing time, its tracking performance is more prominent, especially in speed tracking accuracy.

## 5 Conclusions

In this paper, a relatively new AR model is used to study radar target tracking. Based on AR-IMM algorithm and S-IMM algorithm, IS-AR-IMM algorithm is proposed. The algorithm uses two optimized model sets to track the non-maneuverable and maneuverable state of the target in turn. By comparing the probability characteristics of the models in the current model set, the target maneuver detection is realized at each time to determine whether the model set switching is needed. The simulation results show that, although the IS-AR-IMM algorithm is slightly longer than the AR-IMM algorithm in the operation time, compared with the AR-IMM algorithm and the S-IMM algorithm, the tracking performance has been significantly improved.

## References

1. Li XR, Jilkov VP. Survey of maneuvering target tracking. Part I. Dynamic models. IEEE Trans Aerosp Electron Syst. 2003;39(4):1333–64.
2. Ling Y, Ming L. Research of method for tracking high speed and highly maneuvering target. In: The 6th international conference on ITS telecommunications; 2006. p. 1236–9.
3. Can S, Jianping X, Haozhe L, Wei W, Yanbo Z. S-IMM: switched IMM algorithm for maneuvering target tracking. J Convergence Inf Technol. 2012;14(7):461–8.
4. Jin B, Jiu B, Su T, Liu H, Liu G. Switched kalman filter-interacting multiple model algorithm based on optimal autoregressive model for maneuvering target tracking. IET Radar Sonar Navig. 2015;9:199–209.

# Research on Interval Constraint Range Domain Algorithm for 3D Reconstruction Algorithm Based on Binocular Stereo Vision

Caiqing Wang, Shubin Wang(✉), Enshuo Zhang, and Jingtao Du

College of Electronic Information Engineering,
Inner Mongolia University, Hohhot 010021, China
wangsb09@gmail.com

**Abstract.** For 3D reconstruction algorithm based on binocular stereo vision only approximately restores the 3D shape of the object, but it does not analyze the performance of using 3D reconstruction algorithm based on the model of binocular stereo vision measuring distance device to restore object, an interval constraint range domain algorithm is proposed in this paper. In this paper, this algorithm finds the boundary points of the object, the performance of using 3D reconstruction algorithm to restore object depends on the proportion of the number of points that fall within the constraint interval to the total number of points. The interval constraint range domain algorithm analyzes the performance of the 3D reconstruction algorithm to recover the object, that is, analyzes the measurement accuracy of the algorithm. Under the condition of the resolution of the image sensor used in the algorithm is identical, the simulation results show that the accuracy of the algorithm is determined by the spacing of the image sensor, the position, and shape of the constraint interval.

**Keywords:** Binocular stereo vision · 3D reconstruction · Interval constraint domain algorithm

## 1 Introduction

Binocular stereo vision is an important research direction of computer vision. It is a measurement method to obtain 3D geometric information of an object through two images. The basic principle is to observe the same scene from two viewpoints to obtain perceptual images under different visual angles and calculate the 3D information of the scene by calculating the positional deviation between the pixels of the image through the imaging geometry principle. Because the principle of stereo vision is simple and clear, it has broad application prospects in the field of 3D reconstruction and measurement of space objects. For the measurement of moving objects, the stereo vision method is a more effective measurement method because the image acquisition is completed instantaneously. 3D reconstruction is a computer technology that recovers


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_57

3D information of objects using two-dimensional projection. At present, there are many applications of 3D reconstruction. Xu et al. [1] introduced the experimental study of measuring distance, baseline distance, and baseline direction. Their influence on the accuracy of camera reconstruction was studied. Wang and Li [2] studied the method of measuring the size of a 3D object using a digital camera.

The purpose of this paper is to use a new 3D reconstruction algorithm to approximate the 3D shape of an object based on the binocular stereo vision acquisition device and proposes an interval-constrained domain algorithm to analyze the performance of the 3D reconstruction algorithm to recover the object, that is, to analyze the algorithm's measurement accuracy.

## 2 Binocular Stereo Vision and 3D Reconstruction

Binocular stereo vision is a method of acquiring 3D geometric information of an object from two images. In the 3D movie shooting, two cameras are used to imitate people's eyes at the same time. During the screening, the images captured by the two cameras are projected on the screen at the same time, and the principle of polarized light is used so that the eyes of the person can be seen separately. The camera captures images that give people a real 3D view of the scene. The realization of binocular stereo vision technology can usually be divided into six major steps: image acquisition, camera calibration, image preprocessing, stereo matching, 3D reconstruction, post-processing, and these six parts have an important influence on the realization and accuracy of the system. Most of the existing binocular stereo vision systems use this principle. Gao et al. [3] proposed a method for simultaneous positioning and mapping of autonomous mobile robots in indoor environments using binocular stereo vision systems. Zhu [4] used binocular stereo vision information systems to realize real-time tracking and measurement of moving objects in 3D space distance.

3D reconstruction is an important content and ultimate goal of computer vision research. In simple terms, 3D reconstruction is the restoration of a 2D image of an object into the 3D shape of a visible surface object. The basic principle of 3D reconstruction algorithm based on binocular stereo vision is to observe the same object from two different angles, then obtain two two-dimensional images from two different angles, and finally obtain 3D information of the object through the positional deviation of the two two-dimensional images. To design an accurate computer stereoscopic vision system, we must consider various factors in an integrated manner to ensure that each link has a high degree of accuracy. Jia et al. [5] studied the construction of 3D temperature distribution reconstruction system based on binocular vision technology.

# 3 Interval Constraint Range Domain Algorithm for 3D Reconstruction Algorithm Based on Binocular Stereo Vision

## 3.1 The Model of 3D Reconstruction Algorithm Based on Binocular Stereo Vision

The model of the binocular stereoscopic 3D reconstruction algorithm is shown in Fig. 1, which is a model of binocular stereo vision distance measuring device. It should be noted here that the coordinates of all points are established in the space coordinate system. $A_1B_1C_1D_1$ and $A_2B_2C_2D_2$ are two photosensitive elements of rectangular image sensors. $O_1(x_{01}, 0, 0)$ and $O_2(x_{02}, 0, 0)$ are projections of the intersection of the plane of the field of view angle on the floor. $E_1$, $E_2$ are the projection coordinate of $E$ on the photosensitive element of the image sensor.

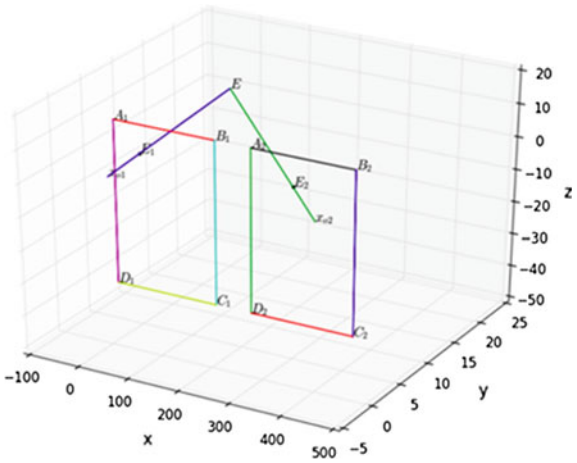



**Fig. 1.** Model of binocular stereo vision measuring distance device

In order to find the spatial information of the measured object, in Sect. 3.2, this paper proposes a 3D reconstruction algorithm based on binocular stereo vision.

## 3.2 Interval Constraint Range Domain Algorithm

Before presenting the interval constraint range algorithm, this paper first proposes a 3D reconstruction algorithm based on binocular stereo vision through the model of binocular stereo vision measuring distance device. This algorithm is used to calculate the coordinates of the boundary points of the measured object, and it can restore the 3D shape of the object by using the boundary points of the object. The algorithm of the 3D reconstruction based on binocular stereo vision in this paper is as follows. The case of

four points is used as an example to illustrate the situation of multiple points. Take the four boundary points of the object: $P_1(x_{01}, y_{01}, z_{01})$, $P_2(x_{02}, y_{02}, z_{02})$, $P_3(x_{03}, y_{03}, z_{03})$, $P_4(x_{04}, y_{04}, z_{04})$. In addition, the projection of these four points on the quadrilateral $A_1B_1C_1D_1$ is $E_{11}(x_{11}, y_{11}, z_{11})$, $E_{12}(x_{12}, y_{12}, z_{12})$, $E_{13}(x_{13}, y_{13}, z_{13})$, $E_{14}(x_{14}, y_{14}, z_{14})$, respectively. The projection of these four points on the quadrilateral $A_2B_2C_2D_2$ is $E_{21}(x_{21}, y_{21}, z_{21})$, $E_{22}(x_{22}, y_{22}, z_{22})$, $E_{23}(x_{23}, y_{23}, z_{23})$, $E_{24}(x_{24}, y_{24}, z_{24})$, respectively. According to the model of binocular stereo vision measuring distance device, the intersection of space lines can determine a unique point, then the space line equation for the boundary point $P_1$ is as follows Eq. (1), and the coordinate of the boundary point $P_1$ is derived from Eq. (1) as Eq. (2). Among them, $x_{11}$, $x_{21}$, $z_{21}$, $y_{11}$ is equivalent to $x_1$, $x_2$, $z_2$, $y_1$.

$$\begin{cases} \frac{x-x_{01}}{x_1-x_{01}} = \frac{y}{y_1} = \frac{z}{z_1} \\ \frac{x-x_{02}}{x_2-x_{02}} = \frac{y}{y_2} = \frac{z}{z_2} \\ y_1 = y_2 \end{cases} \tag{1}$$

$$\begin{cases} x_{01} = \dfrac{x_{11}x_{02} - x_{21}x_{01}}{(x_{02} - x_{01}) + (x_{11} - x_{21})} \\ y_{01} = \dfrac{(x_{02} - x_{01})y_{11}}{(x_{02} - x_{01}) + (x_{11} - x_{21})} \\ z_{01} = \dfrac{(x_{02} - x_{01})z_{21}}{(x_{02} - x_{01}) + (x_{11} - x_{21})} \end{cases} \tag{2}$$

Since the calculation method for the coordinate of each target point is identical, the coordinate of the target point $P_2$, $P_3$, $P_4$ can be obtained from Eq. (2). Since the above-mentioned 3D reconstruction algorithm based on binocular stereo vision only approximates the 3D shape of the restored object, the performance of the 3D reconstruction algorithm to restore the object under the model of binocular stereo vision measuring distance device is not analyzed. Therefore, this paper proposes an interval constraint range algorithm to analyze the performance of the 3D reconstruction algorithm to recover the object. In this paper, the object to be measured is based on the binocular stereo vision distance measuring device model, and a new 3D reconstruction algorithm is used to find the coordinate of the boundary point. The 3D reconstruction algorithm can obtain the coordinates of the boundary points of an object through the projection of boundary points on the image sensor. The 3D reconstruction algorithm can obtain the coordinates of some points on the boundary of the object through the boundary point projection on the image sensor. Since these points are only a part of the boundary of the object, the closed graph formed by the coordinates of these boundary points approximately represents the 3D geometry of the object. People are familiar with the size of a conventional object, so this article constructs the size of a regular object through a constraint interval; that is, the range of $x$, $y$, $z$ can approximate the size of a regular object. A number of boundary points of the object are obtained from the above 3D reconstruction algorithm, and the number and coordinates of the points that fall within the specified constraint interval are calculated. The point that falls within the specified constraint interval is the point that satisfies the requirement. The interval

constraint range algorithm is used to calculate the number and coordinates of points in a given constraint interval, then finding the number and coordinates of the points that meet the requirements, and calculating the proportion of the points that meet the condition among all points to evaluate the performance of binocular stereo vision measuring distance device. The set of coordinates of points within a given constraint interval is called constraint range domain. 3D reconstruction algorithm finds the boundary points of the object, where points that fall within a given constraint interval are points that satisfy the condition. The ratio of the number of points satisfying the condition to the total number of points is called the constraint proportion, and then the performance of the 3D reconstruction algorithm to restore the object is evaluated according to the constraint proportion.

## 4 Simulation and Analysis

That the multiple boundary points of an object adopt the interval constraint domain algorithm is similar to the case of a boundary point, therefore, we only study one boundary point using the interval constraint range algorithm in this paper. It is known that the coordinate of the left and right image sensors is $x_{01} = 24$, $x_{02} = 1048$, respectively.$y_1 = 2$, at this time, taking three constraint intervals: $x \in [0, 1000]$, $y$, $z \in [-300, 300]$. The simulation result is shown in Fig. 2.

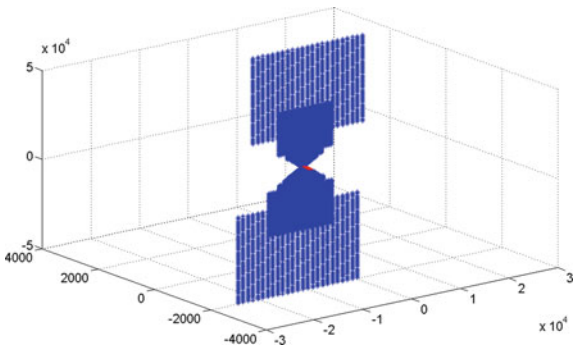



**Fig. 2.** Points that fall within the constraint interval for all points obtained by the 3D reconstruction algorithm

The blue area consists of points found by a 3D reconstruction algorithm. Points in the blue region that satisfy the interval constraint form a red region. From the simulation result, it can be seen that the proportion of the red part is small, but the density is higher in the range. Based on the new 3D reconstruction algorithm, the coordinates of 109440 points can be obtained. Among them, 11604 points fall into the constraint

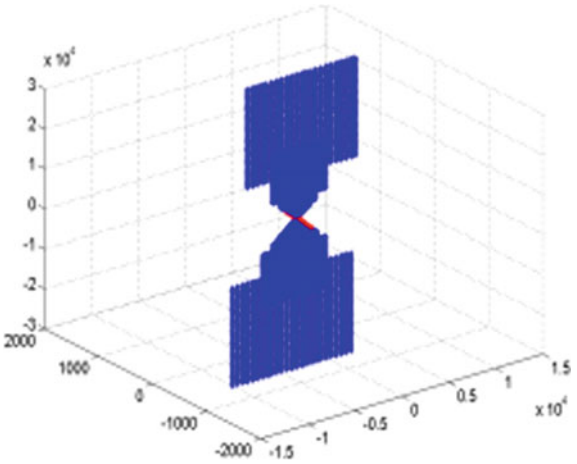



**Fig. 3.** Influence of image sensor's spacing on constraint proportion

interval; that is, 11604 points on the object to be measured are acquired, and the constraint ratio is 10.6%.

Now we discuss the effect of the variation of the pitch of the image sensor on the proportion of the constraint when the resolution of the image sensor is not changed. It is known that the coordinate of the left and right image sensors is $x_{01} = 24$, $x_{02} = 548$, respectively, $y_1 = 2$, at this time, $x \in [0, 1000]$, $y$, $z \in [-300, 300]$. It means that the pitch of the image sensor is reduced, the simulation result is shown in Fig. 3. The coordinate of 109440 points can be obtained. Among them, 24633 points satisfy the interval constraint; that is, 24633 points on the object to be measured are acquired, the constraint proportion is 22.51%, so the constraint proportion increases. Therefore, reducing the pitch of the image sensor will improve the performance of binocular stereo vision measuring distance device.

When the constraint interval size is unchanged and the position of the constraint interval changes, the simulation result is shown in Fig. 4. It is known that the coordinate of the left and right image sensors is $x_{01} = 24$, $x_{02} = 548$, respectively, $y_1 = 2$, at this time, $x \in [0, 1000]$, $y \in [0, 600]$, $z \in [-300, 300]$. Then you can get the coordinate of 109440 points based on the detailed simulation data. Among them, 860 points meet interval constraint, and constraint proportion is 0.79; that is, after the position of the constraint interval changes, the constraint proportion decreases. Therefore, the position of the constraint interval will affect binocular stereo vision measuring distance device.

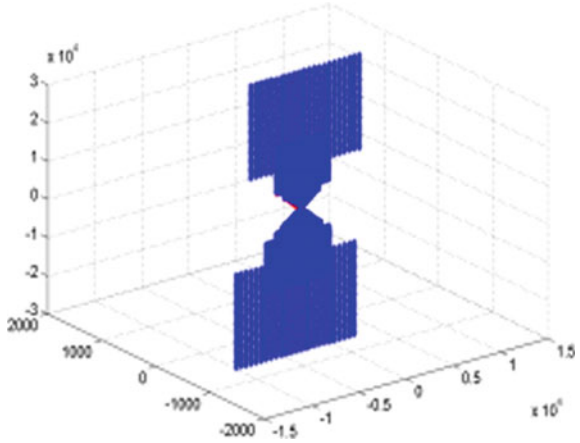



**Fig. 4.** Influence of constraint interval's position on constraint proportion

This shows that the constraint proportion is affected by other factors than the image sensor pitch. From the simulation data above, we know that the shape of the constraint domain is an important factor. It needs to be studied from the perspective of shape. It is known that the coordinate of the left and right image sensors is $x_{01} = 24$, $x_{02} = 1048$ respectively, $y_1 \in [0, 600]$, the value of $x$, $z$ is infinite, the simulation result is shown in Fig. 5. Then you can get the coordinate of 109440 points. There are 10080 points that satisfy the interval constraint; that is, 10080 points on the object to be measured are acquired, the constraint ratio is $9.21\%$, and at this time, $x \in [1092.5, 6068]$ and $z \in [44.5217, 12288]$. That is, the tested object must be in the above area before the algorithm can work. As can be seen from Fig. 5, when the distance of the measured object from the image sensor is longer, the points obtained by the algorithm are sparser; that is, the accuracy is lower. Therefore, reducing the pitch of the image sensor can increase its measurement range. After using the interval constraint range algorithm, the size of the constraint proportion means the performance of the 3D reconstruction algorithm. The greater the constraint proportion is, the better the performance of the 3D reconstruction algorithm is.

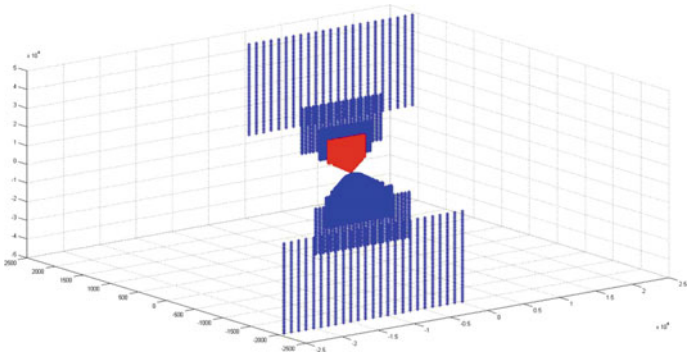


**Fig. 5.** Influence of constraint domain shape on constraint proportion

## 5 Conclusion

Under the condition of binocular stereo vision acquisition device, the interval constraint range algorithm proposed in this paper can analyze the performance of the 3D reconstruction algorithm to recover the object; that is, it can analyze the measurement accuracy of the algorithm. The simulation results show that the 3D reconstruction algorithm based on binocular stereo vision has the same resolution of the image sensor used, and the accuracy of the algorithm is determined by the spacing of the image sensor, the position, and shape of the constraint interval. The smaller the distance between the image sensors is, the larger the measurement radius is, but measurement accuracy will decrease with increasing measuring distance. The performance of a 3D reconstruction algorithm to recover an object depends on the ratio of the number of points that fall within the defined constraint interval to the total number of points in the boundary point of the object obtained by the algorithm, that is, the performance depends on constraint proportion. After using the interval constraint range algorithm, the size of the constraint proportion means the performance of the 3D reconstruction algorithm. The greater the constraint proportion is, the better the performance of the 3D reconstruction algorithm is.

**Acknowledgments.** Shubin Wang (wangshubin@imu.edu.cn) is the correspondent author and this work was supported by the National Natural Science Foundation of China (61761034, 61261020), and the Natural Science Foundation of Inner Mongolia, China (2016MS0616), and the Enhancing Comprehensive Strength Foundation of Inner Mongolia University (No. 10000-16010109-57).

## References

1. Xu G, Li X, Su J, Pan H, Tian G. Precision evaluation of three-dimensional feature points measurement by binocular vision. J Opt Soc Korea. 2011;15(1):30–7.
2. Wang W, Li HL. 3D reconstruction based on digital camera images. J Univ Shanghai Sci Technol. 2005;27(5):429–32.
3. Gao LF, Gai YX, Fu S. Simultaneous localization and mapping for autonomous mobile robots using binocular stereo vision system. In: 2007 IEEE international conference on mechatronics and automation; 2007. p. 326–30.
4. Zhu K. Real-time tracking and measuring of moving object based on binocular vision. Beijing: Beijing Jiaotong University; 2008. p. 1–103.
5. Jia T, Tu M, Jiang Y, Zhang S. 3D temperature distribution model based on vision method. In: 2016 IEEE international conference on robotics and biomimetics (ROBIO); 2016. p. 852–5.

# Improved Distributed Compressive Sensing Basing on HEVC ME and BM3D-AMP Algorithm

Zejin Li, Shaohua Wu(✉), Jian Jiao, and Qinyu Zhang

College of Electronic and Information Engineering, Harbin Institute of Technology, Shenzhen, Shenzhen 518055, Guangdong, China
lizejin@stu.hit.edu.cn
{hitwush, jiaojian, zqy}@hit.edu.cn

**Abstract.** The distributed compressive video sensing (DCVS) system greatly reduces the pressure on the encoder by transferring the computational complexity to the decoder, which is suitable for the limited-resource video sensing and transmission environment, in the meantime, get the better performance from key (K) frames and non-key (CS) frames. In this paper, we use the approximate message passing (AMP) algorithm reconstruct the K-frames. In order to improve the quality of the reconstructed K-frames, we add the block-matching 3D filtering (BM3D) denoising strategy based on the AMP algorithm. For the CS-frames, we improve the reconstructed CS-frames by improving the accuracy of side information (SI) frames by proposing a new high efficiency video coding (HEVC) motion estimation (ME) algorithm with motion vector (MV) prediction method. After we obtain the SI frames and combine the compressed value of the CS-frames with the side information (SI) fusion algorithm based on the difference compensation algorithm, the high accuracy SI frame is integrated into the reconstruction algorithm of the CS-frames. The experimental results demonstrate that our algorithms achieve higher subjective visual quality and peak signal-to-noise ratio than the traditional methods.

**Keywords:** Distributed compressive video sensing · High efficiency video coding · Approximate message passing · Block-matching 3D filtering

## 1 Introduction

With the developing of the Internet of Vehicles (IoV), the video application scene whose encoder is resource-limited is attracting more and more attentions. Conventional video codec standards such as MPEG series and H.26x standards do not apply to the scenario due to a high computational complexity caused by motion estimation (ME) [1]. Distributed compressive video sensing (DCVS)


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_58

[2] system can fill in the shortcomings of traditional video codec in the scenario. The DCVS system combines advantages of compressive sensing (CS) [3] and distributed video coding (DVC) [4]. The CS theory was first proposed by Cands and Donoho et al. in 2004 which break the limits of the Nyquist sampling rate and improve the compression efficiency of the encoder. The DCVS system successfully reduces the complexity of the encoder by moving the ME to decoder. In 2009, T. T. Do et al. also introduced the theory of CS into DVC and proposed distributed compressed video sensing (DISCOS) [5]. Soon, Kang L. W. et al. proposed the distributed compressive video sensing (DCVS) [1] system. The DCVS use CS technique both in the key (K) frames and the non-key (CS) frames. The reconstruction algorithm of K-frames and CS-frames is mainly divided into convex optimization algorithms and iterative algorithms based on greedy thinking. The convex optimization algorithms such as gradient pursuit for sparse reconstruction (GPSR) [6] algorithm feature a small reconstruction error, and the latter algorithm such as approximate message passing (AMP) [7] algorithm is widely used as CS reconstruction method due to its computational efficiency. Later, there appears the AMP algorithm based on denoising method. In this paper, we choose the block-matching 3D filtering (BM3D) [8] which is one of the best denoising algorithms based on the AMP algorithm to improve the quality of video frames. For CS-frames, many scholars work to improve the quality by obtaining more accurate side information based on traditional motion estimation. Reference [9] used the log search algorithm as judging criteria to estimate motion vector (MV) and proposed the enhanced side information (ESI) algorithm. Reference [10] found the best matching block as the center of the new searching window by expanding the search window size. In our works, we proposed a new high efficiency video coding (HEVC) ME algorithm with motion vector (MV) prediction method which makes use of detail information and motion correlation of the video frames.

In this paper, the overall structure of the article is following. Section 2 introduces compressive sensing theory and the DCVS system. Section 3 presents the works to improve the quality of the decoder. Section 4 displays our results and analysis in our experiment. The last section is the conclusion.

## 2 Compressive Sensing and the DCVS Framework

### 2.1 Compressive Sensing and the AMP Algorithm

A frame X in a video sequence is said to be sparse in the transform domain $\Psi$ if it can be represented by the transform coefficients $s$ ($x = \Psi s$) which are most zeros or close to zero. The video frame X is first expressed by a vector $x \in R^{N\times 1}$and then compressed by through the compression matrix $\Phi_{M\times N}(M \ll N)$. The CS reconstruction process is to recovery $x$ from $y$. The decoding process can be described as Eq. (1),

$$\min \|s\|_1 \; s.t. y = \Phi x = \Phi\Psi s, \tag{1}$$

Iterative framework is popular in CS theory for its simple iterative process and low computational complexity, such as AMP, iterative hard thresholding

(IHT) [11], etc. We choose the fast iterative framework, AMP, as the reconstruction algorithm for both K-frames and CS-frames. The AMP algorithm solves compressive sensing through the iterative method:

$$x^{t+1} = \eta_t(A^* z^t + x^t), \tag{2}$$

$$z^t = y - Ax^t + \frac{1}{\delta} z^{t-1} \left\langle \eta_t^{'}(A^* z^{t-1} + x^{t-1}) \right\rangle, \tag{3}$$

where $A = \Phi\Psi$ or $A = \Phi$ is the compression matrix, $A^*$ is the transposition of $A$, the function $\eta_t(.)$ is defined as: $\eta_t(x) = (|x| - \tau) + sign(x)$, where $sign(.)$ is the sign function. $\eta_t^{'}(.)$ is the derivative of $\eta_t(.)$.

### 2.2 The DCVS Framework

The DCVS system is characterized by low computational complexity and high efficiency. The framework of the DCVS is shown in Fig. 1. At encoder, the video sequence is divided into group of pictures (GOPs). The first frame of a GOP is K-frame, and others are CS-frames. Both K-frames and CS-frames are compressed by CS method at encoder. We defined the compression ratio of a frame $x_t$ as $CR = M/N$, where $M$ is the number of pixels in compressed frames and $N$ is in original frames. At the decoder, the K-frames are directly reconstructed by approximate message passing (AMP) algorithm-based BM3D denoising strategy. Then, the reconstructed K-frames of the adjacent GOPs are used to generate SI frames by HEVC ME algorithm. To make full use of the side information (SI) frames and the compression result of CS-frames, the difference compensation algorithm is exploited to reconstruct CS-frames.

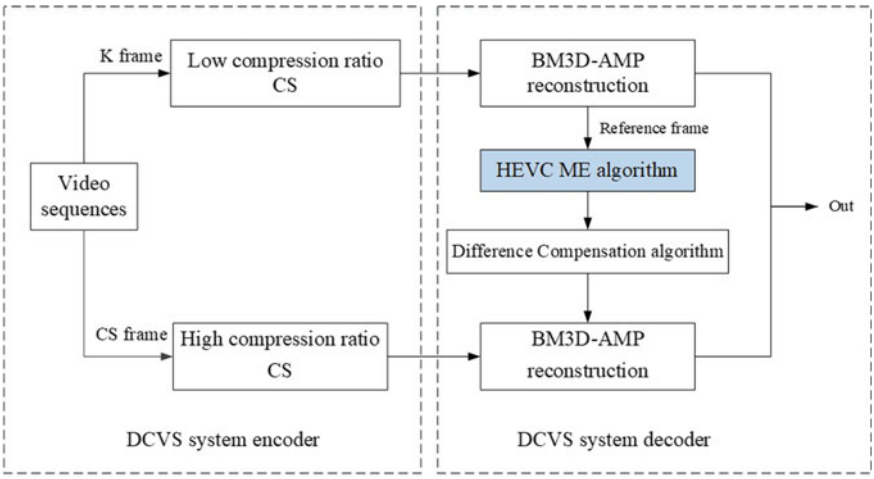



**Fig. 1.** DCVS system.

## 3 Improved DCVS System

### 3.1 The AMP Algorithm-Based BM3D Model

In order to improve compressed video frames recovery, the denoiser methodology is proposed into AMP algorithm. We utilize the AMP algorithm-based BM3D model in our research. The BM3D model is one of the best denoising algorithms in image processing which exploits the non-local self-similarity among pixels in global positioned patches. At the beginning of the process, BM3D utilizes non-local means (NLM) ideas to divide the video frames into several overlapped or non-overlapping blocks and find the smallest distance block based the Euclidean distance in a window of a certain size as a similar block. The second operation is to group similar blocks into stacks and then perform 3D transform on the group. The inverse transform is performed after collaborative filtering. And then the algorithm performs coefficient aggregation to obtain the final estimate of video frame. The algorithm is divided into two major steps as Fig. 2 shows: The first step estimates the denoised frames using hard thresholding during the collaborative filtering. The second step is based both on the original noisy image and on the basic estimate obtained in the first step by using Wiener filtering.

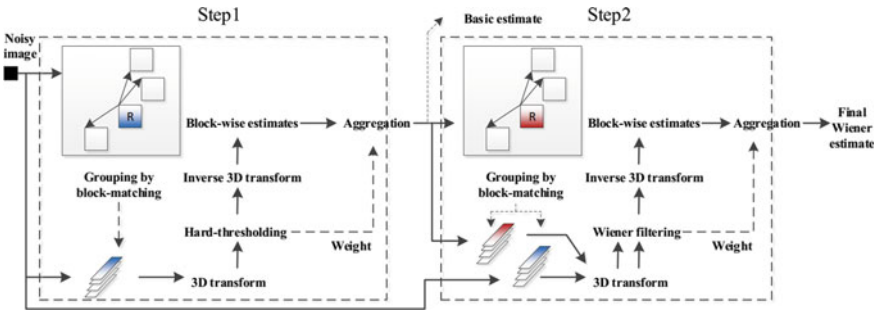



**Fig. 2.** Scheme of the BM3D algorithm.

### 3.2 HEVC ME Algorithm-Based MV Prediction Algorithm

As to further improve the reconstructed CS-frames, we focus on the side information generation method for CS-frames based on the two adjacent K-frames and proposed HEVC ME algorithm. As Fig. 4 shows, compared with the traditional motion estimation algorithm [12], the HEVC motion estimation method takes advantages of the detailed information by using unequal sizes of coding units (CUs) which like a quadtree partitioning method to fulfill motion estimation in the reference frames. The HEVC ME algorithm is mainly divided into four steps: Firstly, the video frames are divided into CUs with size of $8 \times 8$. Then, we perform motion estimation for every CU and get the motion vector (MV). The core

operation of motion estimation is block-matching algorithm which finds the most similar CU in the reference frames of current CU. The criterion to decide which CU is the most similar CU is SATD function, which is also called Hadamard transform. The function calculates the difference between the original CU and the candidate matched CU into a matrix and then multiplies the Hadamard matrix on both side. Thirdly, we expand the size for two times longer than current CU and operate the same motion estimation and choose the smaller the SATD result of two motion estimation process. Finally, continue to expand the CU size and repeat motion estimation, we find the most smallest SATD result and the corresponding MV pointing to the best matching CU. In addition to the choice of CU size, the choice of search scope will also affect the effect of motion estimation. We proposed an MV prediction algorithm as shown in Fig. 3 to select a more accurate search range for CUs in the reference frame. Because the motion between adjacent CUs in the same video frame has a correlation, the algorithm uses the MVs of the CUs around the current CU that has obtained the motion estimation results and determined the search range. The method is to find the same position in the reference frame as the start coordinate position of the current CU at first and then compensate the MVs of the neighboring CUs onto this position. Therefore, we obtain multiple different sets of search scopes. The best matching CU is corresponding to the smallest SATD function value when performing motion estimation search within these search ranges.

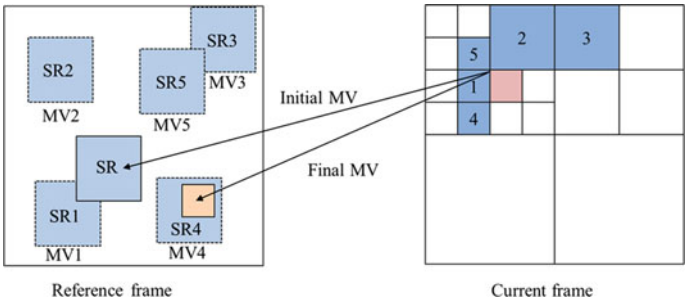



**Fig. 3.** Motion vector prediction algorithm.

## 4 Results and Discussion

In order to prove the advantages of BM3D-AMP, we compare the reconstruction quality of BM3D-AMP and GPSR algorithm in which we used in the early days. Foreman video sequences with BMP format ($144 \times 176$) are used for testing. We test the first 25 frames of the video sequence. The GOP size is set as 2, in other words, the first frame of a GOP is K-frame and the other one is CS-frame.

Except for the reconstruction algorithm, other simulation conditions are the same. The recovery results of video frames are evaluated by peak signal-to-noise ratio (PSNR).

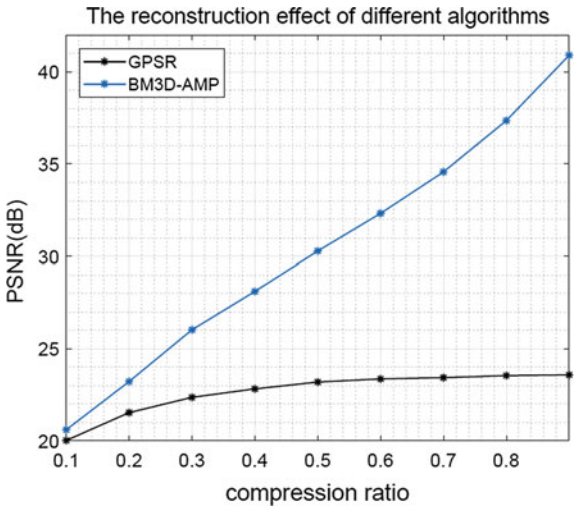



**Fig. 4.** Reconstruction effect of different algorithm.

The recovery quality of the reconstructed K-frames using BM3D-AMP and GPSR algorithm is represented in Fig. 4. The reconstruction quality of BM3D-AMP is greatly improved than GPSR algorithm. The improvement of reconstruction quality is mainly due to the addition of the BM3D denoising method that filters out noise interferences during reconstruction. Simulation results in Fig. 5 express the overall performance of the DCVS system compared to the GPSR and BM3D-AMP algorithm. HEVC ME algorithm is utilized to generate SI frames in every simulation. Using BM3D-AMP can obtain 1.5 2 dB higher than GPSR. The reconstruction effect of DCVS system has greatly improved through the improvement of recovery quality of every frame. Figure 6 indicates the simulation results about the superiority of HEVC ME over conventional ME. The reconstruction algorithm of video frames is BM3D-AMP. Only the algorithms for generating SI frames are different. The PSNR result of using the HEVC ME algorithm can improve DCVS system reconstruction about 5 dB compared to using the GPSR algorithm.

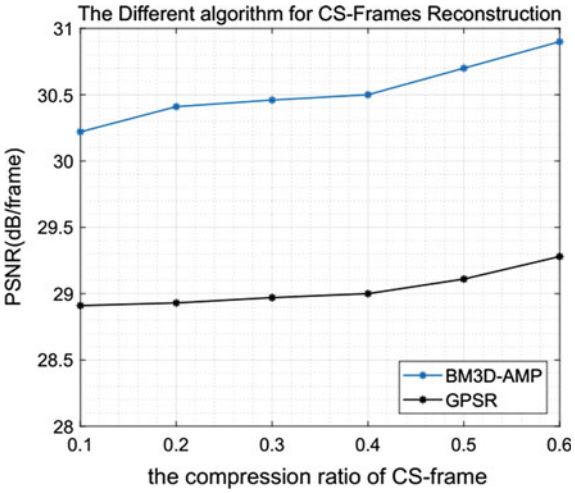



**Fig. 5.** Different algorithm for CS-frames reconstruction.

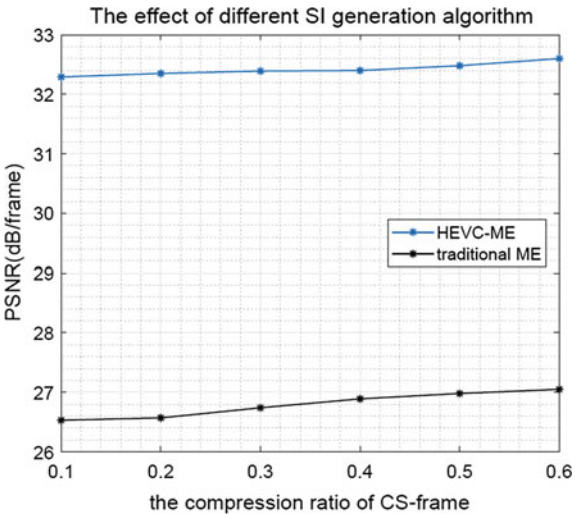



**Fig. 6.** Effect of different SI generation algorithm.

## 5 Conclusion

In this paper, we used AMP algorithm-based BM3D denoising strategy for the reconstruction algorithm for video frames to improve the performance of the DCVS. We also proposed a new HEVC motion estimation algorithm to generate more precise side information based on motion estimation prediction algorithm. The algorithm improves the quality of reconstructed CS-frames.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China under Grant 61371102, 61771158 and in part by the Shenzhen Municipal Science and Technology Plan under Grant JCYJ20170811160142808, JCYJ20170811154309920.

## References

1. Schwarz H, Marpe D, Wiegand T. Overview of the scalable video coding extension of the H. 264/AVC standard. IEEE Trans Circuits Syst Video Technol. 2007;17:1103–20.
2. Kang L-W, Lu C-S. Distributed compressive video sensing. IEEE Int Conf Acoust Speech Signal Process. 2017;7744:52–66.
3. Donoho DL. Compressed sensing. IEEE Trans Inf Theory. 2006;52:1289–300.
4. Girod B, Aaron AM, Rane S, Monedero DR. Distributed video coding. Proc IEEE. 2005;93:71–83.
5. Do TT, Chen Y, Nguyen DT, Nguyen N, Gan L, Tran TD. Distributed compressed video sensing. IEEE Int Conf Image Process. 2009;11:1393–6.
6. Figueiredo MA, Nowak RD, Wright SJ. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. IEEE J Sel Top Signal Process. 2007;1:586–97.
7. Donoho DL, Maleki A, Montanari A. Message-passing algorithms for compressed sensing. Proc Nat Acad Sci. 2009;10:18914–9.
8. Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans Image Process. 2007;16:2080–95.
9. Angayarkanni V, Akshaya V, Radha S. Distributed compressive video coding using enhanced side information for WSN. In: IEEE international conference on wireless communications, signal processing and networking, 2016. p. 1133–6.
10. Chen J, Xue F, Kuo Y. Distributed compressed video sensing based on key frame secondary reconstruction. Multimedia Tools Appl. 2017;1–17.
11. Blumensath T, Davies ME. Iterative thresholding for sparse approximations. J Fourier Anal Appl. 2008;14:629–65.
12. Cuevas E, Osuna V, Oliva D. Motion estimation. Berlin: Springer International Publishing; 2017. p. 95–116.

# Research on Consensus Mechanism for Anti-mining Concentration

Kailing Sui, Chaozhi Yang, and Zhihuai Li(✉)

Information college, Dalian Maritime University, Dalian 116000, China
{suikailing,qhlee}@dlmu.edu.cn

**Abstract.** Based on the analysis of the existing consensus mechanism of blockchain, this paper proposes a POWS mechanism based on the adjustment of workload. In the POW consensus mechanism, the concept of coinage is introduced to adjust the mining difficulty of different nodes. The POWS adjusts the mining difficulty through two factors: calculation force and coinage. The POWS is compared with the POW and the POS in two aspects, the basic performance and the ability to resist the mining pool. The experimental results show that the POWS consensus mechanism can meet the performance requirements of normal blockchain system, at the same time it can better diminish the impact of calculation force and coinage on block generate efficiency, narrow the efficiency gap between the mining pool nodes and the non-mineral pool nodes, and reduce the interest of the mining pool to non-mineral pool nodes.

**Keywords:** Blockchain · Consensus mechanism · Mining pool · POW · POS · POWS

## 1 Introduction

Blockchain technology is a decentralized recording technology [1], and it can also be called a distributed general ledger technology based on P2P network [2]. Blockchain technology is originally applied to decentralized bitcoin electronic cash transaction system [3], whose main features are decentralized and detrusting. The most harmful and significant technical defect of bitcoin is that the consensus mechanism itself can easily cause the man-made centralization of mining pools [4]. That has violated the original intention of blockchain technology, becoming the main problem of the development of blockchain technology.

One of the core technologies of blockchain technology is consensus mechanism technology [5]. The consensus mechanism solves the problem of credible communication in peer-to-peer networks, named the Byzantine failures [6]. There are two main types of consensus mechanism in the employment of the public chain:


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_59

One is the POW consensus mechanism [7], and the other is the POS consensus mechanism [8]. The POW consensus mechanism will artificially generate the mining pool with the concentration of calculation force, while the POS consensus mechanism will produce the pool with the concentration of coinage. Therefore, to improve the consensus mechanism is the primary method of solving the problem of concentration of mining pools.

Based on the analysis of the existing consensus mechanism, this paper proposes a hybrid consensus mechanism of POW+POS called proof of work based on adjusted stake, or POWS. The POWS consensus mechanism is designed to prevent the concentration of the mining pool, increase the operation cost of the mining pool effectively, and reduce the interest drive of the mining pool to the common miner node.

Experiments to compare the basic performance and the ability to resist the mining pool of the POWS consensus mechanism, the POW consensus mechanism, and the POS consensus mechanism were conducted in a laboratory environment. The experimental results show that the POWS consensus mechanism can meet the performance requirements of normal blockchain system, well reduce the impact of the calculation force or coinage on the block generation efficiency, and weaken the interest drive of mining pool on common nodes.

## 2 Design of POWS Consensus Mechanism

The consensus on the whole network is one of the core challenges in distribution systems [9]. With the idea of a hybrid consensus mechanism of POW+POS, this paper introduces the concept of equity (coinage) in the POS consensus mechanism to adjust the mining mode of the POW consensus mechanism and designs a workload proof mechanism based on equity adjustment called the POWS [10].

### 2.1 Overall Structure and Node Design of POWS

**Overall structure design of the POWS consensus mechanism** The POWS consensus mechanism system designed in this paper consists of four parts: application layer, extension layer, network layer and storage layer. The function of the application layer is to provide users with specific block links and to meet the specific needs of users. The most representative applications include wallet client, trading platform, communication software, and so on; the function of the extension layer is to introduce the side chain applications such as intelligent contract and lightning network. The function of the storage layer is to store complete blockchain data; the network layer realizes the consistency of the local blockchain data of all nodes through the consensus mechanism. The consensus mechanism adopted in the network layer is the POWS consensus mechanism.

**Synchronization of node communication and block** When a connection is established between nodes, the node sending the request first sends its system version information to the corresponding node. The corresponding node will also send its system version information to the newly added node after receiving the

information. The two nodes which have established communication connection will both confirm each other's time stamp. If their time stamps are corresponded, they will send back the confirmation and establish the communication connection successfully; otherwise, the communication connection will not be established. When the newly added node in the system maintains the communication connection with other nodes, the block information can be synchronized. The blockchain system will provide all nodes with the function of searching the mining node in the current P2P network, so as to record the IP address of the corresponding node and conduct block information synchronization and mining. In the system, only full nodes can participate in mining and maintaining blockchain system. Light nodes are only using the blockchain system, rather than participating in maintaining the blockchain system.

### 2.2 POWS Decentralized Consensus Process

The consensus mechanism designed by Dorian S. Nakamotoi is a mechanism which operates spontaneously under decentralized conditions [11]. The charac-

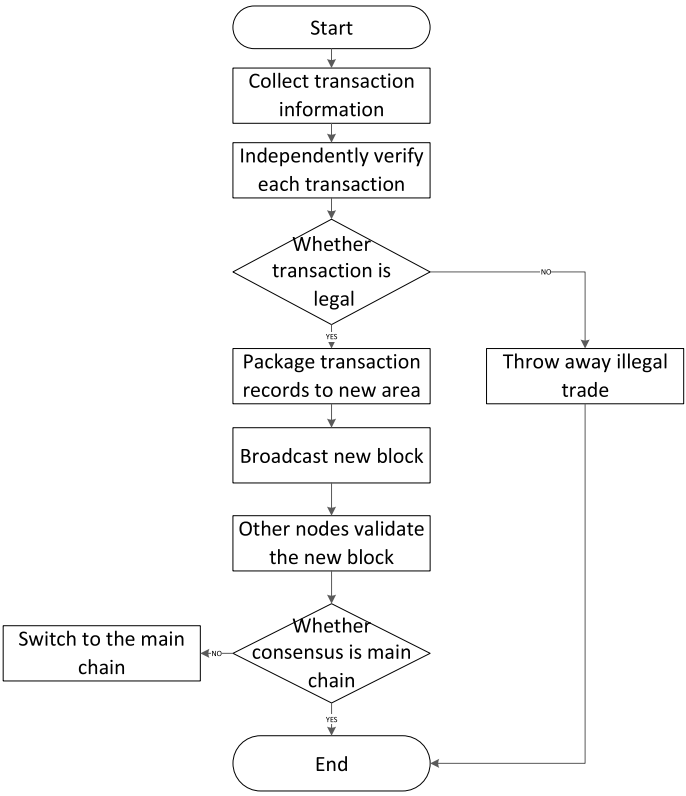



**Fig. 1.** POWS consensus process

teristic of the mechanisms operating spontaneously is embodied in the accomplishment of asynchronous consensus when the explicit election has not been done or the consensus cannot be reached, that is to say each individual node in blockchain system under the rule of established program interacts asynchronous and form consensus spontaneously. The process of the POWS consensus mechanism in this paper is as shown in Fig. 1.

### 2.3 POWS Transaction

**Independent verification of POWS transaction** Blockchain wallet will construct transaction information through a series of operations including collecting transaction input UTXO application system, providing the correct UTXO unlock script, and creating UTXO output used by those who output in the transaction to receive information. Then, the wallet sends the transaction log to the other nodes in the blockchain application system, allowing the transaction to circulate in the blockchain network and to be added to the miner node's transaction pool.

Each miner node needs to independently verify the transaction after receiving the blockchain transaction record. After receiving the transaction record, the miner node will inspect the transaction information independently according to a certain standard. After successful inspection, the miner node will continue to broadcast the transaction record to the whole network, and add the transactions which have been independently verified to the transaction pool in the order that they are received.

**Independent packaging of POWS transactions** The miner node saves all the transaction records that are verified in the local transaction pool. When the miner node packages the transaction records in the transaction pool to the candidate block, the priority order is decided by the currency coinage of UTXO. The calculation formula of the coinage of a single UTXO is as shown in Formula 1.

$$Coinage = Value * InputAge \tag{1}$$

The transaction priority is acquired when the total UTXO coinage divides the total length of the transaction. The Priority is as shown in Formula 2.

$$\text{Priority} = \sum \text{CoinAge}/\text{TransactionSize} \tag{2}$$

where TransactionSize represents the total length of the transaction. The unit of UTXO that trades in is cong, the unit of age is the block number, the unit of coinage is the byte, and the unit of the transaction record is the byte.

**POWS coinbase transaction** Miner nodes must create a coinbase transaction to package into the candidate block before packaging the transaction records in the transaction pool into candidate blocks. The POWS designed in this paper improves the function of coinbase transaction, increasing the function of clearing coinage which is the sum of all UTXO's coinage in the blockchain address of the miner node. The calculation formula of total coinage is as shown in Formula 3.

$$\text{SumCoinAge} = \sum_{i}^{n} (\text{Value} * \text{InputAge}) \tag{3}$$

The UTXO value of the coinbase transaction output is the sum of the UTXO values entered, the total transaction costs of all other transactions in the block, and the sum of three awards of the block miners entered. The temporary bonus of the POWS is 1. The coinbase transaction of POWS has three roles: The total coinage using to participate in mining is cleared; transaction fees in the block are paid to miner nodes; the reward for generating the block is paid to the miner node.

**The generation and mining algorithm of POWS block** The block is mainly divided into the blockhead and the block body. And the main record of the block body is the transaction record. The blockhead of the miner node needs to package the segmentation field, the length of the block, equities record, transaction counter, transaction log details, and so on, generating the candidate block. The miner nodes need to try to calculate a different random number Nonce to find the appropriate random number Nonce, making the calculated block hash values conform to the rules, which means that the hash value of the block is less than the target difficultly. The POWS consensus mechanism generates blocks for the hash calculation to find the specific process of the conditional block hash value as shown in algorithm 1.

*Proof of Work Based on Adjusted Stake (POWS)*

```
max_nonce = 2 ^ 32
//The upper limit of the random number is about 4 billion.
def proof_of_work_based_on_adjusted_stake (header,SumCoinAge):
        target=coefficient*2^(8*(exponent-3))*(SumCoinAge)^(1/2)
        for nonce in xrange(max_nonce):
            hash_result=hashlib.sha256(str(header)+
            str(nonce)).hexdigest()
            if long(hash_result,16)<target:
            return(hash_result,nonce)
        return nonce
if __name__=__main__:
    nonce=o
    hash_result=
    start_time=time.time()
    //Verification time stamp
    new_block=test block with transactions+hash_result
    (hash_result,nonce)=proof_of_work_based_on_adjusted
    _stake(new_block,SumCoinAge)
    //Find the random number nonce in the new bolck.
    end_time=time.time()
    //End time
    Eta=end_time-start_time
```

```
//Calculate the time to create blocks
if eta>0
    hash_power=float(long(nonce)/eta)
    //Calculate the hash rate.
```

The POWS consensus mechanism introduces the concept of the total coinage and the concept of miner node to adjust the target in the calculation formula. The calculation of the target value is shown in Formula 4.

$$\text{Target} = \sqrt{\text{SumCoinAge}} * \text{cofficient} * 2 \wedge (8 * (\text{exponent} - 3)) \tag{4}$$

The parameter SumCoinAge in Formula 4 is the sum of all UTXO's coinage in the address of the miner node, which is used to adjust the mining difficulty; parameter exponent and parameter coefficient indicate the difficulty of mining.

## 3 Experimental Result

The experiment is constructed by constructing the P2P network and simulating blockchain network in the laboratory server. The version information of the major software used in the experiment is as follows: Bitcoin Core version is v14.02; Peercoin version is v0.6.1; Qt Creator version is 3.1.0; the Berkeley DB version is 6.4.9. The data set used in the load test was the transaction data of the bitcoin system. The experiment selected all transaction data of the bitcoin system from the creation block to the 495,000 blocks. And the transaction data is slightly normalized to accommodate the transaction structure of the other two blockchain consensus mechanisms.

### 3.1 Average Block Time

The average block time of the three consensus mechanisms is as shown in Fig. 2. The average length of the POW consensus mechanism is the longest. The average time of the POWS consensus mechanism is in the middle level; the average time of the POS consensus mechanism is the shortest. The average block time of the three blockchains is relatively stable.

### 3.2 Anti-load Comparison Experiment

As shown in Fig. 3, under the condition of heavy concurrency, the block size limit of the POW consensus mechanism is 1M, POWS consensus mechanism changes the size limit to 4M, so the POWS consensus mechanism of block has stronger ability to deal with transaction. The transaction handling capacity of the three blockchain consensus mechanisms decreases with the increase of transaction concurrency.

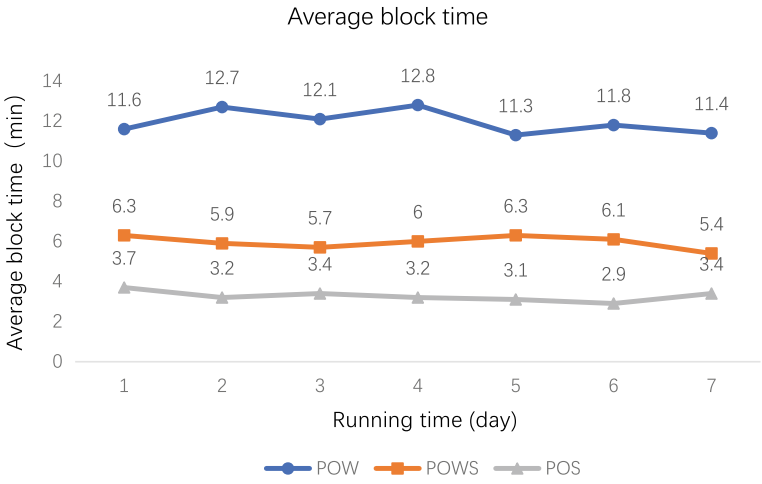



**Fig. 2.** Average time to make block

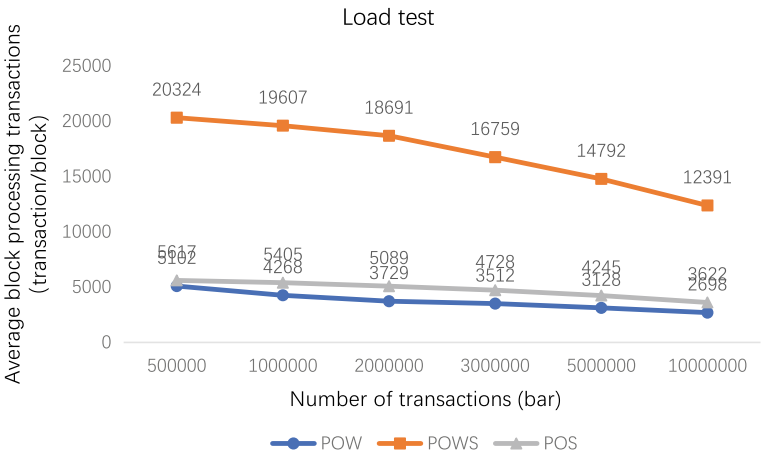



**Fig. 3.** Load experiment

### 3.3 Comparison Experiment of Average Block Efficiency with the Change of Calculation Force

As shown in Fig. 4, the effect of the proposed POWS consensus mechanism is less than that of the POW consensus mechanism. The reason for this phenomenon is that the POWS consensus mechanism limits more in coinage compared with the POW consensus mechanism.

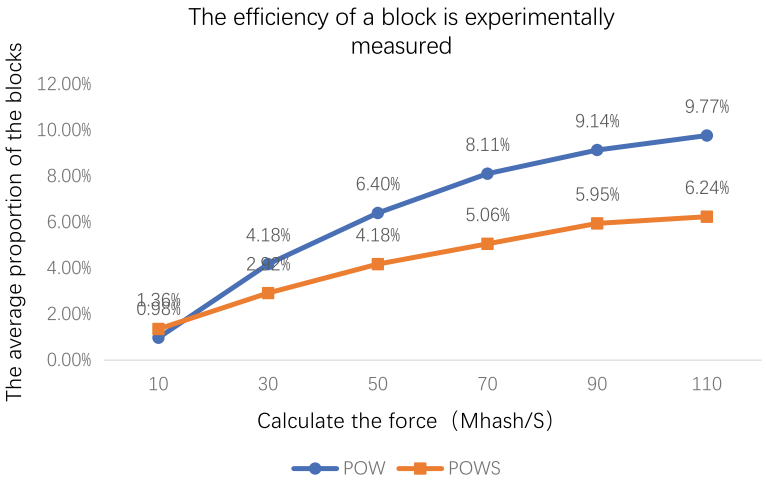



**Fig. 4.** Block efficiency and calculation of power experiment

### 3.4 Comparison Experiment of Average Block Efficiency with the Change of Coinage

As shown in Fig. 5, the effect of the proposed POWS consensus mechanism is less than that of the POS consensus mechanism. The reason for this phenomenon is that the POWS consensus mechanism is more powerful in calculation force compared with the POS consensus mechanism.

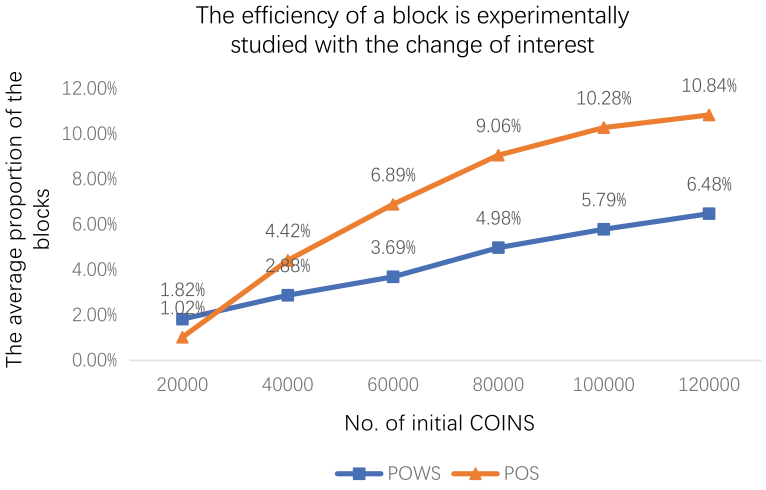



**Fig. 5.** Block efficiency and stake experiment

### 3.5 The Comparison Experiment on the Profit Driving of Non-mineral Pool Nodes

As shown in Fig. 6, in this paper, in the designed POWS consensus mechanism, the average efficiency ratio of non-mineral pool nodes and miner nodes is highest. The average efficiency ratio of the non-mineral node and miner node in the POW consensus mechanism and the POS consensus mechanism decreases with the increase of the ore pool calculation force or the coinage.

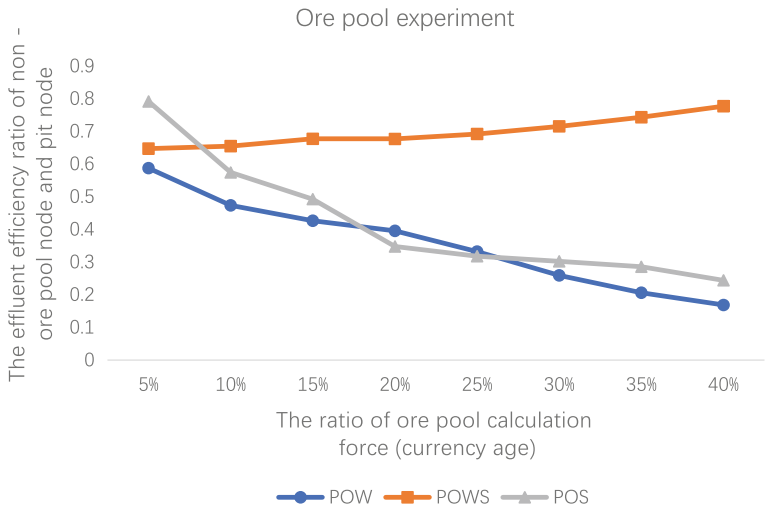



**Fig. 6.** POWS mine pool experiment

## 4 Conclusion

Based on the proposed POWS consensus mechanism, this paper carried out an experiment from two aspects: the basic performance and the ability to resist the mining pool, and reached the following conclusions:

(1) Compared with the POW consensus mechanism and the POS consensus mechanism, the POWS consensus mechanism has a higher average efficiency ratio of non-mineral pool nodes and mining pool nodes, and the mining pool has less interest drive to the non-mineral pool nodes.
(2) The average block efficiency of the POWS consensus mechanism is slowing down with the growth of the calculation force or the age of the currency.

The above conclusion proves that the POWS consensus mechanism proposed in this paper is reasonable and feasible, and the mechanism can achieve the goal of effective inhibition of the concentration of the mining pool.

## References

1. Hurlburt G. Might the blockchain outlive bitcoin? IT Prof. 2016;18(2):12–6.
2. Janusz J. Blockchain technology in the chemical industry: Machine-to-machine electricity market. Appl Energy. 2017;195:234–46.
3. Xu X, Pautasso C, Zhu L, et al. The blockchain as a software connector. In: Software architecture; 2016; New York: IEEE; p. 182–91.
4. Schrijvers O, Bonneau J, Dan B, et al. Incentive compatibility of bitcoin mining pool reward functions, 2016.
5. Xia Q, Zhang FJ, Zuo C. Review for consensus mechanism of cryptocurrency system. Comput Syst Appl, 2017.
6. Dolev D, Lamport L, Pease M, et al. The Byzantine generals [C]. In: Concurrency control & reliability in distributed systems; 1987.
7. Aste T. The fair cost of bitcoin proof of work. Social Science Electronic Publishing; 2016.
8. Spasovski J, Eklund P. Proof of stake blockchain: Performance and scalability for groupware communications. In: The international conference on management of digital ecosystems; 2017.
9. Kosba A, Miller A. Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In: 2016 IEEE symposium on security and privacy, San Jose, California, USA; 2016.
10. Bitcoin Wiki. Proof of Stake [EB/OL]. https://en.bitcoin.it/wiki/Proof_of_Stake, 2015
11. Göbel J, Krzesinski AE, Keeler HP, et al. Bitcoin blockchain dynamics: The selfish-mine strategy in the presence of propagation delay. Perform Eval. 2015;104(1):23–41.

# Application of Multitaper Method (MTM) to Identify the Heterogeneity of Karst Aquifers

Baoju Zhang[1(✉)], Lixing An[1], Wei Wang[1], Yonghong Hao[2], and Yi Wang[1]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
wdxyzbj@163.com
[2] Tianjin Key Laboratory of Water Resource and Environment, Tianjin Normal University, Tianjin 300387, China

**Abstract.** Karst aquifers supply drinking water for 25% of population on the earth. Better understanding of the heterogeneity of karst aquifers can help us develop karst groundwater sustainably. Karst hydrological processes include precipitation (rainfall) infiltration, groundwater wave propagation in karst aquifers, and spring discharge. The processes of precipitation signals' transformation into spring discharge signal are mainly affected by heterogeneity of karst aquifers. Analysis of relations between spring discharge and precipitation can identify heterogeneity of karst aquifers. This paper explores the periodic characteristics of spring discharge and precipitation, and the heterogeneity of karst aquifers in Niangziguan Springs (NS) using multitaper method (MTM). The results show that both spring discharge and precipitation exist in the same period of one year. Cross-correlation function is used to calculate the time lags between the reconstructed spring discharge and precipitation in different areas of the NS basin. The results indicate that the response time of the spring discharge to precipitation is different at different areas. The time lag between the spring discharge and precipitation is 3 months at Pingding County; 4 months at Yu County, Yangquan City, Xiyang County, and Heshun County; and 27 months at Shouyang County and Zuoquan County. The results reflect the heterogeneity of the NS basin and are consistent with the geological structure of the NS basin. MTM is robust in identification of heterogeneity of karst aquifers.

**Keywords:** Multitaper method (MTM) · Cross-correlation function · Heterogeneity · Time lag

## 1 Introduction

Multitaper method (MTM) is a spectral estimation and signal reconstruction method, which is especially suitable for the diagnosis of quasi-periodic signals in short-sequences and high-noise environment. MTM was originally proposed by Thomson in 1982 when analyzing seismic data [1]. In this method, a low-variance, high-resolution spectral estimation is obtained by combining the best window function with multitaper


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_60

smoothing. Based on MTM, Park [2] established signal reconstruction techniques by time domain inversion of multitaper spectral decomposition information, which can simulate the evolution of the amplitude and phase of quasi-periodic signals over time. Recently, Mann and Park [3] combined this method with multivariate statistical methods and proposed a frequency domain-singular value decomposition method.

Due to the best trade-off between variance and spectral resolution, MTM is suitable for the diagnosis of weak signals and time–space-dependent signals in nonlinear climate systems. MTM has been applied to geophysical signal analysis to solve practical problems. Speech enhancement [4] and cognitive radio based on MTM [5] have also been applied.

Karst terrain occupies 10% of the earth's continental surface. It characterizes by sinkholes, conduits, and caves, and provides large spaces for groundwater storage. Karst groundwater is the main water resource for drinking, irrigation, and industrial and municipal usage. The main source of karst groundwater recharge is precipitation. Because karst aquifer is highly heterogeneous with hierarchical porous media, the groundwater wave propagates through conduits will be faster than that through pores. In other words, time lags between spring discharge and precipitation at the fracture and conduit zones will be shorter than in pore zones. Therefore, the heterogeneity of karst aquifer is reflected by the lags between spring discharge and precipitation, and then helps us identify the structure of karst aquifer.

In this study, the purpose is to identify the heterogeneity of karst aquifer taking Niangziguan Spring (NS), China, as a case study. Firstly, the periodicities of precipitation and spring discharge are detected via MTM. Subsequently, the time lag between spring discharge and precipitation is estimated. Finally, the heterogeneity structure of the karst aquifer is analyzed.

## 2 Method

### 2.1 MTM Spectral Estimation

MTM can be used to estimate the singular components and the continuous component of the spectrum. Once the tapers $w(t)$ are calculated for a selected frequency bandwidth, the total power spectrum can be obtained by averaging the individual spectra given by each tapered version of the data set [6]. If $K$ orthogonal windows are added to the data, $K$ eigenspectrum are obtained.

$$s(f) = \frac{\sum_{k=1}^{k} \mu_k |Y_k(f)|^2}{\sum_{k=1}^{k} \mu_k} \tag{1}$$

where $\mu_k$ is the weights, $Y_K(f)$ is the discrete Fourier transform (DFT) of $X(t)_{W_K(t)}$, and $|Y_K(f)|^2$ is the $k$th eigenspectrum. It can be seen that the high-resolution multitaper spectrum is the weighted sum of the $K$ eigenspectra.

### 2.2 Quasi-periodic Signal Detection

Geophysical signals are often quasi-periodic and have a strong spectral component in the low-frequency part and present a strong red noise background. For spectrum estimation based on MTM, Mann and Lees [7] proposed a signal detection method with strong red noise background.

In narrow sense, the statistics model of red noise can be seen as a first-order autoregressive {AR (1)} process. The power spectrum of the AR(1) process is written by [7]:

$$S_r(f) = \varepsilon_0 \frac{1 - r^2}{1 - 2r\cos(2\pi f / f_N) + r^2} \tag{2}$$

where $\varepsilon_0$ is the mean of the power spectrum, while $r$ is the lag-one autocorrelation, and $f_N$ is the Nyquist frequency. Using the sampling distribution theory of spectrum estimation, the critical values of red noise spectrum with different confidence limits are obtained.

### 2.3 Signal Reconstruction of Multitaper Spectrum

Once significant oscillation signals are detected in multitaper spectrum estimation, the information from the multitaper decomposition can be used to reconstruct the related signals [6]. The reconstructed signals have a slowly changing envelope characteristic. The reconstructed signal is given at a peak-centered frequency $f_0$:

$$X(t) = \vartheta\{A(t)\mathrm{e}^{-i2\pi f_0 t}\} \tag{3}$$

or, for the discrete case at hand,

$$X(n\Delta t) = \vartheta\{A_n\mathrm{e}^{-i2\pi f_0 n\Delta t}\} \tag{4}$$

where $A(\mathrm{t})$ and $A(n)$ are the envelope functions, and $\vartheta$ is the real part of the Fourier term.

### 2.4 Cross-correlation Function

A method is used to estimate correlated degree of two series, {$X(i)$ and $Y(i)$}, termed cross-correlation. It can be used to characterize time domain structure with a time-shifted version of two different sequences. The cross-correlation $\rho$ at delay m is written:

$$\rho(m) = \frac{\sum_{i=1}^{N-m}(X_i - x)(Y_{i+m} - y)}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - x)^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - x)^2}} \tag{5}$$

where $x$ and $y$ are the average value of the corresponding series, and N is the value of the series length.

## 3 Study Area and Data

The NS, the largest karst springs in northern China, is located in the Mian River valley, eastern Shanxi province, China. The administrative districts involved include the counties of Yuxian, Shouyang, Pingding, Heshun, Zuoquan, Xiyang, and Yangquan city (Fig. 1). The main aquifers of the NS basin are composed of Ordovician karstic limestone, porous sediments, and Quaternary sandstone. The groundwater in karst aquifers flows toward the Mian River valley in the east, where groundwater perches on low-permeable strata of Cambrian dolomicrite, and eventually intersects the ground surface, thus emerging as the NS.

The monthly precipitation data of the NS is given by the seven meteorological stations in the administrative districts of the NS basin. The stations have been successively built since the 1950s, and the station in Pingding County was built at the last.

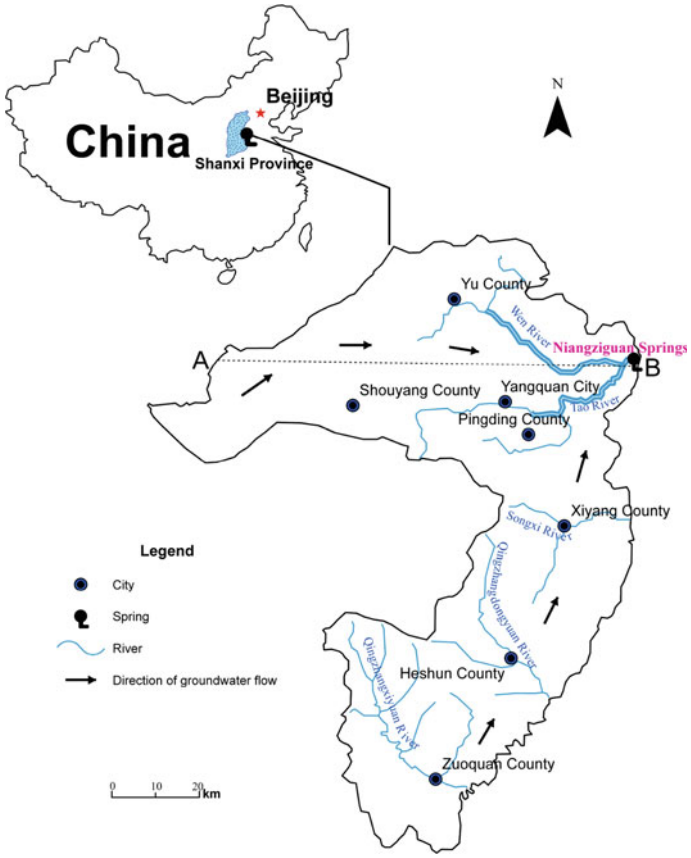



**Fig. 1.** Map of NS basin

In order to make full use of the data from rainfall stations, the calculation of precipitation in NS basin will be carried out in sections: from 1959 to 1971 (six meteorological stations except Pingding station) and from 1972 to 2015 (all seven meteorological stations). Thiessen polygon method is used to obtain the precipitation weight of Niangziguan Spring, and the average monthly precipitation of the NS basin is obtained from 1959 to 2015 (Fig. 2). The monthly spring discharges of the NS basin from January 1959 to December 2015 are obtained from the NS station (Fig. 3).

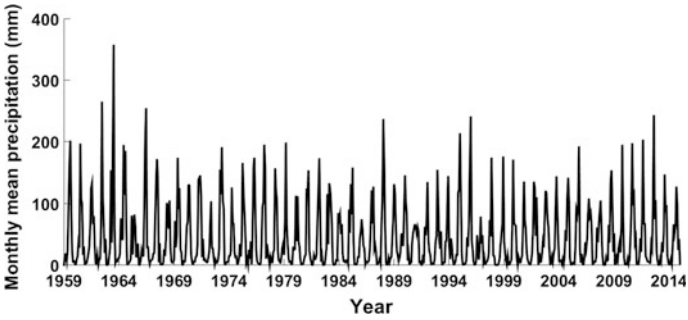



**Fig. 2.** Monthly mean precipitation in the NS basin

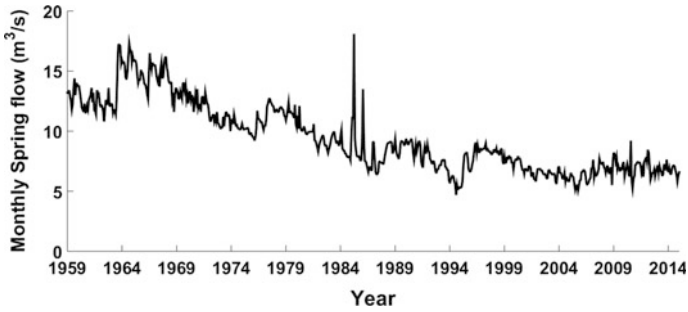



**Fig. 3.** Monthly discharge in the NS basin

## 4 Results

This study will identify significant periodicities of monthly mean precipitation and spring discharge in NS from 1959 to 2015 by using the multitaper method. MTM spectrum estimation of the monthly mean precipitation series and spring discharge in NS are shown in Figs. 4 and 5. The corresponding frequencies of the peaks associated 99% confidence lever are shown in Tables 1 and 2. MTM is also used to detect the significant periodicities of monthly precipitation time series of seven rainfall stations (Table 3).

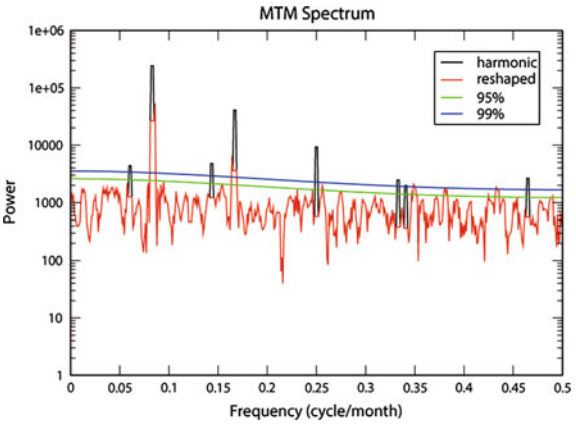



**Fig. 4.** MTM spectrum of precipitation

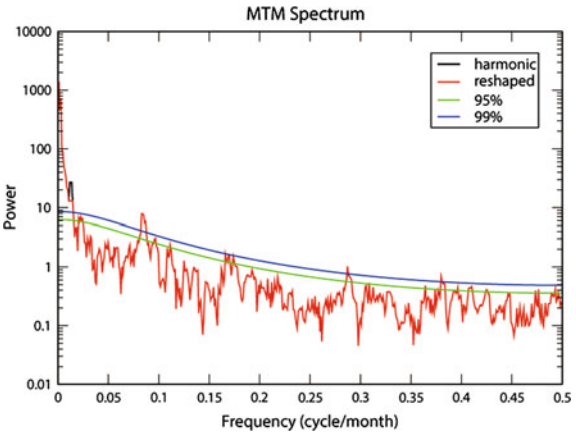



**Fig. 5.** MTM spectrum of spring discharge

**Table 1.** Significant periodicity detected of the mean precipitation by MTM

| Frequency | 0.08 | 0.14 | 0.17 | 0.25 | 0.33 | 0.35 | 0.43 | 0.47 |
|---|---|---|---|---|---|---|---|---|
| Periodicity (month) | 11.77 | 7.01 | 6.06 | 4 | 2.99 | 2.87 | 2.31 | 2.15 |

**Table 2** Significant periodicity detected of the spring discharge by MTM

| Frequency | 0.0029 | 0.0068 | 0.0839 | 0.2871 | 0.3798 |
|---|---|---|---|---|---|
| Periodicity (month) | 342.47 | 146.20 | 11.91 | 3.48 | 2.63 |

**Table 3** Significant periodicity detected of monthly precipitation of subregions of the NS basin by MTM

| City | Periodicity (month) |
|---|---|
| Yu County | 17.07, 11.77, 5.99, 4, 2.63 |
| Shouyang County | 11.77, 6.06, 4.00, 2.87, 2.3 |
| Yangquan city | 11.77, 5.99, 4.00, 2.99, 2.20, 2.16 |
| Pingding County | 12.34, 5.92, 3.97, 3.64 |
| Xiyang County | 11.77, 6.87, 5.99, 4, 2.99 |
| Heshun County | 17.07, 11.77, 6.87, 6.06, 4, 2.29 |
| Zuoquan County | 11.77, 5.99, 4.00, 3.00 |

**Table 4** Time lags and max correlation coefficient for reconstructed sequence of spring discharge and precipitation at the seven rainfall stations

| City | Yu County | Shouyang County | Yangquan County | Pingding County | Xiyang County | Heshun County | Zuoquan County |
|---|---|---|---|---|---|---|---|
| Time lag (month) | 4 | 27 | 4 | 3 | 4 | 4 | 27 |
| Coefficient | 0.6847 | 0.6715 | 0.6923 | 0.5226 | 0.6973 | 0.6835 | 0.6589 |

Significant periodicity of 1 year was found in both precipitation and spring discharge (Tables 1, 2, and 3). To further explore the relationship of them, the common significant periodic components are reconstructed at one-year periodicity.

A lag correlation is useful for investigating the phase shift of the dependent time series that results in the strongest correlation [8]. The cross-correlation function (CCF) is used to calculate the time lag and the maximum correlation coefficient based on the reconstructed sequences of spring flow and precipitation at seven rainfall stations in the NS basin. As is shown in Table 4, the numbers reflected the lags in months at which max positive correlation is attained. At the reconstructed one-year periodicity, the correlation between precipitation of Yu County, Yangquan city, Xiyang County, Heshun County and spring discharge is the highest when the time lag is 4 months. Shouyang County and Zuoquan County have the highest correlation between precipitation and spring discharge when the time lag is 27 months. The precipitation of Pingding country is significantly corrected with spring discharge with the lag of 3 months.

In NS basin, the counties of Zuoquan and Shouyang lie in the southwest and northwest of the spring area, respectively. The area is in the edge of the runoff area with the long runoff distance [9]; it is characterized by weak karst development. Therefore, the time lag of precipitation infiltrate to groundwater will be relatively long. Yu County, Yangquan city, Xiyang County, and Heshun County are the mainstream area of karst water in the region. The area is characterized by its low water level, dynamic stability, and the higher karst development degree; it leads to the shorter time lag. Pingding County is in the strong runoff zone, and the degree of karstification is highest. The time lag at Pingding County is the shortest between precipitation and spring discharge.

## 5 Conclusion

The multiperiodicities exist in the spring discharge and precipitation in the NS basin. But only the one-year periodicity of precipitation can penetrate through karst aquifer and inherit in the spring discharge. Other periodicities of precipitation are altered by karst aquifer, when precipitation signals transport in the aquifer. So the relations between the spring discharge and precipitation at one-year time scale can reflect the feature of karst aquifer. One can identify the heterogeneity of karst aquifer by time lags between the spring discharge and precipitation in the NS basin. In other words, the response time of the spring discharge to the rainfalls in karst-developed zones is shorter than those in karst undeveloped zones.

In the NS basin, the time lag of the spring discharge with precipitation at Zuoquan County and Shouyang County at karst undeveloped zones is 27 months. The time lags of Yu County, Yangquan city, Xiyang County, and Heshun County at middle karst-developed zones are 4 months. The time lag at Pingding County at karst-developed zone is 3 months. The results are consistent with the geological structure of the NS basin.

The multitaper method (MTM) is robust in identification of heterogeneity of karst aquifer by analyzing the relations between spring discharge and precipitation in karst spring basin.

**Acknowledgments.** This paper is supported by Natural Youth Science Foundation of China (61501326, 61401310), the National Natural Science Foundation of China (61731006), and Natural Science Foundation of China (61271411). It was also supported by Tianjin Research Program of Application Foundation and Advanced Technology (15JCZDJC31500), Tianjin Natural Science Foundation (18JCZDJC39500), and Tianjin Science Foundation (16JCYBJC16500). This work was also supported by the Tianjin Higher Education Creative Team Funds Program.

## References

1. Thomson DJ. Spectrum estimation and harmonic analysis. Proc IEEE. 1982;70:1055–96.
2. Park J. Envelope estimation for quasi-periodic geophysical signals in noise In: A multitaper approach, in statistics in the environmental and earth sciences. London: Edward Arnold Press;1992. p.189–219.
3. Mann ME, Park J. Spatial correlations of interdecadal variation in global surface temperatures. Geophys Res Let. 1993;20:1055–8.
4. Hu Y, Loizou PC. Speech enhancement based on wavelet thresholding the spectrum. IEEE Trans Speech Audio Proc. 2004;12(1):59–67.
5. Lan L. Research on the detection of spectrum holes in cognitive radio; 2008.
6. Ghil M, Allen MR, Dettinger MD, Ide K, Kondrashov D, Mann ME. Advanced spectral methods for climatic time series. Rev Geophys. 2002;40(1):1–41.
7. Mann ME, Lees JM. Robust estimation of background noise and signal detection in climatic time series. Clim Change. 1996;33:409–45.

8. Dickinson JE, Hanson RT. HydroClimate—hydrologic and climatic analysis toolkit. U.S: Geological Survey Techniques and Methods; 2013, p. 4–A9.
9. Hao Y, Huang D, Liu J, et al. Study on time lag of precipitation supply in Niangzi Guanquan region. J China Karst. 2003;22(2):92–5.

# Similarity Analysis on Spectrum State Evolutions

Jiachen Sun[1], Ling Yu[1], Jingming Li[2], and Guoru Ding[1,3(✉)]

[1] College of Communications Engineering,
Army Engineering University, Nanjing, China
sun_jiachen@outlook.com, yl2811284434@126.com, dr.guoru.ding@ieee.org
[2] Nanjing University of Aeronautics and Astronautics, Nanjing, China
lijingmingjlu@163.com
[3] National Mobile Communications Research Laboratory,
Southeast University, Nanjing, China

**Abstract.** The correlations between spectrum state evolutions, as a kind of similarity measure, have been revealed to optimize the spectrum usage model or improve the performance in spectrum prediction. However, most existing similarity analyses only end up with the superficial similarity phenomenon. It is of great need for us to conduct the deep investigation and analysis on the similarity of spectrum state evolutions. Firstly, we design a similarity index for spectrum state evolutions based on the Euclidean distance. Then, a network of spectrum state evolutions in the frequency domain can be formed for further analysis by comparing the proposed similarity indexes of frequency points with the decision threshold. Experiments with real-world spectrum data prove the feasibility and rationality of the above similarity analysis.

**Keywords:** Spectrum state evolution · Similarity analysis · Complex network · Scale-free

## 1 Introduction

Studies on the dynamic spectrum access have increasingly drawn attention of researchers, and the results become blooming and flourishing, including the large-scale precise spectrum measurements [1,2], the adaptive spectrum sensing [3], the robust spectrum prediction [4,5], etc.

Either in modelling the spectrum usage or in proposing a new spectrum prediction scheme/algorithm, the inherent correlations between spectrum state evolutions in the time domain and the frequency domain are revealed to optimize the results or improve the performance. For example, the high spectral and temporal correlations of various real-world spectrum datasets are shown in [6], which motivates researchers to restudy the spectrum prediction from a time–frequency two-dimensional perspective. The complex interdependencies in spectrum use


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_61

between the various bands, together with their spatial correlation structures, are described in [7], by presenting extensive cross-correlational analyses.

However, the most existing correlation analysis in spectrum only illustrates the correlations and points out the direction of utilizing the correlations, such as mining the spectrum data evolution from an overall matrix view instead of a vector view [6]. On the one hand, in terms of the frequency domain, how much the correlation between one certain frequency point and another can differ from the correlations between other different frequency points? How can we utilize the correlations in spectrum to make some specific adjustments on the access scheme?

On the other hand, since the objects in the above work are all the spectrum state evolutions in some domain, the correlation mentioned in this paper is a kind of similarity measure, namely a measure that assesses the degree of similarity between the linear evolutions of two objects. The correlation has its limitations and is not the only relationship in the similarity measure. Beyond that, there are many other methods, such as the Euclidean distance, the Manhattan distance, the cosine similarity, the Jaccard similarity coefficient and so on, which are respectively appropriate to different datasets. We can select a suitable method for analysis by the properties of the spectrum state data.

These considerations motivate us to conduct the deep investigation and analysis on the similarity of spectrum state evolutions, and this paper will focus on the similarity in the frequency domain, namely the similarity of spectrum state evolutions of different frequency points in a particular location during the same period. For the similarity is the relationship between any two frequency points, evaluating the similarities among a large number of frequency points in a wide band will remind us the adjacency matrix, which can store relationships of vertices in graph theory. Thus, one frequency point can be assumed as a node in the undirected graph in this paper and massive frequency points of a wide band form a network based on the undirected graph. Similar to making analyses on the structure and characteristics of the traffic load correlation model of base stations coupling with the theory of complex networks in [8], we can invoke the theory of complex networks to analyse the similarity of spectrum state evolutions in the frequency domain.

The main contributions of this paper can be summarized as follows:

1. We design a similarity index for spectrum state evolutions based on the Euclidean distance.
2. We form a network of spectrum state evolutions in the frequency domain by comparing the proposed similarity indexes with the decision threshold.
3. We analyse the property of the degree distribution and the impact of the decision threshold on the degree distribution, then explain the clustering phenomenon of nodes in the spectrum state evolution network.

The remainder of this paper is organized as follows. Section 2 will propose our similarity index for spectrum state evolutions. Section 3 will introduce the formulation of the spectrum state evolution network and some common network

parameters. The properties of the degree distribution and the clustering phenomenon of nodes in the spectrum state evolution network will be presented in detail in Sect. 4. Conclusions are drawn in Sect. 5.

## 2 The Proposed Similarity Index for Spectrum State Evolutions

This paper studies the similarity of spectrum state evolutions in the frequency domain. For each frequency point, its state series, made up of the measured power spectrum density (PSD) values during a period of time, can describe its evolution characteristics. Similarity measure can be conducted on the spectrum state dataset consisting of series of frequency points in the wide band to obtain the similarity relationships between any two frequency points.

### 2.1 Data Preprocessing: Z-Score

Most methods of the similarity measure are sensitive to outliers. To eliminate the noise effects without prejudice to further statistical treatment, the computation of the Z-score on the original series can be conducted as data preprocessing. For each time series, the new elements $\widehat{x}_{ki}$'s are obtained by (1) from the original elements $x_{ki}$'s,

$$\widehat{x}_{ki} = \frac{x_{ki} - \mu_k}{\sigma_k}, \tag{1}$$

where $\mu_k$ is the mean value of the series $X_k$ and $\sigma_k$ is the standard deviation of the series $X_k$.

### 2.2 The Similarity Measure of Time Series: The Euclidean Distance

Whether two sequences are similar or not mainly depends on whether their evolution trends are consistent. The correlation coefficient is usually used to describe the strength and direction of the linear relationship between evolutions of two series in [6]. However, the correlation coefficient is applicable to series with the obvious fluctuation, but sensitive to outliers. The value of the correlation coefficient ranges from $-1$ to $+1$, where the bigger the absolute value is, the stronger the correlation is, which will increase the difficulty of selecting the threshold. The long length of the time series will also lead to the small absolute values of correlation coefficients. Then, we shift our attention to another similarity measure method, the Euclidean distance, which indicates the similarities in one direction and is more applicable to long time series.

The Euclidean distance originally refers to the distance of two points in the $n$-dimensional space in mathematics. When it comes to the series, the Euclidean distance, ED for short, can be also used to evaluate the similarity of two series, which can be denoted as follows:

$$ED\left(X_1, X_2\right) = \sqrt{\sum_{i=1}^{n} \left(\widehat{x}_{1i} - \widehat{x}_{2i}\right)^2}, \tag{2}$$

where $\widehat{x}_{1i}$ and $\widehat{x}_{2i}$ are the elements of two time series $X_1$ and $X_2$ with the same length of $n$ after data preprocessing, respectively.

### 2.3 Normalization Operation

For the long length of series, namely the high dimensions in the computation of distances, values of the Euclidean distance are big and discrete. We can normalize the Euclidean distances between series to rescale the range of the distance indexes to $[0, 1]$ according to (3). To keep similarities between any one series and all the others constant in the global, we conduct the operation of normalization when taking all series in the dataset as a whole. The $ED_{\max}$ in (3) is the maximum among distances between any two series in the dataset, and the $ED_{\min}$ in (3) is the minimum among all except the distance between one series and itself.

$$d\left(X_1, X_2\right) = \frac{ED\left(X_1, X_2\right) - ED_{\min}}{ED_{\max} - ED_{\min}} \tag{3}$$

Thus, the $d\left(X_1, X_2\right)$ denotes the proposed index of similarity measure between the series $X_1$ and $X_2$, similarity for short in the following sections. The smaller the similarity index $d$ is, the more similar the two series are to each other. The similarity index between one series and itself is set to 0.

## 3 Network of Spectrum State Evolutions

In this section, the formulation of the network for analysing the similarity of spectrum state evolutions in the frequency domain will be introduced in detail. Parameters and properties in the theory of the complex network will be also included.

### 3.1 Network Formulation

With all the similarity indexes between any two series in the dataset, we can obtain the adjacency matrix, that describing the similarity relationships between frequency points which the series correspond to respectively, by the decision threshold $t$. If $d\left(X_i, X_j\right)$ is smaller than $t$, we think there exists a link between the series $X_i$ and $X_j$, namely their corresponding $i$-th and $j$-th frequency points, then the element $w_{ij}$ of the adjacency matrix $W$ is set to 1; if not, there does not exist a link between the $i$-th and $j$-th frequency points and the element $w_{ij}$ is set to 0. The decision process can be described in mathematical language as (4).

$$\begin{cases} \text{if} \quad d\left(X_i, X_j\right) < t, & w_{ij} = 1 \\ \text{if} \quad d\left(X_i, X_j\right) \geq t, & w_{ij} = 0 \end{cases} \tag{4}$$

The adjacency matrix $W$ is a symmetric matrix whose all diagonal elements are 0. The undirected graph of a network can be generated from the adjacency

matrix. Thus, in the network of the similarity analysis on spectrum state evolutions, each frequency point included in the spectrum state dataset is assumed as a node and each $w_{ij}$ equal to 1 represents the edge between the $i$-th frequency point and $j$-th frequency point. Then, we can calculate some parameters of the network to investigate the properties of the network.

Actually, the selection of the decision threshold $t$ is of great significance to the size and properties of the network. On the one hand, we should avoid that the threshold $t$ is too small to ensure the relatively few isolated nodes. On the other hand, to prevent the weakly similar nodes from being linked, the threshold $t$ cannot be too large. We should set a suitable threshold according to the distribution of similarity indexes.

### 3.2 Network Parameters

**Network Size**: The network size is the total number of nodes in the network, denoted by $N$. The node mentioned here is the node which is connected to another node in the network at least. When calculating the network size, the isolated nodes are usually excluded.

**Network Diameter**: The network diameter is defined as the maximum distance between any two nodes connecting with each other in the network, denoted by $D$. We can regard the distance between two adjacent nodes directly connecting with each other as 1. When two nodes connect with each other via other nodes, namely linked by multiple hops, the distance between them is equal to the number of hops. Then, the maximum value among all distances is the diameter of the network.

**Degree of Node**: In the undirected graph, the degree of a node refers to the number of edges connected to the node itself, denoted by $m$. In the research of networks, the degree distribution of all nodes causes much attention. The degree distribution of nodes in the network is widely recognized as the classification basis for regular networks, random networks and scale-free networks. For example, the degree distribution of scale-free networks is usually depicted by a power-law function [9], i.e. $P(m) \sim m^{-\lambda}$, where $\lambda$ is the degree exponent.

## 4 Real-World Data Experiment and Analysis

In this section, we conduct the similarity analysis of spectrum state evolutions in the frequency domain with the real-world spectrum dataset.

### 4.1 Experiment Setup

All the measured PSD values are from the open-source real-world spectrum dataset of the well-known RWTH Aachen University spectrum measurement campaign [1]. The time series, namely the PSD values of massive frequency points involved in this paper, are the measurements of the GSM900 downlink band from 930 to 960 MHz during 3 days, which are collected in a residential

area in Aachen, Germany. The total number of the frequency points in the band is 163, for the frequency interval is about 180 KHz. The inter-sample time is about 1.8 s, then the length of the series of each frequency point is 140,000.

Based on the above real-world spectrum dataset, we can calculate the proposed similarity indexes between any two frequency points of all 163 frequency points. With the decision threshold, a network whose size is 163 at most can be formed. Analyses on the properties of the degree distribution and the clustering phenomenon of nodes will be illustrated.

### 4.2 Power-Law Distribution of Degree

According to the calculated similarity indexes, we would like to set the decision threshold $t$ to the value ranging from 0.54 to 0.70. Setting the decision threshold $t$ to 0.64 as an example, a network, whose size is 162, is formed for analysis, where only one node of the total 163 nodes is isolated. Then, the cumulative distribution function of the degree of all nodes is shown in Fig. 1. More than half of the nodes are with the degree not exceeding 3 and less than 10% of nodes are with high degree, connected with many nodes.

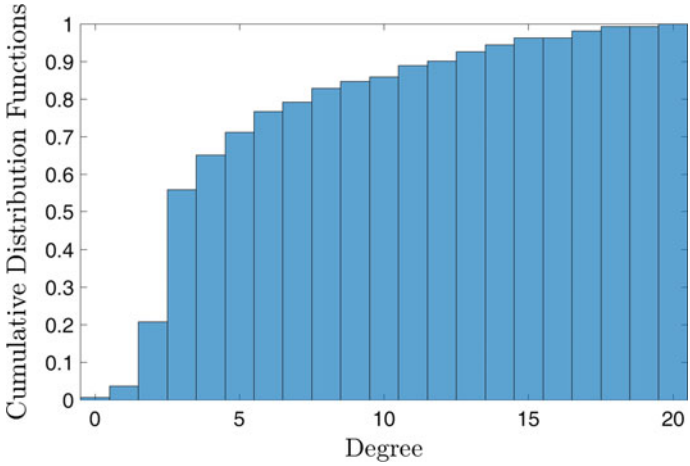
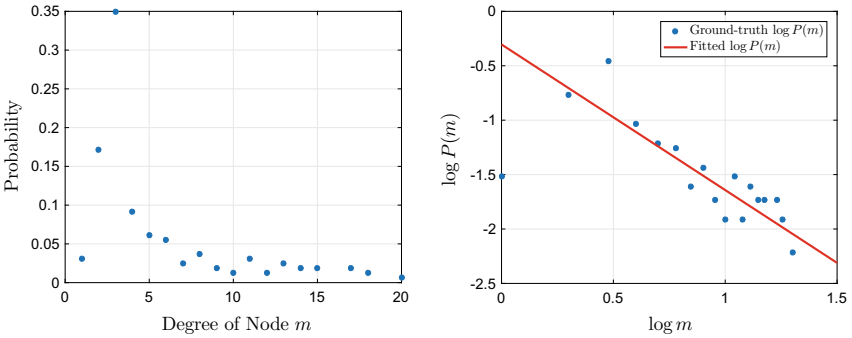



**Fig. 1.** Cumulative distribution functions (CDFs) for the degree of nodes in the spectrum state evolution network with threshold 0.64.

Figure 2 shows the probability distribution of degree in rectangular coordinates and logarithmic coordinates, respectively. The degree distribution in Fig. 2a exhibits the approximate power function property, and the fitted red line in Fig. 2b verifies that the degree distribution can be well depicted by a power-law function after the coordinate transformation, namely the scale-free property of the spectrum state evolution network.




(a) Degree distribution of the spectrum state evolution network with threshold 0.64

(b) Degree distribution of the spectrum state evolution network with threshold 0.64 (logarithmic coordinates)

**Fig. 2.** Degree distribution and its fitting result of the spectrum state evolution network with threshold 0.64.

### 4.3 Impact of the Decision Threshold on the Degree Distribution

The decision threshold $t$ is changed to discuss the difference in the network formulation. As the results shown in Table 1, the selection of the decision threshold within a reasonable range will have an impact on the size of the network, but no change on the scale-free property of the network.

**Table 1.** The network parameters and the degree distribution under different decision threshold.

| Thresholds $t$ | Network size $N$ | Number of isolated frequency points | Rate of isolated frequency points (%) | Degree exponent $\lambda$ |
|---|---|---|---|---|
| 0.54 | 155 | 8 | 4.91 | 2.844 |
| 0.58 | 157 | 6 | 3.68 | 2.189 |
| 0.62 | 162 | 1 | 0.61 | 1.538 |
| 0.66 | 163 | 0 | 0.00 | 1.049 |
| 0.70 | 163 | 0 | 0.00 | 0.965 |

### 4.4 Clustering Phenomenon of Nodes

For the scale-free property, there must exist several important nodes with high degree that connected to many nodes with low degree, which leads to the clustering phenomenon in the network. Figure 3 illustrates the global connections of the spectrum state evolution network with threshold 0.64, where each round represents a frequency point and numbers on the rounds are serial numbers of

163 frequency points in the band from 930 to 960 MHz. The big red ones represent frequency points with high degree, and small blue ones represent frequency points with low degree. All edges with different colours in the graph are equal-weighted.

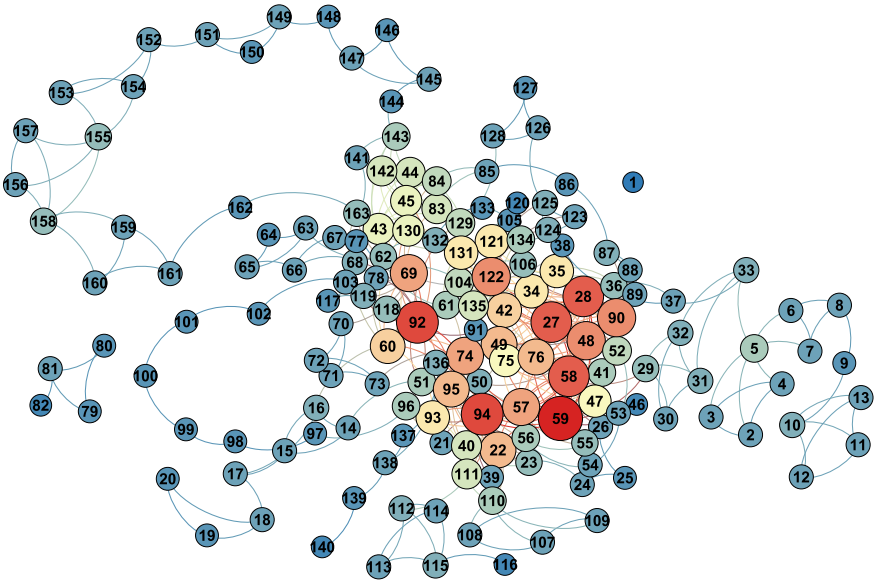


**Fig. 3.** Undirected graph of the spectrum state evolution network with threshold 0.64.

Obviously, adjacent frequency points usually have similarities on spectrum state evolutions and several frequency points will have similarities with more frequency points except for their adjacent ones. This phenomenon may attribute to the fragmented spectrum allocation to different communication carriers. However, the connections in our spectrum state evolution network are a bit different from the radial connections of typical scale-free networks.

## 5 Conclusion

In this paper, we design a similarity index for spectrum state evolutions based on the Euclidean distance. By comparing similarity indexes with the decision threshold, we form a network of spectrum state evolutions in the frequency domain. Then, we conduct a series of analysis on the spectrum state evolution network based on the real-world spectrum data. The degree distribution is verified to be depicted by a power-law function, and the selection of the decision threshold will only influence the network size. At last, the undirected graph of the network illustrates the connection properties in our spectrum state evolution network. For the future work, we would conduct further research on the network

properties in radio spectrum evolutions by invoking the theory of complex networks.

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation of China (Grants No. 61501510 and No. 61631020), Natural Science Foundation of Jiangsu Province (Grant No. BK20150717), China Postdoctoral Science Foundation Funded Project (Grant No. 2016M590398) and Jiangsu Planned Projects for Postdoctoral Research Funds (Grant No. 1501009A).

## References

1. Wellens M. Empirical modelling of spectrum use and evaluation of adaptive spectrum sensing in dynamic spectrum access networks. Ph.D. Dissertation, RWTH Aachen University, May 2010.
2. López-Benítez M, Casadevall F. Spectrum usage models for the analysis, design and simulation of cognitive radio networks. In: Cognitive radio and its application for next generation cellular and wireless networks. The Netherlands: Springer; 2012. p. 27–73.
3. Axell E, Leus G, Larsson EG, Poor HV. Spectrum sensing for cognitive radio: State-of-the-art and recent advances. IEEE Signal Process Mag. 2012;29(3):101–16.
4. Xing X, Jing T, Cheng W, Huo Y, Cheng X. Spectrum prediction in cognitive radio networks. IEEE Wireless Commun. 2013;20(2):90–6.
5. Ding G, Jiao Y, Wang J, Zou Y, Wu Q, Yao Y, Hanzo L. Spectrum inference in cognitive radio networks: Algorithms and applications. IEEE Commun Surveys Tuts. 2018;20(1):150–82.
6. Ding G, Wu F, Wu Q, Tang S, Song F, Vasilakos AV, Tsiftsis TA. Robust online spectrum prediction with incomplete and corrupted historical observations. IEEE Trans Veh Technol. 2017;66(9):8022–36.
7. Palaios A, Riihijärvi J, Holland O, Mähönen P. Detailed measurement study of spatial similarity in spectrum use in dense urban environments. IEEE Trans Veh Technol. 2017;66(10):8951–63.
8. Yuan C, Zhao Z, Li R, Li M, Zhang H. The emergence of scaling law, fractal patterns and small-world in wireless networks. IEEE Access. 2017;5:3121–30.
9. Kim JS, Goh KI, Kahng B, Kim D. Fractality and self-similarity in scale-free networks. New J Phys. 2007;9(6):177.

# A Data-Aided Method of Frequency Offset Estimation in Coded Multi-h CPM System

Cong Peng[1(✉)], Shilian Wang[1], Penghui Lai[1], and Chun Yang[2]

[1] College of Electronic Science and Engineering, National University of Defense Technology, Changsha, China
pengcongl994@163.com, wangsl@nudt.edu.cn
[2] Institute of Electronic Engineering, China Academy of Engineering Physics, Mianyang, China

**Abstract.** The carrier frequency offset is a common problem in communication systems, but in most case, phase-locked loop (PLL) can only track small frequency shift, so an estimation of initial frequency offset is necessary for CPM synchronization. In coded multi-h CPM system, it is a convenient method to use the information of frame header for frequency estimation. In this paper, we combine the pulse-pair technique and fast Fourier transform (FFT) calculating method for our purpose. According to the simulation results, it can achieve accurate estimation for comparative great frequency offset and perform steady even in low signal-to-noise ratios (SNRs).

**Keywords:** Fast Fourier transform (FFT) · Frequency offset estimation · Frame header · Pulse-pair technique

## 1 Introduction

Frequency offset of multi-h CPM signals may cause bad bit error performance, so the carrier frequency compensation is needed in digital receivers to cope with oscillator instabilities or Doppler effects [1]. Frequency offset is usually caused by Doppler frequency shift or difference of the carrier frequency between transmitter and receiver, which can lead to increasing phase shift.

The common methods for frequency offset estimation contain data-aided (DA), non-data-aided (NDA), and decision-directed (DD) synchronization [2]. In most transmission, we have no idea about the data we received so that we can only calculate what we get and try to figure out information about signals as much as possible, which is the common situation that we use NDA algorithm to estimate the frequency. Usually, NDA algorithm is not that complicated and can be implemented easily, but the accuracy is not enough for great Doppler frequency shift. While DD algorithm considers accuracy, it will expense extra resources unless decisive part of demodulation can be reused. But when we have some knowledge of data transmitted, DA algorithm can work in an easy way.

In our discussion, we mainly consider a serially cascaded CAM system with coded data affected by additive white Gaussian noise (AWGN) channel. Actually, in multi-h CPM system, we have previously proposed an algorithm based on barycenter of


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_62

frequency spectrum [3]. This method has great anti-noise capability and tracking accuracy but needs lots of calculation and cost comparatively big resources. But fortunately, the information transformed in frame structure in coded system is known, so we can use these data of frame header to estimate its frequency in the proposed method conveniently when the header of received signals is caught using the algorithm described in [4].

This paper is organized as follows. In Sect. 2, CPM signal model is briefly discussed. And our DA estimating method is presented in Sect. 3. In Sect. 4, we give some simulation results of estimation accuracy and effects to bit error performance, while analyzing the implementation of the algorithm. Section 5 is a conclusion of the paper.

## 2 CPM Signal Model

The complex baseband CPM signal can be presented as

$$s(t;\alpha) = \sqrt{\frac{E_s}{T}}\exp\{\mathrm{j}\psi(t;\alpha)\} \tag{1}$$

where $E_s$ is the energy during a symbol period, $T$ is symbol period, and $\psi(t;\alpha)$ is the phase of CPM signals and usually represented as

$$\psi(t;\alpha) = 2\pi\mathrm{h}\sum_{i=0}^{L-1}\alpha_i q(t - iT_\mathrm{s}) \tag{2}$$

In multi-h CPM, $\alpha_i$ is the sequence of M-ary information symbols selected from the set of $[\pm 1,\ \pm 3\ \cdots \pm (M-1)]$. The variable h is the modulation index which can vary from symbol to symbol; the waveform $q(t)$ is the phase response of CPM and in general is represented as the integral of the frequency pulse $g(t)$ with a duration of $LT_s$ [5].

Assuming the transmission is over AWGN channel, the complex baseband representation of the received signal is

$$r(t) = s(t-\tau;\alpha)\cdot \mathrm{e}^{j(2\pi f_d t+\theta)} + w(t) \tag{3}$$

where $\theta$ is the unknown carrier phase, $f_d$ is the frequency offset, $w(t)$ is the complex baseband AWGN with zero mean and power spectral density $N_0$, and $\tau$ is timing offset caused by multi-accesses and inconsistency of clock rates [6].

As CPM signal has the constant envelope, and almost all information is modulated to the phase part. So not only AWGN can deteriorate the performance of our demodulation, but also $f_d$ and $\theta$ can. Meanwhile, big frequency offset may seriously affect timing synchronization. As a result, it is quite important to estimate the frequency offset accurately and correct the carrier frequency quickly.

## 3 Data-Aided Method of Frequency Estimation

As mentioned in the last two parts, we use the information of frame header to find out the frequency offset. A set of data frame contains frame header and information data. Once the length of header is chosen according to the standard of CCSDS [7], which is unchanged during transmission and accept by both transmitter and receiver. So we can use this known information to modulate a length of known CPM baseband signals and get its conjugation for pulse-pair operation [8]. Assuming that timing offset $\tau$ has been corrected, the expression of matched result is shown as:

$$\Lambda(\widetilde{f_d}, \widetilde{\theta}) = \left[s(t;\alpha)\cdot e^{j(2\pi f_d t+\theta)} + w(t)\right]\cdot s^*(t;\alpha_h), \qquad t \in [0, T_l] \tag{4}$$

where $\alpha_h$ is a sequence of frame header with a length of $T_l$, and $s^*(t;\alpha_h)$ is conjugation of the local wave signal that we have known: it is presented as:

$$s^*(t;\alpha_h) = \sqrt{\frac{E_s}{T}}\exp\{-j\psi(t;\alpha_h)\} \tag{5}$$

As $s^*(t;\alpha_h)$ has little correlation with $w(t)$, (4) can be rewritten as:

$$\Lambda(\widetilde{f_d}, \widetilde{\theta}) = \frac{E_s}{T}e^{j(2\pi\widetilde{f_d}t+\widetilde{\theta}-\theta_0)},\ t \in [0, T_l] \tag{6}$$

It is easy to find that $\Lambda(\widetilde{f_d}, \widetilde{\theta})$ is positively related to a single frequency signal even with the effect of $w(t)$. For this consideration, we choose to do FFT for $\Lambda(\widetilde{f_d}, \widetilde{\theta})$ and pick up maximum peak of the spectrum, which indicates the position of $f_d$, which is faded back to $r(t)$ in (3) by a NCO. The result $\widehat{r}(t)$ is defined as:

$$\widehat{r}(t) = s(t-\tau;\alpha)\cdot e^{j(2\pi\Delta f t+\Delta\theta)} + \hat{w}(t) \tag{7}$$

where $\Delta f = f_d - \widetilde{f_d}$ represents the remaining frequency offset and $\Delta\theta = \widetilde{\theta} - \theta_0$ represents the phase offset, both of which will be tracked and corrected by a second-order PLL. The whole structure is demonstrated as Fig. 1.

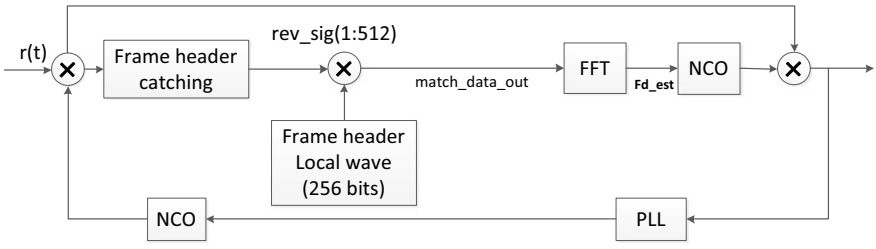
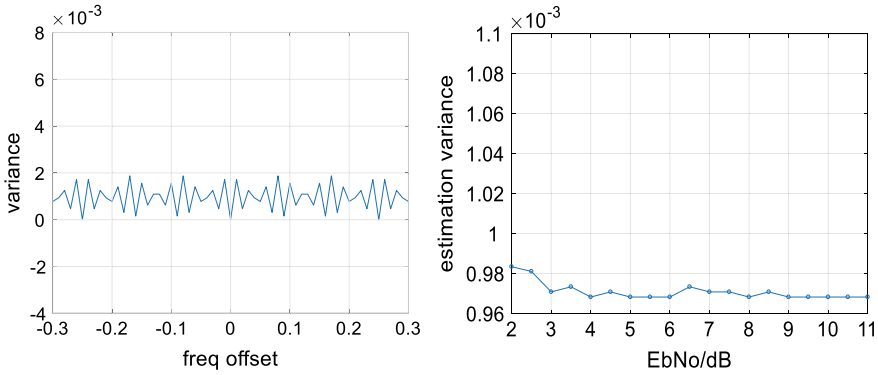



**Fig. 1.** Structure of phase and frequency tracking loop

## 4 Simulation Results

In the previous sections, we have introduced the CPM signals and the proposed method for frequency offset estimation (FOE). During this part, we will do some simulation of our algorithm and give the results to illustrate the accuracy and effect to bit error performance.

First of all, we give an evaluation of the accuracy of the frequency estimation. In MATLAB simulation, the length of frame header is selected as 256 bits, and we set EbN0 = 5 dB (EbN0 refers to signal energy per bit-to-power density of noise ratio) and allow frequency offset $f_d$ ranging from 0.01 $R_b$ to 0.1 $R_b$ with 0.001 $R_b$ step. Figure 2 shows the variance.



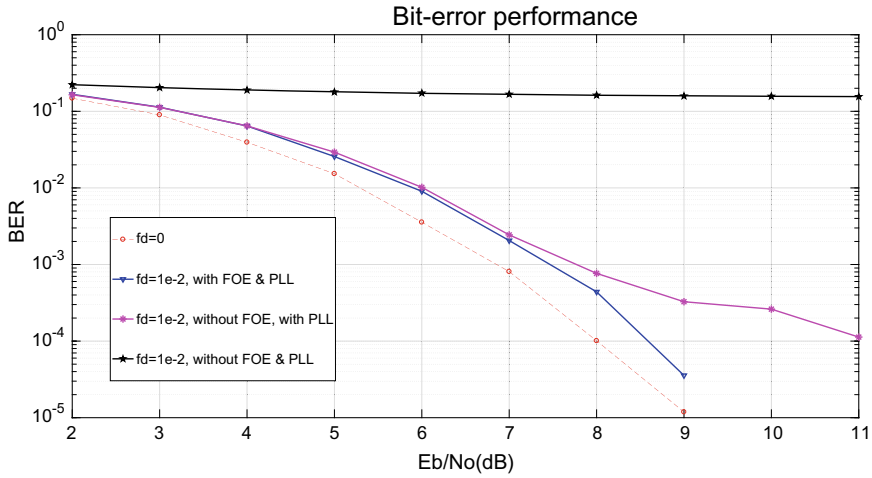

(a) frequency offset changing (b) EbN0 changing

**Fig. 2.** Variance of frequency offset estimation

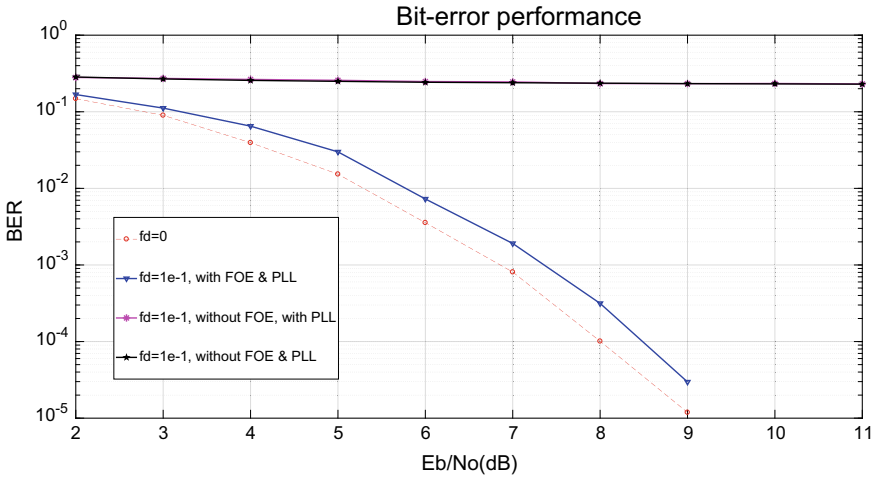
From Fig. 2a, we can see that the variance is in a low degree around 1e-6 with frequency varying, and b shows that it changes little when signal-to-noise ratios (SNRs) deteriorate, which means the proposed method can realize steady estimation. Meanwhile, we have to indicate the importance of estimation to our system.

For such purpose, we set $f_d = 0.01\ R_b$ and $f_d = 0.1\ R_b$ and then give bit error rate (BER) simulation results without FOE and PLL, without FOE, and with both.

Figure 3 shows that our FOE module can successfully correct frequency offset and decode with little loss compared to ideal situation. Even $f_d = 0.1\ R_b$ (a relative big offset), the loss is only 0.5 dB, but the system cannot work without this module. In aspect of implementation, we list several main modules required for the calculation in Table 1. It is clear that the resources consumption is less than 1%, and modules utilized are common in implementation.




(a) $f_{\mathrm{d}} = 0.01R_b$




(b) $f_{\mathrm{d}} = 0.1R_b$

**Fig. 3.** BER in different frequency offsets

**Table 1.** Implementation consumption of the proposed method

| Module | Num | Function | Resources (%) |
|---|---|---|---|
| FFT | 1 | Calculate spectrum | 0.36 |
| Multiplier | 3 | Complex multiply;<br>Power calculation | 0.64 |
| Adder | 1 | Quadratic sum | 0.014 |
| Divider | 1 | Estimate frequency | 0.036 |

## 5 Conclusions

In this paper, we have introduced a simple but effective method for frequency offset estimation. The proposed method can work mainly due to special structure of data, which actually contribute to the system. And we analyze its mathematical correction and draw up the front part of receiver. For convincing illustration, we do a simulation to indicate estimation accuracy and effect to bit error performance. As the results suggest, our method works well in big frequency offset and cost little loss of performance.

But there is still a lot of work to do. As our method just estimates once, when the frequency offsets change too fast, the estimating results may have big biases. And such spectrum based method is greatly affected by point numbers of FFT.

## References

1. D'Andrea AN, Ginesi A, Mengali U. Digital carrier frequency estimation for multilevel CPM signals, Pisa, Italy.
2. Mengali U, D'Andrea A. Synchronization techniques for digital receiver. New York: Plenum Press; 1997.
3. Xie S, Zhhong S, Fan J, Yang C. Carrier frequency synchronization for large dynamic PCM/FM signals. Telecommun Eng. 2017;57(3):276–81.
4. Hosseini E, Perrins E. Timing, carrier, and frame synchronization of burst-mode CPM. IEEE Trans Commun. 2013;61(12).
5. Hosseini E, Perrins E. Maximum likelihood synchronization of burst-mode CPM. Commun Theor Symp Globecom; 2013.
6. Wang W, Abeysekera SS. Data aided phase tracking and symbol detection for CPM in frequency-flat fading channel. In: ICASSP; 2011.
7. Standard IRIG. Telemetry standards, range commanders council 106–15 (Part I) Appendix R. 2015.
8. Abeysekera SS. Efficient frequency estimation using the pulse-pair method at various lags. IEEE Trans Commun. 2006;54(9).

# Optimization Study on Outlet Pressure of Water Supply Pumping Station Based on Relative Entropy Theory

Zhenfeng Shi[1,2(✉)], Xinran Li[2], and Cuina Zhang[1,3]

[1] Heilongjiang Key Laboratory of Underground Engineering Technology, Harbin, China
szf@hit.edu.cn
[2] School of Science, Harbin Institute of Technology, 92 West Dazhi Street, Nan Gang District, Harbin, Heilongjiang province, China
[3] School of Technology, Harbin University, Harbin, China

**Abstract.** Based on the definition of relative entropy and the principle of minimum relative entropy, this paper takes the pressure of the node as the main research object and establishes an optimization model about the water pressure of the pipe network running in the water supply pump station. Then we apply genetic algorithm to solve our model and analyze the water pressure of the actual water supply pumping station in FS city. According to the model established in this paper, we can optimize the relevant decision variables in the water distribution network and provide a new method basis for the explicit pump scheduling.

**Keywords:** Relative entropy · Water pressure of water supply pump station · Information entropy

## 1 Introduction

In the energy input of the entire pipeline network, the importance of the water supply pumping station is irreplaceable, but with the accelerated process of industrialization and urbanization in China, the energy consumption of the water distribution network (WDN) shows an upward trend. Although in recent decades, Chinese pipeline network scheduling technology has been greatly developed, SCADA runs in many waterworks, because of the lack of professional technicians, it has not been able to update, maintain and manage them. Thus, the model is lagging behind the actual pipe network, and the scheduling of many water pumping stations is still mainly based on the experience of the staff.

The optimization of WDN means that the objective function of the system runs to obtain the optimal solution or better solution under the constraint conditions related to the operation of WDN [1]. In the paper [2], the authors propose


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_63

two Bayesian optimization approaches which resolve black box and costly optimization processes. A graph-theoretic approach which assesses the resilience of large-scale WDN was proposed in [3]. In recent ten years, all disciplines have penetrated into each other, and information entropy is more and more applied in the pipe network [4,5]. The paper [6] builds up the early warning decision system based on the information entropy principle which is used to check early risk warning index and signals. An approach which aims to bridge the gap between the dominant entropy-based design optimization models and water distribution practice was proposed in [7]. Genetic algorithm is regarded as a leap in pipe network optimization. However, GA needs to evaluate the adaptive function before convergence, which is slow in computation and premature convergence. Therefore, improving the computational performance of genetic algorithm is the research direction of many researchers [8]. An approach which improves the algorithm to make the calculation more efficient combines fuzzy set theory with GA to analyze the uncertainty of water supply network in [9]. The paper [10] improves the productivity of genetic algorithms by applying heuristic domain knowledge in the algorithm.

In this paper, we study the optimization of water supply pumping station based on the research characteristics of water supply network operation optimization and information entropy theory. We combine the relative entropy model with the lowest energy consumption of the water pumping station during the operation of the network, and use genetic algorithm to solve the model. Finally, the hourly optimal water pressure of the water pumping station is obtained and verified by practical engineering cases.

## 2 The Pressure Model of Water Supply Pumping Station

Minimizing the total energy consumption of all pumping stations in the entire pipe network plays the most important role in the optimization of water supply pumping stations. According to the two-level optimization theory, we should first optimize the overall water pressure of the pumping station, then the scheduling and operation of the pump within the pumping station will be optimized according to the optimized water pressure. If we can get the overall differences between the best or approximate optimal water pressure and the actual water pressure of all nodes, it can measure the running state of the whole pipe network based on the theory of relative entropy.

### 2.1 Precondition of the Model

We consider the establishment of an optimization model for turbocharging station. We assume that:

1. During the period after optimization of the water supply pumping station, it is established that the operation of the water supply pumping station is optimal;

2. At any time between the end of the last optimization and the beginning of the next optimization, there is no change in water consumption and network conditions of upstream and downstream nodes;
3. The node data of the whole pipe network is accurate and reliable, and the comprehensive efficiency of the whole pumping station is known, besides there is no modification in an optimization period.

### 2.2 The Objective Function

In the optimization of water supply pumping stations, the minimum relative entropy is one of the goals we're going to achieve; accordingly, the first objective function is as follows:

$$\min L(P,Q) = \min \left[ \frac{N_s}{n} \ln \left( \frac{N_s}{n(1-s)} \right) + \sum\nolimits_{\substack{i=1 \\ i \neq s}}^{m} \frac{N_i}{n} \ln \frac{(m-1)N_i}{ns} \right] \quad (1)$$

where $N_s$ is the number of data falling into group $s$, $0 \leq N_s \leq n$, $n$ is the total number of nodes in the whole pipe network, Å is the accuracy of calculation, besides it is a positive number close to 0, $s$ is the serial number in the interval corresponding to $q_s = 1 - \varepsilon$, $1 \leq s \leq m$.

When the energy consumption of all water supply pumping stations is minimum, the energy consumption of the whole pipe network system is at the lowest level, so there is the following objective function:

$$\min \mathrm{E} = \sum_{k=1}^{M} \frac{P_k Q_k}{\eta_k} \quad (2)$$

where $M$ is the number of water supply pumping stations in the whole pipe network, $Q_k$ is the yielding water in $k$ pump, $P_k$ is the water pressure of $k$ water supply pump station, $\eta_k$ is the comprehensive efficiency of $k$ water pump.

### 2.3 Constraints and Formulation of Optimization Model

For water supply pumping stations, four constraints must be met first:

(1) Continuity equation of nodes

For every node in the pipe network, it must satisfy the continuity equation, accordingly the inflow must be equal to the outflow:

$$\sum_{i=1}^{NP_j} (\pm Q_k) + q_j = 0 (j \in n) \quad (3)$$

where $Q_i$ is the flow of $i$ in tube section, $q_j$ is the flow of node $j$, $\mathrm{NP}_j$ is the total number of nodes in the pipe network model, $\pm$ represents the directed summation of the pipe segments associated with the node $j$. When the pipe section flows into the node, it takes a negative value and takes positive value when it flows out.

(2) Energy equation

For any pipe section, the head loss of the pipe network at the upper end point and the lower end point must follow the pressure drop formula:

$$h_{ij} = H_i - H_j \tag{4}$$

where $H_i$ and $H_j$ are water head of pipe section $i$ at upper and lower endpoints, respectively; $h_ij$ is the pressure drop of pipe section $ij$.

(3) Water pressure constraint of node

To the customer, we must meet the water pressure requirements for each user of the node:

$$H_j^{\min} \leq H_j \leq H_j^{\max} (j \in n) \tag{5}$$

where $H_j$ is actual water pressure of node $j$, $H_j^{\min}$ and $H_j^{\max}$ are minimum and maximum service water pressure of node $j$, respectively.

(4) Velocity constraint in pipe section

The flow rate in the pipe network should not be too large or too small. The following requirements should be met:

$$v_{ij}^{\min} \leq v_{ij} \leq v_{ij}^{\max} (i \in \mathrm{NP}) \tag{6}$$

where $v_{ij}^{\min}$ and $v_{ij}^{\max}$ are the minimum and maximum flow velocity defined by the pipe section $ij$, $v_{ij}$ is the actual flow velocity of pipe section $ij$, NP is the number of pipe sections.

In addition to the above four conditions, the water pressure of the water supply pumping station must be satisfied at the same time. For the whole pipe network, the pressure of each pumping station has its own maximum and minimum value, which is related to the pump scheduling. The exit pressure of the pump station is selected to be restricted here:

$$P_k^{\min} \leq P_k \leq P_k^{\max} (k \in m) \tag{7}$$

where $P_k$ is the actual water pressure provided by the pump station $k$, $P_k^{\min}$ and $P_k^{\max}$ are the minimum water pressure and maximum water pressure provided by the pumping station $k$, respectively; $M$ is the number of water supply pumping stations in the whole pipe network.

To sum up, the optimal operation model of the water supply network is shown in the following expressions:

$$\begin{cases} \min L(P,Q) = \min[\frac{N_s}{n} ln(\frac{N_s}{n(1-\varepsilon)}) + \sum\limits_{\substack{i=1 \\ i \neq s}}^{m} \frac{N_i}{n} ln(\frac{(m-1)N_i}{n\varepsilon})] \\ \min E = \sum\limits_{k=1}^{m} \frac{P_k Q_k}{\eta_k} \end{cases}$$
$$\begin{aligned} s.t. \quad & \sum_{i=1}^{NP_j} (\pm Q_k) + q_j = 0 \ (j \in n) \\ & H_j^{\min} \leq H_j \leq H_j^{\max} (j \in n) \\ & v_{ij}^{\min} \leq v_{ij} \leq v_{ij}^{\max} (i \in \mathrm{NP}) \\ & P_k^{\min} \leq P_k \leq P_k^{\max} (k \in m) \\ & h_{ij} = H_i - H_j \end{aligned} \tag{8}$$

## 3 GA-Based Solution for the Model

### 3.1 Coding Method

The encoding method we choose is the most commonly used binary encoding. We assume that there are two water supply pumping stations 1 and 2 and the corresponding range of the water pressure is 32–42 m and 31–40 m, respectively. If the final optimization accuracy is 0.01 m, there should be at least 900 and 1000 code numbers. Because of $2^9 < 900 < 1000 < 2^{10}$, the encoding length is 10. But $210 - 1 = 1023$, the 23 codes and 123 codes are restarted, that is to say, there are 23 chromosomes corresponding to 2 codes and 123 chromosomes corresponding to 2 codes. It can be specifically expressed as:

$$A : 0000000000 = 32,\ 0000000001 = 32.01,\ \ldots\ldots$$

$$B : 0000000000 = 31,\ 0000000001 = 31.01,\ \ldots\ldots$$

Then A and B are encoded jointly with multiple parameters:

$$0000000000|0000000000 = \{32, 31\}$$

$$0000000000|0000000001 = \{32, 31.01\}$$

$$\ldots\ldots$$

So the length of the whole code is 20, the search space required is $2^{20}$, and it is prone to implement in computer.

The coding precision of the pump station 1:

$$\Gamma_1 = \frac{42 - 32}{2^{10} - 1} = 0.0098\,\mathrm{m}$$

The coding precision of the pump station 2:

$$\Gamma_2 = \frac{40 - 31}{2^{10} - 1} = 0.0088\,\mathrm{m}$$

Thus, coding precision meets scheduling requirements.

### 3.2 Removal Constraint Processing of Objective Function and Solving Process

In the five constraint conditions, it contains two equality constraints and three inequality constraints. For equality constraints, it is bound to satisfy the computation process of EPANET. For constraint conditions (5), the water pressure constraint of the pumping station is automatically met in the coding process. Consequently, we only need to deal with the constraints $H_j^{\min} \le H_j \le H_j^{\max} (j \in n)$ and $v_{ij}^{\min} \le v_{ij} \le v_{ij}^{\max} (i \in \mathrm{NP})$. Then we remove the constraint conditions according to the penalty function method.

According to the equation $U = \mu \cdot L + \omega \cdot E$, the double-objective function of operation optimization of water supply pumping station is processed:

$$U = \mu \cdot L + \omega \cdot E \tag{9}$$

After obtaining the penalty function, the joint (10) is processed to obtain an unconstrained optimization problem, as shown in Eq. (10)

$$\min F = \mu \cdot L + \omega \cdot E + \frac{1}{NP_j} \sum_{j=1}^{NP_j} f_{1,j} + \frac{1}{\mathrm{NP}} \sum_{ij=1}^{\mathrm{NP}} f_{2,ij} \tag{10}$$

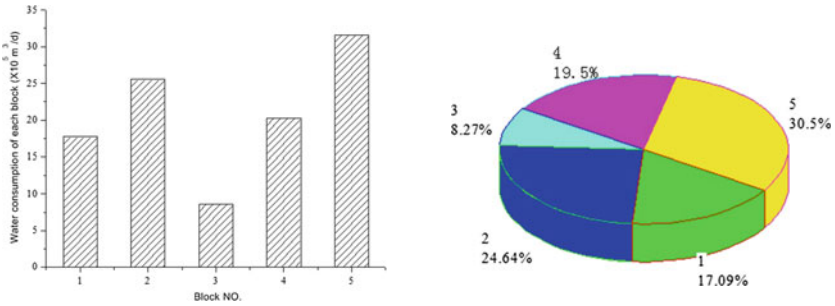
where $\mu$ and $\omega$ are the appropriate proportionality coefficient between values (0, 1).

In order to solve the optimization model of the water pressure of the water supply pumping station by genetic algorithm, the decision variable is the water pressure, it is required to encode it and solve the model with genetic algorithm.

## 4 Experimental Analysis

FS city is located in south of the Five Ridges, east to Guangzhou, South to Hong Kong and Macao. It has a mild climate and a lot of rain. It belongs to the subtropical monsoon humid climate with an annual average temperature of 23.2 °C, the average minimum temperature in January is about 15 °C, and the average maximum temperature in July is about 34 °C and the total water supply population is 170,860,000. FS city has a flat topography, a well-developed manufacturing industry, and a large proportion of industrial water in water supply. The water supply area is divided into five parts: area 2, area 5, area 4, area 3, and area 1; the proportion of population and the amount of water used in July are shown in Fig. 1. Water supply is mainly concentrated in several large town water companies, such as Dali DL waterworks, LC waterworks, GY waterworks, and XT waterworks. Its total water supply accounts for 65.86% of the total water supply in July.

Select 24 h as a research cycle, at the end of each hour, the pressure of the pipe network is counted, and the water pressure of the water supply pumping station is optimized. We assume that the water pressure of the waterworks $B$




**(a)** Water consumption of every block in FS City

**(b)** Population proportion of each block in FS City

**Fig. 1.** Daily water consumption and population ratio in each block

and $A$ are respective $P_1$ and $P_2$, the initial population size is 80, the crossover probability is 0.6, the mutation probability is 0.001, the number of generations is 100, and the initial temperature $T_0$ of the fitness function transformation is 100. It is found in Fig. 2 that the fitness function converges to the maximum after 100 iterations. At this point, the minimum value of the objective function is obtained, and the corresponding variable value is the requirement.

The change curve of water pressure before and after optimization of water plant $B$ and water plant $A$, relative entropy, and energy consumption parameters are listed in Fig. 3 respectively. Obviously, the unoptimized pipe network pressure

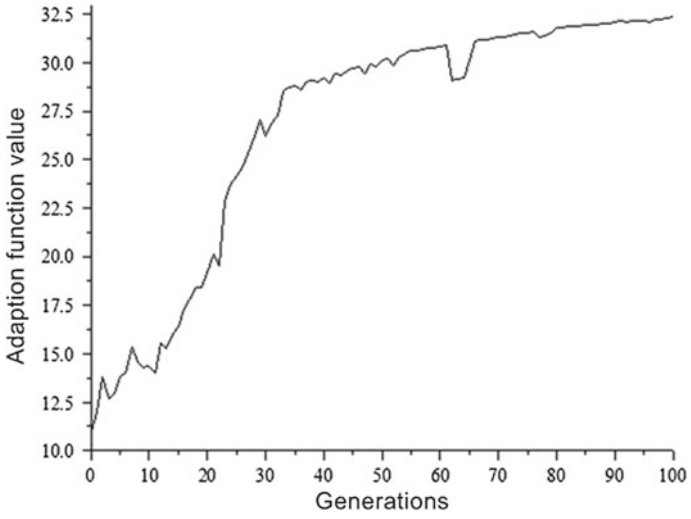
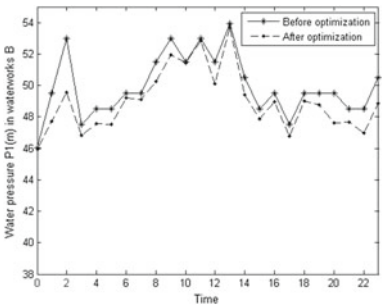
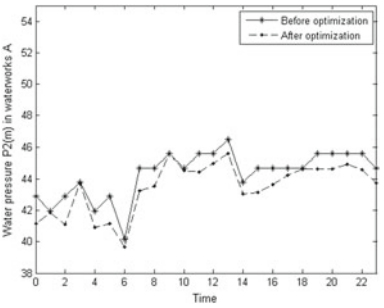


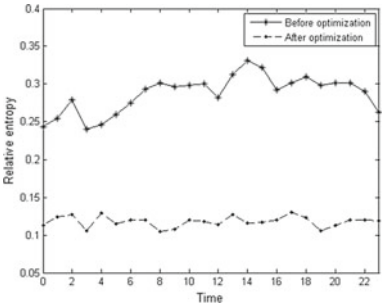

**Fig. 2.** Convergence process of genetic algorithm for optimal model of water supply pumping station



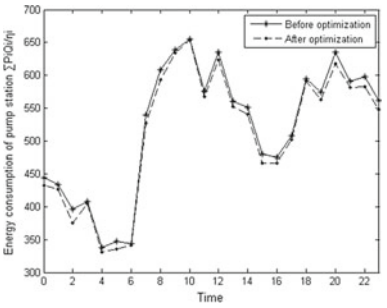

**(a)** Water pressure $P_1(m)$ in waterworks B




**(b)** Water pressure $P_2(m)$ in waterworks A




**(c)** Relative entropy




**(d)** Energy consumption of pump station $\sum P_i Q_i /_i$

**Fig. 3.** Change curve of before and after optimization of water supply pumping station

is higher, which is not conducive to energy saving. $\sum P_i Q_i/\eta_i$ before optimization is 12491.70, and $\sum P_i Q_i/\eta_i$ after optimization is 12252.58; according to this, the energy saving can be obtained:

$$\frac{12491.70 - 12252.58}{12491.70} \times 100\% = 1.91\%$$

## 5 Conclusion and Future Work

In this paper, we first discuss the necessity and goal of optimal operation of water supply pumping station, then take relative entropy and pumping station energy consumption as objective functions, and establish an optimization model of outlet water pressure of supply pumping stations based on information entropy. Genetic algorithm is used to solve the optimization model of the water supply pumping station. Besides, we explain in detail the coding method and the constraint condition processing method for the water pressure optimization of the water supply pumping station. It provides the basis of theory and method for solving the actual pipe network afterward.

## References

1. Agathokleous A, Christodoulou C, Christodoulou SE. Influence of intermittent water supply operations on the vulnerability of water distribution networks. J Hydroinform. 2017;19(6):838–52.
2. Candelieri A, Perego R, Archetti F. Bayesian optimization of pump operations in water distribution systems. J Global Optim. 2018;71(1):213–35.
3. Herrera M, Abraham E, Stoianov I. A graph-theoretic framework for assessing the resilience of sectorised water distribution networks. Water Resour Manage. 2016;30(5):1685–99.
4. Greco R. Resilience and entropy as indices of robustness of water distribution networks. J Hydroinform. 2012;14(3):761–71.
5. Saleh S, Tanyimboh TT. Multi-directional maximum-entropy approach to the evolutionary design optimization of water distribution systems. Water Resour Manage. 2016;30(6):1885–901.
6. Wan F, Yuan W, Li Q. Research on risk and early warning system of inter-basin multi-reservoir water transfer-supply operation with consideration of uncertain factors. Irrig Drainage. 2018;67:461.
7. Tanyimboh TT, Czajkowska AM. Joint entropy based multi-objective evolutionary optimization of water distribution networks. Water Resour Manage. 2018;32(8):2569–84.
8. Han R, Liu J. Spectral clustering and genetic algorithm for design of district metered areas in water distribution systems. Proc Eng. 2017;186:152–9.
9. Haghighi A. Uncertainty analysis of water supply networks using the set theory and NSGA-II. Eng Appl Artif Intell. 2014;32:270–82.
10. Bi W, Dandy GC, Maier HR. Improved genetic algorithm optimization of water distribution system design by incorporating domain knowledge. Environ Model Softw. 2015;69:370–81.

# Knowledge-Aware VNE Orchestration in Cloud- and Edge-Mixed Data Center Networks

Cunqian Yu, Rongxi He(✉), Bin Lin, Li Zhang, and Jingyu Li

College of Information Science and Technology,
Dalian Maritime University, Dalian, China
hrx@dlmu.edu.cn

**Abstract.** Emerging applications, such as mobile cloud computing and WSN, bring lots of challenges to current data center networks, and many of them need to construct their own virtual networks (VNs) for the flow-line calculation. Besides, delay-sensitive applications, like IoV and mobile 5G broadband application, require to move the computing resource and some parts of databases closed to customer sites. Thus, the computing resource near application fields, the so-called edge computing, becomes more and more in short supply; meanwhile, the backbone network is under an increasing pressure on the long-haul communication workload. In this paper, we first make a knowledge-aware of VN and design a mapping orchestration based on the QoS requirement features and workload matching. Here, we distinguish the computing requirement from each virtual node in VN and allocate them into edge or cloud DC, respectively. Then, we formulate this resource allocation problem and propose a fast and efficient algorithm. Finally, the numerical results verified the advantages of our algorithm in terms of average computing latency and average transmission latency.

**Keywords:** Optical data center · Edge computing · Virtual network embedding · 5G · Mapping orchestration

## 1 Introduction

The emerging applications from broadband access networks and mobile cell networks lead to both a large computing capacity consumption and a high bandwidth consumption in data center networks (DCNs) [1]. Even that the elastic optical network (EON) provides high-speed data transmission, some sensitive demands such as autonomous driving and WSN are hard to be satisfied due to the long-haul route and the remote computing resource in data centers (DCs). DCs contain thousands of high-capacity servers and are often placed in outer suburbs or even alpine regions. Yet sometimes, the aid is too slow in coming to be of any help, and the spectrum resource of optical backbone network encounters the bottleneck. Many transmission loads produced for cloud data centers (CDCs) crowd into several main and conspicuous routing paths [2]. Service providers approach some of computing capacity closer to the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_64

application sites and construct the smaller-size and lighter-weight edge DCs (EDCs) to provide convenient and frequently accessed services.

Both software and hardware technology progress brings a novel kind of mixed computing circumstance into being, i.e., the cloud- and edge-mixed DCN (CEM-DCN). Thus, there are two kinds of different computing resources as the candidate mapping targets for virtual node mapping, followed by two kinds of resource provisioning styles subject to the constraints of diverse delay upper limits and device rent costs. Currently, researches focus on the edge computing and the fog computing [1–3]. But few of them studied the virtual network embedding (VNE) in the CEM-DCN associated by EON, and the differential computing resource requirement of virtual nodes in the same VN demand is usually forgotten, leading to either the suboptimal state of the resource allocation or the unsatisfactory of overall delay resulting from computing and communication process.

Take Fig. 1 as an illustration. VNE usually facilitates to build customized virtual networks and resource sharing. Some flow-works are contained in VN to complete different parts of the work, and probably, one of the flow-works needs to be finished in an extremely low latency. CDC is hard to satisfy this work in low latency because of its coherent shortage in physical distance between customers and DC geosites. At this time, a high-efficient EDC could provide basic computing ability to facilitate customers. And the placement of EDCs is more flexible.

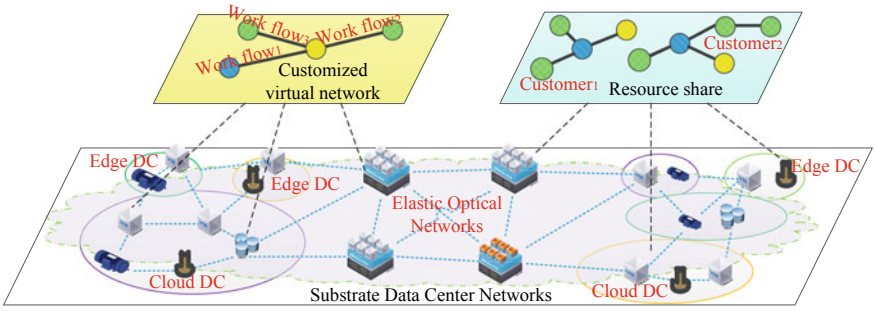



**Fig. 1.** Differential VNE processes in CEM-DCN

In this paper, we first formulate these series of problem into the mathematical model and make a comprehensive analysis of the root season of exorbitant delay and resource provision unbalance. We propose a knowledge-aware VNE orchestration that makes the virtual nodes in the same VN map into EDC/CDCs with differential delay and computing resource requirement strategies. Besides, a heuristic algorithm is proposed to bring a fast VNE approach in CEM-DCN and achieve a near-optimal computing- and communication-resource usage.

## 2 Motivation and Knowledge-Aware VNE Orchestration

### 2.1 Motivation

Currently, the large-size CDCs make many outstanding contributions to ICT development. However, emerging services ram the spectral resource and computing resource in traditional DCNs, such as 5G applications and AR/VR. Actually, the edge computing as an expected alternative scheme is about to solve some of the inherent problems based on its equipped lightweight computing capacity. More importantly, EDC which is closer to the front end of the service has the abilities of task recognition and workload offloading. It can slice the origin task into two pieces and push the heavy workload that is unnecessary to be solved at the front end to CDC. That makes the resource allocation more effective and alleviates the spectrum pressure of core networks.

VN adapts virtualization and cloud computing technology and can facilitate the customized virtual private network and substrate resource sharing. Generally, a VN consists of virtual nodes (actually virtual machine requirement, and here we use VM for short) and virtual links (VLs). Traditionally, VMs should be mapped into several individual servers in the large-scale DC and VLs are mapped onto the substrate routing path in core networks. Nowadays, with the introduction of EDC into the entire DCN system, some of VMs are more compatibly mapped into edge servers to gain faster reactions and cache temporary coarse data. Thus, correspondingly, the physical length of VLs is substantially shortened and their composition is also more complex, maybe the combination of wired and wireless links.

Our objective is to minimize the latency sum from both data transmission and computing process during the overall VNE procedures, meanwhile without losing resource allocation efficiency. Correspondingly, there are some inescapable constraints and pending problems, detailed as follows:

(1) The resource capacity constraints of both core network spectrum and DC computing capacity.
(2) Which VM should be or is preferred to be mapped into EDC requires an accuracy judgment mechanism.
(3) Except the VM mapped into EDC, other VMs should be mapped into CDCs under the constraints of computing capacity and process latency.
(4) In the same VNE request, diverse VMs must be mapped into different DCs; in other words, one DC can and can only accommodate one of VMs in the same VNE request.
(5) When mapping VLs, the latency constraint must be subjected to.
(6) In the same VNE request, VLs must be link-disjoint; in other words, their substrate transmission channel cannot overlap each other.

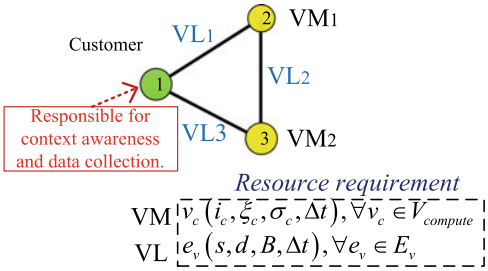
### 2.2 Knowledge-Aware VNE Orchestration

Traditionally, the VNE strategy trends to map virtual nodes into the most computing-resource sufficient DCs and map virtual links into the shortest route paths or the most

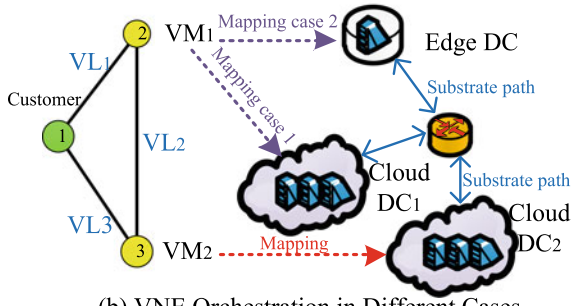
spectrum-resource sufficient paths to achieve load balance. However, actually, with the growth in VN demands, the embedding behavior of VNs demands looks forward the dedifferentiated services to satisfy diverse QoS requirements.

Based on SLA, the information of VNE demands is of foreknowledge, such as the delay upper limit and various kinds of resource requirements. That assists us to judge whether this VNE request is mapped into CDC or into EDC.

Usually, service request starts from the customer who is at the edge of networks and is responsible for context awareness and data collection, as shown by node 1 in Fig. 2. And this request needs the assistant of computing and data storage from high-performance server, and by virtue of generalized virtualization technology and server slice technology, these computing and storage workloads can be abstract into the form of virtual machine (VM), as shown by nodes 2 and 3 in Fig. 2a. Under the situation of abundant resource in networks, $VM_1$ and $VM_2$ could be mapped into two diverse remote CDCs for professional and cheap service, as shown in Case 1 of Fig. 2b. However, if the request needs the high QoS guarantee like low latency, one of VMs (or even all VMs) must search for the service terminal with closer distance, and in this paper, the constructed EDC is an excellent choice. As shown in Case 2 of Fig. 2b, the $VM_1$ is mapped into EDC, due to its knowledge property and the more stressful resource environment.




(a) VNE request




(b) VNE Orchestration in Different Cases

**Fig. 2.** Knowledge-aware VNE illustration

On the other hand, in the extreme case of tenser resource, EDC could provide the function of computing task offload. EDC leaves the necessary workload in its own site and slices the original workload and push the rest part of workload, usually heavy task, to CDC. Thus, from the state of VN knowledge-aware to the state of detailed mapping procedures, the entire orchestration of knowledge-aware VNE can be summarized in Fig. 3.

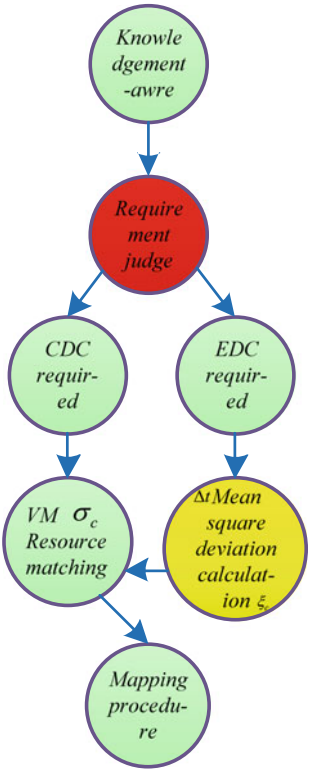



**Fig. 3.** Knowledge-aware VNE orchestration

## 3 Problem Statement

We design an orchestration for VNE request recognition and mapping approach assignment. Since different customers may have diverse QoS requirements, the homogeneous VN constructing method would lead to VNE request dissatisfaction and resource mismatch.

To implement the differential mapping procedure, in one VNE request, we denote the VM as $v_c(j_c, \sigma_c, \xi_c, \overline{t_c}), \forall v_c \in R_{\text{vm}}$ where $j_c$ represents the VN knowledge-aware judgment result (binary, $j_c = 1$ means the current VN needs EDC), besides $\sigma_c$ is the

computing resource requirement, $\overline{t_c}$ is the computing latency upper bound, and $\xi_c$ is the mean square deviation (MSD) between other VMs in the same VNE request. And then we denote the VL as $e_v(s, d, b_{\text{vl}}, \overline{t_{\text{vl}}}), \forall e_v \in R_{\text{vl}}$ where $s$ is the source node, $d$ is the destination node, $b_{\text{vl}}$ is the bandwidth requirement, and $\overline{t_{\text{vl}}}$ is the delay upper bound.

Our objective is to minimize the total latency $\psi$ of both VMs and VLs, represented by Eq. 1, where $\Delta t_c$ and $\Delta t_l$ denote respectively the actual latencies resulted from allocated computing resource and allocated spectrum resource in CEM-DCN. And $\Delta t_c$ and $\Delta t_l$ can be calculated by Eqs. 2 and 3, where $\sigma_{\text{dc}}$ is the computing resource actually allocated by CDC/EDC and $b_f$ is the spectrum slot number actually allocated by the core network.

$$\textbf{\textit{Minimizing}}\ \psi = \sum_{\forall \text{VM} \in \text{VN}} \Delta t_c + \sum_{\forall \text{VL} \in \text{VN}} \Delta t_{\text{vl}}, \forall \text{VN} \in R_{\text{VNE}} \tag{1}$$

$$\Delta t_c = \frac{\overline{t_c} \cdot \sigma_c}{\sigma_{\text{dc}}}, \Delta t_c \leq \overline{t_c}, \sigma_c \leq \sigma_{\text{dc}} \tag{2}$$

$$\Delta t_{\text{vl}} = \frac{\overline{t_{\text{vl}}} \cdot b_{\text{vl}}}{b_f}, \Delta t_{\text{vl}} \leq \overline{t_{\text{vl}}}, b_{\text{vl}} \leq b_f \tag{3}$$

Since the computing resource in CDC is relatively richer than that in EDC, thus for the calculation of computing latency sum, the delay caused by EDC plays a decisive role. On the other hand, both the distance sum between EDC and all the selected CDCs and the bandwidth that may be allocated to the routing path (result of the VL mapping) impact the calculation of transmission latency sum.

$$\xi_c = \frac{\sum_{k=1}^{|R_{\text{vm}}|} \left(\sigma_{c,k} - \sigma_{c,i}\right)^2}{|R_{\text{vm}}|}, \quad \text{if } j_c = 1, \forall v_{c,i} \in R_{\text{vm}} \tag{4}$$

$$v_{c,\text{EDC}} = \text{Min}\left\{\xi_{c,1}, \xi_{c,2}, \ldots, \xi_{c,|R_{\text{vm}}|}\right\} \tag{5}$$

Thereby, which VM could be mapped into EDC seems particularly important, and we design Eqs. 4 and 5 to select which VM to be mapped into EDC. Equation 4 uses the idea of MSD to minimize the actual difference of computing delay among all VMs, and Eq. 5 compares MSDs to pick the minimum. The calculation of Eq. 4 is grateful in the scenario of either massive VNEs or network congestion, since the resource EDC is essentially scarce. And the physical location selection of candidate CDCs and the spectrum slot allocation of VLs is about to be carried out in Sect. 4, due to the NP-hard feature of the multi-commodity flow problem.

Besides, we still require some constraints on the computing capacity of DCs, the spectrum slot capacity of fiber links, and the data flow conservation. However, limited by the length of space, it is not necessary to show them all here, and please see the references [4].

## 4 Heuristics

We need a faster and effective approach for VNE in EDM-DCN. We proposed a novel knowledge-aware and integrated VNE algorithm (KIVA), and in KIVA, the VM requirement recognition, VNE orchestration, EDC/CDC selection, and the optimal allocation of resource are taken into account.

The pseudocode of KIVA is shown in Table 1. First, we sort the sufficient resource CDCs according to their residual resource in descending order, and then, we select the top CDC to undertake VMs. And we search multiple paths from the source CDC to the target EDC and put them into a candidate path set. We map a VL to the shortest and spectrum slot richest path in the candidate set (based on the product of hops and slots), and the above procedures will facilitate to find these optimal CDC locations and to decrease the transmission latency sum.

**Table 1.** Pseudocode of KIVA

*Algorithm : KIVA*

**Input**: $G_{substrate}$ and $R_{VNE}$ .

**Output**: $MAP_{node}$ and $MAP_{link}$ .

$V^{available} \leftarrow \{G(V_{DC}), descending, C_{DC}(computing\ resource)\}$ ;

$E^{available} \leftarrow \{G(Fiber\ links), C_{slot}(e) \geq b_{vl}\}$ ;

**while** $V^{available} \neq \varnothing$ **do**

  $V_{DC} = V^{available}.top()$ ;

  $MAP_{node}(v_c, V_{DC})$ ;

  $P^{candidate}_{CDC,EDC} \leftarrow \{p_{CDC,EDC} = shortestpath[V_{DC}, \forall V_{EDC} \in V^{avaiable}_{EDC}], B(p_{CDC,EDC}) \geq b_{vl}\}$ ;

  $P^{candidate}_{CDC,EDC} \leftarrow \{p_{CDC,EDC}, descending, B(p_{CDC,EDC})\}$ ;

  **for** $\forall e_V \in E_{vl}$ **do**

    $MAP_{link}[e_v, P^{candidate}_{CDC,EDC}.top()]$ ;

    $MAP_{node}(v_c, V_{DC}) \leftarrow Vertices(p_{EDC,EDC})$ and $P^{candidate}_{CDC,EDC}.pop()$ ;

  **end for**

  **if** all $\{MAP_{node}, MAP_{link}\}$ successful **then return** $\{MAP_{node}, MAP_{link}\}$ ;

  **else then** $V_{DC}.pop()$ ;

  **end if**

**end while**

## 5 Simulation

We take VC++ 2015, at the 64-bit Core i7 2.50 GHz computer with 8 GB RAM, to achieve our KIVA. The benchmark is the ordinary first-fit VNE algorithm (FFVA) tailored to CEM-DCNs [5]. We take the substrate topology in Fig. 4 as the simulation topology, and each edge is a bidirectional fiber link with a given length in kilometers. We assume that the bandwidth of each FS is 12.5 GHz and each fiber link can accommodate 358 FSs. The metrics are the average computing latency and the average

routing path latency per VNE (*average latency/path* for short and 4.868 μs/km [6]). Moreover, each CDC contains 2000 units of computing resource, and each EDC contains 100 units of that.

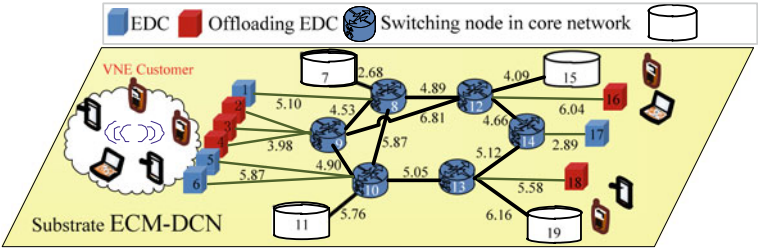



**Fig. 4.** Simulation topology

In Fig. 5, with the increasing VNE requests, both curves increase, and the average computing latency of KIVA is lower than that of FFVA. And even at the load 140, KIVA can achieve a 31% deduction of average computing latency. That is because we consider both the MSD of computing requirement of VMs when mapping EDC and the load balancing when selecting CDCs.

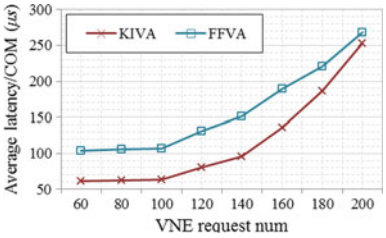



**Fig. 5.** Average computing latency comparison

In Fig. 6, the increasing VNE requests, two curves increase, and KIVA performs better than FFVA does. And KIVA can achieve a nearly 40% deduction in average path latency since KIVA considers both geodistance closer EDCs and the optimal geolocations of CDCs.

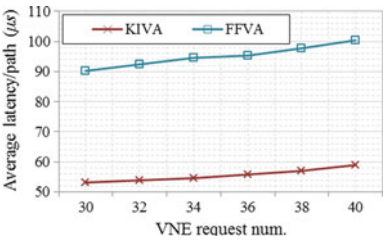



**Fig. 6.** Average path latency comparison

## 6 Conclusion

Emerging applications bring lots of challenges in latency and resource provisioning. We designed a knowledge-aware VNE orchestration based on the analysis of QoS requirement features. We formulate this resource allocation problem and propose a fast and efficient algorithm to facilitate the distinguished VNE process. Finally, the numerical results verified the advantages of our algorithm in terms of average computing latency and average path latency.

**Acknowledgments.** The work in this paper is funded by Fundamental Research Funds for Central Universities (Grant No. 3132016318, 3132017078 and 3132018181) and National Natural Science Foundation of China (Grant No. 61371091).

## References

1. Yang X, Chen Z, Li K, Sun Y, Liu N, Xie W, Zhao Y. Communication-constrained mobile edge computing systems for wireless virtual reality: Scheduling and tradeoff. IEEE Access. 2018;6:16665–77.
2. Tan Z, Yu FR, Li X, Ji H, Leung VCM. Virtual resource allocation for heterogeneous services in full duplex-enabled SCNS with mobile edge computing and caching. IEEE Trans Veh Technol. 2018;67(2):1794–808.
3. Wang W, Zhao Y, Tornatore M, Gupta A, Zhang J, Mukherjee B. Virtual machine placement and workload assignment for mobile edge computing. In: Proceedings of CloudNet, Prague, Czech Republic; 2017. p. 1–6.
4. Yu C, Guo L, Hou W. Novel elastic optical network embedding using re-optimized VCAT framework accompanied by hitless PPSM function. IEEE J Lightwave Technol. 2016;34 (22):5199–213.
5. Tzanakaki A, Anastosopoulos MP, Peng S, Rofoee B, Yan Y, Simeonidou D, Landi G, Bernini G, Ciulli N, Riera J F, Escalona E, Garcia-Espin JA, Katsalis K, Korakis T. A converged network architecture for energy efficient mobile cloud computing. In: Proceedings of ONDM, Stockholm, Sweden; 2014. p. 120–5.
6. Huang S, Rai S, Mukherjee B. Survivable differential delay aware multi-service over SONET/SDH networks with virtual concatenation. In: Proceedings of OFC/NFOEC, Anaheim, CA; 2007. p. 1–3.

# Improved Preamble-Based Symbol Timing Synchronization Schemes for GFDM Systems

Chenglong Yang, Ying Wang(✉), Zhongwen Zhang, and Bin Lin

College of Information Science and Technology, Dalian Maritime University, Dalian 116026, China
wangying@dlmu.edu.cn

**Abstract.** Generalized frequency division multiplexing (GFDM) is a flexible multicarrier modulation scheme and has been considered as a candidate for the physical layer of the 5G mobile communication system. The bit error rate performance of GFDM is affected by the symbol timing synchronization errors; however, the existing preamble-based symbol timing synchronization scheme for GFDM systems suffers from the timing-ambiguous plateau problem. In this paper, two improved preamble structures are constructed according to the characteristics of GFDM symbols to eliminate the timing-ambiguous plateau. The simulation results under AWGN and multipath fading channels show that the proposed autocorrelation timing synchronization schemes have better timing MSE performance compared to the existing preamble-based timing synchronization schemes.

**Keywords:** 5G · GFDM · Synchronization · Autocorrelation

## 1 Introduction

With the increasing complexity of communication environment and the diversification of the application requirements, the mobile communication technology is being developed to the fifth generation (5G) wireless communication. The requirements of the new generation communication systems for the physical layer interface have also changed. At present, OFDM systems are widely adopted in modern wireless communication systems because of its easy implementation and robustness against frequency selective multipath channels. However, some problems with OFDM make it inappropriate to be utilized in some 5G service scenarios. For example, in cognitive radio systems, the second users using OFDM access fragmented and dynamic spectrum holes may cause severe interference to the primary user due to the serious out-of-band (OOB) radiation problem of OFDM [1]. In addition, OFDM adopts a large cyclic prefix to combat inter-symbol interference caused by the multipath fading channel; as a result, it is not an appropriate choice for the tactile Internet where a time delay smaller than 1 ms should be satisfied. Recently, a new alternative multicarrier modulation scheme, namely generalized frequency division multiplexing (GFDM) [2], has been proposed as a candidate for the next generation wireless communication systems. The GFDM has the advantages of low OOB radiation and high spectrum efficiency.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_65

Although GFDM can tolerate a relative larger timing offset (STO) and carrier frequency offset (CFO) in comparison with OFDM, the synchronization errors can still have a serious influence on the bit error rate performance of GFDM systems, so synchronization algorithm should be carefully designed for the GFDM receiver. There are currently three major data-assisted synchronization schemes for GFDM. Ivan [3] proposed a scheme of synchronization using a preamble with two identical subsymbols. This scheme integrates several classical OFDM synchronization methods and limits OOB radiation with window function, but this scheme requires additional computations to eliminate the timing-ambiguous plateau. In [4], an embedded midamble synchronization scheme is proposed, and this scheme needs preliminary demodulation to obtain the information coefficient vector for synchronization. In order to improve the bandwidth utilization [5], proposed a superimposed synchronization scheme to meet the OOB requirements of GFDM system. The main work of this paper is to solve the problem of ambiguous plateau for the timing synchronization scheme in [3] and to reduce the computational complexity.

The paper is organized as follows: Sect. 2 describes the model of GFDM system, and Sect. 3 proposes two improved preamble synchronization schemes for GFDM. Simulation results and analysis of the proposed schemes presented in Sects. 4 and 5 provide conclusions.

## 2 The Model of GFDM

Suppose a GFDM symbol consists of $K$ subcarriers and $M$ subsymbols; thus, the baseband transmitted signal can be described by [1]

$$x[n] = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} g_{k,m}[n] d_{k,m} \tag{1}$$

where $g_{k,m}[n]$ denotes the time-shifted and frequency-modulated version of a prototype filter $g[n]$,

$$g_{k,m}[n] = g[(n - mK) \bmod N] \cdot \exp\left(j2\pi \frac{k}{K} n\right) \tag{2}$$

For convenience, Eq. (1) can also be expressed in matrix form [1],

$$\mathbf{x} = \mathbf{A}\mathbf{d} \tag{3}$$

where $\mathbf{A}$ is $KM \times KM$ modulation matrix, $\mathbf{d}$ is $KM \times 1$ data vector.

The received signal through an $L$-path multipath channel at the receiver can be given by

$$z[n] = \sum_{l=0}^{L-1} h[l] x[n - l] + w[n] \tag{4}$$

where $h[l]$ is the channel impulse response, and $w[n]$ denotes the additive white Gaussian noise(AWGN), $w[n] \sim CN(0, \sigma_w^2)$. When there exist STO and CFO between the GFDM transmitter and receiver, consequently, the baseband signal at the receiver can be expressed as

$$r[n] = z[n-\theta]\mathrm{e}^{j2\pi\frac{\varepsilon}{K}n} \tag{5}$$

where $\theta$ denotes the STO, and $\varepsilon$ denotes the CFO normalized to the subcarrier spacing. Before the receiver estimates the transmitted data, STO and CFO need to be estimated and compensated to achieve timing and frequency synchronization.

## 3 The Proposed Timing Synchronization Schemes

### 3.1 Conjugate Symmetric Preamble-Based Synchronization

The pulse shaping filters of a GFDM system with $K$ subcarriers and $M$ subsymbols are obtained by cyclic time shifting and frequency modulating a prototype filter with a length of $N = KM$. When a GFDM symbol consists of two subsymbols, i.e., $M = 2$, the filters will conjugate symmetric about their centers. So we can construct a training vector with the following structure

$$\mathbf{d}_{p1} = (\mathbf{c}_1^T, \mathbf{c}_1^H, \mathbf{c}_1^H, \mathbf{c}_1^T)^T \tag{6}$$

where $(\cdot)^H$ denotes conjugate transpose, and $(\cdot)^T$ denotes transpose. $\mathbf{c}_1$ is a PN sequence of length $K/2$, and $\mathbf{c}_1 = (c_1[0], c_1[1], \ldots, c_1[K/2-1])^T$.

Then, the proposed conjugate symmetric preamble can be generated by

$$\mathbf{x}_{p1} = \mathbf{A}_{2K\times 2K}\mathbf{d}_{p1} \tag{7}$$

where $\mathbf{x}_{p1} = \left(x_{p1}[0], x_{p1}[1], \ldots, x_{p1}[2K-1]\right)^T$, $\mathbf{A}_{2K\times 2K}$ is a $2K \times 2K$ GFDM modulation matrix. The preamble generated by (7) has the following properties,

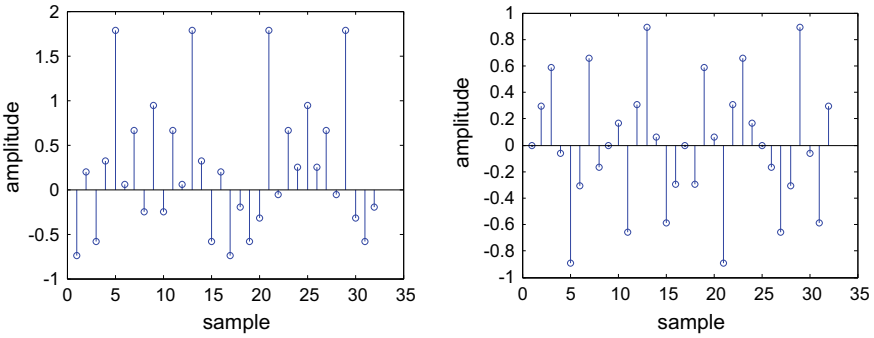
$$x_{p1}[n] = x_{p1}[K-n]^*, \quad n = 0, \ldots, K/2 \tag{8}$$

$$x_{p1}[K+n+1] = x_{p1}[2K-n-1]^*, \quad n = 0, \ldots, K/2-1 \tag{9}$$

$$x_{p1}[n] = x_{p1}(n+K)\cdot(-1)^{n+K}, \quad n = 0, \ldots, K-1 \tag{10}$$

where $[\cdot]^*$ denotes conjugate.

To illustrate the structure of the proposed preamble, an example with the number of subcarriers being 16 is given. The corresponding real and imaginary parts of the preamble are shown in Fig. 1.




(a) Real part (b) Imaginary part

**Fig. 1.** First improved preamble symbol

To make use of the conjugate symmetric property of the preamble, the improved conjugate symmetric preamble-based synchronization (CSPBS) metrics can be defined as

$$P_N[n] = \frac{|P_A[n]|^2}{(R_A[n])^2} \tag{11}$$

where

$$P_A[n] = \sum_{k=0}^{K-1} r[K+k+n] \cdot r[K-k+n] \cdot (-1)^{K+k+n} \tag{12}$$

$$R_A[n] = \sum_{k=0}^{K-1} |r[K+k+n]|^2 \tag{13}$$

Figure 2 shows the improved timing metric under ideal channel condition, where the number of subcarriers is $K = 128$ and the length of CP is $N_{\mathrm{CP}} = 32$. The proposed scheme has an impulse-shaped timing metric and eliminates the timing-ambiguous plateau.

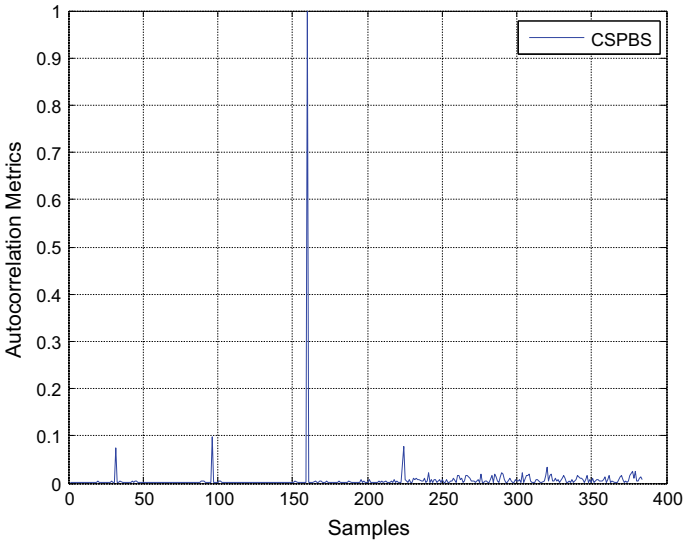



**Fig. 2.** CSPBS's autocorrelation metric

### 3.2 Cyclic Preamble-Based Synchronization

The two subsymbols of the conjugate symmetric preamble are not exactly the same, so this paper proposes another cyclic preamble-based synchronization (CPBS) scheme to solve the timing-ambiguous plateau problem. The vector used to generate a cyclic preamble is designed to be

$$\mathbf{d}_{p2} = \left( \underbrace{0,\ldots,0}_{n_1}, \mathbf{c}_2^T, \underbrace{0,\ldots,0}_{n_2}, \underbrace{0,\ldots,0}_{n_1}, \mathbf{c}_2^T, \underbrace{0,\ldots,0}_{n_2} \right)^T \tag{14}$$

where $\mathbf{c}_2 = (c_2[0], c_2[1], \ldots, c_2[K - n_1 - n_2 - 1])^T$ denotes a real PN sequence, $n_1$ and $n_2$ represent the number of zeros in the PN sequence, and in this paper $n_1 = n_2 = 1$.

The cyclic preamble with two subsymbols can be obtained by the following modulation process,

$$\mathbf{x}_{p2} = \mathbf{A}_{2K\times 2K}\mathbf{d}_{p2} \tag{15}$$

where $\mathbf{x}_{p2} = \left(x_{p2}[0], \ldots, x_{p2}[2K-1]\right)^T$.

To clearly explain the characteristics of the cyclic preamble, Eq. (1) can be rearranged as

$$x[n] = \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} g[(n - mK) \bmod N] \cdot \exp\left(j2\pi\frac{k}{K}n\right) \cdot d_{k,m} = \sqrt{K}\mathbf{W}_K^H\mathbf{d}_k \tag{16}$$

where $\mathbf{W}_K$ is the $K \times K$ dimension DFT matrix, $\mathbf{d}_k = (d_k[0], d_k[1], \ldots, d_k[2K-1])^T$, and

$$d_k[n] = \sum_{m=0}^{M-1} g[(n - mK) \bmod N] \cdot d_{k,m} \tag{17}$$

It can be seen from (16) that up-conversion operation in the GFDM modulation process can be regarded as an IFFT conversion, so the resulted cyclic preamble has the following characteristics:

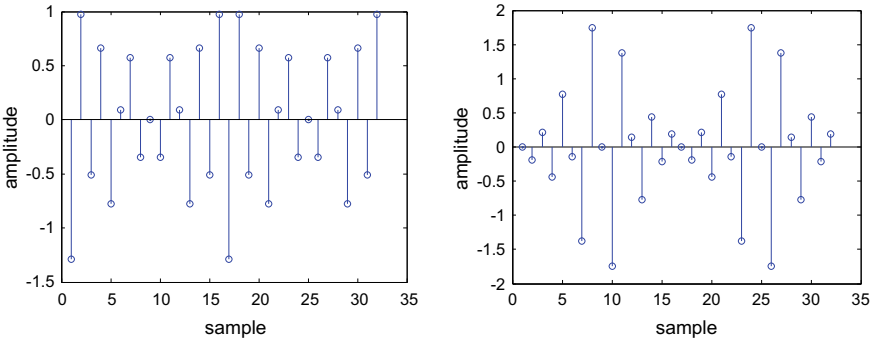
$$x_{p2}[n] = x_{p2}[K-n]^*, \quad n = 0, \ldots, K/2 \tag{18}$$

$$x_{p2}[K+n+1] = x_{p2}[2K-n-1]^*, \quad n = 0, \ldots, K/2-1 \tag{19}$$

And $\mathbf{x}_{p2}$ is composed of two identical PN sequences, so

$$x_{p2}[n] = x_{p2}(n+K), \quad n = 1, \ldots, K \tag{20}$$

When the number of subcarriers is $K = 16$, the real and imaginary parts of the cyclic preamble are shown in Fig. 3.




(a) Real part (b) Imaginary part

**Fig. 3.** Second improved preamble symbol

For the cyclic preamble, the autocorrelation timing metrics can still be calculated using (10), but the autocorrelation needs to be changed to

$$P_A[n] = \sum_{k=1}^{K-1} r[K+k+n] \cdot r[K-k+n] \tag{21}$$

Compared to (12) and (21) is more convenient for calculating the autocorrelation.

The timing metrics curve for the CPBS scheme is shown in Fig. 4. It can be seen from the figure that this algorithm also solves the problem of the timing-ambiguous plateau in the existing preamble synchronization scheme.

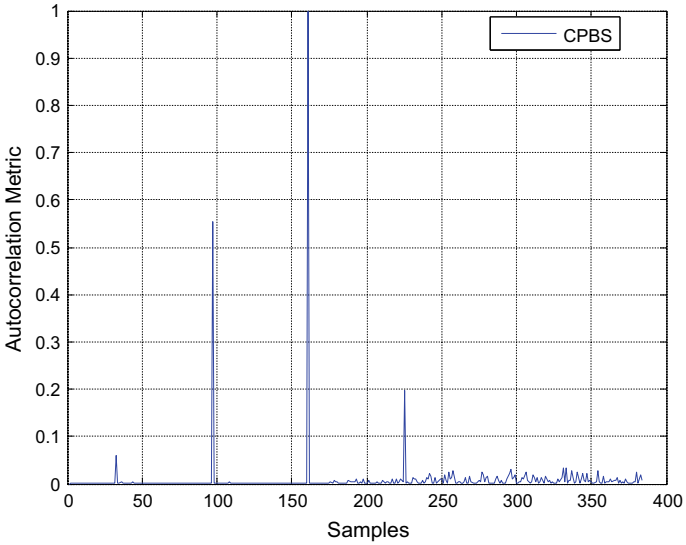



**Fig. 4.** Autocorrelation metric of the CPBS synchronization scheme

## 4 Simulation Results

To analyze the performance of the proposed synchronization scheme, computer simulations were performed. In the simulation, QPSK is utilized to get transmitted data, the RC filter with roll-off factor of 0.1 is used, and the number of subsymbols is 5 in GFDM system. At the same time, AWGN channel and multipath fading channel were considered. As shown in Table 1, the multipath channel is an Extended Pedestrian A (EPA) [6] channel model with a maximum delay of 410 ns. The EPA is an LTE channel model defined by 3GPP Technical Recommendation (TR) 36.104.

**Table 1.** LTE EPA channel parameters

| Tap | Relative delay (ns) | Average power (dB) |
|---|---|---|
| 1 | 0 | 0.0 |
| 2 | 30 | −1.0 |
| 3 | 70 | −2.0 |
| 4 | 90 | −3.0 |
| 5 | 110 | −8.0 |
| 6 | 190 | −17.2 |
| 7 | 410 | −20.8 |

Figures 5 and 6 show the mean square error (MSE) of STO estimation using autocorrelation metrics in AWGN channel and the EPA channel, respectively. Under AWGN channel condition, $K = 64$, $N_{CP} = 16$, while under EPA channel condition, $K = 256$, $N_{CP} = 32$. As can be seen from these two figures, compared with the existing preamble synchronization scheme in [3], the two preamble-based schemes

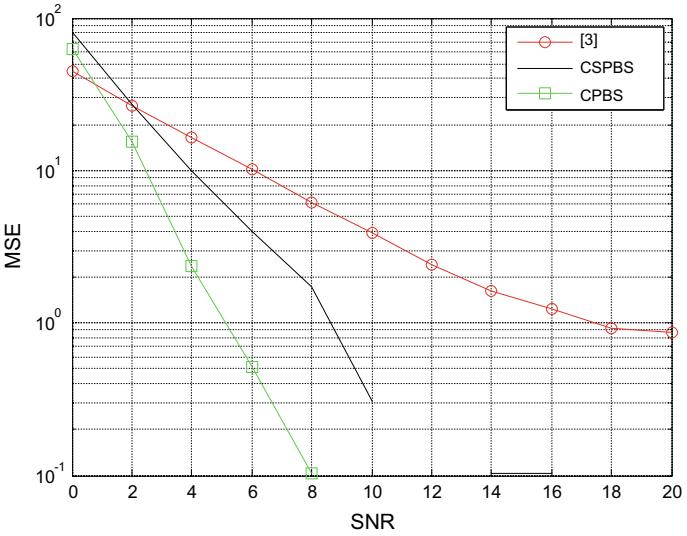



**Fig. 5.** MSE of STO estimation using autocorrelation metric in AWGN channel

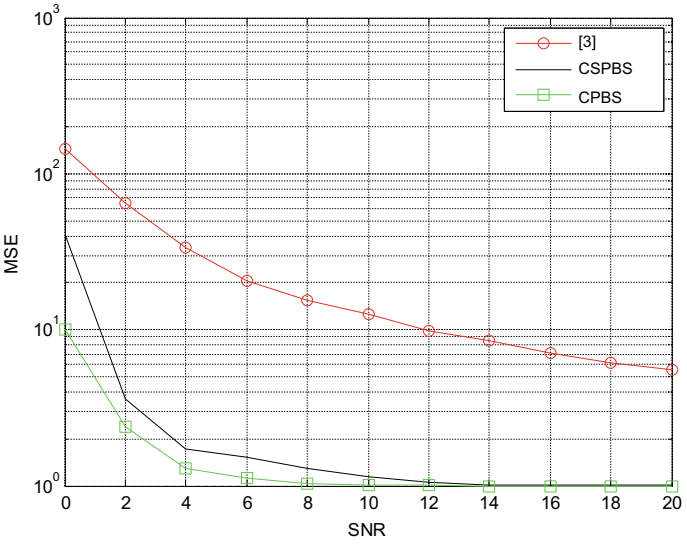



**Fig. 6.** MSE of STO estimation using autocorrelation metric in EPA channel

proposed in this paper have better MSE performance. And the simulations also indicate that the CPBS scheme is better than the CSPBS scheme. The first subsymbol of the CSPBS is not equal to the second subsymbol, and the autocorrelation timing metric requires an additional operation of $(-1)^n$, so compared with the CPBS scheme, the robustness of the CSPBS scheme is relatively weak.

## 5 Conclusion

In this paper, in order to solve the timing-ambiguous plateau in the existing preamble synchronization scheme for GFDM system, two new structures of the preamble are redesigned, and the corresponding autocorrelation timing metrics are given. The simulation results show that neither of the two preamble synchronization schemes proposed in this paper generate a timing uncertain plateau, which reduces the computational complexity and simplifies the receiver design. And compared with the existing preamble synchronization scheme, the proposed autocorrelation timing metrics have better MSE performance of the STO estimation.

**Acknowledgments.** This work was supported in part by Fundamental Research Funds for Central Universities under grant No. 3132016318.

## References

1. Nicola M, Maximilian M, Ivan SG, et al. Generalized frequency division multiplexing for 5th generation cellular networks. IEEE Trans Commun. 2014;62(9):3045–61.
2. Gerhard F, Marco K, Steffen B. GFDM—Generalized frequency division multiplexing. In: IEEE 69th vehicular technology conference; 2009. p. 1–4.
3. Ivan SG, Luciano LM, Nicola M, Gerhard F. A synchronization technique for generalized frequency division multiplexing. EURASIP J Adv Signal Process. 2014;67:1–10.
4. Ivan SG, Gerhard F. An embedded midamble synchronization approach for generalized frequency division multiplexing. In: IEEE global communications conference; 2015. p. 1–5.
5. Chenglong Y, Ying W, Tingting Y, Bin L. Superimposed training for time synchronization in GFDM systems considering out of band emission. In: The 4th international conference on systems and informatics (ICSAI); 2017. p. 944–948.
6. Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception 2015. http://www.etsi.org/deliver/etsi_ts/136100_136199/136104/12.06.00_60/ts_136104v120600p.pdf.

# Research on Modeling of Maritime Wireless Communication Based on Bayesian Linear Regression

Yanshuang Han(✉) and Sheng Bi

Information Science and Technology School,
Dalian Maritime University, Dalian 116000, China
hysh0105@126.com, bisheng@dlmu.edu.cn

**Abstract.** Maritime wireless communication is in the stage of continuous development. Medium and short wave communication has always been the main method of maritime communications. Realizing rapid, efficient, and reliable signal transmission is the urgent need for the development of current situation. In this paper, aiming at MF, HF, or VHF radio waves emitted by land, Bayesian linear regression model is used to solve the reflection problem of calm sea surface, and then EM (expectation maximization) algorithm is further used to solve the complex Bayesian model to solve the reflection problem of rough sea surface. Furthermore, the advantages of the Bayesian linear regression model over other models, such as the Longley-Rice model, are obtained.

**Keywords:** Maritime wireless communication · Bayesian linear regression · EM algorithm

## 1 Introduction

For a long time, medium and short wave communication has been widely used in various departments such as government, diplomacy, meteorology, and commerce for its advantages of strong flexibility, strong anti-damage ability, and long communication distance, which is used for the transmission of information such as languages, characters, images, and data [1]. At the same time, it is also an essential means of communication for high-altitude flight and maritime navigation, especially in the military field; it has always been one of the important means of military command.

With the development of science and technology, the communication efficiency between ships and shores is higher required. At present, China has established a number of medium and short wave shore stations. Many ships already have medium and short wave radio stations and radio navigation equipment that can quickly realize effective positioning. Communication facilities between ships and shores have developed from a single modulation method and analog signal transmission to various communication methods such as satellite communication and frequency modulation radio telephone [2], among which, medium and short wave communication is an important means of long-distance communication at sea.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_66

At present, maritime radio wave transmission models can be roughly divided into three categories: empirical model, semi-empirical and semi-deterministic model and deterministic model. Empirical model is a formula concluded after analyzing a large number of measurement data. It is a relatively simple radio wave transmission model, which is more convenient to use and has fewer restrictions. However, it is usually only suitable for predicting the propagation of small scenes and short-range radio waves, and the accuracy of predicting the path loss is also relatively low [3]. Semi-empirical and semi-deterministic model is a formula generated by putting deterministic methods into certain specific environments and is a statistical model. Among these models, the more classical ones are Longley-Rice model, Walfish-Bertoni model, Ikegami model, etc. [4]. This model is usually used to analyze and predict the radio wave transmission in some specific terrain environment, and the prediction accuracy is not very high. The model is suitable for predicting the propagation of radio waves in urban areas and suburbs [5]. The deterministic model is a mathematical expression derived from Maxwell equations based on electromagnetic theory. At present, most of the deterministic models are obtained by ray tracing electromagnetic method [6]. Due to the differences in the propagation environment of radio waves, there is no universal radio wave propagation model suitable for all radio wave transmissions [7].

Among them, the Longley-Rice radio wave transmission model is based on the radio wave propagation theory and combines the measured data in many actual environments. It is precise because the model is based on the theory of radio wave propagation and adds a lot of actual measurement data that the model has been widely used [8]. According to the relevant documents found, it can be known that the sea surface reflection intensity obtained by Longley-Rice radio wave transmission model and the reflection intensity of the calm sea surface is greater than that of the rough sea surface.

In this paper, Bayesian linear regression model is used to model and simulate maritime radio transmission, and the differences between this model and Longley-Rice radio wave transmission model can be obtained.

## 2 Bayesian Linear Regression Model

When carrying out radio wave propagation on the ocean surface, it is assumed that the noise is white Gaussian noise and the transmission loss is the loss of free space. Since ships receive weather and traffic report when moving on the sea, so they will inevitably keep away from some dangerous and severe-condition area. Therefore, there is no need for considering extremely violent sea condition when we simulate the environment in the model. The undulating waves model is simulated and shown with the three-dimensional wave in Fig. 1.

The radiation intensity of the reflected wave is related to the reflectivity $R$ and the transmittance $T$, but the two variables are not independent. Since the model focuses on

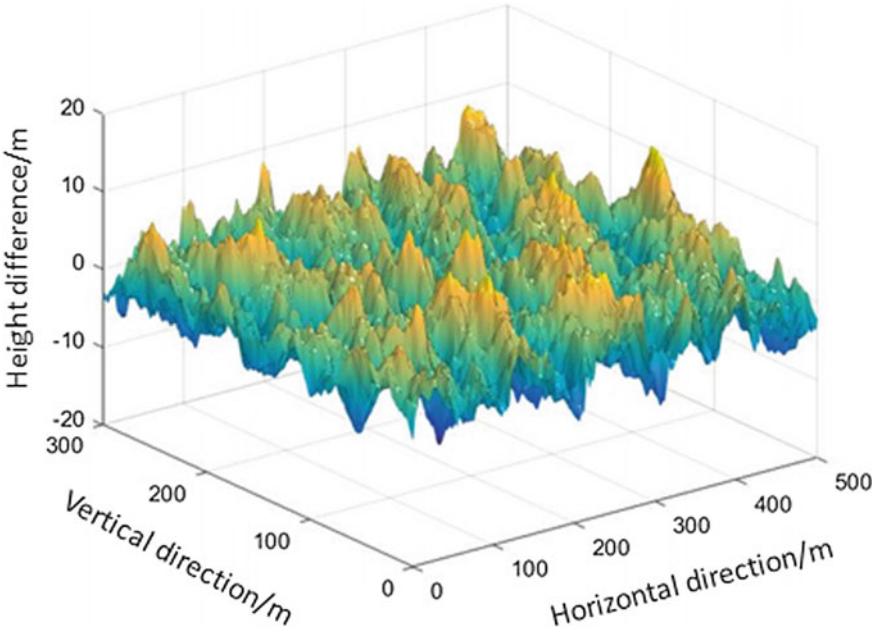



**Fig. 1.** Three-dimensional wave simulation

the light reflections, only $R$ needs to be taken into consideration. And it is assumed that $R$ obeys Gaussian distribution (normal distribution), that is:

$$R \sim N\left(\mu_r, \sigma_r^2\right) \text{ or } N\left(R|\mu_r, \sigma_r^2\right) \tag{1}$$

Then the relationship needs to be determined between reflection strength $y$ and some independent variables $x$ ($x$ is a $n \times 1$ vector) which influence $y$. Firstly, the simplest regression model is used to describe their relationship:

$$y(x, \omega) = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n \tag{2}$$

It is seen from the equation that it is suitable for linear relation, but in most cases, elements do not just meet the linear relation, so the expression of the equation above and $x$ is altered as:

$$y(x, \omega) = \omega_0 + \sum_{j=1}^{M-1} \omega_j \phi_j(x) \tag{3}$$

where $\phi_j(x)$ is called basis function. When nonlinear basis function is used, $y(x, \omega)$ will become a nonlinear function. There are many forms for $\phi_j(x)$, such as $\phi_j(x) = \exp\left\{-\frac{(x-\mu_j)^2}{2S^2}\right\}$.

The most important part of the model is the Bayesian linear regression model. It can self-adjust parameters according to the scale of data. In practical application, the model

can effectively reduce the risk of overfitting. The known data is used to train the model, and then the trained Bayesian linear regression model is used to predict new $y$. It is assumed that object variable $t$ is provided by function $y(x, \omega)$. Because actual data collected are unavoidably accompanied by Gaussian noise, so the equation can be written as:

$$t = y(x, \omega) + \epsilon \tag{4}$$

where $\epsilon$ is a Gaussian random variable. Its mean is 0, and accuracy (Reciprocal of variance) is $\beta$. So the equation is:

$$P(t|x, \omega, \beta) = N\left(t|y(x, \omega), \beta^{-1}\right) \tag{5}$$

Now if a data set $X = \{x_1, \ldots, x_n\}$ is entered and corresponding target value is $t_1, \ldots, t_n$. It is assumed that these data points are selected independently, and then the expression of likelihood function ($\omega, \beta$ is adjustable) can be written as:

$$P(t|X, \omega, \beta) = \prod_{n=1}^{N} N\left(t_n|\omega^T \phi(x_n), \beta^{-1}\right) \tag{6}$$

Then the log of the likelihood above is taken, that is:

$$\begin{aligned} \ln P(t|\omega, \beta) &= \sum_{n=1}^{N} \ln N\left(t_n|\omega^T \phi(x_n), \beta^{-1}\right) \\ &= \frac{N}{2}\ln\beta - \frac{N}{2}\ln(2\pi) - \beta E_D(\omega) \end{aligned} \tag{7}$$

where sum of square error function is defined as:

$$E_D(\omega) = \frac{1}{2}\sum_{n=1}^{N} \left\{t_n - \omega^T \phi(x_n)\right\}^2 \tag{8}$$

The gradient of (7) is shown as follows:

$$\nabla \ln P(t|\omega, \beta) = \beta \sum_{n=1}^{N} \left\{t_n - \omega^T \phi(x_n)\right\}^2 \phi(x_n) \tag{9}$$

Let the value of (9) equals 0:

$$0 = \sum_{n=1}^{N} t_n \phi(x_n)^T - \omega^T \sum_{n=1}^{N} \left(\phi(x_n)\phi(x_n)^T\right) \tag{10}$$

Then it can be made out that:

$$\omega_{ML} = \left(\Phi^T \Phi\right)^{-1} \Phi^T t \tag{11}$$

where $\Phi$ is a design matrix, whose elements are $\phi_j(x_n)$. Equation (11) shows the results of the weight of independent variables on a calm sea surface.

To avoid overfitting when studying turbulent sea surface, regularized least squares method [9] is used. The error function that needs to be minimized is modified as ($\lambda$ is regularization factor):

$$E_D(\omega) + \lambda E_W(\omega) \tag{12}$$

where

$$E_W(\omega) = \frac{1}{2}\omega\omega^T \tag{13}$$

Then Eq. (10) is solved again, and it can be figured out that:

$$\omega = \left(\lambda I + \Phi^T \Phi\right)^{-1} \Phi^T t \tag{14}$$

which is suitable for a turbulent ocean.

Next, the problem of parameter distribution needs to be solved. The prior probability of $\omega$ is introduced, and it is assumed that $\beta$ is known. Because the likelihood function defined by Eq. (6) $P(\omega|t)$ is the exponential form of a quadratic function, whose corresponding conjugate prior probability distribution obeys Gaussian distribution. The expression is as follows:

$$P(\omega) = N(\omega|m_0, S_0) \tag{15}$$

Posterior probability distribution is as follows:

$$P(\omega|t) = N(\omega|m_N, S_N) \tag{16}$$

where

$$m_N = S_N\left(S_0^{-1} m_0 + \beta \Phi^T t\right) \tag{17}$$

$$S_N^{-1} = S_0^{-1} + \beta \Phi^T \Phi \tag{18}$$

To simplify related calculation, a specific form of Gaussian prior probability is considered. To be specific, that is isotropic Gaussian distribution (mean is 0). The distribution is controlled by accuracy parameter (It is a hyper-parameter) $\alpha$.

$$P(\omega|\alpha) = N\left(\omega|0, \alpha^{-1} I\right) \tag{19}$$

The corresponding posterior probability distribution is provided by Eq. (16), where

$$m_N = S_N \beta \Phi^T t \tag{20}$$

$$S_N^{-1} = \alpha I + \beta \Phi^T \Phi \tag{21}$$

Regular term is corresponding to (12), where $\lambda = \frac{\alpha}{\beta}$.

Gathering all the results that are got, the value of following $t$ can be finally predicted.

$$P(t|t, \alpha, \beta) = \int P(t|\omega, \beta) P(\omega|t, \alpha, \beta) d\omega \tag{22}$$

## 3 Reflection Intensity

For reflection of the calm sea surface, reflectivity $R$ is the main reason that influences the reflection intensity. There is only one independent factor in the equation. It is discussed that $R$ obeys Gaussian distribution.

For calculation about related parameters, $\omega_i$ is calculated according to (11) and (14). $\omega_0$ is calculated according to the formula below:

$$\omega_0 = \bar{t} - \sum_{j=1}^{M-1} \omega_j \overline{\phi_j} \tag{23}$$

It is finally figured out that $\omega_0 = 0.3821$, $\omega_1 = 0.1330$.

To solve the problem of $R$ distribution, MATLAB simulation method is used and the result is shown in Fig. 2.

The dotted line shows the mean average level of reflection intensity, and shaded area means the range of fluctuations around the mean level ($\sigma^2 = 1$). Mean value is around 0.4, and fluctuation range is about 0.263–0.638.

With regard to turbulent sea surface, the calculation is more complicated.

There are mainly five influencing factors (independent variable): electromagnetic gradient of seawater, local permittivity, permeability of the ocean, height of the reflection surface, and angle of the reflection surface. The form of these nonlinear basis functions (They obey circular normal distribution) is:

$$P(\theta|\theta_0, m) = \frac{1}{2\pi I_0(m)} \exp\{m\cos(\theta - \theta_0)\} \tag{24}$$

The parameters of these variables $\theta_0, m$ are:
electromagnetic gradient of seawater: $\theta_1 = \frac{\pi}{2}, m_1 = 3$.
local permittivity: $\theta_2 = \frac{\pi}{3}, m_2 = 2$.
permeability of the ocean: $\theta_3 = \frac{\pi}{4}, m_3 = 1$.
height of the reflection surface: $\theta_4 = \frac{3}{2}\pi, m_4 = 0.1$.
angle of the reflection surface: $\theta_5 = \frac{9}{5}\pi, m_5 = 0.3$.

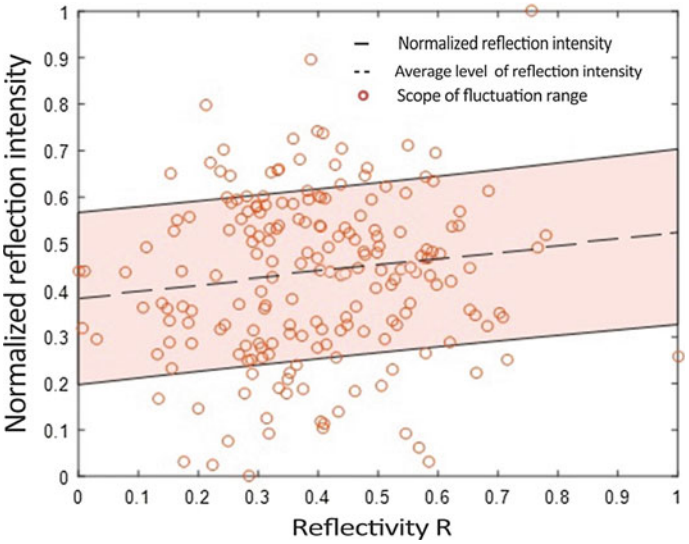



**Fig. 2.** Reflectivity and reflection intensity distribution

These parameters are used to simulate the five circular normal distribution independent variables.

As for $\alpha, \beta$, that introduces EM (Expectation Maximization) algorithm. Complete-data log-likelihood function is:

$$\ln P(t, \omega|\alpha, \beta) = \ln P(t|\omega, \beta) + \ln P(\omega|\alpha) \tag{25}$$

where $P(t|\omega, \beta)$ is (6), $P(\omega|\alpha)$ is (19).

The posterior probability distribution of $\omega$ is taken the expectation as follows:

$$\begin{aligned} E[\ln P(t, \omega|\alpha, \beta)] = \frac{M}{2}\ln\left(\frac{\alpha}{2\pi}\right) - \frac{\alpha}{2}E\left(\omega\omega^T\right) + \frac{N}{2}\ln\left(\frac{\beta}{2\pi}\right) \\ - \frac{\beta}{2}\sum_{n=1}^{N} E\left\{t_n - \omega^T\phi(x_n)\right\}^2 \end{aligned} \tag{26}$$

Let the derivative of (23) with respect to $\alpha$ equals 0, it can be got:

$$\alpha = \frac{M}{E(\omega\omega^T)} = \frac{M}{m_N^T + T_r(S_N)} \tag{27}$$

The initial $\alpha = 0.02, \beta = 0.5$. After iterative computation, (12) is minimized, and it is figured out that $\alpha = 1.14669$, $\beta = 3.3592$. And $\omega_0 = 0.5693, \omega_1 = 0.0872, \omega_2 = -0.8761, \omega_3 = 1.0837, \omega_4 = 0.9798, \omega_5 = 0.7011$.

Putting the calculation results and simulation results above together, the range of reflection intensity of the turbulent sea surface is shown in Fig. 3.

It can be seen that the range of reflection intensity of the turbulent sea surface is 0.35–0.6, which is smaller than the reflection intensity of the calm sea surface.

From the above, the simulation result of Bayesian linear regression model is consistent with that of Longley-Rice radio wave transmission model, which is suitable for maritime radio transmission and is superior to Longley-Rice radio wave trans-

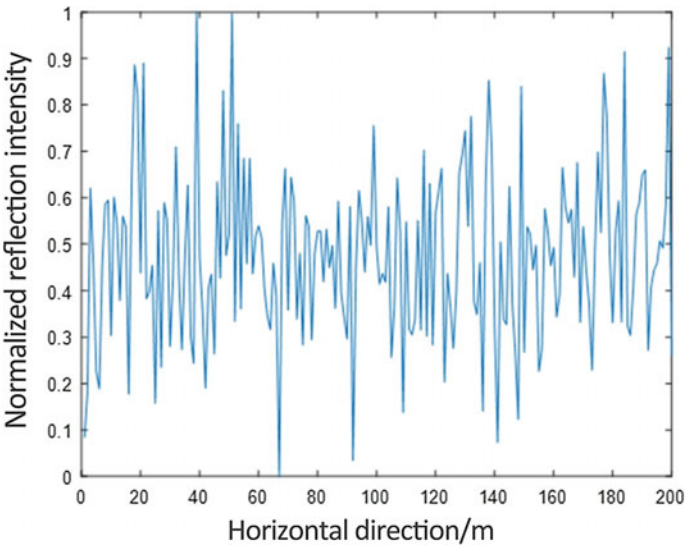



**Fig. 3.** Range of reflection intensity

mission model. The Longley-Rice radio wave transmission model doesn't take into account the influence of the specific marine environment, such as electromagnetic gradient of seawater, local permittivity, permeability of the ocean, height of the reflection surface, and angle of the reflection surface. Moreover, the use of data does not represent all situations and is relatively limited. However, Bayesian linear regression model obtains training data based on a large number of actual data, which is relatively representative and takes more marine environment into consideration. Because the utilization rate of data samples is 100%, the accuracy of the result is relatively high.

## 4 Conclusion

Bayesian linear regression can not only solve the problem of over-fitting existing in maximum likelihood estimation, but also its utilization rate of data samples is high. Only using training samples can effectively and accurately determine the complexity of the model, thus obtaining a more accurate result than ordinary linear regression models. Moreover, the model can simulate the changes of marine environment well by using

various forms of appropriate basis functions to be given a form in the above. The result obtained by this model is verified to be in line with reality and superior to some other models. This model gives a good simulation process for the reception of signals on the sea surface when medium and short wave communication is used by ships traveling, which is of great significance to the development of maritime wireless communication.

**Acknowledgments.** This research was supported by the Fundamental Research Funds for the Central Universities (Grant no. 3132016317).

## References

1. Hu Z, Zeng F, He J, Zhang X. Development of HF adaptive communication and construction of wireless specialty in our hospital. J Chongqing Inst Commun. 2004.
2. Ren J. How to improve the efficiency of shortwave frequency resources in maritime communication. New Obs. 2015.
3. Shao X, Chu X, Wang J. Application of two-dimensional atmospheric model in maritime radio wave propagation. J China Ocean Univ. 2011;09:109–13.
4. Yi Q. Study on electromagnetic wave propagation model in sea area. Hainan University, 2015.
5. Zhao X. Study on propagation characteristics and application of electromagnetic wave in atmospheric waveguide environment. Xidian University, 2008.
6. Liu Y, Zhou X, Jin H. Analysis and research on radio wave propagation prediction model. Nav Electron Eng. 2011;07:84–6.
7. Shao L. Research on AIS sea wave propagation model. Dalian Maritime University, 2014.
8. Wu Y, Wang C, Xu C. Research on radio wave propagation path loss model in wireless communication system. Foreign Electron Meas Technol, 2009; (08).
9. Bishop C. Pattern recognition and machine learning. Berlin: Springer; 2007.

# Differential Evolution FPA-SVM for Target Classification in Foliage Environment Using Device-Free Sensing

Yi Zhong[1], Yan Huang[1(✉)], Eryk Dutkiewicz[1], Qiang Wu[1], and Ting Jiang[2]

[1] School of Electrical and Data Engineering,
The University of Technology, Sydney, Australia
{Yi.Zhong, Yan.Huang-3}@student.uts.edu.au
{Eryk.Dutkiewicz, Qiang.Wu}@uts.edu.au
[2] School of Information and Communication Engineering,
Beijing University of Posts and Telecommunications, Beijing, China
tjiang@bupt.edu.cn

**Abstract.** Target classification in foliage environment is a challenging task in realistic due to the high-clutter background and unsettled weather. To detect a particular target, e.g., human, under such an environment, is an indispensable technique with significant application value. Traditional method such as computer vision techniques is hardly leveraged since the working condition is limited. Therefore, in this paper, we attempt to tackle human detection by using the radio frequency (RF) signal with a device-free sensing. To this end, we propose a differential evolution flower pollination algorithm support vector machine (DEFPA-SVM) approach to detect human among other targets, e.g., iron cupboard and wooden board. This task can be formally described as a target classification problem. In our experiment, the proposed DEFPA-SVM can effectively attain the best performance compared to other classical multi-target classification models and achieve a faster convergent speed than the traditional FPA-SVM.

**Keywords:** Target classification · Human detection · Foliage environment

## 1 Introduction

Human detection is inherently an indispensable technique in computer vision community that has been investigated for many years [12,14,22]. It is a challenging task and underpins crucial application such as target tracking [1,10], recognition [9,13] and re-identification [5–8,17]. However, the traditional vision-based approaches and synthetic-aperture radar-based systems for human detection still


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_67

have limitations under complex cases. Among them, outdoor environment, in particular, foliage environment, is intractable due to challenges of cluttered background, seasonal variations, and unsettled weather. Figure 1 lists the challenges for human detection in foliage environment. In this case, the target (human) becomes indistinguishable since trunks, branches, and leaves of trees produce huge disturbance that virtually overwhelms the target information. Additionally, the cost of using such systems is relatively high. As far as we know, an alternative solution is to use hybrid-sensor-based systems [2,3]. However, the accuracy can be dramatically dropped compared to the other two approaches.

To achieve outdoor human detection with cluttered background and weather variations, in this paper, we concentrate on the study of an emerging technique, known as device-free sensing (DFS) [15,22]. DFS is based on wireless sensor network (WSN) which is cost-effective since no additional hardware is needed but only the RF transceivers [19]. Therefore, it is a natural choice to use DFS for human detection not only for the low cost but also for the data communication capability it has. Considering the potential capability of the RF signal, recently, researches also attempt to use RF in outdoor human detection [18,20,21]. Two scenarios are presented. The first one is human and animal classification since animal like dogs often produce false alarm in foliage environment [18,20]. The second one considers the impact of weather variations [21]. In this paper, we follow the latter one to utilize the RF signal for human detection under different weather conditions.

In the previous attempt [21], the impulse radio ultra-wideband (IR-UWB) signals are utilized for sensing different targets, i.e., human, iron cupboard, wooden board, and no target. To generalize the performance of learned classification model, data from four weather conditions is collected, i.e., sunny, rainy, snowy, and foggy. To classify the above-mentioned different targets under the variational weathers, high-order cumulant (HOC) is extracted from the original signal as features. The support vector machine (SVM), which has been successfully applied to target classification [4,11,16], is utilized in this task. To effectively optimize the parameters of SVM and ensure a superior classification result, the flower pollination algorithm (FPA) is introduced in [21] in conjunction with the traditional SVM. However, although the performance is promoted, FPA may easily get trapped in local minima because of pollen diversity loss. Therefore, we introduce a differential evolution (DE) algorithm to maintain the diversity of population and to explore local search. Also DE does not have a mechanism to memorize the previous process and uses the global information about the search space, so it easily results in a waste of computing power leading to premature convergence. By utilizing DEFPA, we further improve the performance of classification, and on the other hand, a faster convergence speed is achieved.

## 2 Methodology

### 2.1 Data Acquisition

Followed by [21], we use the targets collected from high-clutter foliage environment with variational weather conditions as our experimental data. These targets are categorized into four classes (1) no target, (2) human, (3) iron cupboard, and (4) wooden board collected from four weather conditions: (1) sunny, (2) rainy, (3) snowy, and (4) foggy. Figure 1 shows the foliage environment that utilized for data acquisition. Figure 2 provides the topological structure of data collection. The IR-UWB transceiver is used with the operating frequency from 3.1 to 5.3 GHz and the center frequency at 4.2 GHz. The transceiver is combined with planar elliptical dipole antennas to form a prototype. Also, the overall topological structure is connected to a PC to record data. The position of targets is placed at eight different locations according to the position of receiver and transmitter (see Fig. 2).

### 2.2 The Proposed DEFPA-SVM

The overall pipeline of our target classification can be divided into three stages: data acquisition, feature extraction, and the proposed DEFPA-SVM model optimization. Suppose there are $m$ scenarios and $n$ targets. The $i$th target signal in different scenarios is denoted as $\{S_{i1}, S_{i2}, \ldots, S_{im}\}$. We extract the HOC feature [21] from the signal (data) we collect denoted as $\{\mathrm{Cum}_S^n | n \in 1, \ldots, n\}$, where n represents the number of targets, in our case, $n = 4$.

After getting the feature with four different target classes, SVM is utilized to train the classification model. We use the FPA to optimize parameters of

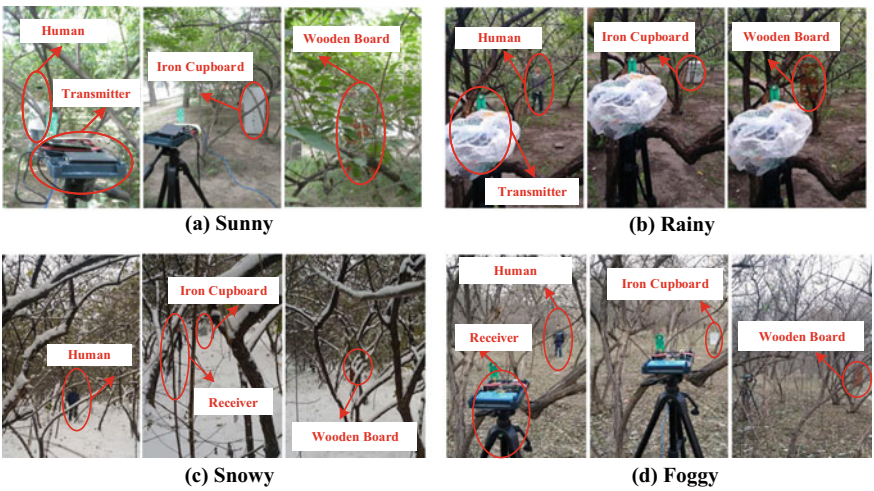



**Fig. 1.** Human detection in foliage environment with weather variations.

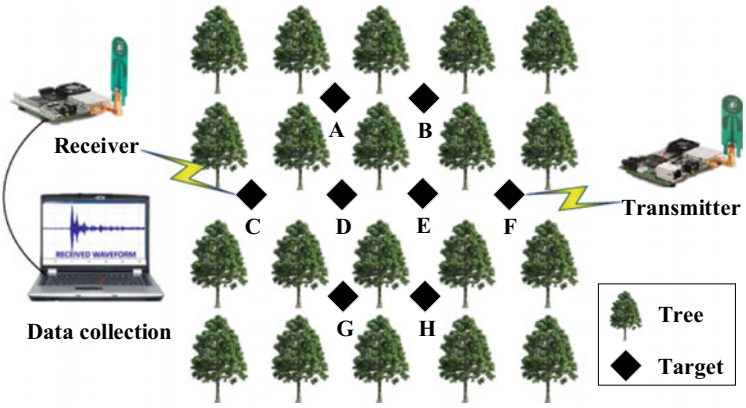



**Fig. 2.** Overall topological structure for data acquisition.

SVM in training. However, due to the inferior diversity of pollen in FPA, the training easily get trapped in local minima. To tackle such problem, we investigate and use the DE algorithm for parameter optimization in our classification task. DE was introduced in [15] that is a powerful stochastic population-based search method for global optimization problems. DE includes three important operations, i.e., mutation, crossover, and selection. Among them, mutation and crossover are used to generate new vector, and selection is used to determine whether the generated vector can survive in the next generation. The detail of the procedure of DE can be found in [15].

The motivation of the proposed DEFPA algorithm is to integrate DE into FPA algorithm, thus increasing the diversity of the population to escape the local minima of FPA-SVM. The general process of proposed DEFPA is presented as follows:

- Step 1: Initialize a population of $N$ pollen gametes with random solutions, i.e., $\{x_i | i = 1, 2, \ldots, N\}$.
- Step 2: Evaluate the fitness function of each solution by calculating its corresponding objection function in the population, and select the best solution as global solution $g^*$ in the initial population.
- Step 3: For all the solutions in populations, we take the following steps to evolve by FPA algorithm:
  (1) If (rand$<p$), flower constancy in FPA is used for the global pollination (for detail, see [21]), else conducting the local pollination, where rand$[0, 1]$ is a random number for each solution $x_i$, and $p$ is the switch probability.
  (2) Evaluate each new solution in the population and update the solutions of the population if their fitness values are superior.
- Step 4: The updated solutions calculated by FPA are regarded as the initial point for DE.
- Step 5: Performing mutation, crossover, and selection of DE operations to obtain the new solutions.

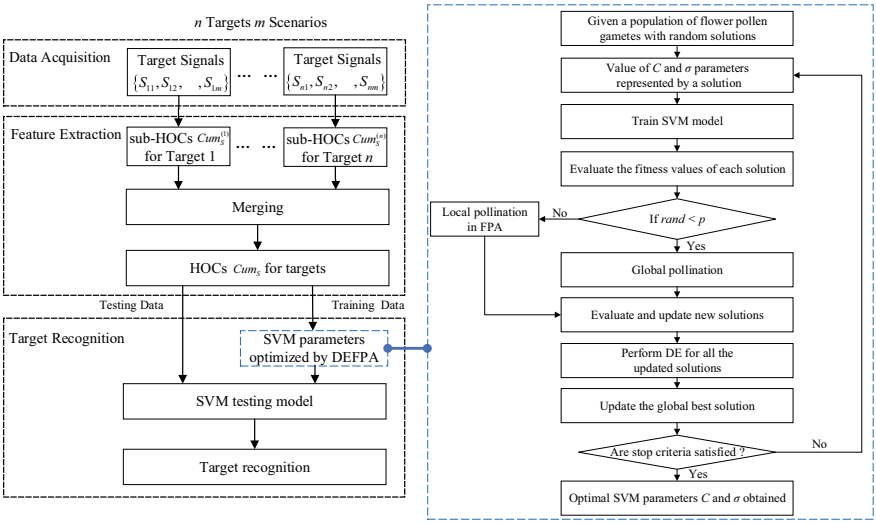



**Fig. 3.** Pipeline of proposed DEFPA-SVM algorithm.

- Step 6: Evaluate each new solution in the population, and update the solutions in the population if their fitness values are superior.
- Step 7: Rank the solutions and pick up the best one. If the current best solution is better than $g^*$, then $g^*$ is replaced by it.
- Step 8: The procedure proceeds until stopping criteria are satisfied; otherwise, loop to Step 3.

In this study, we use RBF kernel to construct SVM classifier, and the selection of the width $\sigma$ and penalization parameters $C$ are significantly influenced by the performance of SVM. The proposed DEFPA is applied to determine the parameters. Figure 3 presents the process of optimizing the SVM parameters with DEFPA; all the steps are introduced above.

## 3 Experiments

We use the data collected in foliage environment to verify the effectiveness of the proposed DEFPA-SVM. Section 2.1 provides the details of our data collection. Specifically, four targets are involved; each target contains 3200 signals (data) under four weather conditions averagely (800 for each weather). HOC features are extracted to represent each signal. We randomly select 640 data (each weather with 160) of each target to consist the testing dataset, totally 2560 data. For training, we randomly select 480 data (average 120 per weather) for each target, totally 1920 data.

Table 1 lists the confusion matrix of our proposed DEFPA-SVM approaches. We can observe that a majority of target can be classified into the correct category. Table 2 presents the values of the classification accuracy. In addition to our

proposed DEFPA-SVM, we also provide the result of the tradition SVM and its improved version FPA-SVM. Also, two classical approaches are utilized as our comparison tasks, i.e., the BPNN and the KNN. Apparently, our DEFPA-SVM achieves the best performance compared to the others, not only the average accuracy of the four targets, but also accuracies of four different targets. Compared to BPNN and KNN, SVM-based methods demonstrate better performance due to its merits in classification tasks.

**Table 1.** Test results of four target types each including 640 samples

| Target type of samples | Test results | | | |
|---|---|---|---|---|
| | No target | Target type I | Target type II | Target type III |
| No target | 610 | 14 | 4 | 12 |
| Target type I | 6 | 600 | 24 | 10 |
| Target type II | 6 | 14 | 604 | 16 |
| Target type III | 4 | 4 | 24 | 608 |

**Table 2.** Summary of classification accuracy using different classifiers

| Target type | Classification rate (%) | | | | |
|---|---|---|---|---|---|
| | DEFPA-SVM | FPA-SVM | SVM | BPNN | KNN |
| No target | 95.28 | 97.34 | 96.72 | 96.25 | 95.16 |
| Target type I | 93.78 | 91.09 | 86.56 | 81.09 | 80.78 |
| Target type II | 94.31 | 92.34 | 88.44 | 87.03 | 86.41 |
| Target type III | 95.03 | 94.78 | 90.31 | 88.91 | 87.34 |
| Average | 94.60 | 93.78 | 90.51 | 88.32 | 87.42 |

Figure 4a shows the results when we change the number of training data to train the three different SVM-based approaches. The training data ranges from 160 to 480 per target, and the gap is 40. We can observe that the DEFPA-SVM demonstrates its robustness when different training data is involved in our experiments. In the meantime, FPA-SVM also shows promising results, but still interior than our proposed DE approach. This is because by using DE in optimizing the parameters of FPA-SVM, the learned parameters can be converged faster; then, a superior result can be achieved.

Figure 4b shows the fitness performance for the proposed DEFPA-SVM and the classical FPA algorithm. Apparently, the proposed DEFPA approach is much faster than FPA in the speed of convergence. This results in a superior performance by using DE algorithm in optimizing FPA-SVM under the same number of training iterations.

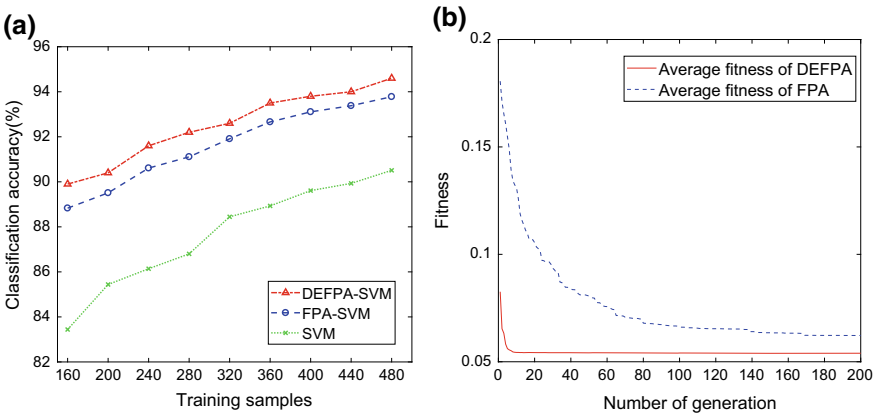



**Fig. 4.** **a** Average classification rates of DEFPA-SVM, FPA-SVM, and SVM under different training samples. **b** Fitness performance for the proposed DEFPA-SVM and classical FPA algorithm.

## 4 Conclusion

In this paper, we propose a DEFPA-SVM to achieve target classification in foliage environment for human detection. The task is challenging due to the high-clutter background and unsettled weather conditions. We use the data that involves four weather conditions with four targets in foliage environment to train the proposed DEFPA-SVM approach. Experiment results demonstrate that, by using the proposed methods, we achieve the best performance compared to other classical approaches we used. In the meantime, we attain a faster convergent speed than the traditional FPA-SVM method.

**Acknowledgments.** This research is supported by NSFC 61671075 and NSFC 61631003.

## References

1. Andriyenko A, Schindler K, Roth S. Discrete-continuous optimization for multi-target tracking. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR). New York: IEEE; 2012. p. 1926–33.
2. Chan T-K, Kuga Y, Ishimaru A. Experimental studies on circular SAR imaging in clutter using angular correlation function technique. IEEE Trans Geosci Remote Sens. 1999;37(5):2192–7.
3. Frolind P-O, Gustavsson A, Lundberg M, Ulander LM. Circular-aperture VHF-band synthetic aperture radar for detection of vehicles in forest concealment. IEEE Trans Geosci Remote Sens. 2012;50(4):1329–39.
4. Geng Y, Chen J, Fu R, Bao G, Pahlavan K. Enlighten wearable physiological monitoring systems: on-body RF characteristics based human motion classification using a support vector machine. IEEE Trans Mob Comput. 2016;15(3):656–71.

5. Huang Y, Sheng H, Liu Y, Zheng Y, Xiong Z. Person re-identification by unsupervised color spatial pyramid matching. In: International conference on knowledge science, engineering and management. Berlin: Springer; 2015. p. 799–810.
6. Huang Y, Sheng H, Xiong Z. Person re-identification based on hierarchical bipartite graph matching. In: 2016 IEEE international conference on image processing (ICIP). New York: IEEE; 2016. p. 4255–9.
7. Huang Y, Sheng H, Zheng Y, Xiong Z. Deepdiff: learning deep difference features on human body parts for person re-identification. Neurocomputing. 2017;241:191–203.
8. Huang Y, Xu J, Wu Q, Zheng Z, Zhang Z, Zhang J. Multi-pseudo regularized label for generated data in person re-identification. IEEE Trans Image Process. 2019;28(3):1391–1403.
9. Joon Oh S, Benenson R, Fritz M, Schiele B. Person recognition in personal photo collections. In: Proceedings of the IEEE international conference on computer vision. 2015. p. 3862–70.
10. Milan A, Roth S, Schindler K. Continuous energy minimization for multitarget tracking. IEEE Trans Pattern Anal Mach Intell. 2014;36(1):58–72.
11. Muñoz-Marí J, Bovolo F, Gómez-Chova L, Bruzzone L, Camp-Valls G. Semisupervised one-class support vector machines for classification of remote sensing data. IEEE Trans Geosci Remote Sens. 2010;48(8):3188–97.
12. Nguyen DT, Li W, Ogunbona PO. Human detection from images and videos: a survey. Pattern Recognit. 2016;51:148–75.
13. Oh SJ, Benenson R, Fritz M, Schiele B. Faceless person recognition: privacy implications in social media. In: European conference on computer vision. Berlin: Springer; 2016. p. 19–35.
14. Satpathy A, Jiang X, Eng H-L. Human detection by quadratic classification on subspace of extended histogram of gradients. IEEE Trans Image Process. 2014;23(1):287–97.
15. Storn R, Price K. Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. J Global optim. 1997;11(4):341–59.
16. Tang Y, Zhang Y-Q, Chawla NV, Krasser S. Svms modeling for highly imbalanced classification. IEEE Trans Syst Man Cybern Part B (Cybern). 2009;39(1):281–8.
17. Zheng Z, Zheng L, Yang Y. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro. arXiv preprint arXiv:1701.07717, 3, 2017.
18. Zhong Y, Dutkiewicz E, Yang Y, Zhu X, Zhou Z, Jiang T. Internet of mission-critical things: human and animal classification: a device-free sensing approach. IEEE Int Things J. 2017.
19. Zhong Y, Yang Y, Zhu X, Dutkiewicz E, Shum KM, Xue Q. An on-chip bandpass filter using a broadside-coupled meander line resonator with a defected-ground structure. IEEE Electron Device Lett. 2017;38(5):626–9.
20. Zhong Y, Yang Y, Zhu X, Dutkiewicz E, Zhou Z, Jiang T. Device-free sensing for personnel detection in a foliage environment. IEEE Geosci Remote Sens Lett. 2017;14(6):921–5.
21. Zhong Y, Yang Y, Zhu X, Huang Y, Dutkiewicz E, Zhou Z, Jiang T. Impact of seasonal variations on foliage penetration experiment: a WSN-based device-free sensing approach. IEEE Trans Geosci Remote Sens. 2018.
22. Zhou F, De la Torre F. Spatio-temporal matching for human detection in video. In: European conference on computer vision. Berlin: Springer; 2014. p. 62–77.

# Optimal Caching Placement in Cache-Enabled Cloud Radio Access Networks

Ruyu Li, Rui Wang(✉), and Erwu Liu

School of Electronics and Information Engineering,
Tongji University, Jiading Qu, China
ruiwang@tongji.edu.cn

**Abstract.** In this paper, we first consider the random caching and carefully analyze the signal-to-interference-plus-noise ratio from the perspective of a typical user, deriving the expression of successful transmission probability. Then, our objective is to maximize the successful transmission probability by optimizing the caching distribution in high SNR using proposed subgradient descent method. Finally, simulations illustrate that the caching distribution optimized by our proposed algorithm achieves a gain in successful transmission probability over existing caching distributions.

**Keywords:** Caching · Cloud radio access network · Optimization

## 1 Introduction

Cloud radio access network (C-RAN) is a very promising network architecture in the future fifth-generation wireless systems improving efficiently network capacity. In previous C-RAN, BSs are connected to the content provider (CP) via digital backhaul link [1]. Because wireless caching can reduce peak traffic and relieve backhaul load effectively, so the problem of limited backhaul capacity can be well solved by deploying caches at BS side [2]. In summary, we consider a cache-enabled C-RAN architecture.

In previous work [3–5], authors consider the stochastic natures of channel fading and geographic location of BSs and mobile users, studying BS cache designs. Specifically, in [3], authors assume that each BS is equipped with the same cache, i.e., all BSs store the most popular files. In [4], authors consider a uniform distribution of random caching. Note that in [3], because all BSs cache the same popular files, they cannot provide file diversity at different BSs and therefore cannot fully utilize storage resources. The random caching design in [4] can provide a variety of files. However, Tamoor-Ul-Hassan et al. [4] does not specify how to efficiently store multiple different files at each BS. The caching


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_68

design in [5] addresses the above problem, at the cost of a huge computation and optimization complexity.

In general, this paper analyzes and optimizes the caching distribution to maximize the whole successful transmission probability of the entire network. Our main contributions and results are summarized below.

Firstly, based on file combinations, we present a random caching distribution. Then, we derive a tractable expression in high SNR. Next, we optimize the given random caching distribution to satisfy the maximum successful transmission probability. In order to reduce complexity and maintain excellent performance, we develop an iterative optimization algorithm under high SNR conditions to achieve a local optimal cache design. Finally, numerical simulations show that caching distribution optimized by our proposed algorithm performs better than existing caching distributions, achieving a significant gain in successful transmission probability.

## 2 System Model

Consider a cache-enabled C-RAN with one CP, some BSs and mobile users, where the BSs and users are both distributed as an independent homogeneous Poisson point process (PPP) with density $\lambda_b$ and density $\lambda_u$ ($\lambda_b \gg \lambda_u$), respectively. For example, in Fig. 1, the CP stores all files and the bandwidth of CP is $W^{'}$. We assume that the CP has one transmit antenna with transmission

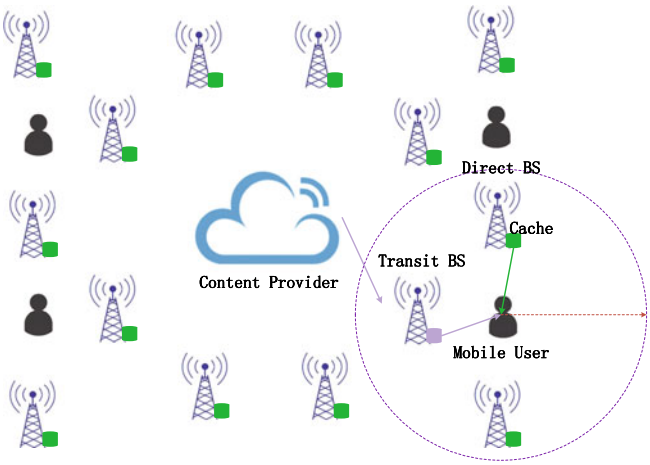



**Fig. 1.** System Model. We consider a cache-enabled C-RAN with one CP, some BSs and mobile users. We assume $u_0$ is located at the origin, within the distance radius $R$. We analyze the performance from the perspective of a typical user $u_0$. When $u_0$ randomly requests file $n$, it will first be served by the direct BS nearest to it, which caches file $n$, denoted as $B_{n,0}$; If there is no BS caches file $n$, then, the user will be served by the transit BS, denoted as $B_{a,0}$. The transit BS transmits the corresponding request to CP to obtain file $n$.

power $P_{cp}$. Due to the limited memory of BS, in this paper, each BS caches one file. Similarly, each BS is equipped with one transmit antenna with transmission power $P$. And the total bandwidth of BS is $W$. Each user has one receive antenna. Here, we assume both large-scale fading and small-scale fading. Transmitted signals are attenuated by a factor $D^{-\alpha}$, and consider the Rayleigh fading channels [6].

Supposed that there are totally $N$ files, and let $\mathcal{N} \triangleq \{1, 2, \ldots, N\}$ represent the group of files. All the files are of the same size and certain popularity for ease of illustration. The file popularity distribution $\boldsymbol{a} \triangleq (a_n)_{n\in\mathcal{N}}$ is identical to all of the users. Here, $a_n \in (0, 1)$ is the probability that each user requests file $n \in \mathcal{N}$ randomly, and $\sum_{n\in\mathcal{N}} a_n = 1$. Furthermore, given that $a_1 \geq a_2 \cdots \geq a_N$.

Each BS randomly caches one file, and we denote $p_n$ as the probability, which also means file n is cached in a BS. On the opposite, $p_{-n}$ means the opposite. $p_n$ satisfying

$$0 \leq p_n \leq 1, n \in \mathcal{N} \tag{1}$$

$$\sum_{n\in\mathcal{N}} p_n = 1 \tag{2}$$

Define the caching distribution $\mathbf{p} \triangleq (p_n)_{n\in\mathcal{N}}$.

As illustrated in Fig. 1, When $u_0$ is served by the direct BS $B_{n,0}$, the signal-to-interference-plus-noise ratio (SINR) of $u_0$ is

$$\mathrm{SINR}_{n,0} = \frac{D_{0,0}^{-\alpha}|h_{0,0}|^2}{\sum_{\ell\in\Phi_b\setminus B_{n,0}} D_{\ell,0}^{-\alpha}|h_{\ell,0}|^2 + \frac{N_0}{P}} \tag{3}$$

where $D_{0,0}$ is the distance between $u_0$ and $B_{n,0}$, $D_{\ell,0}$ is between $u_0$ and other BSs $\ell$, $h_{0,0} \overset{d}{\sim} \mathcal{CN}(0, 1)$ is the small-scale channel between $u_0$ and $B_{n,0}$, $h_{\ell,0} \overset{d}{\sim} \mathcal{CN}(0, 1)$ is between $u_0$ and other BSs $\ell$, and $n_0 \overset{d}{\sim} \mathcal{CN}(0, N_0)$ is the complex additive white Gaussian noise of power $N_0$. $\alpha > 2$ is the path loss exponent.

When $u_0$ is served by transit BS $B_{a,0}$,

$$\mathrm{SINR}_{a,0} = \frac{D_{a,0}^{-\alpha}|h_{a,0}|^2}{\sum_{\ell\in\Phi_b\setminus B_{a,0}} D_{\ell,0}^{-\alpha}|h_{\ell,0}|^2 + \frac{N_0}{P}} \tag{4}$$

where $D_{a,0}$ is the distance between $u_0$ and $B_{a,0}$, $h_{a,0} \overset{d}{\sim} \mathcal{CN}(0, 1)$ is small-scale channel between $u_0$ and $B_{a,0}$.

The signal-to-interference-plus-noise ratio (SINR) of $B_{a,0}$ is

$$\mathrm{SINR}_{cp,a} = \frac{D_{cp,a}^{-\alpha}|h_{cp,a}|^2}{\frac{N_0}{P_{cp}}} \tag{5}$$

where $D_{cp,a}$, $h_{cp,a} \sim \mathcal{CN}(0, 1)$ are the distance and small-scale channel between CP and $B_{a,0}$.

Under multicasting scheme, files are transmitting from $B_{n,0}$ to all users, the corresponding channel capacity of $u_0$ is $C_{n,K,0} \triangleq W\log_2(1+\mathrm{SINR}_{n,0})$. When the transmission of file $n$ at rate $\tau$ can be decoded at $u_0$ properly, $C_{n,K,0} \geq \tau$, we call it successful transmission. Thus, define the successful transmission probability of file $n$, which is transmitted from $B_{n,0}$ to $u_0$, as

$$q_{n,0}(\mathbf{p}) \triangleq \Pr\left[W\log_2(1+\mathrm{SINR}_{n,0}) \geq \tau\right] \tag{6}$$

Similarly, when file $n$ is transited from $B_{a,0}$ to $u_0$ is

$$q_{cp,n}(\mathbf{p}) \triangleq \Pr[W\log_2(1+\mathrm{SINR}_{a,0}) \geq \tau]\Pr\left[\frac{W^{'}}{N}\log_2(1+\mathrm{SINR}_{cp,a}) \geq \tau'\right] \tag{7}$$

Therefore, it can be seen that the caching distribution $\mathbf{p}$ will affect the successful transmission probability.

## 3 Problem Formulation

Based on total probability theorem, the total successful transmission probability of cache-enabled C-RAN can be written as

$$q(\mathbf{p}) \triangleq \sum_{n\in\mathcal{N}} a_n\Big(q_{n,0}(\mathbf{p})p_n^{\pi R^2\lambda_b} + q_{cp,n}(\mathbf{p})p_{-n}{}^{\pi R^2\lambda_b}\Big) \tag{8}$$

According to [7], let $X_v \in \{0,1\}$ imply that whether file $v \in \mathcal{N}$ in $B_{a,0}$ is requested by mobile user. If $X_{v,i} = 0$, then it means that file $v$ in $B_{a,0}$ is not requested by mobile user. The probability $p_{none}$ means there is no mobile user requests a file from a transit BS. Thus,

$$p_{none} = \Pr[X_v = 0] = \Big(1 + 3.5^{-1}\frac{a_v\lambda_u}{T_v\lambda_b}\Big)^{-3.5} \tag{9}$$

After carefully analyzing the interference, Eq. (6) can be rewritten as

$$\begin{aligned} q_{n,0}(\mathbf{p}) = 2\pi p_{none}p_n\lambda_b\int_0^{\infty}&\exp\Big(-\frac{2\pi}{\alpha}\Big(1-p_{none}\Big)p_n\lambda_b s^{\frac{2}{\alpha}}B'\Big(\frac{2}{\alpha},1-\frac{2}{\alpha},\frac{1}{1+sd^{-\alpha}}\Big)\Big)\\ &\times\exp\Big(-\frac{2\pi}{\alpha}\Big(1-p_n+p_{none}p_n\Big)\lambda_b s^{\frac{2}{\alpha}}B\Big(\frac{2}{\alpha},1-\frac{2}{\alpha}\Big)\Big)\\ &\times\exp\Big(-(2^{\frac{k\tau}{W}}-1)d^{\alpha}\frac{N_0}{P}-\pi p_{none}p_n\lambda_b d^2\Big)\mathrm{d}d \end{aligned} \tag{10}$$

where the Beta function is $B(x,y) \triangleq \int_0^1 u^{(x-1)}(1-u)^{(y-1)}\mathrm{d}u$, the complementary incomplete Beta function is $B'(x,y,z) \triangleq \int_z^1 u^{(x-1)}(1-u)^{(y-1)}\mathrm{d}u$, and $s = (2^{\frac{\tau}{W}}-1)d^{\alpha}$.

Similarly, (7) is given by

$$q_{cp,n}(\mathbf{p}) = 4\pi^2 p_{none}\lambda_b \int_0^\infty d^2 \exp\Big( -\Big(2^{\frac{N\tau'}{W}} - 1\Big) d^\alpha \frac{N_0}{P_{CP}}\Big) \times \exp\Big( -\pi d^2 - \frac{2\pi}{\alpha}\Big(1 - p_{none}\Big)\lambda_b s^{\frac{2}{\alpha}} B\Big(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}\Big)\Big) \times \exp\Big( -\frac{2\pi}{\alpha} p_{none}\lambda_b s^{\frac{2}{\alpha}} B'\Big(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}, \frac{1}{1 + sd^{-\alpha}}\Big)\Big) \times \exp\Big( -\Big(2^{\frac{\tau}{W}} - 1\Big) d^\alpha \frac{N_0}{P} - \pi p_{none}\lambda_b d^2\Big) \mathrm{d}d \tag{11}$$

## 4 Optimization

In the following part, we aim to maximize the successful transmission probability by first assigning a specified caching distribution, then carefully optimizing the caching distribution. It is known that $q(\mathbf{p})$ increases with SNR $\frac{P}{N_0}$ and $\frac{P_{cp}}{N_0}$, thus for ease of illustration, in the high SNR region, we get

$$q_\infty(\mathbf{p}) = \sum_{n\in\mathcal{N}} \left( \frac{a_n f {p_n}^{1+\pi R^2\lambda_b}}{f p_n + c_1 p_n - c_1 f p_n + c_2 - c_2 p_n + c_2 f p_n} + \frac{a_n f (1-p_n)^{\pi R^2\lambda_b}}{f + c_1 f + c_2 - c_2 f} \right) \tag{12}$$

where $f = p_{none} = (1 + 3.5^{-1}\frac{a_n\lambda_u}{p_n\lambda_b})^{-3.5}$, $c_1 = \frac{2}{\alpha}(2^{\frac{\tau}{W}} - 1)^{\frac{\alpha}{2}} B'(\frac{2}{\alpha}, 1 - \frac{2}{\alpha}, 2^{-\frac{\tau}{W}})$ and $c_2 = \frac{2}{\alpha}(2^{\frac{\tau}{W}} - 1)^{\frac{\alpha}{2}} B(\frac{2}{\alpha}, 1 - \frac{2}{\alpha})$.

The constraints in (1) and (2) are linear. Thus, we formulate a non-convex function over a convex set:

$$\max_{\mathbf{p}} \; q_\infty(\mathbf{p}) \tag{13a}$$

$$s.t. \;\; (1), (2), \tag{13b}$$

Based on [8], we propose a subgradient descent method. At each iteration, the method removes both the inequality constraints (1) and equality constraints (2). First, objective function can be rewritten as a form of Lagrange function:

$$\mathcal{L} = -q_\infty(\mathbf{p}) + \mu(t)(\sum_{n=1}^{N} p_n(t) - 1) \tag{14}$$

Therefore, we eliminate the equality constraints (2). Then, the inequality constraints (1) can be transformed to $p_n \geq 0$, we eliminate inequality constraints as follows:

$$\mathbf{p}_n(t) = max\{\arg\min\nolimits_{n\in\mathcal{N}}[\mathcal{L}], 0\} \tag{15}$$

Note that, $\mu(t)$ is the dual variable for equality constraints (2); Initially, it is given a fixed value, at each iteration, the value changes according to the following rules:

$$\mu(t+1) = \mu(t) + c(\sum_{n=1}^{N} p_n(t) - 1) \tag{16}$$

$c > 0$ is a step size, which is also a constant.

In summary, we update the values of $\mathbf{p}_n(t)$ and $\mu(t)$ at each iteration until $|\mathbf{p}_n(t+1) - \mathbf{p}_n(t)| < \delta$. $\delta$ is a customizable threshold. Procedures are given in Algorithm 1.

## 5 Simulation Results

In this section, we explore the successful transmission probability under different caching strategies. In our C-RAN scenario, we set density $\lambda_b = 0.1$, $\lambda_u = 0.02$, respectively. For large-scale fading, $\alpha = 4$ is the path loss exponent. The bandwidth is set as $W = 10$ MHz. $N = 100$ is the total number of files, $\tau = 0.1$ Mbps is the transmit rate. For popularity, we assume it follows Zipf distribution, i.e., $a_n = \frac{n^{-\gamma}}{\sum_{n \in \mathcal{N}} n^{-\gamma}}$, where $\gamma = 0.6$. Unless otherwise stated, above simulation parameters remain constant.

**Algorithm 1:** Subgradient Descent Method

**Input**: $\mathbf{p} = (p_n)_{n \in \mathcal{N}}$

**repeat**

  Update $\mathbf{p}_n(t)$:

  $\mathbf{p}_n(t) = max\{\arg\min_{n \in \mathcal{N}}[-q_\infty(\mathbf{p}) + \mu(t)(\sum_{n=1}^{N} p_n(t) - 1)], 0\}$;

  Update dual variable $\mu(t)$:

  $\mu(t+1) = \mu(t) + c(\sum_{n=1}^{N} p_n(t) - 1)$;

  Set $t = t + 1$;

**until** *convergence criterion is met*;

**Output**: $\mathbf{p}^* = (p_n^*)_{n \in \mathcal{N}}$

Figure 2 compares the successful transmission probability versus the total number of file $N$ under different caching distributions. "Average" line refers to uniform caching. "Popularity" line is simulated when each BS selects popular files to cache, the popularity follows Zipf distribution. "Subgradient" line refers to proposed algorithm.

It is observed that, in general, as the total number of files $N$ increases, the successful transmission probability increases as well. We can also find that the performance of our proposed algorithm acts the best. The second is the performance of popularity-based caching distribution. The worst performance comes from the average caching distribution. It is noted that as the total number of files increases, the curve for average cache distribution first rises quickly and then declines slowly. This is because we assume that $\lambda_b \gg \lambda_u$, so when the total number of files increases, it will cause a slight decrease in overall successful transmission probability due to ignoring the popularity of files.

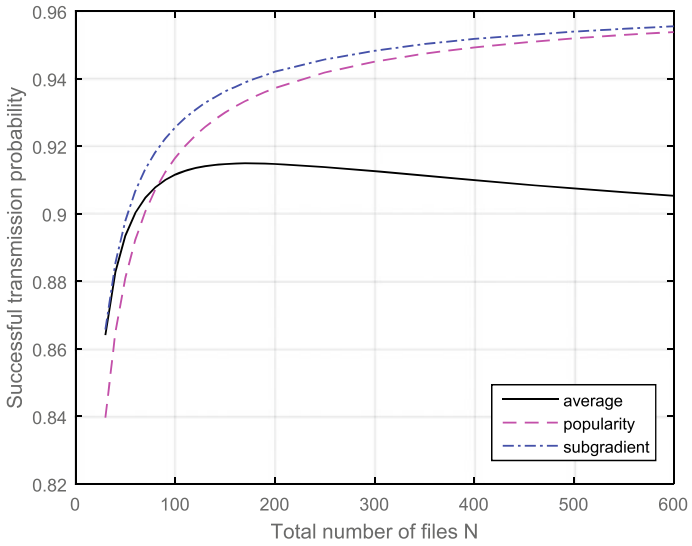



**Fig. 2.** Successful transmission probability versus N for different caching strategies.

## 6 Conclusions

In our work, we first derive the successful transmission probability, considering the random caching and carefully analyzing the interference and signal-to-interference-plus-noise ratio, from the perspective of a typical user. Due to computational complexity, then, we maximize the successful transmission probability in high SNR. Therefore, we propose an optimization method to maximize the successful transmission probability in cache-enabled C-RAN by optimizing the caching distribution. Finally, after simulation, it is observed that the optimized caching distribution performs better than existing caching distributions, which also have lower computational complexity than prior work. This work only serves as the first step toward the global C-RAN in content placement. Future work can provide more promising content placement in general region.

**Acknowledgments.** This work was supported by the National Science Foundation China under Grant 61771345.

## References

1. Zhou H, Tao M, Chen E, Yu W. Content-centric multicast beamforming in cache-enabled cloud radio access networks. In: IEEE global communications conference; 2015. p. 1–6.
2. Tao M, Chen E, Zhou H, Yu W. Content-centric sparse multicast beamforming for cache-enabled cloud RAN. IEEE Trans Wireless Commun. 2016;15(9):6118–31.

3. Baśtuǵ E, Bennis M, Debbah M. Cache-enabled small cell networks: modeling and tradeoffs. In: International symposium on wireless communications systems; 2014. p. 41.
4. Tamoor-Ul-Hassan S, Bennis M, Nardelli PHJ, Latva-Aho M. Modeling and analysis of content caching in wireless small cell networks. 2015;57(1):56–60.
5. Cui Y, Jiang D, Wu Y. Analysis and optimization of caching and multicasting in large-scale cache-enabled wireless networks. IEEE Trans Wireless Commun. 2016;15(7):5101–12.
6. Singh S, Dhillon HS, Andrews JG. Offloading in heterogeneous networks: modeling, analysis, and design insights. IEEE Trans Wireless Commun. 2013;12(5):2484–97.
7. Yu SM, Kim SL. Downlink capacity and base station density in cellular networks. 2012;14(2):119–24.
8. Yu H, Neely MJ. On the convergence time of dual subgradient methods for strongly convex programs. IEEE Trans Autom Control. 2015;99:1.

# Spectral Characteristics of Nitrogen and Phosphorus in Water

Meiping Song[1,2(✉)], En Li[1], Chein-I Chang[1,3], Yulei Wang[1], and Chunyan Yu[1]

[1] Dalian Maritime University, Dalian, China
smping@163.com
[2] State Key Laboratory of Integrated Services Networks, Xian, China
[3] Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD, USA

**Abstract.** The concentration of nitrogen and phosphorus in the waters is an important indicator to affect water quality and determine the degree of water pollution. The development of hyperspectral remote sensing technology makes it possible to monitor the concentrations of nitrogen and phosphorus in different water areas. In this experiment, the spectral curves of different concentrations of nitrogen and phosphorus solutions were collected by using an imaging spectrometer under laboratory conditions. Then compare the spectral curves of different concentrations of sodium phosphate solution by PSR spectrometer, to analyze the sensitive bands of nitrogen and phosphorus. The experimental results show that for phosphorus, its concentration as a whole is positively correlated with the spectral reflectance. In the wavelength range of 450–630 nm, there is a strong positive correlation between the concentration of phosphorus and the spectral reflectance. The correlation coefficient is above 0.8, and the maximum positive correlation is 0.9 at 550, 603, and 740 nm. For nitrogen, its concentration as a whole is negatively correlated with the spectral reflectance. In the wavelength range of 560–850 nm, the correlation coefficient fluctuates within the range from −0.6 to 0.95, and the maximum negative correlation of 0.95 is achieved at 603, 670, and 807 nm.

**Keywords:** Hyperspectral remote sensing · Nitrogen and phosphorus elements · Correlation coefficient

## 1 Instruction

Water is the source of our life. For water bodies, the most serious problem is eutrophication due to excessive levels of nitrogen and phosphorus. Therefore, it is very important to monitor the content of nitrogen and phosphorus in water. The conventional measurement method requires long time of high temperature and high-pressure digestion, and the temperature, time, and reagents of the digestion have a great influence on the measured results. The entire operation is cumbersome, time-consuming, and labor intensive [1]. As a new technology, hyperspectral remote sensing can accurately identify materials through the reflectivity information of each substance, because the spectral characteristics reflect the nature and state of the substances, and


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_69

different substances have different spectral characteristics. Due to the high spectral resolution of hyperspectral data, it is increasingly used for monitoring water quality [2].

At present, there are not so many studies on inverting remote sensing data to content of nitrogen and phosphorus. The commonly used methods are mainly divided into indirect methods and direct methods. The indirect method is mainly based on relationship between nitrogen and phosphorus elements and some photosensitive substances such as chlorophyll and other substances to establish a remote sensing inversion model. The direct methods use least-squares error criterion to obtain the relationship between concentration of nitrogen and phosphorus and spectral data to achieve remote sensing inversion of nitrogen and phosphorus elements [3]. Both the direct method and the indirect method need to find the spectral characteristic bands of nitrogen and phosphorus elements. Gong Shaoqi and others measured the reflectance spectra of nitrogen and phosphorus at different concentrations through laboratory conditions and found that the nitrogen has a reflection peak at 404 and 477 nm, and the phosphorus has a distinct reflection peak at 350 nm [4]. Liu Zheng and others used spectral data collected in the field and found that the first-order differential reflectance of the band at 595 and 873 nm was well correlated with the concentration of total nitrogen and total phosphorus [5]. Huai Hongyan and others found that for nitrogen, the concentration is positively correlated with the spectral emissivity, and the correlation coefficient is taken to the maximum at 682 nm. For phosphorus, the concentration is negatively correlated with the spectral reflectance, and the correlation coefficient is taken to the maximum at 576 nm [6]. Duan Hongtao and others found that phosphorus is negatively correlated with the spectral reflectance between 400 and 443 nm, and nitrogen is negatively correlated with the spectral reflectance between 400 and 490 nm. In the positive correlation, nitrogen has two high correlation coefficient areas, 510–665 and 705–890 nm; however, phosphorus has only a high correlation coefficient range of 705–890 nm [7]. Guo Shaomeng found that there was a maximum negative correlation between total nitrogen concentration and spectral reflectance at 400 and 470 nm, and the highest positive correlation at 622 and 614 nm. Total phosphorus showed a maximum negative correlation at 495 nm, with the highest positive correlation at 717 and 723 nm [8].

In general, due to the different conditions of water body, the characteristic bands of nitrogen and phosphorus elements are quite different, and they can't provide a reliable theoretical basis for hyperspectral remote sensing monitoring. At present, the research on the spectral characteristics of nitrogen and phosphorus is still in the primary exploration stage. Under the ideal conditions of the laboratory, this experiment explored the spectral characteristics of nitrogen and phosphorus using the relationship between the measured hyperspectral data and the concentration of nitrogen and phosphorus in water.

## 2 Collection of Experimental Data

### 2.1 Design of Experiment

The purpose of this experiment is to study the relationship between the concentrations of nitrogen and phosphorus in water and their reflectance characteristics under ideal conditions, so as to provide theoretical basis for the hyperspectral remote sensing monitoring of nitrogen and phosphorus in inland water bodies such as inland rivers and lakes.

In the experiment, nitrogen can be calculated using ammonium nitrate reagent dissolved in water and phosphorus can be calculated as sodium phosphate reagent dissolved in water. The relationship between the concentration of nitrogen and phosphorus and the characteristics of water reflection spectra was compared by measuring the pure water and measuring the spectral curves after adding different amounts of solutions of ammonium nitrate and sodium phosphate. Due to the strong absorption of water in the near-infrared region [9], the wavelength selected in this experiment is in the visible range. The experimental environment is the dark box and light source of the hyperspectral sorter, excluding the interference from the external environment and changes in light source intensity.

### 2.2 Preparation of Instrument Materials

Instrument: (1) Gaia Field-F-V10E (Fig. 1), spectral range from 400 to 1000 nm, spectral resolution is 2.8 nm; (2) Hyperspectral sorter (Fig. 2), providing stable light source and darkroom; (3) Analytical balance; (4) Beaker with 1 L capacity; (5) Plastic dropper; (6) Glass rod; (7) Black plastic barrel with 2 L capacity to simulate the environment in water, reducing interference.

**Fig. 1.** Visible spectrometer


**Fig. 2.** Hyperspectral sorter

### 2.3 Experimental Procedure

(1) Fill 2 L of distilled water into a black bucket and use an imaging spectrometer to measure the spectral curve of pure water; (2) Weigh 0.2 g of sodium phosphate with an analytical balance and add it to a bucket. Stir the glass rod and let it settle until the solution is stable. Spectral imager was used to measure the spectrum of the sodium phosphate solution in the bucket; (3) Repeat the previous operation until the spectral measurement of the 0.1–0.6 g/L sodium phosphate solution is completed; (4) The procedure for measuring the spectral curve of nitrogen is similar to phosphate (Fig. 3).

## 3 Experiment Result and Analysis

### 3.1 Data Processing

The image captured by the spectral imager is a spectral image containing the feature information. Each point on the spectral image has a corresponding spectral reflectance curve. When light hits the solution, the reflection at the bottom of the bucket will have a greater impact on the spectral curve of the solution. The uneven material at the bottom of the bucket will have a large impact on the result. Therefore, in the process of data processing, the spectral reflectances of 25 pixels at the same position are taken from the spectral image and averaged, and the spectral curves of different concentrations of nitrogen and phosphorus solutions are shown in Figs. 4 and 5.

**Fig. 3.** Measurement spectral curve experiment scene

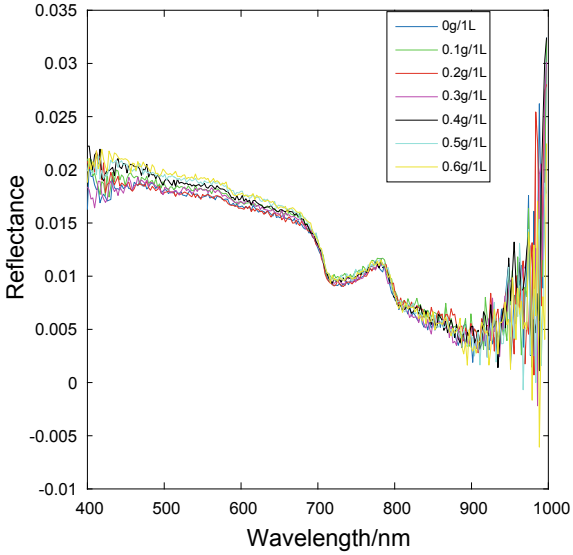



**Fig. 4.** Phosphorus reflectivity curve

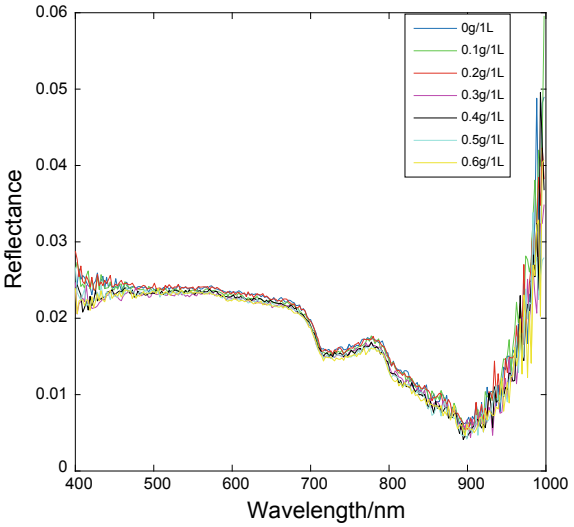



**Fig. 5.** Nitrogen reflectance curve

### 3.2 Spectral Reflectance Data Analysis

From the spectrum diagrams of Figs. 4 and 5, it can be found from the reflectance spectrum curve in the figure that the overall reflectance curves of sodium phosphate and ammonium nitrate solutions are similar to the pure water. The spectral curve has a high reflectance around 400 nm, an absorption peak near 710 nm, a reflection peak near 780 nm, and a decrease in reflectivity, which is consistent with the strong absorption of water in the near infrared.

Comparing different concentrations of sodium phosphate solution, it is found that as the concentration of sodium phosphate solution increases, the reflectance of the solution also increases, which may be the characteristics of phosphorus in water. Comparing the spectral reflection curves of different concentrations of ammonium nitrate solutions, it shows that as the concentration of ammonium nitrate solution increases, the reflectivity of the solution decreases, which may be the characteristics of nitrogen in water.

Correlation analysis measures the relative closeness of two variable elements. From the correlation curves of nitrogen and phosphorus elements and wavelengths (Figs. 6 and 7), it can be found that for total phosphorus, its concentration has a positive correlation with the spectral reflectance. The correlation coefficient in the 450–630 nm wavelength range is in the range of 0.8–0.9. The correlation coefficient of the wavelength range from 630 to 780 nm gradually becomes smaller, and the wavelength range of 780–1000 nm is significantly affected by noise. Among them, the maximum positive correlation 0.9 was reached at 550, 603, and 740 nm. For total nitrogen, its concentration is negatively correlated with the spectral reflectance as a whole. The correlation coefficient in the 560–850 nm wavelength range is in the range from −0.6 to −0.95. Among them, the maximum negative correlation 0.95 was reached at 603, 670, and 803 nm. The correlation coefficient in the wavelength range of 850–1000 nm is significantly affected by noise.

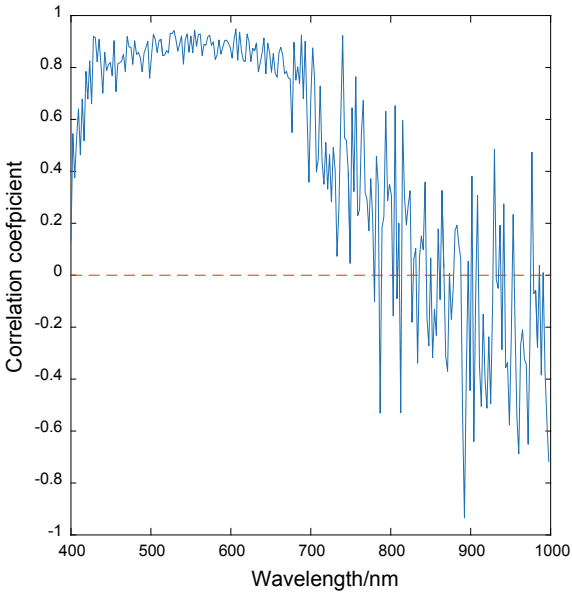



**Fig. 6.** Phosphorus correlation curve

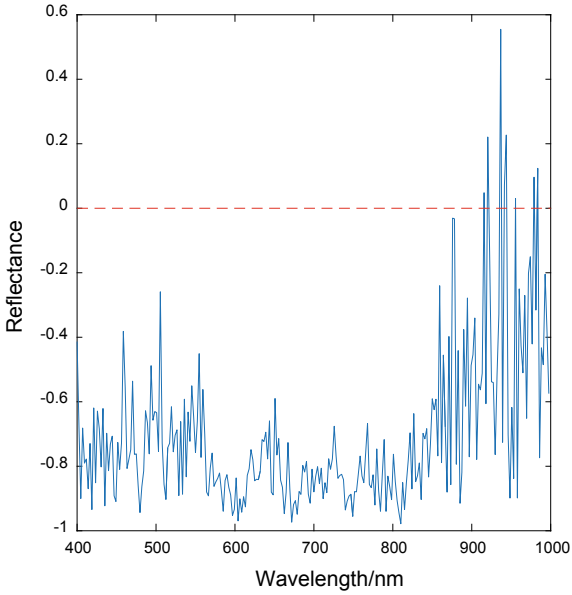



**Fig. 7.** Nitrogen correlation curve

### 3.3 Comparative Test Results Analysis

The PSR spectrometer was used for comparison experiments. The wavelength range is from 340 to 2500 nm, and the spectral resolution is 2.1 nm. Under the same experimental conditions, spectral curves of different concentrations of sodium phosphate solution were collected. According to the analysis of the spectrum diagram of Fig. 8, the results of the PSR spectrometer are similar to the results of the imaging spectrometer. The reflectance curve of sodium phosphate solution is similar to the pure water, and the total phosphorus concentration is positively correlated with the reflectance of the spectrum. The spectral reflectance distributions of different concentrations of sodium phosphate solutions between 400 to 800 nm and 1000 to 1700 nm are relatively obvious. The water reflectivity of the visible range (360–780 nm) is higher than that of the near-infrared range (780–2500 nm). According to the correlation graph of Fig. 9, it can be seen that comparing the PSR spectrometer and the spectral imager in the visible light range, the curve of the correlation coefficient of phosphorus is relatively similar. In the wavelength range of 450–630 nm, there is a strong correlation between the concentration of phosphorus and the spectral reflectance, and the correlation coefficient is above 0.8. At the same time, in the vicinity of the wavelength of 400 nm and the wavelength range of 900–1000 nm, the correlation coefficient varies greatly, which may be related to the noise of the instrument itself and the strong absorption of water in the near infrared.

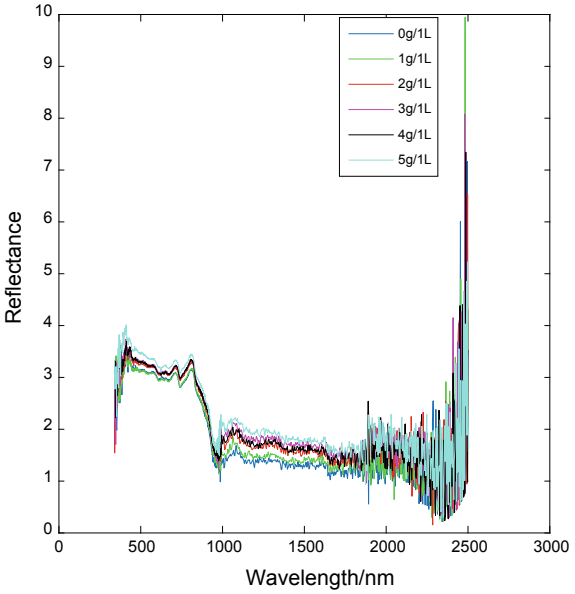



**Fig. 8.** Phosphorus reflectivity curve

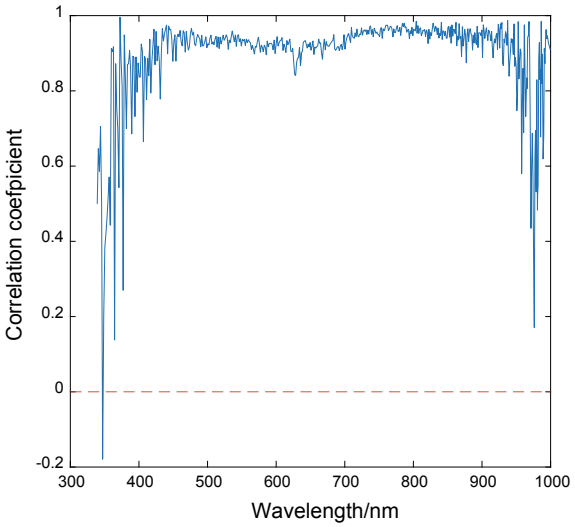



**Fig. 9.** Phosphorus correlation curve

## 4 Conclusion

Through the spectral analysis of nitrogen and phosphorus elements, it can be found that the spectrum curve of nitrogen and phosphorus solution is the same as that of distilled water. The total phosphorus concentration is positively correlated with the reflectance of the spectrum, with a maximum positive correlation of 0.9 at 550, 603, and 740 nm. The total nitrogen concentration is negatively correlated with the reflectance of the spectrum, with a maximum negative correlation of 0.95 at 603, 670, and 807 nm. The lack of sample data in this experiment is insufficient, and the accidental error data may have a relatively large impact on the value of the final correlation coefficient.

**Acknowledgments.** This work is supported by the National Nature Science Foundation of China (61601077), Fundamental Research Funds for the Central Universities (3132016331, 3132018196) and the Open Research Fund of Key Laboratory of Spectral Imaging Technology, Chinese Academy of Sciences (LSIT201707D).

## References

1. State Environment Protection Administration of China. Method of monitoring and analysis for water and wastewater. 4th edn. Beijing: China Environmental Science Press; 2002. p. 243.
2. Clouds EA. Hyper spectral geological remote sensing: evaluation of analytical techniques. Int J Remote Sens. 1996;17(12):2215–42.
3. Silio-C'alzada A, Bricaud A, Gentili B. Estimates of sea surface nitrate concentrations from sea surface temperature and chlorophyll concentration in upwelling areas: A case study for the Benguela system. Remote Sens Environ. 2008;112:3173–80.

4. Shaoqi G, Jiazhu H, Yunmei L, et al. Preliminary exploring of hyperspectral remote sensing experiment for nitrogen and phosphorus in water. Spectrosc Spectr Anal. 2008;28(14): 839–42.
5. Liu Z, He J, Peng L, et al. Correlations between reflectance spectra and contents of TN and TP in Huangbizhuang Reservoir, Hebei Province. J Shijiazhuang Univ. 2009;11(3):45–55.
6. Hongyan H, Junxin L, et al. Estimation of TN and TP concentration in Dianshan Lake based on field hyperspectral measurements. Henan Sci. 2015;33(11).
7. Duan H, Zhang B, et al. Hyperspectral monitoring model of eutrophication in Lake Nanhu Changchun. Lake Sci. 2005;17(3):282–8.
8. Guo S. A study of lake pollution based on hyperspectral data. Shandong Normal University, 2013.
9. Wang H, Hao ZD, et al. Advance in remote sensing of water quality. Mar Environ Sci. 2012;31(2):285–8.

# Big Data-Based Precise Diagnosis in Space Range Communication Systems

Yuan Gao[1,2,4(✉)], Hong Ao[2], Weigui Zhou[2], Su Hu[4], Wanbin Tang[4], Yunzhou Li[3], Yunchuan Sun[5], Ting Wang[2], and Xiangyang Li[2]

[1] Academy of Military Science of the PLA, Beijing 100000, China
yuangao08@tsinghua.edu.cn
[2] Xichang Satellite Launch Center, Sichuan 615000, China
[3] State Key Laboratory on Microwave and Digital Communications, National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, China
[4] University of Electronic Science and Technology of China, Sichuan 611731, China
husu@uestc.edu.cn
[5] Business School, Beijing Normal University, Beijing 100875, China

**Abstract.** With the increase of aerospace launch density, the stability of the firing range measurement and control system and the network communication system in the range is particularly important. The potential failure of the range information system needs more attention, including the aging of the line, the destruction of animals, human damage, and the influence of viruses. Electromagnetic interference, etc., may cause serious problems such as delay in launching missions, errors in receiving and monitoring signals, and inability to issue satellite in-orbit control commands, even causing major accidents involving star destruction. In order to adapt to the load capacity during the high-density task period, to enhance the cognitive ability of the new load launch, and to improve the ability of the range to perform difficult tasks, it is necessary to accurately diagnose and maintain the launch system of the space range.

**Keywords:** Big data · Precise diagnosis · Space range · Communication system

## 1 Introduction

The government of the United States has relied on the Air Force Space Command, DARPA, and the Ministry of National Defense and other military-related agencies to carry out informatization and intelligent changes in the space range in the early 1920s [1], and gradually adopted the policy support of military-civilian fusion space to bring the most advanced technology was introduced into the field of spaceflight launch [2] and a representative US Air Force informatization space range was constructed. The US military's existing Cape Canaveral launch site launched an information-renewing project for the range in 2015. It is jointly funded by DARPA and the US Air Force [3]. SpaceX's participating USD 400 million funds are allocated to meet future


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_70

high-density launches, military and civilian integration launches, battlefield rapid response satellite reissues, and other highly sophisticated tasks. Since the launch of the renewal project, the range has launched a large-thrust launch vehicle several times in complex weather conditions in 2017 and has cooperated with SpaceX and other companies to develop a recyclable rocket project [4]. Its range of intelligence and informatization has led the world in space range. With the support of the US Department of Defense, the US military has also made significant progress in opening up private enterprises to participate in the space launch [5]. For example, cooperation between SpaceX and NASA has updated the US space launch payload platform, monitoring and control platforms, etc., and established the world's first place using big data analysis technology. An information range fault diagnosis and analysis system is quite important to support future business. As an important part of the intelligent platform, the range intelligent failure diagnosis system will be able to help the US military perform the space launch mission again within a 12-h interval [6]. Before that, the official US military data was one week. From the existing US military patents and scientific research institutions published papers, etc., we can learn that the US military's large-data fault diagnosis system for space ranges can rely on autonomous unmanned computer systems to launch the network, spectrum, fuel, and weather needed for space launch. Multiple nodes and units such as flight control, vehicles, personnel, and computer systems are implemented under 24-h control to ensure the precise control of space launch [7–9]. Due to the successful experience of the US military–civilian integration in the GPS system as well as the military and civilian integration space development plan led by the Obama administration and the Trump administration, the US space information systems often use the most advanced information processing technologies, such as Google's big data and artificial intelligence technology, SpaceX's aerospace launch monitoring and control technology, and Amazon's automatic control and robotics technology. Major Japanese space agencies such as Japan's JAXA and the European Space Agency (ESA) have also successively disclosed their plans for building a space astronomy range [10].

With the rapid development of modern communication networks, the maintenance and operation modes of networks have become extremely complicated, and the probability of network failures has greatly increased. Although regular maintenance of network administrators can reduce the possibility of problems and failures in the network, it cannot guarantee that the network will not fail, nor can it guarantee that the failure can be handled and repaired at the earliest time. For space ranges, due to delays or interruptions in space launch caused by various faults will have serious consequences, so fault management has always been a very important network management function and has been highly valued by the base-related business office. With the increasing number of services carried by the network [11], distributed processing applications and remote access have higher and higher requirements for network service reliability, and troubleshooting time is also required to be shortened. Therefore, fault diagnosis and disposal work become particularly important (Fig. 1).

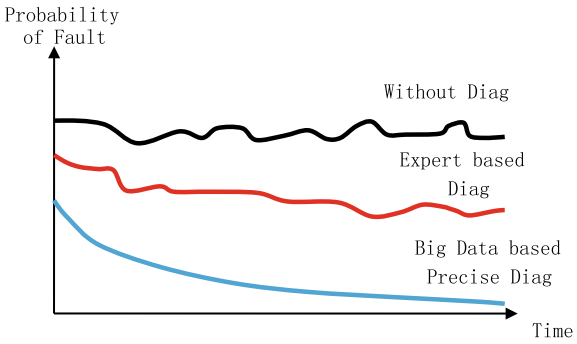



**Fig. 1.** Benefits of big data-based precise diagnosis

## 2 The Hadoop-Based Platform

The stable and efficient measurement and control environment for the space shooting range cannot be separated from the support of daily management and maintenance work. Due to the need to incorporate multiple types of data into the big data processing platform, the big data fault diagnosis platform needs to rely on Hadoop to perform forward-looking and autonomously controllable new designs and integrate multiple resources into the aerospace fault diagnosis system. Unlike ordinary computer network environments or data centers, the space shooting range fault diagnosis platform based on Hadoop has the characteristics of large number of nodes and complex components and applications. Hadoop is designed to run on low-cost computers and regards faults as the norm [12]. Hadoop covers a wide range of functions and uses complex distributed parallel computing architectures, which also brings great challenges to the operation and maintenance of Hadoop clusters. At present, the development of Hadoop technology is still in the stage of improving the capabilities of the technology itself. A large number of scholars have provided many valuable research results in the storage mechanism and computer system of Hadoop. However, in many important aspects, the development of Hadoop still lags behind, such as data security and clusters, maintenance and management, ease of use of the cluster, and more. In order to understand the performance of the traffic analysis system in real time, improve the analysis efficiency, and optimize the system configuration, this project plans to design a system for monitoring and managing distributed clusters suitable for a space shooting range to understand server resource allocation, tracking Hadoop operation status, and alarms. Abnormal conditions, based on this discovery system resource, bottleneck and optimize performance.

The fault diagnosis platform mainly includes the following functions:

- Performance monitoring: The purpose of performance monitoring is to keep abreast of the latest state of the platform, identify resource bottlenecks, and improve the operating efficiency of the platform. Its functions include the collection and analysis of performance data and the prediction of future platform performance based on platform status and business changes. When performance of a computing resource

reaches a bottleneck, performance management should be able to issue a fault alert through the fault management component.

- Fault management: The purpose of fault management is to discover platform faults in time to maintain the normal operation of the platform. Its basic functions include monitoring the status of the platform, discovering faulty nodes, fault processes, and fault services, quickly repairing faults, recording faults, and processing information. Different grades of failures are recommended to be dealt with by different groups of management personnel, even suggesting treatment options.
- Configuration management: The purpose of configuration management is to simplify platform configuration operations, organize and coordinate computing resources, and complete the provision of service capabilities for one or more platforms. Its specific functions include acquiring and storing configuration parameters of computing resources, planning and deploying service components, supporting and tracking changes in platform service capabilities, and providing ancillary means or automated methods to simplify user configuration operations.
- Extended applications: In order to respond to various monitoring needs, an excellent management system should have flexible scalability. Managers can add new monitoring items quickly and easily, which is a test for the underlying architecture of the management platform of the fault diagnosis platform. Among them, performance monitoring and troubleshooting are the most basic functions.

## 3 Deep Learning in Precise Diagnosis

Neural networks (NN), also known as artificial neural networks (ANN), or neural computing (NC), are abstractions and modeling of human brains or biological neural networks, with environments from the environment. The ability to learn adapts to the environment in a biologically similar way. Neural network is an important part of intelligent science and computational intelligence [13]. Based on the research results of brain science and cognitive neuroscience, it expands the methods of intelligent information processing and provides an effective way to solve complex problems and automatic control. Neural network is a nonlinear large-scale adaptive power system composed of a large number of processing units. It is based on the results of modern neuroscience research. It tries to design a new kind of machine with the ability to process and memorize information by simulating the brain neural network so that it has the human brain-like information processing capabilities. At the same time, the study of this neural network will further deepen our understanding of thinking and intelligence (Fig. 2).

A typical deep learning network that is suitable for the fault diagnosis system of a shooting range includes an input layer, an output layer, and several hidden units. The hidden units may or may not be layered. If they are layered, they are called multilayer feedforward networks. The input and output neurons of the network generally take a linear function as their activation function and a non-linear function as the hidden unit. The input unit of the feedforward network accepts signals from the external environment, and after processing, weights the output signals and transmits them to the

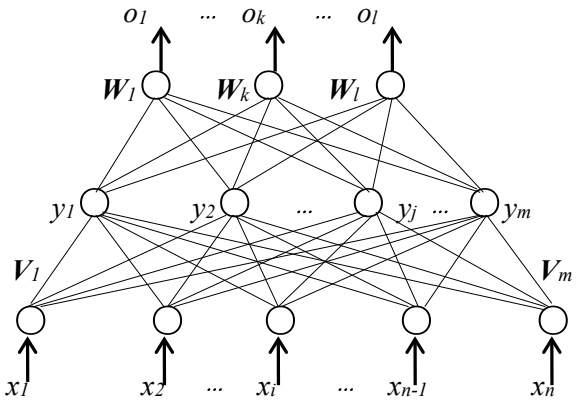



**Fig. 2.** Neural networks in intelligent diagnosis systems

neurons in the projection domain. Implicit elements or output units in the network accept net input from the accepted domains, and then its projection domain sends the output signal, which can be any differentiable function. In the network, the input vector could be expressed as: $X = (x_1, x_2, \ldots, x_i, \ldots, x_n)^{\mathrm{T}}$, similarly, the hidden layer output vector is expressed as: $Y = \left(y_1, y_2, \ldots, y_j, \ldots, y_m\right)^{\mathrm{T}}$, then the output layer vector: $O = (o_1, o_2, \ldots, o_k, \ldots, o_l)^{\mathrm{T}}$, the desired output vector: $D = (d_1, d_2, \ldots, d_k, \ldots, d_l)^{\mathrm{T}}$, the weight matrix between input layer and hidden layer: $V = \left(V_1, V_2, \ldots, V_j, \ldots, V_m\right)$, and the weight matrix between hidden layer and output layer: $W = (W_1, W_2, \ldots, W_k, \ldots, W_l)$.

For a miniaturized network fault diagnosis system, define the diagnostic system to consider six types of problem expressions, two kinds of network faults, and three kinds of disposal solutions

Manifestation: Collect three kinds of information for each form of expression: Yes (1), None (-1), No record (0)

Network failure: Collect three kinds of information for each network failure: Yes (1), None (-1), No record (0)

Disposal plan: Collect two kinds of information for each treatment plan: Yes, No

The information of each fault constitutes a training example. A training example is used to train the network (BP algorithm). Assuming that the neural network shown in the above figure is obtained, the input is the problem representation form, and the fault type is determined in the middle. The final output is the disposal plan.

An autoencoder is a forward-propagation neural network that uses unsupervised learning, uses a back-propagation algorithm, and has a target value equal to the input value. Let the input sample be the vector $x$. The activation of the hidden layer and output layer unit can be obtained by the following function:

$$y = S(Wx + b) \tag{1}$$

$$z = S(W^{\mathrm{T}}x + b') \tag{2}$$

$S(x)$ is a Sigmoid function, weight matrix $W$ and $W^{\mathrm{T}}$ in the two formulae are transposed, and $b$ is the offset. Unlike the supervised learning algorithm of the traditional neural network, the goal of the self-encoder learning is to make the output value equal to the input value, that is, $z$ is equal to $x$ as much as possible. Use KL divergence (Kullback–Leibler divergence) as a loss function,

$$L(X, Z) = \sum_{i=1}^{n} \mathrm{KL}(x_i || z_i) \tag{3}$$

In the above equation, $\mathrm{KL}(x_i||z_i)$ represents the KL divergence between the input vector and the output vector. The automatic encoder uses a stochastic gradient descent algorithm to train, and the weight update formula is as follows:

$$W \leftarrow W - \eta \frac{\partial L(X, Z)}{\partial W} \tag{4}$$

When using an automatic encoder, it is common practice to introduce additional constraints to limit the number of activations of hidden neurons, so that only a few hidden neurons are activated for a sample. This is one of the ways to sparsely encode data. Sparse coding has been shown to have the potential to improve model performance, after introducing the sparse coding constraint for the hidden layer, the loss function expression looks like this:

$$L(X, Z) = \sum_{i=1}^{n} \mathrm{KL}(x_i || z_i) + \beta \sum_{j=1}^{b} \mathrm{KL}(\rho || \rho_i) \tag{5}$$

The self-encoding neural network is formed by stacking a plurality of trained sparse autoencoders. The parameters of the self-coding neural network are obtained by using a layer-by-layer greedy training method. That is, use the original input to train the first layer of the network, get its parameters $W$ (1, 1), $W$ (1, 2), $b$ (1, 1), $b$ (1, 2); then the first layer of the network will. The original input is transformed into a vector of hidden element activation values (assuming the vector is $A$), then $A$ is used as the input of the second layer, and training is continued to obtain the parameters of the second layer $W$ (2, 1), $W$ (2, 2), $b$ (2, 1), $b$ (2, 2); Finally, the same strategy used for the following layers, that is, the output of the previous layer is used as the input method of the next layer in turn.

## 4 Conclusion

In this paper, we discuss the big data-based diagnosis method in space range, we develop the architecture and deep learning method to proceed with the space data. Through deep learning, we can find out the source of the fault and adopt effective operations to deal with the time-tolerance system.

**Acknowledgments.** This work is funded by National Natural Science Foundation of China (61701503), the work of Su Hu was jointly supported by the MOST Program of International S&T Cooperation (Grant No. 2016YFE0123200), National Natural Science Foundation of China (Grant No. 61471100/61101090/61571082), Science and Technology on Electronic Information Control Laboratory (Grant No. 6142105040103) and Fundamental Research Funds for the Central Universities (Grant No. ZYGX2015J012/ZYGX2014Z005). We would like to thank all the reviewers for their kind suggestions to this work.

## References

1. Bowler S. SpaceX success. Astron Geophys. 2014;55(1):1.9. https://doi.org/10.1093/astrogeo/atu026.
2. Kou H, Chen P. An ensemble signature-based approach for performance diagnosis in big data platform. In: 2018 IEEE symposium on service-oriented system engineering (SOSE), Bamberg. 2018, p. 106–15.
3. Chen M, Yang J, Zhou J, Hao Y, Zhang J, Youn CH. 5G-smart diabetes: Toward personalized diabetes diagnosis with healthcare big data clouds. IEEE Commun Mag. 2018;56(4):16–23.
4. Xu Y, Sun Y, Wan J, Liu X, Song Z. Industrial big data for fault diagnosis: Taxonomy, review, and applications. IEEE Access. 2017;5:17368–80.
5. Benayas F, Carrera A, Iglesias CA. Towards an autonomic Bayesian fault diagnosis service for SDN environments based on a big data infrastructure. In: 2018 fifth international conference on software defined systems (SDS), Barcelona, 2018, p. 7–13.
6. Wang B, Xiao X, Xu Y, Li Y. Research on fault diagnosis of power equipment based on big data. In: 2017 IEEE international conference on energy internet (ICEI), Beijing, 2017, p. 193–7.
7. Gao Y, et al. Review of wireless big data in 5G: From physical layer to application layer. In: 2016 2nd IEEE international conference on computer and communications (ICCC), Chengdu, 2016, p. 23–7.
8. Babar MI, Jehanzeb M, Ghazali M, Jawawi DNA, Sher F, Ghayyur SAK. Big data survey in healthcare and a proposal for intelligent data diagnosis framework. In: 2016 2nd IEEE international conference on computer and communications (ICCC), Chengdu, 2016, p. 7–12.
9. Hu H, Tang B, Gong X, Wei W, Wang H. Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks. IEEE Trans Industr Inf. 2017;13(4):2106–16.
10. Ren R, Jia Z, Wang L, Zhan J, Yi T. BDTUne: Hierarchical correlation-based performance analysis and rule-based diagnosis for big data systems. In: 2016 IEEE international conference on big data (Big Data), Washington, DC, 2016, p. 555–62.
11. Pasupathi C, Kalavakonda V. Evidence based health care system using big data for disease diagnosis. In: 2016 2nd international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB), Chennai, 2016, p. 743–7.

12. Zhang J-H, Zou Q. Group learning analysis and individual learning diagnosis from the perspective of big data. In: 2016 IEEE international conference on cloud computing and big data analysis (ICCCBDA), Chengdu, 2016, p. 15–1.
13. Gao Y, Ao H, Wang K, Zhou W, Li Y. The diagnosis of wired network malfunctions based on big data and traffic prediction: An overview. In: 2015 4th international conference on computer science and network technology (ICCSNT), Harbin, 2015, p. 1204–8.

# Joint Kurtosis–Skewness-Based Background Smoothing for Local Hyperspectral Anomaly Detection

Yulei Wang[1,2,3], Yiming Zhao[1(✉)], Yun Xia[4], Chein-I Chang[1,5], Meiping Song[1,2], and Chunyan Yu[1]

[1] Dalian Maritime University, Dalian 116026, China
zym@dlmu.edu.cn
[2] State Key Laboratory of Integrated Services Networks, Xian, China
[3] Key Laboratory of Spectral Imaging Technology, Chinese Academy of Sciences, Xian, China
[4] Research and Development Center for China Academy of Launch Vehicle Technology, Beijing, China
[5] Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD, USA

**Abstract.** Anomaly detection becomes increasingly important in hyperspectral data exploitation due to the use of high spectral resolution to uncover many unknown substances which cannot be visualized or known a priori. The RX detector is one of the most commonly used anomaly detections algorithms, where both the global and local versions are studied. In the double window model of local RX detection, it is inevitable that there will be abnormal pixels in the outer window where the background information is estimated. These abnormal pixels will cause great interference to the detection result. Aiming at a better estimation of the local background, a joint kurtosis–skewness algorithm is proposed to smooth the background and get better detection results. The skewness and kurtosis are three and four order statistics respectively, which can express the non-Gaussian character of hyperspectral image and highlight the abnormal information of the target. The experimental results show that the proposed detection algorithm is more effective for both synthetic and real hyperspectral images.

**Keywords:** Hyperspectral image · Local anomaly detection · Background estimation · Skewness · Kurtosis

This work is supported by the Fundamental Research Funds for the Central Universities (3132016331, 3132018196), National Nature Science Foundation of China (61601077), and the Open Research Fund of Key Laboratory of Spectral Imaging Technology, Chinese Academy of Sciences (LSIT201707D).


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_71

## 1 Introduction

Anomaly detection (AD) has received considerable interest in hyperspectral image processing. Due to the high resolution of the spectral bands, the hyperspectral images can effectively detect small objects which cannot be found in multispectral images. The global RX detector proposed by Reed and Yu is considered as the baseline of AD algorithms. By calculating the martensitic distance between the current pixel and the mean of the background data in the local sliding window, the local anomaly detection (LAD) is proposed using a sliding dual window. LAD is more effective when the anomaly targets are extremely small or only abnormal in a local background but buried in the global background. However, it is often happened that there exist abnormal pixels in the outer window, which will lead to an unfaithful estimation of the background and ultimately lead to a worse detection of the anomaly targets. To solve this problem, this paper proposes an anomaly detection algorithm based on kurtosis–skewness combined background smoothing to improve the local RX algorithm.

## 2 Traditional RX Algorithm

Traditional RX detector can be expressed as follows:

$$\delta_{\mathrm{RX}}(\mathbf{y}) = (\mathbf{y}-\boldsymbol{\mu})^{\mathrm{T}} \sum_{\mathbf{b}}^{-\mathbf{1}} (\mathbf{y}-\boldsymbol{\mu}) \begin{cases} \geq \eta \; H_1 \\ < \eta \; H_0 \end{cases} \tag{1}$$

where $\mathbf{y}$ is the spectral vector to be detected, $\boldsymbol{\mu}$ is the background mean, $\boldsymbol{\Sigma}_{\mathbf{b}}^{-1}$ is the background covariance matrix, and $\eta$ is the decision threshold.

In practical applications, there exists a great variety of ground objects in one hyperspectral imagery. And the targets could be abnormal in a local environment but buried in the global data. In order to solve this problem, a local RX algorithm based on local sliding dual window is proposed. As shown in Fig. 1, it is divided into outer window and inner window. The pixel in the center represents the sample to be detected. The outer window limits the range of the estimated background from the neighborhood, which ensures sufficient information to estimate the local background. Local RX (LRX) detector is as follows:

$$\delta_{\mathrm{LRX}}(\mathbf{y}) = (\mathbf{y}-\boldsymbol{\mu})^{\mathrm{T}} \sum_{\mathbf{local}}^{-\mathbf{1}} (\mathbf{y}-\boldsymbol{\mu}) \begin{cases} \geq \eta \; H_1 \\ < \eta \; H_0 \end{cases} \tag{2}$$

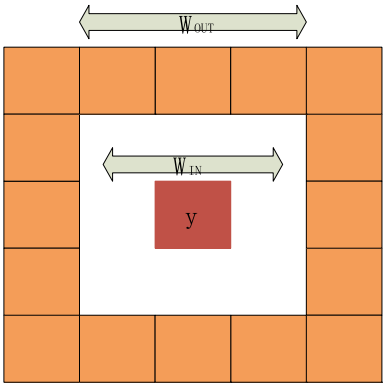



**Fig. 1.** Dual-windows RX detector

## 3 Local RX Detection Based on Background Smoothing

### 3.1 Problems of the Traditional Local RX Detection

As shown in Fig. 2, there exist some pixels in the outer window, whose characteristics are similar to the target to be detected in the center, see pixels in dark red color. This is of common occurrence that the background area is mixed with abnormal pixels. This will cause great interference to the background estimation and have a great effect on the detection performance.

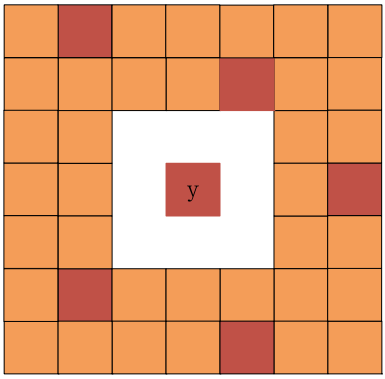



**Fig. 2.** Background anomaly schematic diagram

### 3.2 Skewness and Kurtosis

The statistics parameters skewness and kurtosis are usually used as the statistics description of the data distribution.

Skewness is a measure of skewed direction and degree of statistical data distribution and is a numerical characteristic of asymmetry of statistical data distribution. The mathematical definition of skewness is as follows:

$$\delta^{\mathrm{S}} = \frac{E(\mathbf{X}-\boldsymbol{\mu})^3}{\sigma^3} \tag{3}$$

Kurtosis is a state that measures the probability distribution of real number random variables. The kurtosis $K$ of a random variable is defined by the ratio of the four central matrices of a random variable to the square of variance. The mathematical definition is as follows:

$$\delta^{\mathrm{K}} = \frac{\mathbf{E}(\mathbf{X}-\boldsymbol{\mu})^4}{\sigma^4} - 3 \tag{4}$$

where $\mathbf{E(x)}$ is the expectation of random variable $\mathbf{X}$, $\boldsymbol{\mu}$ is the mean of $\mathbf{X}$, $\sigma$ is the standard deviation of random variable $\mathbf{X}$.

The larger the value of skewness, the more asymmetric the distribution of random variables is.

### 3.3 Joint Skewness–Kurtosis-Based Background Smoothing of Local RX Detection

For standard normal distribution, the larger skewness and kurtosis are, the more deviant the distribution of random variables is, the more abundant the information is. In this paper, the product of skewness and kurtosis coefficient, which is expressed in formula (5), is used to decide whether the pixels in the background window should be included for background estimation (named as Background Smoothing Local RX detector, BS-LRX).

$$\delta^{\mathrm{BS}} = \delta^{\mathrm{S}} \bullet \delta^{\mathrm{K}} \tag{5}$$

In the outer window where the background is estimated, the larger the value of BS-LRX, the more the pixel should be excluded in background estimation. After restraining the $\mathbf{T}$ (the percentage of the excluded pixels in the background) in the background of the outer window, we use the remaining pixels to construct the background statistics and use the LRX detection algorithm to detect the anomalies.

$$\delta_{\mathrm{BS-LRX}}(\mathbf{y}) = (\mathbf{y}-\boldsymbol{\mu}_{\mathbf{0}})^{\mathrm{T}} \sum_{\mathbf{local}}^{\prime -\mathbf{1}} (\mathbf{y}-\boldsymbol{\mu}_{\mathbf{0}}) \begin{cases} \geq \eta\ H_1 \\ < \eta\ H_0 \end{cases} \tag{6}$$

where $\boldsymbol{\mu}_{\mathbf{0}}$ is the background mean of the outer window pixels after excluding abnormal pixels by formula (5) and $\boldsymbol{\Sigma}'^{-1}_{\mathrm{local}}$ is the corresponding covariance matrix.

# 4 Experimental Results and Analysis

## 4.1 Experimental Data Description

In this paper, both synthetic and real hyperspectral images are used to verify the performance of the proposed algorithm. The simulation environment is Matlab 2015b software on WIN7 system with 3.4 GHz processor and 8G memory. The total size of the synthetic hyperspectral image is 200 * 200 pixels with 189 bands. As for the real hyperspectral image, the spatial size is 64 * 64 pixels with a total of 169 spectral bands. The receiver operating characteristic curve (ROC) is used as an evaluation criterion of the detection performance, where the area under curve (AUC) of ROC is also calculated as the quantitative analysis to evaluate the average performance of different algorithms.

## 4.2 Effect of Sliding Window Size on Detection Performance

In this section, the synthetic image is used to discuss the effect of sliding window size on detection performance. Uniform band selection is firstly used on the synthetic data and 10 bands are selected. In order to see how much the outer and inner window sizes influence the detection results, the outer window is selected as 11 * 11, 15 * 15, and the inner window is selected as 3 * 3, 5 * 5, and 7 * 7, respectively. Table 1 represents the AUCs of the traditional local RX algorithm for simulating hyperspectral data under different windows. Combined with grayscale detection maps in Fig. 3, it could be proved that the window size has a large impact. The selection of the size of the double window model has great influence on the detection results; the detection performance is accompanied by a large fluctuation.

**Table 1.** AUC in different window with simulated hyperspectral data

| Window size | 5, 11 | 3, 15 | 5, 15 | 7, 11 | 7, 15 |
|---|---|---|---|---|---|
| AUC | 0.5564 | 0.5468 | 0.5751 | 0.8341 | 0.8309 |

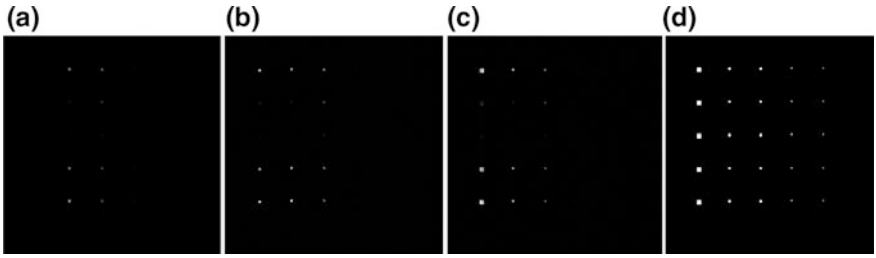



**Fig. 3.** Background local RX with different windows. **a** 3, 15. **b** 5, 15. **c** 7, 15. **d** Ground truth

### 4.3 Performance Evaluation of the Proposed BS-LRX

In order to verify the validity of the proposed BS-LRX algorithm, real hyperspectral image is used for experiment, where uniform band selection is also used (17 bands selected in this experiment). Keep the size of the external window unchanged. The size of the selected window is (3, 11), (5, 11), and (7, 11), respectively. According to the AUC values in Table 2, from horizontal comparison of BS-LRX and LRX with the same window sizes, it could be shown that the proposed BS-LRX algorithm gains better detection performance than the traditional LRX algorithm. On the other hand, if we compare vertically rather than horizontally in the first column of BS-LRX detection results, it is clearly shown that different window sizes will result in different detection performance. According to Table 2, the better detection results can be obtained when the selected inner window is 3. Figure 4 also plots the ROC curves of the BS-LRX and traditional LRX algorithms under different window sizes, for the purpose of comparing detection powers under different false alarms.

**Table 2.** AUC in different window with real hyperspectral data

| Window size | AUC | |
|---|---|---|
| | BS-LRX | LRX |
| 3, 11 | 0.8707 | 0.8285 |
| 5, 11 | 0.8536 | 0.8251 |
| 7, 11 | 0.8684 | 0.8414 |
| 3, 13 | 0.8229 | 0.7378 |
| 3, 15 | 0.8095 | 0.6449 |
| 3, 9 | 0.8641 | 0.8110 |

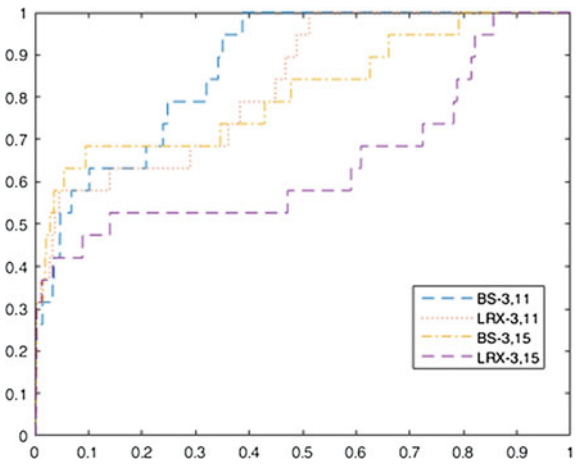



**Fig. 4.** ROC of BS-LRX detection operator in different windows

## 5 Conclusion

Based on the traditional local RX detection algorithm, a new joint kurtosis–skewness background smoothing method is proposed for the purpose of excluding abnormal pixels in the background area. This will result in a more accurate background estimation in LRX algorithm. Through theoretical analysis and experimental discussion, the effectiveness of the proposed BS-LRX algorithm is verified. The detection results are better than the traditional LRX algorithm. Even if the size of the inside and outside window is not exactly selected, it can still get a stable detection result.

## References

1. Liu J. Hyperspectral imagery classification and anomaly detection based on sparse representation. Beijing: Beijing University of Chemical Technology; 2016. p. 43–54.
2. Wang Q, Yang G. Band selection based on local joint skewness and Kurtosis for hyperspectral image. Laser Optoelectron Prog. 2017;54(11):115–22.
3. Wang Y, Chen S-Y, Liu C, Chang C-I. Background suppression issues in anomaly detection for hyperspectral imagery. In: Conference on satellite data compression, communications, and processing X, Baltimore, MD, USA, 2014.
4. Zhu L, Wen G. Hyperspectral anomaly detection via background estimation and adaptive weighted sparse representation. Remote Sens. 2018;10(2).
5. Vafadar M, Ghassemian H. Anomaly detection of hyperspectral imagery using modified collaborative representation. IEEE Geosci Remote Sens Lett. 2018;1–5.
6. Zhao C, Wang Y, Qi B. Global and local real-time anomaly detectors for hyperspectral remote sensing imagery. Remote Sens. 2015;7(4):3966–85.

# Research on Sub-band Segmentation and Reconstruction Technology Based on WOLA-Structured Filter Banks

Yandu Liu(✉), Yiwen Jiao, and Hong Ma

Space Engineering University, Beijing 101416, China
liuyandu1029@126.com

**Abstract.** Sub-band segmentation and reconstruction technology is the core technology of the antenna array signal full spectrum combine scheme. Based on the principle of complex exponential modulation filter banks, the sub-band segmentation and reconstruction technique based on the Weighted OverLap-Add-(WOLA)-structured filter bank is studied. A filter bank with a sub-band number of 256 and an oversampling factor of 1.45 are designed. Compared with the multiphase DFT-structured filter banks currently used, the filter can not only realize the segmentation and reconstruction of performance and considerable, but also break through the oversampling factor must be integer constraints, the structure is more efficient and more flexible parameter settings.

**Keywords:** Sub-band segmentation and reconstruction · Filter bank · WOLA · Wideband signal

## 1 Introduction

At present, the antenna array technology mainly adopts full spectrum combine (full spectrum combining (FSC)) [1–4]. The FSC scheme includes two implementation methods: time-domain combine and frequency-domain combine. The time-domain combine method has the disadvantages of large hardware cost, low delay compensation accuracy and time-delay estimation ambiguity [5, 6]. While the frequency-domain combine method can effectively solve the above problems and fits well with the development trend of large-scale broadband antenna arrays [7, 8]. Compared with narrowband signals, the FSC scheme of broadband signals must consider the phase non-coherence caused by the difference of phase frequency characteristics. Therefore, the wideband signal is channelized, and the low sampling rate sub-band signal is usually needed. Phase difference between different sub-band signals is obtained by phase difference estimation algorithm. Then, estimate and compensate the residual delay and residual phase between wideband signals and the signal of each antenna sub-band is synthesized. Finally, through the reconstruction of the sub-band synthesized by each antenna, the wideband synthetic signal with original sampling rate is restructured.

In the research of frequency-domain combine system, the complex exponential modulation filter bank is the most original, which concept is directly derived from the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_72

sub-band segmentation and reconstruction of frequency domain, but its implementation structure is complex and used only in theoretical research [9]. The sub-band segmentation and reconstruction technology based on WOLA-structured filter banks was first used by JPL. Meanwhile, few research institutions have studied the multiphase DFT design and realization method of the filter bank structure, but the study of WOLA-structured filter band has not been reported in the literature. The multiphase DFT-structured filter bank and the WOLA-structured filter bank are all efficient implementations of the complex exponential modulation filter banks, but the former requires that the oversampling factor must be an integer. The WOLA-structured can not only achieve the equivalent sub-band segmentation and reconstruction performance, but also break through the limitation that the oversampling factor must be an integer. The structure is more efficient and the parameter setting is more flexible. The sub-band analysis and reconstruction technologies based on WOLA-structured filter banks have been studied and a filter bank with 256 sub-banks and a decimation rate of 1.45 has been designed.

First, the basic principle of the complex exponential modulation filter bank is introduced. Then, the design method of the WOLA structure filter bank is studied, including the design of the analysis filter bank and the synthetic filter bank. Finally, the design of the WOLA-structured filter banks is simulated and analyzed.

## 2 Principle of Complex Exponential Modulation Filter Banks

The complex exponential modulation filter bank consists of two parts: the analysis filter bank and the synthetic filter bank. The former is used to divide the signal into multiple sub-banks. The latter is used to integrate the sub-banks to reconstruct the input signal. The overall block diagram is shown in Fig. 1.

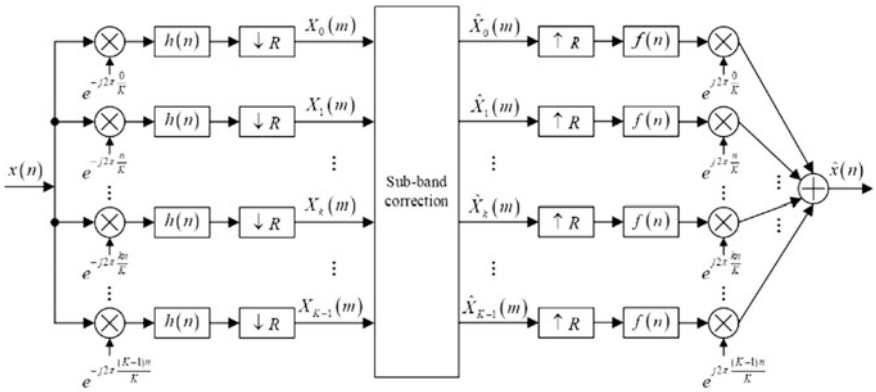



**Fig. 1.** Overall structure block diagram of the complex exponential modulation filter bank

In Fig. 1, $x(n)$ indicates the input time-domain signal, usually the real signal. The analysis filter bank is composed of $K$ down-converter branch. Each down-converter branch is cascaded by complex exponential $\exp(-j2\pi kn/K)$, prototype low-pass filter $h(n)$, and $R$ times decimation. The synthetic filter bank is composed of $K$ up-converter. Each up-converter branch is composed of $R$ times interpolator, prototype low-pass filter $f(n)$ and complex exponential $\exp(j2\pi kn/K)$.

### 2.1 Analysis Filter Bank

The purpose of the analysis filter bank is to divide $x(n)$ evenly into $K$ frequency-domain sub-band signal $X_k(m)(0 \le k \le K-1)$. The frequency range of the signal $x(n)$ is $0 \sim F_s$, and the normalized frequency is represented as $0 \sim 2\pi$. The processing of filter banks can be divided into three steps.

- Step1 The input signal $x(n)$ and the complex exponential $\exp(-j\omega_k n)$ multiplied. From the frequency domain, this process makes the whole spectrum frequency move $-\omega_k$, so the frequency center of the sub-band $k$ is below the zero frequency.
- Step2 The signal after the down conversion is filtered by low-pass filter, and the filter is commonly called analytical filter whose cutoff frequency is $\omega_\Delta/2 = \pi/K$.
- Step3 The filtered signal is extracted by R times, and the $k$ sub-band signal $X_k(m)$ can be obtained.

According to the practical application, it is necessary to amend each sub-band to get the modified sub-band signal $\hat{X}_k(m)$.

### 2.2 Synthetic Filter Bank

The purpose of the synthetic filter bank is to synthesize the $k$ sub-band $\hat{X}_k(m)$ and get the reconstructed time-domain signal $\hat{x}(n)$. The integrated process can be regarded as the inverse process of the analysis process, which can be divided into four steps.

- Step1 Interpolation of sub-band signals will result in spectral imaging at every $2\pi/R$ frequency.
- Step2 A low-pass filter $f(n)$, which generally called a comprehensive filter, is used to filter out the resulting image. After filtering, the reserved signal bandwidth is $\omega_\Delta/2$, that is, $\pi/K$.
- Step3 The complex exponential $\exp(j\omega_k n)$ is used to modulate the sub-band signal to its original frequency $\omega_k$.
- Step4 By adding all the $k$ sub-band signals, the reconstructed signal $\hat{x}(n)$ can be obtained.

### 2.3 Analysis of Filter Parameters

The sub-band number $K$ determines the cutoff bandwidth $\omega_c$ of sub-band signal. The larger the sub-band number $K$ is, the closer the sub-band division is, the smaller the cutoff frequency of the analytical filter and the synthetic filter. The extraction factor $R$

determines the transmission bandwidth $\omega_s$ of the sub-band signal that $\omega_s = 2\pi/R$. The larger the extraction factor $R$ is, the narrower the bandwidth of the sub-band is, the smaller the stopband frequency of the analytical filter and the synthetic filter. The steepness of the filter transition zone is determined by the oversampling factor OS (OS $= K/R$). In fact, the decimation factor $R$ can be selected in the range of $R = 1$ (without extraction) to $R = K$ (critical sampling interval). It is independent of the number of sub-band number $K$, so the value range of OS is $1 \leq \text{OS} \leq K$. In practical applications, OS should be as small as possible to reduce the sub-band transmission rate. However, when OS = 1 is used, the transition bandwidth of the filter is 0, which is not feasible. Document [30] pointed out that when OS $< 1.2$ filter cannot achieve higher stopband attenuation, will cause larger aliasing distortion, and when OS $> 1.5$, it will not significantly reduce aliasing distortion, so the reasonable range of OS is 1.2–1.5. Analyzing the filter length *La* and the synthetic filter length *Ls* determines the performance of the filter. The larger the *La* and *Ls*, the higher the quality of filtering. However, it will lead to larger filter delay and hardware consumption.

These parameters directly determine the stack, group delay, computation, and bandwidth performance of mixed filter, so in the filter design, selection of these parameters is the key to the design.

## 3 Design of WOLA-Structured Filter Bank

### 3.1 Analysis Filter Bank Implementation Structure

As shown in Fig. 1, the analysis procedure of the $k$ sub-band signal can be described by formula.

$$X_k(m) = \left[x(n)\mathrm{e}^{-j\omega_k n}\right] * h(n)|_{n=mR} = \sum_{i=-\infty}^{\infty} h(mR-i)x(i)\mathrm{e}^{-j\omega_k i} \tag{1}$$

where $\omega_k = 2\pi k/K$, $k = 0, 1, \ldots, K-1$. Variable substitution Formula (1), so that $q = i - mR$, has

$$X_k(m) = \sum_{q=-\infty}^{\infty} h(-q)x(q+mR)\mathrm{e}^{-j\omega_k(q+mR)} = \mathrm{e}^{-j\omega_k mR}\tilde{X}_k(m) \tag{2}$$

According to Formula (2), $X_k(m)$ can be regarded as the complex modulation of $\tilde{X}_k(m)$.

$$\tilde{X}_k(m) = \sum_{q=-\infty}^{\infty} h(-q)x(q+mR)\mathrm{e}^{-j\omega_k q} = \sum_{q=-\infty}^{\infty} x_m(q)\mathrm{e}^{-j\omega_k q} \tag{3}$$

where $x_m(q) = h(-q)x(q+mR)$. It can be seen from Formula (3) that $\tilde{X}_k(m)$ is actually the discrete Fourier transform of $X_k(m)$. When the $m$ changes, analysis filter $h(-q)$ is unchanged, while the signal $x(q+mR)$ changes in the form of a packet, each sampling

point is $R$. For a fixed time $m$, the analysis filter $h(-q)$ is multiplied with the signal sequence $x(q+mR)$ by point-to-point. Because the $h(-q)$ coefficient is finite, the number of points in the input sequence $x_m(q)$ of DFT is $L$. It should be noted that the number of points in the output sequence of DFT at each time is $K$. According to this characteristic, when $K \leq L$, the periodicity of $e^{-j\omega_k q}$ (periodicity is $K$) can be used to fold $x_m(q)$ into $K$ point sequence. Thus, the efficient operation of FFT can be used (the number of sub-bands usually takes an index of 2). Based on this, set $q = p + lK$, where $p = 0, 1, \ldots, K-1$ and $l$ are integers, then

$$\tilde{X}_k(m) = \sum_{p=0}^{K-1} \left[ \sum_{l=-\infty}^{\infty} x_m(p+lK) \right] e^{-j\omega_k p} = \sum_{p=0}^{K-1} \tilde{x}_m(p) e^{-j\omega_k p} \tag{4}$$

where

$$\tilde{x}_m(p) = \sum_{l=-\infty}^{\infty} x_m(p+lK) = \sum_{l=-\infty}^{\infty} x(mR+p+lK) h(-p-lK) \tag{5}$$

The implementation structure of the WOLA analysis filter bank can be obtained from the above derivation process, as shown in Fig. 2.

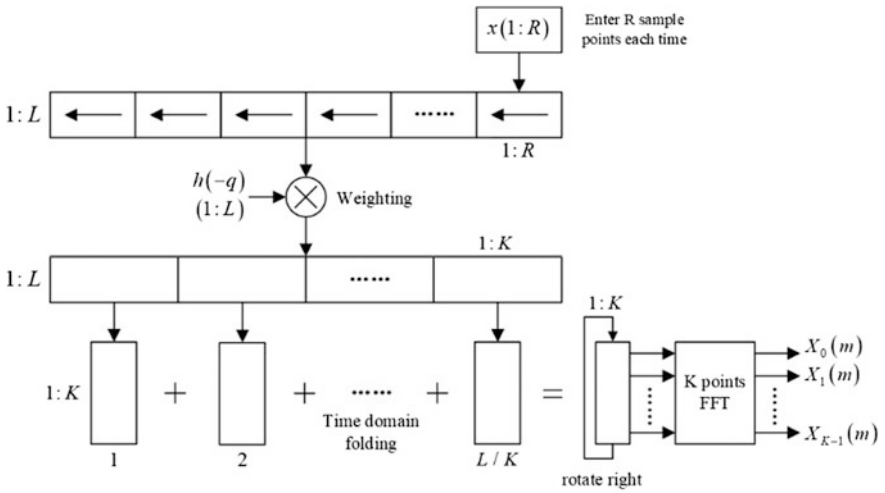



**Fig. 2.** Implementation structure of WOLA analysis filter bank

From Fig. 4, it can be found that the analysis filter banks based on WOLA structure process data in a piecewise manner. Input $R$ sampling points at each time and output $K$ sub-band sampling points. Specifically, the whole analysis end signal processing flow can be divided into five parts.

**Data segmentation**
First, the single channel signal $x(n)$ is divided into several segments according to the sampling order, and the length of each segment is extracted factor $R$. So, each time corresponds to $R$ sampling points, and the data rate is reduced by $R$ times. Then, a shift register with a length of $L$ is constructed. In every moment, there was a new data from the end of the input shift register, and a piece of data from the other end of this out, each operation preserves the previous time $L - R$ sampling points.

**Data weighting**
At any time, the data of the shift register is multiplied with the prototype filter $h(-q)$ and then the data of the output is stored in the register.

**Overlapping accumulation**
According to Formula (5), the weighted $L$ data is divided into a group of $K$ samples and each group is divided into $L/K$ groups. Then, accumulate the $L/K$ data and output $K$ sampling point data.

**Circular shift**
It is known by Formula (4) that $\tilde{X}_k(m)$ is the $K$ point DFT operation of $\tilde{x}_m(p)$, so signals $X_k(m)$ and $\tilde{X}_k(m)$ are both considered as frequency-domain signals, and $X_k(m) = \mathrm{e}^{-j\omega_k mR}\tilde{X}_k(m)$ can be seen as a multiplying in the frequency domain. According to the theory of signal processing, the complex modulation operation in frequency domain can be equivalent to the cyclic shift operation in time-domain.

$$Y(\omega) = X(\omega)\mathrm{e}^{-j\omega k_0} \Leftrightarrow y(k) = x\big((k - k_0)_{\mathrm{mod}\ K}\big) \tag{6}$$

where, $Y(\omega)$ and $X(\omega)$ are frequency-domain signals corresponding to time series $y(k)$ and $x(k)$, respectively. The number of points in the sequence is $K$, and the $k_0$ is the shift. According to Formula (6), the operation $X_k(m) = \mathrm{e}^{-j\omega_k mR}\tilde{X}_k(m)$ can be equivalent to the shift operation in time-domain. Specifically, the cyclic shift operation is performed before FFT, and the number of shifted samples is the norm of $-Rm$ and $K$. Thus, the cyclic shift operation takes the place of the complex modulation operation after FFT and has higher computing efficiency.

**FFT**
Because the number of sub-bands generally takes two of the index, the solution of a can be realized by FFT. Considering that the input signal is generally a real number, the output signal $X_k(m)$ after FFT must be symmetric. So, we only need to intercept half of the sequence data. Here, the sequence data $k = 0, 1, \ldots, K/2 - 1$ is selected.

### 3.2 Synthetic Filter Bank Implementation Structure

The process of integrating $K$ sub-band signals into $\hat{x}(n)$ can be described as follows:

$$\hat{x}(n) = \frac{1}{K}\sum_{k=0}^{K-1}\left[Y_k(n) * f(n) \times \mathrm{e}^{j\omega_k n}\right] = \frac{1}{K}\sum_{k=0}^{K-1}\sum_{m=-\infty}^{\infty}\hat{X}_k(m) f(n - mR)\mathrm{e}^{j\omega_k n} \tag{7}$$

where

$$Y_k(n) = \begin{cases} \hat{X}_k(m), & n = mR \\ 0, & \text{other} \end{cases} \tag{8}$$

Considering the length of the prototype low-pass filter is finite, so the range of the value of $m$ must be $0 \le n - mR \le L - 1$. Set $n = r + MR$, where $r = 0, 1, \ldots, R - 1$, and M is an integer. The range of the value of $m$ is $M - \lceil L/R \rceil \le m \le M$, where $\lceil \cdot \rceil$ represents the smallest integer that is not less than $L/R$, and $\alpha = \lceil L/R \rceil$, thus

$$\hat{x}(r + MR) = \sum_{m=M-\alpha}^{M} f(r + (M - m)R)\frac{1}{K}\sum_{k=0}^{K-1}\tilde{\hat{X}}_k(m)\mathrm{e}^{j\omega_k(r+(M-m)R)} \tag{9}$$

where $\tilde{\hat{X}}_k(m) = \hat{X}_k(m)\mathrm{e}^{j\omega_k mR}$, and the formula corresponds to the inverse complex modulation process of the analytical end [Formula (2)]. For the convenience of analysis, set

$$U_m(r + (M - m)R) = \frac{1}{K}\sum_{k=0}^{K-1}\tilde{\hat{X}}_k(m)\mathrm{e}^{j\omega_k(r+(M-m)R)} \tag{10}$$

Because of the periodicity of $\mathrm{e}^{j\omega_k(r+(M-m)R)}$ (the period is $K$), so $\mathrm{e}^{j\omega_k(r+(M-m)R)} = \mathrm{e}^{j\omega_k(r+(M-m)R)_{\mathrm{mod}\,K}}$, Formula 2 can be rewritten as

$$U_m(r + (M - m)R) = \frac{1}{K}\sum_{k=0}^{K-1}\tilde{\hat{X}}_k(m)\mathrm{e}^{j\omega_k(r+(M-m)R)_{\mathrm{mod}\,K}} = U_m\left((r + (M - m)R)_{\mathrm{mod}\,K}\right) \tag{11}$$

then

$$\begin{aligned}\hat{x}(r + MR) = {} & f(r)U_M(r) + f(r + R)U_{M-1}\left((r + R)_{\mathrm{mod}\,K}\right) \\ & + f(r + 2R)U_{M-2}\left((r + 2R)_{\mathrm{mod}\,K}\right) + \cdots + f(r + \alpha R)U_{M-\alpha}\left((r + \alpha R)_{\mathrm{mod}\,K}\right)\end{aligned} \tag{12}$$

Because of $r = 0, 1, \ldots, R - 1$ and $R \le K$, so $r = r \mod K$ in Formula (12). The implementation structure of the WOLA synthetic filter bank can be obtained from the above derivation process, as shown in Fig. 3.

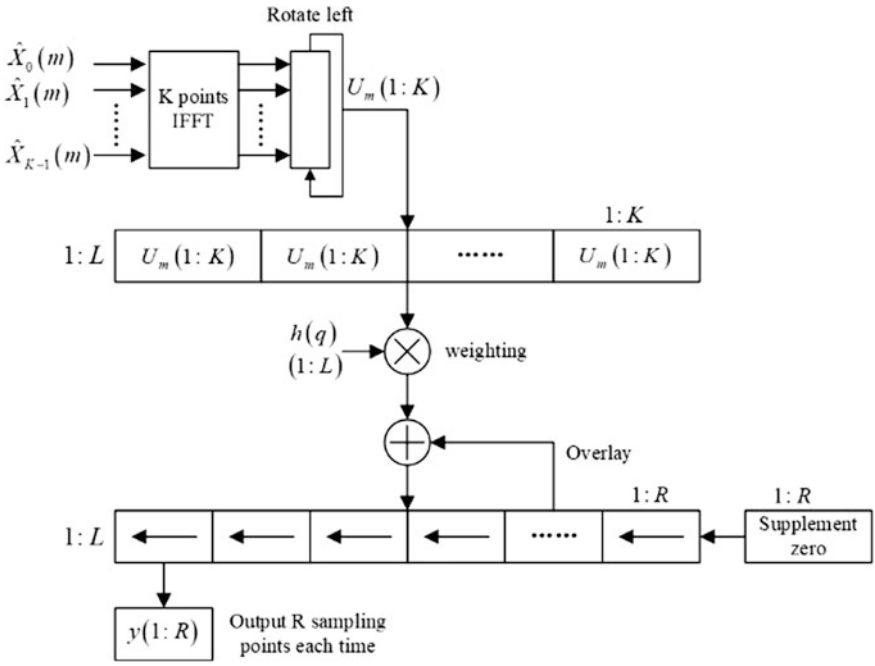



**Fig. 3.** Implementation structure of WOLA synthetic filter bank

Figure 3 shows that data processing in parallel based on group integrated WOLA filter structure, and each time input $K$ sub-band signal sampling, $R$ output signal reconstruction of sampling points. Specifically, the whole signal processing flow can be divided into five parts.

**IFFT**

When the output of the analysis terminal is selected, only the sequence data of $k = 0, 1, \ldots, K/2 - 1$ is selected. Therefore, before the IFFT operation, $\hat{X}_k(m), k = 0, 1, \ldots, K - 1$ is needed to reconstruct the sequence according to the conjugate symmetry properties of the spectrum.

**Reverse circulation shift**

This operation is used to implement $\tilde{\hat{X}}_k(m) = \hat{X}_k(m)e^{j\omega_k mR}$, the points number of cyclic shift is the norm of $Rm$ and $K$.

**Data weighting**

First, copy $L/K$ copies of the $K$ points $U_m(1:K)$, the reverse cycle shift output, and deposit in a register with a length of $L$.Then, the $L$ data is multiplied with the prototype low-pass filter $f(n)$ point-to-point and output the $L$ point data.

**Overlapping accumulation**

A shift register with length $L$ is constructed. At each $m$-time, the weighted $L$-point data is first added to the $L$-point data in the shift register point-to-point. Then, all zero data of

length $R$ is input from one end of the shift register, and $R$ samples are removed from the other end. The $R$ sample points are the sampling points of reconstructed signal at $m$-time.

## 4 Simulation Results

The parameters of the WOLA-structured filter bank are shown in Table 1.

The typical signal in the frequency-domain synthesis system is the wideband BPSK signal. The input BPSK signal is a real baseband signal without carrier modulation. The source adopts bipolar code and obeys uniform distribution. The number of sampling points per symbol is 2. The SQRT filter is used in the shaping filter, and the filter order is 256, and the roll-off factor is 0.5.

**Table 1.** Parameters of the WOLA-structured filter bank

| | |
|---|---|
| Sampling rate $F_s$ (MHz) | 1280 |
| Number of sub-band $K$ | 256 |
| Extract factor $R$ | 176 |
| Length of filter $L$ | 11,264 |
| Oversampling factor OS$(= K/R)$ | 1.45 |
| Roll-off factor $\alpha(=$ OS $- 1)$ | 0.45 |

Figure 4(left) gives the reconstructed signal after removing the previous sample and compares it with the input signal, $\tau_{\mathrm{WOLA}} = 11,088$. For clear display, only the sampling points in the 10,000–10,100 interval are intercepted here. It can be seen that the reconstructed signal basically coincides with the input signal. To further investigate the reconfiguration performance, subtract the two signals and get the difference signal between them. Figure 4(right) shows the amplitude difference between the reconstructed signal (after removing the delay sample) and the input signal. It can be seen that the reconstruction error is $1 \times 10^{-3}$ within the scope and can be ignored. In conclusion, the WOLA-structured filter banks have better reconstructed performance.

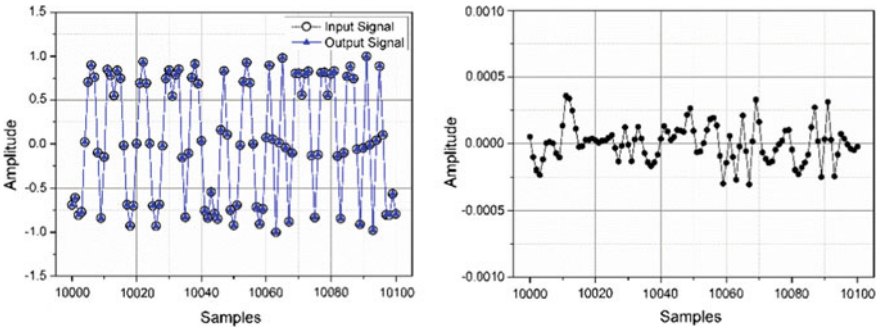



**Fig. 4.** Reconstruction performance of wideband BPSK signal

## 5 Conclusion

In this paper, sub-band segmentation and reconstruction techniques based on WOLA-structured filter banks are studied. Compared with the multiphase DFT-structured filter banks currently used, the WOLA structure has the following advantages: (1) after the data grouped, the sampling rate of the whole structure is reduced by $R$ times. (2) with the help of efficient FFT operation, the hardware consumption of the system is reduced. (3) there is no strict relationship between the extraction factor $R$ and the number of sub-bands $K$. The structure is more efficient and more flexible parameter settings.

## References

1. Lee CH, Vilnrotter V, Satorius E, et al. Large-array signal processing for deep-space applications. California: JPL IPN Progress Report 42-150, 2002.
2. Million S, Shah B, Hinedi S. Comparison of two maximal ratio combining techniques in antenna arraying. IEEE Trans Commun. 1996;44(11):1599–609.
3. Mileant A, Hinedi S. Overview of arraying techniques for deep space communications. IEEE Trans Commun. 1994;42(2/3/4):1856–65.
4. Million S, Shah B, Hinedi S. A comparison of full-spectrum and complex-symbol combining techniques for the Galieo S-band mission. California: JPL TDA Progress Report 42-116, 1994.
5. Navarro R, Bunton J. Signal processing in the deep space array network. California: JPL IPN Progress Report 42-157, 2004.
6. Shang Y, Feng XT. MLC-SUMPLE algorithm for aligning antenna arrays in deep space communication. IEEE Trans Aerosp Electron Syst. 2013;49(4):2828–34.
7. Bai YF, Wang XH, Gao CJ, et al. Adaptive correlation algorithm for aligning antenna arrays in deep space communication. Electron Lett. 2013;49(12):733.
8. Shen C, Yu H. Time-delay alignment technique for a randomly distributed sensor array. IET Commun. 2011;5(8):1068–72.
9. Pham TT, Jongeling AP, Rogstad DH. Enhancing telemetry and navigation performance with full spectrum arraying. In: 2000 IEEE aerospace conference proceedings, California, 2000.

# A Photovoltaic Image Crack Detection Algorithm Based on Laplacian Pyramid Decomposition

Dai Sui[1] and Dongqing Cui[2(✉)]

[1] Guangdong Testing Institute of Product Quality Supervision, Foshan 528300, Guangdong, China
daisui@qq.com
[2] Information and Technology College, Dalian Maritime University, Dalian, China
864020528@qq.com

**Abstract.** Aiming at detecting cracks in photovoltaic images, a crack detection algorithm of photovoltaic images based on Laplacian pyramid decomposition is studied in this paper. Firstly, in order to suppress noise from the crack area, the image is subjected to a filtering process and contrast enhancement operation. Then, the multi-scale edge detection based on Laplacian pyramid decomposition is applied to the processed image to extract the edge of the image. The results of the extracted fractures are optimized to eliminate the influence of partial noise. Through tests and comparisons, the algorithm is proved effective on crack detection for photovoltaic image.

**Keywords:** Photovoltaic image · Crack · Filtering · Laplacian pyramid

## 1 Introduction

In today's world, non-renewable resources are being consumed more and more, energy reserves are getting less and less, and renewable resources are inexhaustible. Therefore, replacing non-renewable resources with renewable resources has drawn great attention. As one of the safe and clean energy sources, solar energy has received extensive attention. Solar panels can convert light energy into electrical energy and achieve the utilization of renewable resources, but there will be some defects due to various reasons during the use of batteries, such as cracks, debris, and black edges. A defective solar cell will seriously affect the quality of the solar cell and affect the efficiency of the photoelectric conversion. Therefore, it is very important to realize fast and efficient defect detection for solar cells. This paper is mainly about the research on the detection algorithm for the crack defect of the electroluminescent solar cell. At present, there are few research methods in this field. Most of the tests are detected by artificial naked eye, which not only has low detection efficiency, but also is easy to leak detection. With the development of computer and image processing technology, it is possible to make use of the existing technology to automatically detect the crack of solar cell chip.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_73

At present, the crack detection algorithms are mainly aimed at the road surface cracks, including the wavelet modulus maxima [1, 2], the tensor voting algorithm [3], and so on. However, because the structure and surface characteristics of solar panels are much more complex than those of the ground, it is not possible to extend these algorithms directly to the detection of cracks in solar panels. In this paper, a crack detection algorithm based on Laplacian pyramid [4, 5] decomposition is proposed. The proposed algorithm is compared with the commonly used edge detection algorithm and pavement crack detection algorithm, which shows the advantages of the proposed algorithm in crack detection, especially for complex background situations.

## 2 Fracture Detection Algorithm Based on Laplacian Pyramid Decomposition

Photovoltaic imaging crack detection algorithm is shown as Fig. 1, and the main steps are as follows:

Step 1: Preprocessing of the photovoltaic image, including filtering, contrast enhancement, and dividing the battery into small pieces [6].
Step 2: The image is decomposed by Laplacian pyramid [4], and the image subsets at different scales are extracted.

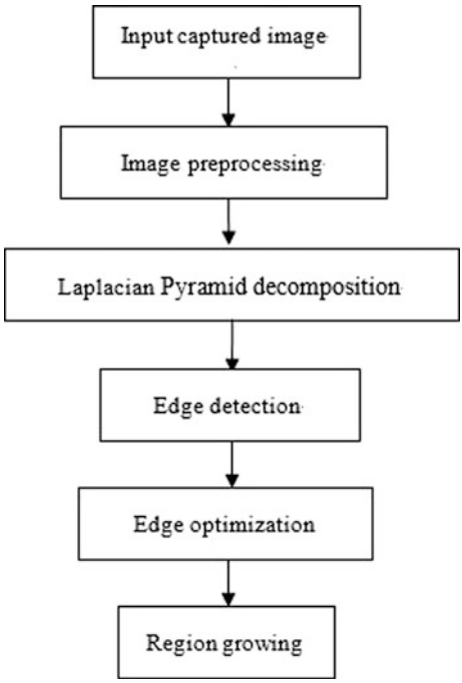



**Fig. 1.** Flowchart of crack detection algorithm based on Laplacian pyramid decomposition

Step 3: The edge of the processed image is detected, and the crack profile is extracted [7].
Step 4: The results are optimized by edge, and the non-fractured region is removed.
Step 5: The use of regional growth to complete or disconnected edges [8].

### 2.1 Image Preprocessing

Because the near-infrared camera used for shooting does not achieve high cooling function, the photovoltaic image taken (Fig. 2) is affected by a lot of noise. It is necessary to de-noise the image before the crack detection is carried out on the image. This paper mainly through the filtering operation of noise reduction, for some types of random noise median filter has a very good noise reduction capability. In the noise reduction processing, the noise point is directly ignored, and the noise reduction caused by the fuzzy effect is low, so this paper selects the median filter [9] for noise reduction. Meanwhile, in order to better detect cracks, we need to contrast the image after noise reduction. After enhanced processing, the target area and background have higher contrast, and the crack area is more obvious.

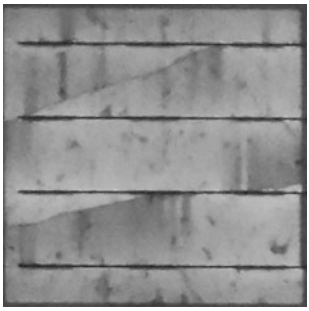


**Fig. 2.** Photovoltaic images

### 2.2 Laplacian Pyramid Decomposition

One of the difficulties in edge detection is the uncertainty of the image spatial scale. Laplacian pyramid decomposed the image into subsets of different scales. Each size image provided some edge information, so that we could achieve more precise edge detection.

Laplacian pyramid image decomposition is implemented as follows:

First, the original image as the bottom image $G_0$ (Gaussian pyramid layer 0), and then the Gaussian kernel ($5 \times 5$) is used to convolute $G$). Then, the image after the convolution is down to be sampled (removing the even number of rows and columns) to get the upper layer of the image $G_1$. Taking this image as input, repeat convolution and down-sampling operation to get a more advanced image and repeat it many times

to form a pyramid-shaped image data structure, namely Gauss pyramid (Fig. 3). The construction process is as follows:

Suppose Gauss pyramid's layer $l$ image is $G_l$:

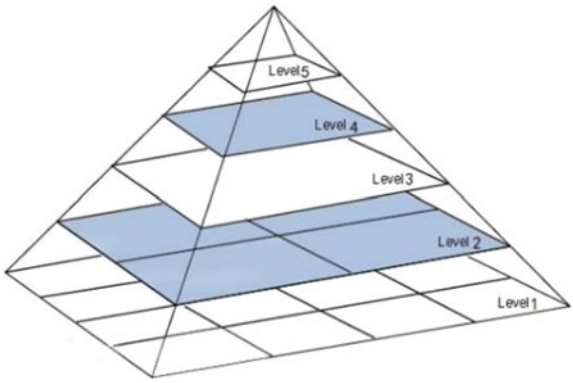



**Fig. 3.** Gauss's pyramid demonstration

$$G_l(i,j) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} \tilde{\omega}(m,n) G_{l-1}(2i+m, 2j+n) \quad (0 \le l \le N, 0 \le i \le R_l, 0 \le j \le C_l) \tag{8}$$

where $N$ is the Gauss pyramid top layer number; $R_l$ and $C_l$ are the number of rows and columns of the Gauss pyramid layer l; $\tilde{\omega}(m,n) = h(m) \cdot h(n)$ is a two-dimensional detachable $5 \times 5$ window function, $h$ is taken as Gaussian density distribution function.

Subtracting the front image from image of each layer of Gaussian pyramid to get a series of difference images is the Laplacian decomposition image.

The $G_l$ is interpolated to get the magnified image $G_l^*$, so that the size of the $G_l^*$ is the same as that of the $G_{l-1}$. The specific practice is:

Step 1: The image is expanded to two times in each direction, and the new rows and columns are filled with 0.

Step 2: Using the previous kernel and the enlarged image to convolution to obtain the approximate value of the "new pixel"

Expressed as,

$$G_l^*(i,j) = 4 \times \sum_{m=-2}^{2} \sum_{n=-2}^{2} \tilde{\omega}(m,n) G_l\left(\frac{i+m}{2}, \frac{j+n}{2}\right)$$
$$(0 \le l \le N, 0 \le i \le R_l, 0 \le j \le C_l) \tag{9}$$

where,

$$G_l^*\left(\frac{i+m}{2}, \frac{j+n}{2}\right) = \begin{cases} G_l\left(\frac{i+m}{2}, \frac{j+n}{2}\right), & \frac{i+m}{2}, \frac{j+n}{2} \in Z \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

$Z$ is an integer set, let,

$$\mathrm{LP}_l = \begin{cases} G_l - G_{l+1}^*, & 0 \le l < N \\ G_N, & l = N \end{cases} \tag{11}$$

In the formula, $N$ is the Laplacian pyramid top layer number, $\mathrm{LP}_l$ is the $l$ layer image Laplacian pyramid decomposition. The Laplacian pyramid can be fused at each level, the methods are as follows:

$$\mathrm{LP} = \sum_{i=1}^{N} \mathrm{LP}_i \tag{12}$$

### 2.3 Edge Detection

In order to detect the crack defect of the photovoltaic image accurately, the edge detection of the processed image is needed. In the image edge detection, the suppression of noise and the precise location of the edge can't be met at the same time, in the image edge detection and suppression of accurate positioning noise and the edge is not satisfied at the same time, some of the edge detection algorithm by smoothing filter to remove noise and increase in edge location uncertainty, and improve the sensitivity of edge detection operator on edge at the same time, but also improve the sensitivity to noise. The Canny operator can accurately locate the edge at the same time, so the Canny [7] operator is used for edge detection in this paper.

### 2.4 Edge Optimization

Because the result of edge detection is inevitably influenced by noise and other factors, according to the results of detection and the analysis of the characteristics of cracks in photovoltaic images, we can effectively remove the influence of noise and other factors, operation is as follows:

Step 1: Before the need to optimize the edge of the edge detection results are an expansion and refinement, connecting part of the edge off, to prevent accidentally deleted edges and simultaneously eliminate parallel cracks.

Step 2: In order to find edge endpoints, you need to disconnect the intersection point. It traverses every pixel in the connected domain through 3 * 3 window. Finally, count the number $z$ of pixel points that each pixel $(x, y)$ is an edge point in 8($x \pm 1$, $y \pm 1$) neighborhoods inside the window, and when $z > 2$, it indicates that the point is an intersection and the pixel in the point becomes 0, Disconnect the edge.

Step 3: The threshold is set to remove the smaller edge of the length, and the threshold is determined according to the actual situation.

Step 4: The two endpoints of each connected domain are obtained. A second step method is used to find the point of the $z = 1$ in the pixel point, which is the endpoint of the connected domain.

Step 5: Through the coordinates of the two endpoints, find the distance between two points, through the distance and the number of connected pixels to compare, when the two values close to that is approaching the vertical or horizontal edge, remove it, When the distance is much less than the sum of the numbers, the edge curvature is relatively large, and it is also cleared.

## 3 Results and Analysis

In order to prove the validity and accuracy of the proposed multi-scale crack detection algorithm based on Laplacian pyramid decomposition, the algorithm proposed in this paper is compared with the wavelet modulus maxima [10].

Multi-scale wavelet modulus maxima method of photovoltaic image for crack detection result is shown in Fig. 4. Among them, the (a), (e), (i) is the photovoltaic image segmentation into small pieces. (b), (c), (d), (f), (g), (h) and (j), (k), (l) are test results of the diagram of the original image (a), (e), (i) under different scale $s = 1, 2, 3$. It can be seen in both the noise and the edge of the integrity and so on, wavelet modulus maxima method of photovoltaic image detection effect is not very ideal.

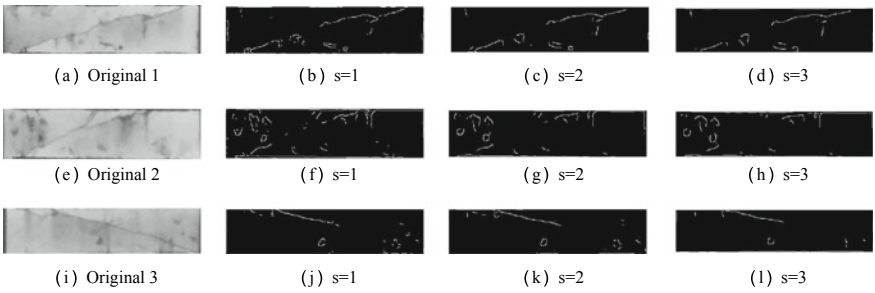



**Fig. 4.** Multi-scale wavelet modulus maximum method photovoltaic imaging crack test results

The results of the crack detection in the multi-scale Laplace pyramid are shown in Fig. 5, where (a), (f), and (k) are small pieces of the photovoltaic image. (b), (g), and (l) are the results of direct detection of the edge detection by the original image. (c), (h), (m) is the original image edge detection based on Laplacian pyramid decomposition of image, (d), (I), (n) is on (c), (h), (m) for the optimized results. Comparing (b), (g), (l) and (c), (h), (m), it can be seen that the proposed edge detection algorithm in suppressing noise and the integrity of the edge detection has a better effect. Through the (d), (I) and (n) as you can see, the edge optimization can effectively remove the noise, the influence of such factors as remove mistakenly identified, make the detection more accurate.

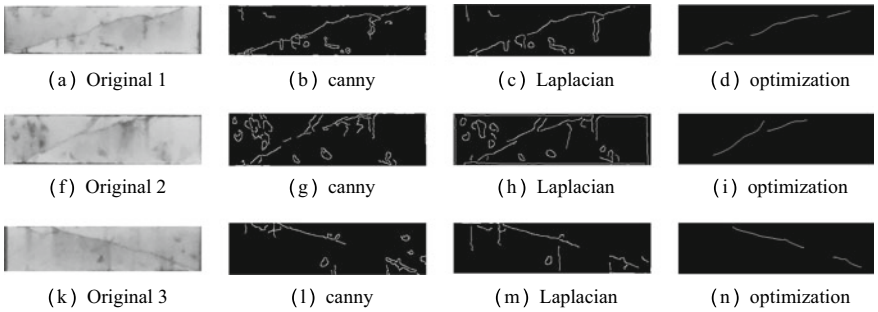



**Fig. 5.** Multi-scale Laplacian pyramid decomposition of photovoltaic image crack test results

Compared with Figs. 4 and 5, when the two algorithms detect images, it is very obvious that the image is affected by noise. The detection result of wavelet modulus maximum method is very unsatisfactory. The algorithm proposed in this paper can effectively suppress noise and the edge detection is complete.

Solar cells studied in this paper is provided by the Guangdong Testing Institute of Product Quality Supervision, respectively, 30 pieces of photovoltaic images without cracks and 30 pieces of photovoltaic images with cracks are selected. The statistical results of the two groups of images are shown in Table 1.

**Table 1.** Result

| Number of pictures | Number of cracks | Detection number | Correct detection | Error detection | Missing detection | Accuracy rate (%) | False rate (%) |
|---|---|---|---|---|---|---|---|
| 30 | 35 | 31 | 24 | 7 | 11 | 69 | 20 |

It can be seen from the table that the methods proposed in this paper are missed and false checked to a certain extent, but this method basically meets the requirements without high accuracy.

## 4 Conclusion

This paper presents a photovoltaic image crack detection method based on the decomposition of Laplacian Pyramid, the image is decomposed into multi-scale image, and achieve a reasonable balance between noise reduction and edge location, through the experimental results show that the algorithm for the detection of PV image can reach a high accuracy at the same time, the algorithm extended to the pavement crack detection way detection can also very good face.

## References

1. Yang XN, Huang H, Xiao-Yu XU, et al. Edge extraction based on multi-scale wavelet modulus maxima. J Guizhou Univ, 2013.
2. Jun WF. Multiscale edge detection based on wavelet modulus maxima. Chin J Sci Instrum, 2006.
3. Qian B, Tang Z, Xu W. Pavement crack detection based on improved tensor voting. In: International conference on computer science & education. New York: IEEE; 2014. p. 397–402.
4. Dong HY, Wang L, Li JC, et al. Multiscale edge detection based on Laplacian pyramid. Opto-Electron Eng. 2007;34(7):135–40.
5. Sandic-Stankovic D, Kukolj D, Callet PL. Image quality assessment based on pyramid decomposition and mean squared error. In: Telecommunications forum telfor. New York: IEEE; 2015. p. 740–3.
6. Gonzalez RC, Woods et al. Digital image processing. Publishing House of Electronics Industry, 2010.
7. Ding L, Goshtasby A. On the Canny edge detector. Pattern Recogn. 2001;34(3):721–5.
8. Jiang HY, Yue-Peng SI, Luo XG. Medical image segmentation based on improved Ostu algorithm and regional growth algorithm. J Northeast Univ. 2006;27(4):398–401.
9. Zhe LV, Wang FL, Chang YQ, et al. Region-adaptive median filter. J Syst Simul. 2007;19 (23):5411–4.
10. Zou Q, Cao Y, Li Q, et al. CrackTree: Automatic crack detection from pavement images. Pattern Recogn Lett. 2012;33(3):227–38.

# Semantics Images Synthesis and Resolution Refinement Using Generative Adversarial Networks

Jian Han(✉), Zijie Zhang, Ailing Mao, and Yuan Zhou

School of Electrical and Information Engineering, Tianjin University, Tianjin, China
{jianhan,zhouyuan}@tju.edu.cn

**Abstract.** In this paper, we proposed a method to synthesizing a super-resolution image with the given image and text descriptions. Our work contains two parts. Wasserstein GAN is used to generate low-level resolution image under the guidance of a novel loss function. Then, a convolution net is followed to refine the resolution. This is an end-to-end network architecture. We have validated our model on Caltech-200 bird dataset, Oxford-102 flower dataset, and BSD300 dataset. The experiments show that the generated images not only match the given descriptions well but also maintain detailed features of original images with a higher resolution.

**Keywords:** Generative Adversarial Networks (GANs) · Semantics images synthesis · Resolution refinement

## 1 Introduction

It is feasible to integrate pictures with natural language. Reed's work [1] reveals that the synthesis between images and text semantics through the neural network is practical, but the results may contain noise and lose some details. Our contributions can be divided into two parts. First, inspired by the loss function proposed in [2], we propose a novel loss function to evaluate the output, which made the network pay more attention to the details. Second, we introduced a subnet to refine the generated images resolution. The results generated by the network look more realistic. There are three parts in our network: generator G, discriminator D, and enhancer E. The input of the generator is made up of the origin image and text semantics, which encoded by a pretrained text model. The goal of the generator is to decode the combined feature representations to a synthesized image. Then, the discriminator judges the outputs legality. The legality means whether the output has the features described by the text semantics. The

Han and Zhang—Equal contribution.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_74

enhancer's goal is to improve the resolution of the output of the generator. How to combine the image with other format of information is an important research problem in computer vision. By adopting deep neural networks, a large number of models have been developed to achieve this goal recently, such as variational autoencoders (VAE) [3], autoregressive models and Generative Adversarial Networks (GANs) [4]. Zhang et al. [5] proposed StackGAN, which has successfully generated the corresponding image with text semantics by decomposing the process into two stages. Reed et al. [1] proposed a two-step method that can synthesize images by given query images and text descriptions, denoted style transfer in their paper. Donahue et al. [6] proposed models that train the generator and encoder network via adversarial learning. Larsen et al. [7] combined VAE and GAN directly by collapsing the generator of GAN and decoder of VAE into one network and replaced the element-wise metric of original VAE with feature-wise metric imposed by the discriminator. Dong et al. [2] proposed a novel network architecture by using residual network with conditional GAN. On their paper, they achieve better results, but at some points, they lose some details.

## 2 Background

### 2.1 Generative Adversarial Networks

Generative Adversarial Network is proposed by Ian Goodfellow [4], which consists of two networks: a generative network $G$ and a discriminator network $D$. The generator, or the generative network, generates fake samples $G(z)$ from noise $z$, and the discriminator, or the discriminative network, takes both real samples x and fake samples $G(z)$ and calculated probability $D(x)$ or $D(G(z))$ to distinguish them. Both generator and discriminator are trained mutually: The generator attempts to fool the discriminator, while the discriminator attempts distinguish between real data and fake samples synthesized by the generator. This training process can be described mathematically by a minmax function over the value function $V(D, G)$:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))] \quad (1)$$

### 2.2 Baseline Method

We will use the method proposed by Reed et al. [1] as a baseline approach. More specifically, Reed's method is that they trained a style encoder network $S$ to invert the generator network $G$, such that the synthesized images $x$ generated by G were mapped back to latent variable $z$.

## 3 Method

Our advanced model consists of three parts, a generator used to generate synthesized image, a discriminator used to judge the output's legality, and an enhancer

used to refine the output through super-resolution. GANs, or specifically conditional GANs, rely on two conditions, image and text descriptions. For generator $G$, we applied an encoder–decoder architecture. The main goal of the encoder is to extract the features in the image and text semantics, while the decoder's goal is to synthesize the features together and produce a sophisticated image. The discriminator $D$ judges whether the output meets the description of the text. After the generation step, the well-performed image will be sent into our enhancer $E$ to refine the resolution. The enhancer can be trained parallelly on different dataset.

### 3.1 Network Architecture

The architecture of our network is illustrated in Fig. 1. We can see from it that the generator $G$ is made up of three parts, encoder, residual blocks, and decoder. The encoder is made up of three convolutional layers, which are used to extract features in the input image. These three convolutional layers are critical to our model performance, and their parameters need to be finely tuned. The input image size is 64 * 64 * 3, while the size of output of the encoder is 16 * 16 * 512. ReLU activation function is used after every convolutional layer. And batch normalization is performed in all layers except the first layer. Additionally, we can also replace the encoder with a much deeper network, e.g., VGG-19. For text descriptions, we use a pretrained text encoder $\phi$ to extract literal features. We adopted the method by Zhang et al. [5], which performs a text embedding augmentation method on $\phi$(t). After the augmentation, we duplicate the text embeddings to be 16 * 16 * 128 and concatenate it with the output of the encoder. After the residual blocks, the synthesized features will be into decoder, which will produce a synthesized image. The decoder's structure is very symmetry with the encoder. Similarly, ReLU activation function and batch normalization strategy are performed on every layer except the last layer. During the training process, the output of the decoder will directly go into the discriminator D,

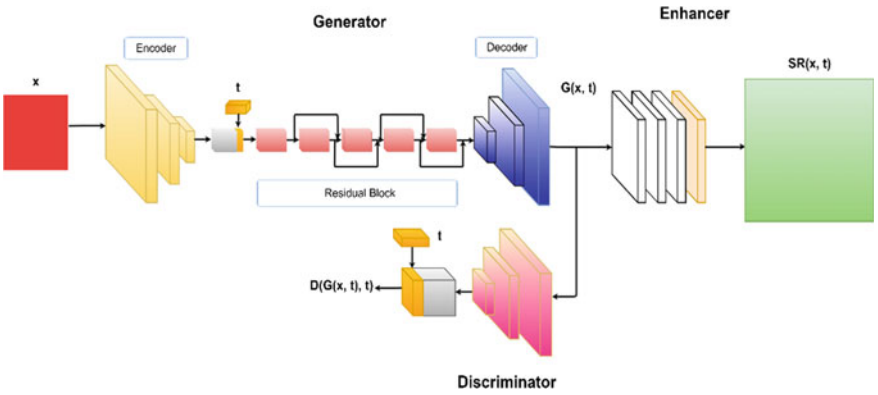



**Fig. 1.** The architecture of our network

rather than into enhancer E. This tactic will not only make the architecture easier but also reduce the computational complexity. For discriminator D, we use one convolutional layer to downsample the images. Then, we concatenate the text embeddings with the downsampled feature representations. After that, two convolutional layers are applied to produce judged probabilities.

### 3.2 Resolution Refinement Subnetwork

This enhancer is called ESPCN by Shi [8]. The structure of the enhancer is very simple. From Fig. 2, enhancer consists of three convolution layers and a subpixel convolutional layer. The first three layers will extract features and the current output size is the same as the input size, but the number of feature channels will be the square of the scale factor r. Thus, current shape of the output of $r^2 \cdot C \cdot H \cdot W$. In the subpixel convolutional layer, we will use interpolation method to convert the $r^2 \cdot C \cdot H \cdot W$ features representations to $C \cdot rH \cdot rW$ output. Compared with the method that first converts into rH · rW shape, then performs convolutional operations, this method achieves higher PSNR and computes more efficiently.

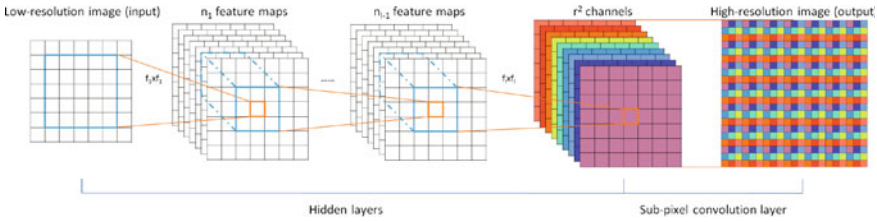



**Fig. 2.** Structure of enhancer

### 3.3 Loss Function

In the proposed method, discriminator D is fed with four types of input pairs, and outputs the independent probabilities of the input pair being real and matching with the text:

$$\begin{aligned} S_{r+} &\leftarrow D(x, \phi(t)), \\ S_{w-} &\leftarrow D(x, \phi(\bar{t})), \\ S_{s-} &\leftarrow D(\hat{x}, \phi(\hat{t})), \\ S_{d-} &\leftarrow D(\hat{x}, \phi(t)) \end{aligned} \tag{2}$$

Among them, $S_{r+}$ is a common input that adopted in vanilla cGAN, which is a real image(represented as x) and corresponding text(represented as t). $S_{w-}$ (pictures and irrelevant text descriptions represented as $\bar{t}$) is proposed by [5], which

can greatly enhance the discriminator D to distinguish whether the text and the image correspond to each other, further promoting generator G to synthesize an image that better matches the text description. The term $S_{s-}$, proposed by [2] (input pictures and relevant text represented as $\hat{t}$ to $G$ and then feed $D$ with the output of $G$ and the text), cannot only reduce the generated image overfitting the text, resulting in the loss of image details, but also reduce overfitting the image, resulting in mismatching with the text description. We proposed $S_{d-}$, that is, input pictures and corresponding text to $G$ and then feed $D$ with the output and the text. Although [2] et al. argue that the corresponding text fed to $G$ in the training phase will cause $G$ to overfit the image and lose the ability to synthesize semantically consistent images. We think that their argument is only established when the input of $G$ is all $S_{d-}$ or $S_{d-}$ to some extent. We believe that introducing $S_{d-}$ to an appropriate degree will not overfit the image, but will make the network generate more realistic images. Since in many cases, $G$ fits well to the text but loses the information of the original image, which accounts for the synthesized images lost a lot of details and look fake. In order to make the network converges faster and improve its robustness to $G's$ structure, we use WGAN [9]. According to the discussion above, we formulate the loss function as follows:

$$\begin{aligned} L_D = & E_{(x,t)\sim P_{data}}[D(x, \phi(t))] \\ & - E_{(x,\bar{t})\sim P_{data}}[D(x, \phi(\bar{t}))] \\ & - E_{(x,\hat{t})\sim P_{data}}[D(G(x, \phi(\hat{t})), \phi(\hat{t}))] \\ & - \lambda E_{(x,t)\sim P_{data}}[D(G(x, \phi(t)), \phi(t))] \\ L_G = & E_{(x,\hat{t})\sim P_{data}}[D(G(x, \phi(\hat{t})), \phi(\hat{t}))] \\ & + \lambda E_{(x,t)\sim P_{data}}[D(G(x, \phi(t)), \phi(t))] \end{aligned} \tag{3}$$

$\lambda$ is the weight of $S_{d-}$, and it equals to 1 in all experiments. The loss functions make the generator capable to fit the distribution of real data $P_{data}(x, t)$.

## 4 Experiment

### 4.1 Experimental Setting

For training, we use Caltech-200 bird dataset [10] and Oxford-102 flower dataset [11] to train our GAN model. The Caltech-200 bird dataset has 11,788 images with 200 classes of birds. We split it into 150 training classes and 50 testing classes. The Oxford-102 flower dataset has 8189 images with 102 classes of flowers and divides it to 82 training classes and 20 testing classes. To train the generator and discriminator, we set initial learning rate of 0.0002. The learning rate is reduced by 0.5 every 100 epochs. Batch size is 32, and the training epochs is 300.

### 4.2 Qualitative Experiment

From Figs. 3 and 4, we can see that our model can handle much complex edges and details. These details can be seen clearly after the super-resolution. Additionally, we noticed that Reed's result may mess the content and the background. Due to our novel loss function, we could overcome this problem. Here is the result.

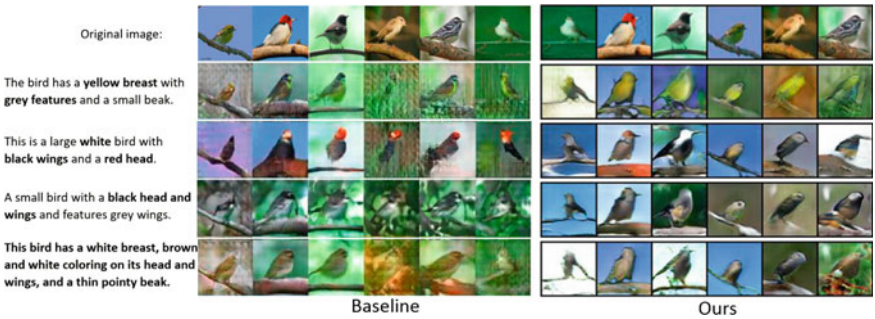



**Fig. 3.** Qualitative comparison in Caltech-200 dataset

**Fig. 4.** Some examples of our proposed method

There is no quantified criterion that has been widely adopted for the evaluation of generative models. Inspired by [2], we conducted a human evaluation to compare the baseline method with ours. We first randomly selected 20 different text descriptions and original images from the test set. Then, we fed them to the proposed method and the baseline to synthesize 20 images, respectively. Ten volunteers of different ages and genders were recruited and presented with former 20 synthesized image pairs without knowing which method they came from. These volunteers were asked to score according to the following criteria:

**Table 1.** Human judgement result

| | Baseline | Ours |
|---|---|---|
| Pose | 0.42 | 0.58 |
| Background | 0.32 | 0.68 |
| Text matching | 0.38 | 0.62 |
| Average | 0.37 | 0.63 |

(1 means better, 0 means worse). Then, we averaged the scores of the volunteers on these three items and got the following results as shown in Table 1.

The results show that compared to the baseline, the proposed method synthesizes more realistic and more semantically matching images while retaining more background details.

### 4.3 Quantitative Experiment

Although human scores are very intuitively and effectively, this method has introduced great randomness and uncertainty. To avoid this shortcoming, we used inception score to measure the difference between the distribution of real data and generated images. The inception score, proposed by [12], has been widely adopted in the evaluation of GAN although existing controversial. Its core idea is that good generated images should be a good representation of its class while remain diversity. The inception score formulates as follows:

$$Inception\ Score: \quad I = \exp(E_x D_{KL}(p(y|x) \mid\mid p(y)) \tag{4}$$

where $x$ is a generated sample, and $y$ is its predicted label. Since the images we synthesized are mainly flowers and birds, which fit well into ImageNet [13] classes, we use the Inception model directly. The result is shown in Table 2.

**Table 2.** Inception score comparison

| Inception score | | |
|---|---|---|
| | Flowers | Birds |
| Real data | 2.97(0.48) | 3.28(0.42) |
| Baseline | 2.09(0.11) | 2.28(0.47) |
| TGAN | 2.13(0.27) | 2.51(0.21) |
| TGAN(RF) | 2.64(0.14) | 3.02(0.18) |
| TGAN(SI) | 2.30(0.26) | 2.97(0.32) |
| TGAN(RF+SI) | 2.81(0.18) | 3.19(0.28) |

As can be seen, the scores for the proposed method are closer to that of real data distribution for both datasets, which demonstrate that the proposed method outperforms the baseline.

More detailed experiments were carried out to prove the effectiveness of our method. In Table 2, TGAN represents our method without refinement network and $S_{d-}$ input, TGAN(RF) is our method with refinement, TGAN(SI) is our method with $S_{d-}$ input, and TGAN(RF+SI) is our method with both techniques. According to the results above, our bare network is slightly better than the baseline. But when equipped with refinement network and Sd- input, the proposed method outperforms the baseline with large advantage. Besides, on Oxford-102(flowers), TGAN(RF) is greatly improved compared to TGAN. On both datasets, TGAN(SI) and TGAN(RF+SI) have shown better performance than those without $S_{d-}$ input pairs.

## 5 Conclusion

In this work, we proposed an improved way to synthesize images based on text description. Using multiple input pairs and a novel loss function, the proposed network could generate more plausible images. Then through a resolution refinement subnetwork, we further make a contribution on the generated images' quality. Besides, higher resolution images could be generated as well. We evaluate our model on Caltech-200 bird dataset and Oxford-102 flower dataset. The experiment's results reveal that our method achieves the desired results and outperforms the baseline method.

## References

1. Reed S, et al. Generative adversarial text to image synthesis. In: International conference on machine learning, 2016. p. 1060–9 (JMLR.org)
2. Dong H, Yu S, Wu C, Guo, Y. Semantic image synthesis via adversarial learning. In: IEEE international conference on computer vision (ICCV). New York: IEEE; 2017. p. 5707–15.
3. Kingma D, Ba J. Adam: A method for stochastic optimization. In: ICLR, 2014.
4. Goodfellow IJ, et al. Generative adversarial nets. In: Neural information processing systems, 2014. p. 2672–80.
5. Zhang, H, Xu, T, Li, H. StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. In: International conference on computer vision. New York: IEEE; 2017. p. 5908–16.
6. Donahue J, Krahenbuhl P, Darrell T. Adversarial feature learning. In: International conference on learning representations, 2017.
7. Larsen, ABL, Larochelle, H, Winther O. Autoencoding beyond pixels using a learned similarity metric. In: International conference on machine learning, 2016. p. 1558–66 (JMLR.org).
8. Shi W, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Computer vision and pattern recognition. New York: IEEE; 2016. p. 1874–83.
9. Arjovsky M, Chintala S, Bottou L. Wasserstein GAN. arXiv, 2017.
10. Wah C, et al. The Caltech-UCSD Birds-200-2011 Dataset. In: Advances in water resources, 2011.

11. Nilsback M, Zisserman A. Automated flower classification over a large number of classes. In: Indian conference on computer vision, graphics and image processing, 2008. p. 722–9.
12. Salimans T, et al. Improved techniques for training GANs. In: Neural information processing systems, 2016. p. 2234–42.
13. Russakovsky O, et al. ImageNet large scale visual recognition challenge. Int J Comput Vis. 2015;115:211–52.

# Bat Algorithm with Adaptive Speed

Siqing You[1(✉)], Dongjie Zhao[1], Hongjie Liu[2], and Fei Xue[1]

[1] School of Information, Beijing Wuzi University, Beijing 101149, China
93028603@qq.com
[2] Beijing Advanced Innovation Center for Future Internet Technology, Beijing 100124, China

**Abstract.** As a famous heuristic algorithm, bat algorithm (BA) simulates the behavior of bat echolocation, which has simple model, fast convergence and distributed characteristics. But it also has some defects like slow convergence and low optimizing accuracy. Facing the shortages above, an optimization bat algorithm based on adaptive speed strategy is proposed. This improved algorithm can simulate the bat in the process of search based on adaptive value size and adaptive speed adjustment. His approach can improve the optimization efficiency and accuracy. Experimental results on CEC2013 test benchmarks show that our proposal has better global searchability and a faster convergence speed, and can effectively overcome the problem convergence.

**Keywords:** Heuristic optimization algorithm · Bat algorithm · Convergence · Adaptive speed

## 1 Introduction

In order to solve complex computational problems in engineering optimization, computational science and other fields, many heuristic optimization algorithms [1–3] are proposed. According to the theory of the no free lunch theorem (NFL), it is shown that all optimization algorithms can not deal with all optimization problems [4, 5]. Therefore, researchers continue to propose various new heuristic intelligent algorithms [6–8].

In 2010, Yang proposed BAT algorithm based on bat echolocation behavior, which has good parallelism and robustness [9, 10]. Bat algorithm has excellent performance, but there are still some problems such as falling into local optimum and low accuracy. A new bat algorithm with adaptive speed strategy is proposed in this paper. The strategy can effectively prevent the algorithm from falling into local optimum and improve the searchability. The CEC2013 test function [11] is used to test the proposed algorithm, which shows that the proposed algorithm is more efficient than the other three comparison algorithms.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_75

Other parts of this paper: Part 2 outlines the basic bat algorithm. Part 3 introduces the adaptive speed strategy bat optimization algorithm proposed in this paper. Section 4 gives the experimental conclusion. The last is the conclusion of the paper.

## 2 Standard Bat Algorithm

In order to make use of the bionic principle of bat echolocation, three idealized rules need to be assumed in the basic bat algorithm [10]:

(1) All bats use echolocation to sense distances and distinguish between prey and background obstacles.
(2) Each bat flies randomly at position $\overrightarrow{x_i}(t)$ at velocity $\overrightarrow{v_i}(t)$, searching for prey at fixed frequency $f_i$, loudness $A_i(t)$, and rate of emission pulse $r_i(t)$.
(3) The loudness changes from the maximum to the fixed minimum as the number of searches increases (from $A_{\max}$ to $A_{\min}$).

For bat $i$ at time $t$ + 1, the velocity $\overrightarrow{v_i}(t+1)$ and position $\overrightarrow{x_i}(t+1)$ show as follow:

$$\overrightarrow{x_i}(t+1) = \overrightarrow{x_i}(t) + \overrightarrow{v_i}(t+1) \tag{1}$$

$$\overrightarrow{v_i}(t+1) = \overrightarrow{v_i}(t) + (\overrightarrow{x_i}(t) - \overrightarrow{p_i}(t)) * f_i \tag{2}$$

$$f_i = f_{\min} + (f_{\max} - f_{\min}) * \beta \tag{3}$$

where $\overrightarrow{P}(t)$ is the global optimum position of time $t$, and $\beta \in [0, 1]$ is a random following the normal distribution.

The new location update formula of bat algorithm is as follows:

$$\overrightarrow{x_i}(t) = \overrightarrow{p_i}(t) + \varepsilon A_{\text{mean}}(t) \tag{4}$$

where $\varepsilon \in [-1, 1]$ is a random number and $A_{\text{mean}}(t)$ is equal to the average of all loudness at time $t$.

With the bats approach their target, the loudness $A$ decreases and $r$ increases. They are following the formulas:

$$A_i(t+1) = \alpha A_i(t) \tag{5}$$

$$r_i(t+1) = r_i(0)[1 - \exp(-\gamma t)] \tag{6}$$

where $r_i(0)$ is the initial value of pulse rate, $\gamma$ and $\alpha$ are constant.

## 3 Bat Algorithm with Adaptive Speed

Bat individual will inherit the previous generation of speed, and the best individual flight departing. When the algorithm progresses to the later stage, most bat individuals fall into local optimization, and the bat's speed cannot make it escape local optimization. To solve this problem, we introduced the concept of average speed.

$$\overline{v^t} = \sum_{i=1}^{n} v_i^t \tag{7}$$

which $\overline{v^t}$ represents the first generation of the average speed of all the bat, $n = 100$ as the number of the population, the standard algorithm bat sphere function as an example of the change in average speed will be shown in Fig. 1.

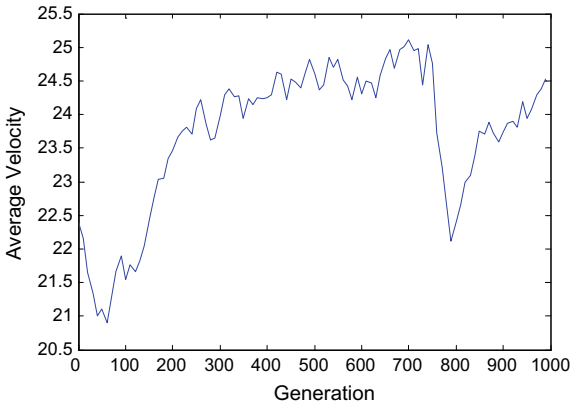



**Fig. 1.** Bat average speed change map algorithm

In view of this shortcoming, we propose an adaptive algorithm bat speed. This particular improvement is based on the fitness of each individual (optimal distance from the closer adaptation higher the value, the lower the farther away) its speed plus a weight in accordance with the right to a strategy. Obviously, if the individual is closer to the optimal solution than others, its adaptation value is higher. We should slow down the individual rapidly, so that the individual converges to the optimal solution. In this case, we give a smaller weight value. On the contrary, if the individual's adaptation value is very low, it means that the individual is far from the optimal solution. We need to give it a larger weight value.

According to the above analysis, weight calculation method is designed as follows:

$$w_i^t = \frac{e_i^{\text{score}_i^t}}{e_i^{\text{score1}} + e_i^{\text{score2}} + \cdots + e_i^{\text{score}n}} \tag{8}$$

where the symbol $e_i^{\text{score}_i(t)}$ is defined as:

$$\text{score}_i(t) = \begin{cases} 1, & \text{if } (f_{\text{worst}} = f_{\text{best}}) \\ \frac{f_{\text{worst}} - f_i}{f_{\text{worst}} - f_{\text{best}}}, & \text{otherwise} \end{cases} \tag{9}$$

## 4 Simulation Results

To test the performance of ASBA, the CEC2013 test functions [11] are employed, witch is a set of 28 shifted and rotated benchmark problems. In the laboratory, the parameter $\alpha = 0.95$, and the parameter $\gamma = 0.9$. We set the population size of all algorithms equal to 100 and the dimensionality equal to 30. The algorithms run 30 times independently in the test. The stop condition for all algorithms is that the number of functions evaluated reaches a maximum of 50 dimensions.

Comparison results of the mean error at CEC2013 test functions with BA are presented in Table 1. In this table, *w* means that ASBA better than BA on *w* functions, *l* mean ASBA worse on *l* functions, *t* means that they are tied. From the results, ASBA performs better than BA on 16 functions, while BA outperforms ASBA on 12 functions.

**Table 1.** Comparison of the CEC 2013 benchmark problems

| Fun | BA | ASBA | Fun | BA | ASBA |
|---|---|---|---|---|---|
| 1 | 1.96E+00 | 1.37E+00 | 16 | 2.16E+00 | 1.98E+00 |
| 2 | 3.69E+06 | 1.07E+06 | 17 | 8.92E+02 | 4.11E+02 |
| 3 | 3.44E+08 | 3.70E+08 | 18 | 9.44E+02 | 3.50E+02 |
| 4 | 3.20E+04 | 1.43E+03 | 19 | 6.07E+01 | 2.54E+01 |
| 5 | 5.86E−01 | 3.03E−01 | 20 | 1.44E+01 | 1.42E+01 |
| 6 | 5.63E+01 | 5.00E+01 | 21 | 3.38E+02 | 3.13E+02 |
| 7 | 2.16E+02 | 1.77E+02 | 22 | 5.94E+03 | 5.88E+03 |
| 8 | 2.09E+01 | 2.09E+01 | 23 | 5.77E+03 | 5.99E+03 |
| 9 | 3.57E+01 | 3.52E+01 | 24 | 3.15E+02 | 3.22E+02 |
| 10 | 1.32E+00 | 1.12E+00 | 25 | 3.49E+02 | 3.53E+02 |
| 11 | 4.07E+02 | 5.13E+02 | 26 | 2.00E+02 | 2.30E+02 |
| 12 | 4.06E+02 | 5.35E+02 | 27 | 1.28E+03 | 1.28E+03 |
| 13 | 4.37E+02 | 4.92E+02 | 28 | 3.42E+03 | 4.22E+03 |
| 14 | 4.78E+03 | 4.82E+03 | w/t/l | 16\12\0 | |
| 15 | 4.89E+03 | 4.93E+03 | | | |

**Table 2.** Friedman test

| Algorithm | Ranking |
|---|---|
| ASBA | 1.43 |
| BA | 1.57 |

**Table 3.** Wilcoxon test

| Algorithm | *p*-values |
|---|---|
| ASBA versus BA | 0.946 |

Table 2 presents the Friedman test results, as we can seen, the ranks donate that the performance of ASBA is better than BA. Table 3 is the *p*-value achieved by Wilcoxon test; this can show that our proposal has better searchability.

## 5 Conclusion

This paper proposed an improved bat algorithm with adaptive speed strategy, which can simulate the bat in the process of search based on adaptive speed adjustment. This approach can improve the optimization efficiency and accuracy by providing a balance between exploitation and exploration. Experimental results on CEC2013 test benchmarks show that our proposal has better global searchability and a faster convergence speed, and can effectively overcome the problem convergence. Furthermore, research topics include some applications.

**Acknowledgments.** This work was supported by Beijing Key Laboratory (No: BZ0211) and Beijing Intelligent Logistics System Collaborative Innovation Center.

## References

1. Eberhart R, Kennedy J. A new optimizer using particle swarm theory. In: New York: IEEE; 1995. p. 39–43.
2. Holland JH. Genetic algorithms and the optimal allocation of trials. SIAM J Comput. 1973;2 (2):88–105.
3. Yang X. Nature-inspired metaheuristic algorithms. Luniver press, 2010.
4. Wolpert DH, Macready WG. No free lunch theorems for optimization. IEEE Trans Evol Comput. 1997;1(1):67–82.
5. Ho Y, Pepyne DL. Simple explanation of the no-free-lunch theorem and its implications. J Optim Theory Appl. 2002;115(3):549–70.
6. Dorigo M, Birattari M, Stutzle T. Ant colony optimization. IEEE Comput Intell Mag. 2006; 1(4):28–39.
7. Karaboga D, Akay B. A survey: Algorithms simulating bee swarm intelligence. Artif Intell Rev. 2009;31(1–4):61–85.
8. Whitley D. A genetic algorithm tutorial. Stat Comput. 1994;4(2):65–85.

9. Yang X, Hossein Gandomi A. Bat algorithm: A novel approach for global engineering optimization. Eng Comput. 2012;29(5):464–83.
10. Yang X. A new metaheuristic bat-inspired algorithm. In: Nature inspired cooperative strategies for optimization (NICSO 2010), 2010. p. 65–74.
11. Liang JJ, Qu BY, Suganthan PN, et al. Problem definitions and evaluation criteria for the CEC 2013 special session on real-parameter optimization. Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou, China and Nanyang Technological University, Singapore, Technical Report. 2013, 201212: 3–18.

# Marine Environment Information Collection Network Based on Double-Domain Compression Sensing

Qiuming Zhao, Hongjuan Yang, and Bo Li(✉)

Harbin Institute of Technology (Weihai), Weihai, China
libol983@hit.edu.cn

**Abstract.** This paper proposes a dual-domain compression sensing (DCS) data collection scheme. The scheme requires only through some nodes for data collection, and it uses the multi-user detection algorithm based on spatial sparse compressive sensing to perform node active state and data detection at the receiver, and then uses the sparsity of the frequency domain for information recovery, thereby further saving the control overhead of the sink node's downstream sending address frame. Through the comparison of simulation experiments, it is found that the scheme proposed in this paper is better than the previous IDMA multiple access detection schemes in terms of bandwidth while guaranteeing the reconstruction performance of the marine environment monitoring network.

**Keywords:** Compressed sensing · Underwater wireless sensor network · Multi-user detection · Multiple access

## 1 Introduction

With the development of wireless communication technology and embedded systems, the research of wireless sensor networks is promoted. In 1993, the United States Office of Naval Research (ONR) established an intelligent underwater information collection network for the physical characteristics of the marine environment and collected data through autonomous underwater vehicles [1]. In recent years, some domestic universities and research institutions have achieved certain achievements.

For land-based sensor networks, it has matured, but its network framework cannot be directly applied to underwater sensor networks. This is mainly due to the complexity and particularity of the underwater communication channel environment. Most underwater communication uses sound waves as the carrier. In view of characteristics of the underwater sensor network, in general, such as route optimization and network topology optimization, can be used to improve the efficient data collection and transmission capabilities of the network. However, the area to be monitored by the underwater sensor network is large, and it is necessary to distribute a large number of sensor nodes to achieve complete information collection. Therefore, it is difficult to fundamentally improve the system performance through routing and topology optimization methods. The fundamental goal of the underwater sensor network is to obtain


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_76

the ocean data. Therefore, how to use the structural characteristics of the original data to improve the data collection and transmission rate, reduce the network energy consumption and save bandwidth resources has become a problem that needs to be solved [2].

In summary, this paper uses compressed sensing theory to guide a new type of underwater sensor network information acquisition framework, including information acquisition, transmission, and data reconstruction process [3]. Compressive sensing theory fundamentally changes the way of the underwater sensor network obtains information, greatly improves the utilization rate of the node's collected data, reduces the overhead of the information acquisition process, and reduces the amount of data transmission and signal processing complexity of the sensor node [4, 5]. In addition, the sensor theory can be combined with information acquisition methods, multiple access, multi-user detection, and other technologies, which greatly promotes the development and innovation of marine science.

The rest of this paper is organized as follows. Section 2 introduces the system model. In Sect. 3, a marine environment information acquisition scheme based on dual-domain compressed sensing (DCS) is proposed. System simulation results are shown in Sect. 4 to help us look into the effectiveness of the algorithm. Finally, conclusions are drawn in Sect. 5.

## 2 System Model

The basis of the marine environment monitoring sensor network is the underwater wireless sensor network, which is composed of wireless sensor nodes and a sink node in the monitoring area. In this paper, the sensor nodes are deployed in the underwater two-dimensional plane, and the sink nodes are deployed in shallow seawater depths, as shown in Fig. 1.

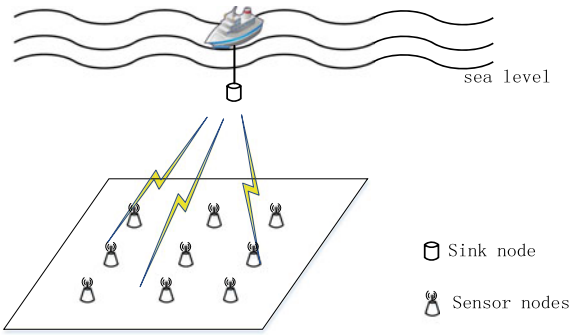



**Fig. 1.** Underwater 2D planar sensor network

In Fig. 1, the sensor nodes are evenly distributed on the two-dimensional detection plane to collect the required marine environment information and then transmitted to

the sink node through the uplink multiple access channel. Since the information acquired by the sensor nodes has certain spatial correlation, the data of some sensor nodes can be collected at the sink node according to the compressed sensing theory and then the original data can be reconstructed. Therefore, this paper mainly studies an efficient data acquisition and transmission scheme in a large ocean monitoring underwater sensor network.

## 3 Proposed Scheme

In this chapter, we use the sparseness of sensor node data in frequency domain and airspace and propose a marine environment information collection scheme based on dual-domain compressed sensing (DCS). Through acquisition probability, some sensor nodes are determined to participate in the acquisition. At the same time, the multi-user detection algorithm based on compressed sensing is used to jointly detect the active state of nodes and data at the receiver. The sink node can reconstruct the original information by successfully receiving the information from the active nodes, which greatly saves the control overhead of sending address frames in the downlink of the sink node. In the proposed scheme, the sink node receives the superimposed signals after each active node passes through the noise. Therefore, the aggregation node first needs to recover the spatial sparse signal, that is, the identity information and data of the active node through multi-user detection based on compressed sensing. Then restore the ocean information through the compressed sensing reconstruction technology. The system block diagram of DCS scheme is shown as in Fig. 2.

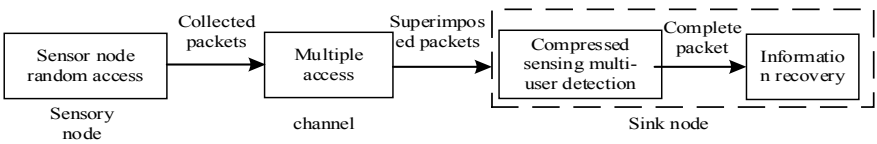



**Fig. 2.** DCS system block diagram

## 4 Simulation and Analysis

This section systematically evaluates the reconstruction error and energy consumption of the network bandwidth based on the dual-domain compressed sensing ocean environment information collection scheme and compares it with the scheme proposed in previous studies. This section uses actual measured data of oceanic currents in the Monterey Bay area of the USA. Detection location 100 m from the water, the detection range covers, the longitude [−122.8E, −122.6E], latitude [36.6N, 36.8N], and the minimum resolution is 0.01 * 0.01. This section compares the IDMA-based CS scheme proposed in [6–8] with the DCS scheme. Set the acquisition probability $p_a = 0.45$, $E_b/N_0 = 8$ dB, $RTT = 5$, under the different number of sub-frames, the

simulation curves of the minimum digital bandwidth and reconstruction error of the underwater channels required by the two schemes are shown in Fig. 3.

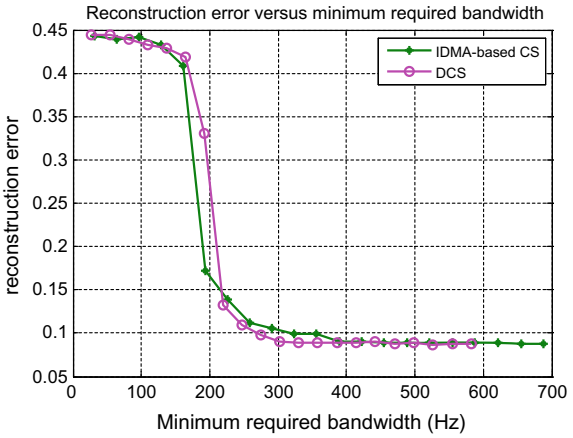



**Fig. 3.** Relation between minimum bandwidth and reconstruction error needed for underwater channel of DCS scheme

From Fig. 3, we can see that when the accuracy of ocean information recovery does not exceed 0.1, the DCS scheme requires a bandwidth of 274.6 Hz, IDMA-based CS requires a bandwidth of 323.6 Hz, and so the DCS scheme saves about 15.14% of the bandwidth resources.

## 5 Conclusions

Due to the complexity of the underwater environment, the underwater sensor networks in the ocean have the characteristics of limited energy and bandwidth. This paper studies multi-user detection based on compressed sensing theory in sporadic communication scenarios, continuously exploiting the sparseness of sensor nodes' data in frequency domain and airspace, proposes a dual-domain compressed sensing data collection scheme to construct a high-efficiency multi-point stable communication network. The simulation results show that the DCS scheme can not only ensure the reconstruction error performance of the marine environment monitoring network, but also outperforms the IDMA-based CS scheme in bandwidth performance. The design of this paper is of great significance to the research and development of marine environmental information monitoring and provides an important reference for the next generation of large-scale, three-dimensional, high-resolution marine environment monitoring sensor network construction. In the simulation of this paper, only the effects of underwater acoustic attenuation and noise are considered. Therefore, the subsequent study can further improve the underwater acoustic channel.

**Acknowledgments.** This work is supported in part by National Natural Science Foundation of China (No. 61401118, and No. 61671184), Natural Science Foundation of Shandong Province (No. ZR2018PF001 and No. ZR2014FP016), the Fundamental Research Funds for the Central Universities (No. HIT.NSRIF.201720 and HIT.NSRIF.2016100) and the Scientific Research Foundation of Harbin Institute of Technology at Weihai (No. HIT(WH)201409 and No. HIT (WH)201410).

## References

1. Chen YS, Lin YW. Mobicast routing protocol for underwater sensor networks. J Sens. 2013;13(2):2737–49.
2. Liu G, Kang W. Underwater sparse sensor network information acquisition technology based on compressed sensing. J Instrum Instrum. 2014;3(2):253–60.
3. Donoho DL. Compressed sensing. IEEE Trans Inf Theory. 2006;52(4):1289–306.
4. Liu G, Kang W. IDMA-based compressed sensing for ocean monitoring information acquisition with sensor networks. Math Probl Eng. 2014;1:1–13.
5. Fazel F, Fazel M, Stojanovic M. Random access compressed sensing for energy-efficient underwater sensor networks. IEEE J Sel Areas Commun. 2011;29(8):1660–70.
6. Li B, Du R, Kang W, Liu G. Multi-user detection for sporadic IDMA transmission based on compressed sensing. Entropy, 2017.
7. Wang B, Dai L, Mir T, et al. Joint user activity and data detection based on structured compressive sensing for NOMA. IEEE Commun Lett. 2016;20(7):1473–6.
8. Liu G, Kang W. IDMA-based compressed sensing for ocean monitoring information acquisition with sensor networks. Math Prob Eng. 2014;2014(1):1–13.

# Frequency Estimation by Two-Layered Iterative DFT with Re-Sampling Under Transient Condition

Hui Li[1(✉)] and Jiong Cao[2]

[1] College of Information Science and Technology, Hainan University, Haikou, China
hitlihuilll2@163.com

[2] Marine Communication and Network Engineering Technology Research Center of Hainan Province, Hainan University, Haikou, China
953125669@qq.com

**Abstract.** Frequency deviation incurred by sudden changes of frequency introduces harmonics and inter-harmonics in the power system, which influences the accuracy of frequency estimation with the method of discrete Fourier transform (DFT). A two-layered iterative DFT (TLI-DFT) with re-sampling was presented to measure the frequency in non-steady states. A simple frequency estimation method named exponential sampling is amended to calculate the initial sampling frequency in the inner-layered process of the DFT iteration. TLI-DFT can track the frequency in non-steady states that is contaminated by decaying direct current offsets. Mean squared error of measured frequency and rate of change of frequency indicate that the proposed algorithm is valid and more accurate than the traditional one under a transient condition in the power system.

**Keywords:** Phasor measurement · Frequency tracking · Transient condition · DFT · Exponential re-sampling

## 1 Introduction

Protecting and controlling of the smart grid require accurate and timely estimation of frequency. Measurement of frequency provides information of state estimation in the power networks around the wide area measurement system. The signal in power systems is easily to be distorted by harmonics, inter-harmonics, decaying direct current (DC) offset and to be modulated in dynamic states. Therefore, methods of frequency estimation and measurement should have the ability of frequency tracking under noisy, distorted, and distributed circumstances.

In the past several decades, researchers have put emphasis on frequency measurement and analysis, and kinds of frequency estimation methods have been reported, such as zero crossing [1, 2], least error squares [3], Newton algorithm [4], Kalman filter [5–7], prony approach [8], artificial neural networks [9], discrete Fourier transform (DFT)/fast Fourier transform [1, 10, 11], and demodulation technique [1]. DFT algorithm is the most well-known method for its capabilities of harmonics rejection and


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_77

implementation of recursion. It is preferable for its availability, understandability, and simplicity when it is implemented by advanced digital signal processing chips.

Full-cycle DFT is a kind of window-based method requiring integral samples in each cycle. If the number of samples in a window is not an integer, which is common in reality due to off-nominal components, DFT provided certain errors [12]. DFT is able to approximate the instantaneous frequency continuously, suppress the harmonics, and smooth the noise by least squares method [13] in steady state. Unfortunately, the fundamental frequency may change and the signal would contain many off-nominal components as listed above in a power system, especially the decaying DC offset, which cannot be eliminated by simple DFT. The fixed sampling rates are adopted by most data acquisition systems including the power system. It contains some drawbacks in frequency estimation and measurement.

As a uniform sampling method, the sampling period of traditional DFT is constant. A simple frequency estimation method via exponential sampling and dyadic rational was provided in [14]. Other kinds of non-uniform samplings methods, *e.g.*, long-time sampling [15], extended staggered under-sampling [16], logarithmic sampling [17], and near-optimal sampling [18] were introduced. These indicate that designing of phase-locked loop is important [19, 20]. Methods such as higher-order lags of the sample autocorrelation [21] and high-order Yule-Walker estimation [22] were utilized to calculate the sinusoidal frequency, and a variable-window-based algorithm for frequency tracking and phasor estimation was narrated in [23].

## 2 Frequency Estimation by DFT

### 2.1 Algorithm of DFT

DFT can provide the amplitude and phase angle of phasor and the frequency by way of differentiating the phase angles. Suppose the nominal voltage or current signal of a power system is

$$x(t) = A\cos(\omega t + \varphi_0) \quad (-\infty < t < \infty) \tag{1}$$

where $A$ and $\omega$ are the amplitude and angular frequency, respectively, $\omega$ is supposed to be $2\pi f_0$, $f_0$ is the nominal frequency, and $\varphi_0$ is the initial phase angle. In order to carry on DFT calculation, signal represented by (1) should be truncated and finite samples are taken by certain kinds of windows. The simple rectangular window is adopted in this paper. DFT converts these equally spaced and uniform samples into the list of a finite combination of complex sinusoids, ordered by their frequency. According to the DFT, phasor is calculated by

$$X_k = \sum_{n=0}^{N-1} x_n \cdot \mathrm{e}^{-j\frac{2\pi}{N}kn} \tag{2}$$

where $N$ is number of samples, $x(n) = A\cos(2\pi n/N + \varphi_0)$, ($n$ = 0, 1, 2, …, $N-1$) are the samples taken uniformly within the length of the rectangular window, $k$ is the order

of harmonic, and especially $k = 1$ stands for the phasor of the nominal component. Amplitudes and phase angles of $k$th harmonics are given by

$$\begin{aligned} |X_k| &= \sqrt{(X_{k\_\text{real}})^2 + (X_{k\_\text{imag}})^2} \\ \arg(X_k) &= \text{atan}(X_{k\_\text{imag}})/X_{k\_\text{real}}) \end{aligned} \tag{3}$$

where atan stands for the arctan function. $X_{k\_\text{real}}$ and $X_{k\_\text{imag}}$ are real part and imaginary part of $X_k$. According to (2) and (3), for $k = 1$, the phasor of the nominal signal is calculated by

$$\begin{aligned} X_1 &= \frac{2}{N}\sum_{n=0}^{N-1} x(n)\mathrm{e}^{-j\frac{2\pi}{N}\cdot n} \\ X_{1\_\text{real}} &= \frac{2}{N}\sum_{n=0}^{N-1} x(n)\cos\left(\frac{2\pi}{N}\cdot n\right) \\ X_{1\_\text{imag}} &= -\frac{2}{N}\sum_{n=0}^{N-1} x(n)\sin\left(\frac{2\pi}{N}\cdot n\right) \end{aligned} \tag{4}$$

where $X_{1\_\text{real}}$ is the real part of the phasor of nominal component, and $X_{1\_\text{imag}}$ is the imaginary part. The multiplier $2/N$ in front of the summation sign is used to generate the normalized amplitude of DFT, i.e., 1 p.u. The phasor of nominal signal is expressed as

$$X_1 = A\mathrm{e}^{j\varphi_0} \tag{5}$$

## 2.2 Frequency Estimation by DFT

The synchronous phasor (synchrophasor) is defined as a complex number, representing the fundamental frequency component of a voltage or current, with an accompanying time-tag defining the time instant for which the phasor measurement is performed [24]. Serial synchrophasors can be calculated by shifting windows at each sampling time $t_r = rT_s$, $r = 0, 1, 2, \ldots, +\infty$ where $T_s$ is the sampling interval, $T_s = 1/f_s$. A synchrophasor at time $t_r$ is represented as

$$\begin{aligned} X_1^r &= \frac{2}{N}\sum_{n=0}^{N-1} x(n)\mathrm{e}^{-j\frac{2\pi}{N}\cdot n}\cdot \mathrm{e}^{j\frac{2\pi}{N}\cdot r} \\ &= A\mathrm{e}^{j(\varphi_0 + 2\pi r/N)} = X_1\mathrm{e}^{j2\pi r/N}, \quad r = 0, 1, 2, \ldots, +\infty \end{aligned} \tag{6}$$

Amplitudes of serial synchrophasor at steady state are $A$, if the signal is a nominal one. The phase angles of serial synchrophasor are $\{\varphi_0,\ \varphi_0 + 2\pi/N,\ \varphi_0 + 4\pi/N, \ldots,\ \varphi_0 + 2\pi r/N, \ldots\}$. In the following, we use $X^r$ as the nominal phasor instead of $X_1^r$ for

simplicity. Frequency is defined as the speed of rotation of a phasor. Then for two consecutive measured phases, the frequency is calculated by

$$f^r = (\varphi^{r+1} - \varphi^r)/(2\pi T_s) \tag{7}$$

where $\varphi^r$ is the phase angle calculated by DFT. If $\varphi^r$ is not the nominal one, $f^r$ would be inaccurate.

## 3 Iterative DFT with Re-sampling

### 3.1 Inner Layer DFT Iteration by Re-sampling

From the analysis above, the measured frequency $f_{\text{meas}}$ would approach $f_0 + \Delta f$ more and more closely if we change the sampling frequency $f_s = Nf_{\text{meas}}$ consecutively. Especially when the sampling frequency becomes close enough to $N(f_0 + \Delta f)$, DFT calculated iteratively would give a measured frequency $f_{\text{meas}}$ as the input off-nominal frequency $f_0 + \Delta f$ exactly. It means that there are integral samples in one cycle of the input frequency again. This process of iterative DFT algorithm within one cycle is shown in Fig. 1.

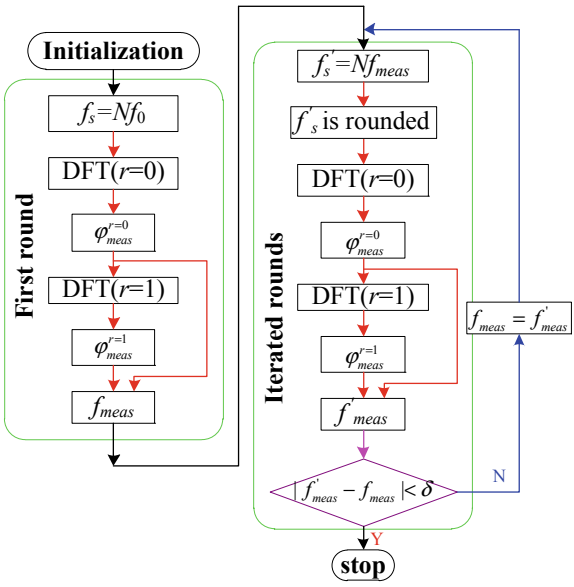



**Fig. 1.** Process of iterative DFT in one cycle

In the first round of DFT calculation, the sampling frequency $f_s$ is set to be $Nf_0$ and two phasors are calculated to get the frequency $f_{\text{meas}}$. Then in the following rounds, new $f'_{\text{meas}}$ is gotten according to the sampling frequency $f'_s = Nf_{\text{meas}}$, until the difference

of two successive frequencies $f'_{\text{meas}}$ and $f_{\text{meas}}$ is less than a threshold $\delta$. The samples are taken by a rectangular overlapping shifting window at time $t_r = rT_s$ ($r = 0$ and 1).

### 3.2 Outer Layer DFT Iteration Between Cycles

#### 3.2.1 Determination of Initial Frequency by Amended Exponential Sampling

Exponential sampling is a kind of simple frequency estimation algorithm, which can simplify the process of sampling by exponential sampling and needs only few samples distributed exponentially along the time [14]. But in the dynamic states or under the circumstance of low signal to noise ratio (SNR), the input frequency may change in every cycle. It introduces lots of harmonics and spectrum leakage and also limits the application of exponential sampling in such situation. Frequency is estimated based on a modified exponential sampling method, which is amended to be used in unsteady states. For traditional exponential sampling, the sample is taken exponentially at

$$t_p = 2^{p-Q-1}\ (s), \quad p = 1, 2, \ \ldots, \ P \tag{8}$$

where $P$ defines the bit accuracy, and $Q$ defines the maximal frequency $f_{\max} = 2^Q$. The samples are

$$x(t_p) = A\cos(2\pi f t_p + \varphi_0) \tag{9}$$

where $f$ is the instantaneous frequency at $t_p$ in the dynamic state which is not a constant anymore. According to [14], the signal is a sinusoidal signal. If the cosine signal is used in accordance with (9), we have

$$s(t_p) = \pm\sqrt{1 - x(t_p)^2} \tag{10}$$

where symbols of "+" and "−" are taken according to the quadrants of phases of the signal. To guarantee the frequency to be an accuracy of $2^{Q-P}$ Hz, we get

$$f_{\text{meas}} = f_{\max}\sum_{p=1}^{P} b_p 2^{-p} \tag{11}$$

where

$$\begin{cases} b_p = 0, & \text{if } s(t_p) > 0 \\ b_p = 1, & \text{if } s(t_p) < 0 \end{cases} \tag{12}$$

If $s(t_p) = 0$ for some $p = p_0 \leq P$, $b_{p_0} = 1$, (11) is rewritten as

$$f_{\text{meas}} = f_{\max}\sum_{p=1}^{p_0} b_p 2^{-p} \tag{13}$$

where $p_0$ is a terminator of exponential sampling calculation. For example, a power system with a nominal frequency of 60 Hz, and the dynamic frequency range is [−5, +5] Hz. It indicates that off-nominal frequency may be 65 Hz > 64 = $2^6$ Hz. Hence, we set $Q = 7$, and $P = 7$ for an accuracy of 1 Hz. We get $x(t_p) = \{-0.9809, 0.9239, 0.7071, 0, -1, 1, 1\}$, and suppose $\varphi_0 = 0$, $s(t_p) = \{0.1951, -0.3827, -0.7071, -1, 0, 0, 0\}$, and $b = \{0, 1, 1, 1, 1, 0, 0\}$ where $p_0 = 5$ without noise.

Two factors influence the accuracy or even correctness of frequency estimation by exponential sampling: First of all, if at the first sampling time $t_1 = 1/128$ (s), the system frequency is suddenly changed to 65 Hz, we have $s(t_1) = -0.0491$, which means $b_1 = 1$ and introduces frequency estimation error of $2^{Q-1} = 64$ Hz; secondly, if the SNR is not big enough, values of samples may change from negative to positive (*e.g.*, $b_2 = -0.3827$ may change to a positive value because of noise), which introduces frequency error of $2^{Q-2} = 32$ Hz. So in the algorithm of amended exponential sampling (AES), we set $b_0 = 0$ and $b_1 = 1$ constantly for $P = Q = 7$.

### 3.2.2 Frequency Tracking Cycle by Cycle

Flowchart of DFT iteration is shown in Fig. 2, where $L$ is the total number of cycles to be generated and $f^l_{\text{meas}}$, $l = 1, 2, \ldots, L$ are the measured frequency in the $l$th cycle by inner layer calculation. In the following, we named the algorithm as "TLI-DFT aidedc by AES." In Fig. 2, inner layer iteration processes are implemented by iterative DFT in one cycle as shown in Fig. 1.

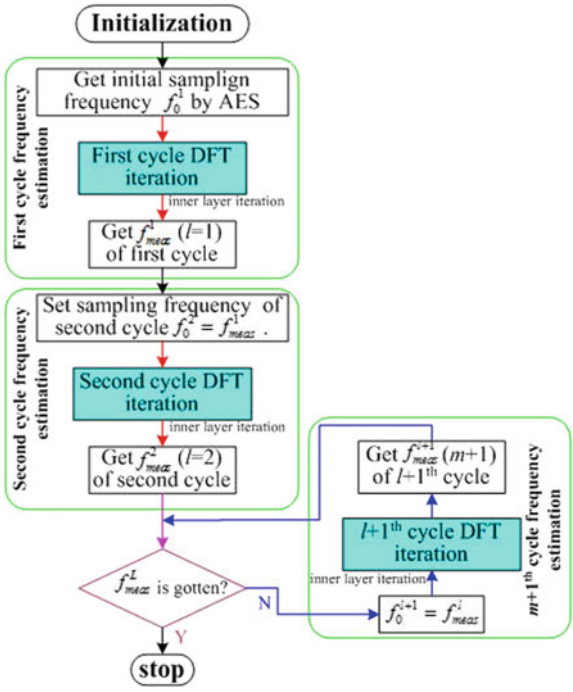



**Fig. 2.** Iteration process of TLI-DFT aided by AES

### 3.2.3 Rate of Change of Amplitude and Frequency

The rate of change of frequency (ROCOF) is calculated by taking the time-derivative of the estimated frequency numerically

$$\mathrm{ROCOF}(r) = (f^{r+1} - f^{r})/T_s \tag{14}$$

## 4 Simulation Results and Data Analysis

Traditionally, an adaptive sampling algorithm with varied sample interval $T_s$ is adopted. It is based on a feedback system until sampling frequency $f_s$ equals to the frequency of the incoming signal, provided that the incoming signal does not fluctuate in frequency [25]. New algorithm is compared with the traditional one. Parameters used in simulation are listed in Table 1.

**Table 1.** Parameters used in simulation

| Symbol | Quantity | Value | Unit |
|---|---|---|---|
| $f_0$ | Fundamental frequency | 60 | Hz |
| $\omega$ | Angular frequency of fundamental signal | – | radian |
| $A$ | Amplitude of fundamental signal | 1 | p.u. |
| $\varphi_0$ | Initial phase angle | 0 | degree |
| $N$ | Number of samples per cycle | 12 | – |
| $f_s$ | Sampling frequency | – | Hz |
| $T_s$ | Sampling interval | – | second |
| $L$ | Length of signal (in cycles) | 1000 | cycles |
| $T$ | Length of one cycle | 0.0167 | second |
| $\Delta f$ | Shifting frequency off the nominal one | [−5, 5] | Hz |
| SNR | Signal to noise ratio | [10, 40] | dB |

Weighted mean value of frequency $\overline{f}_r$ in (15) is adopted by iteration process. Several more complicated scenarios are considered to show the performance of these algorithms in the following text.

$$\overline{f}_r = 0.6 \times f_{r+2} + 0.3 \times f_{r+1} + 0.1 \times f_r \ (r \geq 0) \tag{15}$$

When a fault or a disturbance occurs, the current signal consists of exponentially decaying direct current (DC) offsets in electrical power system. The decaying rates depend on the time constants determined by the inductive reactance to resistance ratio (X/R ratio) of the system. The larger the X/R ratio, the slower the DC component decays. Signal with the nominal component and decaying DC offsets is represented as

$$x(t) = A\cos(2\pi f_0 t + \varphi_0) - \sum_{i=1}^{N_{\mathrm{DC}}} I_i \mathrm{e}^{-t/\tau_i} + N_{\mathrm{noise}} \; (0 \le t \le T_{\mathrm{DC}}) \tag{16}$$

where $N_{\mathrm{DC}}$ is the number of DC offset components, $I_i$ and $\tau_i$ are the amplitude and time constant, respectively, of the $i$th DC offset component. $T_{\mathrm{DC}}$ is the effecting period of decaying DC offsets, since the decaying DC offsets last only in several cycles and then disappear. DC offset is a non-periodic signal, and its frequency spectrum encompasses all the frequencies and cannot be removed by anti-aliasing low pass filter.

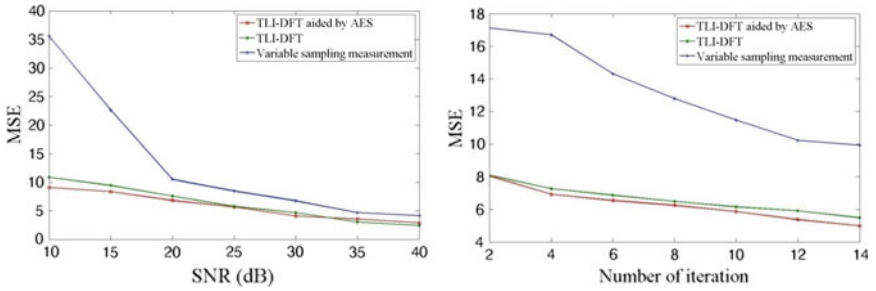
A digital mimic filter was proposed to suppress the effect of decaying DC offsets over a broad range of time constants [26, 27]. This kind of filter needs exact values of the time constants for eliminating DC offsets. It is usually impractical in power system. Kalman filter also needs the time constants exactly to obtain a good performance of filtering [5–7, 28]. DFT-based techniques are generally used for removal decaying DC offset from phasor estimation. On the other hand, the decaying DC component affects the accuracy of the DFT algorithm greatly. Different windows have been suggested, and half-cycle DFT was used to get the phasor in the case of decaying DC offsets. Full-cycle DFT is still the most widely adopted one. Suppose there are two decaying DC offset components (i.e., $N_{\mathrm{DC}} = 2$). Their parameters are listed in Table 2. Speed of decaying of DC offset 1 is faster than that of DC offset 2. But the absolute value of amplitude of DC offset 2 is smaller.

**Table 2.** Parameters of decaying DC offset components

| Number | Amplitude $I_i$ (p.u.) | Time constant $\tau_i$ (cycle) |
|---|---|---|
| 1 | −1 | 0.5 |
| 2 | −0.125 | 5 |

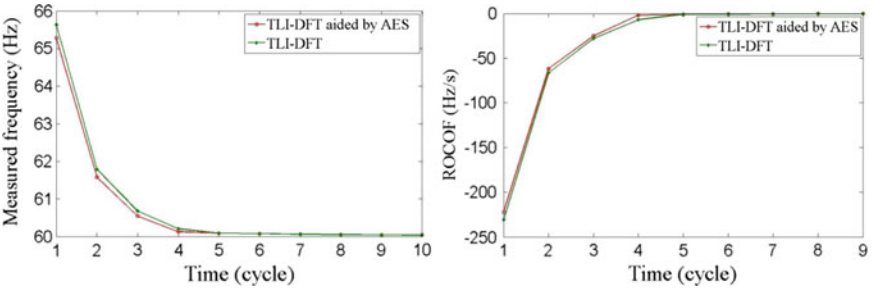
In Fig. 3, SNR = 20 dB is about the point of inflexion, although curves of MSE are not as smooth as desirable owing to the limitation of number of estimated frequencies (i.e., $T_{\mathrm{DC}}$ = 10 cycle). MSE of our proposed algorithms would not obtain much more gain when the number of iterations is more than 3 or 4 in Fig. 3b.

Frequency tracking and ROCOF by TLI-DFT are plotted in Fig. 4. Performance of algorithm TLI-DFT aided by AES is almost the same as that of TLI-DFT, because AES is used only once in the initial frequency estimation of inner layer iteration in the first cycle.




(a) MSE vs. SNR (Iteration process stops at 3) (b) MSE vs. Number of iteration (SNR=20dB)

**Fig. 3.** Performance of three algorithms with decaying DC offsets (considering the time constants of two decaying DC offsets. $T_{DC}$ = 10 cycle)




(a) Frequency measured cycle-by-cycle (b) ROCOF calculated cycle-by-cycle

**Fig. 4.** Frequency tracking and ROCOF of TLI-DFT with decaying DC offsets

## 5 Conclusions

A new approach of two-layered iterative DFT is proposed to track the input frequency in dynamic states. Inner-layered DFT iteration is adopted in every cycle, and frequency estimated in the previous cycle is used as the initial frequency in the following one. And in a simple and fast method, exponential sampling is amended to adapt the non-static states. This algorithm is more accurate than the traditional one. New algorithms are tested with different simulation parameters, such as different SNR and maximal numbers of iteration predefined to stop the iteration process.

In practical, variable sampling algorithms are highly connected with phase-locked loop (PLL) of the frequency generator and the digital signal processor. Multiple-rated structure of adaptive PLL and time-varying PLL-based sampling methods is the main work for implementation of the proposed TLI-DFT algorithm.

**Acknowledgments.** This work was supported by Hainan Provincial Key R. & D. Projects of China (ZDYF2016010 and ZDYF2018012) and the National Natural Science Foundation of China (No. 61661018).

## References


1. Begovic MM, Djuric PM, Dunlap S, Phadke AG. Frequency tracking in power networks in the presence of harmonics. IEEE Trans Power Del. 1993;8(2):480–6.
2. Nguyen CT, Srinivasan K. A new technique for rapid tracking of frequency deviation based on level crossings. IEEE Trans Power App Syst. 1984;103(8):2230–6.
3. Sachdev MS, Giray MM. A least error squares technique for determining power system frequency. IEEE Trans. Power App Syst. 1985;104(2):437–444.
4. Terzija VV, Djuric MB, Kovacevic BD. Voltage phasor and local system frequency estimation using Newton type algorithm. IEEE Trans Power Del. 1994;9(3):1368–74.
5. Wood HC, Johnson NG, Sachdev MS (1985) Kalman filtering applied to power system measurements for relaying. IEEE Trans Power App Syst. 1985;104(12):3565–73.
6. Routray A, Pradhan AK, Rao KP. A novel Kalman filter for frequency estimation of distorted signals in power systems. IEEE Trans Instrum Meas. 2002;51(3):469–79.
7. Siavashi EM, Afshania S, Bina MT, Zadeh MK, Baradar MR (2009) Frequency estimation of distorted signals in power systems using particle extended Kalman filter. In: 2nd international conference PEITS; 2019. p. 174–8.
8. Lobos T, Rezmer J. Real-time determination of power system frequency. IEEE Trans Instrum Meas. 1997;46(4):877–81.
9. Vianello R, Prates MO, Duque CA, Cequeira AS, da Silveira PM, Ribeiro PE. New phasor estimator in the presence of harmonics, DC-offset and interharmonics. In: 14th ICHQP; 2010. p. 1–5.
10. Yang JZ, Liu CW A new family of measurement technique for tracking voltage phasor, local system frequency, harmonics and DC offset. In: IEEE PES summer meeting; 2010. p. 1327–32.
11. Zeng B, Teng ZS, Cai YL, Guo SY, Qing BY. Harmonic phasor analysis based on improved FFT algorithm. IEEE Trans Smart Grid. 2011;2(1):51–9.
12. Karimi-Ghartemani M, Ooi B, Bakhshai A. Investigation of DFT-based phasor measurement algorithm. In: IEEE PES General Meeting; 2010. p. 1–6.
13. Phadke AG, Thorp JS, Adamiak MG (1983) A new measurement technique for tracking voltage phasor, local system frequency and rate of change of frequency. IEEE Trans Power App Syst. 1983;102(5):1025–38.
14. Kay S. Simple frequency estimation via exponential samples. IEEE Signal Process Lett. 1994;1(5):73–5.
15. Olkkonen H, Olkkonen JT. Log-time sampling of signals: Zeta transform. Open J Discrete Math. 2011;1(2):62–5.
16. Sadinezhad I, Agelidis VG. Extended staggered undersampling synchrophasor estimation technique for wide area measurement systems. In: IEEE PES ISGT; 2011. p. 1–7.
17. Rusu C, Kuosmanen P. Phase approximation by logarithmic sampling of gain. IEEE Trans Circuits Syst II Analog Digit Signal Process. 2003;50(2):93–101.
18. Trittle S, Hamprecht FA. Near optimum sampling design and an efficient algorithm for single tone frequency estimation. Digit Signal Process. 2009;19(4):628–39.
19. Yen CS. Phase-locked sampling instruments. IEEE Trans Instrum Meas. 1965;14(1/2):64–8.
20. Karimi H, Karimi-Ghartemani M, Iravani MR. Estimation of frequency and its rate of change for applications in power systems. IEEE Trans Power Del. 2004;19(2):472–80.
21. Elasmi-Ksibi R, Besbes H, Lopez-Valcarce R, Cherif S. Frequency estimation of real-valued single-tone in colored noise using multiple autocorrelation lags. Signal Process. 2010;90 (7):2303–7.

22. Stoica P, Moses RL, Soderstrom T, Li J. Optimal high-order Yule-Walker estimation of sinusoidal frequencies. IEEE Trans Signal Process. 1991;39(6):1360–8.
23. Hart D, Novosel D, Hu Y, Smith B, Egolf M. A new frequency tracking and phasor estimation algorithm for generator protection. IEEE Trans Power Del. 1997;12(3):1064–73.
24. Gomes S, Martins N, Stankovic A. Improved controller design using new dynamic phasor models of SVC's suitable for high frequency analysis. In: IEEE PES transmission and distribution conference and exposition; 2006. p. 1436–44.
25. Benmouyal G. An adaptive sampling-interval generator for digital relaying. IEEE Trans Power Del. 1989;4(3):1602–9.
26. Benmouyal G. Removal of DC-offset in current waveforms using digital mimic filtering. IEEE Trans Power Del. 1995;10(2):621–30.
27. Yu CS. A discrete Fourier transform-based adaptive mimic phasor estimator for distance relaying applications. IEEE Trans Power Del. 2006;21(4):1839–46.
28. Girgis AA, Brown RG. Application of Kalman filtering in computer relaying. IEEE Trans. App. Syst. 1981;100(7): 3387–97.

# Dimension Reduction Based on Effects of Experienced Users in Recommender Systems

Bo Chen[1(✉)], Xiaoqian Lu[2], and Jian He[3]

[1] School of Information and Software Engineering, University of Electronic Science & Technology of China, Chengdu 610054, China
bochen@uestc.edu.cn

[2] Information Center, University of Electronic Science & Technology of China, Chengdu 610054, China
xqlu@uestc.edu.cn

[3] School of Automation Engineering, University of Electronic Science & Technology of China, Chengdu 610054, China
hejian@uestc.edu.cn

**Abstract.** The paradox of huge volume with high sparsity of rating data in collaborative filtering (CF) system motivates the present paper to utilize information underlying sparsity to reduce the dimensionality of data. This difference in user experiences resembles factor underlying widely used term frequency weighting scheme in information retrieval. Hypothesis of Rational Authorities Bias (H-RAB) is proposed, supposing that higher prediction accuracy can be attained to emphasize referential users with higher experiences. Dimension reduction suggests pruning all referential users with less experience than a given maturity threshold. Empirical results from a series of experiments on three major public available CF datasets justify the soundness of both modifications and validity of H-RAB. A few open issues are also proposed for future efforts.

**Keywords:** Collaborative filtering recommender · Sparsity · Rater maturity · Rational Authorities Bias · Dimension reduction

## 1 Introduction

The present paper concerns with two of the major problems encountered by memory-based CF algorithms in the case of annotation-in-context:

(1) Prediction Accuracy

For annotation-in-context recommender, prediction accuracy is the quality of how precisely the predicted ratings for items match the true ratings. A well-known weighting scheme in IR community, term frequency—inverse document frequency (TF-IDF) [1], is a source of inspiration for improving the prediction accuracy. Inverse user frequency (IUF) weighting scheme is the counterpart of IDF in CF context, proposed in [2], leading to fairly good improvements.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_78

While mapping the TF part from IR to CF context, the frequency of a specific term present in documents transforms to item amount a user rated. This factor is generally known as user experience. The term "user experience" has two different meanings in human–computer interaction: (a) user knowledge and familiarity, built-up through a series of interactions; (b) user's psychological activity such as feelings and impressions while using the system, which is the case in McLaughlin and Herlocker [3]. To avoid confusion, in the present paper, "rater maturity" is used to represent above factor, bearing the first meaning of "user experience." The motivation of utilizing rater maturity comes from following phenomena: As the number of rated items increases for a given user, the user is more experienced in using the CF system, more familiar with subtle intended meaning of rating scales, and with more background knowledge of items. Then, the users with more frequent rating actions would have more confidence in the rating scores they gave. This leads to an intuition to give higher weight to votes from more experienced users while predicting ratings. The present paper is intended to empirically justify a counterpart scheme of TF in CF context, viz. to take account of rater maturity, so as to achieve higher accuracy than standard memory-based algorithms.

(2) Scalability for Huge Volume of Data with High Sparsity

In realistic CF applications, it is common for there to be tens of thousands of users voting on several thousands of items, which make two aspects of a paradox: The existing data volume is huge, usually on a million scale; on the other hand, the number of rated items rated by a single user is quite small, leading to very sparse rating matrix.

In memory-based CF systems, the response time is crucial, for the user is waiting online, whereas huge data volume makes it hard to achieve. This calls for dimension reduction techniques in multivariate data analysis, such as principal component analysis (PCA) and singular value decomposition (SVD). PCA is applied in Eigentaste to implement a constant time model-based CF algorithm in [4]. SVD is used to implement Latent Semantic Indexing (LSI) [5, 6], which is influential in IR communities. LSI/SVD is used to solve dimension reduction in CF-based systems in Billsus and Pazzani [7].

A syndrome of high sparsity is that only a few commonly rated items are present when evaluating the similarity between two users. An extreme example is: When there is only one item co-rated by user1 and user2 with the same rating, the system may conclude that user1 and user2 are most similar. Significance weighting [8] makes use of implicit information underlying the sparsity to alleviate above negative influence; in this case, the information is the number of commonly rated items. This motivates the authors of the current paper to consider a way to take full advantage of implicit information underlying sparsity to achieve both high accuracy and performance. Apparently, the sparsity is a natural result when most of the users rated only a few items, viz. most of the users has low rater maturity. The dimension reduction methods based on PCA and SVD both involve latent variables/components, which are not explicitly explained with concrete features of data. The present paper aims to utilize the rater maturity to explicitly reduce the dimensionality of dataset.

## 2 Rater Maturity and Dimension Reduction

It is not uncommon, in using a CF-based recommender system, such as MovieLens, that the users are required to give sample ratings to a small number of movies when they register and log on for the first time (the number is 15 for MovieLens). This is a way to solve cold start problem. Then, appropriate recommendations can be made based on similarities of their sample ratings to existing ratings from other users. This measure is not pleasant for newcomers, for a great deal of patience is needed for them to browse tens of pages to find movies they ever watched in the randomly given film lists, not to say to further choose from the watched ones 15 times to give ratings, in a totally unacquainted system. This may lead to a rush and restless action by arbitrarily voting on any movies with random ratings, even if they never knew about them.

Despite the widely existed reluctance, in real CF systems, there are still some users who contribute to substantial portion of the total votes. As the number of votes from one rater increases, the user becomes more and more familiar with operations on the specific CF system and the subtle meaning of rating scales and builds up considerable background knowledge on the items being rated, all of which give rise to rater maturity with respect to the system. And more mature, the user has more confidence in their ratings to items.

Since CF mechanism is based on the premise that rating correlation reflects user similarity in personality or interest, if the rating is given in an unconfident status, the deviation of final similarity may be distorted to an unexpected extent. User experience is quantified as simple rater maturity, SRM.

Three public available CF datasets are frequently used in the literature for experiments, including EachMovie, MovieLens, and Jester. Figure 1 illustrates global distributions of SRM in above datasets.

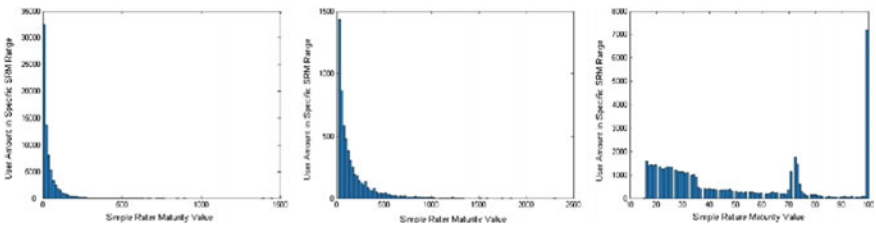


**Fig. 1.** Histograms for SRM distributions in the three datasets (EachMovie, MovieLens, Jester)

We aim to empirically validate *Hypothesis of Rational Authorities Bias* (H-RAB): Due to familiarities with the system behavior and strong background knowledge, it is more credible and rational to emphasize ratings from more mature referential users in making prediction for active users. *RAB-aware dimension reduction* (RAB-DR) is defined as: For a specified maturity threshold $(\theta)$, the users are classified as mature and immature. The prediction takes account of votes from only a subset of the $n$ referential users to $m$ items. It is equivalent to prune votes from immature raters; hence, the dimensionality of the original $n$-*by*-$m$ rating matrix is reduced to $n_1$-*by*-$m$, where $n_1$ is

the number of mature users with respect to $\theta$. The set of mature raters in the reduced matrix is referred to as a top rater selection (TRS):

$$p_{a,k} = \bar{r}_a + \kappa \sum_{u_i \in \text{TRS}} w(a,i)(r_{ik} - \bar{r}_i) \tag{1}$$

# 3 Experiments

## 3.1 General Experiment Settings

### 3.1.1 Test Bed

A public available MATLAB implementation for CF methods, C/MATLAB toolkit contributed by *Guy Lebanon*, is modified to incorporate RAB-aware designs proposed in former sections. Since C/MATLAB toolkit is a well-known implementation in CF community, the original results for memory-based algorithms can be credibly used as a baseline to identify improvements brought by RAB-aware modifications.

### 3.1.2 Baselines and Success Criteria

For the original CF methods, based on the type of similarity, there are two generic methods, called Generic PCC (G/PCC) and VSS (G/VSS) in the present paper. Effectiveness of above two generic CF methods is evaluated in original C/MATLAB by comparing with non-CF-based approaches, such as average prediction (AP). AP makes prediction with the average value of existing ratings from current active user, without taking account of any information from referential users. Improvement ratio for values of mean absolute error (MAE) brought by generic CF methods with respect to AP can be regarded as a baseline ratio for evaluating salience of subsequent modifications.

For RAB-DR, CF algorithms are expected to benefit from it on two perspectives, to achieve higher prediction quality in less processing time. To be rigorous, success criteria for RAB-DR are required to achieve salient progresses on both accuracy and efficiency. However, it should be reminded that even if only marginal accuracy improvement is attained, RAB-DR is still desirable when processing time is reduced substantially.

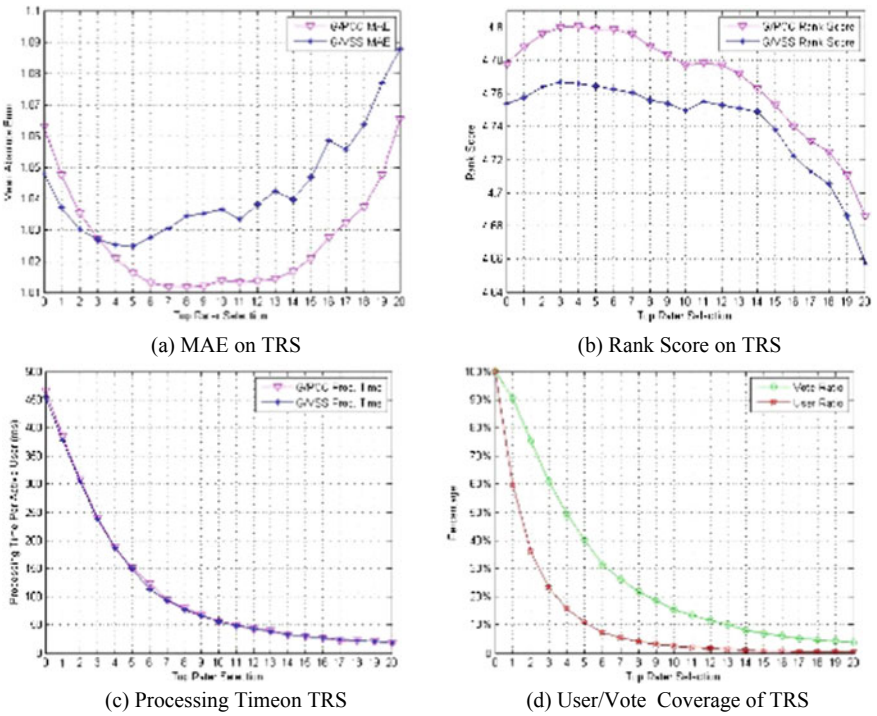
As a metric used in original version of C/MATLAB, the half-life utility metric, rank score, is also calculated and analyzed in the experiments. Since rank score focuses on special systems where the user behavior is characterized by the half-life pattern, it is sensitive to operational context and not used as widely as MAE to evaluate quality of CF algorithms [9]. It is desirable for RAB-aware version of CF algorithms to improve rank score, which is not adopted as an essential requirement for the validity of H-RAB. On the other hand, the observed properties presented by rank score in experiments may give rise to new issues that deserve future efforts.

In each test, the users in subset of datasets are split into 50% active users and 50% referential users. For active users, *GivenHalf* protocol is used to withhold half of each active user's existing votes, and the other half is used to calculate similarity, viz. PCC and VSS with referential users. Then predicted ratings are compared to the withheld

votes to evaluate metrics. For rank score metric, there are two more arguments to set: coefficient $\alpha$ as 0.5 and half-life parameter $k$ as 10. This valuation would not affect validity/invalidity of H-RAB in terms of MAE.

### 3.2 Results

In the stepwise exploration of RAB-DR, performances of G/PCC and G/VSS on each of the first 21 TRS are shown in Fig. 2. It should be reminded that TRS_0 stands for the entire set of 5000 referential users, performance on which is the baseline calculated in unmodified algorithms in original C/MATLAB.




(a) MAE on TRS

(b) Rank Score on TRS

(c) Processing Timeon TRS

(d) User/Vote Coverage of TRS

**Fig. 2.** Positive test of RAB-DR on 10,000-user EachMovie

As expected, MAE for G/PCC in subfigure a of Fig. 2 decreases for the first nine TRS (TRS_0 to TRS_8) and then increases but it does not exceed the baseline value until the last TRS in the test. Upon baseline presented on TRS_0, it means any TRS in the range of TRS_1 to TRS_19 can achieve no worse MAE. Such range is referred to as favorable range. MAE for G/VSS does not benefit from RAB-DR so much as G/PCC, where TRS_1 to TRS_15 is the favorable range.

To highlight the improvement brought by RAB-DR, from the stepwise exploration, a fairly good TRS (FG_TRS) can be identified in the favorable range, on which there is the lowest MAE. The SRM for the lower boundary of FG_TRS is the approximate

value for maturity threshold. As in this case, for G/PCC, FG_TRS is TRS_8, with approximation for MT = 160; for G/VSS, FG_TRS is TRS_5, with approximation for MT = 100. The concrete improvements are shown in Table 1. With one-sixth of original processing time, G/PCC on TRS_8 achieves 8.44% improvement ratio on MAE. The accuracy further attained 4.63% (8.44−3.81%) by RAB-DR is salient enough compared to the baseline ratio 3.81%. With one-third of processing time, G/VSS achieves higher prediction accuracy too. The effectiveness of RAB-DR is far more impressive than that of Naïve RAB-WS, where the latter is regarded as only fine tuning to original algorithms.

**Table 1.** Improvements on FG_TRS for RAB-DR on EachMovie

| Method | MAE | | Rank score | | Processing time (PAU) | |
|---|---|---|---|---|---|---|
| | Value | Imp. ratio | Value | Imp. ratio | Value (ms) | Imp. ratio |
| AP | 1.1050 | Baseline | 4.2684 | Baseline | – | – |
| G/PCC on TRS_0 | 1.0629 | 3.81% | 4.7775 | 11.93% | 466.11 | Baseline |
| G/PCC on TRS_8 | 1.0117 | 8.44% | 4.7886 | 12.19% | 78.95 | 83.06% |
| G/VSS on TRS_0 | 1.0480 | 5.16% | 4.7539 | 11.37% | 453.5 | Baseline |
| G/VSS on TRS_5 | 1.0246 | 7.28% | 4.7645 | 11.62% | 150.37 | 66.84% |

Rank score for G/PCC in subfigure (b) increases for the first five TRS, and it is not worse than the baseline value until TRS_10. Similar desirable pattern persists for rank score of G/VSS.

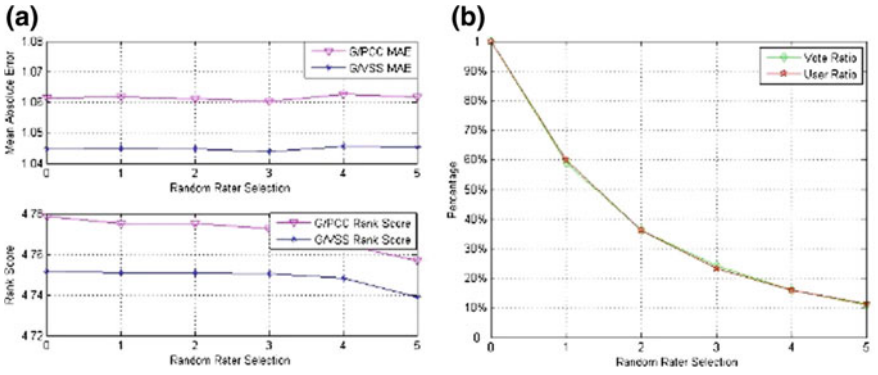



**Fig. 3** Negative test of RAB-DR on 10,000-user EachMovie. **a** MAE/Rank score on RRS. **b** User/Vote coverage of RRS

In Fig. 3, results for negative tests on RRS corresponding to user percentage of first six TRS are visualized. The fluctuation of MAE does not show the pattern presented in RAB-DR (nor does rank score). This denies the false hypothesis that the accuracy comes from purely decreasing coverage of referential users.

Comparing the above three results on different datasets, the accuracy further achieved is not as good as the other two datasets. A reasonable interpretation is that in a more mature system where more mature raters present, the baseline prediction is more accurate with generic methods, leading to less margin for subsequent improvement. It should not be interpreted as a simple consequence of average SRM, since this measurement for Jester is 48.62, which is similar to that of EachMovie, 45.90. Normalized system maturity (NSM), proposed in Definition 3, is more applicable.

To compare prediction accuracy on different datasets with drastically dissimilar rating scales, the normalized mean absolute error (NMAE) proposed in Goldberg et al. [4] can be used. The comparison in Table 2 shows, for datasets with higher normalized system maturity, the initial accuracy (e.g., G/PCC on TRS_0) is better. This is one of the causes lying behind the lower improvement ratio of Jester.

**Table 2.** Influence of NSM to prediction accuracy

| | EachMovie | MovieLens | Jester |
|---|---|---|---|
| $r_{max}$–$r_{min}$ | 5 | 4 | 20 |
| NSM | 0.0229 | 0.0419 | 0.4862 |
| NMAE of AP | 0.2127 | 0.2020 | 0.1933 |
| NMAE of G/PCC on TRS_0 | 0.2060 | 0.1994 | 0.1803 |
| NMAE of FG_TRS | 0.1969 | 0.1949 | 0.1763 |

## 4 Concluding Remarks

The overwhelming data volume in CF-based recommenders has been a bottleneck for memory-based CF algorithms to be efficient. Based on the empirical results in last section, there is a simple but nontrivial idea to alleviate the problem: Since no worse accuracy can be achieved on a subset of referential users, RAB-DR is reasonable to be a standard preprocessing step to reduce the actual referential data used online, not to mention that the experimental results also display salient progresses for accuracy.

As emphasized, the efforts of the present paper are in an explorative manner, to validate the existence of *Rational Authorities Bias* in various datasets. Some detailed deviations from expectations in empirical results and open issues are left for future investigations, including:

- Differences and their causes in RAB sensibility for methods on PCC and VSS.
- Differences and their causes in RAB sensibility for different evaluation metrics.
- More effective heuristics or methods, compared to the stepwise exploration in this paper, to determine optimal value for maturity threshold for RAB-DR.
- Definition of more intricate maturity measures than SRM and more effective modifications.
- Applying RAB-aware modifications to systems with implicit rating, where the ratings are not given explicitly, represented by user interaction with the system, such as browsing, ordering, or canceling an item.

- Applying RAB-aware modifications for model-based CF algorithms, e.g., to train the model with data from TRS.
- Theoretically justification H-RAB and identification of relationships between explicit RAB-aware approach with latent factors.

One precaution is mentioned as the last concluding remark. Since the item-based CF algorithms are proposed in Sarwar et al. [10], the symmetry of user and item is widely used to establish new CF methods. The counterpart of rater maturity centered on items is the item popularity. Matrix composed of more popular items would surely have higher density, but it is not preferred to prune items in a brute force way, since it would lead to undesirable bias to ignore less rated items, eventually decrease coverage of recommended items.

**Acknowledgments.** This work is partly funded by UESTC Fundamental Research Funds for Central Universities (ZYGX2015J069), Grant: JSEB-201303 of e-Commerce Key Laboratory of Jiangsu Province, and Grant 2017GZYZF0014 of Sichuan Sci-Tech Plan.

## References

1. Sparck Jones K. Index term weighting. Inf Storage Retrieval. 1973;9:619–33.
2. Breese JS, Heckerman D, Kadie C. Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the 14th conference on uncertainty in artificial intelligence; 1998. p. 43–52.
3. McLaughlin MR, Herlocker JL. A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In: Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval (Sheffield, United Kingdom, July 25 – 29, 2004). SIGIR'04; New York: ACM Press; 2004. p.329–36.
4. Goldberg K, Roeder T, Gupta D, Perkins C. Eigentaste: a constant time collaborative filtering algorithm. Inf Retrieval. 2001;4(2):133–51.
5. Berry MW, Dumais ST, O'Brien GW. Using linear algebra for intelligent information retrieval. SIAM Rev. 1995;37(4):573–95.
6. Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R. Indexing by latent semantic analysis. J Am Soc Inform Sci. 1990;41(6):391–407.
7. Billsus D, Pazzani MJ. Learning collaborative information filters. In: Proceedings of the 15th international conference on machine learning; 1998. San Francisco, CA: Morgan Kaufmann. p. 46–54.
8. Herlocker Jl, Konstan JA, Borchers A, Riedl J. An algorithmic framework for performing collaborative filtering. In: Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval; 1999; Berkeley, California, United States. p. 230–7.
9. Herlocker JL, Konstan JA, Terveen LG, Riedl J. Evaluating collaborative filtering recommender systems. ACM Trans Inf Syst. 2004;22(1):5–53.
10. Sarwar B, Karypis, G, Konstan J, Riedl J. Item-Based collaborative filtering recommendation algorithms. In: Proceedings of the 10th international world wide web conference; 2001. p. 285–95.

# Part III
# Radar and Sonar Signal Processing

# Electromagnetic Compatibility Analysis of Radar and Communication Systems in 35 GHz Band

Zebin Liu[1(✉)] and Weixia Zou[1,2]

[1] Key Laboratory of Universal Wireless Communications, MOE, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
lzbxtu@163.com
[2] State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, People's Republic of China

**Abstract.** The radar system and communication system in the military and civilian are developing to millimeter-wave band; in order to coordinate the frequency use of communication system and the radar system in this band, it is necessary to analyze the electromagnetic compatibility between them. In this paper, the LFM pulse compression radar in the 35 GHz band is taken as an example and the interference system model is constructed to analyze the electromagnetic compatibility with digital cellular mobile communication system in the same frequency band. The physical-layer simulation is carried out on the millimeter-wave radar and the communication system, and the corresponding interference limits are obtained by combining the path loss, bandwidth, and equipment space distribution of the specific system. Then, an effective link-level interference evaluation model is established, thus obtaining the system transmission power threshold and equipment distribution isolation distance threshold reference value. The research results can provide some technical references for the relevant departments of radio management in China.

**Keywords:** Millimeter wave · Radar · Electromagnetic compatibility · Physical layer

## 1 Introduction

With the millimeter-wave band having good prospects in terms of bandwidth, detection capability, transmission quality, security, etc., modern communication systems and radar systems are developing to the millimeter-wave band [1–4]. Electromagnetic compatibility analysis of the existing communication systems and radar systems is analyzed at low frequency bands [5–9]. There are two general ideas. One is to analyze from the perspective of power antagonism, according to the mutual interference power of radar and communication systems to get the result. This method is only from the power point of view, skimming the specific features of the system, and the result is often not ideal. And another idea is to analyze the electromagnetic compatibility issues from the perspective of the modulated signal waveform. However, this analysis does


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_79

not take into account the signal amplitude, signal transmission, and reception technologies, and it cannot necessarily accurately simulate the overall interference of the system. In this paper, the BER of the system is calculated by comparing the sending sequence and the receiving sequence one by one, and the transmission quality of the communication system is represented by the BER.

## 2 Simulation Model

### 2.1 Millimeter-Wave Communication System Simulation Model

In this paper, the millimeter-wave communication system in physical-layer simulation model is shown in Fig. 1, and the details of the system parameters are summarized in Table 1.

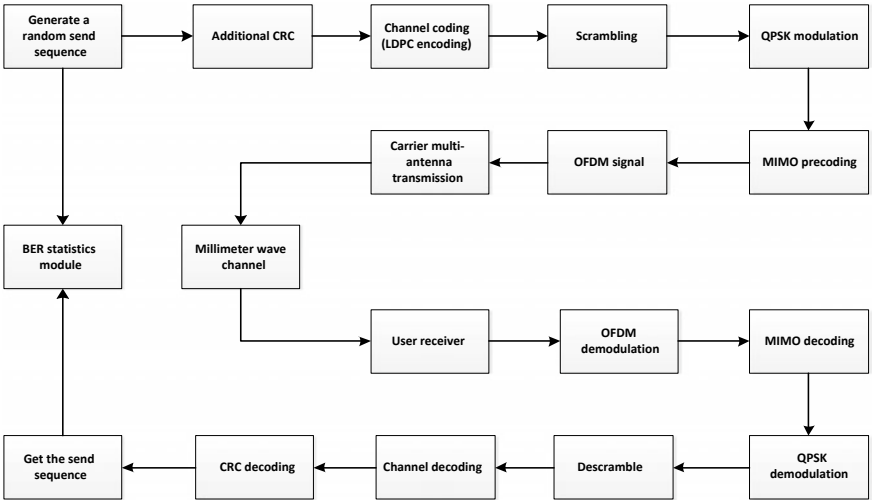



**Fig. 1.** Millimeter-wave communication system model in physical layer

**Table 1.** Millimeter-wave communication system parameters

| Symbol | Parameter | Value |
|---|---|---|
| $fc$ | Carrier frequency | 35 GHz |
| $N$ | Number of pulses | 3240 |
| $\Delta f$ | Subcarrier separation | 125 kHz |
| $T_p$ | Pulse repetition period | 8 us |
| $T$ | Pulse width | 1 us |
| $B$ | Total signal bandwidth | 400 MHz |
| $P_t$ | Transmit power | 19 db |
| $G_t$ | Transmit antenna gain | 20 dbi |
| $G_r$ | Receive antenna gain | 20 dbi |
| $d$ | Distance of BS and MS | 100 m |

### 2.2 LFM Radar System Simulation Model

The LFM signal is selected here, and signal expression is shown in (1).

$$S(t) = \text{rect}\left(\frac{t}{\tau}\right) e^{j2\pi(f_0 t + \frac{1}{2}ut^2)} \tag{1}$$

where $\tau$ is the pulse width and $f_0$ is the carrier frequency. $\text{rect}(\frac{t}{\tau})$ is a rectangular signal, and $u = \frac{B}{\tau}$ is the modulation slope.

The ideal echo expression of the radar signal is shown in (2).

$$\begin{aligned} s_r(t) &= A_0 \text{rect}\left(\frac{t - t_i}{\tau}\right) e^{j2\pi(f_0(t-t_i) + \frac{1}{2}u(t-t_i)^2)} \cdot \exp(jw) \\ A_0 &= \sum_{i=1}^{M} \sigma_i^{\frac{1}{2}} \left(\frac{P_t G_t^2 \lambda^2}{(4\pi)^3 R_i^4 L_{\text{sp}}}\right)^{\frac{1}{2}} \end{aligned} \tag{2}$$

where $M$ represents the number of targets, $\sigma_i$ is the radar cross-sectional area of the $i$th target, $t_i$ is the time between the speed of the light and the radar and the target, $R_i$ is the relative distance between the $i$th target and the radar, and $P_t$ is the radar The transmit power $G_t$ is the radar transmit antenna gain, $\lambda$ is the wavelength, Lsp is the decline of the radar signal, and w is the phase shift when the signal is reflected by the target.

Radar system simulation model is shown in Fig. 2.

The radar system is widely used in the 35 GHz frequency band, and the typical LFM radar system parameters are shown in Table 2. This paper mainly analyzes the electromagnetic compatibility of the same frequency band communication system with four typical radar systems in the table.

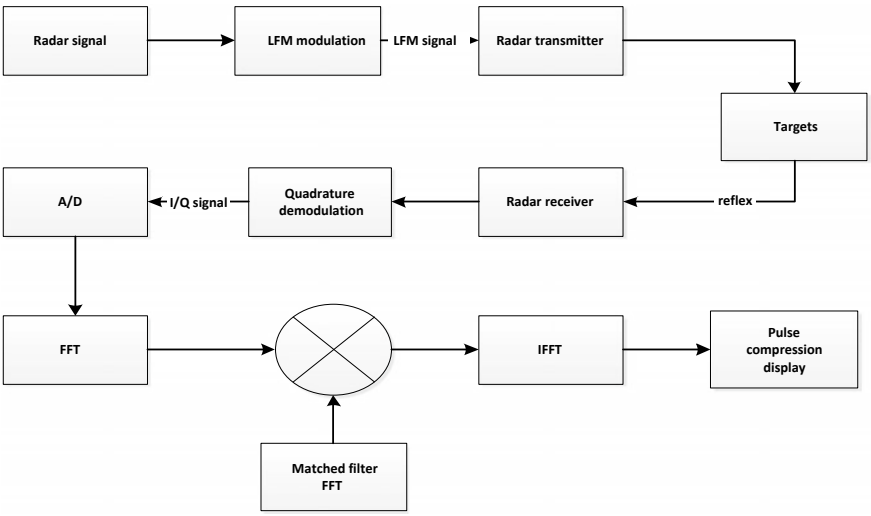



**Fig. 2.** Radar system simulation model in physical layer

**Table 2.** Typical radar parameters in the 35 GHZ band

| Radar type | Model A | Model B | Model C | Model D |
|---|---|---|---|---|
| Peak power (kW) | 100 | 10 | 200 | 300 |
| Antenna gain (dbi) | 43 | 50 | 46 | 50 |
| Pulse width (μs) | 1 | 10 | 0.5 | 2 |
| Duty cycle (%) | Max 0.2 | Max 0.1 | Max 0.5 | Max 1 |
| Working distance (km) | 8 | 8 | 9 | 14 |

## 3 Research Communication System's Interference to Radar

### 3.1 Separation Distances of Communication and Radar Systems

Under the premise of determining radar system parameters and communication system parameters, the relationship between different separation distances and radar peak-to-average ratio between the two systems can be analyzed, and the minimum separation distance between the systems can be obtained without affecting the normal operation of the radar. The changes in the peak-to-average ratio of the signals of the A, B, C, and D radars at different separation distances are shown in Fig. 3.

It can be seen from Fig. 3 that without affecting the normal working conditions of the radar, the type A radar needs to be separated from the base station at a minimum distance of 15 km, the type B radar needs to be separated by 8 km, the type C radar has a minimum distance of 12.5 km, and the type D radar needs an interval of 9 km.

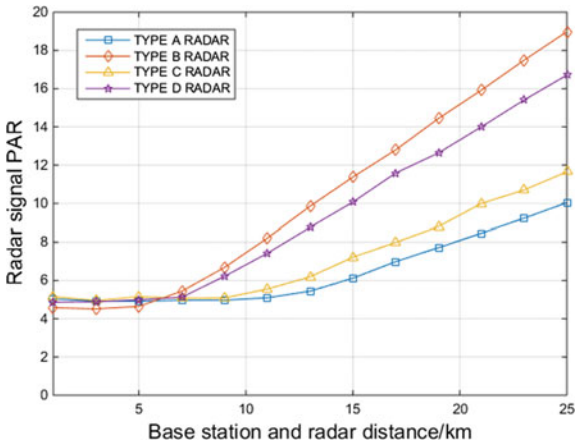



**Fig. 3.** Relationship between radar peak-to-average ratio and separation distance

### 3.2 Research Base Station Power's Interference to Radar System

When the separation distance between the base station and the radar system and the corresponding radar system parameters are selected, the transmission power of the base station can be changed to ensure the normal operation of the radar system. Assuming that the radar system and the communication system are deployed at a separation distance of 20 km, the peak-to-average ratio changes of the A, B, C, and D radars at different base station transmission powers are shown in Fig. 4.

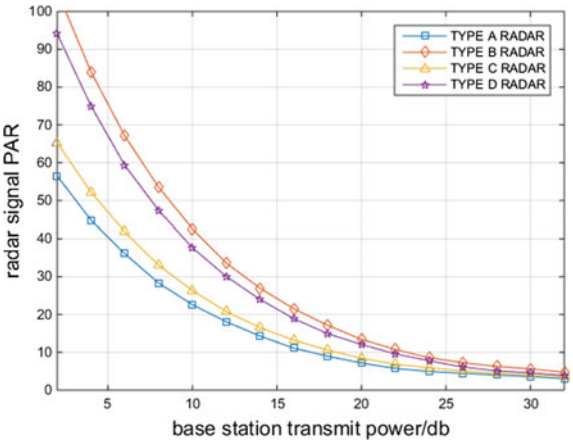



**Fig. 4.** Relationship between radar peak-to-average ratio and transmit power

From Fig. 4, it can be seen that when the transmission power of the base station is greater, the peak-to-average ratio of the radar target signal is smaller, and the probability of a missed radar is greater. When the deployment interval is 20 km, in the normal working range of the radar, type A radar allows the maximum transmit power of the base station to be 22 dB, type B is 30 dB, type C is 24 dB, type D is 26 dB. If the base station gradually increases from the radar, then the radar allows the maximum transmission power of the base station to gradually increase.

## 4 Research Radar System's Interference to Communication System

### 4.1 Separation Distances of Communication and Radar Systems

To ensure the normal operation of the communication system, the threshold of the radar system will also be analyzed from the two main aspects: distance between systems and power. When the communication parameters are certain, the change in communication bit error rate of the A, B, C, and D radars at different separation distances is shown in Fig. 5.

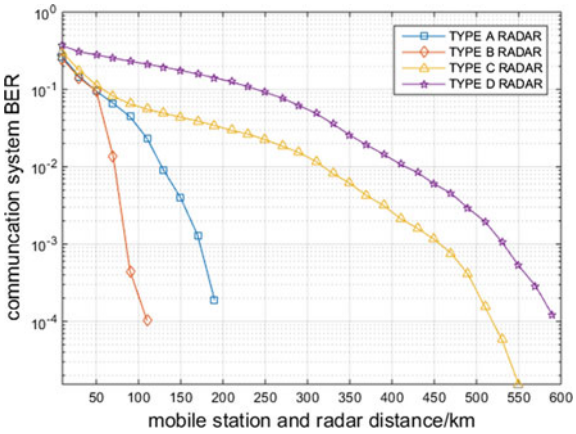



**Fig. 5.** Relationship between bit error rate and separation distance

In this paper, when the bit error rate is higher than $10^{-3}$, it is considered that the radar signal constitutes an effective interference to the communication signal. From Fig. 5, we can see that when guaranteeing the normal operation of the communication system, the minimum separation distance between type A radar and communication systems is 170 km, type B is 80 km, type C is 460 km, and type D is 530 km.

### 4.2 Research Rader Power's Interference to Communication System

When the separation distance of systems and the corresponding communication system parameters are selected, the radar transmission power can be changed to ensure the normal operation of the communication system. Assuming that the radar system and the mobile station are deployed at a separation distance of 20 km, the bit error rate changes of the signals of the A, B, C, and D radars at different radar transmit powers are shown in Fig. 6.

As can be seen from Fig. 6, when the transmission power of the radar is larger, the bit error rate of the system is bigger. Under the same conditions, because type A radar has the least antenna gain, it has the least interference to the communication system, and the type D radar has the largest. When the bit error rate of the communication system is $10^{-3}$, type A radar transmission power needs less than 26 db, type B is 26.5 db, type C is 27 db, and type D is 32 db.

## 5 Conclusion

Figures 3 and 5 comprehensively analyze the minimum distance between the two. When the transmission power of the base station is 19 db, the minimum distance between the typical radar and the base station should be obtained when both the radar system and the communication system work normally, as shown in Table 3.

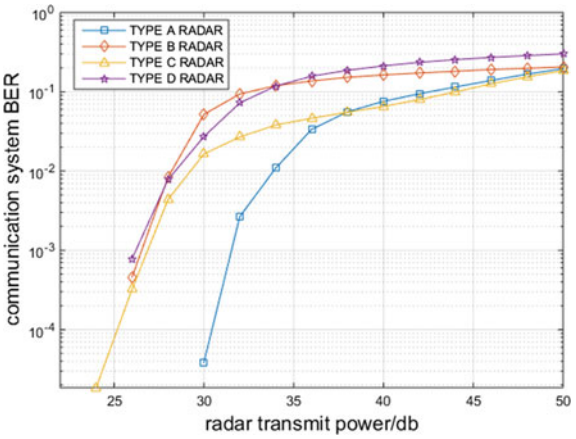



**Fig. 6.** Relationship between bit error rate and radar transmit power

**Table 3.** Minimum separation distance when coexist

| Radar type | Model A (km) | Model B (km) | Model C (km) | Model D (km) |
|---|---|---|---|---|
| The minimum distance required for normal radar operation | 15 | 8 | 12.5 | 9 |
| Minimum distance required for normal operation of communication systems | 170 | 80 | 460 | 530 |
| Minimum separation distance when radar and communication systems coexist | 170 | 80 | 460 | 530 |

Figures 4 and 6 comprehensively analyze the transmission power of the base station. When the distance between the base station and the radar is 20 km, the power distribution of the base station and radar when the communication system and the radar system coexist is shown in Table 4.

**Table 4.** Base station and radar power when radar and communication systems coexist

| Radar type | Model A (dB) | Model B (dB) | Model C (dB) | Model D (dB) |
|---|---|---|---|---|
| Radar normal work allows the maximum power of the BS | 22 | 30 | 24 | 26 |
| Radar maximum power for normal operation of the communication system | 26 | 27.5 | 27 | 32 |

The coexistence of the lightning system and the communication system can be realized by increasing the separation distance between the base station and the radar or appropriately reducing the transmission power of the radar according to the demand.

Comprehensive analysis shows that radar and communication systems that do not have an integrated design for radar communications require a high level of coexistence. The coexistence of radar communication systems can be achieved by further studying radar communication integrated waveform design.

## References

1. Li Y. Research status and development of millimeter wave communication technology. In: Sichuan provincial institute of communications academic annual conference proceedings; 2010.
2. Wang J, Ge J, Zhang Q, et al. Study of aircraft icing warning algorithm based on millimeter wave radar. Chin J Meteorol. 2017;31(6):1034–44.
3. Cheng Y, H Fei, H Feng, et al. Millimeter-wave emission characteristics of bilayer radar-infrared compound stealth material. Chin Opt Express. 2016;14(6):100–4.
4. Yue G, Wang Z, Chen L, et al. Demonstration of 60 GHz millimeter-wave short-range wireless communication system at 3.5 Gbps over 5 m range. Sci Chin Inf Sci. 2017; 60 (8):080306.
5. Lei W, Shuguo X. Radar interference analysis method for spread spectrum communication system. J Beihang Univ. 2012;38(9):1167–71.
6. Jing W, Dacheng Y, Ruiming Z, et al. Simulation research on mutual interference system between land mobile communication system and radionavigation radar. Mod Telecommun Technol. 2010;40(z1):74–80.
7. Zhao T, Du P, Wang Y. Analysis of interference of GSM mobile communication system to LFM radar in the same frequency band [J]. J Air Force Early Warning Coll. 2006;20(3): 171–3.
8. Sanders FH, Carroll JE, Sanders GA, et al. Effects of radar interference on LTE base station receiver performance. 2014.
9. Safavi-Naeini HA, Ghosh C, Visotsky E, et al. Impact and mitigation of narrow-band radar interference in down-link LTE. In: IEEE international conference on communications. IEEE; 2015:2644–9.

# SAR Image Denoising Via Fast Weighted Nuclear Norm Minimization

Huanyue Zhao[1], Caiyun Wang[1(✉)], Xiaofei Li[2], Jianing Wang[1], Chunsheng Liu[2], Yuebin Sheng[2], and Panpan Huang[1]

[1] College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, Jiangsu, People's Republic of China
wangcaiyun@nuaa.edu.cn

[2] Beijing Institute of Electronic System Engineering, Beijing 100854, People's Republic of China

**Abstract.** A new synthetic aperture radar (SAR) image denoising method based on fast weighted nuclear norm minimization (FWNNM) is proposed. SAR image is firstly modelled by a logarithmic additive model for modelling of the speckle. Then, the non-local similarity is used for image block matching. Next, according to the framework of the low-rank model, randomized singular value decomposition (RSVD) is introduced to replace the singular value decomposition (SVD) in weighted nuclear norm minimization (WNNM) for approximating the low-rank matrix. Finally, the gradient histogram preservation (GHP) method is employed to enhance the texture of the image. Experiments on MSTAR database show that the proposed approach is effective in SAR image denoising and the edge preserving in comparison with some traditional algorithms. Moreover, it is three times faster than WNNM method.

**Keywords:** Image denoising · Synthetic aperture radar · Nuclear norm · Singular value decomposition

## 1 Introduction

Synthetic aperture radar (SAR) image is inherently affected by speckle noise, which reduces radiometric resolution and affects the tasks of human interpretation and image analysis. In recent decades, scholars have designed a lot of algorithms to remove speckle noise, representative conventional methods are Lee filter, Frost filter, Kuan filter and GammaMAP [1] and so on. They all have a problem of losing texture details because of over-smoothing effect. Then, wavelet transform has been applied to SAR image denoising [2], however, which cannot effectively express the image edge information. In 2005, Buades applied non-local model to image denoising and designed non-local means denoising method [3], which is not effective in image edge area. Subsequently, K.Dabov proposed a block matching and 3D filtering algorithm, which has good denoising effect but has a high complexity [4]. In 2006, Terrence Tao proposed low-rank matrix approximation [5] and brought it into image denoising. Low-rank matrix approximation can be achieved by nuclear norm minimization (NNM) [6]. There are many similar image blocks which have similar features and data structure in


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_80

SAR image. So the matrix composed of similar image blocks is considered as approximately low rank [7]. Hence, NNM is able to be used for SAR image denoising.

However, NNM treats singular values as equivalents in the calculation process, which would cause deviations in practical application. To suppress deviations, spatially adaptive iterative singular value thresholding (SAIST) [8] is proposed, and the WNNM denoising method is also proposed [9]. Nevertheless, WNNM algorithm requires SVD in each iteration and consumes a lot of time. In addition, WNNM will over-smooth texture details.

This paper presents a fast WNNM (FWNNM) method to despeckle SAR images. SAR image is firstly modelled by a logarithmic additive model for modelling of the speckle. Then, the non-local similarity is used for image block matching. Next, according to the framework of the low-rank model, randomized RSVD [10] is introduced to replace the SVD in WNNM for approximating the low-rank matrix. Finally, GHP [11] method is employed to enhance the texture of the image. Experiments on MSTAR database show that the proposed approach is effective in SAR image denoising and the edge preserving in comparison with some traditional algorithms. Moreover, it is three times faster than WNNM method.

## 2 Proposed SAR Image Denoising Approach

The proposed approach is carried out in three steps. Figure 1 shows the process chart.

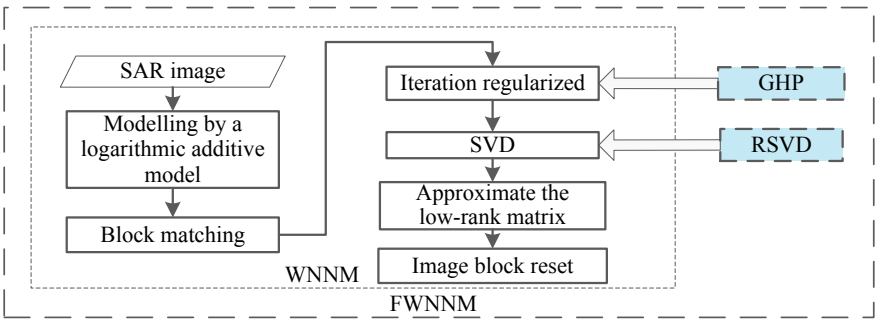



**Fig. 1.** Process chart of the proposed approach

### 2.1 Weighted Nuclear Norm Minimization (WNNM) Algorithm

The purpose of image denoising is to reconstruct the original image $\boldsymbol{x}$ from the noise image $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}$, where $\boldsymbol{n}$ is assumed to be additive Gaussian white noise, whose mean value is 0, variance is $\sigma_n^2$. Recently, the pioneering contribution of non-local self-similarity (NSS) [3] has led to extensive research on the method of denoising based on non-local self-similarity images. NSS reveals that there are many similar structures in a natural image, and the aggregation of non-local similar blocks to a given block can help to reconstruct a given block.

For the local block $\boldsymbol{y}_j$ in noise image $\boldsymbol{y}$, the non-local similar block can be searched in the image by block matching. These non-locally similar blocks are superimposed into a matrix and can be expressed as $\boldsymbol{Y}_j = \boldsymbol{X}_j + \boldsymbol{N}_j$, where $\boldsymbol{X}_j$ and $\boldsymbol{N}_j$ are the original image block and the noise block matrix, respectively. In fact, $\boldsymbol{X}_j$ should be an approximately low-rank matrix, and the low-rank matrix approximation can be implemented by NNM method. Therefore, the WNNM algorithm is able to be used for low-rank matrix recovery. WNNM can be described as follows

$$\widehat{\boldsymbol{X}}_j = \arg\min_{\boldsymbol{X}_j} \frac{1}{\sigma_n^2} \left\| \boldsymbol{Y}_j - \boldsymbol{X}_j \right\|_F^2 + \left\| \boldsymbol{X}_j \right\|_{w,*} \tag{1}$$

where $\left\| X_j \right\|_{w,*} = \sum_i \left| w_i \sigma_i(X_j) \right|_1$, $\boldsymbol{w} = [w_1, \ldots, w_n]$, $w_i \geq 0$ is a non-negative weight assigned to $\sigma_i(\boldsymbol{X}_j)$. Equation (1) can be decomposed by SVT to obtain the global optimal solution

$$\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \mathrm{svd}(\widehat{\boldsymbol{X}}_j) \tag{2}$$

$$\boldsymbol{\Lambda}_{\boldsymbol{w}} = S_{\boldsymbol{w}}(\boldsymbol{\Sigma}) = \max(\sigma_i(\boldsymbol{X}_j) - \boldsymbol{w}, 0) \tag{3}$$

The calculation formula of weight $w_i$ is as follows

$$w_i = \mathrm{c}\sqrt{q} / (\sigma_i(\boldsymbol{X}_j) + \varepsilon) \tag{4}$$

where $\mathrm{c} > 0$ is a constant, $q$ is the number of similar blocks in $\boldsymbol{Y}_j$ and $\varepsilon = 10^{-16}$ is the number that avoids being divided by zero.

Assuming that noise is uniformly distributed in $\boldsymbol{U}$ and $\boldsymbol{V}$ subspaces, $\sigma_i(\boldsymbol{X}_j)$ can be estimated by

$$\sigma_i(\boldsymbol{X}_j) = \sqrt{\max(\hat{\sigma}_i(\boldsymbol{Y}_j)) - n\sigma_n^2, 0} \tag{5}$$

where $\sigma_i(\boldsymbol{Y}_j)$ is the $i^{\text{th}}$ singular value of $\boldsymbol{Y}_j$, $\sigma_n^2$ is the noise variance.

The WNNM algorithm has the disadvantages of low efficiency, over-smooth image edge and texture details. In order to overcome these disadvantages, FWNNM is proposed in this paper.

### 2.2 Randomized Singular Value Decomposition (RSVD)

SVD is performed in every iteration of the WNNM algorithm. The complete SVD has a large amount of calculation and takes a long time. And the sum of the first 10% or even 1% singular values accounts for more than 99% of the sum of all the singular values. So in practical applications, it is usually only necessary to truncated SVD [12]. To further improve the operation speed while ensuring the accuracy of the operation, a faster RSVD is proposed to replace the truncated SVD in [10]. Therefore, RSVD is introduced to replace SVD in the WNNM algorithm which can greatly increase the speed of operation.

The decomposition of the random matrix mainly includes random sampling algorithm that extracts a number of columns or rows from the original matrix for decomposition and random projection, in which the high-dimensional matrix is projected into a low-dimensional decomposition operation. For RSVD, the second random projection algorithm is more suitable [13].

RSVD algorithm based on randQB [14] process is an efficient algorithm. The algorithm firstly generates a $n \times l$-dimensional Gaussian random matrix $\boldsymbol{\Omega}$, and $l$ is far less than the smaller values of $m$ and $n$. To construct an $m \times l$-dimensional sample matrix $\boldsymbol{Y} = \boldsymbol{A\Omega}$, and $\boldsymbol{\Omega}$ is multiplied by a matrix $\boldsymbol{A}$ of $m \times n$ dimensions to be decomposed. Then, QR decomposition is performed on $\boldsymbol{Y}$ to obtain an orthogonal matrix $\boldsymbol{Q}$ of $m \times l$ dimensions. Then, construct an $l \times n$-dimensional matrix $\boldsymbol{B} = \boldsymbol{Q}^{\mathrm{T}} \times \boldsymbol{A}$ and finally perform SVD on $\boldsymbol{B}$, that is

$$\boldsymbol{B} = \boldsymbol{S\Sigma V}^{\mathrm{T}} \tag{6}$$

Let $\boldsymbol{QS} = \boldsymbol{U}$ to get the SVD of $\boldsymbol{A}$,

$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^{\mathrm{T}} \tag{7}$$

Since the number of rows of the matrix $\boldsymbol{B}$ is much smaller than the smaller value of $m$ and $n$, the time spent on the SVD of $\boldsymbol{B}$ is less than the time consumed for the SVD of the matrix $\boldsymbol{A}$.

### 2.3 Texture Enhanced Image Denoising via Gradient Histogram Preservation

The WNNM algorithm has the disadvantage of over-smoothing the image texture details and reducing the visual quality of the image. The denoising effect is better if the gradient of the denoised image is closer to original image. So a texture enhancement method based on GHP is proposed in [11]. This algorithm first estimates the gradient histogram $\boldsymbol{h_r}$ of the original image $\boldsymbol{x}$ as a reference gradient and takes the updated image as close as possible to the reference gradient histogram as a constraint and finally obtains the resulting image. Among them, the formula of the gradient histogram $\boldsymbol{h_r}$ is as follows,

$$\boldsymbol{h_r} = \arg\min\left\{ \left\| \boldsymbol{h_y} - \boldsymbol{h_x} \otimes \boldsymbol{h_g} \right\|^2 + \mathrm{d} \cdot R(\boldsymbol{h_x}) \right\} \tag{8}$$

where c is a constant, and $R(\boldsymbol{h_x})$ is a regularized term based on the prior information of the natural image gradient histogram. $\boldsymbol{g} = \nabla \boldsymbol{n}$ is the gradient of noise $\boldsymbol{n}$. $\boldsymbol{h_x}$ and $\boldsymbol{h_g}$ are the histogram of the original image and the gradient $\boldsymbol{g}$, respectively.

To make the gradient histogram of the denoising image approximate to the reference histogram $\boldsymbol{h_r}$, the denoising model based on GHP is as follows

$$\boldsymbol{h_r} = \arg\min_{\boldsymbol{h_x}} \left\{ \left\| \boldsymbol{h_y} - \boldsymbol{h_x} \otimes \boldsymbol{h_g} \right\|^2 + \mu \cdot R(\boldsymbol{h_x}) \right\} \tag{9}$$

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x},\boldsymbol{F}} \left\{ \begin{array}{l} \dfrac{1}{2\sigma_n^2} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda R(\boldsymbol{x}) + \\ \mu \|F(\nabla \boldsymbol{x}) - \nabla \boldsymbol{x}\|^2 \end{array} \right\} \text{s.t. } \boldsymbol{h_f} = \boldsymbol{h_r} \tag{10}$$

where $\nabla$ is gradient operation. $F$ is an odd function that is increasing monotonically in $(0, +\infty)$. $\boldsymbol{h_F}$ represents the gradient image after transformation $|F(\nabla \boldsymbol{x})|$ histogram, $\boldsymbol{h_r}$ represents the histogram of the original image, $R(\boldsymbol{x})$ denotes the regularization term and $\lambda$ is a positive constant.

The texture enhancement method based on GHP is introduced into the iterative regularization term of WNNM algorithm, which improves the shortcomings of over-smoothing texture detail in WNNM algorithm.

## 3 Fast SAR Image Denoising Method Based on FWNNM

SAR image is firstly modelled by a logarithmic additive model for modelling of the speckle. Then, the non-local similarity is used for image block matching. Next, according to the framework of the low-rank model, randomized RSVD is introduced to replace the SVD in WNNM for approximating the low-rank matrix. Finally, GHP method is employed to enhance the texture of the image. The steps of the fast denoising method for SAR image based on FWNNM are as follows

**Algorithm: Fast Weighted Nuclear Norm Minimization (FWNNM)**
**Input**: Noise image $\boldsymbol{z}$, the number of iterations $K$
**Output**: Clean image $\tilde{\boldsymbol{x}}^{(K)}$
**Main Iterations:**

1: Logarithmic transformation of noise image $\boldsymbol{z}$ to get image $\boldsymbol{y}$
2: Initialize $\hat{\boldsymbol{x}}^{(0)} = \boldsymbol{y}, \boldsymbol{y}^{(0)} = \boldsymbol{y}$
3: for $k$ = 1: $K$
GHP texture enhancement introduced iterative regularization $\boldsymbol{y}^{(k+1)} = \hat{\boldsymbol{x}}^k + \delta(\boldsymbol{y} - \hat{\boldsymbol{x}}^{(k)}) + \mu \|F(\nabla \boldsymbol{x}) - \nabla \boldsymbol{x}\|^2$ s.t. $\boldsymbol{h_f} = \boldsymbol{h_r}$
4: for each patch $\boldsymbol{y_j}$ in $y^{(k)}$ do
5: Find similar patch group $\boldsymbol{Y}_j$
6: Estimate weight vector $\boldsymbol{w}$
7: Randomized singular value decomposition $[\boldsymbol{U}; \boldsymbol{\Sigma}; \boldsymbol{V}]$ = r$svd(\boldsymbol{Y}_j)$
8: Get the estimation: $\hat{\boldsymbol{X}}_j = \boldsymbol{U}\boldsymbol{S}_{\boldsymbol{w}}(\boldsymbol{\Sigma})\boldsymbol{V}^{\mathrm{T}}$
9: Aggregate $\boldsymbol{X}_j$ to form the clean image $\tilde{\boldsymbol{x}}^{(k)}$
10: **end for**
11: **Output:** Clean image $\tilde{\boldsymbol{x}}^{(K)}$

## 4 Experimental Results and Discussion

In order to evaluate the denoising ability of the proposed method for SAR image, this paper introduces the other two methods to despeckle the image. These two methods are SAIST [8] and WNNM [9]. We use the moving and stationary target acquisition and recognition (MSTAR) three-class public dataset. The dataset consists of X-band SAR images with 0.3 m × 0.3 m resolution of multiple targets, with the polarization mode is HH polarization, and the detection azimuth is 360°. We use a typical three classes of ground targets: BTR70_SNC71 (armoured vehicle), BMP_SN9563 (armoured vehicle) and T72_SN132 (main battle tank).

We chose two experimental SAR images. The first one is the artificially simulated image, which is obtained by adding Gaussian white noise with variance 50 to a clean image. The second one is a SAR image (MSTAR). The noise image is denoised by SAIST, WNNM and the proposed FWNNM method. Figure 2 shows the denoised results of the SAR image of BTR70_SNC71 after denoising by three algorithms.

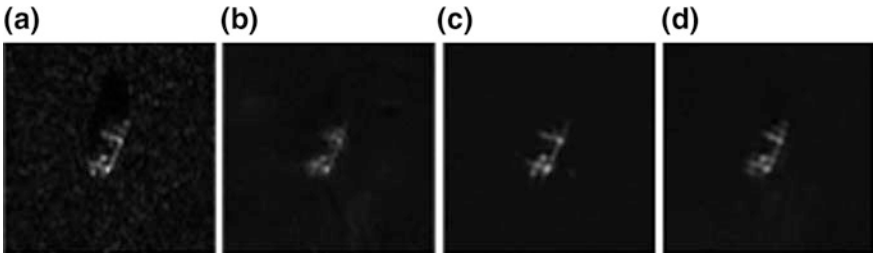



**Fig. 2.** Denoising results of the BTR70_SNC71 image. **a** Original image, **b** SAIST, **c** WNNM, **d** proposed approach

As shown in Fig. 2, compared with the WNNM and the SAIST algorithm, the FWNNM algorithm can better preserve the edge and texture details of the target and has better visual effects.

In order to more accurately describe the effect of despecking objectively, we introduce peak signal-to-noise ratio (PSNR), equivalent number of looks (ENL), and the edge preserve index (EPI). The higher PSNR shows stronger denoising ability. The larger ENL means the better the visual effect of the denoised image. With larger EPI, the algorithm has stronger edge-keeping ability. Table 1 shows the results of three objective evaluation data after three kinds of denoising algorithms.

From Table 1, compared with the other two algorithms, the PSNR of the image denoised with FWNNM is improved significantly, which shows that the denoising ability of FWNNM is relatively stronger. The ENL of image denoised with FWNNM is improved, which indicates that the visual effect of the image denoised by FWNNM is better than the other two algorithms. The EPI of the image after denoising by FWNNM is higher than the other two algorithms, which shows that the algorithm's edge-keeping ability is obviously better than the other two algorithms.

**Table 1.** Performance comparison of three denoising algorithms

| Target | Denoising method | PSNR/(dB) | ENL | EPI |
|---|---|---|---|---|
| BMP_SN9563 | SAIST | 29.37 | 3.18 | 0.26 |
| | WNNM | 29.43 | 2.43 | 0.32 |
| | FWNNM | 29.51 | 3.24 | 0.38 |
| BTR70_SNC71 | SAIST | 30.12 | 1.65 | 0.29 |
| | WNNM | 30.14 | 1.70 | 0.36 |
| | FWNNM | 30.16 | 1.76 | 0.48 |
| T72_SN132 | SAIST | 31.32 | 1.22 | 0.35 |
| | WNNM | 31.27 | 1.21 | 0.44 |
| | FWNNM | 31.37 | 1.28 | 0.46 |

Experiments were carried out in MATLAB R2014a using the MatConvNet toolbox, with an Intel Core CPU at 2.50 GHz. Interestingly, once the testing is over, FWNNM exhibits the lowest run-time complexity as shown in Table 2.

**Table 2.** Run-time comparisons for despeckling an image of size 128 × 128 (t/s)

| Target | SAIST | WNNM | FWNNM |
|---|---|---|---|
| BMP_SN9563 | 55.0 | 40.0 | 12.6 |
| BTR70_SNC71 | 55.0 | 40.7 | 12.6 |
| T72_SN132 | 55.2 | 40.9 | 12.5 |
| Average | 55.1 | 40.5 | 12.5 |

From Table 2, to the same target, the denoising speed of the FWNNM algorithm is four times faster than the SAIST algorithm, three times faster than the WNNM algorithm, and the denoising efficiency is improved significantly.

## 5 Conclusion

In this paper, a new synthetic aperture radar image denoising method based on FWNNM is proposed. First, SAR image is modelled by a logarithmic additive model for modelling of the speckle. Then, the non-local similarity is used for image block matching. Next, according to the framework of the low-rank model, RSVD is introduced to replace SVD in WNNM for approximating the low-rank matrix. Finally, the GHP method is employed to enhance the texture of the image. Experiments on MSTAR database show that the proposed approach is effective in SAR image denoising and the edge preserving in comparison with some traditional algorithms. Moreover, it is three times faster than WNNM method.

**Acknowledgments.** This work was supported in parts by the National Natural Science Foundation of China (no. 61301211), the Postgraduate Education Reform Project of Jiangsu Province (no. JGZZ17_008) and the Postgraduate Research and Practice Innovation Programme of Jiangsu Province (no. KYCX18_0295).

## References

1. Yi Z, Yin D, Hu An Z, et al. SAR image despeckling based on non-local means filter. J Electron Inf Technol. 2012;34(4):950–5.
2. Ray A, Kartikeyan B, Garg S. Towards deriving an optimal approach for denoising of RISAT-1 SAR data using wavelet transform. Int J Comput Sci Eng. 2016;4(10):33–46.
3. Buades A, Coll B, Morel JM. A non-local algorithm for image denoising. In: IEEE computer society conference on, computer vision and pattern recognition. IEEE, vol. 2; 2005. p. 60–5.
4. Chen G, Xie W, Dai S. Image denoising with signal dependent noise using block matching and 3D filtering. In: International symposium on neural networks. Springer International Publishing; 2014. p. 423–30.
5. Zhou M, Song ZJ. Video background modeling based on sparse and low-rank matrix decomposition. Appl Res Comput. 2015;32(10):3175–8.
6. Zhang WQ, Zhang HZ, Zuo WM, et al. Weighted nuclear norm minimization model for matrix completion. Compu Sci. 2015;42(7):254–7.
7. Zhao J, Wang PP, Men GZ. SAR image denoising based on nonlocal similarity and low rank matrix approximation. Comput Sci. 2017;44(s1):183–7.
8. Dong W, Shi G, Li X. Nonlocal image restoration with bilateral variance estimation: a low-rank approach. IEEE Trans Image Process a Public IEEE Sig Process Soc. 2013;22(2):700–11.
9. Gu S, Zhang L, Zuo W, et al. Weighted nuclear norm minimization with application to image denoising, In: IEEE conference on computer vision and pattern recognition. IEEE computer society; 2014. p. 2862–9.
10. Feng X, Li KX, Yu WJ, et al. Fast matrix completion algorithm based on randomized singular value decomposition and its applications. J Comput-Aided Des and Comput Graphics. 2017;29(12).
11. Zuo W, Zhang L, Song C, et al. Texture enhanced image denoising via gradient histogram preservation. In: Computer vision and pattern recognition. IEEE; 2013:1203–10.
12. Sharma LN. Information theoretic multiscale truncated SVD for multilead electrocardiogram. Comput Methods Prog Biomed. 2016;129(C):109–16.
13. Wang P, Cai SJ, Liu Y. Improvement of matrix completion algorithm based on random projection. J Comput Appl. 2014;34(6):1587–90.
14. Larsen RM (2017) PROPACK-Software for large and sparse SVD calculations. http://sun.stanford.edu/∼rmunk/, PROPACK.

# A Real-Valued Approximate Message Passing Algorithm for ISAR Image Reconstruction

Wenyi Wei[1], Caiyun Wang[1(✉)], Jianing Wang[2], Xiaofei Li[2], Yuebin Sheng[2], Chunsheng Liu[2], and Panpan Huang[1]

[1] College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, Jiangsu, People's Republic of China
wangcaiyun@nuaa.edu.cn

[2] Beijing Institute of Electronic System Engineering, Beijing 100854, People's Republic of China

**Abstract.** Compressed sensing (CS) theory describes the signal using space transformation to obtain linear observation data selectively, breaking through the limit of the traditional Nyquist theorem. In this paper, we aim at accelerating the current approximate message passing (AMP) and propose an approach named real-valued AMP (RAMP) for faster and better inverse synthetic aperture radar (ISAR) imaging reconstruction. The azimuth dictionary is first processed with real. We then use matrix processing to solve the AMP vector iterative method, by utilizing the relation between the quantification of matrix product and the Kronecker product. The experimental results are presented to demonstrate the validity of this method.

**Keywords:** Inverse synthetic aperture radar (ISAR) · Image compressed sensing · Approximate message passing

## 1 Introduction

CS theory includes sparse representation of signals, construction of perceptual matrix, and reconstruction method [1–4]. Radar data can usually meet the requirements of sparseness of CS theory [5]. Therefore, CS method is possible to be applied to ISAR.

Numerous CS algorithms have been applied to reconstruct images, including greedy iterative method, convex optimization method, the one based on Bayesian framework, etc. Orthogonal matching pursuit (OMP) is the most commonly used, which belongs to greedy iterative methods. Smoothed l0 norm (SL0) proposed by [6] is a kind of convex optimization method. Compared with the traditional methods, AMP proposed by Donoho et al. [7–10] reduces the complexity of the method and has higher reconstruction quality by utilizing the central limit theorem and the Taylor expansion.

AMP is based on complex operation, the main drawback of which is heavy computation [11]. To solve this problem, we propose RAMP by introducing the ISAR echo azimuth compression sensing model. Simulation results of the middle of the trajectory show that the proposed RAMP performs favorably in terms of computation and speed.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_81

## 2 Azimuth Compressed Sensing Model

The target imaging is equivalent to the turntable model. Let the radar slow time series contain $M$ elements, and $P$ is the number of Doppler sequences. CS uses the azimuth sensing matrix $\boldsymbol{\Phi} \in \boldsymbol{R}^{Q\times M}$ to project the signal into the low-dimensional space, resulting in a series of measured values. When the random noise is introduced, the observation set matrix $\boldsymbol{Y}$ can be expressed as:

$$\boldsymbol{Y} = \boldsymbol{\Phi S} = \boldsymbol{\Phi \Psi \theta} + \boldsymbol{N} \tag{1}$$

where $\boldsymbol{S} \in \boldsymbol{R}^{M\times N}$ is a reconstructed two-dimensional super-resolution echo matrix, and $\boldsymbol{\Psi} \in \boldsymbol{R}^{M\times P}$ is an azimuth dictionary matrix.

## 3 Proposed ISAR Image Reconstruction Approach

Compared with the complex number operation, the real-valued processing can reduce the amount of calculation and improve the speed of signal processing. In this paper, the unitary transformation method is used for the modifiability. The steps are as follows:

(a) In order to facilitate real-valued processing, the slow sampling time zero is placed in the middle.
(b) The azimuth dictionary is modified in real-valued. Pre-multiplying the azimuth dictionary $\boldsymbol{\Psi}$ by a real matrix $\boldsymbol{Q} \in \boldsymbol{R}^{M\times M}$, we can obtain:

$$\bar{\boldsymbol{S}} = \boldsymbol{Q\Psi\theta} = \boldsymbol{A\theta} \tag{2}$$

where $\bar{\boldsymbol{S}} \in \boldsymbol{R}^{M\times N}$ is a real-valued modified super-resolution matrix.

Substituting (2) into (1) gives:

$$\boldsymbol{Y} = \boldsymbol{\Phi Q \Psi \theta} + \boldsymbol{N} = \boldsymbol{\Phi A \theta} + \boldsymbol{N} = \boldsymbol{\Phi}\bar{\boldsymbol{S}} + \boldsymbol{N} \tag{3}$$

Before reconstructing the echo, it is necessary to do the vectorization. vec($\boldsymbol{A}$) represents the quantization of the matrix $\boldsymbol{A}$, and the sign $\otimes$ represents the Kronecker product. Vectorizing (3), we can obtain:

$$\begin{aligned} \mathrm{vec}(\boldsymbol{Y}) &= \mathrm{vec}(\boldsymbol{\Phi}\bar{\boldsymbol{S}} + \boldsymbol{N}) \\ &= \mathrm{vec}(\boldsymbol{\Phi}\bar{\boldsymbol{S}}) + \mathrm{vec}(\boldsymbol{N}) \\ &= (\boldsymbol{I} \otimes \boldsymbol{\Phi})\mathrm{vec}(\bar{\boldsymbol{S}}) + \mathrm{vec}(\boldsymbol{N}) \end{aligned} \tag{4}$$

Let $\boldsymbol{A}' = \boldsymbol{I} \otimes \boldsymbol{\Phi}$, $\boldsymbol{s} = \mathrm{vec}(\bar{\boldsymbol{S}})$, $\boldsymbol{n} = \mathrm{vec}(\boldsymbol{N})$, and $\boldsymbol{y} = \mathrm{vec}(\boldsymbol{Y})$. The vector $\boldsymbol{s}$ is sparse and the matrix $\boldsymbol{A}$ satisfies the isometric constraints, so the problem of sparse reconstruction can be expressed as:

$$\arg\min_{s}\left\{\lambda\|\boldsymbol{s}\|_1+\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{A}'\boldsymbol{s}\|_2^2\right\} \tag{5}$$

Supposing that $\boldsymbol{s}$ obeys the Laplace distribution of the parameter $\lambda$, the corresponding edge probability problem is simplified by the cross-product information transfer method. With product messaging method to simplify the edge of the corresponding probability problem, AMP iterative can be obtained after the approximation and the processing of vectorization. By utilizing the relationship between the quantification of matrix product and the Kronecker product, we have:

$$\begin{cases}\boldsymbol{X}^t=\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{U}'^{t-1}+\boldsymbol{U}^{t-1}\\ \boldsymbol{U}'^t=\boldsymbol{y}-\boldsymbol{\Phi}\boldsymbol{U}^{t-1}+\boldsymbol{U}'^{t-1}\frac{1}{2\delta}\left[\begin{array}{c}\left\langle\dfrac{\partial\mathrm{Re}(\alpha)}{\partial r}(\boldsymbol{X}^t;\gamma^t)\right\rangle\\ +\left\langle\dfrac{\partial\mathrm{Im}(\alpha)}{\partial i}(\boldsymbol{X}^t;\gamma^t)\right\rangle\end{array}\right]\\ \boldsymbol{U}^{t-1}=\alpha(\boldsymbol{X}^t;\gamma^t)\end{cases} \tag{6}$$

where $t$ represents the number of iterations, $\boldsymbol{U}$ denotes the node message mean matrix $\boldsymbol{X}^t=\boldsymbol{U}^{t-1}+\boldsymbol{A}'^*\boldsymbol{U}'^{t-1}$, and $r$ and $i$ are, respectively, the real part and imaginary part of each element of $\boldsymbol{X}^t$.

The method flowchart of RAMP is shown in Fig. 1.

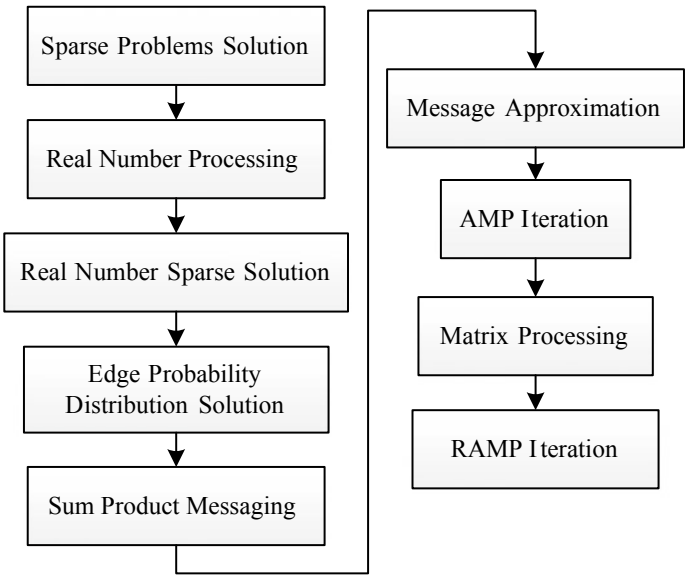



**Fig. 1.** Process chart of the proposed approach

## 4 Experimental Results and Discussion

To evaluate the performance of the proposed methods, we use Fig. 2 as the refactoring object. The proposed method is compared to OMP, SL0, and AMP.

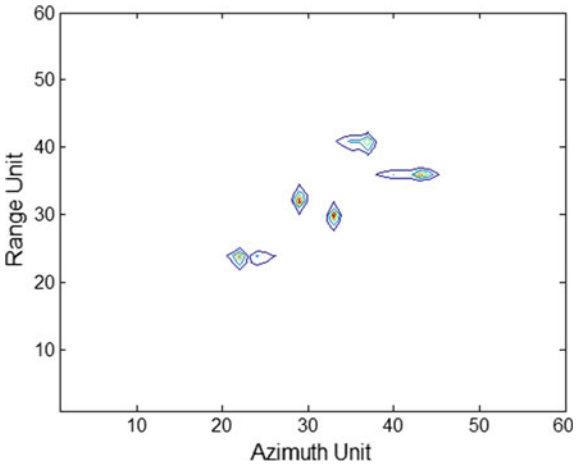



**Fig. 2.** ISAR imaging results of midcourse target

We select the sampling rate of 0.3, 0.35, 0.4, 0.45, and 0.5, and the simulation is performed according to the different reconstruction method. Taking the sampling rate of 0.4 as an example, Fig. 3 shows the imaging results.

Figure 4 shows the relationship between the computing time of the four reconstruction methods and the entropy value of the imaging results versus different sampling rates.

From Figs. 3 and 4, it is clear that the entropy of OMP and SL0 is larger, and the imaging effect is poor, while AMP and RAMP are opposite.

We select the SNR of 2, 4, 6, 8, and 10 dB in turn, and the simulation is performed according to the different reconstruction method. Taking the SNR of 6 dB as an example, Fig. 5 shows the imaging results.

Figure 6 shows the relationship between the reconstruction time of the four reconstruction methods and the entropy value of the imaging results versus different SNR.

It can be seen from Figs. 5 and 6 that the larger the SNR is, the better the imaging effect is. Vs. the same SNR, OMP has a long reconstruction time, while RAMP takes the least time. From the perspective of imaging, OMP and SL0 have higher entropy value, which is difficult to distinguish when the SNR is low. While the reconstructed image of RAMP is significantly better than the first two.

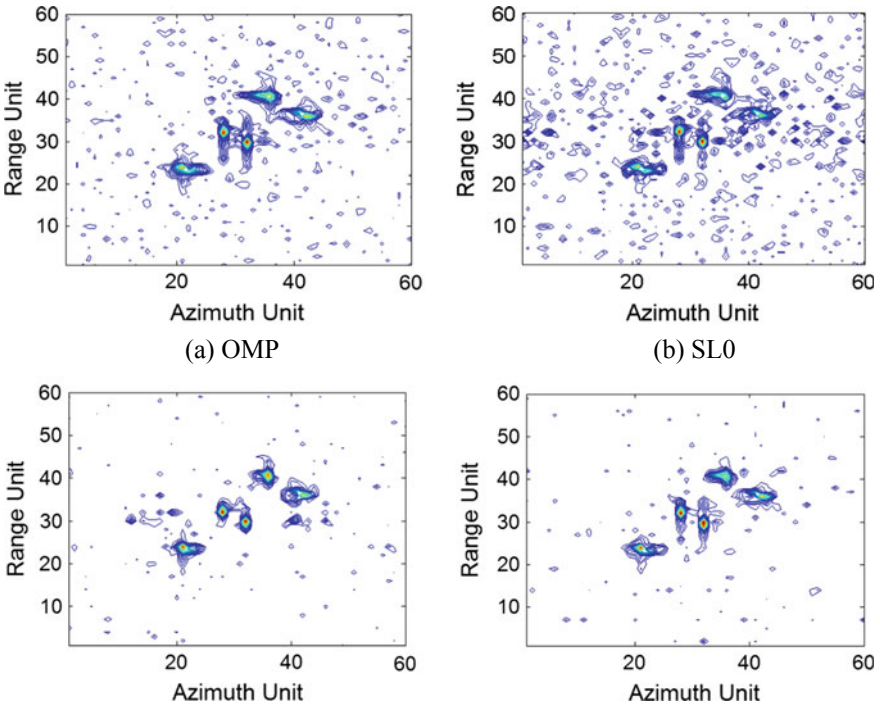



**Fig. 3.** Imaging results of four reconstruction methods

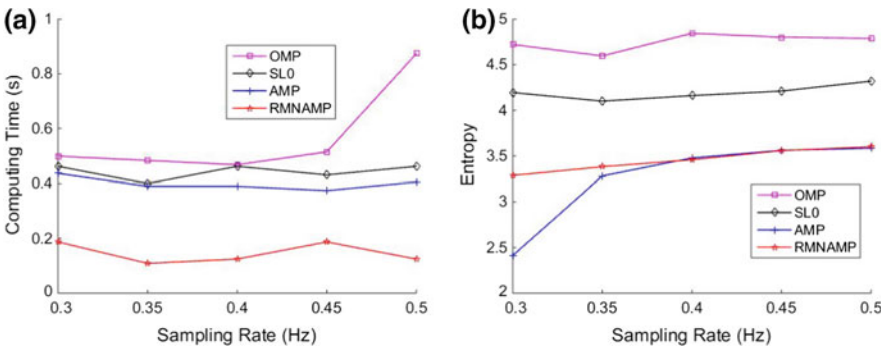



**Fig. 4.** Computing time and entropy versus different sampling rates

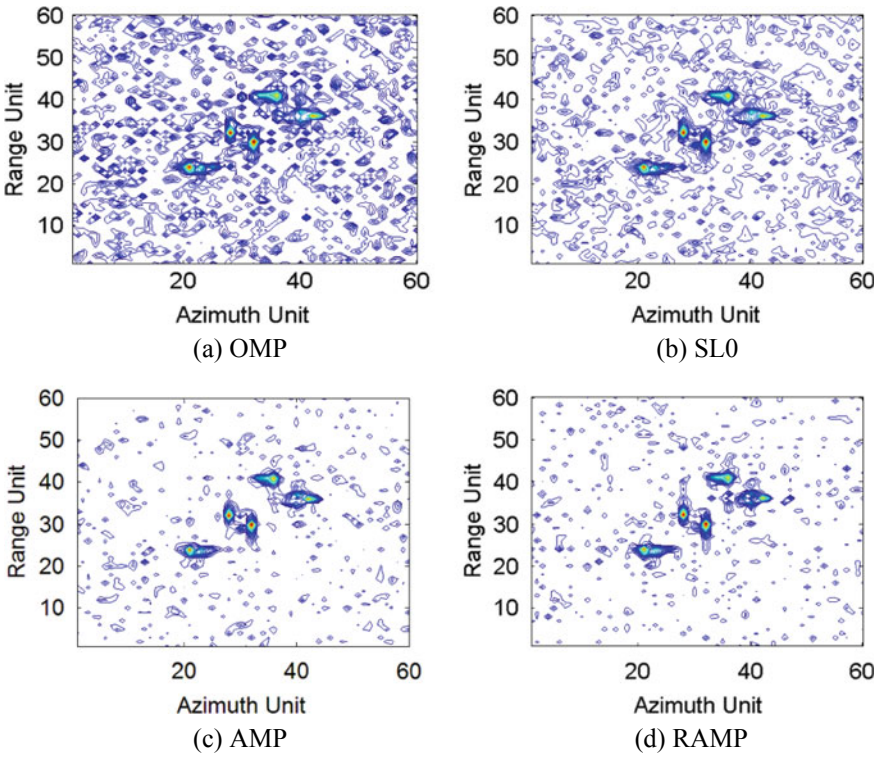



**Fig. 5.** Imaging results of four reconstruction methods versus different SNR

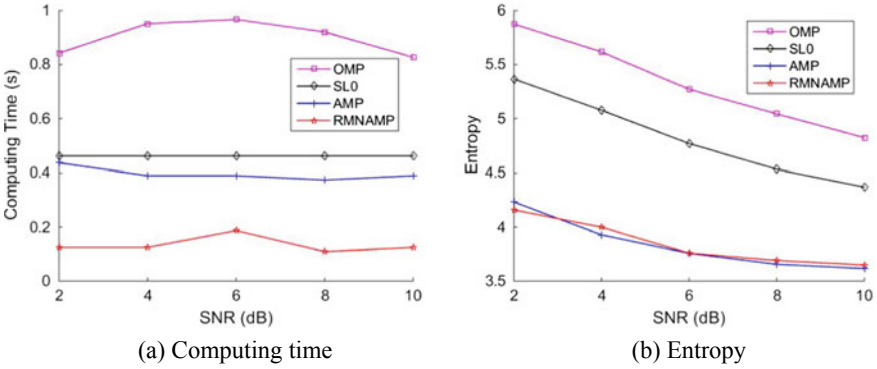



**Fig. 6.** Computing time and entropy versus different SNR

## 5 Conclusions

Compression sensing can reconstruct sparse signals through less observational data. This paper proposes an improved method based on the AMP reconstruction using real-valued processing and matrix calculation. The method reduces the computation of matrix calculation by means of real-valued processing of complex matrix. The matrix processing step is added to the traditional AMP iterative method to further improve the speed of the method reconstruction. Finally, the simulation data show significant improvements in the speed and imaging effect, compared with OMP, SL0, and AMP.

**Acknowledgments.** This work was supported in parts by the National Natural Science Foundation of China (no. 61301211), the Postgraduate Education Reform Project of Jiangsu Province (no. JGZZ17_008), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (no. KYCX18_0295).

## References

1. Qiong-Hai D, Chang-Jun F, Xiang-yang J. Research on compressed sensing. Chin J Comput. 2011;34(3):425–34.
2. Kipnis A, Reeves G, Eldar YC, Compressed sensing under optimal quantization. In: IEEE international symposium on information theory (ISIT), Aachen; 2017. p. 2148–52.
3. Zheng L, Maleki A, Liu Q, An lp-based reconstruction algorithm for compressed sensing radar imaging. In: IEEE radar conference, Philadelphia; 2016. p. 1–5.
4. Donoho D. Compressed sensing. IEEE Trans Inf Theory. 2006;52(4):1289–306.
5. Zhu Q, Volz R, Mathews JD. Coherent radar imaging based on compressed sensing. Radio Sci. 2015;50(12):1271–85.
6. Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed l0 norm. IEEE Trans Signal Process. 2009;57(1):289–301.
7. Tang L, Jiao S. Compressed sensing radar imaging using approximate message passing. CAAI Trans Intell Syst. 2015;10(4):592–8.
8. Maleki A, Anitori L, Yang Z. Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP). IEEE Trans Inf Theory. 2013;59(7):4290–308.
9. Byrne E, Gribonval R, Schniter P, Sketched clustering via hybrid approximate message passing. In: Asilomar conference on signals, systems and computers, Pacific Grove; 2017. p. 410–4.
10. Patterson S, Eldar YC, Keidar I. Distributed compressed sensing for static and time-vary ing networks. IEEE Trans Signal Process. 2014;62(19):4931–46.
11. Musa O, Hannak G, Goertz N, Efficient recovery from noisy quantized compressed sensing using generalized approximate message passing. In: IEEE 7th international workshop on computational advances in multi-sensor adaptive processing (CAMSAP), Curacao; 2017. p. 1–5.

# ISAR Image Formation of Maneuvering Target via Exploiting the Fractional Fourier Transformation

Yong Wang(✉) and Jiajia Rong

Research Institute of Electronic Engineering Technology,
Harbin Institute of Technology, Harbin 150001, China
Wangyong6012@hit.edu.cn

**Abstract.** Inverse synthetic aperture radar (ISAR) imaging based on the range-Doppler (RD) imaging algorithm has been proved to be effective at the case of the target moves smoothly. However, for the maneuvering target, due to the Doppler frequency of the received signals, which can be regarded approximately as chirp signals, is time-varying, the conventional RD imaging algorithm is not appropriate anymore. In consideration of the fractional Fourier transform (FRFT) having an advantage in concentrating the energy of the chirp signal, a novel technique of ISAR image formation of the maneuvering target through exploiting the FRFT is proposed. Different from the traditional Wigner–Ville distribution (WVD) and other bilinear time–frequency distributions, the FRFT is a linear operator so that it cannot be affected by the cross-terms for the multi-component signals. The validity of the introduced method is validated with the results of ISAR imaging for simulated and real data.

**Keywords:** Inverse synthetic aperture radar (ISAR) · Maneuvering target · Fractional Fourier transform (FRFT) · CLEAN technique

## 1 Introduction

ISAR technique is widely investigated for its capability of forming the target's radar images with high resolution [1–3]. It could be supposed that the target moves uniformly within the observation time, the Doppler frequency shift of the received signals is constant and the classical method RD technique could be utilized to obtain 2-D high-resolution image. Specifically, high-range resolution is achieved by the wideband transmitted signal, while the high cross-range resolution is obtained via the accumulation of echoes in the azimuth domain [4]. However, when the target has a maneuvering flight, the Doppler frequency becomes time-varying, and namely the assumption that the conventional RD method depends on is invalid. Therefore, the RD algorithm cannot be applied to achieve high-quality images anymore and the alternative method named as the range-instantaneous-Doppler (RID) algorithm has been presented consequently. For the RID technique, the time–frequency transform is applied to replace the Doppler analysis, mainly including the time–frequency representation (TFR) [4],


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_82

the Wigner–Hough transform [5], and the signal decomposition technique [6]. Here, a novel algorithm for ISAR image formation of the target with non-uniform movement by exploiting the FRFT is proposed. This algorithm is effective especially under the condition that the backscattered signal of each scatterer can be modeled as the chirp signal, and the FRFT is adopted to achieve the parameters of a chirp signal. For multicomponent signals, the CLEAN technique is utilized to separate the different components. For its linear property, the FRFT cannot be affected by the interferences for the multicomponent signals [7], which indicates that the FRFT has an advantage compared with the Wigner–Ville distribution in the ISAR image formation. Results of ISAR imaging for simulated and real data illustrate the performance of the novel method.

## 2 The Date Model of Non-uniform Rotating Scatterers

Consider a special scatterer as an example, and the cross-range coordinate is $y_i$. Assumed that the scatterer is rotating non-uniformly, and the initial velocity is $c_0$, the acceleration is $c_1$. The radial velocity $v_r(t)$ can be represented as:

$$v_r(t) = (c_0 + c_1 t) y_i \tag{1}$$

Therefore, the Doppler frequency is:

$$\omega_i(t) = \frac{4\pi}{\lambda} \quad v_r(t) = \frac{4\pi}{\lambda} y_i (c_0 + c_1 t) \tag{2}$$

where $\lambda$ denotes the wavelength. Specially, if $c_1 = 0$, Eq. (2) turns to the circumstance of smooth rotation, and the Doppler is constant.

It can be assumed that there are $k$ scatterers in a range cell, the complex envelope for the backscattered signal of $i$th scatterer is $A_i\,(i = 1, 2, \ldots k)$, and then the backscattered signal within the range bin is the sum of the backscattered signals from all scatterers:

$$s(t) = \sum_{i=1}^{k} A_i \exp\left[ j \frac{4\pi}{\lambda} y_i (c_0 t + c_1 t^2) \right] \tag{3}$$

Equation (3) denotes the echo of rotation with acceleration, and it can be seen that it is the sum of complex chirp signals. Therefore, the FRFT combined with the CLEAN technique is utilized to achieve the estimated value of multiple chirp signals with the aim to yield the clear ISAR images simultaneously.

# 3 The FRFT and Its Application in ISAR Imaging

## 3.1 The Definition of FRFT

The FRFT of the signal $s(t)$ is shown in [8] as

$$s^{\ell}(\mu) = \int_{-\infty}^{+\infty} K_{\ell}(t,\mu)s(t)\mathrm{d}t \tag{4}$$

and

$$K_{\ell}(t,\mu) = \begin{cases} \sqrt{\frac{1-j\cot\vartheta}{2\pi}}e^{j\left(\frac{1}{2}\mu^2\cot\vartheta-\mu t\csc\vartheta+\frac{1}{2}t^2\cot\vartheta\right)} & \vartheta \neq n\pi \\ \delta(t+\mu) & \vartheta = (2n+1)\pi \\ \delta(t-\mu) & \vartheta = 2n\pi \end{cases} \tag{5}$$

where $\ell$ denotes the FRFT order, $\vartheta = \pi\ell/2$ denotes the rotation angle, and $n$ is an integer.

## 3.2 The Application of the FRFT for the Chirp Signals

For a chirp signal, when $\vartheta$ is identical with the slope angle on the time–frequency plane, the FRFT gets the maximum value. This is the principle for the FRFT to detect chirp signals and estimate the parameters of them, which can be illustrated by the relationship of FRFT and the Radon–Wigner transform (RWT), as shown in Fig. 1.

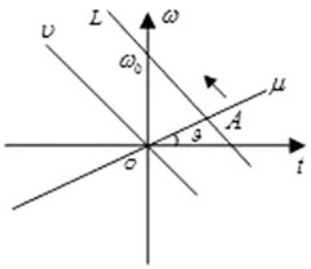



**Fig. 1.** Illustration of RWT

The RWT of the signal $s(t)$ is given as

$$D_s(\mu,\vartheta) = \int_{-\infty}^{+\infty} W_s(\mu\cos\vartheta - v\sin\vartheta, \mu\sin\vartheta + v\cos\vartheta)\mathrm{d}v \tag{6}$$

where $W_s(\cdot)$ is the WVD of the signal $s(t)$.

For a LFM signal $\exp\left[j(mt^2/2+\omega_0 t)\right]$, the RWT gets the maximum value at $(m,\omega_0)$.

Let

$$m=-\cot\vartheta,\quad \omega_0=\mu/\sin\vartheta \tag{7}$$

owing to

$$t=\mu\cos\vartheta-v\sin\vartheta \tag{8}$$

We obtain

$$\mathrm{d}v=-\csc\vartheta\mathrm{d}t \tag{9}$$

Therefore,

$$\begin{aligned} D_s(\mu,\vartheta)&=\csc\vartheta\int_{-\infty}^{+\infty}W_s(t,\omega_0+mt)\mathrm{d}t\\ &=\csc\vartheta\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}W_s(t,\mu)\delta(\omega-\omega_0-mt)\mathrm{d}t\mathrm{d}\omega \end{aligned} \tag{10}$$

According to the WVD of the LFM signal $\exp\left[j(mt^2/2+\omega_0 t)\right]$, we obtain

$$D_s(\mu,\vartheta)=2\pi|\csc\vartheta|\left|\int_{-\infty}^{+\infty}s(t)\mathrm{e}^{-j(\frac{1}{2}mt^2+\omega_0 t)}\right|^2 \tag{11}$$

Based on the definition of FRFT, we obtain

$$\begin{aligned} |s^p(\mu)|^2&=\left|\sqrt{\frac{1-j\cot\vartheta}{2\pi}}\right|^2\left|\int_{-\infty}^{+\infty}s(t)\mathrm{e}^{j(\frac{1}{2}t^2\cot\vartheta-\mu t\csc\vartheta)}\mathrm{d}t\right|^2\\ &=\frac{1}{(2\pi)}|\csc\vartheta|\left|\int_{-\infty}^{+\infty}s(t)\mathrm{e}^{-j(\frac{1}{2}mt^2+\omega_0 t)}\mathrm{d}t\right|^2\Bigg|_{\substack{m=-\cot\vartheta\\ \omega_0=\mu/\sin\vartheta}} \end{aligned} \tag{12}$$

From Eq. (12), it is obvious that the chirp signal parameters can be evaluated through the FRFT. The procedure is as follows:

Step 1: $\ell \in [-1, 1]$, and $\vartheta \in [-\pi/2, \pi/2]$; find the peak of FRFT of the signal and note the peak position as $(\mu, \vartheta)$.

Step 2: The parameters of the chirp signal can be calculated by

$$m = -\cot\vartheta = -\cot(\ell\pi/2) \tag{13}$$

$$\omega_0 = \mu/\sin\vartheta = \mu/\sin(\ell\pi/2) \tag{14}$$

### 3.3 The Application of the FRFT in ISAR Imaging

The technique to detect the multicomponent chirp signal and estimate the parameters of its sub-signal has been presented. But in reality, different sub-signals usually have different energies, and sub-signals with weak energy can often be submerged by sub-signals with strong energy. To deal with it, a new imaging technique using the FRFT method combined with the CLEAN method is proposed. The detailed process is shown as:

Step 1: Initialize $i$ that represents the number of the scatterers in Eq. (3) by letting $i = 1$, and $s(t)$ denotes the data of each range cell as Eq. (3).

Step 2: Calculate the FRFT of the signal $s(t)$, then find the position of the peak of it, and assume it is $(\mu, \vartheta)$. Thereby, the coefficients of the $i$th component are $\hat{m}_i = -\cot\vartheta$, $\hat{\omega}_{0i} = \mu/\sin\vartheta$.

Step 3: Construct a band-pass filter around $\hat{\omega}_{0i}$ to remove the narrow spectrum, then execute the inverse Fourier transform, and the $i$th LFM component can be separated as

$$s_i(t) = \text{IFFT}\{\text{window}_i(\omega) \cdot \text{FFT}[s(t)z_i(t)]\} \cdot z_i^*(t) \tag{15}$$

where $\text{window}_i(\omega) = \begin{cases} 1 & \omega_1 < \omega < \omega_2 \\ 0 & \text{else} \end{cases}$ is a band-pass filter to pass the narrow spectrum whose center frequency is within the frequency range from $\omega_1$ to $\omega_2$, and FFT[·] and IFFT[·] stand for the fast Fourier transform (FFT) operator and inverse FFT operator, respectively. Moreover, $z_i(t) = \exp\left(-j\hat{m}_i t^2/2\right)$.

Step 4: Compute the residual signal as

$$s_{ri}(t) = \text{IFT}\{[1 - \text{window}_i(\omega)]\text{FFT}[s(t)z_i(t)]\} \cdot z_i^*(t) \tag{16}$$

Step 5: Let $s(t) = s_{ri}(t)$, and repeat step 1 to step 3 as long as the termination condition that the energy of the remaining signal is bigger than a threshold is satisfied.

With the aim to obtain the ISAR image of the maneuvering target, after translation compensation of the echo signals, the FRFT could be adopted as an efficient tool to yield the clear radar images.

## 4 Experimental Results

The simulation results through the simulated and real data are given to verify the effectiveness of the new method.

### 4.1 The Simulation Based on B727 Data

The B727 data has 256 pulses, the bandwidth is 150 MHz, and the center frequency is 9 GHz. The translation compensation has been executed on the backscattered signals. Figure 2a is the WVD of a range resolution bin without separating operation, and the interferences are obvious. Figure 2b is the WVD of each component separated by the FRFT, and the interferences have been reduced greatly.

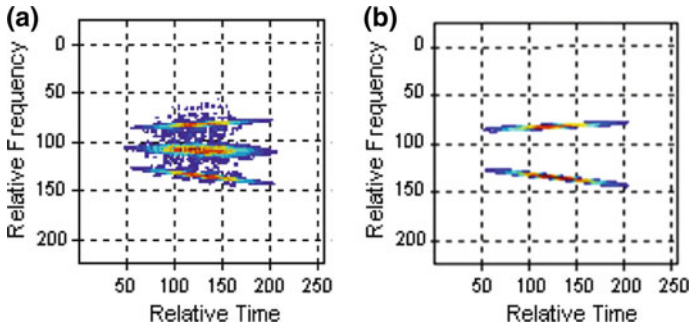



**Fig. 2.** Time–frequency distributions of a range resolution bin: **a** WVD of a range resolution without separating operation and **b** WVD of each component separated by the FRFT

Figure 3 is the ISAR images of B727. Figure 3a denotes the ISAR image with the RD technique, while Fig. 3b, c is obtained with the FRFT algorithm presented in this paper. It is obvious that the image acquired with RD algorithm is blurred seriously and the resolution of the ISAR image acquired with FRFT is improved prominently. What's more, the gesture change of the target is illustrated clearly by comparing the images in Fig. 3b, c.

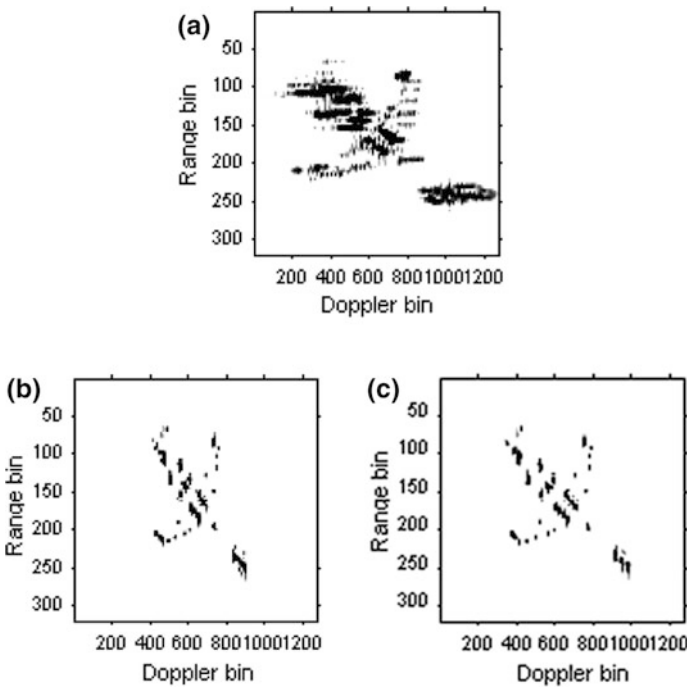



**Fig. 3.** ISAR images of B727: **a** the image yielded by the RD method. **b** The instantaneous image yielded by the FRFT algorithm at $t = 0$. **c** The instantaneous image yielded by the FRFT algorithm at $t = 0.0032$ s

### 4.2 The Experiments Based on the Raw Data of Yark-42 Plane

Part of the raw data of Yark-42 plane is selected. The ISAR image in Fig. 4a is generated by the RD technique while that in Fig. 4b, c is generated by the FRFT algorithm. Figure 4a shows that the Yark-42 plane cannot be recognized in the ISAR image formed with the RD technique. On the contrary, the images in Fig. 4b, c are much clearer and the outline of the plane can be identified readily.

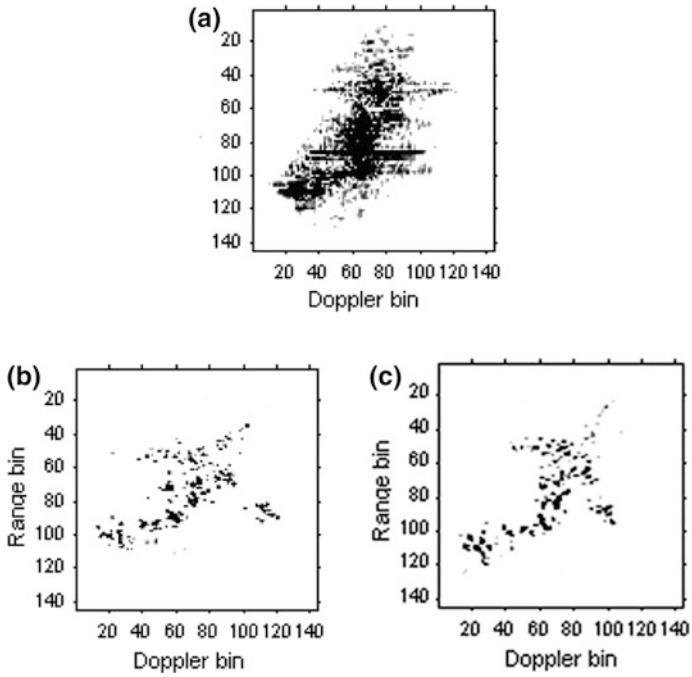



**Fig. 4.** ISAR images of Yark-42 plane: **a** The image yielded by the RD method. **b** Instantaneous image yielded by the FRFT method at $t = 1.344$ s. **c** Instantaneous image yielded by the FRFT algorithm at $t = 2.496$ s

### 4.3 The Experiments Based on the Raw Ship Data

The raw data of a ship is utilized here. The radar works at X-band, and the experiment results are listed in Fig. 5, where Fig. 5a denotes the image through the RD method and Fig. 5b, c shows the instantaneous image through the FRFT approach. The conclusion we can get from Fig. 5 is similar to that obtained from Fig. 4; namely, the proposed method improves the performance greatly when compared with RD algorithm.

## 5 Conclusions

An ISAR image formation method of the maneuvering target is proposed. The echo signal is modeled as a multicomponent chirp signal, and the FRFT method incorporating the CLEAN technique is adopted to evaluate the coefficients of each component.

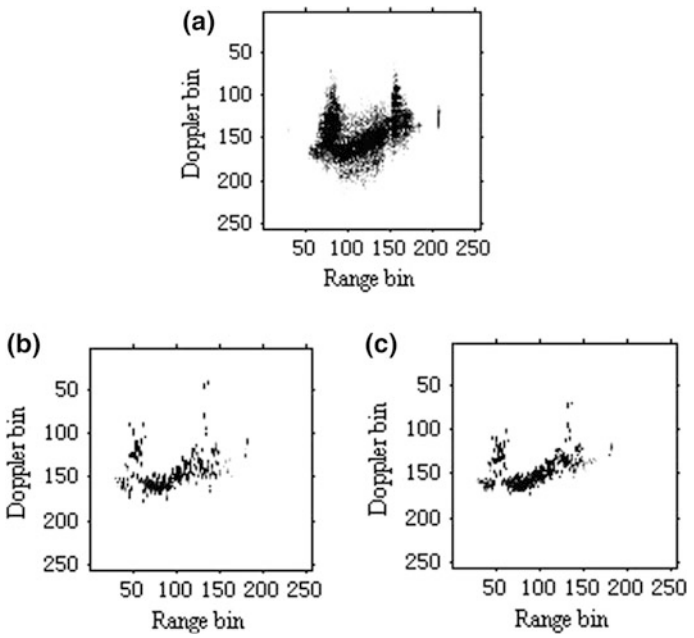



**Fig. 5.** ISAR images of a ship target: **a** the image yielded by the RD method. **b** Instantaneous image yielded by the FRFT algorithm at $t = 0.039$ s. **c** Instantaneous image yielded by the FRFT algorithm at $t = 0.39$ s

The results based on the simulated and real data illustrated the superiority of the technique in this paper.

## References

1. Sparr T. ISAR-Radar imaging of targets with complicated motion. In: International conference on image processing; 2004. p. 5–8.
2. Berizzi F, Mese ED, Diani M, Martorella M. High-resolution ISAR imaging of maneuvering targets by means of the range instantaneous Doppler technique: modeling and performance analysis. IEEE Trans Image Process. 2001;10(12):1880–90.
3. Xing M, Wu R, Bao Z, High resolution ISAR imaging of high speed moving targets. IEEE Proc-Radar Sonar Navig. 2005;152(2):58–67.
4. Chen VC, Qian S. Joint time-frequency transform for radar range-doppler imaging. IEEE Trans Aerosp Electron Syst. 1998;34(2):486–99.
5. Bao Z, Wang GY, Luo L. Inverse synthetic aperture radar imaging of maneuvering targets. Opt Eng. 1998;37(5):1582–8.
6. Wang GY, Xia XG, Chen VC. Three-dimensional ISAR imaging of maneuvering targets using three receivers. IEEE Trans Image Process. 2001;10(3):436–47.
7. Sun HB, Liu GS, Gu H, Su WM. Application of the Fractional Fourier transform to moving target detection in airborne SAR. IEEE Trans Aerosp Electron Syst. 2002;38(4):1416–24.
8. Namias V. The fractional Fourier transform and its application in quantum mechanics. J Inst Math Its Appl. 1980;25:241–65.

# Ship Target Detection in High-Resolution SAR Images Based on Information Theory and Harris Corner Detection

Haijiang Wang(✉), Yuanbo Ran, Shuo Liu, Yangyang Deng, and Debin Su

College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China
whj@cuit.edu.cn

**Abstract.** In order to make up the shortcomings of the traditional CFAR detection algorithm, a ship target detection algorithm based on information theory and Harris corner detection for SAR images is proposed in this paper. Firstly, the SAR image is pretreated, and next, it is divided into superpixel patches by using the improved SLIC superpixel generation algorithm. Then, the self-information value of the superpixel patches is calculated and the threshold $T_1$ is set to select the candidate superpixel patches. And then, the extended neighborhood weighted information entropy growth rate threshold $T_2$ is set to eliminate false alarm detection results of the candidate superpixel patches. Finally, the Harris corner detection algorithm is used to process the detection result, the number of the corner threshold $T_3$ is set to filter out the false alarm patches, and the final SAR image target detection result is obtained. The effectiveness and superiority of the proposed algorithm are verified by comparing the proposed method with the results of CFAR detection algorithm combining with morphological processing algorithm and information theory combining with morphological processing algorithm on the experimental high-resolution ship SAR images.

**Keywords:** SAR image · Ship detection · CFAR · Superpixel · Information theory and Harris corner

## 1 Introduction

Synthetic aperture radar (SAR) imaging is not limited by weather, illumination, or other conditions. To carry out research on ship SAR target detection is very important for marine monitoring management and timely military information acquisition [1].

There are many methods for ship target detection, but the CFAR detection algorithm is most widely used [2]. Scholars have proposed some detection methods based on different statistical distribution models of sea clutter in SAR images, including the Rayleigh distribution model, Weibull distribution model, and so on [3].

CFAR detection methods are based on pixel-level detection, and they have some shortcomings. After the concept of superpixel is proposed, the target detection in SAR


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_83

images based on superpixel has developed rapidly [4]. It uses the superpixel instead of the pixel as the processing unit, and it can obtain a better detection result [5]. Some researchers have used the method based on multiscale variance weighted entropy method to detect the ship targets in the complex background SAR images [6], it can detect the true targets successfully, and however, the original morphology of the ship targets cannot be well preserved.

Aiming to make up the shortcomings of some existing ship target detection algorithms for high-resolution SAR images, a ship target detection algorithm based on information theory and Harris corner detection for high-resolution SAR images is proposed in this paper.

## 2 Detection Algorithm in This Paper

The proposed algorithm consists of three stages, including global detection, local detection, and Harris corner detection, as illustrated in Fig. 1.

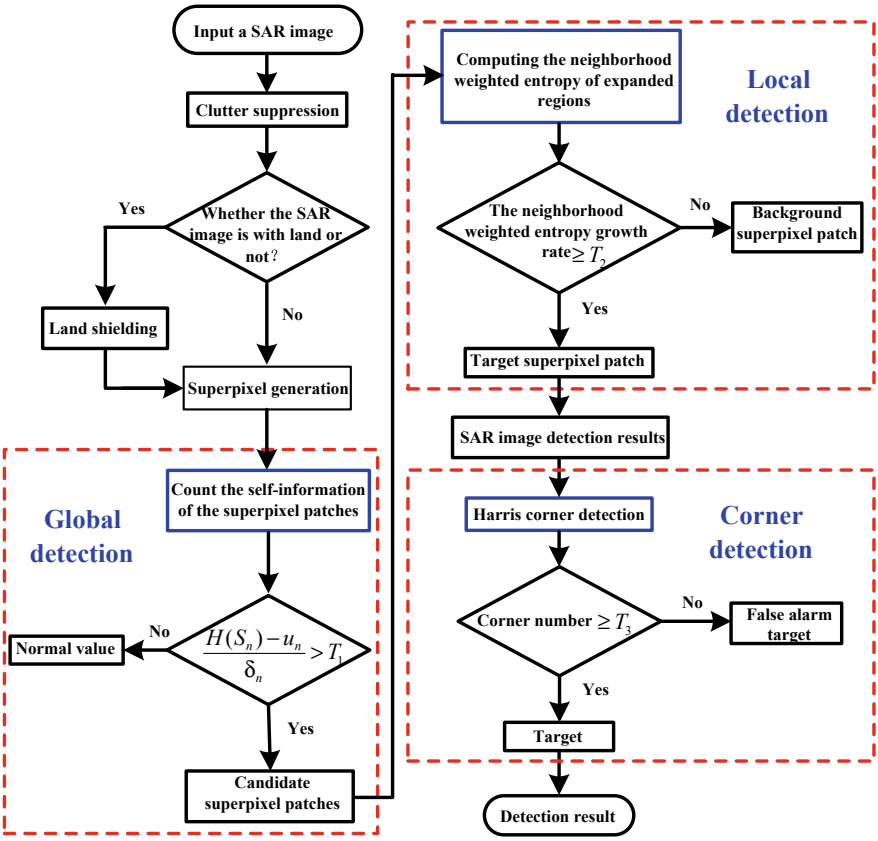



**Fig. 1.** Flowchart of the proposed ship target detection algorithm

### 2.1 Superpixel Generation Algorithm

At present, the superpixel generation algorithms such as watershed algorithm mean shift algorithm and K-means algorithm are widely used in the optical image processing. For ship target detection in SAR images, the subsequent processing will be easier if the edge of the superpixel patches can coincide or almost coincide with the edge of the targets in the SAR image. The simple linear iterative clustering (SLIC) superpixel generation algorithm [7] which is based on clustering can meet the requirement mentioned above. However, because the SLIC algorithm is used for optical image processing, without considering that the SAR image is the gray image, so the color distance of an optical image must be adjusted to the gray scale. The other steps are the same as the SLIC algorithm.

### 2.2 Self-information of SAR Image

In the information theory, the self-information of a random event $e_{\mathrm{i}}$ is defined as [8]:

$$H(e_i) = -\log_m p(e_i) \tag{1}$$

Among the expression, $p(e_{\mathrm{i}})$ is the occurrence probability of the random event $e_{\mathrm{i}}$, and the variable $i$ represents the $i$th possible event.

For any superpixel patch $S_n$ in SAR image $I$, assuming that the superpixel patch is composed of $a \times a$ pixels, the intensity distribution model of $S_n$ can be expressed by the following formula.

$$\begin{pmatrix} r \\ p(r) \end{pmatrix} = \begin{pmatrix} 0, & 1, & \ldots, & 255 \\ p(0), & p(1), & \ldots, & p(255) \end{pmatrix} \tag{2}$$

Among the expression, the variable $p(r)$ represents the probability of the pixel whose grayscale intensity value is $r$. The likelihood function is used to express the probability of each superpixel patch. Its conditional probability value is:

$$P(S_n|\mathrm{I}) = \prod_{k=1}^{a^2} P_I(S_n(k)) \tag{3}$$

In the expression, $P_I(\bullet)$ is the probability distribution of the SAR image I, and $\mathrm{S}_n(k)$ is the corresponding gray value of the $k$th pixel. Because the order of pixels in superpixel patches is not considered, the upper expression can be rerepresented by the probability distribution of superpixel patch $S_n$.

$$P(S_n|\mathrm{I}) = \prod_{i=1}^{255} \left( P_I(i)^{P_{Sn}(r)\cdot a^2} \right) \tag{4}$$

Substituting Eq. (4) to Eq. (1), the corresponding self-information of the superpixel patch $S_{\mathrm{n}}$ can be obtained, as illustrated in the following expression:

$$\begin{aligned} H(S_n|\mathrm{I}) &= -\log_m(P(S_n|\mathrm{I})) = -\log_m\left(\prod_{r=0}^{255}(P_I(r)^{P_{S_n}(r)\cdot a^2}\right) \\ &= -a^2\cdot\sum_{r=0}^{255}(P_{S_n}(r)\cdot\log_m(P_I(r))) \end{aligned} \tag{5}$$

A definition is made that if $P_{S_n}(r)\neq 0$, $P_{S_n}(r)=1$. Then, it can be obtained that

$$H^{'}(S_n|\mathrm{I}) = -a^2\cdot\sum_{r=0}^{255}\log_m(P_I(r)) \tag{6}$$

In this way, the self-information of any superpixel patch $S_n$ can be calculated by using the probability distribution of the grayscale intensity value for the whole image.

## 2.3 Outlier Detection and False Alarm Filtering

### 2.3.1 Outlier Detection

The self-information of all the superpixel patches is counted, and the corresponding distribution histogram is established firstly. The outlier detection expression for the self-information of the superpixel patches is shown as follows:

$$\begin{cases} \frac{H(S_n)-u_n}{\delta_n} > T_1, & \mathrm{n}\in(1,2,\ldots w), H(S_n) \text{ is outlier} \\ \frac{H(S_n)-u_n}{\delta_n} < T_1, & \mathrm{n}\in(1,2,\ldots w), H(S_n) \text{ is normal} \end{cases} \tag{7}$$

In the expression, $H(S_\mathrm{n})$ is the self-information of the superpixel $S_n$ and assume that the number of superpixel patches is $w$. The variables $u_n$ and $\delta_n$ are, respectively, the mean and variance of the histogram distribution of the self-information corresponding to the n superpixel patches, and the variable $T_1$ is the detection threshold. After the global detection, there may be some background superpixel patches in the candidate superpixel patches, and they can be filtered by using the following methods.

### 2.3.2 The False Alarm Filtering Method Based on Weighted Information Entropy

By using $S_T$ as the center, the four extended $3\times 3$ neighborhoods $S_T(d)$ in the vertical direction, horizontal direction, right diagonal direction, and left diagonal direction can be obtained. The extended neighborhood areas are shown in Fig. 2.

The information entropy reflects the average amount of information in the image, and it can be expressed as:

$$H = -\sum_{r=0}^{255}p(r)\log p(r) \tag{8}$$

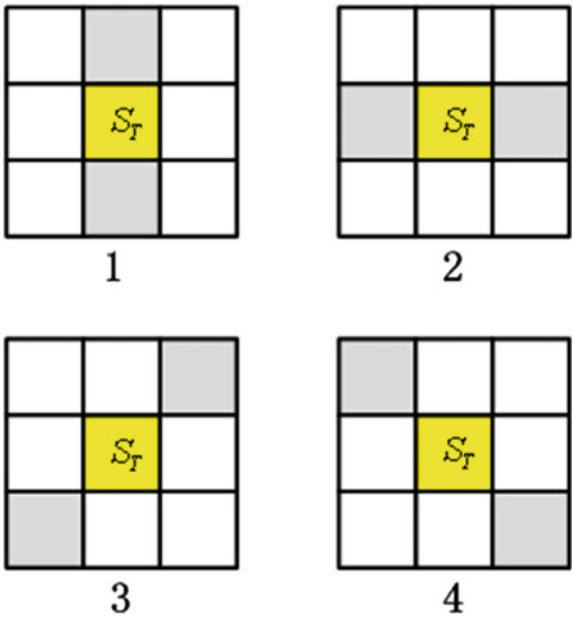



**Fig. 2.** Four extended neighborhoods

When $p(r) = 0$, let $p(r)\ \log p(r) = 0$. Among the expression, $p(r)$ represents the appearance probability of the pixels whose grayscale intensity value is $r$. The weighted information entropy of the testing superpixel patch $S_T$ can be expressed as follows:

$$E(S_T) = -\sum_{r=0}^{255} (r - \overline{r_V})^2 P_{S_T}(r) \log_2(P_{S_T}(r)) \tag{9}$$

The variable $\overline{r_V}$ is the mean grayscale intensity value. The weighted information entropy $E(S_T(d))$ of the extended neighborhood $S_T(d)$ in four directions can be expressed as follows:

$$E(S_T(d)) = -\sum_{r=0}^{255} (r - \overline{r_{ER}(d)})^2 \cdot P_{S_T(d)}(r) \log_2(P_{S_T(d)}(r)) \tag{10}$$

In the equation, $\overline{r_{ER}(d)}$ represents the mean grayscale intensity value, and $P_{S_T(d)}(r)$ is the probability of the pixel whose grayscale intensity value is $r$. The weighted entropy difference between the candidate superpixel patch and the extended neighborhood is:

$$V(S_T) = \arg\min(E(S_T(d))) - E(S_T) \tag{11}$$

Then, the growth rate of weighted entropy can be represented by $\frac{V(S_T)}{E(S_T)}$. Compare it with the second threshold $T_2$ to determine whether the candidate superpixel patch belongs to the target area or the background area, so that the false alarm superpixel patches can be removed.

$$\begin{cases} \frac{V(S_T)}{E(S_T)} > T_2, & S_T \text{ belongs to the target area} \\ \frac{V(S_T)}{E(S_T)} < T_2, & S_T \text{ belongs to the background} \end{cases} \tag{12}$$

#### 2.3.3 The False Alarm Filtering Method Based on Harris Corner Detection

Harris corner detection algorithm is often used in optical image detection; however, it is seldom used for target detection in SAR images [9]. The high-resolution marine ship targets often have some corners, and they can be detected. The Harris corner algorithm is used to detect the corner of ship targets in high-resolution SAR images. Then, the corner number threshold $T_3$ is set to further filter out the false alarm patches and the final target detection result in SAR images can be obtained.

## 3 Simulation Experiment and Result Analysis

A high-resolution marine ship SAR image with land is selected, as shown in Fig. 3. The resolution is 1.25 m $\times$ 1.25 m, and its size is 961 $\times$ 762; moreover, the SAR image contains 17 ships.

**Fig. 3.** Original SAR image

Firstly, Otsu's method combining with the morphological processing method is used to shield the land area in the ship SAR image [10], and the land shielding result is shown in Fig. 4a. Then, the improved SLIC superpixel segmentation algorithm is used to process it. The STEPSIZE of the patch area is set as 13 pixels, and the superpixel generation result is shown in Fig. 4b.

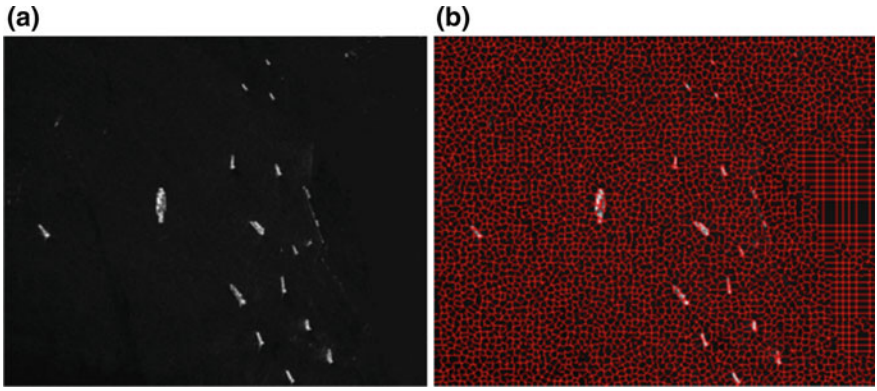



**Fig. 4.** **a** Land shielding result and **b** superpixel segmentation result

The self-information of the superpixel patches is calculated, then set the outlier parameter $T_1$ as 1.4, and the candidate superpixel patches are shown in Fig. 5a. There are some false alarm superpixel patches in the graph. Then, set the growth rate of the neighborhood weighted information entropy $T_2$ as 20%, and the detection result is shown in Fig. 5b.

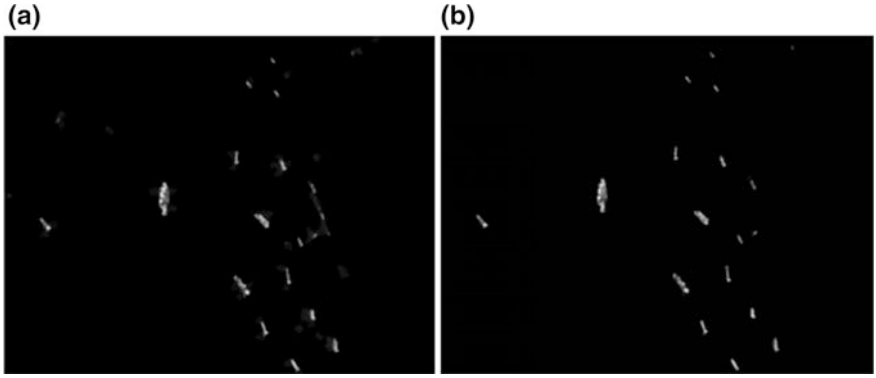



**Fig. 5.** **a** Candidate superpixel patches and **b** the weighted entropy filtering result

Carry out the Harris corner detection on the filtering result as shown in Fig. 5b, and the detection result is shown in Fig. 6a. Then, set the threshold $T_3$ as 1; after filtering out the two false alarm targets, the final detection result can be obtained as shown in Fig. 6b.

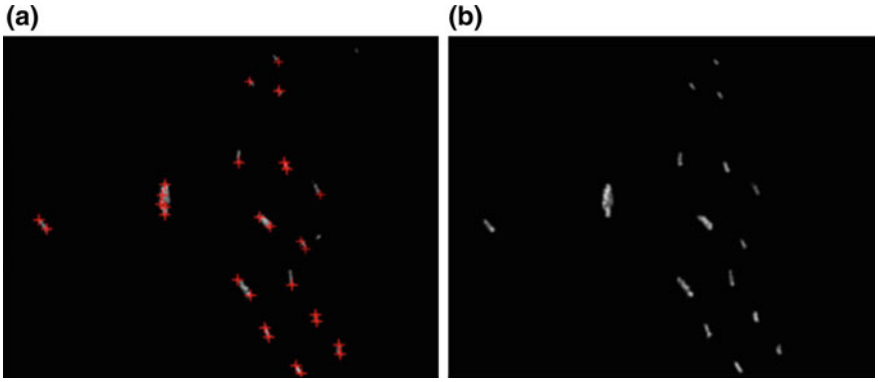



**Fig. 6.** **a** Harris corner detection result and **b** the final detection result

Compare the detection performance of the proposed algorithm with other two detection algorithms, and all experiments are conducted in the same MATLAB R2014a environment and in the same HP Pavilion G4 notebook computer. The detection results based on two detection algorithms for comparison are shown in Fig. 7.

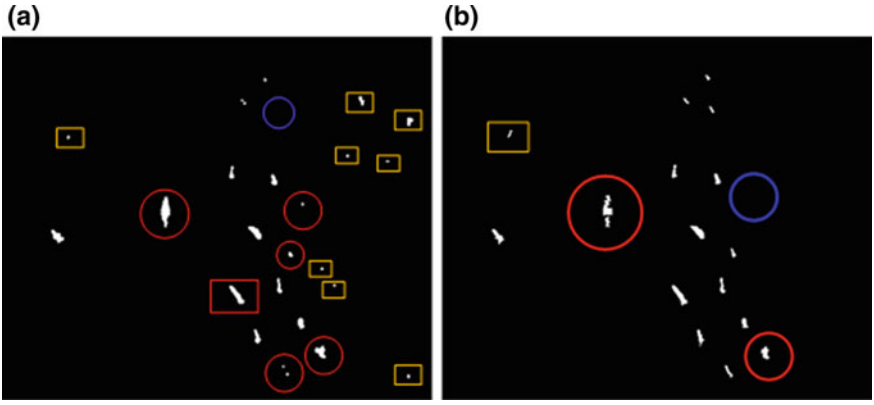



**Fig. 7.** **a** Detection result based on CFAR combined with morphological processing and **b** detection result based on information theory combined with morphological processing

In Fig. 7, the correctly detected target and the false alarm targets are marked by the red rectangle box and the yellow rectangle box, respectively. The targets whose form is

seriously distorted are marked by the red round box, and the missing targets are marked by the blue round box in their corresponding position of the original ship targets. The figure of merit (FoM) is defined as a quality factor of the detection performance, and $\mathrm{FoM} = \frac{N_{\mathrm{tt}}}{N_{\mathrm{gt}} + N_{\mathrm{fa}}}$. In the equation, $N_{\mathrm{tt}}$, $N_{\mathrm{fa}}$, and $N_{\mathrm{gt}}$ represent the number of correctly detected targets, the number of false alarm targets, and the number of real targets, respectively. The newly defined detection distortion rate is the ratio of the distorted target number to the total number of ship targets. The performance of the three detection algorithms is shown in Table 1.

**Table 1.** Comparison of detection performance between the three detection algorithms

| Detection algorithm | Number of correctly detected target | False alarm number | Number of missing targets | Number of distorted targets | Detection distortion rate (%) | FOM | Time (s) |
|---|---|---|---|---|---|---|---|
| CFAR combined with morphology | 14 | 8 | 2 | 5 | 29.41 | 0.560 | 169.3256 |
| Information theory combined with morphology | 15 | 1 | 2 | 2 | 11.76 | 0.833 | 189.6531 |
| The proposed algorithm | 16 | 0 | 1 | 0 | 0 | 0.941 | 131.6881 |

It can be seen that the detection result of the proposed algorithm has a higher detection rate and a lower missing detection rate, and the false alarm rate and detection distortion rate are both lower. In addition, the proposed detection algorithm has a higher operation efficiency, the original shape of the ship targets can be completely retained, and the ship targets in the SAR image can be generally located.

## 4 Conclusion

A ship target detection algorithm based on information theory and Harris corner detection for SAR images is proposed in this paper. A self-information-based global detection and a weighed entropy-based local detection can almost separate the ship targets from the background; moreover, the Harris corner detection algorithm is used to further filter out the false alarm targets. By setting three thresholds and making the subsequent double false alarm filtering, the final detection result can be obtained. The simulation experiments verify the effectiveness and superiority of the proposed algorithm. In subsequent further research, we will combine other features of the ship targets to improve the algorithm for ship target detection in SAR images with the larger scene and richer information.

## References

1. Yeremy M. Ocean surveillance with polarimetric SAR. Can J Remote Sens. 2001;27 (4):328–44.
2. An W, Xie C, Yuan X. An improved iterative censoring scheme for CFAR ship detection with SAR imagery. IEEE Trans Geosci Remote Sens. 2014;52(8):4585–95.
3. Gandhi PP, Kassam SA. Analysis of CFAR processors in homogeneous background. IEEE Trans Aerosp Electron Syst. 2002;24(4):427–45.
4. Ren X, Malik J. Learning a classification model for segmentation. In: Proceedings of the IEEE international conference on computer vision, IEEE, vol.1; 2003. p. 10–7.
5. Yu W, Wang Y, Liu H, et al. Superpixel-based CFAR target detection for high-resolution SAR images. IEEE Geosci Remote Sens Lett. 2016;13(5):730–4.
6. Wang X, Chen C. Ship detection for complex background SAR images based on a multiscale variance weighted image entropy method. IEEE Geosci Remote Sens Lett. 2017;14(2):184–7.
7. Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans Pattern Anal Mach Intell. 2012;34(11):2274.
8. Cao Z, Ge Y, Feng J. Fast target detection method for high-resolution SAR images based on variance weighted information entropy. EURASIP J Adv Signal Process. 2014;2014(1):45.
9. Harris C. A combined corner and edge detector. Proc Alvey Vision Conf. 1988;1988 (3):147–51.
10. Wang Q. Inshore ship detection using high-resolution synthetic aperture radar images based on maximally stable extremal region. J Appl Remote Sens. 2015;9(1):095094.

# Hidden Markov Model-Based Sense-Through-Foliage Target Detection Approach

Ganlin Zhao[1(✉)], Qilian Liang[1], and Tariq S. Durrani[2]

[1] Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX 76019-0016, USA
ganlin.zhao@mavs.uta.edu, liang@uta.edu

[2] Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, Scotland, UK
durrani@strath.ac.uk

**Abstract.** In this paper, we propose sense-through-foliage target detection approach based on Hidden Markov Models (HMMs). Separate Hidden Markov Models are trained for signals containing target signature and no target (clutter), respectively. Less correlated features are selected as input of Hidden Markov Models for training and testing. Foliage data is collected from three different UWB radar locations, and experimental results show that position 1 data gives the best detection result. All three locations have above 0.8 AUC from the ROC curves.

**Keywords:** Radar target detection · Foliage · UWB · HMM

## 1 Introduction

Target detection is an important topic in both military and civilian research and applications. Clutter echoes returned from non-target objects may contain higher power thus obscure the real target and made the detection difficult. For example, in battleground, enemies could utilize background clutter such as foliage, ocean and other environmental obstacles as covers to avoid detection. Therefore, improving target detection accuracy could potentially eliminate such hostile activities.

Foliage is known as a challenging environment when performing such detection task. Foliage has non-stationary nature since the motion of tree branches and leaves can cause severe doppler shift. Also, multi-path effect exists and may have negative impact on detection performance. In this work, our object is to detect target in the strong foliage clutter using ultra-wideband (UWB) radar. Further, we consider data collected from UWB radar has poor signal quality which is caused by using low-amplitude transmitting pulses during the data collection.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_84

Hidden Markov model (HMM) is a statistical approach used for temporal data modeling such as speech recognition [1,2], machine translation [3], and text classification [4,5]. Since radar signal is also temporal, we propose to apply HMM-based approach to radar target detection using real-world UWB data. We construct HMMs for target and non-target cases using training data and validate the detection results using testing sequences as input of pretrained models.

The rest of this paper is organized as follows: In Sect. 2, we briefly introduce Hidden Markov Models. In Sect. 3, we present data measurement and collection. In Sect. 4, we describe HMM-based target detection approach. In Sect. 5, experiment results and analysis are provided. In Sect. 6, we draw the conclusion.

## 2 Introduction to Hidden Markov Models

HMM consists of two random processes. The first process is on-time dimension which is essentially a first-order Markov chain and at each time $t$ ($t \in \{1, T\}$) the system is in state $q_t$ with $N$ possible finite states choices:

$$S = \{S_1, S_2, \ldots, S_N\} \tag{1}$$

In the second random process, each state produces an observation result which is directly visible. However, the states are "hidden" and not observable. In other words, we cannot directly tell the exact state where the observation occurs. The $T$ observation symbols are

$$O = \{O_1, O_2, \ldots, O_T\} \tag{2}$$

A basic HMM is characterized by three sets of parameters, and we use $v_t$ to represent the observation symbol at time $t$:

1. Initial state probability matrix $\pi = \{\pi_i\}$, where each element $\pi_i = P(q_1 = S_i), 1 \leq i \leq N$ is the probability that the system is in state $i$ at initial time $t = 1$.
2. State transition probability matrix $A = \{a_{ij}\}$, where $a_{ij} = P(q_t = S_j \mid q_{t-1} = S_i), 1 \leq i, j \leq N$ is the probability that the system is in state $j$ at time $t$ given that the system is in state $i$ at time $t - 1$.
3. Observation probability matrix $B = \{b_j(t)\}$, where $b_j(t) = P(v_t = O_t \mid q_t = S_j), 1 \leq j \leq N, 1 \leq t \leq T$ is the probability that the output observation symbol is $O_t$ given the system is in state $j$ at time $t$.

The notation $\lambda = \{A, B, \pi\}$ denotes a complete parameter set of HMM. In addition, if the observation symbol probability density function is discrete, an HMM is called discrete Hidden Markov Model. Otherwise, it is called continuous Hidden Markov Model. For continuous case, the state transition probability density functions are typically represented as a form of Gaussian mixtures [6].

## 3 Sense-Through-Foliage Data Measurement

Sense-through-foliage data set is from Air Force Research Laboratory [7]. For UWB radar data collection, "good"-quality signal refers to the data collection using high-amplitude transmitting pulses and enough reflected signals are averaged to produce each collection. "Poor"-quality signal refers to data collection using low-amplitude transmitting pulses, and fewer reflected signals are averaged to produce each collection. Some previous work based this data set can be found in [8–13].

In this experiment, poor signal quality data set is used for sense-through-foliage target detection task. Multiple radar positions data are available. We choose three radar locations data, and each set contains 35 radar echoes in both target and no target scenarios.

## 4 Hidden Markov Model-Based UWB Radar Target Detection

In this work, we use signals collected by UWB radar to perform this detection task. Firstly, standard Baum–Welch algorithm is applied to both target and no target training sequences to construct HMMs for each case. After models are trained, we use testing sequences to get a likelihood score of each HMM. Based on the obtained likelihood scores, we classify a certain sequence as target or no target and compare to the ground truth. In the training stage, we extract signal features as a vector input of each HMM. Assuming the UWB radar signal is the observation results of an underlying Markov chain state transitions, we evenly divide the signal into $K$ windows which are shown in Fig. 1. Therefore, $K$ corresponds to the number of observations and each window with a corresponding state is assumed. Within each window, time–frequency features are extracted, for example, mean, variance, skewness, kurtosis, entropy, Dickey–Fuller test, peak-to-peak ratio, crest factor, energy, RMS, number of abrupt changes, number of small changes, mean frequency. We use such features to form a feature vector. The feature vector of $i$th window is denoted as $\boldsymbol{X^i} = (X_1^i, X_2^i, \ldots, X_L^i)^T$, and $L$ is the total number of extracted features. To make sure each element of feature vector has comparable magnitude, they are normalized to $[-1, 1]$ before using the Baum–Welch algorithm.

For the obtained feature set of the signal, redundancy features may exist. For example, two features could be mutually correlated and using both of them is not efficient. In this work, we eliminate such redundancy features by looking at their covariance matrix

$$\Sigma_{ij} = cov(X_i, X_j) = \frac{E[(X_i - \mu_{X_i})(X_j - \mu_{X_j})]}{\sigma_{X_i}\sigma_{X_j}} > \psi \tag{3}$$

where $\psi$ is the predefined threshold. Based on this criteria, $l$ features are selected with less cross-correlations. Mutual information can also be used to evaluate feature correlation [14].

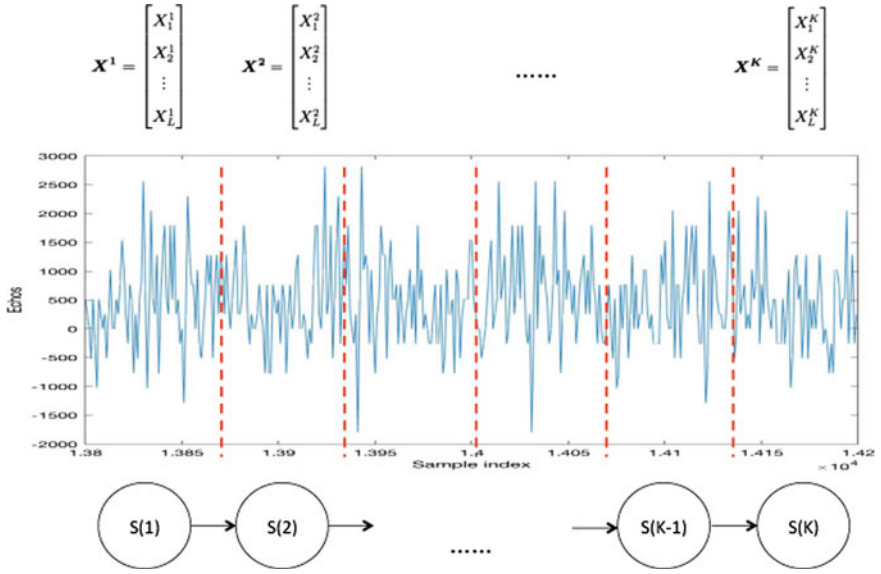



**Fig. 1.** Signal is evenly divided into $K$ windows, and $L$ features are extracted with respect in each window

Selected feature vectors of each signal window $\boldsymbol{X} = \{\boldsymbol{X}^1, \boldsymbol{X}^2, \ldots, \boldsymbol{X}^K\}$ are more effective representation of the original signal thus treated as distinctive observation sequences for HMM training. In this work, we implement continuous HMM assuming that the observation sequences are drawn from a mixture of $M$ $l$-dimensional multivariate Gaussian densities.

For poor quality signal, with a priori knowledge that target on-site location is around 14,000 sample with an approximate duration of 250 samples, we select a wider range of 400 samples around the target location as the area of interest without losing any target information. These 400 samples are subsequently divided into 20 windows. Finally, by comparing cross-correlation of extracted features, we select 13 most non-relevant features to represent the original signal.

At each location, a total number of 35 reflected echoes are collected by the UWB radar. In this experiment, 20 radar echoes are used as training set and left 15 echoes are used as testing set. Two HMMs for target and no target denoted as $\lambda_{Tar}$ and $\lambda_{Notar}$ are trained via Baum–Welch algorithm.

For testing phase, log-likelihood of each sequence $O_{test}$ is calculated using both pretrained models: $L_1 = \log P(O_{test} \mid \lambda_{Tar})$ and $L_2 = \log P(O_{test} \mid \lambda_{Notar})$, the decision rule of judging whether a sequence belongs to target or no target category is the following:

$$H_0 : \frac{L1}{L2} < \delta \quad No\ target \tag{4}$$

$$H_1 : \frac{L1}{L2} > \delta \quad Target \tag{5}$$

where $\delta$ is a predefined threshold. The Baum–Welch training algorithm needs an initial guess of model parameters. These initial parameters can be obtained from uniform distributions. However, different initializations could lead to slightly different trained model parameters which affect the detection probability. Therefore, during the training phase, 100 Monte Carlo simulations are performed to average out the initialization effect.

## 5 Experimental Results

Table 1 shows the confusion matrix of sense-through-foliage target detection results for three different positions. The hypothesis test threshold $\delta$ is chosen to be 1. From Table 1, we can see that position 1 has 88% probability of detection and 0.07% false alarm rate. For position 2 and position 3, lower probability of detection (82.43 and 72%) and higher false alarm rate (24 and 18%) are observed. Figure 2 shows the receiver operating characteristic (ROC) curves of three location detection performance. From the ROC curve, we observe that location 1 has 0.99 area under the curve (AUC), while the other two locations have around 0.80 and 0.83 AUC, respectively. The lower detection performance could result from low radar cross section of these two positions. The experiment is conducted using poor signal quality radar echoes with lower transmitting power. To further improve the detection accuracy, diversity combing schemes could be applied to enhance signal quality.

**Table 1.** Sense-through-foliage target detection results of three radar positions

| Position 1 | | |
|---|---|---|
| | No target | Target |
| No target | 99.03% | 0.07% |
| Target | 11.33% | 88.67% |
| Position 2 | | |
| | No target | Target |
| No target | 76% | 24% |
| Target | 17.57% | 82.43% |
| Position 3 | | |
| | No target | Target |
| No target | 82% | 18% |
| Target | 28% | 72% |

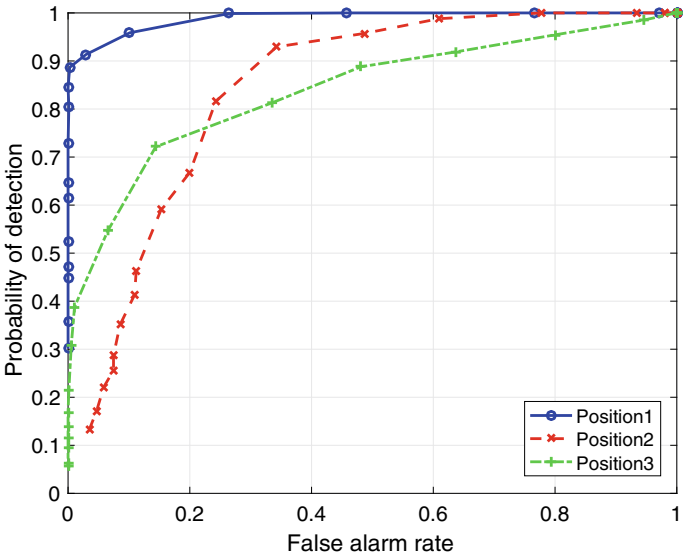



**Fig. 2.** ROC curve of sense-through-foliage target detection results

## 6 Conclusion

In this work, we propose HMM-based sense-through-foliage target detection approach and experiment on real-world UWB radar data. Radar signals are divided into windows for feature extraction. In each window, feature vector serves as distinctive observation for HMM's training and testing. Detection results show that position 1 data gives the best performance in terms of probability of detection and false alarm rate. All three locations radar data have above 0.80 AUC observed from ROC curves.

**Acknowledgments.** This work was supported in part by NSFC under Grant 61731006, 61771342, 61711530132, Royal Society of Edinburgh, and Tianjin Higher Education Creative Team Funds Program.

## References

1. Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE. 1989;77(2):257–86.
2. Lee KF, Hon HW. Speaker-independent phone recognition using hidden Markov models. IEEE Trans Acoust Speech Signal Process. 1989;37(11):1641–8.
3. Hahn S, et al. Comparing stochastic approaches to spoken language understanding in multiple languages. IEEE Trans Audio Speech Lang Process. 2011;19(6):1569–83.
4. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE. 1998;86(11):2278–324.

5. Sauper C, Haghighi A, Barzilay R. Incorporating content structure into text analysis applications. In: Proceedings of the 2010 conference on empirical methods in natural language processing (EMNLP '10). Stroudsburg, PA, USA: Association for Computational Linguistics; 2010. p. 377–87.
6. Gauvain JL, Lee C-H. Maximum a posteriori estimation for multivariate Gaussian mixture observations of Markov chains. IEEE Trans Speech Audio Process. 1994;2(2):291–8.
7. Dill C. Foliage Penetration (Phase II) Field Test: Narrowband versus Wideband Foliage Penetration. Final Report of Contract Number F41624-03-D-7001/04, July 2005 to Feb 2006.
8. Liang J, Liang Q, Samn SW. Foliage clutter modeling using the UWB radar. In: 2008 IEEE international conference on communications, Beijing, 2008; p. 1937–41.
9. Liang Q, Samn SW, Cheng X. UWB radar sensor networks for sense-through-foliage target detection. In: 2008 IEEE international conference on communications, Beijing, 2008; p. 2228–32.
10. Liang Q, Cheng X, Huang S, Chen D. Opportunistic sensing in wireless sensor networks: theory and applications. IEEE Trans Comput. 2014;63(8):2002–10.
11. Liang Q. Situation understanding based on heterogeneous sensor networks and human-inspired favor weak fuzzy logic system. IEEE Syst J. 2011;5(2):156–63.
12. Liang Q, Cheng X, Samn SW. NEW: Network-enabled electronic warfare for target recognition. IEEE Trans Aerosp Electron Syst. 2010;46(2):558–68.
13. Liang Q, Cheng X. KUPS: Knowledge-based ubiquitous and persistent sensor networks for threat assessment. IEEE Trans Aerosp Electron Syst. 2008;44(3):1060.
14. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. 2005;27(8):1226–38.

# Ship Detection via Superpixel-Random Forest Method in High-Resolution SAR Images

Xiulan Tan(✉), Zongyong Cui, Zongjie Cao, and Rui Min

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China
972642539@qq.com

**Abstract.** With the increasing resolution of synthetic aperture radar (SAR), the traditional SAR image target detection methods used for medium-low resolution are not suitable for high-resolution SAR images, which contain detailed information about structure, shape, and weak echoes that are hardly detected in traditional ways. In this paper, we proposed a new method, Superpixel-Random Forest Technique, to detect ships in high-resolution SAR images. The method combines superpixel and random forest algorithms. The superpixel is adopted to divide images into many subregions properly, and the random forest is used for unsupervised clustering these subregions into ships or others. The experimental results show that the algorithm can accurately detect the ship targets.

**Keywords:** SAR · Ship detection · Superpixel · Random forest · Clustering

## 1 Introduction

Synthetic aperture radar (SAR) can provide real-time earth observations, no matter in what kind of weather condition [1]. Different ground objects have different backscatter properties, and SAR images can provide such discriminative features for scene understanding and interpretation. With the increasing resolution, SAR images appear to contain more structure and shape information and weak scatters, which are hard to detect according to traditional ways. Therefore, it is necessary to develop new interpretation algorithms to detect targets efficiently from high-resolution SAR images [2].

As an important maritime application, ship detection has been widely studied. Constant false alarm rate detection (CFAR) is the most traditional SAR image ship detection technology [3]. It is based on the hypothesis testing and attempts to get a self-adaptive threshold on the basis of the background distribution, so that the false alarm rate is maintained at the preset level. CFAR detector performance depends on the design of the sliding window structure, the distribution of the clutter statistical model and the model parameter estimation [4].

In order to detect ship targets accurately, we use the superpixel principle to divide images before clustering and random forest model to measure the rarity of each subregion. The entire working flow of the proposed method is shown in Fig. 1.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_85

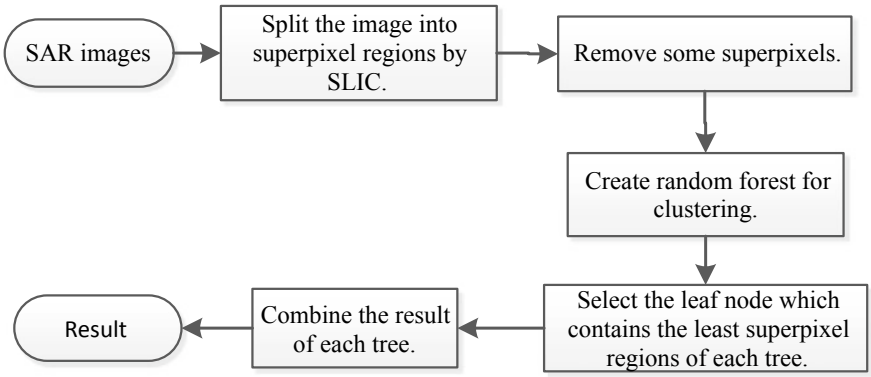



**Fig. 1.** Working flow of the proposed method

## 2 Superpixel Segmentation

Image segmentation is an important step in our method, because we need to process SAR image piece by piece. Superior segmentation algorithms for optical imagery are widely studied, such as mean shift [5] and statistical region merging (SRM) [6]. The simple linear iterative clustering (SLIC) algorithm adopts *k*-means clustering to generate superpixels [7], which is simple to use and understand. The SLIC consists of three basic steps [8]:

(1) Initialization of the cluster centers;
(2) Local *k*-means clustering;
(3) Postprocessing.

Generally, $k$ is the only parameter of the algorithm and also the expected number of superpixels with roughly equal size. The clustering procedure starts with an initialization step where k initial cluster centers $C_i = [l_i\ a_i\ b_i\ x_i\ y_i]^{\mathrm{T}}$ are sampled on a regular grid spaced $S$ pixels apart, in terms of color images in the CIELAB color space [7]. To produce roughly equally sized superpixels, the grid interval is $S = \sqrt{N/k}$. To avoid seeding a center on an edge or a noisy pixel, the center should be moved to the location that is corresponding to the lowest gradient position in a $3 \times 3$ neighborhood.

In the next local k-means clustering step, each pixel $i$ is assigned to the nearest cluster center according to the distance measure. Similar pixels searching is done in a region $2S \times 2S$ around the superpixel center, because an expected superpixel area size is close to $S \times S$.

After each pixel has been assigned to the nearest cluster center, the clustering procedure is completed. However, influenced by the noise, the pixels in the same cluster may be not connective. Therefore, it is necessary to have a postprocessing step to improve connectivity by reassigning the disjointed pixels to nearby superpixels.

Some experiments about superpixels generated by SLIC show that the superpixels can express the image subregions well both in an optical image and a SAR image [7, 8], so we use the SLIC to divide the SAR image for subsequent ship target detection.

## 3 Random Forest

Random forest, a classical ensemble learning model, has been successfully used in many classification and regression problems [9]. In classification regression analysis, the random forest aims to make the elements in each leaf node as pure as possible, and the best condition is that a leaf node contains only one class. Classification and regression are mostly conducted under supervised conditions, which means that a large amount of training data is required. In the unsupervised way, the problem of data volume is not considered any more.

Here, we exploit random forest clustering function to divide the image into several groups and make elements in the same group have similar features. By this way, the rare ship target regions are connected with the number of elements contained in the same group. The smaller the number, the more likely the node will contain the ship targets and vice versa.

In the proposed method, the random forest is formed of a number of trees and each tree clusters the superpixels by randomly choosing features to compare. To reduce the computational burden of constructing forest, some superpixels will be removed before building the forest. Fist, the mean of each superpixel is calculated according to formula

$$m_i = \frac{1}{|X_i|} \sum_{x \in X_i} \sum_{y \in Y_i} I(x, y) \tag{1.1}$$

$m_i$ is the mean of superpixel $i$; $X_i$ and $Y_i$ are coordinate sets of superpixel $i$. $I(x,y)$ is the pixel value at position $(x,y)$. Then, threshold $T$ is set. If the mean of a certain superpixel $m_i$ is greater than $T$, the superpixel is reserved for subsequent random forest correlation calculation. Otherwise, the superpixel is removed.

According to Sect. 2, a SAR image is segmented by SLIC. We store the grayscale values of all the pixels contained in a superpixel in a column vector, and this vector is defined as the corresponding feature of the superpixel. Denote the set of superpixel regions by $S = \{s_1, s_2, \ldots, s_N\}$ and their corresponding feature vectors by $F = \{f_1, f_2, \ldots, f_N\}$, $i = 1, \ldots, N$. Then, we will build a random forest that makes up of several trees based on this feature space. The main process is as follows:

1. Generate two numbers from 1 to $p$ randomly, e.g., $a_1$ and $a_2$; $p$ is the minimum dimension of the feature vectors in $F$.
2. Calculate average feature difference in dimensions $a_1$ and $a_2$ among all the elements in $S$. $F(a_1,i)$ is the value of $F$ in row $a_1$ and column $i$, $|S|$ is the number of elements in node $S$.

$$\delta(a_1, a_2) = \frac{1}{|S|} \sum_{s_i \in S} [F(a_1, i) - F(a_2, i)] \tag{1.2}$$

3. Split the root node $S$ into left and right child nodes, according to the following formula:

$$s_i \in \begin{cases} S_l & \text{if } F(a_1, i) - F(a_2, i) \leq \delta(a_1, a_2) \\ S_r & \text{otherwise} \end{cases} \tag{1.3}$$

   where $S_l$ is left child node and $S_r$ is right child node.
4. Repeat Steps 1, 2, and 3 until reaching the depth of the tree.

Once a tree is built up, we choose the leaf node whose elements number is smallest as the group that may contain ship targets and then synthesize the smallest leaf nodes of other trees as the final test result. Generally, the random forest has a better accuracy than a single tree.

## 4 Results

We use two high-resolution SAR images to validate the efficiency of the proposed method and compare it with the constant false alarm rate (CFAR) detection. All the experiments in the following are conducted on a desktop with a 3.0-GHz CPU and 4.0-GB RAM, and our code is written in Python 3.5.

Figures 2a and 3a show two 3-m-resolution TerraSAR images with size of $226 \times 292, 467 \times 150$ pixels, respectively. They were captured in the Strait of Gibraltar, England. Figure 2b is obtained by segmenting Fig. 2a into 1775 superpixel regions.

Figure 2d is the result of our method. It is shown that all nine ship targets are detected correctly, without any false detection. However, CFAR has some error detection due to the effect of background clutter, as shown in Fig. 2c.

In the previous, the effect was pretty good when the proposed method is applied to a simple sea scenario. In the following, the experiment is more challenging because the sea background is more complicated and there are several land areas. Figure 3b shows 1923 superpixel regions generated by SLIC. All 4 ships are detected by our method in Fig. 3d, without any noise points. However, the result of CFAR appears too much error detection as shown in Fig. 3c.

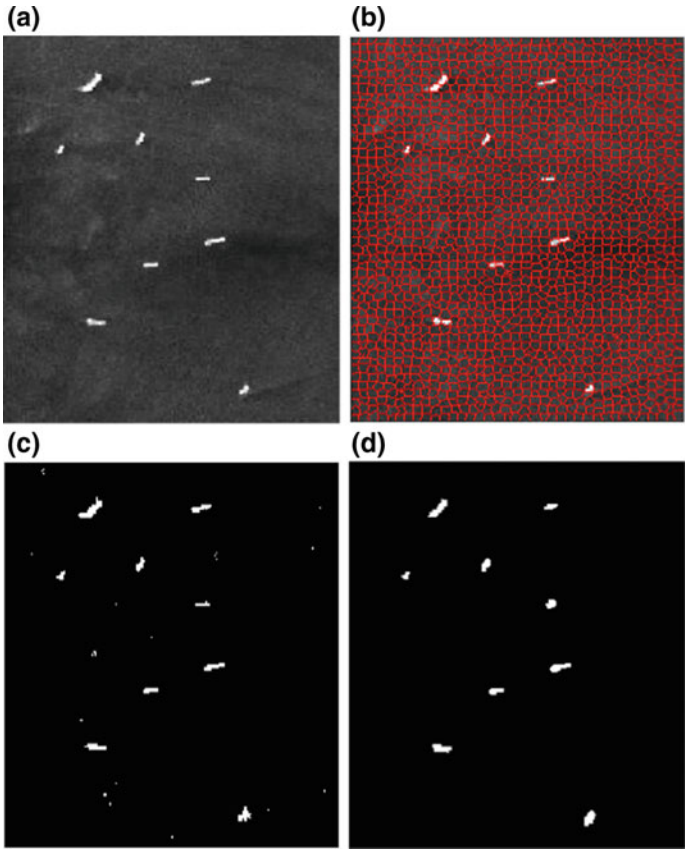



**Fig. 2.** **a** Strait of Gibraltar SAR image 1; **b** result of superpixel segmentation; **c** result of CFAR based on Rayleigh distribution model ($P_{fa} = 0.001$); **d** result of the proposed method

## 5 Conclusion

In this paper, we put forward a method, which is suitable for ship target detection in high-resolution SAR images via superpixel and random forest. The method firstly produces superpixels using SLIC and then establishes the random forest to cluster ship targets. Experiments above show that our method can detect the ship targets accurately. However, when the number of ships increases dramatically and the ships are densely distributed, the performance of the proposed method will decrease because a superpixel may contain both the ship and the ocean background, and a ship may be divided into more than one superpixel.

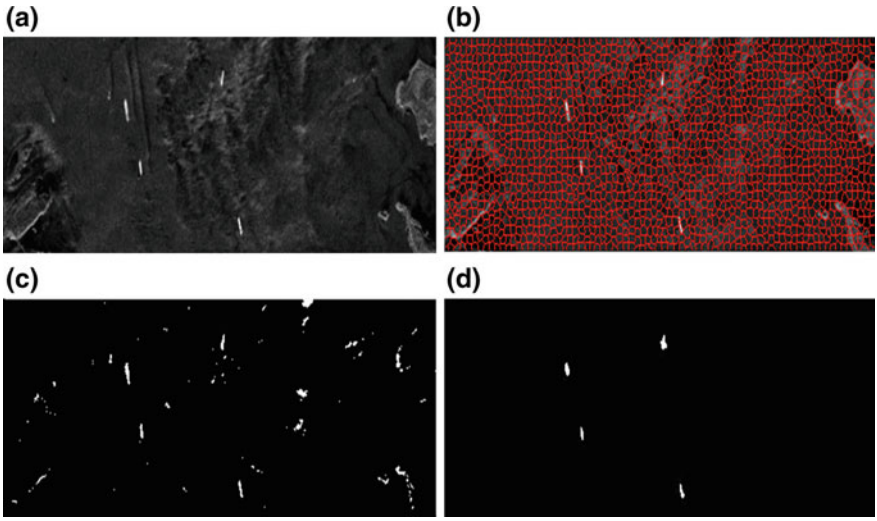



**Fig. 3.** **a** Strait of Gibraltar SAR image 2; **b** result of superpixel segmentation; **c** result of CFAR based on Rayleigh distribution model ($P_{fa}$ = 0.001); **d** result of the proposed method

**Acknowledgments.** This study was supported by the Key Technology R&D Program of Sichuan Province 2015GZ0109, the National Nature Science Foundation of China under Grant 61271287 and Grant U14331.

## References

1. Zhou X, Chang NB, Li S. Applications of SAR interferometry in earth and environmental science research. Sensors. 2009;9(3):1876–912.
2. Hou B, Chen X, Jiao L. Multilayer CFAR detection of ship targets in very high resolution SAR images. IEEE Geosci Remote Sens Lett. 2014;12(4):811–5.
3. Hansen VG. Constant false alarm rate processing in search radars. Radar-present and future 1973.
4. Wang S, et al. New hierarchical saliency filtering for fast ship detection in high-resolution SAR images. IEEE Trans Geosci Remote Sens. 2016;55(1):351–62.
5. Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Mach Intell. 2002;24(5):603–19.
6. Nock R, Nielsen F. Statistical region merging. IEEE Trans Pattern Anal Mach Intell. 2004;26 (11):1452.
7. Achanta R, et al. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans Pattern Anal Mach Intell. 2012;34(11):2274–82.
8. Qin F, et al. Superpixel segmentation for polarimetric SAR imagery using local iterative clustering. IEEE Geosci Remote Sens Lett. 2017;12(1):13–7.
9. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.

# Non-contact Detection of Vital Signs via a UWB Radar Sensor

Zhenzhen Duan(✉), Yang Zhang, Jian Zhang, and Jing Liang

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China
839014412@qq.com

**Abstract.** The detection of vital signs by the ultra-wideband (UWB) radar sensor does not require touching the body. The UWB radar sensor has shorter detection time and less radiation than X-rays. Considering that heartbeat signals and respiratory signals are non-stationary, this paper uses the empirical mode decomposition (EMD) algorithm and the variational mode decomposition (VMD) algorithm to process them. Both respiratory and heartbeat characteristics are obtained with non-contact detection in the paper. The Hilbert transform of vital signs is conducted to reflect the time–frequency information of vital signs. Experimental results show that a highly accurate detection of vital signs can be achieved using the proposed methods.

**Keywords:** UWB radar sensor · Human detection · Mode decomposition · Hilbert transform

## 1 Introduction

The UWB radar sensor can be widely applied in special patient monitoring, rescue after the earthquake, and other fields due to its excellent features [1–3]. Features of heartbeat and respiratory are part of healthcare. They are two important diagnostic foundations for clinical diagnosis. Igor Immoreev and Teh-Ho Tao applied the UWB radar sensor to non-contact patient monitoring and designed a detection system [4]. The system has been tested in the hospital in Moscow [5]. Kumar et al. [6] used the UWB radar sensor to measure the distance for the human target and detected respiratory characteristics. Domenico Zito et al. developed a wearable respiration detection system [7]. Its detection accuracy is close to the contact-type life monitor [8,9].

The 8th International Conference on Communications, Signal Processing, and Systems: The research in this paper is ethical and does not affect any human health. We have already informed the experiment content for all participants. The person in Fig. 2 allows her photograph to be used in this paper and agree to publish.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_86

However, these studies failed to detect both respiratory and heartbeat characteristics using UWB radar sensors. It is necessary to process echo signals with modal decomposition algorithms to obtain the characteristics of respiratory and heartbeat at the same time. This paper successfully extracts heartbeat signal and respiratory signal from original signals via a UWB radar sensor. The contributions of this paper are as follows:

1. A detection system for the human body based on the UWB radar sensor PulseOn 440 is designed and original-type omnidirectional dipole antennas are replaced with quasi-Yagi antennas. The use of quasi-Yagi antennas reduces the inherent antenna coupling noise.

2. Both respiratory and heartbeat motions are obtained from signals that have been preprocessed. Moreover, the respiration detection error between the UWB radar sensor and electrocardiogram (ECG) is less than 1 per minute. In addition, the time–frequency information got by the Hilbert transform verifies that using the VMD algorithm to process echo signals can effectively separate the respiratory and heartbeat motions.

The rest paper is organized as follows: Section 2 describes the processing of echo signals. It includes preprocessing and mode decomposition. Preprocessing of echo signals is used to remove coherent background noise and clutter. Vital signs can be acquired by using modal decomposition algorithms. Section 3 is about the experiment carried out in the hospital and the analysis of experimental results. The experiment includes detection based on the UWB radar sensor and ECG.

## 2 Signals Processing for Human Detection

The flow diagram of signal processing for human detection is shown in Fig. 1. Firstly, establishing a detection system to collect original echo signals. Secondly, preprocessing original echo signals considering the influence of scene noise and clutter. Then, EMD and VMD algorithms are used to extract respiratory and heartbeat signals. In addition, the Hilbert transform can be used to find time–frequency information.

### 2.1 Echo Signals Preprocessing

Due to the influence of the measuring environment and auxiliary equipments, the original echo signals are not ideal. The signals detected by the UWB radar sensor need to be preprocessed. The preprocessing methods include coherent background noise removal and clutter filtering.

**2.1.1 Coherent Background Noise Removal** The background noise of the measurement environment can be eliminated or reduced by using coherent background noise removal. Then, the effect of UWB radar sensor detection can be improved. Collecting environmental detection echo signals $A_b$ without human bodies firstly. Then, gathering echo signals $A_o$ when the human body is present.

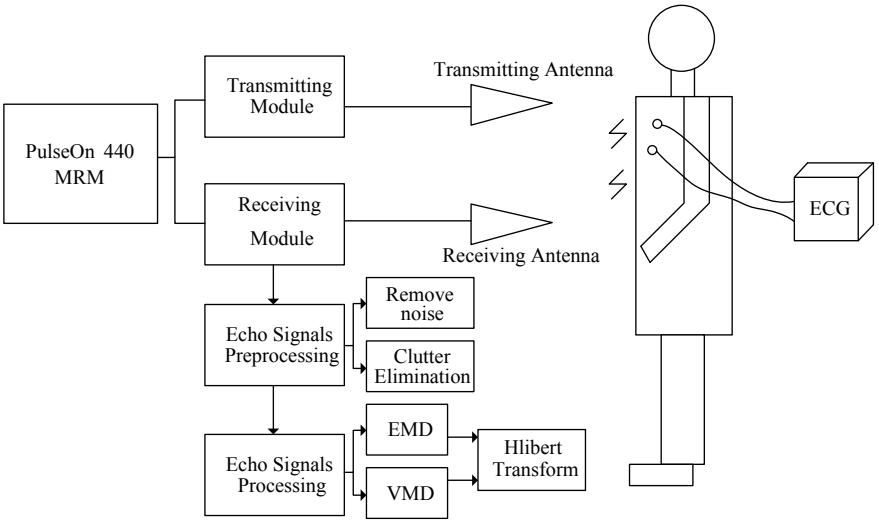



**Fig. 1.** Signals processing for human detection

Finally, we can get the human target echo signals $A$ by deleting the scene information from the original echo signals (1). $A$ is constructed from $N$ scans and the length of each scan is $M$.

$$A = A_o - A_b \tag{1}$$

**2.1.2 Clutter Elimination** Generally speaking, signals of vital signs are weak because of the presence of background clutter. Background clutter usually consists of low frequency, DC components, and some linear slowly changing trends. This paper uses range profile subtraction method to eliminate the constant part of signals. $A_{diff}$ is a $M$ by $(N-1)$ matrix.

$$A = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{m1} & \cdots & s_{mn} \end{pmatrix} \tag{2}$$

$$A_{diff} = \begin{pmatrix} s_{11} - s_{12} & \cdots & s_{1(n-1)} - s_{1n} \\ \vdots & \ddots & \vdots \\ s_{m1} - s_{m2} & \cdots & s_{m(n-1)} - s_{mn} \end{pmatrix} \tag{3}$$

This method is simple and effective for clutter elimination.

Signals of vital signs are stronger than those without preprocessing. Noise can be reduced using coherent background noise removal and clutter can be eliminated using range profile subtraction method.

### 2.2 Echo Signals Processing Algorithms

In order to acquire respiratory and heartbeat signs, the echo signals after preprocessing should be processed by mode decomposition algorithms. Since EMD and VMD algorithms are suitable for non-stationary signals, it is appropriate to use them to extract respiratory and heartbeat signals. EMD algorithm has been successfully applied to marine, mechanical vibration, and other fields.

Different from the EMD algorithm, the VMD algorithm is based on the mathematical theory. Two signals with similar frequency can be successfully separated by VMD algorithm [10]. The steps of the VMD algorithm are constructing the variational problem and solving the variational problem:

Constructing the variational problem:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|^2 \right\} \tag{4}$$

where, $\{u_k\}$ is the intrinsic mode function. $\{u_k(t)\} = \{u_1, \ldots, u_k\}$. $\{\omega_k\}$ is the corresponding center frequency. $\{\omega_k\} = \{\omega_1, \ldots, \omega_k\}$. $\sum_{k=1}^{K}$ is the sum of all intrinsic mode functions. The definition of $u_k(t)$ is as (5).

$$u_k(t) = A_k(t) \cos(\varphi_k(t)) \tag{5}$$

We can clearly see that the intrinsic mode function is an amplitude modulation-frequency modulation signal. The change velocities of envelope $A_k(t)$ and instantaneous frequency $\omega_k(t) = \frac{d\varphi_k(t)}{dt}$ are far less than those of phase $\varphi_k(t)$.

Solving the variational problem:

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k & \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|^2 \\ & + \left\| s(t) - \sum_k u_k(t) \right\|^2 + \left\langle \lambda(t), s(t) - \sum_k u_k(t) \right\rangle \end{aligned} \tag{6}$$

where $\alpha$ is the penalty coefficient. $\lambda(t)$ is the lagrangian multiplier. $s(t)$ is the echo signal. The alternating directions method of multipliers is used to get the optimal solution of the (6).

### 2.3 Time–Frequency Information

Frequency of vital signs can be obtained after preprocessing and processing. In addition, time–frequency information of vital signs is presented form the Hilbert spectrum because it can describe the change rule of the data sequence in real time.

$$Z(t) = u_k(t) + j\tilde{u}_k(t) = a_k(t)\, \mathrm{e}^{j\theta_k t} \tag{7}$$

where $a_k(t) = \sqrt{u_k^2(t) + j\tilde{u}_k^2(t)}$, $\theta_k(t) = \arctan \frac{\tilde{u}_k(t)}{u_k(t)}$.

$Z(t)$ is the complex sequence and $\tilde{u}_k(t)$ is the Hilbert transform of $u_k(t)$. $\omega_k(t) = \frac{d\theta_k(t)}{dt}$ is the instantaneous frequency. So, echo signals can be represented as $s(t) = \mathrm{Re}\sum_{k=1}^{n} a_k(t)\, e^{j\int \omega_k(t)dt}$. Because $\omega_k(t)$ is the function of time, the Hilbert spectrum can describe changes in the time-domain.

## 3 Detection Experiment and Result Analysis

### 3.1 Detection Experiment

As Fig. 2 shows, the experiment to detecting vital signs was carried out in the hospital of UESTC. In the experiment, the PulsOn 440 Monostatic Radar Module (P440-MRM) was used to transmit and receive signals. P440-MRM can provide UWB Gaussian monopulse signals whose center frequency is 4.3 GHz, and the bandwidth is 2.2 GHz. It collects echo signals at a sampling rate of 16.4 GHz. The number of data points in a single scan is 480. Pulse repetition rate is 10.1 MHz.

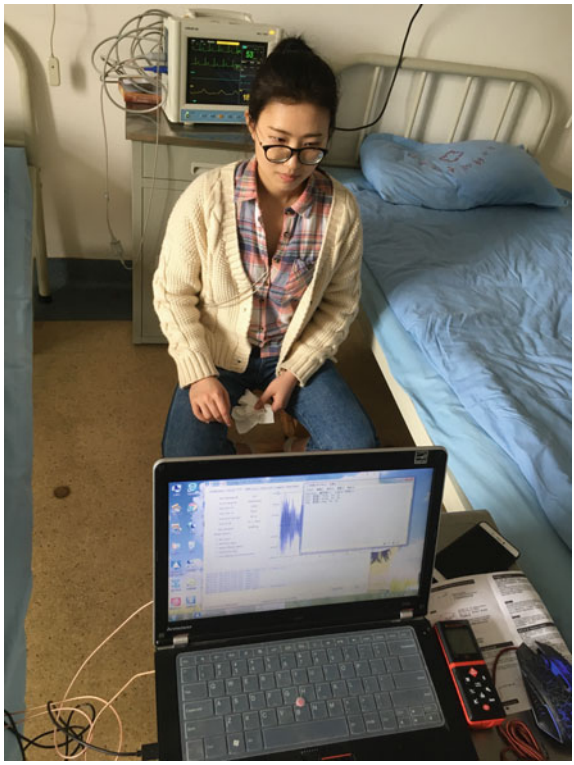
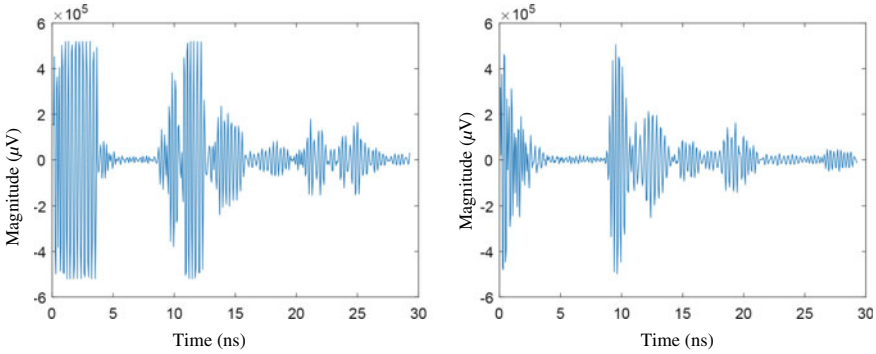


**Fig. 2.** Experiment for human body

Broadband planar quasi-Yagi antennas replaced original-type omnidirectional dipole antennas in the measurement. The group delay of broadband planar quasi-Yagi antennas is not more than 0.82 ns. The antennas have good directivity and wide frequency band. Broadband planar quasi-Yagi antennas can effectively reduce the coupling noise of echo signals compared with original-type omnidirectional dipole antennas. As shown in Fig. 3, coupling noise of broadband planar quasi-Yagi antennas starts to decrease at 2 ns, while the coupling noise of original omni-dipole antennas begins to decrease at 4 ns. Echo signals begin to appear a large amount of information around 9 ns because of the use of 1 m SMA coaxial cable.




(a) The Omnidirectional Dipole Antenna

(b) The Broadband Planar Quasi Yagi Antenna

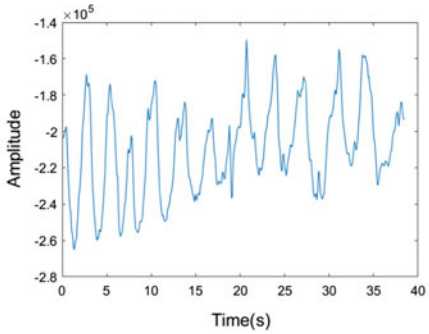
**Fig. 3.** Comparison of radar receipt signals

The experiment conducted 481 scans via the P440-MRM. The human target was 0.822 m away from the antennas. First, choosing the maximum range gate of each scan. The range gate is a 1 by $(N-1)$ vector, and it is the signal that should be processed by processing algorithm. The parameters of VMD algorithm are set as follows: The number of "K" is 4, and it represents the number of intrinsic modes. $\tau$ is the dual ascending step length, and it is 0. "DC" is 1 because the first mode is low frequency. "Init" is 0 to initialize the center frequency. Set "Tol" as $10^{-5}$ and "Tol" is the condition of convergence. $\alpha$ is the convergence rate, and it is 2500.

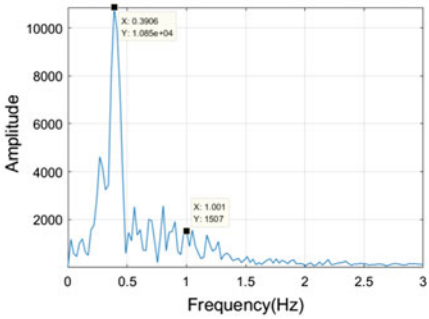
Processing the collected echo signals using preprocessing and modal decomposition methods, the results of Sect. 3.2 can be obtained.

### 3.2 Result Analysis

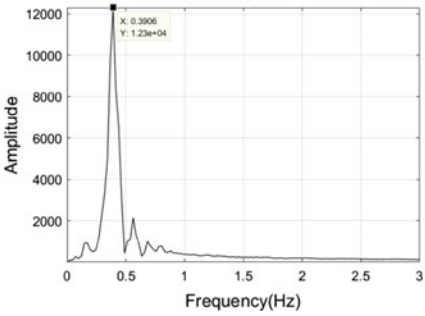
The results of detection of vital signs based on the UWB radar sensor PulseOn 440 are shown in Fig. 4.



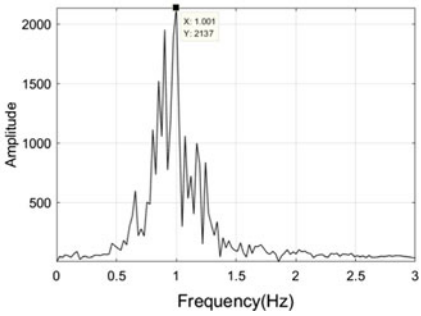

(a) The Time Domain Waveform




(b) The Respiratory and Heartbeat Rates Based on EMD



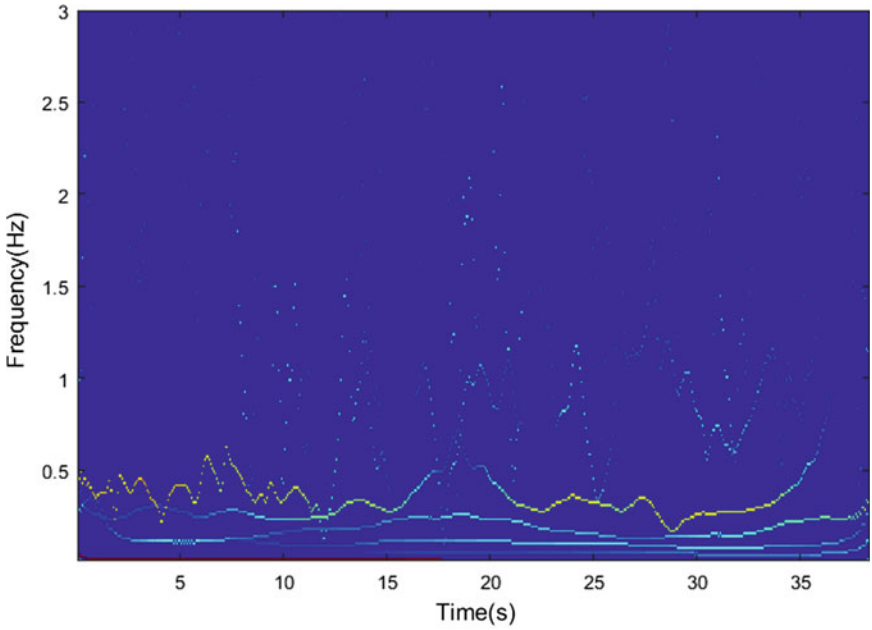

(c) The Respiratory Rate Based on VMD



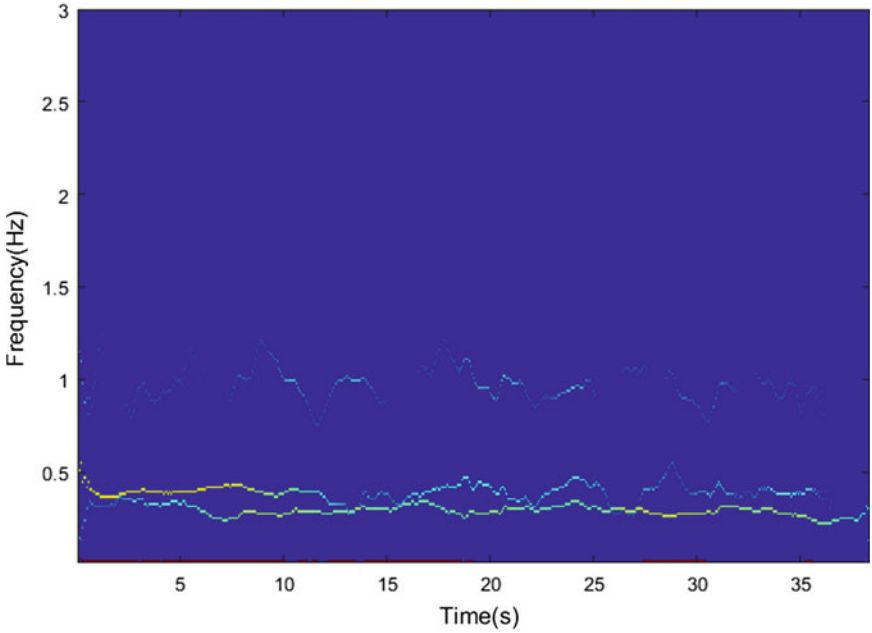

(d) The Heartbeat Rate Based on VMD

**Fig. 4.** Detection results of vital signs

In the experiment, the respiratory rate is 18 beats per minute and the heart rate is 65 beats per minute for the human body detected by ECG. As can be seen from these figures, the respiratory frequency based on EMD and VMD is 0.3906 Hz, which corresponds to 17.58 times per minute. The heartbeat frequency based on EMD and VMD is 1.001 Hz which is equivalent to 60 times per minute. So, we can find that the precision of respiratory is very high and the error is less than one time per minute. The error of heartbeat is less than five times. However, from Fig. 4b, the detection based on the EMD algorithm appears modal aliasing. The respiratory motion is obvious, but the heartbeat motion is hidden in the noise. Analyzing Fig. 4c, d, the heartbeat motion is successfully separated from the respiratory motion. Comparing Fig. 5a with Fig. 5b, the EMD-based Hilbert spectra present modal aliasing, while the Hilbert spectra based on the VMD algorithm do not appear. This demonstrates the superiority of the VMD algorithm again.




(a) Hilbert Spectrum Based on EMD




(b) Hilbert Spectrum Based on VMD

**Fig. 5.** Time–frequency information

## 4 Conclusion

In this paper, the respiratory and the heartbeat motions are detected clearly via the UWB radar sensor. The accuracy is very high, so the detection based on the UWB radar sensor is reliable. The detection system contains PulseOn 440 and broadband planar quasi-Yagi antennas and other equipments. The use of broadband planar quasi-Yagi antennas can reduce the echo signals' coupling noise effectively. Heartbeat signals based on the VMD algorithm can be successfully separated from respiratory signals and other harmonics. Hilbert spectra based on two algorithms show the time–frequency information of vital signs detected by the UWB radar sensor.

The non-contact detection of vital signs via a UWB radar sensor is verified in this paper. As future work, vital signs in different motion states can be detected by the UWB radar sensor. Due to the superiority of the UWB radar signal, detection for multiple human bodies can be conducted in the future.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61671138, 61731006), and was partly supported by the 111 Project No. B17008.

## References

1. Staderini EM. Uwb radars in medicine. IEEE Aerosp Electron Syst Mag. 2002;17(1):13–8.
2. Li C, Cummings J, Lam J, Graves E, Wu W. Radar remote monitoring of vital signs. IEEE Microwave Mag. 2009;10(1):47–56.
3. Kaplan M. Wireless sensing for the respiratory activity of human beings: measurements and wide-band numerical analysis. Int J Antennas Propag. 2013;2013(3):559–62 (2013-4-22).
4. Immoreev I, Tao TH. Uwb radar for patient monitoring. IEEE Aerosp Electron Syst Mag. 2008;23(11):11–8.
5. Ziganshin EG, Numerov MA, Vygolov SA. UWB baby monitor. In: International conference on ultrawideband and ultrashort impulse signals; 2010. p. 159–61.
6. Kumar A, Liang Q, Li Z, Zhang B, Wu X. Experimental study of through-wall human being detection using ultra-wideband (UWB) radar; 2012.
7. Zito D, Pepe D, Neri B, Zito F, De RD, Lanat A. Feasibility study and design of a wearable system-on-a-chip pulse radar for contactless cardiopulmonary monitoring. Int J Telemed Appl. 2008;2008(4):328597.
8. Mincica M, Pepe D, Tognetti A, Lanat A, Rossi DD, Zito D. Enabling technology for heart health wireless assistance. In: IEEE international conference on e-Health networking applications and services; 2010. p. 36–42.
9. Zito D, Pepe D, Mincica M, Zito F. A 90 nm CMOS SoC UWB pulse radar for respiratory rate monitoring. In: Solid-state circuits conference digest of technical papers; 2011. p. 40–1.
10. Dragomiretskiy K, Zosso D. Variational mode decomposition. IEEE Press; 2014.

# Bind Intra-pulse Modulation Recognition based on Machine Learning in Radar Signal Processing

Xiaokai Liu[1(✉)], Shaohua Cui[2], Chenglin Zhao[1], Pengbiao Wang[1], and Ruijian Zhang[1]

[1] Beijing University of Posts and Telecommunications (BUPT), Beijing, China
liu_xiaokai@bupt.edu.cn
[2] China Petroleum Technology and Development Corporation, Beijing, China

**Abstract.** Intra-pulse modulation recognition is one of the radar reconnaissance key technologies; it is especially a hot point of recent researching under low SNR. This thesis propounds a novel way for radar intra-pulse modulation characteristic recognition based on machine learning means of extreme learning machine (ELM), which is widely applied in the region of pattern recognition. As a novel learning framework, the ELM attracts increasing draws in the regions of large-scale computing, high-velocity signal processing, and artificial intelligence. The aim of the ELM is to break the barriers down between the biological learning mechanism and conventional artificial learning techniques and represent a suite of machine learning methods in which hidden neurons need not to be tuned. This algorithm has a trend to provide perfect generalization performance at staggering learning rate. This article focuses on the high frequency (HF) channel environment and Wavelet transform algorithm with the lower computational complexity. The simulation results imply that the ELM could reap a perfectly satisfactory acceptance performance and therefore supplies a substantial ground structure for dealing with intra-pulse modulation challenges in inadequate channel conditions.

**Keywords:** Automatic modulation recognition · Radar signal · HF channel · Wavelet transform · Machine learning

## 1 Introduction

The complex intra-pulse radar signal automatic modulation recognition is a major task for the electronic warfare (EW) receivers in recent decades. Intra-pulse modulation identification is one of radar reconnaissance key technologies.

Xiaokai Liu is with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications (BUPT), Beijing, China.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_87

The process primary bottlenecks lie in the computational complexity of the classifier and the noisy signal conditions. The shortage of a priori knowledge about interest waveform establishes modem EW receiver design with real-time recognition even more challenging.

For the radar signal automatic modulation recognition, there are two ordinary categories of algorithms, likelihood-based (LB) and feature-based (FB) algorithms [1], respectively. The LB ways are based on the likelihood function and construct a decision by comparing the likelihood ratio against a set threshold. The LB algorithms make the results optimal in a quantity of sense, but suffer for computational complexity. On the other hand, in the FB algorithms, several features are commonly utilized for recognition. When designed decently, it is ordinarily simple to be applied with sub-optimal performance. For practical applications, a lot of positive results have been yielded for the intra-pulse modulation with several FB schemes. The researchers have made great efforts to discover new features to improve the performance of FB under poor environments. In [2], the authors present the means systematically applying the Rihaczek distribution and Hough transform features to classify the radar signals. In [3], the author did some research for the AMC of radar signals, exploiting wonderful features based on Choi–Williams time–frequency distributions and Wigner while the SNR is more than 6 dB. In [4], the authors deliver a rapid recognition algorithm to recognize three types of intra-pulse modulation signals, and this algorithm works comfortably when the SNR is above 0 dB. Lopez-Risueno recommended some methods to distinguish four kinds of radar signals, and his algorithms play well around 0 dB for most types of signals [5–7]. In [8], a novel method based on first-order and second-order phase contrast was evolved for intra-pulse modulation recognition; this method need not disclose the phase ambiguity and can get good result around 6 dB. In [9], the author used the spectrogram time–frequency to estimate the instantaneous frequency, and abstracted some characteristic features to separate amplitude shift-keying, FSK and PSK signals around 2 dB. However, how to classify the classical radar signals under negative SNR and decreasing the computational complexity are still challenging problems.

Pattern recognition and machine learning have drawn a lot of researchers' attention. From [10], we can get their geneses in engineering which can be viewed as the same field's two aspects, while machine learning originated from computer science. They all have also suffered considerable evolution during the past ten years. Extreme learning machine (ELM), as a novel learning framework, draws increasing attractions in the regions of large-scale computing, high-velocity signal processing, artificial intelligence, etc.. The aim of ELM is to break the barriers down between the biological learning mechanism and traditional artificial learning techniques and represents a suite of machine learning techniques in which hidden neurons need not to be tuned. Algorithms about ELM theories suggest that "random hidden neurons" capture the essence of some brain learning mechanisms as well as the intuitive sense that brain learning's efficiency need not depend on computing neurons power. Thus, in comparison with support vector machine and traditional neural networks, ELM provides considerable advantages

such as fast learning velocity, comfort of implementation, and minimum human intervention. Because of its implementation efficiency and remarkable generalization performance, ELMs have created tremendous improvements. Following the reference materials, the ELM has been evidenced an effective and efficient learning mechanism for regression and pattern classification [11,12]. Distinct from the existent works, we cover radar signal classification's challenges in following parts:

- This thesis adopts the HF channel with Doppler shift and multipath time delay into the modulation identification. In this model, the negative effect eroded by the HF channel makes the radar signal modulation recognition more complex. To tackle the instance, in the classifier selection procedure the ELM, SAE-ELM algorithms are fortuitously implied to radar signal recognition.
- The ELM algorithms' advantage lie in the random feature mapping, that is, generating the bias and input weights randomly to fulfill what the network require. The ELMs guarantee the real time of the identification system.
- In the feature extraction part, we utilize the wavelet transform to construct the datasets. All the selected features can ensure the differences among all the signals.

The rest of the paper is organized as follows: Section 2 defines the system model and the Watterson channel model. Section 3 presented the proposed modulation identification algorithm, and Sect. 4 concentrates on the recognition simulation results. Finally, conclusions and perspectives of the research work are presented in Sect. 5.

## 2 System Model and Preliminaries

In this section, the signal model is defined; this paper introduced four types of the radar signals: linear frequency modulation (LFM), nonlinear frequency modulation (NLFM), two-phase/four-phase coded (BPSK/QPSK), and frequency encoding (2FSK/4FSK). The signal model is introduced below.

### 2.1 Signal Model

The radar signal model can be portrayed using the following equation:

$$s(t) = a(t)\exp\{j[2\pi f_0 t + \varphi(t) + \phi_0]\} \tag{1}$$

For which the $\varphi(t)$ stands for the coding function, $f_0$ is the carrier frequency, $\phi_0$ is a constant phase, and $a(t)$ is the envelope of the pulse defined by:

$$a(t) = Arect(t/T) \tag{2}$$

where $A$ is the pulse magnitude, $T$ is the pulse span, and the function $rect(t/T)$ is defined to be:

$$rect(t/T) = \begin{cases} 1 & t \leqslant T0 \\ \text{otherwise} \end{cases} \tag{3}$$

we can change the $\varphi(t)$ to form the different modulation type.

## 3 Proposed Modulation Identification Algorithms

The Watterson channel makes the signal corrupting their properties and leads to negative simulation results. Our intentions are to abstract the radar signal modulation type via the pattern recognition and machine learning. The machine learning part should acquire the signal attributes intelligently and separate the signal in real time.

### 3.1 Features Extraction

The fundamental features for modulation classification in pattern recognition method should be elected. These features ought to own robust properties of sensitive with the signal kinds and insensitive with SNR variations [13]. The continuous wavelet transform (CWT) of a signal x(t) is defined as:

$$WT_f(a,b) = |a|^{-1/2} \int_{-\infty}^{+\infty} f(t)\psi^*(\frac{t-b}{a})dt = \int_{-\infty}^{+\infty} f(t)\psi^*_{a,b}(t)dt = \langle f(t), \psi_{a,b}(t)\rangle \tag{4}$$

where $\langle f(t), \psi_{a,b}(t)\rangle$ stands for dot product. $\psi(t)$ is the mother wavelet.

### 3.2 The Machine Learning algorithm

In [14,15], our previous research has present the machine learning algorithm. In this part, this section attempts to summarize the key issues for the ELM and SAE-ELM process. Overall, signal recognition based on ELM involves the following stages

- Step1: Given a training set$\{\mathbf{X}, \mathbf{Y}\} = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^{N}$ , activation function $g(x)$ and hidden node number $\tilde{N}$ .
- Step2: Randomly assign input weight $\mathbf{w}_i$ and bias $b_i$, $i = 1, \ldots, N$.
- Step3: Calculate the hidden layer output matrix $\mathbf{H}$.
- Step4: Calculate the output weight $\boldsymbol{\beta} = \mathbf{H}^T\mathbf{T}$. Where $\mathbf{T} = [t_1, \ldots, t_N]^T$.
- Step5: Test the performance of ELM.

In conclusion, signal recognition based on SAE-ELM consists of the following steps:

**Step1. Initialization**

A set of NP vectors where each one includes all the network hidden node parameters are initialized as the populations of the first generation.

$$\boldsymbol{\theta}_{k,G} = [\mathbf{w}^T_{1,(k,G)}, \ldots, \mathbf{w}^T_{L,(k,G)}, b_{1,(k,G)}, \ldots, b_{L,(k,G)}] \tag{5}$$

where $\mathbf{w}_j$ and $b_j$ $(j = 1, \ldots L)$ are randomly generated, G represents the iterations generation and $k = 1, 2, \ldots, NP$.

**Step2. Calculations of output weights and RMSE**
Figure the network output weight matrix and root mean square error (RMSE) in regard to each population vector with the following equations, respectively.

$$\boldsymbol{\beta}_{k,G} = \mathbf{H}_{k,G}^{\dagger}\mathbf{T} \tag{6}$$

$$RMSE_{k,G} = \sqrt{\frac{\sum_{i=1}^{N} \parallel \sum_{j=1}^{L} \boldsymbol{\beta}_j g(\mathbf{w}_{j,(k,G)},\ b_{j,(k,G)},\ \mathbf{x}_i) - \mathbf{t}_i \parallel}{m \times N}} \tag{7}$$

where $\mathbf{H}_{k,G}$ is demonstrated in Eq. 8, and $\mathbf{H}_{k,G}^{\dagger}$ is the MP generalized inverse of $\mathbf{H}_{k,G}$.

$$\mathbf{H}_{k,G} = \begin{pmatrix} \mathbf{w}_{1,(k,G)},\ b_{1,(k,G)},\ \mathbf{x}_1 & \dots & \mathbf{w}_{L,(k,G)},\ b_{L,(k,G)},\ \mathbf{x}_1 \\ \vdots & \ddots & \vdots \\ \mathbf{w}_{1,(k,G)},\ b_{1,(k,G)},\ \mathbf{x}_N & \dots & \mathbf{w}_{L,(k,G)},\ b_{L,(k,G)},\ \mathbf{x}_N \end{pmatrix} \tag{8}$$

$$\boldsymbol{\theta}_{k,G+1} = \begin{cases} \mathbf{u}_{k,G+1} if RMSE_{\boldsymbol{\theta}_{k,G}} - RMSE_{\mathbf{u}_{k,G+1}} > \varepsilon RMSE_{\boldsymbol{\theta}_{k,G}} \\ \mathbf{u}_{k,G+1} if |RMSE_{\boldsymbol{\theta}_{k,G}} - RMSE_{\mathbf{u}_{k,G+1}}| < \varepsilon RMSE_{\boldsymbol{\theta}_{k,G}} \\ and \quad \parallel \boldsymbol{\beta}_{\mathbf{u}_{k,G+1}} \parallel < \parallel \boldsymbol{\beta}_{\boldsymbol{\theta}_{k,G}} \parallel \\ \boldsymbol{\theta}_{k,G}\ else. \end{cases} \tag{9}$$

In the first generation, the population vector with the nicest RMSE is stored as $\boldsymbol{\theta}_{RMSandbest,1}$ $E_{\boldsymbol{\theta}_{best,1}}$.
**Step3. Mutation and Crossover**
Similar to the [16], for each objective vector in the prevailing generation, the trial vector generation strategy is selected from a candidate pool that is created by a number of strategies according to possibility.
**Step4. Evaluation**
All the trial vectors $\mathbf{u}_{k,G+1}$ produced at the $G+1$th generation are measured employing Eq. 9 where $\varepsilon$ is the preset little certain tolerance rate. Steps 3 and 4 are restated till the preset aim is gotten or the maximal learning iterations are finished.

## 4 Performance Result and Discussion

The proposed algorithm was verified and validated for the radar digital modulation schemes. In our simulations, we consider the following configurations: (1) In the signal choice, we choose the mixed signals: LFM, NLFM, BPSK/QPSK, 2FSK/4FSK. (2) Radar signals transmit though channel with the SNR differs from −10 to 10 dB stepped by 2 dB in interval. (3) The symbol rate was 200 bps, sampling frequency is 1000 Hz. (4) For the Frequency encoding signal, we set the starting frequency to 200 Hz, maximum frequency interval is 50 Hz, the biggest change bandwidth is 100 Hz. (5) Watterson channel maximum Doppler frequency is 10 Hz, the multipath parameters was set to 2, the time delay was set to 0.001 s. (6) A total of 800 groups of signals of radar modulation scheme are generated; each dataset is separated into two no overlapped datasets: the training and the validation datasets.

### 4.1 Performance Results of Wavelet transform

Figures 1, 2, 3, 4, 5, and 6 show the characteristics of radar signal features. From the cD1, cD2, cD3, cD4, cD5, cA5, we can draw some conclusions. For any SNR, the six signal characteristics can reflect the differences perfectly, even if in some features, the distinction was not obvious, we can get good results with other feature help.

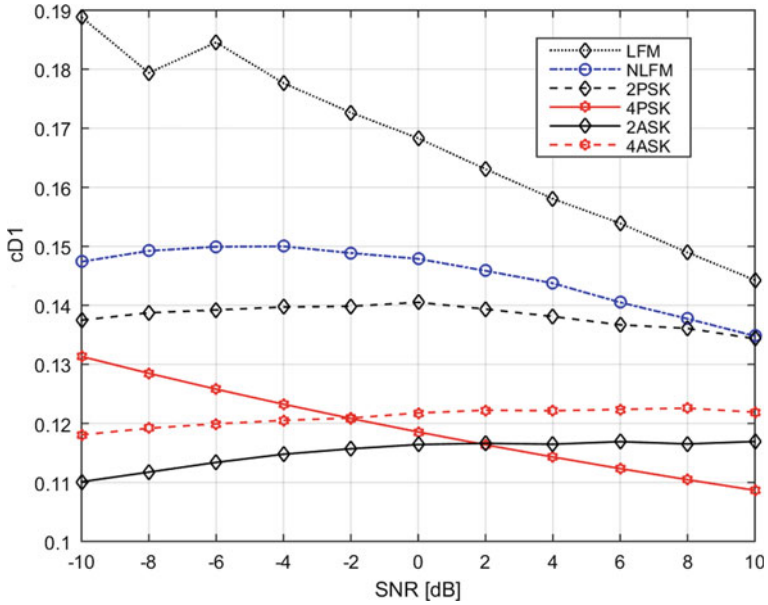



**Fig. 1.** cD1

### 4.2 Blind Classification Performance for the Mixed Signals in the Watterson Channel

For the mixed signal, we reveal the performance results of identification probability for the radar system; we produce signal feature sets to be prepared for the experimentation. Then randomly chose 800 groups of features to train ELMs neural network to obtain the parameters which are necessary for the subsequent testing. In the testing procedure, we randomly pick 800 groups. The system, however, does not realize which kind they belong to. We keep the right results to establish a comparison with the testing ones to acquire the precision rate. Figures 7 and 8 illustrate the results of SAE-ELM and the ELM. From the Figs. 7 and 8, we can gain the following conclusions: (1) For the mixed modulation recognition in the Watterson channel. The ELM and SAE-ELM algorithms present a good result, take the ELM as a example, and when the SNR is −10dB, the

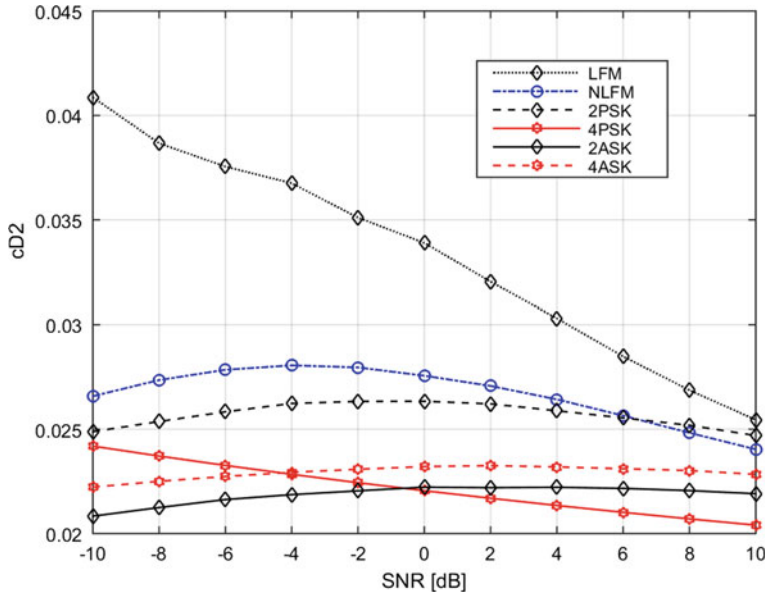



**Fig. 2.** cD2

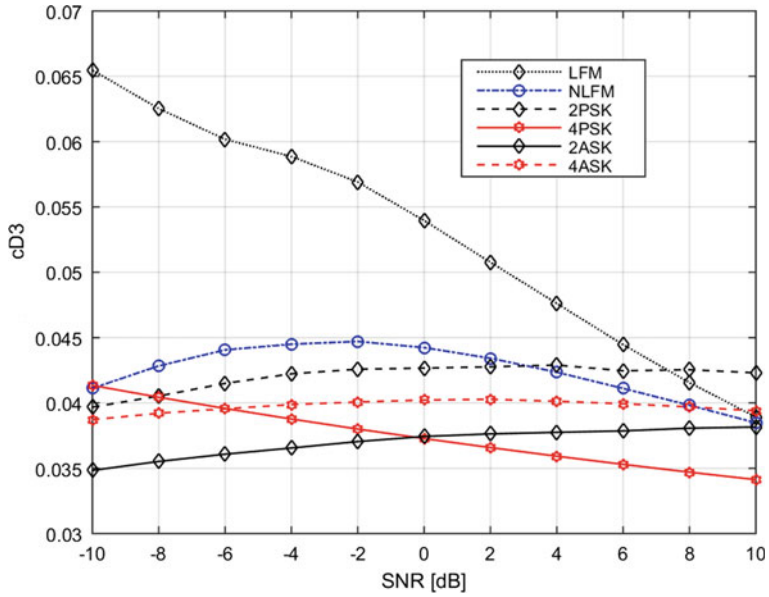



**Fig. 3.** cD3

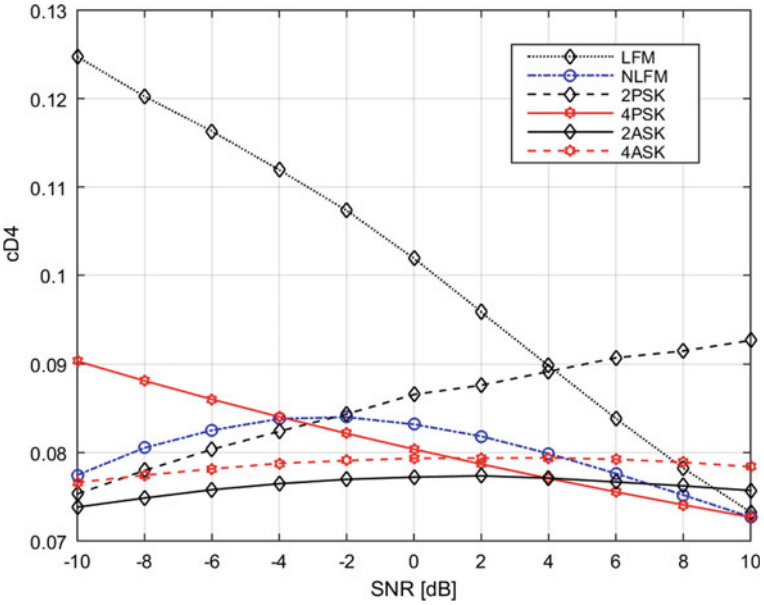



**Fig. 4.** cD4

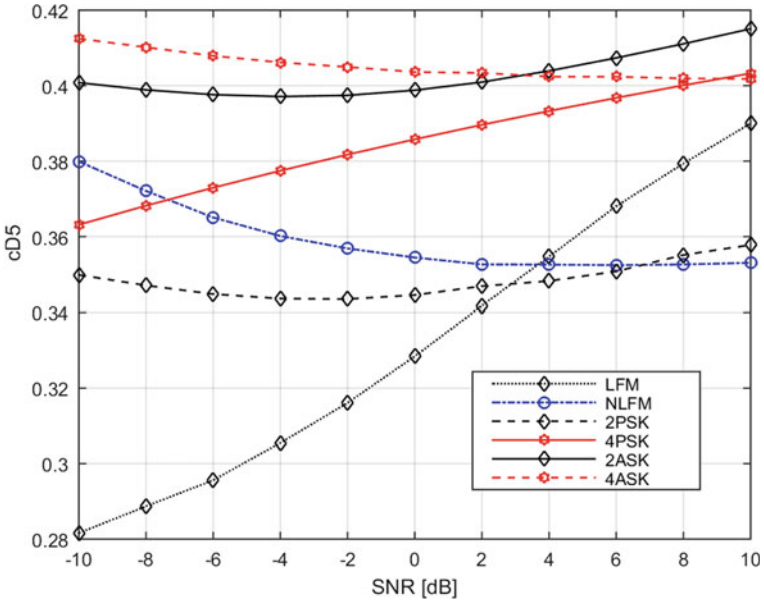



**Fig. 5.** cD5

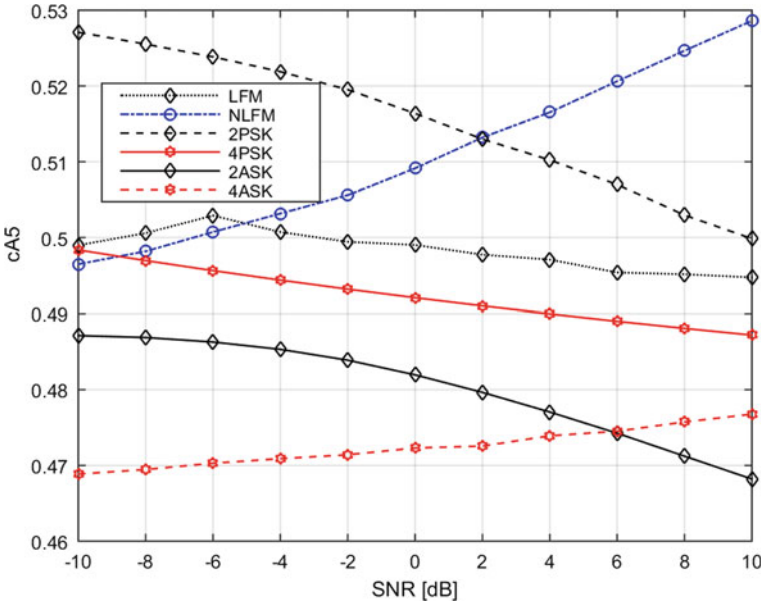



**Fig. 6.** cA5

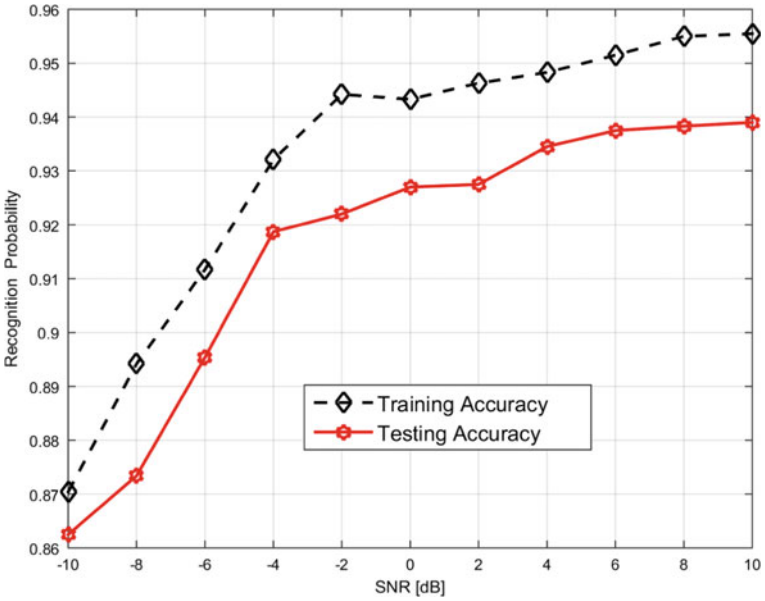



**Fig. 7.** Blind identification probability with ELM

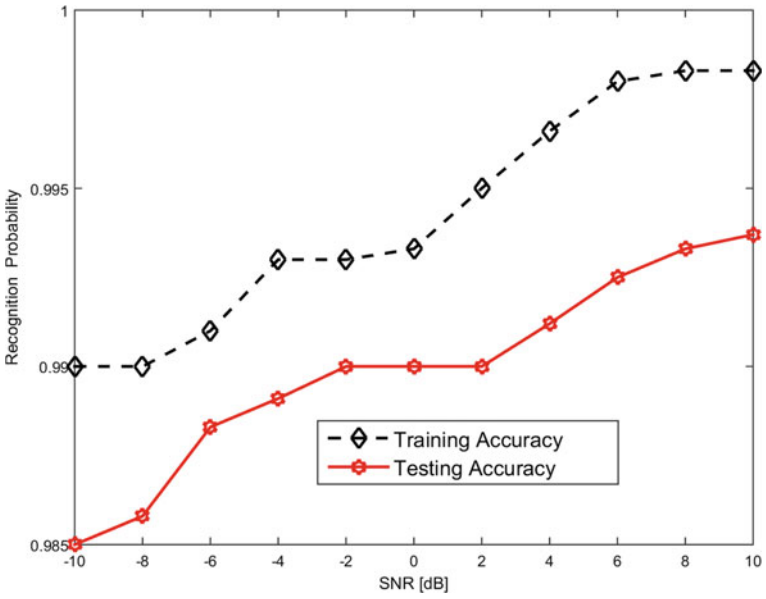



**Fig. 8.** Blind identification probability with SAE-ELM

probability of the system identification reaches more than 86%; when the SNR is 4dB, the probability of identification will reach their upper bound, more than 93%. As a whole, the ELM method, applied to the Watterson channel system identification, obtains an acceptable result, but not perfectly. (2) The contrast of the SAE-ELM and the ELM in a similar way has visually been emerged. From the numeric consequences, the established SAE-ELM model is relatively effective for blind modulation recognition. Utilizing the SAE-ELM algorithms, identification probability is superior to the ELM apparently, specially under the low SNR.

### 4.3 The Delay Effect for the Radar Signals in the Watterson Channel

Figures 7 and 8 have verified that ELM algorithms are an effective method for signal recognition and the derivative SA-ELM can resolve the puzzle for low SNR case. This section introduces the distinction while we alter the delay parameter. In this simulation, we determine the SNR = 0dB adjust the time delay to [0.01,0.03,0.05,0.08,0.10]; in this method, select the ELM algorithm as the classifiers. The performance is pointed out in the Fig. 9. From the results, we can well arrive at the conclusion that when the delay time grows, the right classification probability declines. When the delay time was 0.01 s, the probability of the correct classification reaches 93.9%; this upper bound limitation is serious in the SNR = 0, and the identification performance will not transcend the upper bound; it was as a result of effect caused by the Doppler shifts and noise.

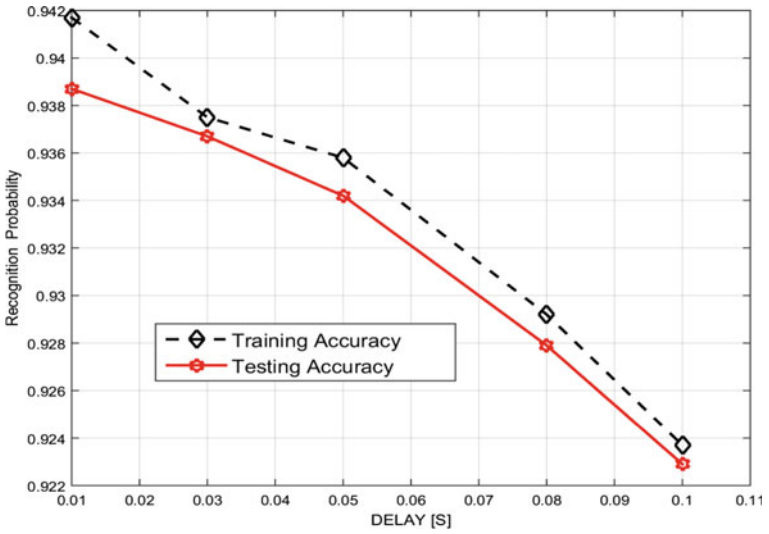



**Fig. 9.** Blind identification probability versus delay time in the SNR = 0

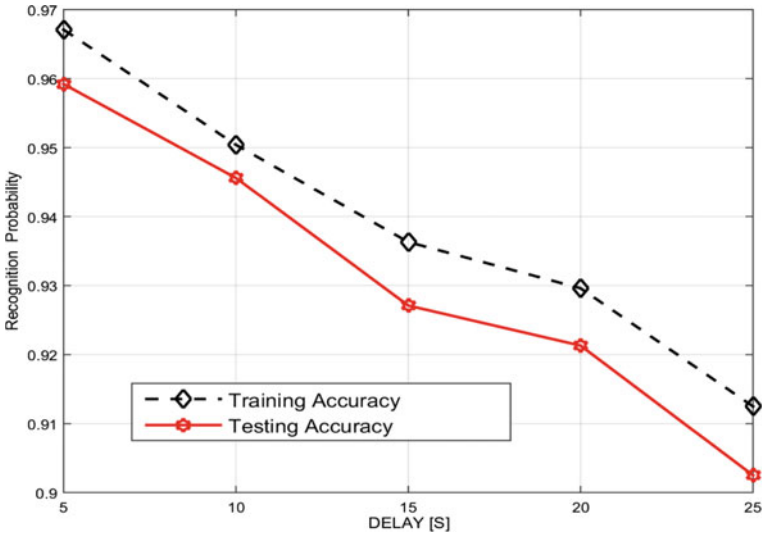



**Fig. 10.** Blind identification probability versus Doppler shifts in the SNR = 0

### 4.4 The Doppler Effect for the Radar Signals in the Watterson Channel

From the Fig. 9, we have gotten some conclusions that the delay time will worsen the classification performance. This thesis ought to be devoted to Doppler shifts. In this simulation, we lay the SNR = 0 dB and adjusted the Doppler shifts to

[5,10,15,20,25] Hz. When the Doppler shift is 5 Hz, we can obtain some excellent results. From Fig. 10, with Doppler shifts' growth, ELM algorithm has turned to be progressively inefficient. When the Doppler shift is 25 Hz, the right classification's probability is over 90%. From the results, we can easily reach the conclusion that the Doppler shifts have a negative impact on the radar signal identification, with the rise of the Doppler shifts. Across the board, the ELMs supply joint robustness against the Doppler shifts.

## 5 Conclusions

At present, signal classification, particularly the complex radar signals, has become the most important bottleneck of signal processing in electronic reconnaissance. Extracting the signal's intra-pulse features is an effective approach to solve and sort the complex radar's problem signals' recognition. In this thesis, recognition and sorting of intra-pulse modulation signals based on WT is researched. We make use of pattern recognition and machine learning for the radar modulation identification. The ELM and SAE-ELM algorithms outperform in all investigated scenarios with a lower computational complexity and faster convergence with no iterative tuning. The suggested algorithms are examined through Watterson channels, and they illustrate a extremely impressive capability of identifying different radar modulation schemes with admirable precision.

## References

1. Dobre OA, Abdi A, Bar-Ness Y, et al. Survey of automatic modulation classification techniques: classical approaches and new trends. IET Commun. 2007;1(2):137–56.
2. Zeng D, Zeng X, Cheng H, et al. Automatic modulation classification of radar signals using the Rihaczek distribution and hough transform. IET Radar Sonar Navig. 2012;6(5):322–31.
3. Lunden J, Koivunen V. Automatic radar waveform recognition. IEEE J Sel Top Sign Proces. 2007;1(1):124–36.
4. Zeng D, Xiong H, Long K, et al. A fast recognition algorithm for three kinds of intra-pulse modulation signals, 2009;(551):1–4.
5. Lopez-Risueno G, Grajal J. Multiple signal detection and estimation using atomic decomposition and EM. IEEE Trans Aerosp Electron Syst. 2006;42(1):84–102.
6. Lopez-Risueno G, Grajal J, Sanz-Osorio A. Digital channelized receiver based on time-frequency analysis for signal interception. IEEE Trans Aerosp Electron Syst. 2005;41(3):879–98.
7. Lopez-Risueno G, Grajal J, Yeste-Ojeda O. Atomic decomposition-based radar complex signal interception. IEE Proc Radar Sonar Navig. 2003;150(4):323–31.
8. Wang F, Huang Z, Zhou Y, et al. A new approach for intra-pulse modulation recognition; 2007. p. 1–5.
9. Tan JL, Sha'Amerr AZ. Automatic analysis and classification of digital modulation signals using spectogram time frequency analysis; 2007. p. 916–20.

10. Howard WR. Pattern recognition and machine learning. Academic Press; 1992. p. 461–2.
11. Huang GB, Zhu QY, Siew CK. Extreme learning machine: theory and applications. Neurocomputing. 2006;70(1):489–501.
12. Huang GB, Zhu QY, Siew CK. Extreme learning machine: a new learning scheme of feedforward. Neural Netw. 2004;2:985–90.
13. Park C S, Jang W, Nah S P, et al. Automatic modulation recognition using support vector machine in software radio applications; 2007. p. 9–12.
14. Liu X, Li R, Zhao C, et al. Robust signal recognition algorithm based on machine learning in heterogeneous networks. J Syst Eng Electron. 2016;27(2):333–42.
15. Liu X, Zhao C, Wang P, et al. Blind modulation classification algorithm based on machine learning for spatially correlated MIMO system. IET Commun. 2017;11(7):1000–7.
16. Wu HC, Saquib M, Yun Z. Novel automatic modulation classification using cumulant features for communications via multipath channels. IEEE Trans Wireless Commun. 2008;7(8):3098–105.

# Transmitting Beampattern Synthesis for Colocated MIMO Radar Based on Pulse-to-Pulse Coding

Feng Xu(✉), Xiaopeng Yang, Fawei Yang, and Xuchen Wu

School of Information and Electronics, Beijing Institute of Technology, and Key Laboratory of Electronic and Information Technology in Satellite Navigation (Beijing Institute of Technology), Ministry of Education, Beijing 100081, China
{fengxu, xiaopengyang}@bit.edu.cn

**Abstract.** Transmitting beampattern synthesis is of great importance in colocated MIMO radar since the system demands a flexible power transmitting. However, most research focuses on the optimization of transmitting waveform covariance matrix and the design of transmitting waveform with the requirement of arbitrary waveform generators. In this paper, a transmitting beampattern synthesis method based on pulse-to-pulse coding is proposed. The waveform diversity is obtained in Doppler domain since each transmitting signal is modulated to a different Doppler frequency. Then by selecting the additional phase properly, the transmitting beampattern is synthesized to emit power to specific region. This proposed method could reduce the cost of system hardware since all transmitting signals are identical except the phase. Comparing to conventional transmitting beampattern synthesis method, the system complexity and computation complexity are reduced. Simulation results verify the efficiency of our proposed method.

**Keywords:** Beampattern synthesis · Colocated MIMO Radar · Pulse-to-pulse coding

## 1 Introduction

Multiple-input and multiple-output (MIMO) radar emits orthogonal waveform via its transmitting array and receives target echoes via its receiving array simultaneously, thus improving the performance of target detection and parameter estimation. Basically, there are two categories of MIMO radar system judging by the configuration of system array, the former is separated MIMO radar system [1] and the latter is colocated MIMO radar system [2].

In MIMO radar system, the first problem is to generate orthogonal transmitting waveform. Existing waveform design methods can be classified as two types, intra-pulse coding [3] and pulse-to-pulse coding [4]. The second problem is to synthesize the transmitting beampattern to emit power into specific regions [5, 6]. These two problems are coupled since the waveform covariance matrix $\mathbf{R}$ determines the transmitting beampattern directly. Thus the problem can be summarized as a two-step constrained optimization. The first step is to synthesize a waveform covariance matrix to achieve


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_88

the desired beampattern and the second step is to design the transmitting waveform corresponding to the covariance matrix. Meanwhile, as a covariance matrix, the first constraint is this matrix **R** should be positive semidefinite. And the second constraint is all diagonal elements should be equal so that multiple radio frequency amplifiers can work at its best efficiency.

Notice that most research in waveform design is implemented in the intra-pulse, or fast-time domain. With all these strict constraints satisfied, the system requires multiple arbitrary waveform generators and digital receivers, which is unacceptable in some cases. To achieve waveform diversity with lower complexity, a MIMO radar implementation based on pulse-to-pulse, or slow-time coding, is introduced in [7]. The orthogonality of transmitting waveform is obtained in slow-time domain by a series of additional phase on each transmitting antenna. Since all transmitting waveforms share the same signal envelope, the constraints above are obviously satisfied.

In this paper, we focus on the design of colocated MIMO radar system based on pulse-to-pulse coding and its angle estimation performance. We propose a method to achieve the desired transmitting beampattern by selecting proper additional phase and the waveform diversity is obtained by several band-pass filters in slow-time domain, i.e., Doppler domain. Furthermore, the angle estimation performance is analyzed. This slow-time MIMO radar system surpasses the conventional MIMO radar system with lower system complexity and flexibility in transmitting beampattern synthesis. Simulation results verify the efficiency of proposed method.

The rest of this paper is organized as follows. The signal model of MIMO radar based on pulse-to-pulse coding is given in Sect. 2. The transmitting beampattern synthesis method is proposed in Sect. 3. In Sect. 4, several simulations are demonstrated to evaluate the proposed method. And finally, Sect. 5 concludes this paper.

## 2 Signal Model

Consider a colocated MIMO radar system with $L$ transmitters spacing at $d_t$ and $N$ receivers spacing at $d_r$. The carrier frequency is $f_0$. Assume that there are $M$ pulses during coherent integration, and each antenna transmits the same LFM signal $u_p(t)$ with an additional phase $\varphi_{lm}$, $l = 1, 2, \ldots, L$ and $m = 1, 2, \ldots, M$, as shown in Fig. 1.

The whole phase matrix is denoted by $\mathbf{\Phi}$, where

$$\Phi = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1M} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{L1} & \varphi_{L2} & \cdots & \varphi_{LM} \end{pmatrix} \tag{1}$$

The pulse recurrence time (PRT) is represented by $T_r$, and $\mathbf{w} = [w_{t1}, w_{t2}, \ldots, w_{tL}]^T$ denotes the transmitting antenna weight vector, where $(\cdot)^{\mathbf{T}}$ is the transpose operator.

Consider a far-field direct-path scatterer located at $\theta_t$ with Doppler shift $f_t$, so the signal received at this target $S(t)$ at time $t$ is given by

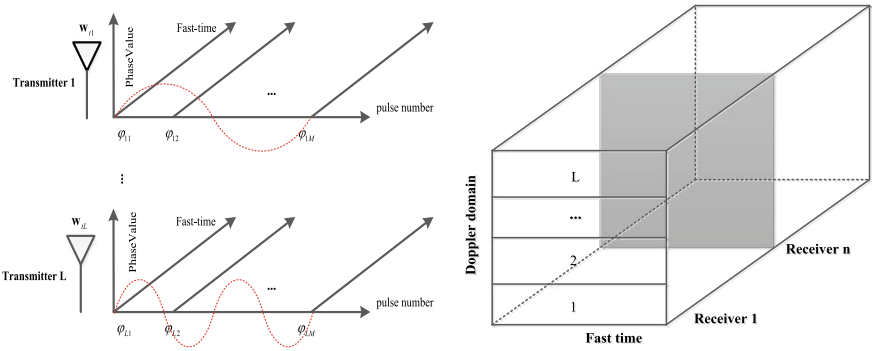



**Fig. 1.** Transmitting waveform with pulse-to-pulse coding (left) and receiving data cube after waveform diversity (right).

$$S(t) = \sum_{l=1}^{L} w_{tl} \mathrm{e}^{-j(2\pi/\lambda) l d_t \sin\theta_t} \sum_{m=1}^{M} u_p(t-(m-1)T_r) \mathrm{e}^{j2\pi(f_0 t + \varphi_{lm})} \tag{2}$$

and the received echo of this target at the $n$th receiver and slow-time $m$ is given by $\{\mathbf{X}_t\}_{nm} = \gamma_t \mathrm{e}^{j(2\pi/\lambda) n d_r \sin\theta_t} \mathrm{e}^{j2\pi f_t m T_r} S(t)$, where $\gamma_t$ represents the target radar cross-section (RCS) and $\lambda$ denotes the working wavelength.

After carrier demodulation and pulse compression, the **XXX** matrix is vectorized as represented by

$$\mathrm{vec}(\mathbf{X}_t) = \gamma_t \boldsymbol{\alpha}(\theta_t) \otimes \sum_{l=1}^{L} \boldsymbol{\chi}((\mu_1 + f_t)T_r) \times (w_{tl} \boldsymbol{\beta}_1(\theta_t)) \tag{3}$$

where

$$\begin{aligned} \boldsymbol{\alpha}(\theta_t) &= \left[1, \mathrm{e}^{-j\frac{2\pi}{\lambda} d_r \sin\theta_t}, \ldots, \mathrm{e}^{-j\frac{2\pi}{\lambda}(N-1) d_r \sin\theta_t}\right]^{\mathbf{T}} \\ \boldsymbol{\beta}(\theta_t) &= \left[1, \mathrm{e}^{-j\frac{2\pi}{\lambda} d_t \sin\theta_t}, \ldots, \mathrm{e}^{-j\frac{2\pi}{\lambda}(L-1) d_t \sin\theta_t}\right]^{\mathbf{T}} \\ \boldsymbol{\chi}(f_t T_r) &= \left[1, \mathrm{e}^{j2\pi f_t T_r}, \ldots, \mathrm{e}^{j2\pi(M-1) f_t T_r}\right]^{\mathbf{T}} \end{aligned} \tag{4}$$

and $\boldsymbol{\beta}_l(\theta_t)$ means the $l$th element of $\boldsymbol{\beta}(\theta_t)$. Here we assume that the additional phase satisfies $\varphi_{lm} = \mu_l m T_r$, and $\otimes$ denotes Kronecker product.

Obviously, $\boldsymbol{\chi}((\mu_l + f_t)T)$ indicates the process of frequency spectrum shifting corresponding to each transmitting antenna in Doppler domain. If $\mu_l$ is chosen properly, then transmitting waveform diversity can be achieved by a series of band-pass filters in Doppler domain. Moreover, these slow-time modulations shift every $L$ pulse to different Doppler domain. To simplify the process, we assume that $M/L$ is an integer, so that each Doppler domain shares the same amount of MIMO channels. So after the

Doppler domain filtering, $M$ columns' echoes on each receiver are broken into a $M/L \times L$ matrix, corresponding to $L$ transmitters.

This is the operation of waveform diversity based on pulse-to-pulse coding. And the steer vector now is denoted by $\mathbf{\Omega} = \boldsymbol{\alpha}(\theta_t) \otimes \overline{\boldsymbol{\chi}}(f_t T_r) \otimes (\mathbf{w} \odot \boldsymbol{\beta}(\theta_t))$, where $\overline{\boldsymbol{\chi}}(f_t T_r) = [1, \mathrm{e}^{j2\pi f_t T_r}, \ldots, \mathrm{e}^{j2\pi(M/L-1)f_t T_r}]^{\mathbf{T}}$, and $\odot$ denotes Hadamard product.

Notice that the Doppler domain is $L$ times smaller than conventional MIMO radar, which implies that the performance of target velocity estimation is degraded. The design of $\mu_l$ will determine the Doppler shift for each transmitter directly. Meanwhile, this additional phase modulated on the transmitting array can cause a periodic influence on the transmitting beampattern. In the following section, we will discuss the problem of designing $\mu_l$ considering waveform diversity and beampattern synthesis.

## 3 Pulse-to-Pulse Coding Method for Colocated MIMO Radar

Notice that the design of $\mu_l$ leads to a Doppler shift for each transmitter with period of $L$ pulse. Consider the transmitting signal at $m$th pulse, $m = 1, 2, \ldots, L$, then we can rewrite (2) as

$$\begin{aligned} S_m(t) &= \sum_{l=1}^{L} w_{tl} \mathrm{e}^{-j(2\pi/\lambda) l d_t \sin\theta_t} u_p(t-(m-1)T_r) \mathrm{e}^{j2\pi(f_0 t + \varphi_{lm})} \\ &= \sum_{l=1}^{L} w_{tl} \mathrm{e}^{-j(2\pi/\lambda) l d_t \sin\theta_t} \mathrm{e}^{j2\pi\varphi_{lm}} u_p(t-(m-1)T_r) \mathrm{e}^{j2\pi f_0 t} \end{aligned} \tag{5}$$

or equivalently $S_m(t) = (\mathbf{w} \odot \overline{\boldsymbol{\beta}}(\overline{\theta}_t))^T \mathbf{U}(t)$,where

$$\overline{\boldsymbol{\beta}}(\overline{\theta}_t) = [e^{j2\pi\varphi_{1m}}, e^{-j\frac{2\pi}{\lambda} d_t \sin\theta_t} e^{j2\pi\varphi_{2m}}, \ldots, e^{-j\frac{2\pi}{\lambda}(L-1) d_t \sin\theta_t} e^{j2\pi\varphi_{Lm}}]^{\mathbf{T}} \tag{6}$$

Assume that $\mu_l = l f_{\text{shift}}$ and put it into (6), we get $\overline{\boldsymbol{\beta}}(\overline{\theta}_t) = \boldsymbol{\beta}(\theta_t) \odot \boldsymbol{\beta}^*(\theta_{\text{shift}})$,where $\sin\theta_{\text{shift}} = m T_r f_{\text{shift}} \frac{\lambda}{d_t}$, and $(\cdot)^*$ means conjugate operator. This implies there is a transmitting beam pointing shift related to the pulse number $m$ and Doppler shift value $f_{\text{shift}}$. The beam pointing offset is determined by $f_{\text{shift}}$ and the beam point shift period is determined by $m$. In real system, the Doppler frequency ranges from $-1/2T_r$ to $1/2T_r$, and the process of frequency spectrum shifting in Doppler domain starts from the center of Doppler domain, i.e., the zero Doppler frequency.

So clearly this is a two-way modulation, to avoid misunderstanding, it is necessary to declare a new parameter $m' = m - L/2$,

$$\sin\theta_{\text{shift}} = m' T_r f_{\text{shift}} \frac{\lambda}{d_t},\ m' = -L/2+1, -L/2+2, \ldots, L/2 \tag{7}$$

For an uniform linear array (ULA) spacing at half the wavelength,

$$\sin\theta_{\text{shift}} = 2m'T_r f_{\text{shift}} \tag{8}$$

Furthermore, we have

$$f_{\text{shift}} \le \min\left\{\frac{1}{2m'T_r}\right\} = \frac{1}{LT_r} \tag{9}$$

when the equation in (9) is satisfied, the whole Doppler domain with Doppler frequency ranging from $-1/2T_r$ to $1/2T_r$ is uniformly divided into $L$ subset with its frequency ranging from $-1/2LT_r$ to $1/2LT_r$, thus MIMO transmitting waveform diversity can be obtained by a series of band-pass filters in Doppler domain, as shown in Fig. 1.

Without losing generality, make $f_{\text{shift}} = 1/\tilde{L}T_r$ thus $\sin\theta_{\text{shift}} = \frac{2m'}{\tilde{L}}$ and $\tilde{L} \ge L$. This equation indicates all the transmitting beam pointings for a colocated MIMO radar based on pulse-to-pulse coding. Apparently, the smaller $f_{\text{shift}}$ is, the more beam pointing exists. And it is possible to select specific beam pointing when the beam pointing is redundancy due to the design of $f_{\text{shift}}$. Although this could sacrifice the scope of measuring velocity, when it comes to the problem of slow speed target detection, this disadvantage is acceptable, not to mention the ability of synthesizing transmitting beampattern precisely.

## 4 Simulation Result

In our simulations, we assume a colocated MIMO radar system with 4 transmitters and 8 receivers. Table 1 summarizes the parameters used in this section.

**Table 1.** Simulation system parameters.

| Parameter | Value |
|---|---|
| Transmitter number | $L = 4$ |
| Receiver number | $N = 8$ |
| Pulse number | $M = 160$ |
| PRT | $T_r$ = 42 μs |
| Carrier frequency | S band |

### 4.1 Doppler Shifting and Waveform Diversity

In simulation 1, let the spacing of transmitting array and receiving array be half the working wavelength, and $\tilde{L} = 8$. The frequency spectrum shifting should appear in 4 different Doppler frequencies, thus creating four peaks in the range-Doppler plane of every receiver. All peaks share the same range bin but different Doppler bin due to the Doppler shifting. Figure 2a gives the profile of receiver result with 160 pulses after carrier demodulation, pulse compressive and coherent integration. And in Fig. 2b, the Doppler slice of this single scatterer is denoted by blue line, we also provide the cutoff frequency corresponding to all band-pass filters with red line. To obtain waveform

diversity, just utilize these band-pass filters after signal processing for every receiver. One of the waveform diversity results is demonstrated in Fig. 2c, which represents the result of first transmitter and first receiver. The measurement of target velocity and range are 9.02 m/s and 3002 m, respectively.

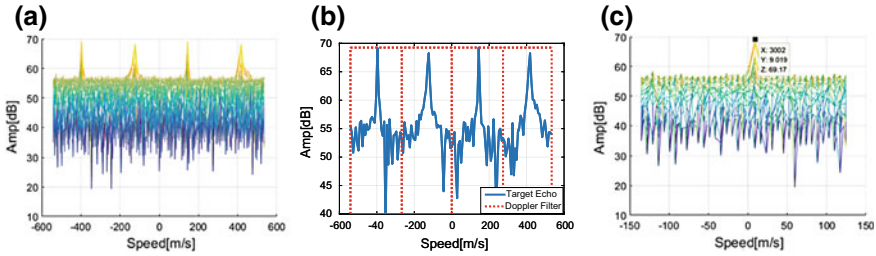



**Fig. 2.** Doppler shifting and waveform diversity result. **a** Profile of range-Doppler result. **b** Doppler slice of single target. **c** Profile of range-Doppler result after filtering.

From Fig. 2a, c, we must notice that the valid scope of velocity measurement is reduced for about 4 times after the process of Doppler filtering. However, this disadvantage can be ignored since the radar system is designed for low-speed target detection.

### 4.2 Transmitting Beampattern Synthesis

In simulation 2, we first discuss the transmitting beam pointing shift problem and the array configuration is exactly the same as simulation 1. From Fig. 3a, b, notice that 4 transmitting beampatterns appear every 4 pulses periodically, and each beampattern corresponds to one particular Doppler frequency shifting which can be found in Fig. 2. The sin value of ideal beam pointing shift should be $\pm 0.5$ and $\pm 1$. In Fig. 3a, the existing beam pointing is clearly $\pm 30\,^{\circ}$, $0\,^{\circ}$, and $90\,^{\circ}$. The sidelobe level of overlapping region is about $-3.7$ dB, thus the transmitting beampattern can be regarded as a wide beam with equal gain at slow-time domain.

Then we increase the spacing of transmitter to 2 times the working wavelength. Obviously grating lobes will be created, but a denser transmitting beampattern is obtained, as presented in Fig. 3c, d. This leads to the elaborate beampattern synthesis. Notice that although grating lobes are created in transmitting beampattern, the overall beampattern is free from angle ambiguity with the corresponding virtual array being a full-array uniform linear array.

As mentioned before, to obtain enough beam pointing, the value of Doppler shifting $f_{\text{shift}}$ should be small. In Fig. 3e, this value is decreased four times as before. And four particular beam pointings are selected to generate special transmitting beampattern. If the location of noncorrelation target is prior known, we can synthesize the transmitting beampattern to reduce the power emits to this region, thus improving the system performance. Figure 3f utilizes a $-20$ dB Chebyshev taper window to control the sidelobe.

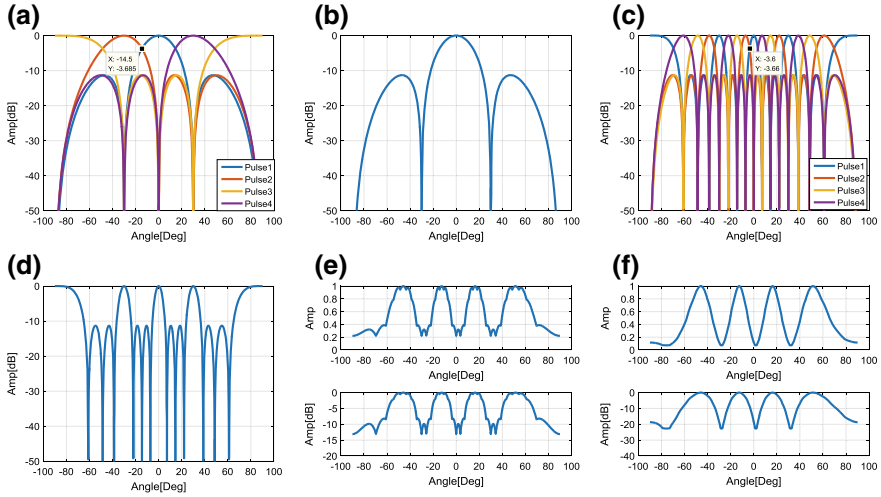



**Fig. 3.** Transmitting beampattern synthesis result. **a** Beampattern shift for 4 pulses. **b** Beampattern for one pulse. **c** Beampattern shift with transmitting antennas spacing at 2 times the working wavelength. **d** Beampattern for one pulse with transmitting antennas spacing at twice the working wavelength. **e** Transmitting beampattern synthesis with multi-beam pointing. **f** Transmitting beampattern with 20 dB Chebyshev taper window.

## 5 Summary

In this paper, a colocated MIMO radar transmitting beampattern synthesis method based on pulse-to-pulse coding is proposed. By utilizing an additional phase on each pulse, the transmitting antennas can work at different Doppler frequencies so that waveform diversity is obtained. This could reduce the cost of system hardware since there is no need to apply arbitrary waveform generators. Furthermore, by designing this phase matrix at slow-time domain, the transmitting beampattern can be synthesized to emit power to specific region, thus enhance the system performance. The only disadvantage of this proposed method is the scope of velocity loss, which is acceptable when it comes to the slow speed target detection. Simulation results verify the efficiency of our proposed method.

**Acknowledgments.** This work was supported by National Natural Science Foundation of China (Grant Nos. 61860206012, 61671065 and 31727901), and 111 Project of China (Grant No. B14010).

## References

1. Haimovich M, Blum RS, Cimini LJ. MIMO radar with widely separated antennas. IEEE Signal Process Mag. 2008;25(1):116–29.
2. Li J, Stoica P. MIMO radar with colocated antennas. IEEE Signal Process Mag. 2007;24 (5):106–14.

3. Babur G, Aubry P, Le Chevalier F. Space-time codes for active antenna systems: Comparative performance analysis. In: IET International Radar Conference 2013, Xi'an, 2013, p. 1–6.
4. Kantor JM, Bliss DW. Clutter covariance matrices for GMTI MIMO radar. In: 2010 Conference record of the forty fourth Asilomar conference on signals, systems and computers; Pacific Grove, CA, 2010, p. 1821–6.
5. Stoica P, Li J, Xie Y. On probing signal design for MIMO radar. IEEE Trans Signal Process. 2007;55(8):4151–61.
6. Ahmed S, Thompson JS, Petillot YR, Mulgrew B. Unconstrained synthesis of covariance matrix for MIMO radar transmit beampattern. IEEE Trans Signal Process. 2011;59(8): 3837–49.
7. Li J, Stoica P. MIMO radar signal processing. Wiley-IEEE Press; 2009.

# A Transfer Learning Method for Ship Target Recognition in Remote Sensing Image

Hongbo Li(✉), Bin Guo, Hao Chen, and Shuai Han

School of Electronic and Information Engineering,
Harbin Institute of Technology, Harbin, China
{drbobo, hit_hao, hanshuai}@hit.edu.cn,
18845750300@163.com

**Abstract.** In this paper, an effective approach of ship target recognition is proposed. This method based on the theory of transfer learning aims at using labeled ships with different imaging angles and different resolutions to help identifying unlabeled ships in a fixed angle. Since training ship samples and test ship samples are imaging in different angles, they obey different distributions. However, in traditional machine learning method, training data and test data obey the same distribution. In order to solve this problem, we proposed a method called mapped subspace alignment (MSA) which is different from other domain adaptation methods. While maximizing the difference between different categories, it first uses Isometric Feature Mapping (Isomap) to generate subspace and uses objective functions to spatial alignment and probabilistic adaptation. This paper focuses on the identification of three types of ships which are destroyers, cruisers, and aircraft carriers basing on MSA. The experimental results show that this method is better than several state-of-the-art methods.

**Keywords:** Ship target recognition · Transfer learning · Domain adaptation

## 1 Introduction

Ships, as important transportation carriers and military targets at sea, have an important significance for recognition. With the increase of the number of satellites, the amount of data based on optical satellite remote sensing images for ship target recognition is increasing. However, the acquisition time, weather conditions, imaging angles and image spatial resolution of the ships in the optical image are usually different. Although the quantity of ship images has increased with the development of the world, the data with labeled is scarce. What's more, labeling data is time-consuming. Therefore, it is necessary to study how to automatically identify the ship target-based few labeled data.

There are many methods of ship recognition. Zhu et al. [10] extracted shape and texture features and classified them using support vector machines. Du et al. [2] used the Hull moments and ART coefficients to extract eigenvalues for the regional shape of the ship targets, classified by KNN. There are also some methods using deep learning. For example, Tang et al. [1] combined compression domain with the deep neural network to classify. Gong et al. [4] used wavelet coefficients extracted from compression techniques and deep neural networks.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_89

However, traditional machine learning methods assume the training data and the test data under the same distribution, but the distribution is always different in the real world. This paper focuses on the identification of three kinds of ships at different imaging angles and resolutions. We propose a novel method called mapped subspace alignment (MSA), which firstly maps the source and target domains to the subspace using the Isometric Feature Mapping (Isomap) [8]. Then, we align the source and target domains in the subspace and maintain the differences between different categories in the target domain and the source domain at the same time. There are two main differences between the proposed method and others: (i) MSA first uses the Isomap method to map the source and the target domain data to the corresponding subspace and then aligns the spatial distribution to achieve a better recognition effect. (ii) The variances of the samples in the source and target domains are kept to a maximum, and the alignment is implemented by using probability distribution adaptation. Experimental results show that the method proposed in this paper is superior to the state-of-the-art methods.

## 2 Related Work

At present, domain adaptation methods can be divided into two categories. One is to perform probability distribution adaptation. For example, Pan et al. [7] proposed transfer component analysis (TCA) which matched marginal distributions in the source and target domains. Long et al. developed [6] joint distribution adaptation (JDA), which adapted both marginal and conditional data distributions. Another approach is to map the source and target domains to a low-lattice subspace and then find the transformation matrix for spatial alignment. For example, geodesic methods [5] find a path along the subspace manifold. In subspace alignment [3, 9], the subspaces can be aligned by computing the linear map that minimizes the Frobenius norm of the difference between them.

## 3 Mapped Subspace Alignment Method

MSA maps the source and target domains to the subspace by Isomap and aligns two domains in the subspace. It maintains the distinguishability of different kinds of ships, so that the same kind of target in the mapped space is more similar.

Supposing the labeled source domain data is $\boldsymbol{X}_s \in \mathbb{R}^{D\times n_s}$, the unlabeled target domain data is denoted as $\boldsymbol{X}_t \in \mathbb{R}^{D\times n_t}$, where $D$ is the dimension of features, and $n_s$ and $n_t$ are the numbers of ship samples in the source and target domains, respectively. Because the source and target domains are ships under different angles, they have different distributions. $P_s(\boldsymbol{X}_s)$ and $P_t(\boldsymbol{X}_t)$ are distributions of the source domain and the target domain, respectively. Then, the source domain is $\{(\boldsymbol{x_i}, \boldsymbol{y_i})\}_{\boldsymbol{i}=1}^{\boldsymbol{n_s}}$, $\boldsymbol{x_i} \in \mathbb{R}^D$ and the target domain data is $\{(\boldsymbol{x}_j)\}_{j=1}^{n_t}$, $x_j \in \mathbb{R}^D$. We want to learn the labels of ships in the target domain by using the data in the source domain.

### 3.1 Feature Subspace Generation

After extracting the HOG feature which is robust to the effects of variations of ships, the Isomap method is used to generate the feature subspace. Isomap is a kind of manifold learning which is often used to identify faces with different angles. In order to achieve a better recognition effect, firstly we use Isomap to get the feature subspaces of the source and target domains. There are four main steps. Firstly, construct neighborhood maps. Secondly, calculate the shortest path. Thirdly, calculate d-dimensional embedding. In this paper, the HOG features in the source and target domains are mapped to the d-dimensional subspace, respectively, using the Isomap method, and then subspace alignment is performed.

### 3.2 Feature Subspace Alignment

In order to keep the category distinguishability maximum, we should keep the distance between class maximum in the source domain and variance maximum in the target domain which has no labels. Hence, the distance between classes and the variance maximization can be achieved as follows.

$$L = \max_{\boldsymbol{T}_s} Tr(\boldsymbol{T}_s^T \boldsymbol{S}_b \boldsymbol{T}_s) + \max_{\boldsymbol{T}_t} Tr(\boldsymbol{T}_t^T \boldsymbol{S}_t \boldsymbol{T}_t) \tag{1}$$

where $\boldsymbol{T}_s$ and $\boldsymbol{T}_t$ are projection matrixes, $\boldsymbol{S}_b = \sum_{c=1}^{C} n_s^{(c)} (\boldsymbol{u}_s^{(c)} - \bar{u}_s)(\boldsymbol{u}_s^{(c)} - \bar{u}_s)^T$ is the between-class scatter matrix of the source domain data, $\boldsymbol{S}_t = \boldsymbol{X}_t \boldsymbol{M}_t \boldsymbol{X}_t^T$ is the target domain scatter matrix, $\boldsymbol{u}_s^{(c)} = \frac{1}{n_s^{(c)}} \sum_{i=1}^{n_s^{(c)}} \boldsymbol{x}_i^{(c)}$, $\bar{u}_s^{(c)} = \frac{1}{n_s} \sum_{i=1}^{n_s} \boldsymbol{x}_i$, $n_s^{(c)}$ is the number of the source samples in class, and $\boldsymbol{M}_t = \boldsymbol{I}_t - \frac{1}{n_t} \boldsymbol{1}_t \boldsymbol{1}_t^T$ is the centering matrix.

MMD distance has been a popular method that is used in many transfer learning approaches [6, 7]. Because the target domain has no labels, we use the class conditional distribution $P_t(\boldsymbol{X}_T|\boldsymbol{y}_t)$ to approximate $P_t(\boldsymbol{y}_t|\boldsymbol{X}_T)$ [6]. And we minimize the joint probability between domains.

$$\mathrm{D}(\boldsymbol{X}_s, \boldsymbol{X}_t) = \min_{\boldsymbol{T}_s, \boldsymbol{T}_t} \sum_{c=0}^{C} \left\| \frac{1}{n_s^{(c)}} \sum_{\boldsymbol{x}_i \in \boldsymbol{X}_s^{(c)}} \boldsymbol{T}_s^T \boldsymbol{x}_i - \frac{1}{n_t^{(c)}} \sum_{\boldsymbol{x}_j \in \boldsymbol{X}_t^{(c)}} \boldsymbol{T}_t^T \boldsymbol{x}_j \right\|_F^2 \tag{2}$$

By mathematical principles of matrix and regularization, Eq. (2) can be formalized as follows.

$$\mathrm{D}(\boldsymbol{X}_s, \boldsymbol{X}_t) = \min_{T_s, T_t} Tr\left( \begin{bmatrix} \boldsymbol{T}_s^T & \boldsymbol{T}_t^T \end{bmatrix} \begin{bmatrix} \boldsymbol{K}_s & \boldsymbol{K}_{st} \\ \boldsymbol{K}_{ts} & \boldsymbol{K}_t \end{bmatrix} \begin{bmatrix} \boldsymbol{T}_s \\ \boldsymbol{T}_t \end{bmatrix} \right) \tag{3}$$

where

$$\boldsymbol{K}_s = \boldsymbol{X}_s\left(\boldsymbol{M}_s + \sum_{c=1}^{C}\boldsymbol{M}_s^{(c)}\right)\boldsymbol{X}_s^T,\ \boldsymbol{M}_s = \frac{1}{n_s^2}\mathbf{1}_s\mathbf{1}_s^T,\ \left(\boldsymbol{M}_s^{(c)}\right)_{ij} = \begin{cases} \frac{1}{\left(n_s^{(c)}\right)^2} & \boldsymbol{x}_i, \boldsymbol{x}_j \in \boldsymbol{X}_s^{(c)} \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{K}_t = \boldsymbol{X}_t\left(\boldsymbol{M}_t + \sum_{c=1}^{C}\boldsymbol{M}_t^{(c)}\right)\boldsymbol{X}_t^T,\ \boldsymbol{M}_t = \frac{1}{n_t^2}\mathbf{1}_t\mathbf{1}_t^T,\ \left(\boldsymbol{M}_t^{(c)}\right)_{ij} = \begin{cases} \frac{1}{\left(n_t^{(c)}\right)^2} & \boldsymbol{x}_i, \boldsymbol{x}_j \in \boldsymbol{X}_t^{(c)} \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{K}_{st} = \boldsymbol{X}_s\left(\boldsymbol{M}_{st} + \sum_{c=1}^{C}\boldsymbol{M}_{st}^{(c)}\right)\boldsymbol{X}_t^T,\ \boldsymbol{M}_{st} = -\frac{1}{n_s n_t}\mathbf{1}_s\mathbf{1}_t^T,\ \left(\boldsymbol{M}_{st}^{(c)}\right)_{ij} = \begin{cases} -\frac{1}{n_s^{(c)} n_t^{(c)}} & \boldsymbol{x}_i \in \boldsymbol{X}_s^{(c)}, \boldsymbol{x}_j \in \boldsymbol{X}_t^{(c)} \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{K}_{ts} = \boldsymbol{X}_s\left(\boldsymbol{M}_{ts} + \sum_{c=1}^{C}\boldsymbol{M}_{ts}^{(c)}\right)\boldsymbol{X}_t^T,\ \boldsymbol{M}_{ts} = -\frac{1}{n_s n_t}\mathbf{1}_t\mathbf{1}_s^T,\ \left(\boldsymbol{M}_{ts}^{(c)}\right)_{ij} = \begin{cases} -\frac{1}{n_s^{(c)} n_t^{(c)}} & \boldsymbol{x}_j \in \boldsymbol{X}_s^{(c)}, \boldsymbol{x}_i \in \boldsymbol{X}_t^{(c)} \\ 0 & \text{otherwise} \end{cases}$$

Therefore, we use different mapping matrices to maximize the difference between different categories and use joint probabilistic adaptation to perform edge and conditional probabilistic adaptations. At the same time, we use the F norm to make the mapped space as similar as possible. It can be expressed as follows.

$$D(\boldsymbol{T}_s, \boldsymbol{T}_t) = \min_{\boldsymbol{T}_s, \boldsymbol{T}_t}\left\|\boldsymbol{T}_s^T - \boldsymbol{T}_t^T\right\|_F^2 \tag{4}$$

Our objective function is to maximize the following formula:

$$F = \max L/(\mathrm{D}(\boldsymbol{X}_s, \boldsymbol{X}_t) + \mathrm{D}(\boldsymbol{T}_s, \boldsymbol{T}_t)) \tag{5}$$

Our goal is to find a mapping matrix that meets the requirements.

$$F = \max_{T_s, T_t} Tr\left(\begin{bmatrix}\boldsymbol{T}_s^T & \boldsymbol{T}_t^T\end{bmatrix}\begin{bmatrix}\boldsymbol{S}_b & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}_t\end{bmatrix}\begin{bmatrix}\boldsymbol{T}_s \\ \boldsymbol{T}_t\end{bmatrix}\right) / Tr\left(\begin{bmatrix}\boldsymbol{T}_s^T & \boldsymbol{T}_t^T\end{bmatrix}\begin{bmatrix}\boldsymbol{K}_s & \boldsymbol{K}_{st} \\ \boldsymbol{K}_{ts} & \boldsymbol{K}_t\end{bmatrix}\begin{bmatrix}\boldsymbol{T}_s \\ \boldsymbol{T}_t\end{bmatrix}\right) \tag{6}$$

where $I \in R^{d\times d}$ is the identity matrix. To optimize (5), we rewrite $[T_s \quad T_t]$ as $T$. By mathematical principles of matrix and regularization, the objective function and corresponding constraints can be written as:

$$\max_{W} Tr\left(\boldsymbol{T}^T\begin{bmatrix}\boldsymbol{S}_b & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}_t\end{bmatrix}\boldsymbol{T}\right) \text{ s.t.} Tr\left(\boldsymbol{T}^T\begin{bmatrix}\boldsymbol{K}_s & \boldsymbol{K}_{st} \\ \boldsymbol{K}_{ts} & \boldsymbol{K}_t\end{bmatrix}\boldsymbol{T}\right) = 1 \tag{7}$$

The Lagrange function of (7) is

$$Z = Tr\left(\boldsymbol{T}^T\begin{bmatrix}\boldsymbol{S}_b & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}_t\end{bmatrix}\boldsymbol{T}\right) + Tr\left(\left(\boldsymbol{T}^T\begin{bmatrix}\boldsymbol{K}_s & \boldsymbol{K}_{st} \\ \boldsymbol{K}_{ts} & \boldsymbol{K}_t\end{bmatrix}\boldsymbol{T} - \boldsymbol{I}\right)\boldsymbol{\Phi}\right) \tag{8}$$

where $\boldsymbol{\Phi} = \mathrm{diag}(\lambda_1, \ldots, \lambda_k)$. By setting the derivative $\frac{\partial Z}{\partial T} = 0$, we can get the optimal transformation matrix $T$ including $T_s$ and $T_t$ can be obtained by solving this equation and finding its $k$ eigenvectors.

## 4 Experiment and Analysis

In this section, we conduct experiments with our state-of-the-art methods and analyze the results. We compare our approach MSA with six related methods for image classification. They are 1 nearest neighbor classifier, transfer component analysis (TCA) [7] +1NN, joint distribution adaptation (JDA) [6] +1NN, geodesic flow kernel (GFK) [5] +1NN, subspace alignment (SA) [3] +1NN, and our proposed method without Isomap. In order to explore the role of Isomap, we set up the controlled trial which is MSA method without Isomap.

The data sources are ships with different angles under different imaging conditions obtained from Google Maps. There are one hundred destroyers, one hundred cruisers, and one hundred aircraft carriers with different views (angled offset) and different resolutions in the source domain. There are forty destroyers, forty cruisers, and forty aircraft carriers with no angle offset in the target domain. We use the labeled images in the source domain to identify the unlabeled images in the target domain. Part of images in data set are shown in Fig. 1. Ships are imaging in a fixed angle in the first line and have different angles in the second line. It can be seen that there are great differences between the sample data under different angles.

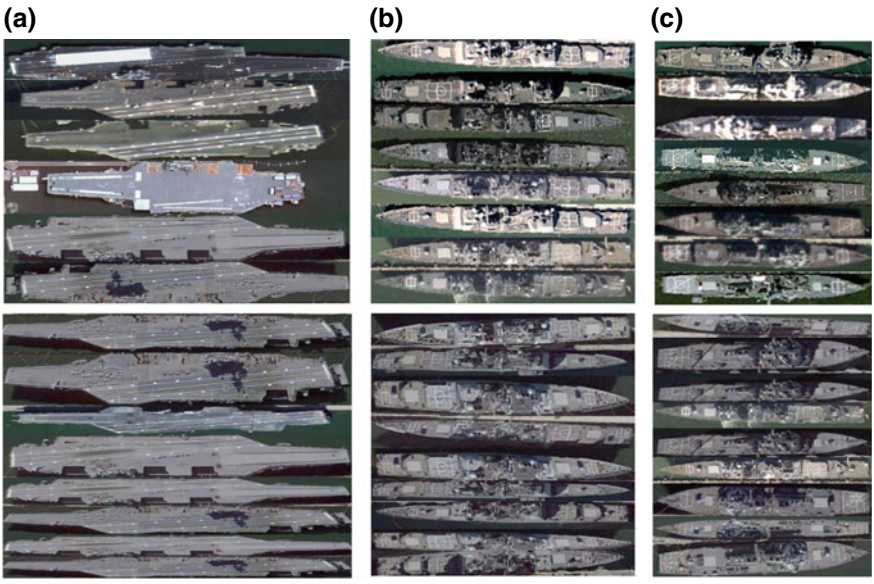



**Fig. 1.** Part of images in data sets. **a** Aircraft carriers. **b** Cruisers. **c** Destroyers

For all the methods proposed above, some of these methods have parameters which can be determined through fivefold cross-validation and the best results will be admitted. In this paper, these methods involve subspace bases $k$ or $d$. We set $k = d = 100$ to take experiments. The iteration number is set to be $T = 15$. Then, we can get

the results. We use MMD distances and classification accuracy which is calculated as the ratio of the number of correctly identified samples to the total number of sample input in the target domain to describe the performance of approaches. The results of all methods acting on the ship data sets are illustrated in Table 1. For better interpretation, the results are visualized in Fig. 3.

**Table 1.** Classification accuracy

| Classification | Cruiser (%) | Destroyer (%) | Aircraft carrier (%) |
|---|---|---|---|
| MSA | 95 | 92.5 | 97.5 |
| JDA | 90 | 87.5 | 92.5 |
| MSA without Isomap | 85 | 82.5 | 92.5 |
| SA | 85 | 80 | 87.5 |
| GFK | 80 | 77.5 | 87.5 |
| TCA | 77.5 | 75 | 85 |
| 1 NN | 65 | 62.5 | 72.5 |

From Table 1 and Fig. 2, we can observe that MSA achieves much better performance than the other six methods on statistical significance. What is more, based on Table 1, the average classification accuracy of the seven methods is 95, 90, 85.83, 84.1, 81.6, 79.1, and 65%, respectively. First of all, the average classification accuracy of MSA on the three ships is 95% and the performance improvement is 5% compared to the method JDA. And JDA performs better than the other methods. So, our method has the best performance in different views and different resolution problems. Next, the average classification accuracy of MSA without Isomap is 85%. Compared to MSA, we can see Isomap plays an important role in the proposed method. Due to the different angles of the samples, Isomap, the nonlinear dimensionality reduction, can map HOG features in the source and target domains to a better feature space.

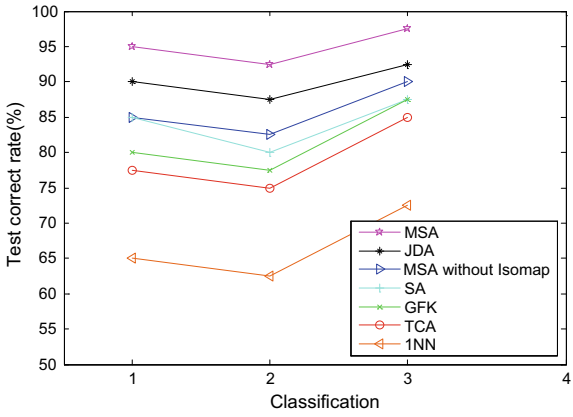



**Fig. 2.** Classification accuracy

In order to further verify the effectiveness of MSA, we discuss their performance by MMD distances. It has shown the MMD distance and the accuracy of these methods with increasing iteration in Fig. 3.

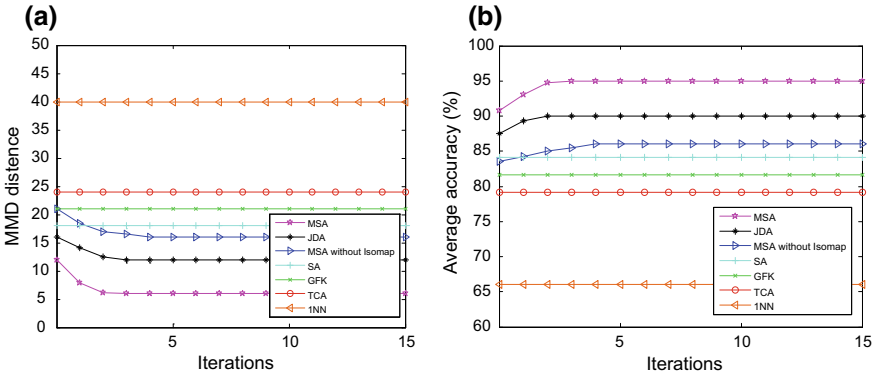



**Fig. 3.** Performance of methods. **a** MMD distance and **b** classification accuracy

The results are shown in Fig. 3, based on which we can draw the following observations. Firstly, the distance between source domain and target domain is large when only using 1NN. Secondly, the transfer learning approaches can reduce the MMD distance. And the MMD distance of MSA is obviously smaller than JDA and the other methods, since MSA that we proposed achieves the best performance for cross-domain problems.

## 5 Conclusion

In this paper, we proposed an MSA approach for transfer learning problem. MSA not only aligns the space between source domain and target domain but also minimizes the marginal and conditional distribution discrepancy. At the same time, the differences between categories are maintained to achieve better classification results. Thus, it can significantly improve the accuracy of cross-domain target recognition problems. Extensive experiments demonstrate that MSA is effective in cross-domain problems and performs better than the state-of-the-art adaptation methods. In the future, we will continue to explore cross-domain identification issues, for example, transfer knowledge from images of different sources, such as SAR images and hyperspectral images.

**Acknowledgments.** This work was supported in part by a grant from the Defense Industrial Technology Development Program (No. JCKY2016603C004).

## References


1. Bi F, Liu F, Gao L. A hierarchical salient-region based algorithm for ship detection in remote sensing images. In: Lecture notes in electrical engineering; 2010, vol. 67. Berlin/Heidelberg, Germany: Springer.
2. Du Q, Zhang Y, Liu W et al. Ship target classification based on Hu invariant moments and ART for maritime video surveillance. In: International conference on transportation information and safety; 2017. p. 414–9, https://doi.org/10.1109/ictis.2017.8047799.
3. Fernando B, Habrard A, Sebban M et al. Unsupervised visual domain adaptation using subspace alignment. In: IEEE international conference on computer vision; 2014, vol. 58, no. 8. p. 2960–7, https://doi.org/10.1109/iccv.2013.368.
4. Gong B, Shi Y, Sha F et al. Geodesic flow kernel for unsupervised domain adaptation. In: IEEE conference on computer vision and pattern recognition; 2012. p. 1–8.
5. Grauman K. Geodesic flow kernel for unsupervised domain adaptation. In: IEEE conference on computer vision & pattern recognition; 2012, vol. 157, no. 10. p. 2066–2073.
6. Long M, Wang J, Ding G et al. Transfer feature learning with joint distribution adaptation. In: IEEE international conference on computer vision; 2014. p. 2200–7.
7. Pan SJ, Tsang IW, Kwok JT, Yang Q, et al. Domain adaptation via transfer component analysis. IEEE Trans Neural Netw. 2011;22(2):199–210.
8. Tenenbaum JB, de Silva V, Langford J. A global geometric framework for nonlinear dimensionality reduction. Science. 2000;290(5500):2319–38.
9. Zhang J, Li W, Ogunbona P. Joint geometrical and statistical alignment for visual domain adaptation. In: IEEE conference on computer vision and pattern recognition; 2017.
10. Zhu C, Zhou H, Wang R, et al. A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features. IEEE Trans Geosci Remote Sens. 2010;48(9):3446–56. https://doi.org/10.1109/tgrs.2010.2046330.

# Through Wall Human Detection Based on Support Tensor Machines

Li Zhang(✉), Wei Wang, Yu Jiang, Dan Wang, and Min Zhang

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, 300387 Tianjin, China
zhangli94_tjnu@163.com, weiwang@tjnu.edu.cn

**Abstract.** Through wall human detection based on ultra-wideband (UWB) radar is a challenging task due to the complex environment. In this case, it is not enough for the research sample that is only with high cost. In this paper, we propose a novel algorithm named support tensor machines (STMs). It avoids the overfitting in pattern recognition. We conduct two groups of experiments on high-dimensional and small-sampling data. The experimental results prove that our method not only achieves the desired results, but also saves plenty of computation time.

**Keywords:** Support tensor machines (STMs) · Through wall human detection · Alternating projection algorithm · High-dimensional and small-sampling data

## 1 Introduction

Through wall human detection has important applications in anti-terrorism, post-disaster search, rescue and other occasions [1]. In the past decades, many researchers have been focusing on studying human detection approaches. M. Chia used UWB radar with the FCC's mark to acquire the non-invasive sensing of heartbeat and breathing rate [2]. In [3], the fast Fourier transform (FFT) and S transform are used to detect and identify the human's vital signs. In [4], a continuous-wave subsurface radar was utilized to describe the theoretical approach and experiments. M Jelen, EM Biebl presented a low-cost method by using Doppler radar sensors to solve the issue. Experimental results showed that the presence of people and the frequency of breathing and heartbeat can be well detected [5].

However, the training samples with high cost and high-dimensional characteristic are not enough due to the special conditions. Through wall human detection becomes a challenging task that might cut down the success ratio for the hostage rescue. To alleviate the problem, the support tensor machines (STMs) algorithm is proposed for through wall human detection.

The paper is organized as follows. Section 2 briefly introduces theoretical tensor knowledge. In Sect. 3, the support tensor machines are presented in detail. Numerical experimental process and results are described in Sect. 4. Finally, Sect. 5 concludes the paper.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_90

## 2 Theory

### 2.1 Tensor Theory and Representation

Tensor was first proposed by Woldemar Voigt in the 1790s [6]. Einstein used tensor to study general relativity. This makes tensor received extensive attention and recognition. Nowadays, tensor has been widely used in many fields such as physics, engineering, and computer science. Mathematically, defined a multidimensional array as a tensor, which is expressed as $A \in R^{L_1 \times L_2 \times,\ldots,L_N}$. The sequence of the tensor $A$ is $N$. For instance, the zero-order tensor is regarded as a scalar, the first-order tensor is considered as a vector, and the second-order tensor is seen as a matrix in the sense of isomorphism. This paper briefly reviews the relevant concepts for matrices and tensors in multilinear algebra.

**Definition 2.1 (*Tensor Product*)** The tensor product $X \circ Y$ of a tensor $X \in R^{I_1 \times I_2 \times,\ldots,\times I_N}$ and another tensor $Y \in R^{I'_1 \times I'_2 \times,\ldots,I'_M}$ is defined by:

$$(X \circ Y)_{i_1,i_2,\ldots,i_N,i'_1,i'_2,\ldots,i'_M} = x_{i_1,i_2,\ldots,i_N} y_{i'_1 i'_2,\ldots,i'_M}. \tag{1}$$

for all values of the indexes.

**Definition 2.2 (*Inter Product*)** The inner product of two same-sized tensors $X, Y \in R^{I_1 \times I_2 \times,\ldots,\times I_N}$ is presented as the sum of the products of all the same elements:

$$\langle X, Y \rangle = \sum_{I_1=1}^{L_1} \cdots \sum_{I_N=1}^{L_N} X_{I_1 \times I_2 \times,\ldots,I_N} Y_{I_1 \times I_2 \times,\ldots,I_N} \tag{2}$$

**Definition 2.3 (*Mode*-d *Product*)** The Mode-$d$ product is a tensor $A \in R^{L_1 \times L_2 \times,\ldots,\times L_N}$ and a matrix $U \in R^{L_d \times L'_d}$ operation, defined by

$$(A \times_d U)_{l_1 \times l_2 \times,\ldots,l_{d-1} \times l'_d \times l_{d+1} \times,\ldots,\times l_N} = \sum_{l'_d} (A_{l_1 \times l_2 \times,\ldots,l_{d-1} \times l_d \times l_{d+1} \times,\ldots,\times l_N} U_{l'_d \times l_d}) \tag{3}$$

where $l_1, l_2, \ldots, l_N$ represent the index of each mode. It is necessary to mention that the size of tensor $A$ in $d$-th mode is equal to the size of the matrix $U$ in the second mode. The mode-$d$ product should be seen as incorporating the message of input matrix and tensor. In the meantime, as the first mode size of $U$ changes, dimension-reduction and data filtering should be implemented by the operation. Figure 1 shows a visual illustration of mode 1 product between a matrix and a third-order tensor.

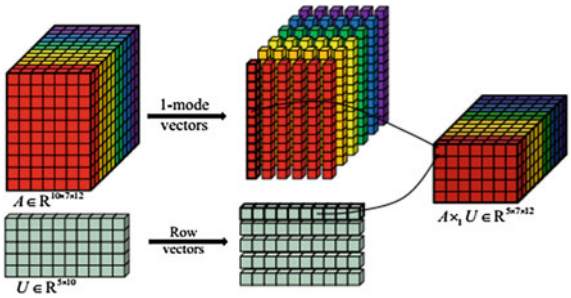



**Fig. 1.** Visual illustration of the mode-$d$ product.

### 2.2 Tensor Decomposition

Generally speaking, there are two particular tensor decompositions using on higher-order extensions of the matrix singular value, which are Tucker decomposition and CANDECOMP/PARAFAC (CP) decomposition. Herein, the CP decomposition shows that a tensor can be expressed as the sum of rank-one tensors. And the Tucker decomposition is an advanced principal component analysis that expresses a tensor as a core tensor multiplied by a matrix along each pattern. Besides, tensor decompositions include PARAFAC2, INDSCAL, CANDELINC and so on. The paper introduces the CP decomposition as a concrete instance.

Given a third-order tensor $X \in R^{I \times J \times K}$, we want to write it as

$$X \approx \sum_{r=1}^{R} a_r \circ b_r \circ c_r \tag{4}$$

Here, $R$ presents a positive integer, and
$a_r \in R^I, b_r \in R^J$, and $c_r \in R^K$, for $r = 1, \ldots, R$
Elementwise, (4) is written as
$x_{ijk} \approx \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr}$, for $i = 1, \ldots, I,\ j = 1, \ldots, J,\ k = 1, \ldots, K$
As is illustrated in Fig. 2.

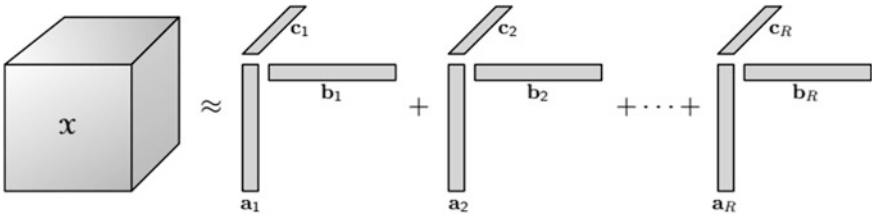



**Fig. 2.** CP decomposition of a three-way array

## 3 Support Tensor Machine Algorithm

The proposed algorithm for the high-dimensional and small-sampling data issue is described in [support tensor machine algorithm]. In this paper, we choose second-order tensor as the instance.

**Algorithm** Support tensor machine algorithm

**Input**: Split the training samples $T = \{(X_1, y_1), (X_2, y_2), \ldots, (X_l, y_l)\}$, where $X_i \in R^{n_1 \times n_2}$ is a second-order tensor and the associated with class label $y_i (y_i \in \{\pm 1\}),\ i = 1, 2, \ldots, l$.

**Output**: Testing for the accuracy.

1. Initialized the vector u is the unit row vector. Let $\beta_1 = \|u\|^2$ and $x_i = X_i^T u$.
2. Solve the parameter $v$:
   The solution of $v$ can be solved by the following least squares support vector machine.

$$\min_{v,b,\xi} \frac{1}{2}\beta_1 \|v\|^2 + C\sum_{i=1}^{l} \xi_i$$
$$\text{s.t. } y_i(v^T x_i + b) \geq 1 - \xi_i, \xi_i \geq 0,\ i = 1, \ldots, l.$$

3. Solve the parameter u: After obtaining the parameter $v$, we let $\beta_2 = \|v\|^2$ and $\tilde{x}_i = X_i v$. Then according to formula of the quadratic programming problem getting the parameter $u$.

$$\min_{u,b,\xi} \frac{1}{2}\beta_2 \|u\|^2 + C\sum_{i=1}^{l} \xi_i$$
$$\text{s.t. } y_i(\langle u, x_i \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0,\ i = 1, \ldots, l.$$

4. Alternately iterative calculate the parameter of $u$ and $v$. Using the formula $W = uv^T$ to obtain the parameter $W$. This first of all satisfies the following conditions: $\|u_i - u_{i-1}\| \leq \varepsilon, \|v_i - v_{i-1}\| \leq \varepsilon$ and $\|b_i - b_{i=1}\| \leq \varepsilon$ iterations stop. Among them, $\varepsilon$ is a very small number close to 0. Otherwise, we need to repeat steps 2 and 3.
5. End loop

## 4 Experimental Process and Results

The section aims at verifying the feasibility and effectiveness of the support tensor algorithm for human target detection. The unmanned data is regarded as the normal data. The status of some people after the wall is regarded as abnormal data. All the samples were normalized and selected certain attribute value. The experiments were implemented in MATLAB (version 2012a) software on a Windows 7 personal computer (Intel(R) Core(TM) i5-4250U CPU (1.30 GHz), memory 4 GB).

To evaluate the effectiveness of the proposed method, we compare it with SAE. Specifically, we performed a similar test in which human target detection for one group is 50 test samples, and another group is with 110 test samples. Generally, people focus on the accuracy and the efficiency of human detection. In this paper, the optimal parameter C is calculated by cross-validation. Besides, the range of parameter $C$ is $\{2^{-6}, 2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^{0}, 2^{1}, 2^{2}, 2^{3}, 2^{4}, 2^{5}, 2^{6}\}$. We need to initialize the vector of $u$ and $v$. The results are presented in Table 1.

**Table 1.** Experimental results of STM and SAE under the same conditions

| The number of training data | | The number of test data | | STM | | SAE | |
|---|---|---|---|---|---|---|---|
| Positive | Negative | Positive | Negative | Accuracy (%) | Time (s) | Accuracy (%) | Time (s) |
| 40 | 5 | 40 | 10 | 86 | 0.12 | 100 | 0.82 |
| | 10 | | | 96 | 0.12 | 100 | 0.87 |
| | 15 | | | 98 | 0.11 | 100 | 0.92 |
| 100 | 5 | 100 | 10 | 91.82 | 0.35 | 100 | 3.38 |
| | 10 | | | 96.36 | 0.36 | 100 | 3.50 |
| | 15 | | | 97.27 | 0.34 | 100 | 2.99 |

## 5 Conclusions

In this paper, we have demonstrated the STM algorithm used in through wall human detection. Numerical experiment shows that it has been widely explored on a variety of high-dimensional and small-sampling data sets and helps in better detecting human targets. Compared with SAE algorithm, it can compute using less time.

In the future, we will explore other effective methods to identify human target through the wall. Not only track the human targets but also can identify human's motions more accurately and efficiency.

**Acknowledgments.** This paper is supported by Natural Youth Science Foundation of China (61501326, 61401310), the National Natural Science Foundation of China (61731006) and Natural Science Foundation of China (61271411). It is also supported by Tianjin Research Program of Application Foundation and Advanced Technology (15JCZDJC31500), and Tianjin Science Foundation (16JCYBJC16500). This work was also supported by the Tianjin Higher Education Creative Team Funds Program.

## References

1. Li J, Zeng Z, Sun J, Liu F. Through-wall detection of human being's movement by UWB radar. IEEE Geosci Remote Sens Lett. 2012;9(6):1079–83.
2. Chia M, Leong S, Sim C, Chan K. Through-wall UWB radar operating within FCC's mask for sensing heart beat and breathing rate. Proc EURA. 2005;3:267–70.
3. He W, Hu X. Research on text classification based on tensor space model. J Hefei Univ. 2010;33(12).
4. Bugaev AS, Chapursky VY, Ivashov SL. Through wall sensing of human breathing and heart beating by monochromatic radar. In: Proceedings of 10th international conference on ground penetrating radar, Delft, The Netherlands; 2004. p. 291–4.
5. Jelen M, Biebl EM. Multi-frequency sensor for remote measurement of breath and heartbeat. Adv Radio Sci. 2006;4:79–83.
6. Yang B. Research and application of machine learning algorithm based tensor representation; 2014.

# Radar Signal Waveform Recognition Based on Convolutional Denoising Autoencoder

Zhaolei Liu[1], Xiaojie Mao[2(✉)], and Zhian Deng[2]

[1] Nanjing Research Institute of Electronics Technology, Nanjing 210039, China
[2] College of Information and Telecommunication, Harbin Engineering University, Harbin 150001, China
15645178051@163.com,1324501764@qq.com

**Abstract.** To solve the problem of the low recognition rate of the existing methods at low signal-to-noise ratio (SNR), we propose a novel method of radar signal waveform recognition. In this method, we extract the time-frequency images (TFIs) of radar signals through Cohen class time frequency distribution. Then, we introduce convolutional denoising autoencoder (CDAE) to denoise and repairs the TFIs. Finally, we build a convolutional neural network (CNN) to identify the TFIs of radar signals. Simulation experiment shows that the proposed method can identify 12 kinds of radar signal waveforms, and the overall probability of successful recognition (PSR) is 95.4% when the SNR is −7 dB.

**Keywords:** Radar signal recognition · Cohen class time frequency distribution · Convolutional denoising autoencoder · Convolutional neural network

## 1 Introduction

Radar signal waveform recognition has very important significance for electronic reconnaissance and anti-reconnaissance [1,2]. However, with the advent of LPI radars, it is very difficult to identify radar signal waveforms using traditional methods. At present, there are two difficult problems in radar signal waveform recognition. First, the signal-to-noise ratio (SNR) of the received signal is getting lower and lower, while the threshold of the SNR of the traditional recognition method is higher. Second, the waveform of radar signal is more and more diverse, which requires the method to identify more types of radar signals.

For these reasons, some scholars have proposed some new radar signal waveform recognition methods. In [3], the Rihaczek distribution (RD) and the Hough transform (HT) are used to extract the features of radar signal, including the ratio of the minimum to the maximum of the HT and the peak number of the HT of the real part of the RD, to identify radar waveforms. This method can


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_91

recognize linear frequency modulation (LFM), frequency shift keying (FSK), binary phase shift keying (BPSK) and monopulse (MP) signals. However, this method requires manually extracting recognition features which are limited and have poor generalization performance. In order to achieve automatic recognition of radar signal waveforms and improve recognition performance, some scholars have introduced deep learning for radar signal waveform recognition. In [4], a deep neural network model composed of multiple restricted Boltzmann machines is designed for radar signal recognition. It has good performance for eight kinds of radar signals (CW, PSK, DPSK, FSK, MP, LFM, and NLFM). In [5], convolutional neural network (CNN) is introduced to identify the time-frequency images (TFIs) of radar signal for radar signal waveform recognition. However, these methods have unstable performance at lower SNR. Therefore, how to further reduce the impact of noise and improve the recognition performance of the method is a problem that needs to be studied.

In this paper, we propose a novel method of radar signal waveform recognition. The proposed method introduces the convolutional denoising autoencoder (CDAE) to denoise and repair the TFI of radar signal, which largely removes the effects of noise on signal waveform recognition. Therefore, it is easy to build a CNN to recognize the TFI of radar signal and complete radar signal waveform recognition at lower SNR. The method can identify 12 types of radar signal waveforms, including LFM, sinusoidal frequency modulation (SFM), 2FSK, 4FSK, dual frequency modulation (DLFM), even quadratic frequency modulation (EQFM), multiple linear frequency modulation (MLFM), BPSK, QPSK, MP and composite modulation (LFM-BPSK, 2FSK-BPSK). Simulation results show that the proposed method achieves an overall probability of successful recognition (PSR) of 95.4% when the SNR is −7 dB.

## 2 Technical Background

### 2.1 Radar Signal Model

The received radar signal is composed of a modulated signal and noise. Its model can be written as

$$s(t) = x(t) + n(t) \tag{1}$$

where $s(t)$ and $x(t)$ are received signal and modulated signal, respectively. $n(t)$ is random noise, which is generally assumed to be additive white Gaussian noise.

The modulated signal $s(t)$ is given by

$$x(t) = A\mathrm{rect}(t/T)\exp\left[-\mathrm{j}\left(2\pi f_c t + \phi(t) + \phi_0\right)\right] \tag{2}$$

where $A$ is the amplitude and $T$ is the pulse width. $f_c$ and $\phi_0$ are the carrier frequency and the initial phase, respectively. $\phi(t)$ is the phase function, which determines the modulation type of the signal.

### 2.2 Cohen Class Time Frequency Distribution

Cohen class time frequency distribution is a kind of time-frequency analysis, which can effectively extract the time-frequency images of the signal. It has strong anti-noise performance and high time-frequency resolution. In this paper, we will use Cohen class time frequency distribution to extract the time-frequency images of the received signals. Cohen class time frequency distribution is defined as

$$C(t,\omega) = \frac{1}{4\pi^2} \iint AF(\tau, v)\phi(\tau, v) \exp(-\mathrm{j}vt - \mathrm{j}\omega\tau)\, dvd\tau \tag{3}$$

$$AF(\tau, v) = \int x\left(u + \frac{\tau}{2}\right) x^*\left(u - \frac{\tau}{2}\right) \exp(\mathrm{j}vu)\, du \tag{4}$$

where $x(u)$ is the received signal. $AF(\tau, v)$ is the ambiguity function. $\tau$ and $v$ are the time delay and the frequency shift, respectively. $\phi(\tau, v)$ is a kernel function.

In this paper, the kernel function is designed as

$$\phi(\tau, v) = \exp\left[-\left(\alpha\tau^2 + \beta v^2\right)\right] \tag{5}$$

where $\alpha$ and $\beta$ are two parameters that can adjust the shape of the kernel function. In this paper, $\alpha = 0.0005$, $\beta = 0.001$.

In the actual signal processing, the signal is discrete, so we use the following method to calculate $C(n, k)$. The discrete autocorrelation function of the signal is expressed as

$$R(m, n) = x(n + m)\, x^*(n - m) \tag{6}$$

Thus,

$$AF(m, k) = \mathrm{FFT}_n\{R(m, n)\} \tag{7}$$

In (7), $\mathrm{FFT}_n$ denotes the Fourier transform integral of the variable $n$. Through (6) and (7), we can obtain $C(n, l)$.

$$C(n, l) = \mathrm{FFT}_m\{\mathrm{IFFT}_k\{\mathrm{FFT}_n\{x(n + m)\, x^*(n - m)\} \cdot \phi(m, k)\}\} \tag{8}$$

Using (8) to calculate $C(n, l)$ can avoid convolution in time domain, and use fast Fourier transform to reduce computation.

## 3 Convolutional Denoising Autoencoder

### 3.1 Denoising Autoencoder

Autoencoder is a kind of neural network. After training, it can try to copy the input to output. Its basic structure is shown in Fig. 1.

In Fig. 1, $\mathbf{x}$ and $\mathbf{y}$ represent input and output, respectively. Autoencoder has a hidden layer $\mathbf{h}$. The process of mapping from $\mathbf{x}$ to $\mathbf{h}$ is called encoding and is represented as $\mathbf{h} = f(\mathbf{x})$. The process of mapping from $\mathbf{h}$ to $\mathbf{y}$ is called decode and is represented as $\mathbf{y} = g(\mathbf{h})$. The purpose of autoencoder is to make output

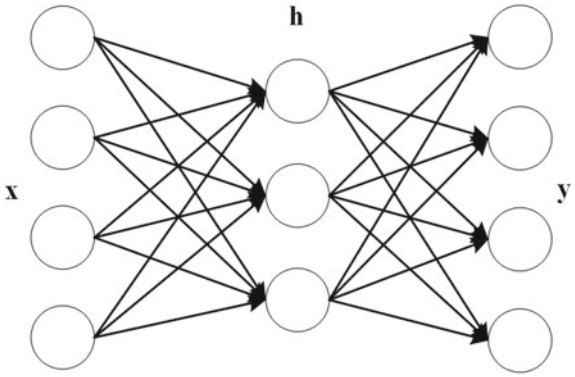



**Fig. 1.** Basic structure of autoencoder.

equal to input as much as possible. Its calculation method is the same as feed-forward neural network. The learning process of autoencoder can be described as minimizing a loss function $L$.

$$\min L\left(\mathbf{x}, g\left(f\left(\mathbf{x}\right)\right)\right) \tag{9}$$

In general, the mean variance function is selected as the loss function $L$. Thus,

$$L\left(\mathbf{x}, g\left(f\left(\mathbf{x}\right)\right)\right) = \left\|g\left(f\left(\mathbf{x}\right)\right) - \mathbf{x}\right\|_2^2 \tag{10}$$

Through gradient descent method to optimize (9), autoencoder will be trained and it is often able to learn the useful features of data.

If the input of autoencoder is replaced with $\tilde{\mathbf{x}}$, that is formed by adding some noise to $\mathbf{x}$, and the output is replaced with $\mathbf{x}$, it becomes a denoising autoencoder. The purpose of denoising autoencoder is to recover clean data $\mathbf{x}$ from the noise data $\tilde{\mathbf{x}}$. Therefore, the learning process of denoising autoencoder can be expressed as

$$\min L\left(\mathbf{x}, g\left(f\left(\tilde{\mathbf{x}}\right)\right)\right) \tag{11}$$

However, if it is used for image denoising, there will be problems with large amount of calculation and long training time. If the connection mode between layer and layer of denoising autoencoder becomes a convolution operation, it becomes a convolutional denoising autoencoder, which can greatly reduce the amount of calculation and improve the denoising performance [6]. In this paper, we will use the convolutional denoising autoencoder to remove the noise of the TFIs of radar signals.

### 3.2 Convolutional Network

Convolutional network has an excellent performance in image classification, speech recognition, object detection, and so on. Convolution operation and pooling operation are the most important parts of the convolutional network.

The output resulting from the convolution operation is called a feature map. The activity of feature map $a$ in layer $k$ is given by

$$I_{k,a} = f\left(\sum_b w_{k,ab} \otimes I_{k-1,b} + \theta_{k,a}\right) \tag{12}$$

where $I_{k-1,b}$ are feature maps that provide input to $I_{k,a}$ and $b$ denotes the $b$-th feature map. $\otimes$ denotes the convolution operation. The function $f$ is the activity function, which is the ReLU $f(x) = \max(0, x)$ in this paper. $\theta_{k,a}$ is a bias parameter.

Pooling operation is called downsampling, that is, in a region to take a specific value as the output value. Its mathematical expression is as shown in Eq. (13).

$$I_{k,a} = down\left(I_{k-1,b}\right) \tag{13}$$

where $down(\cdot)$ is a downsampling function, usually using average downsampling or maximum downsampling.

Convolution and pooling operations will be used in convolutional denoising autoencoder and convolutional neural network designed in this paper. There is also an operation called deconvolution or upsampling, which is the converse operation of the convolution operation. We will use it in the CDAE.

## 4 Radar Signal Waveform Recognition Based on Convolutional Denoising Autoencoder

The framework of the method of radar signal waveform recognition based on convolutional denoising autoencoder has been given in Fig. 2. First, the received radar signal is transformed into the time-frequency domain by Cohen class time frequency distribution, and the TFI of the signal is extracted. Then, the TFI of the signal is resized to $64 \times 64$ and normalized. Next, CDAE is used to denoise and repair the TFI, eliminating the effects of noise. Finally, use a CNN to classify radar signal waveform.

In Fig. 3, we give the structure and parameters of CDAE designed for denoising and repairing TFI. The CDAE includes three convolutional layers and three deconvolution layers. Its convolution kernel size is $3 \times 3$. Each convolutional layer is followed by a batch normalization layer that can speed up the training of the network and prevent the gradient from disappearing.

Figure 4 shows the structure of the CNN used in this paper. CNN has four convolutional layers and a fully connected layer and its output layer is 12 types. Its convolution kernel size is $3 \times 3$.

## 5 Simulation and Analysis

In this section, the performance of the proposed method will be analyzed. There are 12 types of simulated signals, including LFM, MP, SFM, BPSK, QPSK,

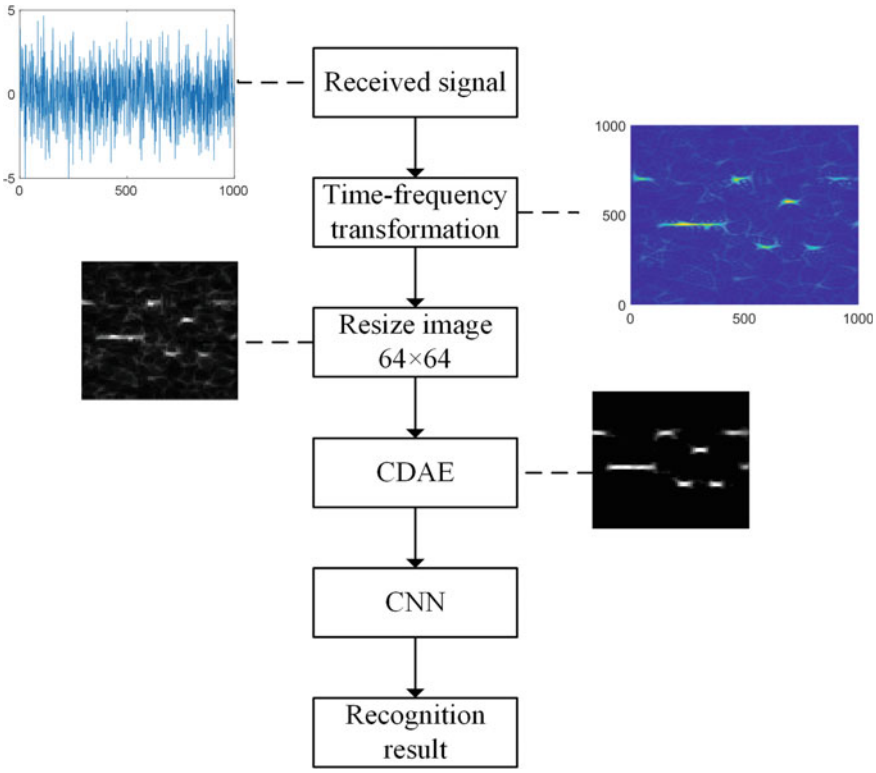



**Fig. 2.** Framework of the method proposed in this paper.

2FSK, 4FSK, DLFM, EQFM, MLFM, LFM-BPSK, and 2FSK-BPSK. When the SNR varies from −6 to 10 dB, each type of signal generates 100 samples per SNR as a training set sample. When the SNR varies from −10 to 10 dB, each type of signal generates 100 samples per SNR as a test set sample.

### 5.1 CDAE Denoising Performance Analysis

In Fig. 5, we show the effect of this CDAE on the processing of the TFI of 4FSK signal. From the figure, it is easy to see that when the TFI of 4FSK signal is processed by the CDAE, the noise is removed and the TFI is repaired. This shows that the CDAE can effectively remove the noise of TFI and can also effectively repair the distortion of TFI caused by noise. This has greatly improved the anti-noise performance of the method proposed in this paper.

### 5.2 PSR of the Proposed Method

In Fig. 6, the relation curve between the PSR of signal and the SNR is given. From Fig. 6, it can be clearly seen that the PSR of signal is positively correlated

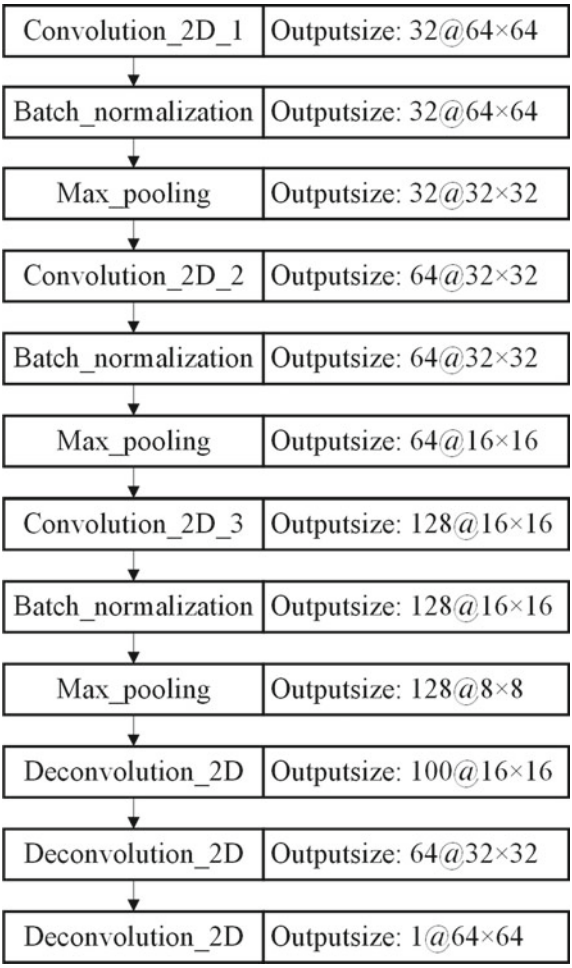



**Fig. 3.** Structure of CADE designed for denoising and repairing the TFI of radar signal.

with the SNR. When the SNR is greater than or equal to −3 dB, the proposed method achieves a PSR of 100% for the recognition of 12 types of signals. When the SNR is greater than or equal to −6 dB, the proposed method achieves a PSR of more than 90% for 12 types of signals. With the further reduction of SNR, the proposed method still maintains more than 90% of PSR with SNR of −8 dB for SFM, 2FSK, MLFM, LFM-BPSK, and 2FSK-BPSK signals. This shows that the proposed method has good anti-noise performance and good generalization performance.

In order to further analyze the performance of this method, we compare the method with Ref. [5]. The average PSRs of the proposed method and Ref. [5] method for 12 types of signals are shown in Table 1. From Table 1, it can be seen

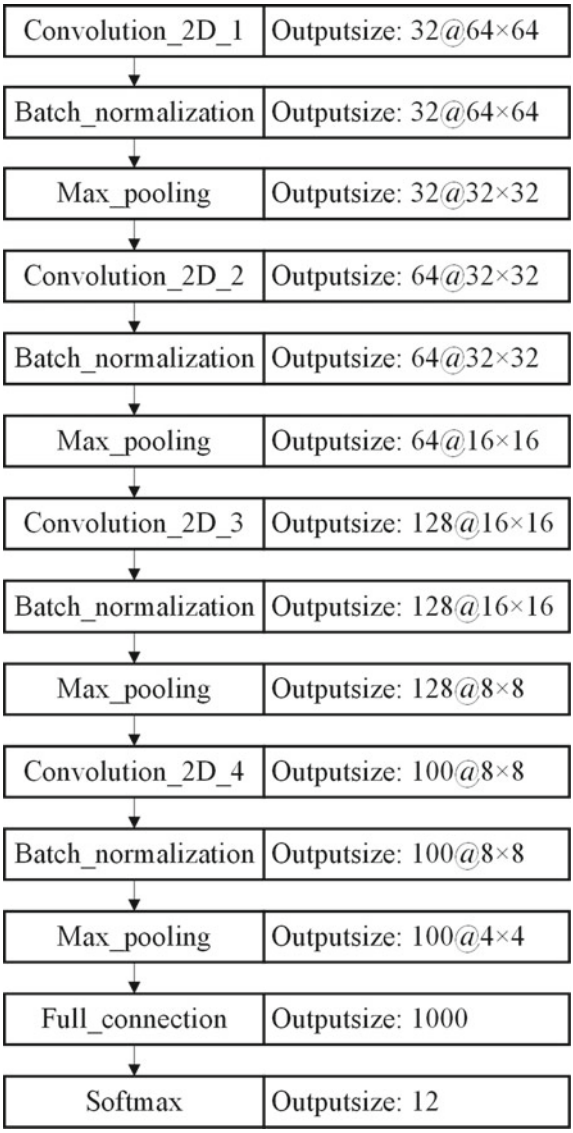
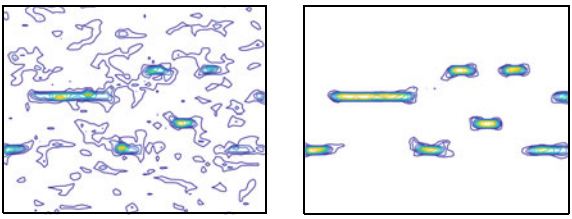



**Fig. 4.** Structure of CNN designed to identify the TFI of radar signal.

that the average PSR of the reference [5] method is lower than 90% when the SNR is less than −2 dB. In comparison, the proposed method still maintains an average PSR of 96% at a SNR of −6 dB. This shows that the proposed method is effective with strong anti-noise performance and high recognition rate.



(a) Before denoising. (b) After denoising.

**Fig. 5.** Effect of CDAE on the TFI of 4FSK signal when the SNR is −6 dB.

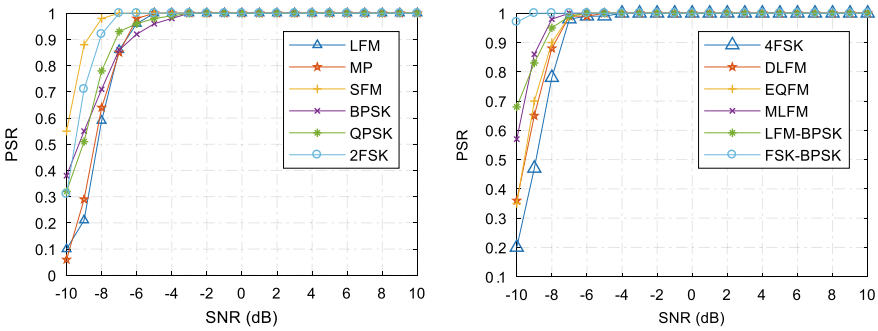



**Fig. 6.** Relation curve between the PSR of signal and the SNR.

**Table 1.** Average PSR for 12 types of signals

| Method | SNR (dB) | | | | | |
|---|---|---|---|---|---|---|
| | −10 | −8 | −6 | −4 | −2 | 0 |
| Reference [5] (%) | 15.3 | 27.8 | 50.6 | 81.2 | 92.7 | 95.7 |
| This paper (%) | 40.4 | 84.3 | 96.2 | 98.3 | 100 | 100 |

## 6 Conclusion

In this paper, we propose a method of radar signal waveforms recognition based on convolutional denoising autoencoder. This method can effectively identify 12 types of radar signals at low SNR. It has strong anti-noise performance and generalization performance. In addition, this paper introduces CDAE to denoise and repair the TFI of signal, which provides a new method for the denoising and repair of TFIs. It can be used to preprocess the TFI of signal in similar method, which can significantly improve the performance of the method.

## References

1. Latombe G, Granger E, Dilkes FA. Fast learning of grammar production probabilities in radar electronic support. IEEE Trans Aerosp Electron Syst. 2010;46(3):1262–89.
2. Gupta M, Hareesh G, Mahla AK. Electronic warfare: issues and challenges for emitter classification. Defence Sci J. 2011;61(3):228–34.
3. Zeng D, Zeng X, Cheng H, Tang B. Automatic modulation classification of radar signals using the Rihaczek distribution and Hough transform. IET Radar Sonar Navig. 2012;6(5):322–31.
4. Zhou D, Wang X, Tian Y, Wang R. A novel radar signal recognition method based on a deep restricted Boltzmann machine; 2017
5. Zhang M, Diao M, Guo L. Convolutional neural networks for automatic cognitive radio waveform recognition. IEEE Access. 2017;5:11074–92.
6. Jain V, Seung HS. Natural image denoising with convolutional networks. In: International conference on neural information processing systems; 2008. p. 769–76

# Soil pH Classification Based on LSTM via UWB Radar Echoes

Tiantian Wang(✉), Fangqi Zhu, and Jing Liang

University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave, Chengdu 611731, China
wtt_inf@sina.com

**Abstract.** This paper proposed a new method to classify soil pH based on long short-term memory (LSTM) via ultra-wideband (UWB) radar echoes. The main contribution of this paper is to provide a solution by incorporating the LSTM into the field experiment related to UWB based on soil pH echoes. Five types of UWB soil echoes with different pH values are collected and investigated using LSTM approach. Finally, the analysis of results shows that LSTM method presents a good classification performance with a short execution time and the data features do not need to be extracted manually. The high accuracy rate also shows that LSTM method is beneficial to the study of other soil parameters.

**Keywords:** Soil pH · UWB radar echoes · LSTM

## 1 Introduction

The emerging of the precision and smart agriculture calls for the requirement of the advanced and automatic retrieval of the soil parameters, and one of the categories is the soil pH. The investigation about the mapping between soil pH and remote sensing signals has been investigated for decades and a common method is to analyze soil reflected echoes[1–3]. In 2004, Lambot [1] proposed a ground-penetrating radar (GPR) model in soil parameter retrieval via the UWB radar echos, which is one of the most significant studies about soil parameters retrieval. From this, the hidden information of soil pH can be collected by collecting soil reflected echoes. Different soil pH values contribute to different soil electric properties. Analyzing the signal inversion in the frequency domain, the soil electric properties are identified from GPR signals [2]. Liu [3] discussed the soil channel model based on different soil pH values and statistical signal processing results show that soil pH can affect the UWB radar signal propagation and the echoes' textures of soil with different pH values are diverse.

However, these previous work did not provide a specific way to obtain the value of soil pH directly, just like [4], X. Liu used ANN (artificial neural network) to analyze soil reflected echoes and retrieved the specific values of soil moisture. Many studies are needed for more information about the soil pH classification


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_92

than before from UWB echoes to push the development of the precision and smart agriculture. Therefore, in our work, the LSTM method is proposed to solve the problem. Our contribution has been summarized as follows:

- We collected five types of UWB soil echoes with different pH values and 2000 samples for each pH are gathered as raw data. The soil pH values are measured manually as the labels and the raw data are preprocessed.
- LSTM is applied on soil pH classification different from the previous study and outstanding classification performance is attained. The curve reflecting the predicted results of training data converges fast and achieves a high accuracy rate.

The rest of the paper is arranged as follows. In Sect. 2, the field experiment is introduced and the instruments for our work are described. In Sect. 3, the processing of raw data sets is described and the classification method LSTM is proposed. In Sect. 4, we analyze the LSTM layers and the training performance. Finally, we conclude in Sect. 5.

## 2 Field Experiment

The field experiment is carried out continuously from March 2018 to April 2018 in Chengdu, Sichuan province of China. We selected bare soil with the area of approximate 50 $m^2$ in the west of Chengdu as shown in Fig. 1a. We took the flat area as field experiment environment to ensure that there is no obvious change in terms of the roughness of the surface of the bare soil. In the experiment, the values of pH are adjusted manually by adding acetin or caustic soda in order to change the soil pH gradually, with the auxiliary data alignment from the pH meter (set as the ground truth data for reference).

The UWB radar module P440 is applied in soil echoes collection due to its portable, low energy consumption and good penetration performance [5]. We mounted the P440 module on the shelf and the antennas are 0.9 m away from the bare soil surface [6]. The two antennas are paralleled to the ground which can make the UWB signal propagates perpendicular to the ground. The apparatus setting is demonstrated in Fig. 1b.

The reflected signals from the surface of the soil were combined in a rake receiver structure and then it was loaded into PC via the USB port. The unit sample in the signal is spaced at a 61 ps increment. As no direct relation can be obtained between the UWB radar echoes and the soil pH value, specifical calibration is necessary. For each UWB radar measurements, we labeled the UWB signal with the pH data collected by the pH 3000. The pH 3000 is employed to measure the soil pH, which is shown in Fig. 1c. There are two reagents, pH 4.0 and pH 7.0, which are used for verifying the probe. During a experiment, we collected different lumps of soil and thin them up, then plugging the probe in diluent. Finally, we obtained the soil pH from the pH 3000's screen. The main parameters of the pH 3000 are listed in Table 1.

During our experiment, we collected 10,000 groups of UWB soil echoes for five different soil pH values (i.e., pH = 5.00, pH = 7.23, pH = 8.19, pH = 8.95 and pH = 11.00), 2000 groups per pH. The discrete time length of each echo is 480 ns.

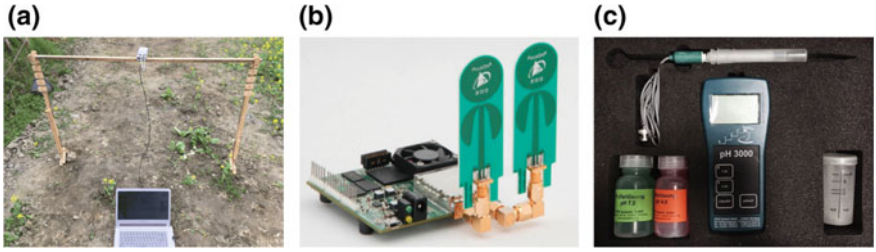



**Fig. 1. a** Field environment; **b** P440-MRM; **c** pH 3000

**Table 1.** The specifications of pH 3000

| Parameters | Values |
|---|---|
| Range | 0–14 pH |
| Precision | 0.01 pH |
| Power Requirements | 9 V |
| Probe | Special inserted glass probe |

## 3 LSTM in Soil pH Classification

### 3.1 Data Preprocessing

The UWB signal propagates 0.5 m in soil, therefore, we set the discrete time length of each echo 480 ns, which is enough for us to carry out our experiment, but the length of each echo is not all the soil surface reflecting echo. We calculated the efficient propagation time and three fragments of the reflected echoes are required to be trimmed as follows. The overall situation is shown in Fig. 2.

The first is the part related to the antenna coupling effects and the propagation in the air. In the air, the propagation speed $v$ is the same as the speed of light $c$ and the propagation distance $s$ is the distance from UWB to soil surface-0.9 m. The time delay $\tau$ of two sample points is 61 ps. Using (1) and (2), we can compute the accuracy count $m$ of sample points about the first fragment.

$$t = \frac{2s}{v} \tag{1}$$

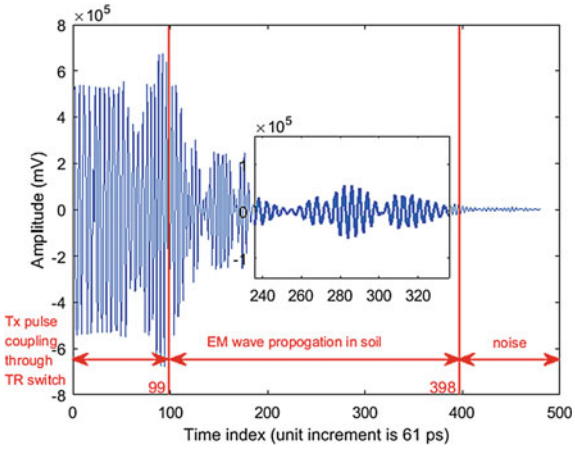



**Fig. 2.** Complete UWB reflected signal (pH = 7.23 of the bare soil as the example).

$$m = \frac{t}{\tau} \tag{2}$$

The second is the efficient fragment, which is used for simulation. According to [7], the relative dielectric constant of soil $\epsilon_r$ is between 4 and 40 when the UWB is working. The propagation speed of UWB signal is updated, shown in formula (3). The propagation distance $s$ under soil is 0.5 m. The same as the first computation process, we can get the number $m$ of sample points about the second fragment. Through computering, the efficient fragment is calculated, from the time index 99 to 398. For each case, the data are normalized as follows (4).

$$v = \frac{c}{\sqrt{\epsilon_r}} \tag{3}$$

$$x(t) = \frac{x(t) - mean(X)}{std(X)} \tag{4}$$

### 3.2 LSTM Algorithm for Soil pH Classification

LSTM is a special RNN, which is an effective tool to simulate the process of selective reading, writing, and forgetting by the gated control mechanism [8]. The LSTM framework used in our model consists of three modules (input, forget, and output) [9–11]. The input module is inspired by selecting the type of information we want to update. The forget module is a selective loss mechanism, for reducing the contribution to the learning task [12]. The output module is a output for current state information. Figure 3 shows the detailed information of LSTM framework.

In the figure, $t$ stands for the current time and $t-1$ stands for the past time. $i_t$, $f_t$, $o_t$ are the output parameters of input module, forget module, and output

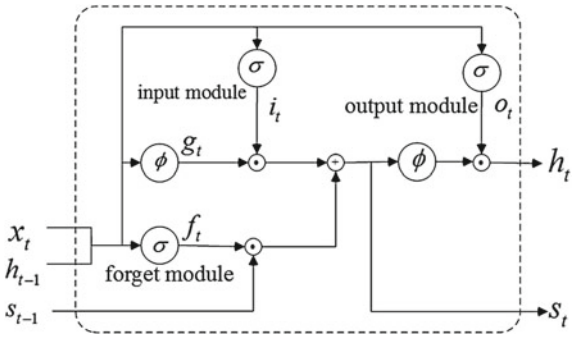



**Fig. 3.** LSTM framework

module and $s_t$ is the state of unit. $h_t$ moves forward through time and $g_t$ moves backward through time. $\sigma$ is the sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$ and $\phi$ is the hyperbolic tangent function, $\phi(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. During our simulation, the input data $x_t$ are the soil echoes and the parameter setting of the LSTM layers are listed in Table 2.

**Table 2.** The parameter setting of LSTM

| Layer | Parameter |
|---|---|
| Sequence input | 300 dimensions |
| BiLSTM | 100 hidden units |
| Fully connected | 5 fully connected layer |
| Softmax | Softmax function |
| Classification output | Crossentropy |

The number of features of the data set, 300, is taken as the size entered by the LSTM and $y$ is a categorical vector of labels "1", "2", "3", "4", "5", which correspond to the five categories echoes with different pH values. Specify a bidirectional LSTM layer with 100 hidden units and output the last element of the sequence. The softmax function is used for the output layer, mapping a scalar output to a probability output, which can be written as follows:

$$\boldsymbol{\sigma}(\boldsymbol{z})_j = \frac{e^{z_j}}{\sum_i^K e^{z_i}} \tag{5}$$

The cross-entry is a loss function for estimate the multi-class case. Let there be $K$ classes and the target variable $y \in 1, 2, \ldots, K$, by leveraging the softmax function, consider the estimation of the distribution for data sets of $K$ classes,

the discrete multi-category loss function can be expressed as follows:

$$J(\boldsymbol{\theta}) = -\frac{1}{M}\left[\sum_{i}^{M}\sum_{j}^{K} I\{y^{(i)} = j\}\log\frac{e^{\theta_j^T x^{(i)}}}{\sum_{l}^{K} e^{\theta_l^T x^{(i)}}}\right] \tag{6}$$

Where $M$ stands for $M$ dimension, $I$ stands for the indicative function, when the value in parentheses is true, the function returns 1, otherwise 0.

## 4 Simulation and Analysis of Soil pH Classification

In this section, we consider all of the five scenarios with different soil pH values. Figure 4 displays five types of soil echoes with different pH values. For each case, we collect 2000 collections and each collection is the sequence with length of 300 ns after data preprocessing. We use the ratio of 80%: 20% as the train-test split, and we input the normalized data as the input and treat the problem as the five-category classification.

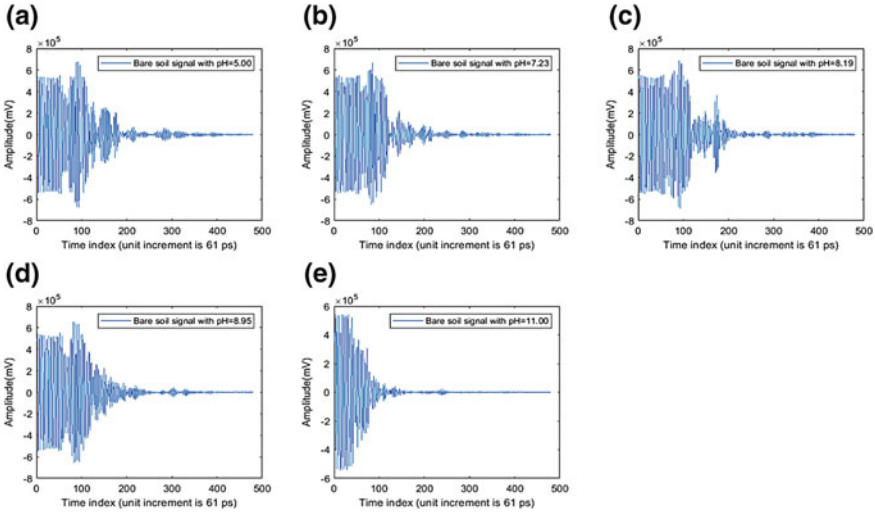



**Fig. 4.** The time index of soil signal with different pH values **a** pH = 5.00; **b** pH = 7.23; **c** pH = 8.19; **d** pH = 8.95; **e** pH = 11.00

The running time of code is 28 s and the training progress is shown in Fig. 5. From the figure, we can see that the curve converges at around 440 iterations (or around 11 epoch) and the accuracy rate of training is **97%**. The iteration is one-time training using the batch size 100 of the training set, while an epoch is a full pass through the entire training data set. The detail classification accuracy

rates of testing for each pH value are listed in Table 3. The accuracy rate is calculated by the list equation:

$$\boldsymbol{Accuracy(\%)} = \frac{\sum_i^M I\{y_{real}^i = y_{predict}^i\}}{M} * 100\% \tag{7}$$

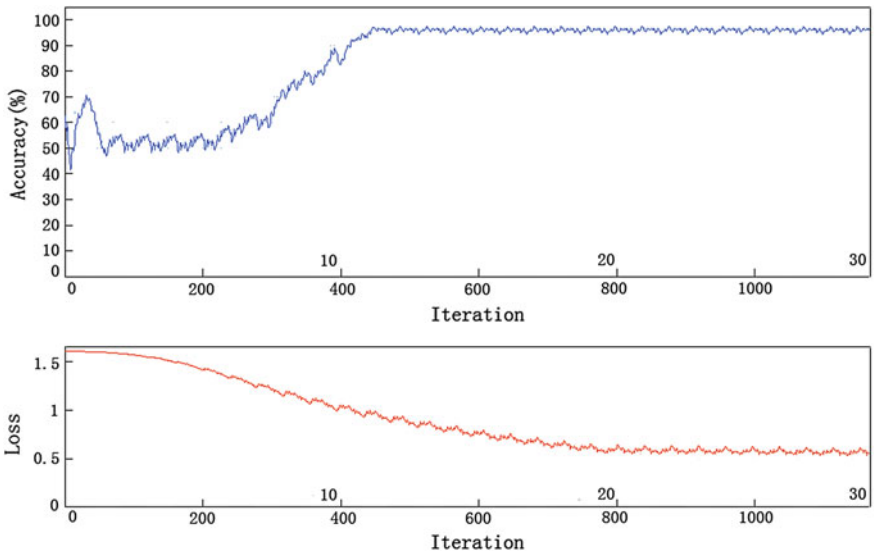



**Fig. 5.** The training progress

**Table 3.** The testing accuracy rates of each pH value

| pH value | 5.00 | 7.23 | 8.19 | 8.95 | 11.00 |
|---|---|---|---|---|---|
| Accuracy rate (%) | 100 | 98.77 | 93.46 | 100 | 80 |

## 5 Conclusion

In this paper, a soil pH classification method is proposed. We totally investigated five values of soil pH and the raw data after normalized are classified by utilizing LSTM. The LSTM approach plays a role that extracts the features of UWB soil echoes for five soil pH values and using these features to classify UWB soil echoes. We collected 2000 groups of UWB soil echoes for per soil pH value and used the

ratio 80%:20% as train-test split. The classification results show that the accuracy rates of soil echoes with different pH values are 100% (pH = 5.00), 98.77% (pH = 7.23), 93.46% (pH = 8.19), 100% (pH = 8.95), and 80% (pH = 11.00). In the future work, this approach can be extended for the complete soil parameters retrieval task and applied to the real-time in situ application.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61671138, 61731006), and was partly supported by the 111 Project No. B17008.

## References

1. Lambot S, Slob EC, van den Bosch I, Stockbroeckx B, Vanclooster M. Modeling of ground-penetrating radar for accurate characterization of subsurface electric properties. IEEE Trans Geosci Remote Sens. 2004;42(11):2555–68.
2. Lambot S, Slob EC, van den Bosch I, Stockbroeckx B, Scheers B, Vanclooster M. Estimating soil electric properties from monostatic ground-penetrating radar signal inversion in the frequency domain. Water Resour Res. 2004;40(4)
3. Liu M, Zhu F, Liang J. Channel modeling based on ultra-wide bandwidth (UWB) radar in soil environment with different pH values. In: 2014 Sixth international conference on wireless communications and signal processing (WCSP); 2014. IEEE, p. 1–6.
4. Liang J, Liu X, Liao K. Soil moisture retrieval using UWB echoes via fuzzy logic and machine learning. IEEE Internet Things J. 2017
5. Dewberry B. Monostatic radar module reconfiguration and evaluation tool (mrm-ret) pulson ®. Time Domain Corp; 2012
6. Liang J, Zhu F. Soil moisture retrieval from UWB sensor data by leveraging fuzzy logic. Accepted for publication on IEEE Access, https://doi.org/10.1109/ACCESS.2018.2840159.
7. Tury W, Horton R. Soil physics. Hoboken: Wiley & Sons Inc; 2004.
8. Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge: MIT press; 2015.
9. Velickovic P, Karazija L, Lane ND, Bhattacharya S, Liberis E, Lio P, Vegreville M. Cross-modal recurrent models for weight objective prediction from multimodal time-series data. ArXiv e-prints; 2017
10. Understanding LSTM Networks. http://colah.github.io
11. Schmidhuber J. Deep learning in neural networks: an overview. Neural Networks. 2015;61:85–117.
12. Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Networks. 2005;18(5–6):602–10.

# Compressed Sensing in Soil Ultra-Wideband Signals

Chenkai Zhao(✉) and Jing Liang

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China
515588290@qq.com

**Abstract.** This paper investigated the compressed sensing (CS) of ultra-wideband (UWB) soil echo signals. When CS is used in the transmission of UWB signals, sampling rate can be effectively reduced and sparse signals can be reconstructed from fewer observations. Therefore, how to apply CS into UWB soil echo signals is of great importance. The proposed approach reveals that UWB signals can be expressed by linear combinations of many atoms from a proper dictionary. In this paper, K-singular value decomposition (KSVD) dictionary and three types of Gaussian pulse dictionaries are designed, and the probability of successful reconstruction can reach 0.95. It is shown that Gaussian first-order derivative dictionary is the most suitable; the root-mean-square error (RMSE) of UWB signals and reconstructing signals is lower than 0.12.

**Keywords:** Compressed sensing · Sparse dictionary · UWB signals

## 1 Introduction

According to Nyquist–Shannon theory, sampling UWB signals with a maximum frequency of 10.6 GHz requires a sampling rate of 21.2 GHz. However, the sampling rate of microcontrollers is less than 2 GHz with the current technology. Therefore, the transmission of UWB signals has a great deal of difficulty. Precision agriculture is a new trend of agricultural development in the world [1], the key is collecting soil information, and UWB system is used to collect and transmit soil echo signals. It is urgent to find a method of signals' sampling to meet the UWB systems' requirements [2].

According to the theory of compressed sensing, UWB signals can become sparse in some transformation basis domains, which is referred to as dictionaries. Then Gaussian matrix is used to compress the sparse signals. Finally, by using orthogonal matching pursuit (OMP) algorithm, the original signals can be reconstructed with a high probability. Therefore, the sampling rate can be effectively reduced [3].

This paper focuses on designing proper dictionaries of UWB soil echo signals. KSVD dictionary and three types of Gaussian pulse dictionaries are designed, and the reconstruction errors of several methods are analyzed.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_93

## 2 UWB Soil Channel Data Acquisition

### 2.1 Equipment and Measurement Environment

The radar sensor, Time Domain P410 is used to collect data. P410 is a radar module with small size, and the pulse sampling interval is 61 ps. As shown in Fig. 1, P410 is secured in the holder and 0.8 m away from the ground. At each point of measurement, 400 sets of data are measured and the sequence length of each data is 480.

**Fig. 1.** Equipment and measurement environment.

### 2.2 Data Acquisition

A vegetable field and a sandy land were selected as experimental site; the experiment began in December 2017 and continued through March 2018. Measurements were made in ten locations with moisture values of 0.084, 0.141, 0.164, 0.186, 0.269, 0.289, 0.294, 0.354, 0.369, and 0.395. See Fig. 2 for parts of the waveform. It can be seen from the figures that the amplitude of signal has an obvious attenuation.

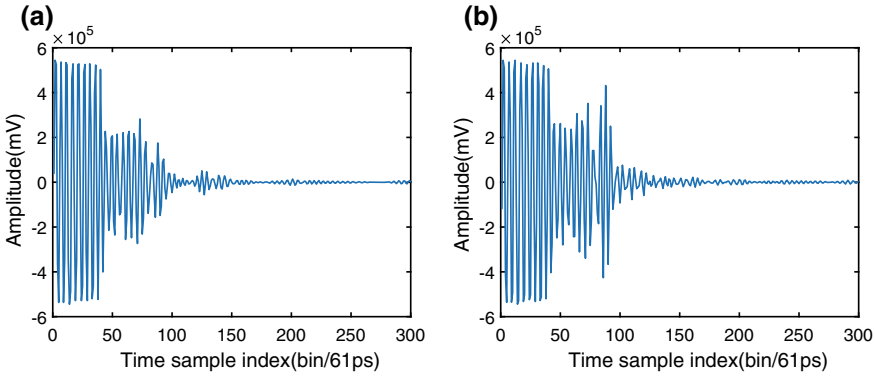



**Fig. 2.** Received signals' waveform diagram contains two types of moisture values. **a** 0.084 and **b** 0.186.

## 3 Design of Sparse Dictionary

According to the theory of compressed sensing, UWB signals can be expressed by a linear combination of atoms from the sparse dictionary, which contains all the information of the signals. In this section, four types of dictionaries are designed.

### 3.1 KSVD Learned Dictionary

KSVD dictionary is trained from some signals and contains information of the signals. The KSVD algorithm is divided into two steps: sparse representation and dictionary updating. First, set up an initialization dictionary $D$ and use orthogonal matching pursuit (OMP) [4] algorithm to make sparse representation of the given data, and then get the coefficient matrix $X$. So that the original signal can be expressed as:

$$Y{=}DX \tag{1}$$

Second, update the dictionary $D$ according to $X$. The update of the dictionary $D$ is carried out by column.

$$\|Y - DX\|_F^2 = \left\| Y - \sum_j d_j x^j \right\|_F^2 \tag{2}$$

In the process of updating dictionary, the coefficient matrix $X$ and the dictionary $D$ are fixed. What is going to be updated is the $d_k$ of the dictionary.

$$\left\| Y - \sum_j d_j x^j \right\|_F^2 = \left\| \left( Y - \sum_{j \neq k} d_j x^j \right) - d_k x^k \right\|_F^2 \tag{3}$$

Then, we can get the error matrix:

$$E_k = \left( Y - \sum_{j \neq k} d_j x^j \right) \tag{4}$$

In order to ensure the sparsity of the matrix $X$, we only preserve the nonzero values in the coefficients and then make singular value decomposition [6]. Then we can obtain the updated dictionary.

Last, use the updated dictionary to update signals again and use the signals to update the dictionary. Finally, the dictionary will be obtained.

### 3.2 Gaussian Pulse Function Dictionary

The soil echo signals are composed by Gaussian pulse and its derivatives, so we use them to construct dictionaries. The Gaussian pulse is as follows; $a$ is pulse-shaping factor.

$$p(t) = \frac{\sqrt{2}}{a} e^{-\frac{2\pi t^2}{a^2}} \tag{5}$$

The time shifts of the pulse are atoms of dictionary. The dictionary is generated by shifting with step $k$. The pulse can be seen as follows:

$$d_j(t) = p(t - jk) \tag{6}$$

Then the dictionary is obtained:

$$D = \{d_0(t), d_1(t), d_2(t), \ldots,\} \tag{7}$$

First-order derivative and second derivative of the Gaussian pulse are also used to construct dictionaries, so there are three types of Gaussian pulse dictionaries.

## 4 Performance Analysis of Dictionaries

In the previous section, four types of dictionaries are obtained. In this section, the dictionaries and OMP algorithm [4] are used to obtain the sparse representation of signals.

$$x = As \tag{8}$$

$A$ is the dictionary, $x$ is the UWB soil echo signal, and $s$ is the sparse signal. Then, the sparse signals are multiplied by the measurement matrix (C), which is Gaussian random $M * N$-dimensional matrix [7]. As $M$ is less than $N$, the dimension of $y$ is less than $s$.

$$y = Cs \tag{9}$$

Finally, another type pf OMP algorithm [5] is used to restructure the signals from $y$.

$$\hat{s} = \arg\min \|x\|_0 \; s.t. \; Ms = y \tag{10}$$

### 4.1 Performance Analysis of Dictionaries

Root-mean-square error (RMSE) between the initial signals and reconstructed signals is calculated to analyze the performance of different dictionaries.

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}\left(x_i-\hat{x}_i\right)^2}{n}} \tag{11}$$

Figure 3 is relationship between the moisture value and RMSE of different dictionaries.

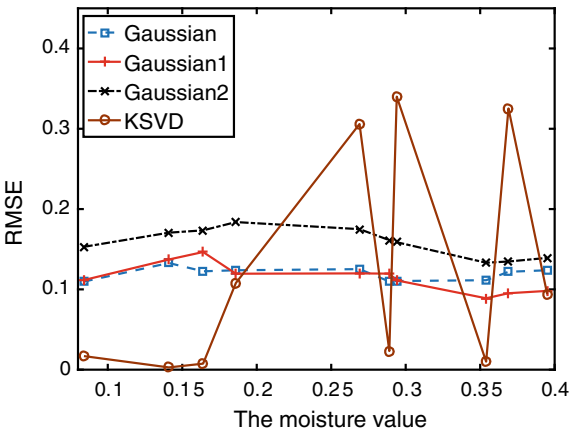



**Fig. 3.** Relationship between moisture value and RMSE.

In Fig. 3, the Gaussian1 of curve represents Gaussian first-order derivative pulse dictionary and Gaussian2 represents Gaussian second derivative pulse dictionary. It can be seen that when using three Gaussian dictionaries, RMSE showed a downward trend with the moisture value rises. On the whole, Gaussian first-order derivative pulse dictionary is more suitable than others. The curve of KSVD dictionary is unstable because the dictionary is more suitable for the signals it has learned from. In conclusion, Gaussian first-order derivative pulse dictionary is most suitable for UWB soil echo signals, and the RMSE is lower than 0.12.

### 4.2 Relationship Between Sparseness and RMSE

In this paper, dictionaries are used to obtain the sparse representation of UWB signals. Proportion of nonzero values in the signals $(k/N)$ has a great impact on the result of reconstruction. Sparseness $k$ is the nonzero number of sparse signals' sequences. The length of the sparse signals is $N$. Figure 4 is relationship between proportion of nonzero values in the signals $(k/N)$ and RMSE.

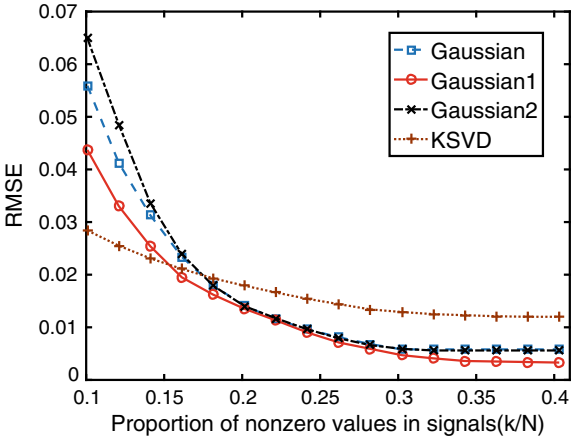



**Fig. 4.** Relationship between proportion of nonzero values and RMSE.

When the proportion is less than 0.16, KSVD dictionary is better than others. With the increase of proportion, RMSEs of all the dictionaries are lower than 0.02. Gaussian first-order pulse dictionary is more suitable than others. The result shows that the UWB signals are formed by the time shift of Gaussian pulse.

### 4.3 Probability Analysis of Successful Reconstruction

Sparse signals can be reconstructed from fewer observations, and the number of observations will affect the result. In order to analyze relationship between success reconstruction and compression ratio, the program experiments are repeated 500 times and probability of successful reconstruction is calculated. When the error satisfies the following equation, the signal is considered to be successfully restructured.

$$\|g - \hat{g}\|_2^2 \leq 0.01 \|g\|_2^2 \tag{12}$$

The relationship between probability of successful reconstruction and compression ratio is depicted in Fig. 5. When the compression ratio is higher than 0.65 for KSVD dictionary and 0.46 for Gaussian first-order derivative pulse dictionary, the probability will reach 0.95. Gaussian first-order derivative dictionary can satisfy the requirements for probability of successful reconstruction.

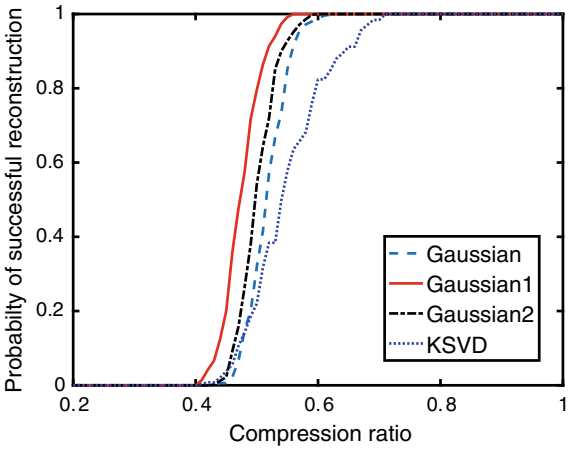



**Fig. 5.** Probability of success reconstruction for four different dictionaries.

## 5 Conclusion and Future Work

This paper investigated the compressed sensing for UWB soil echo signals. Four types of sparse dictionaries have been designed and analyzed. When using Gaussian first-order derivative pulse dictionary, root-mean-square error (RMSE) of UWB signals and reconstructing signals is lower than 0.12 and probability of successful reconstruction can reach 0.95 with compression ratio of 0.46. RMSEs of other dictionaries are higher than 0.15, and the probability of successful reconstruction is lower than 0.8 at the same compression ratio. It can be concluded that Gaussian first-order derivative pulse dictionary is the most suitable for UWB signals. In the future, we will investigate the improved dictionary algorithms and more compressed sensing algorithms which are suitable for UWB soil echo signals.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61671138, 61731006) and was partly supported by the 111 Project No. B17008.

## References

1. Liang J, Liu X, Liao K. Soil moisture retrieval using UWB echoes via fuzzy logic and machine learning. IEEE Internet Things J (Early Access); 2017 Oct 9.
2. Reed JH. An introduction to ultra wideband communication systems. Prentice Hall Communications Engineering and Emerging Technologies Series, NJ: Prentice-Hall; 2005.
3. Donoho DL. Compressed sensing. IEEE Trans Inf Theor. 2006.
4. Pati YC, Rezaiifar R, Krishnaprasad PS. Orthogonal matching pursuit-recursive function approximation with applications to wavelet decomposition. In: Proceedings

of annual asilomar conference signals, systems, and computers; 1993 Nov; Pacific Grove.
5. Tropp JA, Gilbert AC. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theor. 2007;53(12):4655–66.
6. Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm fordesigning overcomplete dictionaries for sparse representation. IEEE Trans Signal Process. 2006;54(11):4311.
7. Paredes JL, Arce GR, Wang Z. Ultra-wideband compressed sensing: channel estimation. IEEE J Sel Top Signal Process. 2007;1(3):383–95.

# Research Progress of Inverse Synthetic Aperture Radar (ISAR) Imaging of Moving Target via Quadratic Frequency Modulation (QFM) Signal Model

Yong Wang[1(✉)], Aijun Liu[2], and Qingxiang Zhang[1]

[1] Research Institute of Electronic Engineering Technology, Harbin Institute of Technology, Harbin 150001, China
wangyong6012@hit.edu.cn
[2] Department of Communication Engineering, Weihai Campus of Harbin Institute of Technology, Shandong 264209, China

**Abstract.** ISAR imaging of dynamic target is very significant in real applications. Plenty of available outcomes have been acquired by the scholars in the near years. Considering the accuracy and the computational complexity, the radar echo can be described as multicomponent QFM signal after envelope alignment with initial phase calibration. Many parametric algorithms in this case have been developed recently. This paper provides a comprehensive summarization of the ISAR imaging approach with QFM signal model in recent years, with the aim to introduce the research progress of it to the researchers and interested readers.

**Keywords:** ISAR · Complex motion · QFM signal

## 1 Introduction

ISAR imaging of dynamic target has great meaning in the military or civilian utilizations [1–4]. The range direction resolution is decided by the radar signal bandwidth, and the azimuth direction resolution is determined by the accumulated imaging interval and the rotational velocity of the target. Before radar imaging course, the translational compensation, which involves the envelope alignment and initial phase calibration, should be operated. Some mature methods have been proposed for the motion compensation [5–10]. Then, the 'turntable' target can be obtained for the ISAR imaging, and the classical Fourier transformation is adopted to finish the azimuth processing to yield a high-quality ISAR image. This is the well-known Range-Doppler (RD) approach, and this mature method is widely used in the past decades [11, 12]. The RD algorithm is effective under the condition of uniform rotation after movement compensation for the dynamic target. Thus, the echo signal can be depicted as multicomponent sinusoidal signal. The condition will be matched for the big size and smooth movement target. Considering the target with a complicated motion, such as the ship with three-dimensional (3D) movement (yaw, roll and pitch), the Doppler frequency has obvious time-varying property. Hence, the Fourier transformation is


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_94

unsuitable to analyze this kind of nonstationary signals. Then, the radar echo should be considered as multicomponent polynomial phase signal (PPS), and the Range-Instantaneous-Doppler (RID) technique should be adopted in order to yield a good quality ISAR image at arbitrary time instant. The nature of RID technique is the instantaneous frequency estimate of multicomponent PPS model. Many effective ISAR imaging algorithms with the RID technique have been proposed in the references [13–51].

We can divide the aforementioned algorithms into two types. One is the non-parametric method, and the high-concentrated time-frequency representation (TFR) with free cross-terms is used in the azimuth processing. The high-quality radar image could be generated consequently. The prevalent TFR includes the Cohen class distributions and some high order distributions [13–20]. The tradeoff of the TFR methods is the high-concentration and the cross-term elimination, and it is still a problem which has not been solved completely. The other is the parametric method; in this case, the radar echo is depicted as the PPS model with a certain order. When the target has a common maneuverability, such as the target has an accelerative rotation with the radar, the radar echo can be depicted as multicomponent linear frequency modulation (LFM) signal. With the parameter estimation for it, we can yield a clear radar image associated with the RID approach. Plenty of radar imaging methods via the LFM signal model have been presented in the past decades, and these algorithms are effective and efficient under some certain conditions [21–23]. In general, the target would have a severe maneuverability; thus, it is not practical to yield a high-quality radar image with LFM signal because of the low accuracy. In this situation, the QFM signal is appropriate, where the third order coefficient for the radar echo is introduced, which has great influence on the azimuth focusing quality. With the QFM signal model in each range cell, the image quality will be enhanced obviously. The ISAR imaging technique via the QFM signal has been a main issue recently, such as the technique in [24–51]. In this paper, the authors would like to provide a comprehensive summarization of the ISAR imaging methods with QFM signal model in recent years, with the aim to introduce the research progress of it to the researchers and interested readers.

The paper is constructed as follows. The QFM signal is established in Sect. 2; Sect. 3 is the principle of ISAR imaging method via the QFM signal model; Sect. 4 brings reference overview of some ISAR imaging methods via the QFM signal model, and the conclusion for the paper is provided in the last section.

## 2 QFM Signal Construction for the Radar Imaging

### 2.1 Construction of Radar Echo Model

The translational compensation is the chief step before the radar imaging process. This involves the envelope alignment and initial phase calibration. The envelope alignment could be implemented by some existing methods, such as the maximum correlation method, the minimum entropy method and the global method. The phase calibration could be operated by the phase gradient approach and the phase error reduction method [5–10]. Then, the imaging geometry is defined in Fig. 1.

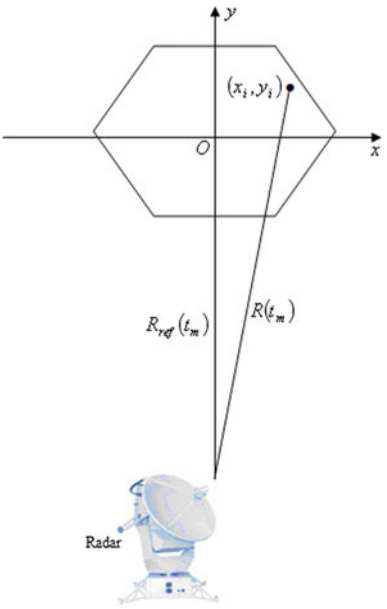



**Fig. 1.** Radar imaging geometry

First, the compensation for the radar echo should be implemented, and then we can assume the distance between the radar and imaging target geometric center $O$ is $R_{\text{ref}}(t_m)$. The arbitrary point scatterer on the target with coordinates $(x_i, y_i)$ is considered, and the range between it and radar at a slow time $t_m$ is $R(t_m)$.

It is supposed that the radar transmits LFM signal as follows:

$$s(\tilde{t}) = \text{rect}\left(\frac{\tilde{t}}{T_p}\right) \exp\left[j2\pi\left(f_c\tilde{t} + \frac{1}{2}k\tilde{t}^2\right)\right] \tag{1}$$

where $\text{rect}(\tilde{t}) = \begin{cases} 1, & |\tilde{t}| \le 1/2 \\ 0, & |\tilde{t}| > 1/2 \end{cases}$, $T_p$ denotes the pulse width, $\tilde{t}$ denotes the fast time, $f_c$ denotes the central frequency, and $k$ denotes the chirp rate. The radar echo for the point $(x_i, y_i)$ could be obtained as

$$s(\tilde{t}, t_m) = \sigma(t_m)\text{rect}\left(\frac{\tilde{t} - 2R(t_m)/c}{T_p}\right) \exp\left\{j2\pi\left[f_c\left(\tilde{t} - \frac{2R(t_m)}{c}\right) + \frac{1}{2}k\left(\tilde{t} - \frac{2R(t_m)}{c}\right)^2\right]\right\} \tag{2}$$

where $c$ denotes the light velocity, and $\sigma(t_m)$ denotes the reflection amplitude.

In order to reduce the sampling rate for the radar echo, the reference signal could be constructed in the radar receiver with the same form as (2), and it is listed as follows:

$$s_{\text{ref}}(\tilde{t}, t_m) = \text{rect}\left(\frac{\tilde{t} - 2R_{\text{ref}}(t_m)/c}{T_p}\right) \exp\left\{j2\pi\left[f_c\left(\tilde{t} - \frac{2R_{\text{ref}}(t_m)}{c}\right) + \frac{1}{2}k\left(\tilde{t} - \frac{2R_{\text{ref}}(t_m)}{c}\right)^2\right]\right\} \tag{3}$$

Then, by the stretch technique with the reference signal, the radar echo can be dechirped

$$\begin{aligned} s_v(\tilde{t}, t_m) = \sigma(t_m)\text{rect}\left(\frac{\tilde{t} - 2R(t_m)/c}{T_p}\right) \exp\left\{j\frac{4\pi}{\lambda}[R_{\text{ref}}(t_m) - R(t_m)]\right\} \\ \cdot \exp\left\{j\frac{4\pi k}{c}\left[\tilde{t} - \frac{2R_{\text{ref}}(t_m)}{c}\right][R_{\text{ref}}(t_m) - R(t_m)]\right\} \end{aligned} \tag{4}$$

where we have neglected the residue video phase term $4\pi k[R(t_m) - R_{\text{ref}}(t_m)]^2 \big/ c^2$. Take the Fourier transformation of (4) with $\tilde{t}$, we will achieve the high-resolution range profile as follows [24]:

$$s_v(r, t_m) = \sigma(t_m)kT_p\text{sinc}\left[\frac{2kT_p}{c}(r - \Delta R(t_m))\right] \cdot \exp\left[j\frac{4\pi}{\lambda}\Delta R(t_m)\right] \tag{5}$$

where $\Delta R(t_m) = R_{\text{ref}}(t_m) - R(t_m)$, $\sin c[\cdot]$ has the form of $\text{sinc}(x) = \sin(x)/x$, and $r$ denotes the range bin.

For the target with complicated movement, $\Delta R(t_m)$ could be approximated as follows [24]:

$$\Delta R(t_m) = a_0 + a_1 t_m + a_2 t_m^2 + a_3 t_m^3 \tag{6}$$

where $a_0$, $a_1$, $a_2$ and $a_3$ denote the undetermined coefficients.

Thus, the radar echo could be shown as:

$$s(t_m) = \sum_{q=1}^{Q} \sigma_q \exp\left[j\frac{4\pi}{\lambda}\left(a_{0,q} + a_{1,q}t_m + a_{2,q}t_m^2 + a_{3,q}t_m^3\right)\right] \tag{7}$$

where $Q$ denotes the scatters number, and $\sigma_q$ denotes the amplitude of the $q$th point.

Hence, the multicomponent QFM signal model could be constructed, where the third order coefficient is introduced, which has great influence on the azimuth focusing quality. With the aim to yield the ISAR images with good quality in this situation, the parameter of multicomponent QFM signal should be evaluated with high precision.

### 2.2 Some Experimental Results

With the aim to illustrate the validity of QFM signal, some experimental results are listed here for the better illustrations.

A set of raw data for the airplane with high maneuverability is adopted here, and the bandwidth is $B = 400$ MHz, which means the range solution is 0.375 m. Now, we

calculate the Wigner–Ville distribution (WVD) for the echo in 140th and 150th range cells, respectively. The corresponding results are shown in Fig. 3. It is easy to conclude the nonlinear time-varying character for the radar echo. This means that the QFM signal is appropriate in this situation (Fig. 2).

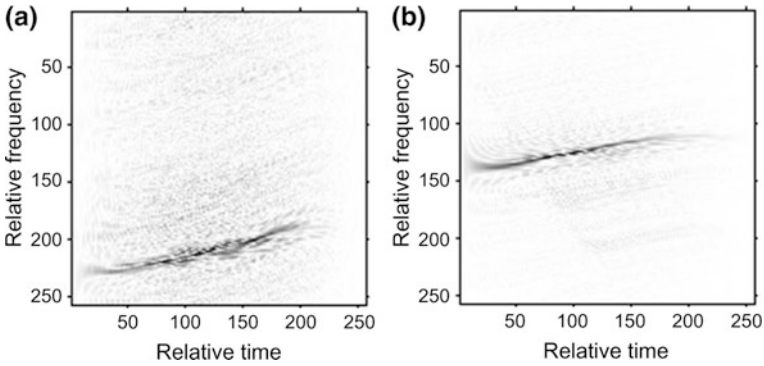



**Fig. 2.** WVD for the echo: **a** 140th range cell, **b** 150th range cell

The ISAR imaging results by the RD method is given in Fig. 3. This image is defocused in the azimuth direction for the complicated movement. Hence, the RD method is inappropriate to yield a good quality image, and some suitable methods have been proposed to solve this problem.

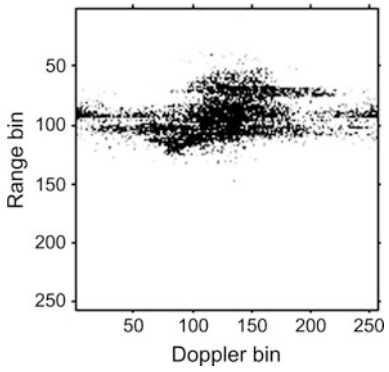



**Fig. 3.** ISAR image by the RD method

## 3 Approach of ISAR Imaging via QFM Signal Model

Here, the principle of ISAR imaging approach via QFM signal model is validated in Fig. 4. After motion compensation of the radar echo, we can model it as multicomponent QFM signal, as listed in (7). Thus, we can yield the instantaneous ISAR image with high quality via the RID approach.

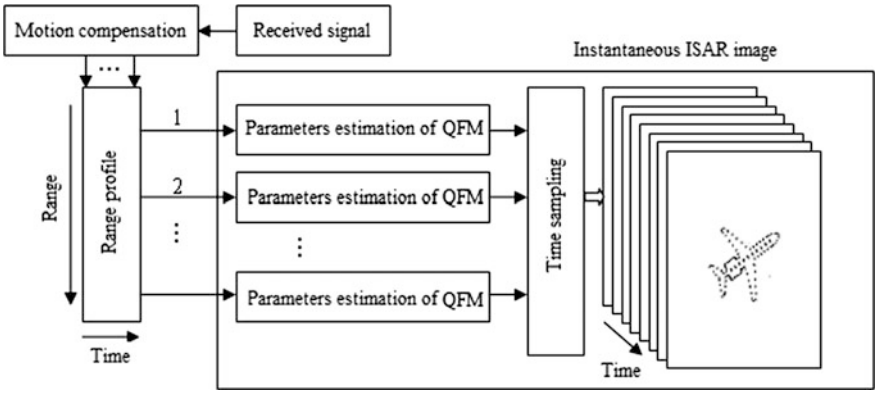



**Fig. 4.** Principle of ISAR imaging method via QFM signal model

## 4 Overview of the ISAR Imaging Method via the QFM Signal Model

In this part, we will overview and analyze some ISAR imaging methods via the QFM signal model, which were proposed in the references recently. These methods are founded on the parameter estimation of multicomponent QFM signal model, and the accuracy and computational load are different. Here, we summarize these algorithms as follows:

### 4.1 Parameter Estimation by the Matched-Phase Transform

This kind of algorithms is based on the matched-phase principle, and the phase coefficients are estimated by the one-dimensional (1D) or two-dimensional (2D) searching procedure. The typical algorithms involve the maximum-likelihood approach [25], the product generalized cubic phase function (PGCPF) method [26], the product high-order matched-phase transform (PHMT) [27], the match Fourier transform (MFT) [28], the local polynomial Wigner distribution (LPWD) and integrated high-order ambiguity function (IHAF) [24, 29], the modified cubic phase function (MCPF) [30], the integrated generalized cubic phase function (IGCPF) [31], the scaled Fourier transform (SCFT)-based method [32], the integrated cubic phase function (ICPF) method [33], the TC-DechirpClean method [34] and the discrete chirp-Fourier transform [35], etc. These algorithms are efficient to evaluate the parameter of QFM signal

with reduced computational load, and via RID method, the ISAR images with good quality can be generated. The drawback of these kinds of methods is that the phase coefficients are estimated gradually, and the error propagation phenomena cannot be avoided. Furthermore, the searching procedure should be implemented by the predetermined range of value for the estimated parameters, and the searching step is another factor to influence the computational load and estimate accuracy.

For the matched-phase transform, the authors also consider the case of QFM signal with time-varying amplitude. The phenomena could be occurred when the target has dihedral or trihedral components or the migration through resolution cell (MTRC) induced by the complicated movement. Some efficient methods for the parameter estimation of time-varying amplitude QFM signal have been proposed in [36, 37], and corresponding ISAR image quality can be increased consequently.

### 4.2 Fast Parameter Estimation Algorithms

In [38–45], Jibin Zheng et al. proposed a series of fast parameter estimation algorithms for multicomponent QFM signal. These algorithms are based on the fast Fourier transform (FFT) and nonuniform fast Fourier transform (NUFFT) technique, and the computational cost can be reduced greatly with high parameter estimation accuracy. Furthermore, such as the chirp rate-quadratic chirp rate distribution method in [39], the searching procedure can be avoided and the anti-noise performance is high. Hence, this kind of fast parameter estimation algorithms is very efficient in the ISAR imaging of target with complicated movement.

### 4.3 Parameter Estimation Algorithms in Low SNR Environment

In [46, 47], Dong Li et al. proposed a set of parameter estimation algorithms for the multicomponent QFM signal in low signal-to-noise-ratio (SNR) situation. These algorithms are based on the coherently integrated modified cubic phase function (CIMCPF) and the coherently integrated modified high-order ambiguous function (CIMHAF), or the coherently integrated generalized cubic phase function (CIGCPF) and the coherently integrated cubic phase function (CICPF). By the coherent integration and NUFFT, these algorithms have good property in the low SNR condition with reduced computational cost, and the corresponding ISAR imaging method in the low SNR conditions was proposed by the authors consequently.

### 4.4 Parameter Estimation Algorithms with the Signal Decomposition Technique

The signal decomposition technique is an efficient way to analyze the multicomponent amplitude modulated and frequency modulated (AM-FM) signal. The most commonly used method for the signal decomposition is the adaptive Chirplet decomposition, where the Chirplet atom has the format of Gaussian envelope chirp signal. Some ISAR imaging methods via the Chirplet decomposition were proposed in [48, 49]. For the multicomponent QFM signal model, the cubic Chirplet decomposition has been proposed to match the complicated signals with higher precision, and the corresponding

ISAR imaging methods were presented in [50, 51]. The signal decomposition method is appropriate for the analysis of nonstationary signal, but the implementation process is complicated for the multi-parameters optimizations.

## 5 Numerical Results for the Raw Data

In this part, the IGCPF approach presented by the authors in [31] is adopted as the example to enhance the quality of radar image for the real data in Sect. 2.2. The corresponding ISAR imaging results at two-time positions are listed in Fig. 5a, b.

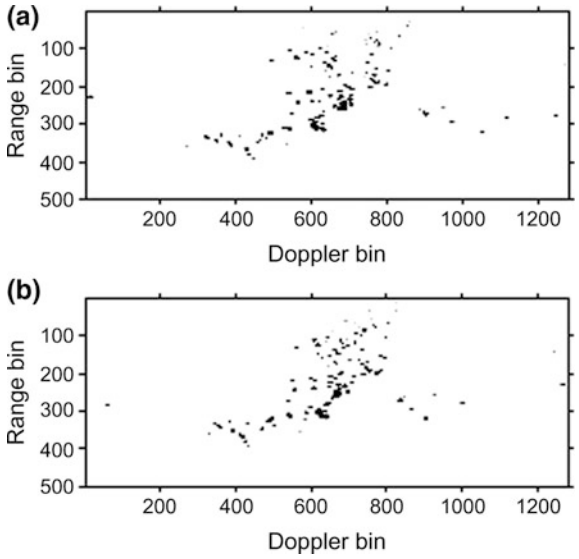



**Fig. 5.** Instantaneous images via IGCPF method in [31]: **a** ISAR image at $t = 0.54$ s, **b** ISAR image at $t = 0.94$ s

The image quality has been enhanced by the comparison with Fig. 3, and the shape of target can be well identified by the radar images. This is very important for the target classification and recognition. Furthermore, the imaging results in Fig. 5 also validate the superiority of the QFM signal model for the echo.

## 6 Conclusion

The contribution of this paper is the overview of the ISAR imaging algorithms with QFM signal model in recent years, and some efficient algorithms are introduced and analyzed. The ISAR imaging method via the QFM signal model is very important, and it has received much attention for the scholars in the world. For the target with

complicated motion, the traditional RD method or the RID algorithm based on the LFM signal is not precise. Thus, the QFM signal is more accurate, and the imaging method can be carried out with high efficiency.

**Acknowledgments.** This work is supported by the National Natural Science Foundation of China under grant 61622107 and 61471149.

## References

1. Berizzi F, Mese ED, Diani M, Martorella M. High-resolution ISAR imaging of maneuvering targets by means of the range instantaneous Doppler technique: modeling and performance analysis. IEEE Trans Image Process. 2001;10(12):1880–90.
2. Pastina D, Bucciarelli M, Lombardo P. Multistatic and MIMO distributed ISAR for enhanced cross-range resolution of rotating targets. IEEE Trans Geosci Remote Sens. 2010;48(8):3300–17.
3. Zhao LF, Wang L, Bi GA, Yang L. An autofocus technique for high-resolution inverse synthetic aperture radar imagery. IEEE Trans GRS. 2014;52(10):6392–403.
4. Xu G, Xing MD, Bao Z. High-resolution inverse synthetic aperture radar imaging of maneuvering targets with sparse aperture. Electron Lett. 2015;51(3):287–9.
5. Liu L, Zhou F, Tao ML, Sun PG, Zhang ZJ. Adaptive translational motion compensation method for ISAR imaging under low SNR based on particle swarm optimization. IEEE J Sel Top Appl Earth Observations Remote Sens. 2015;8(11):5146–57.
6. Wang JF, Liu XZ. Improved global range alignment for ISAR. IEEE Trans AES. 2007;43 (3):1070–5.
7. Pastina D. Rotation motion estimation for high resolution ISAR and hybrid SAR/ISAR target imaging. In: IEEE radar conference; 2008. p. 1–6.
8. Hu JM, Zhou W, Fu YW, Li X, Jing N. Uniform rotational motion compensation for ISAR based on phase cancellation. IEEE Geosci Remote Sens Lett. 2011;8(4):636–40.
9. Yeh CM, Xu J, Peng YN, Xia XG, Wang XT. Rotational motion estimation for ISAR via triangle pose difference of two range-Doppler images. IET Radar Sonar Navig. 2010;4 (4):528–36.
10. Zhu DY, Wang L, Yu YS, Tao QN, Zhu ZD. Robust ISAR range alignment via minimizing the entropy of the average range profile. IEEE Geosci Remote Sens Lett. 2009;6(2):204–8.
11. Walker JL. Range-Doppler imaging of rotating objects. IEEE Trans AES. 1980;16(1):23–52.
12. Rao W, Li G, Wang XQ, Xia XG. Adaptive sparse recovery by parametric weighted L1 minimization for ISAR imaging of uniformly rotating targets. IEEE J Sel Top Appl Earth Observations Remote Sens. 2013;6(2):942–52.
13. Cohen L. Time-frequency distribution-A review. Proceedings of IEEE. 1989;77(7):941–81.
14. Xing MD, Wu RB, Li YC, Bao Z. New ISAR imaging algorithm based on modified Wigner-Ville distribution. IET Proc-Radar Sonar Navig. 2009;3(1):70–80.
15. Xia XG, Wang GY, Chen VC. Quantitative SNR analysis for ISAR imaging using joint time-frequency analysis-Short time fourier transform. IEEE Trans AES. 2002;38(2):649–59.
16. Bao Z, Sun CY, Xing MD. Time-frequency approaches to ISAR imaging of maneuvering targets and their limitations. IEEE Trans AES. 2001;37(3):1091–9.
17. Chen VC, Miceli WJ. Time-varying spectral analysis for radar imaging of maneuvering targets. IEE Proc-Radar Sonar Navig. 1998;145(5):262–8.
18. Trintinalia LC, Ling H. Joint time-frequency ISAR using adaptive processing. IEEE Trans Antennas Propag. 1997;45(2):221–7.

19. Chen VC, Ling H. Time-frequency transforms for radar imaging and signal analysis. Artech House; 2002.
20. Stankovic L, Dakovic M, Thayaparan T. Time-frequency signal analysis with applications. Artech House, Boston; 2013, 3.
21. Wang GY, Bao Z, Sun XB. Inverse synthetic aperture radar imaging of non-uniformly rotating targets. Opt Eng. 1996;35(10):3007–11.
22. Bao Z, Wang GY, Luo L. Inverse synthetic aperture radar imaging of maneuvering targets. Opt Eng. 1998;37(5):1582–8.
23. Wang Y, Jiang Y. ISAR imaging of three-dimensional rotation target based on two-order match Fourier transform. IET Signal Process. 2012;6(2):159–69.
24. Wang Y, Kang J, Jiang YC. ISAR imaging of maneuvering target based on the local polynomial Wigner distribution and integrated high order ambiguity function for cubic phase signal model. IEEE J Sel Top Appl Earth Observations Remote Sens. 2014;7(7):2971–91.
25. Li YC, Xing MD, Su JH, Quan YH, Bao Z. A new algorithm of ISAR imaging for maneuvering targets with low SNR. IEEE Trans AES. 2013;49(1):543–57.
26. Wang Y, Jiang YC. Inverse synthetic aperture radar imaging of maneuvering target based on the product generalized cubic phase function. IEEE Geosci Remote Sens Lett. 2011;8 (5):958–62.
27. Wang Y, Jiang YC. ISAR imaging of a ship target using product high order matched-phase transform. IEEE Geosci Remote Sens Lett. 2009;6(4):658–61.
28. Wang C, Wang Y, Li SB. ISAR imaging of ship targets with complex motion based on match Fourier transform for cubic chirps model. IET Radar Sonar Navig. 2013;7(9):994–1003.
29. Wang Y, Zhao B, Kang J. Asymptotic statistical performance of local polynomial Wigner distribution for the parameters estimation of cubic phase signal with application in ISAR imaging of ship target. IEEE J Sel Top Appl Earth Observations Remote Sens. 2015;8 (3):1087–98.
30. Wang Y, Zhao B. Inverse synthetic aperture radar imaging of non-uniformly rotating target based on the parameters estimation of multi-component quadratic frequency-modulation signals. IEEE Sens J. 2015;15(7):4053–61.
31. Wang Y, Lin YC. ISAR imaging of non-uniformly rotating target via range-instantaneous-Doppler-derivatives algorithm. IEEE J Sel Top Appl Earth Observations Remote Sens. 2014;7(1):167–76.
32. Bai X, Tao R, Wang ZJ, Wang Y. ISAR imaging of a ship target based on parameter estimation of multicomponent quadratic frequency-modulation signals. IEEE Trans GRS. 2014;52(2):1418–29.
33. Wang Y. Inverse synthetic aperture radar imaging of manoeuvring target based on range-instantaneous-Doppler and range-instantaneous-chirp-rate algorithms. IET Radar Sonar Navig. 2012;6(9):921–8.
34. Li YC, Wu RB, Xing MD, Bao Z. Inverse synthetic aperture radar imaging of ship target with complex motion. IET Proc-Radar Sonar Navig. 2008;2(6):395–403.
35. Wu L, Wei XZ, Yang DG, Wang HQ, Li X. ISAR imaging of targets with complex motion based on discrete chirp Fourier transform for cubic chirps. IEEE Trans GRS. 2012;50 (10):4201–12.
36. Wang Y, Zhang QX, Zhao B. Inverse synthetic aperture radar imaging of maneuvering target based on cubic chirps model with time varying amplitudes. J Appl Remote Sens. 2016;10 (1):1–10.
37. Wang Y, Xu RQ, Zhang QX, Zhao B. ISAR imaging of maneuvering target based on the quadratic frequency modulation signal model with time varying amplitude. IEEE J Sel Top Appl Earth Observations Remote Sens. 2017;10(3):1012–24.

38. Zheng JB, Su T, Zhu WT, Liu QH. ISAR imaging of targets with complex motions based on the Keystone time-chirp rate distribution. IEEE Geosci Remote Sens Lett. 2014;11(7):1275–9.
39. Zheng JB, Su T, Zhang L, Zhu WT, Liu QH. ISAR imaging of targets with complex motion based on the chirp rate-quadratic chirp rate distribution. IEEE Trans GRS. 2014;52 (11):7276–89.
40. Zheng JB, Su T, Zhu WT, Zhang L, Liu Z, Liu QH. ISAR imaging of nonuniformly rotating target based on a fast parameter estimation algorithm of cubic phase signal. IEEE Trans GRS. 2015;53(9):4727–40.
41. Zheng JB, Liu HW, Liao GS, Su T, Liu Z, Liu QH. ISAR imaging of targets with complex motions based on a noise-resistant parameter estimation algorithm without nonuniform axis. IEEE Sens J. 2016;16(8):2509–18.
42. Zheng JB, Liu HW, Liao GS, Su T, Liu Z, Liu QH. ISAR imaging of nonuniformly rotating targets based on generalized decoupling technique. IEEE J Sel Top Appl Earth Observations Remote Sens. 2016;9(1):520–32.
43. Zheng JB, Su T, Liao GS, Liu HW, Liu Z, Liu QH. ISAR imaging of fluctuating ships based on a fast bilinear parameter estimation algorithm. IEEE J Sel Top Appl Earth Observations Remote Sens. 2015;8(8):3954–66.
44. Li YY, Su T, Zheng JB, He XH. ISAR imaging of targets with complex motions based on modified Lv's distribution for cubic phase signal. IEEE J Sel Top Appl Earth Observations Remote Sens. 2015;8(10):4775–84.
45. Lv Q, Su T, Zheng JB. Inverse synthetic aperture radar imaging of targets with complex motion based on the local polynomial ambiguity function. J Appl Remote Sens. 2016;10 (1):1–18.
46. Li D, Zhan MY, Zhang XZ, Fang ZP, Liu HQ. ISAR imaging of nonuniformly rotating target based on the multicomponent CPS model under low SNR environment. IEEE Trans AES. 2017;53(3):1119–35.
47. Li D, Gui XG, Liu HQ, Su J, Xiong H. An ISAR imaging algorithm for maneuvering targets with low SNR based on parameter estimation of multicomponent quadratic FM signals and nonuniform FFT. IEEE J Sel Top Appl Earth Observations Remote Sens. 2016;9(12):5688–702.
48. Wang Y, Jiang YC. Approach for high resolution inverse synthetic aperture radar imaging of ship target with complex motion. IET Signal Proc. 2013;7(2):146–57.
49. Wang Y, Jiang YC. ISAR imaging for three-dimensional rotation targets based on adaptive Chirplet decomposition. Multidimension Syst Signal Process. 2010;21(1):59–71.
50. Wang Y, Zhao B, Jiang YC. Inverse synthetic aperture radar imaging of targets with complex motion based on cubic Chirplet decomposition. IET Signal Proc. 2015;9(5):419–29.
51. Wang Y. Radar imaging of non-uniformly rotating target via a novel approach for multi-component AM-FM signal parameter estimation. Sensors. 2015;15(3):6905–23.

# Bistatic SAR Imaging Based on Compressive Sensing Approach

Yong Wang[1(✉)], Hongyuan Zhang[2], and Jing Zhou[1]

[1] Research Institute of Electronic Engineering Technology, Harbin Institute of Technology, Harbin 150001, China
wangyong6012@hit.edu.cn
[2] Beijing Institute of Astronautical System Engineering, Beijing, China

**Abstract.** The compressive sensing (CS) technique has been introduced to the field of synthetic aperture radar (SAR) imaging procedure to reduce the amount of measurements. In this letter, a novel algorithm for bistatic SAR imaging based on the CS technique is proposed. The range profile is reconstructed by the Fourier transform, and the azimuth processing is implemented by the CS method consequently. The proposed algorithm can realize the high-quality imaging with limited measurements efficiently for the missing bistatic SAR radar echoes. Results of simulated data demonstrate the validity of the novel approach.

**Keywords:** CS technique · Bistatic SAR · Limited measurements

## 1 Introduction

The main idea for bistatic SAR image formation is designing a two-dimensional matched filter. The time domain method is a direct matched filtering of the baseband signal which uses the echo signal at each position [4]. The point target spectrum is the key step for most efficient processing methods applying in the two dimensional. However, the derivation of the point target for bistatic SAR imaging exist some difficulties. The echo of the bistatic SAR can be expressed as a double root formula, which makes the computation of the point target spectrum much more complicated and intensive. To simplify the process of bistatic SAR imaging, the bistatic SAR imaging procedure can be equal to the baseline midpoint SAR imaging method and the traditional SAR imaging technique can be used in this case.

Bistatic SAR imaging based on the concept of compressed sensing (CS) technique has attracted much attention for the scholars all around the world. This method can obtain high-resolution imaging from fewer measurements than the Nyquist sampling theorem requires [1–3]. In general, the two-dimensional (2D) undersampling can further reduce the amount of measurement. Hence, a novel algorithm for bistatic SAR imaging based on the CS technique is proposed in this paper. The range profile is reconstructed by the Fourier transform, and the azimuth processing is implemented by the CS method consequently. The proposed algorithm can realize the high-quality imaging with limited measurements efficiently for the missing radar echoes. Bistatic SAR imaging results for simulated data demonstrate the effectiveness of the algorithm in this paper.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_95

## 2 The Bistatic SAR Method Based on Baseline Midpoint SAR Imaging

The general model of the bistatic SAR is shown in Fig. 1 as follows.

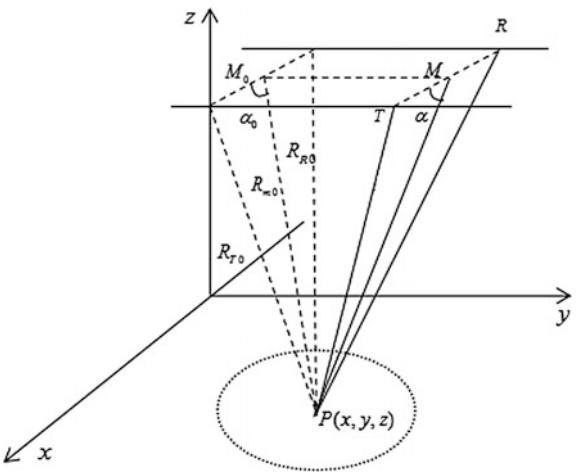



**Fig. 1.** General mode of bistatic SAR

Consider the length of the baseline between the transmitter and the receiver is $L$. The two antennas move at the same speed $V$ along the $y$ direction. At time $t = 0$, it is assumed that the transmitter platform $T$ locates at $(X_{T0}, Y_{T0}, Z_{T0})$, and the receiver platform $R$ locates at $(X_{R0}, Y_{R0}, Z_{R0})$. The distance between the transmitter and the arbitrary point scatterer $P(x, y, z)$ is $R_{T0}$, and the distance between the receiver and the arbitrary point scatterer $P(x, y, z)$ is $R_{R0}$. The midpoint for the baseline at $t = 0$ is $M_0$, and the distance between $M_0$ and $P(x, y, z)$ is $R_{M0}$. The angle between the baseline at $t = 0$ and $M_0P$ is $\alpha_0$. At time $t$, the transmitter platform $T$ locates at $(X_T, Y_T, Z_T)$, while the receiver platform $R$ locates at $(X_R, Y_R, Z_R)$. It is assumed that $M$ is the midpoint for the baseline $TR$. Then, the instantaneous range from the transmitter to the receiver for the point target can be shown as:

$$\begin{aligned} R(\eta) = |TP| + |RP| = &\sqrt{(x - X_T)^2 + (Y_T + Vt - y)^2 + (z - Z_T)^2} \\ &+ \sqrt{(x - X_R)^2 + (Y_R + Vt - y)^2 + (z - Z_R)^2} \end{aligned} \tag{1}$$

From (1), according to the cosine theorem, we obtain

$$\begin{cases} |TP| = \sqrt{(L/2)^2 + |MP|^2 - L|MP|\cos\alpha} \\ |RP| = \sqrt{(L/2)^2 + |MP|^2 + L|MP|\cos\alpha} \end{cases} \tag{2}$$

We assume that

$$|TP| + |RP| = 2a \tag{3}$$

Then, from (2) and (3), we can see that

$$|MP| = a\sqrt{\frac{4a^2 - L^2}{4a^2 - L^2\cos^2\alpha}} \approx a - \frac{\sin^2\alpha}{8a}L^2 \tag{4}$$

If $L << a$, the formula above can be simplified as follows.

$$|TP| + |RP| = 2a \approx 2|MP| + \frac{\sin^2\alpha}{4a}L^2 \tag{5}$$

We can see that the instantaneous range is equivalent to the middle point of the baseline distance with the distance fluctuation. From the derivation results, we can see that this term is not fixed. With the movement of the antenna, the location between the antenna and the target is constantly changing. In the synthetic aperture time, $a$ and $\alpha$ are also changing. When the synthetic aperture is much smaller, and the mapping belt is narrow, it can be considered $\alpha$ as a constant. At this point, $\alpha$ can be approximated as $\alpha_0$. In addition, $a$ can also be approximated by the average distance of the center scene $a_0$. So we can simplify (1) as follows.

$$R(\eta) \approx 2\sqrt{(x - X_m)^2 + (Y_m + Vt - y)^2 + (z - Z_m)^2} + \frac{\sin^2\alpha_0}{4a_0}L^2 \tag{6}$$

where $X_m = (X_T + X_R)/2; Y_m = (Y_T + Y_R)/2; Z_m = (Z_T + Z_R)/2$.

According to the above equivalent, as long as the compensation for the system delay, the bistatic SAR method based on baseline midpoint SAR imaging can be used for the processing.

- Simulation1

It is assumed that the carrier frequency for the transmitted signal is 9 GHz, the bandwidth is 50 MHz. The height of transmit platform and receive platform is 4000 m. The two platforms move at the same speed of 100 m/s. The baseline distance between the two platforms is 500 m. The bistatic SAR imaging results are shown in Fig. 2a, b, respectively. From the simulation results, we can see that the proposed method is suitable for the bistatic SAR imaging.

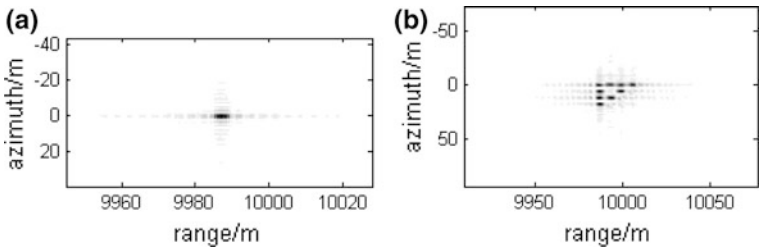



**Fig. 2.** Imaging results of the bistatic SAR: **a** the simulation results of single point target and **b** the simulation results of multiple point target

## 3 Sparse Signals Reconstruction by Compressive Sensing

The compressive sensing is a new theory of signal processing. In recent years, it has been widely concerned. The research of compressive sensing theory is expected to solve the problem of data storage and transmission in imaging radar system. A signal is sparse if the nonzero number of coefficients is much lower than the number of zero samples [5]. Sparse signals can be obtained from a few set of measurements. Compressive sensing is applied to deal with sparse signal reconstruction. In general signal processing problems, signal samples can be considered as observations of linear combinations with transformation coefficients in a transformation domain. Compressive sensing can reduce the measurements of signals while reserving the useful information of the signal.

There exist various algorithms for reconstruction of sparse signals from a reduced set of measurements. The orthogonal matching pursuit (OMP) method is one of the commonly used algorithms.

The discrete time signal $x(n)$ with $N$ samples can be represented by a $N \times 1$ column vector with notation $x = [x(0), x(1), \ldots, x(N-1)]^{\mathrm{T}}$. Its transform coefficients can be defined as $X(k) = T[x(n)]$. If the number of nonzero coefficients is far less than samples of the signal then the signal $x(n)$ is sparse in this transformation domain. If the signal can be shown as follows:

$$x = \sum_{i=1}^{N} \psi_i s_i = \Psi s \tag{7}$$

Then, the signal is sparse or compressible on the base $\Psi$, the base $\Psi$ is defined as the sparse base of the signal. In general, the signal itself is non-sparse, but some transform domain is likely to show the sparsity, and the orthogonal dictionary transform domain is a special transform domain which is used frequently.

In general, the sparsity measure minimization is the key for the reconstruction algorithm. Nonzero transform coefficients computation is the simplest sparsity measure [6]. Then, the calculation can be implemented from the mathematical form named as the $l_0$ norm [4]. The signal reconstruction problem is considered as

$$\min \|X\|_0 \text{ subject to } y = AX \tag{8}$$

In (8), the matrix $A$ is the measurement matrix. In general signal transforms, it corresponds to the inverse transformation matrix where the rows corresponding to the unavailable samples are omitted [2]. Although there are some methods to reconstruct the signal using (8), it is an NP-hard combination optimization problem. That is why using the closest $l_1$ norm of the signal to transform instead the $l_0$ norm in the optimization process, just as follows:

$$\min \|X\|_1 \text{ subject to } y = AX \tag{9}$$

- Simulation2

The compressive sensing algorithm is applied to the simulation data and the measure data processing to verify its effectiveness. In the following simulation processing, the sparse base of the signal is selected from the Fourier orthogonal basis, the signal reconstruction algorithm is based on the orthogonal matching pursuit method. The selected signal is constituted by the four sinusoidal signals with the frequencies are 50, 100, 150, and 200 Hz, respectively. The sampling frequency is 800 Hz. The signal sampling point is 256, while 64 of them are selected for the signal reconstruction. The corresponding results are shown in Fig. 3.

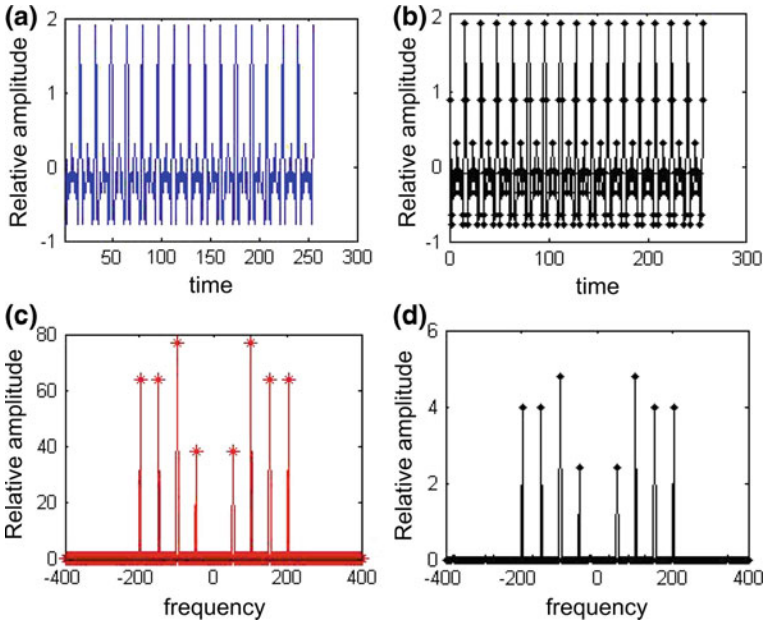



**Fig. 3.** Sinusoidal signals reconstruction by compressive sensing: **a** original signal waveform in time domain, **b** recovery signal by compressive sensing in time domain, **c** original signal waveform in frequency domain, and **d** recovery signal by compressive sensing in frequency domain.

The relative error of signal reconstruction is $1.5129 \times 10^{-14}$. For the sinusoidal signal with few signal components, the compressive sensing can restore the original signal accurately. In order to verify the effectiveness of the compressive sensing, the radar echo signals will be selected. The number of measurements is 256. The corresponding results are shown in Fig. 4.

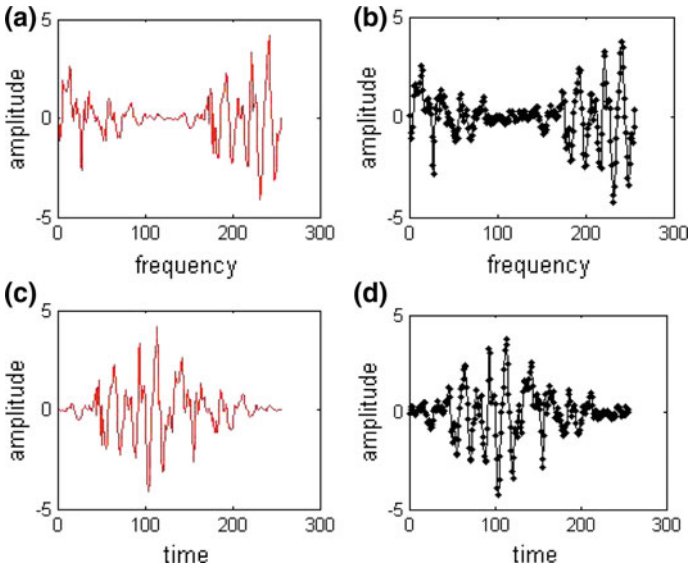



**Fig. 4.** Radar echo reconstruction by compressive sensing: **a** original signal waveform in time domain, **b** recovery signal by compressive sensing in time domain, **c** original signal waveform in frequency domain, and **d** recovery signal by compressive sensing in frequency domain.

The reconstructed error of the signal is 0.6969, so the compressive sensing algorithm can reconstruct the radar signal accurately.

- Simulation3

When the radar azimuth signals are missing, the compressive sensing technique can be used for the signal recovery, which can obtain the high-resolution images. Here, we uniformly select 50 samples from the total 1024 samples in the azimuth domain in simulation 1. The Fourier sparse base is used for the azimuth reconstruction by the compressive sensing technique, and the bistatic SAR imaging results are shown in Fig. 5a, b, respectively.

Figure 5a is the bistatic SAR imaging results by the Fourier transform. We can see that when the radar azimuth signals are missing, the corresponding SAR image is defocused by the Fourier transform. However, by using the compressive sensing technique in the SAR imaging process, the image quality can be improved, as shown in Fig. 5b. The simulated results demonstrate the effectiveness of the method proposed in this paper.

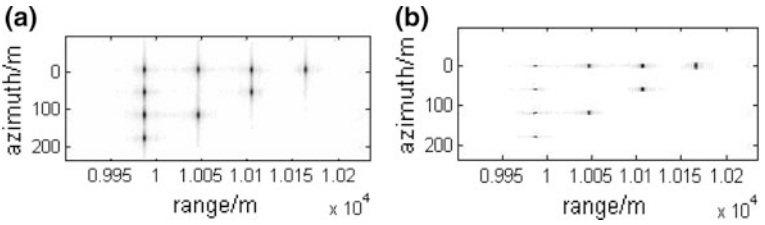



**Fig. 5.** Imaging results of the bistatic SAR by the two methods: **a** imaging results of the bistatic SAR by the Fourier transform. **b** Imaging results of the bistatic SAR by the compressive sensing

## 4 Conclusion

The bistatic SAR imaging can be equivalent to the baseline midpoint SAR imaging in some simple cases, which can simplify the calculation complexity of the algorithm. Then, the compressive sensing approach can be applied to obtain high-quality images for the situation of the bistatic SAR imaging with missing azimuth signals. The results of simulated data demonstrate the effectiveness of the algorithm proposed in this paper.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China under grant 61471149 and 61622107, and the Fundamental Research Funds for the Central Universities.

## References

1. Bu HX, Bai X, Tao R. Compressed sensing SAR imaging based on sparse representation in fractional Fourier domain. Sci China. 2012;55(8):1789–800.
2. Dong X, Zhang Y. A novel compressive sensing algorithm for SAR imaging. IEEE J Sel Top Appl Earth Obs Remote Sens. 2014;7(2):708–20.
3. Bu H, Tao R, Bai X, Zhao J. A novel SAR imaging algorithm based on compressed sensing. IEEE Geosci Remote Sens Lett. 2015;12(5):1003–7.
4. Barber B. Theory of digital imaging from orbital synthetic aperture radar. Int J Remote Sens. 1985;6(6):1009–57.
5. Candès EJ, Wakin MB. An introduction to compressive sampling. IEEE Signal Process Mag. 2008;25(2):21–30.
6. Donoho DL. Compressed sensing. IEEE Trans Inf Theor. 2006;52(4):1289–306.

# A ViSAR Imaging Method for Terahertz Band Using Chirp Z-Transform

Feng Zuo(✉) and Jin Li

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, Sichuan, China
zf_uestc@163.com

**Abstract.** Video synthetic aperture radar (ViSAR) as a new imaging mode has received more attention in recent years. In this mode, ViSAR system needs to persistently form a sequence of images while collecting data. For ViSAR, the frame rate needs to exceed 5 Hz to track maneuvering ground targets. Therefore, ViSAR system generally works in terahertz band. To avoid the 2-D interpolation of the polar format algorithm (PFA), a ViSAR imaging method is proposed in this paper. Based on the small synthetic angle characteristic of THz ViSAR, this method decomposes 2-D interpolation into range and azimuth interpolation, which is implemented by chirp Z-transform (CZT). The effectiveness of this method is validated by the simulation results in 0.3 THz.

**Keywords:** ViSAR · CZT · PFA · Terahertz

## 1 Introduction

Video synthetic aperture radar (ViSAR) is a new imaging mode. In this mode, the antenna beam invariably points to the center of the scene when radar platform circles the region of interest (ROI) [1]. ViSAR system forms a sequence of images while collecting data. Generally, the frame rate of ViSAR should exceed 5 Hz to track maneuvering ground targets. A sequence of images is continuously formed at the ViSAR frame rate, the process of which is viewed as forming a video. This is the origin of the name of ViSAR.

For the ViSAR frame rate of 5 Hz, it requires that the process of forming each SAR image needs to be accomplished within 0.2 s or less. Therefore, the processing efficiency plays a key role in ViSAR system. In general, there are two algorithms which are applied to form SAR images of circular flight path are the polar format algorithm (PFA) and the backprojection algorithm (BPA), respectively. The PFA is more computationally effective than the BPA [2], which result in the consequent that the PFA is more suitable for the reconstruction of ViSAR images. However, the 2-D interpolation of the PFA causes the imaging process to begin after the data collection of each frame is accomplished. It wastes time between each pulse. Therefore, the traditional PFA should be improved to adapt to the ViSAR imaging process in real time. Utilizing the small synthetic angle characteristic of THz ViSAR, we assume that the azimuth wave


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_96

numbers are evenly distributed after range resampling. Based on the assumption, the 2-D interpolation is decomposed into range and azimuth resample. This method can complete range resampling after receiving each pulse to improve the real-time performance of the algorithm. Meanwhile, this method implements the range and azimuth resampling using CZT, which reduces the complexity of the algorithm to improve the processing speed on the hardware, such as DSP, FPGA.

The remainder of this paper is organized as follows. In Sect. 2, the signal model of ViSAR is introduced. Section 3 presents the proposed imaging method of THz ViSAR. Sections 4 and 5 present the simulation results and the conclusion of the study, respectively.

## 2 Signal Model

Figure 1 illustrates the typical geometry of ViSAR where the flight path is circular with the radar beam providing continuous surveillance over a region of interest [3]. As shown in Fig. 1, $r_a$ denotes the slant distance between the antenna phase center and the center of the scene, $\varphi$ denotes the elevation angle, and $\theta$ denotes the azimuth angle.

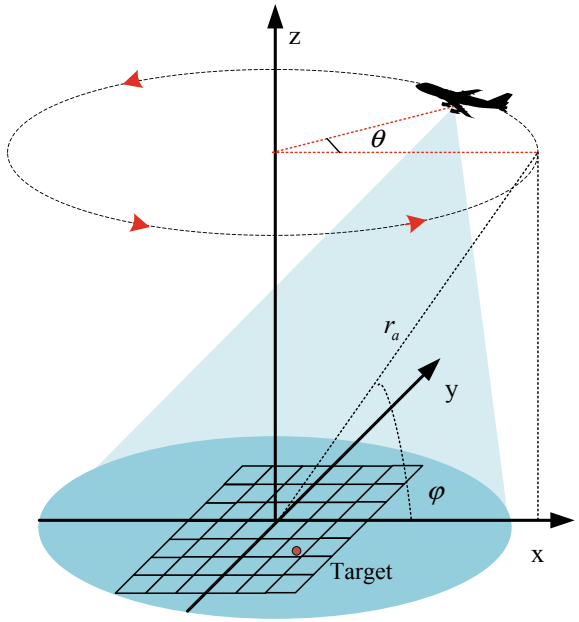



**Fig. 1.** ViSAR geometry with the platform traveling along a circular flight path.

Suppose that the radar transmits linear frequency-modulated (FM) signal at a constant linear FM rate and then the received de-chirped echo after removing the residual video phase from a point target is

$$s(\hat{t},t)=\mathrm{rect}\left[\frac{\hat{t}-2R(t)/c}{T_r}\right]\mathrm{e}^{-jK_r(R(t)-r_a)} \tag{1}$$

where $\hat{t}$ and $t$ denote the fast and slow time, respectively.$R(t)$ is the instantaneous range between the radar and the target. $T_r$ denotes the sweep time, $c$ denotes the speed of light. $K_r$ is the signal wavenumber, and it is defined as

$$K_r=\frac{4\pi}{c}(f_c+\gamma\hat{t}) \tag{2}$$

where $f_c$ is the carrier frequency and $\gamma$ is the chirp rate of the transmitted pulse.

Ignore some higher-order error terms and the range difference quantity $R(t)-r_a$ can be expanded to

$$R(t)-r_a\approx -x_0\cos\varphi\cos\theta-y_0\cos\varphi\sin\theta \tag{3}$$

where $(x_0,y_0)$ is the spatial coordinates of the point target. We can define the range and azimuth wavenumber components as

$$\begin{aligned}K_x&=K_r\cos\varphi\cos(\theta+\pi)\\K_y&=K_r\cos\varphi\sin(\theta+\pi)\end{aligned} \tag{4}$$

And the signal can be rewritten as

$$s(\hat{t},t)=\mathrm{rect}\left[\frac{\hat{t}-2R(t)/c}{T_r}\right]\mathrm{e}^{-j\left(x_0K_x+y_0K_y\right)} \tag{5}$$

For the phase history data, the inverse Fourier transform to yield a spatial image would be

$$I(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}s(\hat{t},t)\mathrm{e}^{j\left(xK_x+yK_y\right)}\mathrm{d}K_x\mathrm{d}K_y \tag{6}$$

When dealing with the sampled phase history data, the 2-D inverse fast Fourier transform is the efficient method for inverse Fourier transformation. However, the 2-D inverse Fourier transforms require uniformly spaced samples on a rectangular grid [4]. Therefore, it is necessary for the 2-D wavenumber domain data to interpolate from the polar grid to a rectangular grid as shown in Fig. 2.

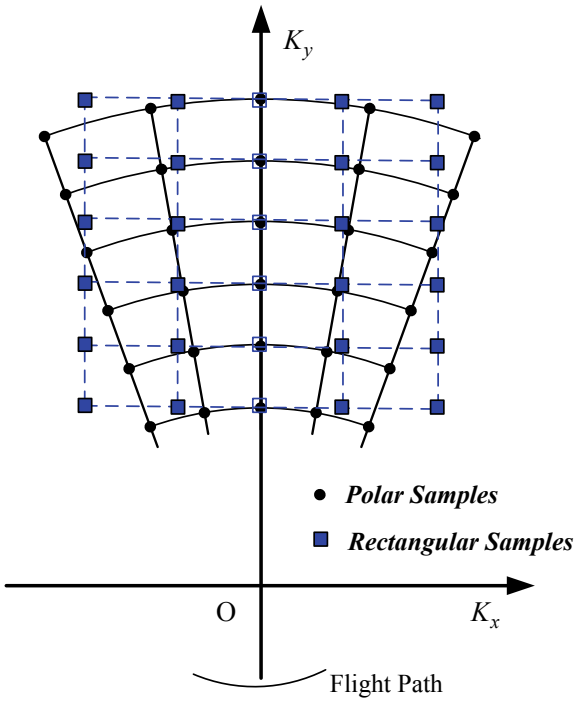



**Fig. 2.** Two-dimensional resampling process from polar to rectangular grid.

## 3 ViSAR Image Algorithm Based on CZT

For the spotlight mode with the linear flight path, the 2-D resampling in the PFA can be decomposed as the range and azimuth 1-D interpolation in order. For every interpolation, the phase history data are uniformly spaced samples on the range and azimuth wavenumber grid, respectively. In [5], the authors proposed an imaging method using CZT to replace the interpolation for the linear spotlight mode. The method is totally free of interpolation, for the sake of reducing the algorithm complexity.

However, the phase history data are non-uniformly spaced samples on the azimuth wavenumber grid after the range resampling. For the small spread of polar angles, we can approximate [6]

$$\cos(n\Delta\theta + \frac{\pi}{2}) \approx -n\Delta\theta \tag{7}$$

where $\Delta\theta$ is the azimuth angular increment in $\theta$ between pulses, and $n$ $(-N/2 \le n \le N/2 - 1)$ is the index of pulse number, $N$ is the number of pulses.

Fortunately, the synthetic angle is small that the approximation is accurate enough for terahertz band. However, the approximation will fail horrible with the wavelength of the radar increasing as shown in Fig. 3.

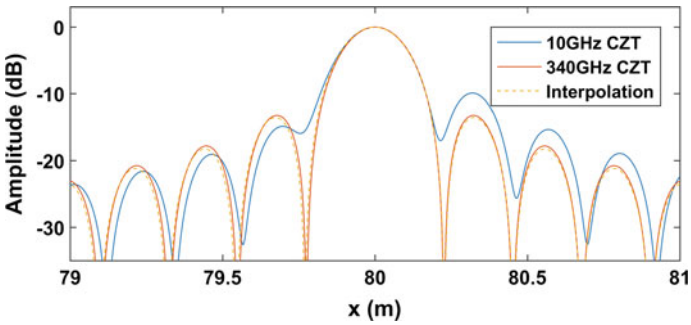



**Fig. 3.** Azimuth profiles of point target.

Figure 3 shows the azimuth profiles of point target located at (80, 80) in 10 GHz and 340 GHz, whose resolution is 0.2 m, using CZT to implement azimuth resampling based on the assumption of (5). Comparing with the result of azimuth non-uniform interpolation, the azimuth imaging quality almost the same with it in 340 GHz, unlike the imaging quality drops seriously in 10 GHz.

The CZT is the Z-transform along a spiral contour. Unlike the FFT, the CZT can perform a spectral zoom operation that the spectral zoom can start and end at any arbitrary frequency defined by $A$ and $W$.

$$A = \exp(j\omega_s), \qquad W = \exp(-j\Delta\omega) \tag{8}$$

where $\omega_s$ donate the start digital frequency and $\Delta\omega$ is the digital frequency increment in the CZT output digital frequency. Therefore, the CZT is often used for implementation of resampling from uniform grid to uniform grid.

The corresponding relation between the digital frequency and the range wavenumber is shown in Fig. 4. From the relationship, it is easy to gain $\omega_s$ and $\Delta\omega$. Then, the parameters of range CZT can be yield

$$\begin{aligned} A_r &= \exp\left(j\left(2\pi - 2\pi\frac{K_{\text{ymin}}/\cos(\theta+\pi) - K_{\text{ymin}}}{K_{\text{ymax}} - K_{\text{ymin}}}\right)\right) \\ W_r &= \exp(-j\frac{\Delta K_y}{\cos(\theta+\pi)}\frac{-2\pi}{K_{\text{ymax}} - K_{\text{ymin}}}) \end{aligned} \tag{9}$$

where $K_y$ is the range wavenumber, it can be defined by

$$K_{\text{y}} = \frac{4\pi}{c}(f_c + \gamma\hat{t})\cos\varphi \tag{10}$$

$K_{\text{ymin}}$ and $K_{\text{ymax}}$ donate the minimum and maximum, respectively. $\Delta K_y = \frac{4\pi}{c}\frac{B}{M}$ is the range wavenumber increment and $M$ is the number of samples, $B$ is the signal bandwidth.

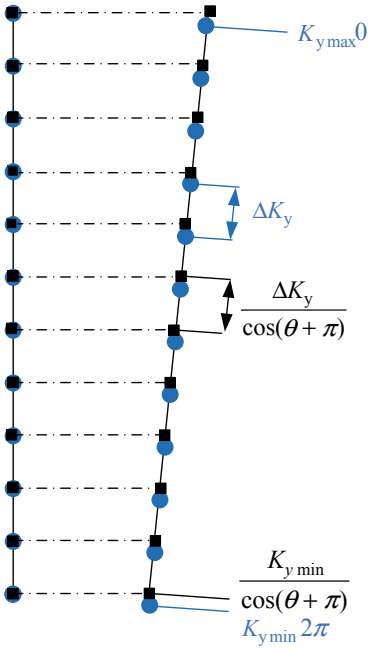



**Fig. 4.** Relationship between the digital frequency and range wavenumber.

Similarly, the azimuth CZT parameters can be obtained through the corresponding relation. A conclusion is given here, and the process of derivation is no longer repeated.

$$\begin{aligned} A_a &= \exp\left(j\frac{K_y}{K_{yc}}\pi\right) \\ W_a &= \exp\left(j\frac{K_y}{K_{yc}}\frac{2\pi}{N_I}\right) \end{aligned} \tag{11}$$

where $K_{yc} = \frac{4\pi f_c}{c}\cos\varphi$ is the range center wavenumber and $N_I$ is a scalar that specifies the length of CZT.

Therefore, the complete flow of THz ViSAR imaging algorithm is shown in Fig. 5.

## 4 Simulation Results

The simulation parameter is listed in Table 1, where $\rho$ donates the image resolution. The position of point targets is shown in Fig. 6. Figure 7a, b shows the imaging result with different center angle. From these imaging results, we can see that the position relationship between every point target is not in accordance with the actual position relationship. It is caused by the plane wave assumption in (3). It has been described in many related papers and it can be compensated by 2-D resampling in the spatial image domain, such as [7]. Therefore, we do not repeat the discussion and ignore it.

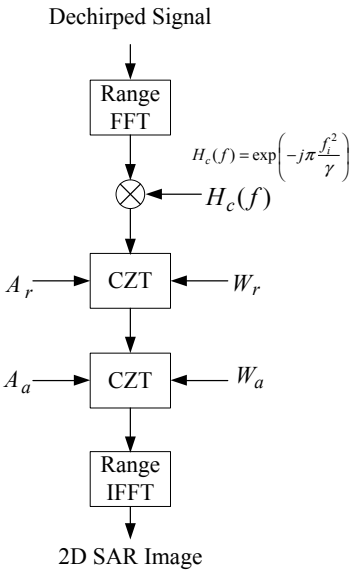



**Fig. 5.** Flow diagram of THz ViSAR imaging algorithm.

**Table 1.** Simulation parameters.

| $f_c$ | $\varphi$ | $r_a$ | $\rho$ |
|---|---|---|---|
| 340 GHz | 45° | 3 km | 0.3 m |

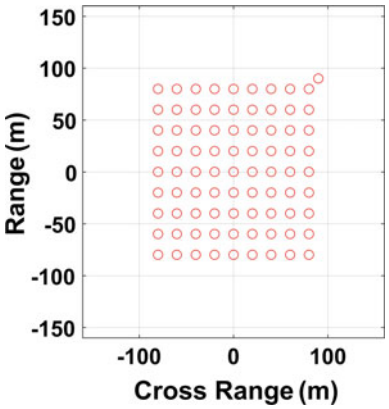



**Fig. 6.** Position of the point targets.

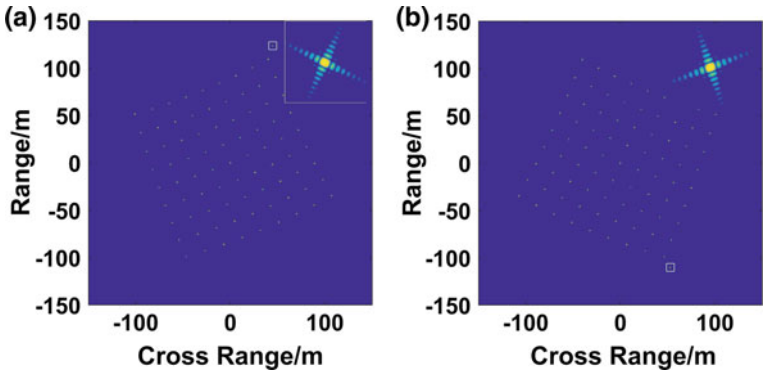



**Fig. 7.** Imaging results with different center angle. **a** $-5/8\pi$; **b** $\pi/8$.

For the convenience of observation imaging results, we zoom in the point target imaging result located in the edge of ROI. The zoomed imaging results illustrate that point targets have an excellent imaging performance and it verifies the effectiveness of the THz ViSAR algorithm based on CZT.

## 5 Conclusion

A THz ViSAR imaging algorithm is proposed in this paper. Utilizing the feature of THz ViSAR, the 2-D resampling is decomposed into twice 1-D resampling. In this method, we use the CZT to realize 1-D resampling, which results that the method is totally free of interpolation to reduce the algorithm complexity. In the end, the simulation verifies the effectiveness of the proposed method.

**Acknowledgments.** This work was supported by the 2011 Project for Collaborative Innovation Center of Information Sensing and Understanding.

## References

1. Song X, Yu W. Processing video-SAR data with the fast backprojection method. IEEE Trans Aerosp Electron Syst. 2016;52(6):2838–48.
2. Gorham LRA, Moore LJ. SAR image formation toolbox for MATLAB. In: Algorithms for synthetic aperture radar imagery XVII. International Society for Optics and Photonics; 2010, vol. 7699. p. 769906.
3. Zuo F, Li J. A persistent imaging method for video-SAR in terahertz band. In: 2017 International applied computational electromagnetics society symposium (ACES). IEEE; 2017. p. 1–2.
4. Jakowatz CVJ, Wahl DE, Eichel PH, et al. Spotlight-mode synthetic aperture radar: a signal processing approach: a signal processing approach. Springer Science & Business Media; 2012. p. 133–6.

5. Zhu D, Zhu Z. Range resampling in the polar format algorithm for spotlight SAR image formation using the chirp z-transform. IEEE Trans Signal Process. 2007;55(3):1011–23.
6. Doerry AW. Basics of polar-format algorithm for processing synthetic aperture radar images. Sandia National Laboratories report SAND2012-3369, Unlimited Release; 2012.
7. Garber WL, Hawley RW. Extensions to polar formatting with spatially variant post-filtering. In: Algorithms for synthetic aperture radar imagery XVIII. International Society for Optics and Photonics; 2011, vol. 8051. p. 805104.8.

# The Effect of Ground Screen System on High-Frequency Surface Wave Radar Antenna Directivity

Linwei Wang, Changjun Yu(✉), and Yi Huang

The School of Information and Electrical Engineering, Harbin Institute of Technology at Weihai, Weihai, China
yuchangjun@hit.edu.cn

**Abstract.** For quarter wave vertical antenna in high-frequency band, a ground screen system is indispensable. The ground screen is laid to reduce the effects by earth whose conductive is nonuniform and poor. High-frequency surface wave radar (HFSWR) system demands antenna directivity stricter than the radio station, so the system needs a high-quality ground screen system. This work examines the performance of the ground screen system for high-frequency antenna and finds applicable size and shape for antenna fabrication. Different ground screen sizes, densities and shapes are compared in the simulation for a 9.8 m tall monopole antenna in 7 MHz. The simulation results show that the size and density of the ground screen are important parameters for antenna, and radial ground screen performs better than latticed when the size of the ground screen is large. To find appropriate ground screen parameters for HFSWR, the analysis takes the effect of sea and coast into account. Furthermore, the results may provide guidance for ground screen laying in practical applications.

**Keywords:** High-frequency surface wave radar · Ground screen system · Antenna directivity

## 1 Introduction

High-frequency surface wave radar (HFSWR) operates in 3–30 MHz, with the corresponding radio wavelength 10–100 m. In HFSWR, vertically polarized monopoles are generally utilized. Theoretically, the conductive coast or sea water is considered as the reflect plane of the antenna. In practical, however, the conductivity of the dry coast or rocks is inhomogeneous and poor. In order to reduce the effects on antenna performance by the resistance of earth, ground screen system (GSS) is indispensable.

Brown [1] analyses the GSS and presents a system consisting of 120 radial wires, each one-half wave long, is desirable as early as 1937. Bekefi [2] gives the theoretical derivation and numerical solving method of impedance of the ground plane. Benjamin [3] shows that a GSS with 120 radials is a very conservative design, smaller systems can also produce acceptable performance. Trainotti [4, 5] analyses the effect of GSS with a practical situation in theoretical arithmetic. Jin [6] discusses the inhomogeneous ground's effect on the dipole antenna radiation field and gives a formula applied to


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_97

complicated structure. Christman [7] designs a transmitting antenna for commercial medium wave broadcast stations.

However, limited by the technique, these works can only appropriately analyse the performance in theory or in actual measurement. Along with the development of computer technology, we can conveniently analyse the antenna performance by simulation for more aspects of the GSS in different situations.

For monopole antenna, infinite perfect electric conduction ground plane is required in theory, but in practice, we can only fabrication a ground screen in finite space and resources. According to the engineering experience of the radio station, the ground screen should consist of 60–120 radial electric conduction wires whose lengths are half of wavelength. In some specific scene which needs higher quality of antenna radiation such as radar, the engineering experience maybe not reliable enough. To make antenna performance better in finite resource, analysis of ground screen is necessary.

In this paper, we analyse the GSS in aspects of screen sizes, densities and shapes, for a simulated system whose parameters is the same with a real HFSWR. The rest of the paper is organized as follows. In Sect. 2, we introduce the geometry structure of the system. Section 3 provides the simulation results. Conclusions are given in Sect. 4.

## 2 Geometry

The simulations performed by HFSS consider a monopole antenna that is 9.8 m height, whose radius is 10 cm, and material is PEC, operating in 7 MHz. The antenna is on the infinite ground plane, whose conductivity is 0.003 S/m, and relative dielectric constant is 4. Several electrical wires with 6 mm radius are laid on the ground constructing the GSS. The electrical wires length $L$ changes from $\lambda/16$ to $\lambda$, and the number of wires $N$ changes from 8 to 72. Figure 1 shows the HFSS model.

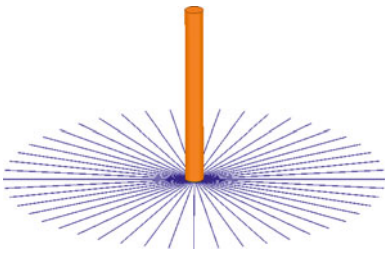


**Fig. 1.** 9.8 m monopole antenna and 48 wires radial GSS model in HFSS

In an ideal condition, the ground should be a perfect electric conductor, in which case, the radiation field in theory is

$$|E| = \frac{60I_A}{r \sin kh} \frac{\cos(kh \sin\theta) - \cos kh}{\cos\theta} \tag{1}$$

where $I_A$ is the current fed into the antenna, $r$ is the distance from antenna, $h$ is the height of the antenna. The antenna directivity can reach about 5 dBi when $h = \lambda/4$. However, in practice, due to the soil impedance, displacement currents that flow between the antenna and the ground terminal of the monopole are very large which obviously reduces the performance of the antenna. The GSS provides a low-impedance return path to reduce the currents.

## 3 Results

### 3.1 Effect of Ground Electrical Parameters

Table 1 shows common geology electrical parameters, including dry soil, wet soil, freshwater and sea water. In different conditions of the ground, the antenna exerts different performance. Figure 2 is the antenna gain in different conductivity conditions without GSS, which indicates that as conductivity increasing the antenna gain also increases, but the increasing speed is decreasing.

**Table 1.** Common geology electrical parameters

| Geology | Relative dielectric constant | Conductivity (S/m) |
|---|---|---|
| Dry soil | 4 | 0.003 |
| Wet soil | 80 | 0.03 |
| Freshwater | 80 | 0.001 |
| Sea water | 80 | 4 |

When conductivity is 1 S/m or 4 S/m, the antenna gain values are similar to ideal condition, in these parameters, the ground can be regarded as perfect electricity conductor.

### 3.2 Effects of Ground Wire Length

The simulated directivity from the model constituted by 48 wires radial GSS is shown in Fig. 3, which indicates that the effect of ground wire length approaches an inverse proportional function. When wire is longer than 3/4 of wavelength, the effect becomes negligible.

The reflection coefficient variation shows some periodicity in the simulation, as shown in Fig. 4. When the wire length equal to 3/8 of the wavelength, the reflection coefficient performs the best.

### 3.3 Effects of Ground Wires Number

The density, which can be described as wires number, is another factor effecting antenna performance. Figure 5 shows the antenna directivity that changes with different number of ground wires. When ground wire length is short, wire number does

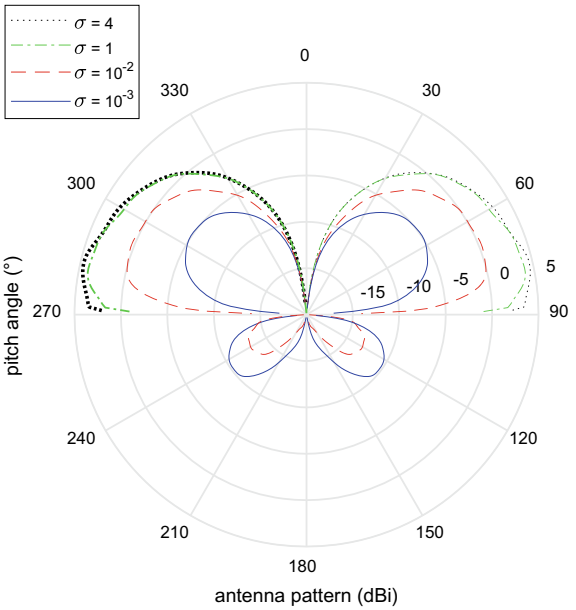



**Fig. 2.** Different conductivity impact on antenna pattern

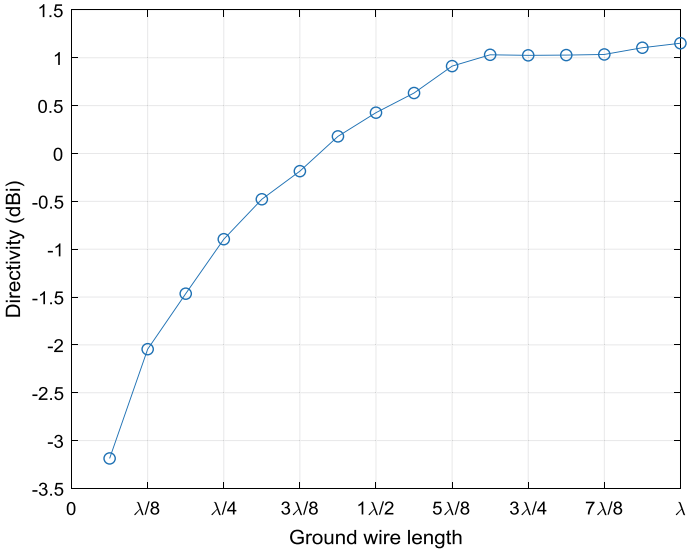



**Fig. 3.** Directivity affected by ground wire length

not work well, wire density has no influence on the directivity of the antenna. While, when ground wire length increases, GSS with less wires number reaches the limitation earlier. In other word, the number of ground wires determines the upper limit value.

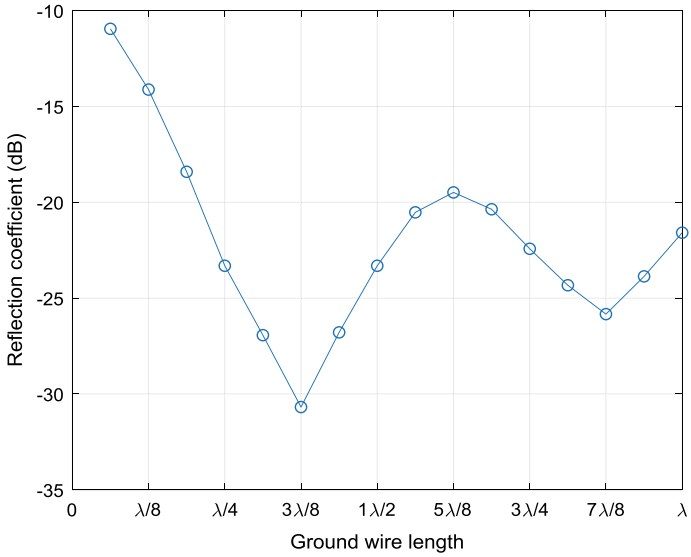



**Fig. 4.** Reflection coefficient affected by ground wire length

GSS with eight wires supports −1 dBi directivity in maximum, in spite of the wire length, while 48 wires support more than 1 dBi directivity when length is longer than 3/4 wavelength.

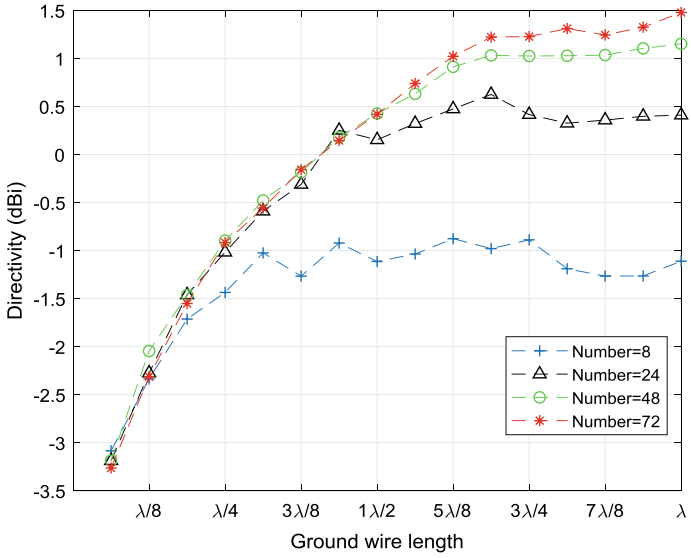



**Fig. 5.** Directivity affected by ground wire number

Combined with ground wire length and number, a function to describe the variety trend can be found with curve fitting tool in MATLAB. Equation (2) shows the result, where $D$ is the antenna directivity, $x$ is the ratio between wire length and wavelength. The wavelength in high frequency is 10–100 m, which means that the antenna with GSS needs more than 200 $m^2$ for an ideal directivity performance. A half wavelength GSS can offer about 0.5 dBi directivity at most. In this condition, 48 wires are enough according to Fig. 5.

$$D = \frac{4.98x - 2.002}{x + 0.4694} \tag{2}$$

### 3.4 Latticed GSS

In some scene, latticed GSS is also usually used, since it fabricates easier. However, simulation shows that, in the same area and wire length, radial GSS has better performance than latticed GSS. These two GSSs have same performance when wire length is small, but the performance of latticed GSS is limited by density when GSS size enlarged. For seam space, latticed GSS needs higher density. When wire number is 72, latticed GSS only gets −2 dBi directivity, while radial GSS can get more than 1 dBi directivity.

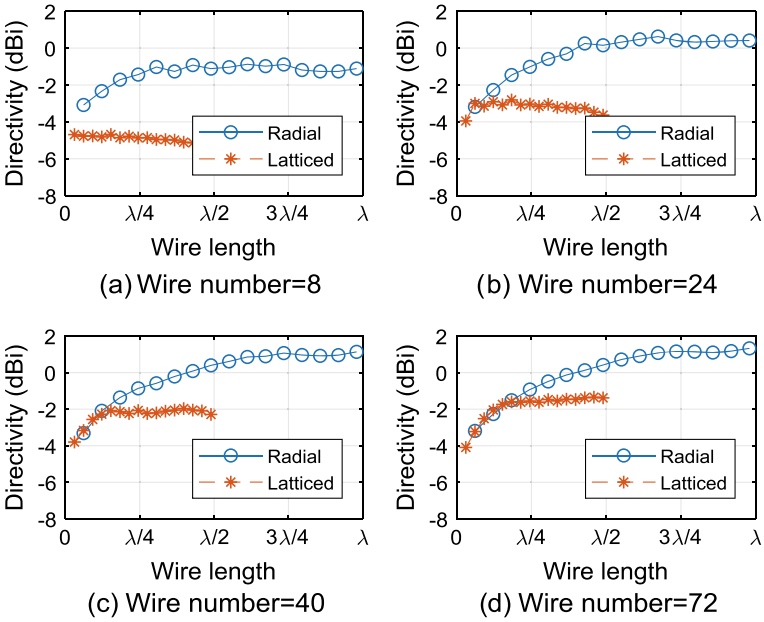



**Fig. 6.** Comparison between radial GSS and latticed GSS

### 3.5 Effects of Sea

According to Sect. 3.1, sea water can be considered as good conductor, which means that the sea can be regarded as a part of the GSS. Considering the effects of the sea, the performance of the antenna improves significantly, since the size of sea much larger than radio wavelength (Fig. 6).

Figure 7 shows the model that consists of antenna, GSS, ground and sea. The antenna is 15 m far from the sea on the coast. In this model, the conductivity of sea water is 4 S/m, and the relative dielectric constant is 81. Considering that the moisture of soil near the sea is high, the conductivity of soil is 0.03 S/m, relative dielectric constant is 80. The GSS is constructed by 48 wires, whose radius and length are 6 mm and quarter wavelength, respectively.

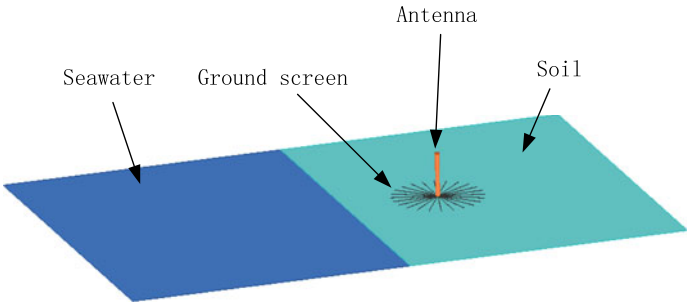



**Fig. 7.** HFSS model consists of antenna, GSS, ground and sea

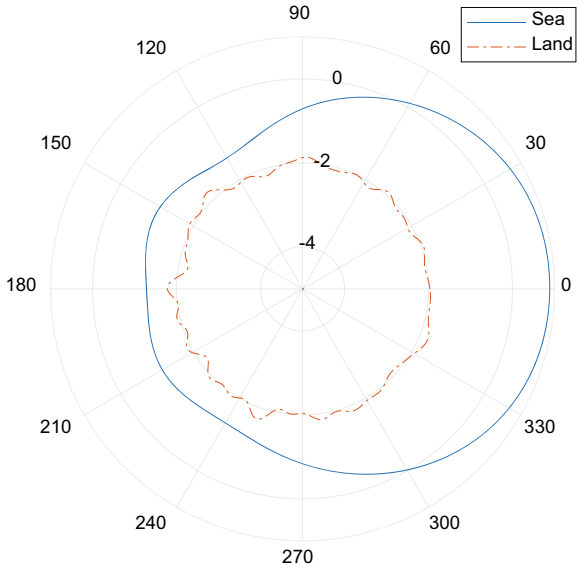



**Fig. 8.** Horizontal plane pattern considering the effects of sea

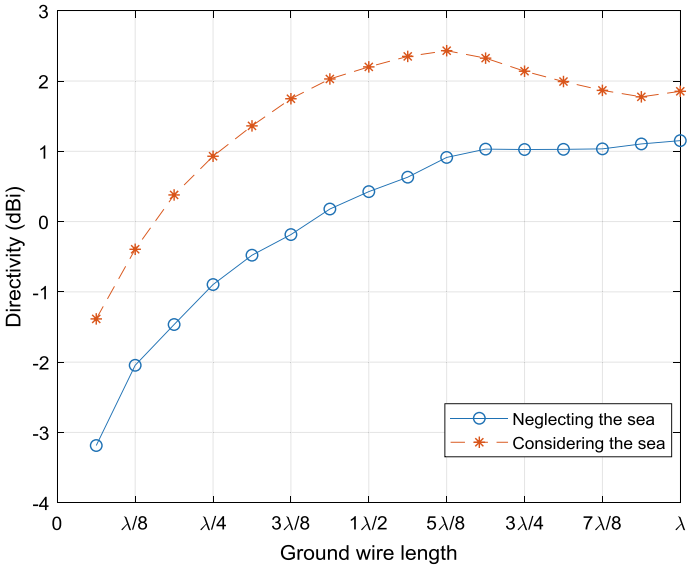



**Fig. 9.** Directivity affected by ground wire length and sea

Sea water leads to inhomogeneous of the ground system; therefore, the antenna pattern has changed. As shown in Fig. 8, 0° is the direction from the antenna to the sea. Antenna pattern is a circle when neglecting the effects of the sea, while the value of pattern increases when considering the effects of the sea. The increment in the direction of the sea is 3 dB while the negative direction increases 1 dB.

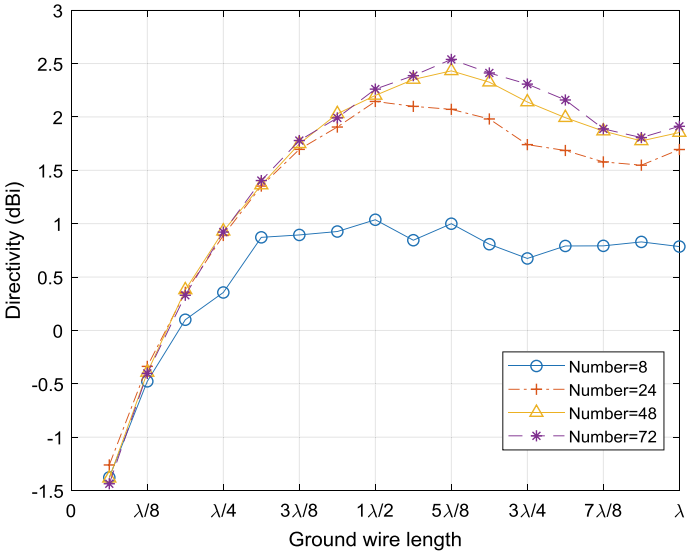



**Fig. 10.** Directivity affected by ground wire number and sea

In different ground wire length, antenna directivity changes as shown in Fig. 9. Compared with results in Sect. 3.2, the directivity increases about 2 dB.

Recalculating all condition of ground wire length and number, we have Fig. 10. Comparing with Figs. 5 and 10 illustrates that the effect of the sea improves the performance of almost all condition in 2 dB, while the trend of curves still the same. When the wire length and number is large, the result may be inaccurate because of the lack of computing resource.

## 4 Conclusion

In this paper, we have analysed the parameters that influence the performance of the GSS. Based on the results, we can conclude that, in order to gain a higher antenna directivity, larger GSS size is preferred. For HFSWR, it is extremely difficult to reach the theoretical value, due to the limited field. In finite space, excessive increasing density makes no sense. The maximum of directivity in different GSS size is given by (2). In practice, for an area in which the ground wire length less than a half wavelength, a GSS with 48 wires is satisfied. Considering the effect of the sea, the GSS with 48 wires in a half wavelength offers 2 dBi antenna directivity.

**Acknowledgments.** This work was supported in part by the National Key R&D Program of China under Grant 2017YFC1405202, in part by the National Natural Science Foundation of China under Grant 61571159 and Grant 61571157, and in part by the Public Science and Technology Research Funds Projects of Ocean under Grant 201505002.

## References

1. Brown GH, Lewis RF, Epstein J. Ground system as a factor in antenna efficiency. Proc Inst Radio Eng. 1937;25(6):753–87.
2. Bekefi G. the impedance of an antenna above a circular ground plate laid upon a plane earth. Can J Phys. 2011;32(3):205–22.
3. Dawson BF, Stephen SL. Revisiting medium-wave ground-system requirements. IEEE Antennas Propag Mag. 2008;50(4):111–4.
4. Trainotti V, Dorado LA. Short low- and medium-frequency antenna performance. IEEE Antennas Propag Mag. 2005;47(5):66–90.
5. Trainotti V, Figueroa G. Vertically polarized dipoles and monopoles, directivity, effective height and antenna factor. IEEE Trans Broadcast. 2010;56(3):379–409.
6. Jin YS, Cai ZY, Qin JP. Influence of ground inhomogeneity on performance of DF antenna system. Chin J Radio Sci. 2008;23(5):873–6.
7. Christman A. A very large antenna for the 1.8 MHz band. IEEE Antennas Propag Mag. 2002;42(4):75–9.

# A Rateless Transmission Scheme for Underwater Acoustic Communication

Fan Bai[1,2], Zhiyong Liu[1,2(✉)], and Yinyin Wang[1,2]

[1] School of Information and Electrical Engineering, Harbin Institute of Technology (Weihai), Weihai 264209, People's Republic of China
694683139@qq.com, lzyhit@aliyun.com, wcuteyy@163.com
[2] Key Laboratory of Science and Technology on Information Transmission and Dissemination in Communication Shijiazhuang, Shijiazhuang 050081, People's Republic of China

**Abstract.** In view of characteristics of channel parameters changing with time and space of underwater acoustic channel, a rateless transmission scheme for adaptive transmission over underwater acoustic channels is studied in this paper. The scheme utilizes the characteristics of LT coding without fixed coding rate, and the communication efficiency is further improved through single-carrier frequency-domain equalization. Simulation results show that the scheme we proposed can effectively reduce the computational complexity compared to the rateless coding system under classical time-domain equalization. With achieving the same reliable information transmission in the same underwater acoustic channel, the redundancy required by the proposed scheme is about 1/3 of the classical retransmission method, which improves the efficiency of information transmission.

**Keywords:** Underwater acoustic channel · Adaptive transmission · Single-carrier frequency-domain equalization · LT coding · Redundancy

## 1 Introduction

Underwater acoustic communication is known as to be very difficult telecommunication, which also is faced with problems such as low signal-to-noise ratio and time-varying and space-variant problems of the underwater acoustic channels. Additionally, the condition of the feedback channel is the cause of that the communication efficiency is seriously affected by the long propagation distance and the long feedback delay time.

The problems existing in the above underwater acoustic channels are generally improved by time-domain retransmission [1] method, but the communication efficiency is not high in many cases. At the same time, the network of underwater acoustic communication has become a trend [2] with the development of underwater acoustic communication. A higher encoding rate flexibility and a lower correlation with the channel and the communication nodes are required in the encoding scheme applied to the underwater acoustic communication network. The traditional methods are difficult


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_98

to meet the requirements, and then, the emergence of rateless coding provides a new way to solve this problem. In 1998, Byers et al. [3] proposed the concept of rateless codes. In 2002, Luby [4] proposed the first LT coding that is consistent with the concept of rateless codes and has practical value.

Existing studies are generally directed at rateless codes in time-domain equalization (TDE) underwater acoustic systems [5, 6]. However, the frequency-domain equalization (FDE) technique has much less computational complexity than the TDE technique under underwater acoustic communications. Therefore, the single-carrier frequency-domain equalization (SC-FDE) technique is used in this paper to replace the classical TDE method, and the scheme improves the efficiency of communication while introducing the advantages of LT coding, also improves the reliability of underwater acoustic communication.

### 1.1 System Model

The system mode is illustrated in Fig. 1. At the transmitter, the information bit data is sent in LT encoding, and then, the symbol mapping is performed to complete modulation, the modulated data $x_n$ is subjected to modulation conversion and a cyclic prefix (CP) is inserted in each data block, and finally, the signal $x_n'$ is transmitted.

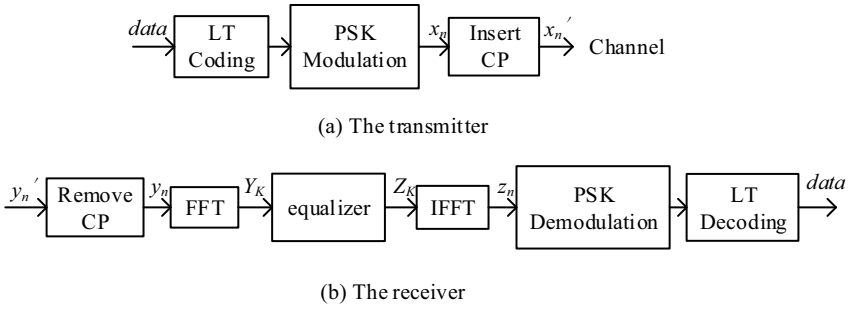



**Fig. 1.** Transmission system mode

The signal $x_n'$ passes through the underwater acoustic channel and arrives at the receiver. At the receiver, the received signal $y_n'$ is first subjected to remove cyclic prefix. Then, after the cyclic prefix is removed for each data block $y_n$ is converted to the frequency domain for equalization processing. After equalization is completed, the signal $Z_K$ is converted to the time-domain demodulation and to obtain the final *data* after decoding. The following will describe the system mode mathematically. In the description process, let the length of the signal block be $N$, the transmission rate is $1/T$, the transmission signal adopts the form of adding the cyclic prefix, the signal of the receiving end is expressed as

$$y_n = h_n \otimes x_n + v_n, \quad n = 0, 1, \ldots, N-1 \tag{1}$$

where $h_n$ represents the impulse response of the channel, and $v_n$ represents additive white Gaussian noise with a mean of zero and a variance of $n^2$. Convert the above equation to the frequency domain,

$$Y_K = H_K \times X_K + V_K, \quad k = 0, 1, \ldots, N-1 \tag{2}$$

Set the FDE coefficient to $W_K$, then the signal after FDE can be expressed as,

$$Z_K = Y_K \times W_K = H_K X_K W_K + V_K W_K \tag{3}$$

Using minimum mean square error (MMSE) equilibrium,

$$W_K = \frac{H_k^*}{|H_K|^2 + \sigma^2/P}, \quad k = 0, 1, \ldots, N-1 \tag{4}$$

where $P$ is the power of the signal.

### 1.2 Rateless Transmission Scheme

Underwater acoustic channel is a typical multipath channel, the system performance improvement is restricted because the complexity of the time-domain equalizer is proportional to the maximum delay spread of the channel in practice. Previous studies have shown that the SC-FDE technique effectively combines the advantages of single-carrier modulation and OFDM modulation [7], which gives it a very broad application prospect in the future of high-speed underwater acoustic communication. On this basis, we have proposed an LT coded scheme based on SC-FDE for underwater acoustic communication.

#### 1.2.1 Encoding LT Codes

Assume that the existing $k$ symbols are to be transmitted. Then, each LT code symbol is generated as follows.

1. According to the distribution of degrees, an integer $d$ greater than 0 is generated as the degree of the current encoding by using a random generator;
2. With another generator, $d$ integers between 1 and $k$ are uniformly distributed, which is the sequence number of $d$ information;
3. According to the $d$ number, select the corresponding symbol and perform XOR operation to obtain the code symbol to be sent.

#### 1.2.2 Adaptive Frequency-Domain Equalizer

For the linear equalization mode (MMSE) described above, the channel needs to be relatively accurately channel estimated, and then, the equalization coefficient is calculated based on the estimated value to compensate for the signal, and the adaptive equalizer can quickly generate the appropriate equalization coefficients according to

the change of the channel. The adaptive frequency-domain equalizer is an adaptive algorithm that incorporates signal processing in the design of the equalizer.

The least mean square (LMS) algorithm [8] is a more commonly used adaptive iterative algorithm. The specific process of the adaptive frequency-domain LMS algorithm is as follows: The input signal $x(n)$ and the desired signal $d(n)$ form an $N$-point data block and then perform an $N$-point fast Fourier transform (FFT) and the output of each FFT is composed of $N$ complex points $X(K)$ and $D(K)$, and the output of the time-varying equalizer is $Y(K)$,

$$Y(k) = X(k)W(k) \tag{5}$$

In the simulation, the overlap retention method can be used to implement FFT. In this method, the length of the overlapped portion can be set to an arbitrary value, but the 1/2 overlap is the highest computational efficiency. Since 1/2 data overlap is the most effective, that is, $M$ zeros are added after the $M$-point equalizer tap coefficients, and then, $N$-point FFT is performed, where $N = 2M$, i.e.,

$$\begin{aligned} W^T(k) &= [W_0(k), \ldots, W_{2M-1}(k)] \\ &= \mathrm{FFT}\left[W^T(k), 0, \ldots, 0\right] \end{aligned} \tag{6}$$

Let $X(K)$ be the diagonal matrix formed by the $2M$ (i.e., $N$) point FFT after inputting the $k-1$th block and the $k$th block, i.e.,

$$X(k) = \mathrm{diag}\left\{ \mathrm{FFT}\left[ \underbrace{x(kM - M), \ldots, x(kM - 1)}_{(k-1)\mathrm{th}}, \underbrace{x(kM), \ldots, x(kM + M - 1)}_{k\mathrm{th}} \right] \right\} \tag{7}$$

$$\begin{aligned} y^T(k) &= [y(kM), \ldots, y(kM + M - 1)] \\ &= \mathrm{IFFT}[X(k)W(k)] \end{aligned} \tag{8}$$

The output value of the $k$th block is the last $M$ term of formula (8). Because the first $M$ elements are circularly convolved, only the last $M$ elements need to be retained. The $k$th block expectation response vector is,

$$d(k) = [d(kM), d(kM + 1), \ldots, d(kM + M - 1)]^{\mathrm{T}} \tag{9}$$

The corresponding error function is,

$$\begin{aligned} e(k) &= d(k) - y(k) \\ &= [e(kM), e(kM + 1), \ldots, e(kM + M + 1)]^{\mathrm{T}} \end{aligned} \tag{10}$$

In the process of linear convolution in (8), the starting M elements have been discarded from the output, and the error function $e(k)$ vector can be transformed into the frequency domain,

$$E(k) = \mathrm{FFT}\left\{ \left[ \underbrace{0, \ldots, 0}_{M}, e^{T}(k) \right]^{\mathrm{T}} \right\} \tag{11}$$

Since linear correlation is actually a form of operation similar to linear convolution, memory retention methods can be used in vectors $\psi(k)$, it can be calculated using the first $M$ terms of the IFFT Eq. (12).

$$\psi(k) = \mathrm{IFFT}[X^{*}(k)E(k)] \tag{12}$$

Finally, the frequency-domain weight vector adjustment formula is,

$$W(k+1) = W(k) + 2\mu\mathrm{FFT}\left\{ [\psi(k), 0, \ldots, 0]^{\mathrm{T}} \right\} \tag{13}$$

Therefore, the flowchart [1] of the adaptive LMS frequency-domain block algorithm is shown in Fig. 2.

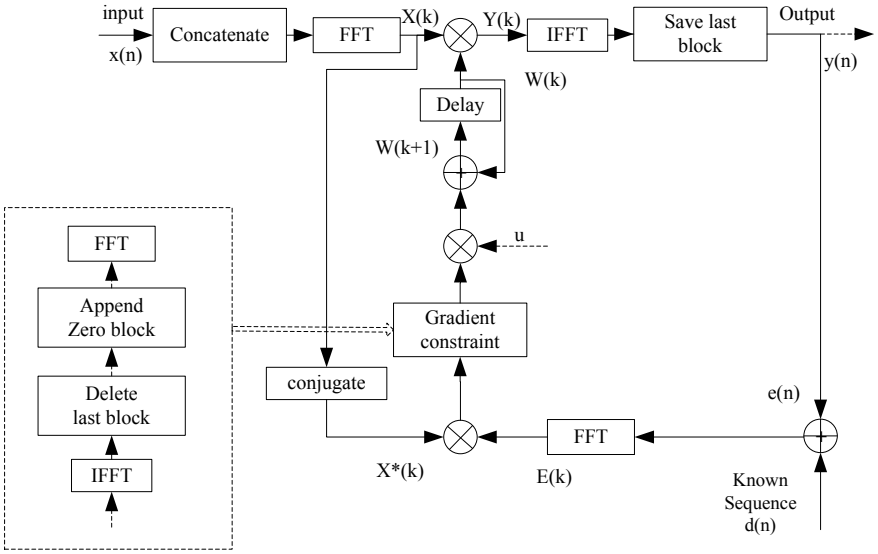



**Fig. 2.** Flowchart of the adaptive LMS frequency-domain block algorithm

### 1.2.3 Decoding LT Codes

The decoding of the LT code needs to start with a code symbol with a degree of 1, and the specific process is as follows:

1. When the user receives $n$ ($\geq k$) coded symbols, look for coded symbols of degree 1 among the $n$ coded symbols;

2. If a code symbol with a degree of 1 can be found, all code symbols of the degree 1 are recovered from the corresponding seed, and the coded symbols of degree 1 are deleted; otherwise, go to section 1;
3. The encoded symbols that have not been decoded are XORed to remove the recovered symbols; the encoding symbol whose degree is 1 continues to be searched, and goes to section 2; if there are no encoded symbols and there are unrecovered transmitted symbols, then to section 4, otherwise it is successful decoding, then stop.
4. Continue to receive more code symbols;
5. For the newly received code symbol, remove the recovered symbol from the code symbol by XOR operation according to the recovered symbol and then transfer to section 2.

### 1.3 Simulation Results

Simulation is built on BELLHOP underwater acoustic channel model. The carrier frequency is 12 kHz. The distance between transmitter and receiver is 100 m, and they all located at a depth of 10 m. Wave height is 0.2 m. The basic system parameters are set as follows: The source is 0, 1-bit data stream, one packet is 1000 bits, and 100 packets are sent; using BPSK modulation data, the front and back of each frame are, respectively, inserted with a unique word of length 64.

First, we compare the performance of the LT coding system with FDE and the system with TDE in the existing research. On the one hand, the BER performance of the two is compared, and the simulation results are shown in Fig. 3. It can be seen that two methods can achieve reliable decoding and the minimum required input-to-noise ratio for both is about 6 to achieve BER low at $10^{-3}$, which means that there is not much gap in the error performance of the two methods.

On the other hand, we compare the performance of the two algorithms by analyzing the computational amount. Calculating the calculation amount of two equalizer algorithms by using real multiplication as the standard, the ratio of the computational complexity of the LMS algorithm with FDE and the time-domain LMS algorithm can be expressed as

$$f(N) = \frac{5N\log_2 2N + 16N}{2N^2} = \frac{5\log 2N + 13}{N},$$

where $N$ is the number of points for each input sequence. According to the formula, the simulation curve is shown in Fig. 4, and it can be found obviously that the computational complexity with FDE is much less than TDE with lager $N$. However, the required filtering order is large under underwater acoustic channel, so the FDE is more suitable for underwater acoustic communication. In summary, the scheme proposed in this paper is more suitable for underwater acoustic channel than the system under classical TDE in existing studies.

In order to further verify the performance of rateless coding transmission scheme in terms of coding flexibility and communication reliability, rateless coding transmission scheme and classical time retransmission method are compared. A sufficient number of

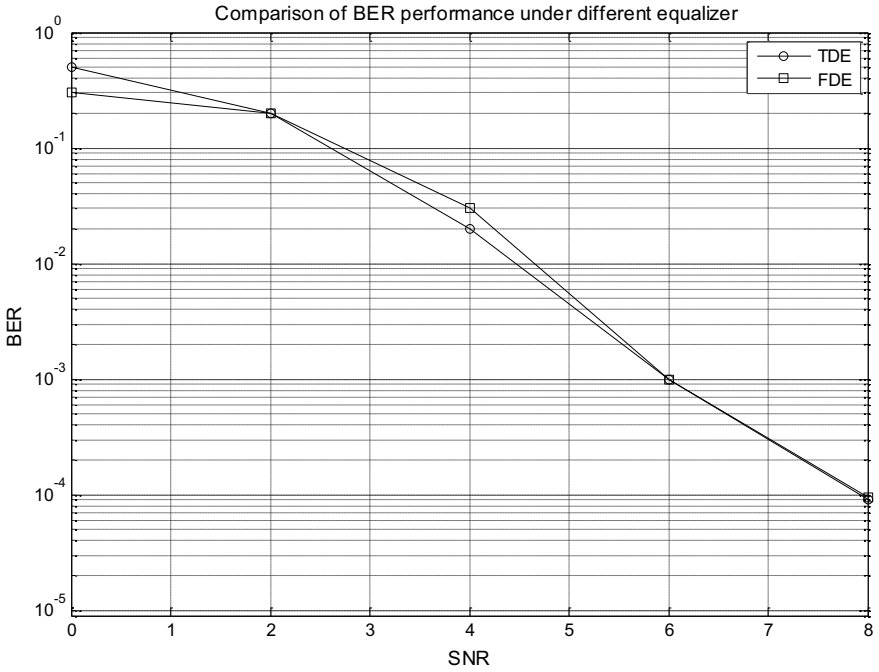



**Fig. 3.** Comparison of BER performance under different equalizer

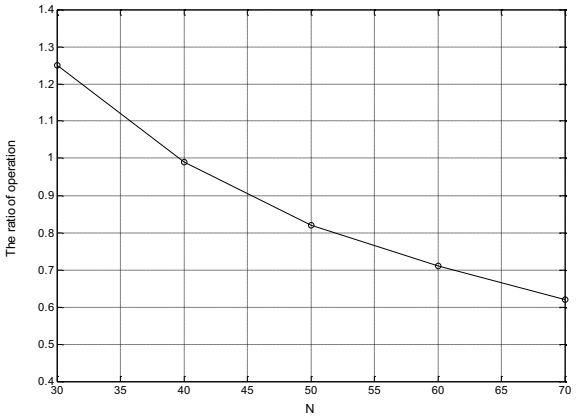



**Fig. 4.** Ratio of curve of the computational complexity

LT encoding symbols are transmitted at different SNRs, and the minimum number of LT encoding symbols required with an error rate of less than $10^{-5}$ is recorded at each SNR.

In order to be more intuitive, the redundancy (number of transmitted bits/effective number of bits) is used in the simulation to measure the performance. The minimum coding redundancy comparison is shown in Fig. 5. It can be seen that the rateless coding scheme in FDE can achieve reliable transmission of information by sending fewer codes compared to the traditional retransmission method in the underwater acoustic channel. The redundancy required by the proposed method is about 1/3 of the classical time-domain retransmission method, which improves the transmission efficiency of information.

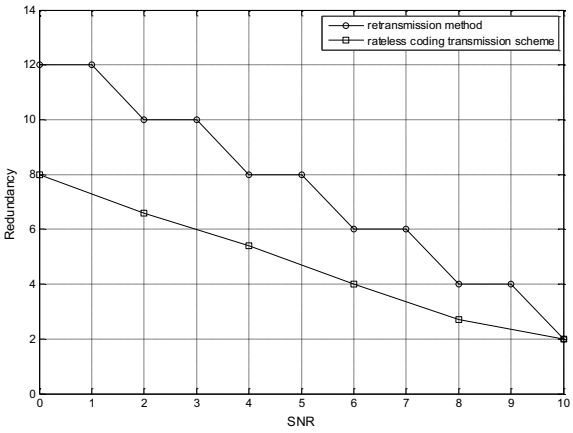



**Fig. 5.** Performance of coding redundancy

## 2 Conclusion

Based on the status of the research on underwater acoustic communication, we propose an LT coding adaptive transmission scheme based on SC-FDE. Simulation results show that the proposed method has greater performance advantages than the traditional time-domain equalization system and retransmission method and is more suitable for underwater channels with the characteristics of channel parameters varying with time and space and multipath delay spread.

## References

1. Youcef A, Laot C, Amis K. Adaptive frequency-domain equalization for underwater acoustic communications. In: OCEANS. Santander: IEEE; 2011. p. 1–6.
2. Akyildiz IF, Pompili D, Melodia T. Underwater acoustic sensor networks: research challenges. Ad Hoc Netw. 2005;3:257–79.

3. Byers JW, Luby M, Mitzenmacher M, et al. A digital fountain approach to reliable distribution of bulk data. In: Proceedings of the ACM SIGCOMM 98 conference. New York: ACM; 1998. p. 56–67.
4. Luby M. LT codes. In: Proceedings of 43rd foundations of computer science. Vancouver: IEEE; 2002. p. 271–82.
5. Casari P, Rossi M, Zorzi M. Towards optimal broadcasting policies for HARQ based on fountain codes in underwater networks. In: Proceedings of wireless on demand network systems and services. Garmisch-Partenkirchen: IEEE; 2008. p. 11–9.
6. Tai Y, Wang H. A novel LT-Turbo equalization for long-range deep-water acoustic communication. Sci Sin-Phys Mech Astron. 2016;46(9):094313.
7. Kang Y, Kim K, Park H. Efficient DFT-based channel estimation for OFDM systems on multipath channels. IET Commun. 2007;1(2):197–202.
8. Yuan Z, Songtao X. Application of new LMS adaptive filtering algorithm with variable step size in adaptive echo cancellation. In: International conference on communication technology (ICCT). IEEE; 2017. p. 1715–19.

# The Impact of Rotary Joint on Deviations of Amplitude and Phase and Its Calibration for Dual-Polarization Weather Radar

Shao Nan[1], Han Xu[1], Bu Zhichao[1(✉)], Chen Yubao[1], Pan Xinmin[2], and Qin Jianfeng[3]

[1] CMA Meteorological Observation Centre, Beijing 100086, China
buzhichao@163.com
[2] Henan Meteorological Observation Data Center, Zhengzhou 450003, China
[3] Wuhan Meteorological Agency, Wuhan 430040, China

**Abstract.** The variation of rotary joint is one of the primary reasons that result in dynamic deviation of $Z_{DR}$ and $\Phi_{DP}$ as polarimetric parameters, and external instruments are needed to detect and calibrate to ensure the amplitude and phase consistency for dual-polarization weather radar. This paper analyzes the impact of rotary joint of dual-polarization weather radar on $Z_{DR}$ and $\Phi_{DP}$, and it proposes a detection and calibration method using external instrument based on baseline curve of deviation. Then this method is used to test and calibrate a S-band dual-polarization weather radar produced by Beijing Metstar Radar Co., Ltd., and the results are analyzed. It is shown that $Z_{DR}$ and $\Phi_{DP}$ deviation introduced by rotary joint can satisfy requirement of relevant technical specifications, and this method can reduce deviation of amplitude and phase through calibration to enhance the reliability of radar observation data effectively.

**Keywords:** Rotary joint · $Z_{DR}$ · $\Phi_{DP}$ · Detection · Calibration

## 1 Introduction

Dual-polarization weather radar transmits horizontal and vertical polarized wave alternately or simultaneously and adopts various signal processing methods to echo signals in two polarized direction to get several polarimetric parameters such as differential reflectivity ($Z_{DR}$), differential propagation phase ($\Phi_{DP}$), linear depolarization ratio (*LDR*), correlation coefficient ($\rho_{HV}$) [1, 2]. These polarimetric parameters reflect features of hydrometeor particles in the air such as size, shape, phase, and orientation, which promotes the development of research area such as hydrometeor classification and quantitative precipitation estimation [3]. Compared with routine Doppler weather radar, dual-polarization weather radar enhances the ability of quantitative precipitation estimation and hydrometeor classification, and it is also an important tool in fields such as weather modification, aviation warning, and disaster monitoring [4–8]. However, the quality of dual-polarization radar data is affected by different factors such as beam blockage, system internal and environmental noises, and signal attenuation. It is necessary to evaluate the data quality before application [9].


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_99

Polarimetric variables such as $Z_{DR}$ and $\Phi_{DP}$ are primary parameters for reflecting targeted features of the cloud and rain, and it has wide application in fields such as quantitative precipitation estimation and hydrometeor classification. The accuracy of $Z_{DR}$ and $\Phi_{DP}$ is greatly significant to dual-polarization weather radar. For instance, when the consistency deviation of dual channel for $Z_{DR}$ is controlled within 0.2 dB, the error of quantitative precipitation estimation is controlled within 18%. When the consistency deviation of $Z_{DR}$ is within $\pm 0.1$ dB, the deviation of light precipitation measurement is also controlled between 10 and 15%. At the same time, only when the deviation of $\Phi_{DP}$ is controlled within 3°, the data have higher reliability [10, 11].

In recent years, many routine Doppler weather radars in China, European and USA have been upgraded to dual-polarization weather radar, and some scholars have made numerous studies to the deviation calibration of dual-polarization data. In 1999, Gorgucci, et al. adopted light rain method to calibrate system deviation of $Z_{DR}$ and $\Phi_{DP}$ caused by the whole channel including transmitter, antenna and receiver, but this method must be applied under the condition of stratiform precipitation with very little rain intensity [12]. In 2002, Hubbert, et al. suggested a method of deviation calibration to $Z_{DR}$ with natural scattering echoes, and the essence of the method is calibrating receiving channel of dual-polarization radar with echo data produced by solar radiation [13]. At the same time, it can also calibrate transmitting channel with cross-polar measurement. But as a signal source, the signal-to-noise ratio of the sun is low, and the calibration result of $Z_{DR}$ and $\Phi_{DP}$ is not satisfactory, and it can only be applied in a limited period of time every day [14].

In 1996, Scarchilli, et al. suggested that $Z$, $Z_{DR}$, and $\Phi_{DP}$ have strong self-consistency in precipitation, and any one of these polarization parameters can be represented by the other two polarization parameters [15]. That is to say, one parameter can be calibrated by the other two polarization parameters. However, in 2003, Vivekanandan, et al. indicated that the condition of using this calibration method is that the deviation of two polarization parameters must be controlled within 0.5–1 dB [16].

After dual-polarization upgrade of WSR-88D weather radar, the front end of the receiver moves up. RF signal is down-converted to IF, through rotary joint, it enters digital intermediate frequency receiver in machine room for A/D conversion, and then it is sent into RVP processor. The advantage of the method is that the effect cable and joint on IF signal transmission is small. The amplitude consistency of antenna, feeder, and receiver is calibrated with solar testing method periodically (every month). The deviation of dual-channel of the receiver (RCB) is calibrated online through internal signal between two consecutive volume scans, so that the consistency of the dual-channel can be guaranteed better. In the WSR-98D and CINRAD/SA dual-polarization after being upgraded in China, the receiver is located in the machine room, and this configuration can reduce the disturbance of environment temperature to radar signal; thus, it is helpful to keep the consistency of the amplitude and phase of dual channel of the receiver, but the deviation of the amplitude and phase consistency for dual channels will be affected by rotary joint.

This paper proposes a detection and calibration method using external instrument based on baseline deviation curve and applies the method in a S-band dual-polarization weather radar produced by Beijing Metstar Radar Co., Ltd. to test, calibrate and analyze the deviation caused by rotary joint. It is shown that $Z_{DR}$ and $\Phi_{DP}$ data accuracy after being calibrated are improved obviously.

## 2 Analysis on Systematic Deviation of $Z_{DR}$ and $\Phi_{DP}$

### 2.1 Analysis of $Z_{DR}$ and Its Systematic Deviation

Dual-polarization weather radar transmits horizontal and vertical polarized wave and obtains horizontal and vertical reflectivity factors $Z_H$ and $Z_V$, respectively, while different reflectivity factor $Z_{DR}$ can be expressed as:

$$Z_{DR} = 10 \cdot \log\left(\frac{Z_H}{Z_V}\right) \tag{1}$$

It can be known from the Eq. (1) that $Z_{DR}$ is related to the size and axis ratio of hydrometeor particle (axis ratio is defined as $a/b$, and $a$ is the radius of horizontal axis, and $b$ is the radius of vertical axis of particle). It reflects physical information of hydrometeor particle such as the size and shape [17, 18]. At the same time, $Z_{DR}$ is obtained from horizontal and vertical polarized wave signal, and the data precision is closely related to dual-channel consistency of dual-polarization weather radar.

Data precision of polarimetric variable $Z_{DR}$ is not only related to noise inside radar system and interference between signals, but also related to hardware of radar system, so it has significance in calibrating deviation produced by radar hardware [19]. An important function of dual-polarized weather radar is enhancing the capacity of quantitative precipitation estimation. Currently, the most common formula of quantitative precipitation estimation from dual-polarized radar is proposed by NOAA National Severe Storms Laboratory, and it is applied in JPOLE observation experiment in USA [11]. The algorithm is:

$$R = \begin{cases} R(Z)/f_1(Z_{DR}) & R(Z) < 6 \text{ mmh}^{-1} \\ R(K_{DP})/f_2(Z_{DR}) & 6 < R(Z) < 50 \text{ mmh}^{-1} \\ R{=}R(K_{DP}) & R(Z) > 50 \text{ mmh}^{-1} \end{cases} \tag{2}$$

$$R(Z) = 1.7010^{-2} Z^{0.174} \tag{3}$$

$$R(K_{DP}) = 44.0|K_{DP}|^{0.822}\text{sign}(K_{DP}) \tag{4}$$

$$f_1(Z_{DR}) = 0.4 + 5.0|Z_{dr} - 1|^{1.3} \tag{5}$$

$$f_2(Z_{DR}) = 0.4 + 3.5|Z_{dr} - 1|^{1.7} \tag{6}$$

It can be seen from the above equations that rainfall rate is calculated only through reflectivity factor Z in case of light rain, while $Z_{DR}$ and $K_{DP}$ are used to calculated rainfall rate in case of small to moderate rain, and only $K_{DP}$ is used to estimate rainfall rate in heavy rain.

Rainfall rate and precipitation accumulated estimate deviation brought by $Z$ and $Z_{DR}$ can be evaluated from the Eq. (2) by the differential algorithm, and it can be written as:

$$\varepsilon_1 = \alpha \Delta Z^{(b)} - \beta \Delta Z_{DR}^{(b)} \tag{7}$$

$$\varepsilon_2 = -\beta_2 \Delta Z_{DR}^{(b)} \tag{8}$$

$$\alpha = \frac{1}{R(Z)} \frac{dR(Z)}{dZ} \tag{9}$$

$$\beta_{1,2} = \frac{1}{f_{1,2}(Z_{DR})} \frac{df_{1,2}(Z_{DR})}{dZ_{DR}} \tag{10}$$

Evaluating deviation of rainfall rate can be divided into two parts, including systematic deviation and standard deviation, which means:

$$\delta^2 = \varepsilon^2 + \sigma^2 \tag{11}$$

The value of standard deviation is mainly decided by factors such as the variation rate of disdrometer and statistic error of radar variable, and the value is obtained under the condition that $Z$ and $Z_{DR}$ are calibrated completely.

For JOPLE data system, if the rainfall rate is lower than 5 mm, the standard deviation is 68%; if the rainfall rate is larger than 5 mm but lower than 30 mm, the standard deviation is 38%; if the rainfall rate is larger than 30 mm, the standard deviation is 22%. It can be deducted from Eq. (7) that when the deviation of $Z_{DR}$ is within ±0.2 dB, the estimate deviation of quantitative precipitation estimation can be controlled below 18%. When the deviation of $Z_{DR}$ is within ±0.1 dB, the deviation of precipitation measurement of light rain can be controlled within 10–15%.

### 2.2 Analysis of $\Phi_{DP}$ and Its Systematic Deviation

Differential propagation phase is defined as

$$\Phi_{DP} = \Phi_{HH} - \Phi_{VV} \tag{12}$$

In the formula, $\Phi_{HH}$ and $\Phi_W$ are the two-way phase angle in a certain distance in a round trip to antenna when radar signal is in horizontal and vertical polarization. In nature, they are the phase difference caused by different propagation phase constant of horizontal and vertical polarized wave in different precipitation regions. In a real observation process, meteorological target environment and its own disturbance as well

as instability in radar system will impact $\Phi_{DP}$ [20]. Most deviations of $\Phi_{DP}$ caused by antenna and feeder are fixed which can be measured through external instruments regularly.

Rotary joint links static and dynamic waveguide end face through axis to realize flexible turning and microwave signal transmission. The coaxiality of static and dynamic waveguide end face has a large impact on conductive property of microwave. In operation, especially with the increasing of operating time and wearing of mechanical bearings, coaxiality of two-end faces change dynamically in one cyclic period, which causes that deviations of $\Phi_{DP}$ ($Z_{DR}$ deviation) caused by rotary joint are dynamically changing and uncertain. It needs to adopt external instrument to detect and evaluate for long term to ensure the reliability of observation data of dual-polarization weather radar.

## 3 Detection Analysis and Calibration Method of the Impact of Rotary Joint Amplitude and Phase Consistency

S-band dual-polarization weather radar produced by Beijing Metstar Radar Co., Ltd. adopts "simultaneous horizontal and vertical transmission and reception" mode (SHV) that is transmitting and receiving horizontal and vertical polarized wave simultaneously. This mode avoids the problem that high-power ferrite microwave switch is affected by environment. Under the condition of guaranteeing the same data precision, it can enhance the rotating speed of antenna, which is favorable for observation of the weather process that is changing rapidly, and the correlation of observation target is good. But there is cross-coupling between horizontal and vertical channels, so it requires that isolation of dual channels has to be better than −40 dB [21].

The dual-channel receivers of this radar are located inside the machine room, so it is impossible to detect the influence of the antenna and feeder, especially the double-channel rotary joint on the amplitude and phase consistency by using the internal signal. Therefore, it must detect and calibrate through external instrument. There are two kinds of detection and analysis methods for the impact of the rotary joint to of the amplitude and phase consistency: one is receiver chain method, and the other is whole transmitting and receiving chain method. The principles of these two methods are the same: which is to adopt continuous wave baseline testing signal of stable amplitude and phase generated by high-performance radio frequency signal source, injecting into receiving feeder, or whole transmitting and receiving chain, then select specific scanning mode to control the moving or static of azimuth rotary joint and elevation rotary joint while signal passes, and acquire and process the output signal of receiver, whose amplitude and phase is compared with baseline testing signal, to analyze the dynamic change of amplitude and phase. In the chain of transmitting and receiving feeders, other feeders or devices can seldom introduce a dynamic change of amplitude and phase except for rotary joint, so dynamic change of amplitude and phase of receiving feeder chain or whole chain is the effect of the rotary joint.

### 3.1 Receiving Chain Detection and Analysis Method

The method is to inject testing signal above antenna elevation rotary joint with external instrument, through attenuator, power divider, and cross-directional coupler to dual-channel waveguide (the signal-to-noise ratio entering LNA is about 67 dB) which is to test the impact of the rotary joint on $Z_{DR}$ and $\Phi_{DP}$ with the change of azimuth and elevation. Without high voltage on transmitter, VCP21D (apply elevation of 0.5° and azimuth of 0–360° data in analyzing) and RHI (azimuth of 0° and elevation of 0–20°) scanning mode are operated in RCW. These two scanning modes save base data separately. Then the changes of $Z_{DR}$ and $\Phi_{DP}$ are calculated according to dual-polarization parameter in different azimuths (different elevations for RHI) to test the impact of azimuth and elevation rotary joint on the amplitude and phase consistency and make an analysis of the data before and after calibration. The principle of the detection and analysis method is shown in Fig. 1.

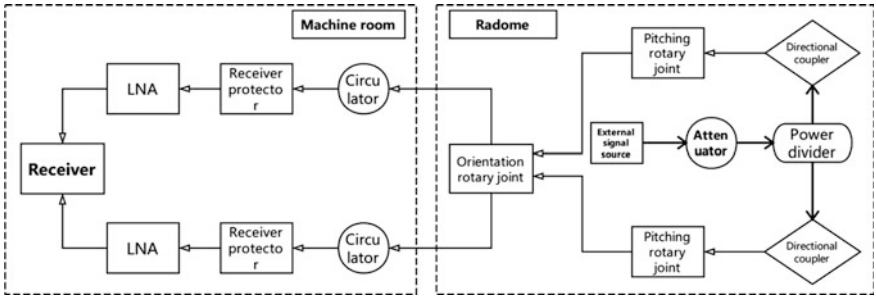



**Fig. 1.** Principle chart of receiving chain method for testing the impact of rotary joint on the amplitude and phase consistency

### 3.2 Whole-Chain Detection and Analysis Method

The soft waveguide at the upper end of elevation rotary joint is disconnected, instead it is connected with waveguide coaxial conversion and short circuit, then a continuous wave testing signal from external signal source is injected from the reverse port of the upper coupler 1DC1 above the transmitter. The test signal is reflected through the transmitting branch by the waveguide coaxial converter at the elevation rotary joint port and the reflected signal is injected into receiving channel, which can simulate the whole chain of transmitting and receiving branch (without antenna and radome). At reverse port of directional coupler 1DC1, a signal with 0 dBm is injected (signal-to-noise ratio entering LNA is about 70 dB). Without high voltage on transmitter, VCP21D (applies elevation 0.5° and azimuth 0–360° in analysis) and RHI (azimuth 0° and elevation 0–20°) scanning mode are operated, and these two scanning modes store base data separately. Then the changes of $Z_{DR}$ and $\Phi_{DP}$ are calculated according to dual-polarization parameter in different azimuths (different elevations for RHI), to test

the impact of the rotary joint of under whole-chain working condition on amplitude and phase consistency, and analyze data before and after the calibration. In addition, at reverse port of directional coupler 1DC1, a continuous wave signal with 0 dBm is injected (signal to noise ratio entering LNA is about 70 dB) to test the impact of the rotary joint on the amplitude and phase consistency for 24 h. The principle of testing and analyzing method is as shown in Fig. 2.

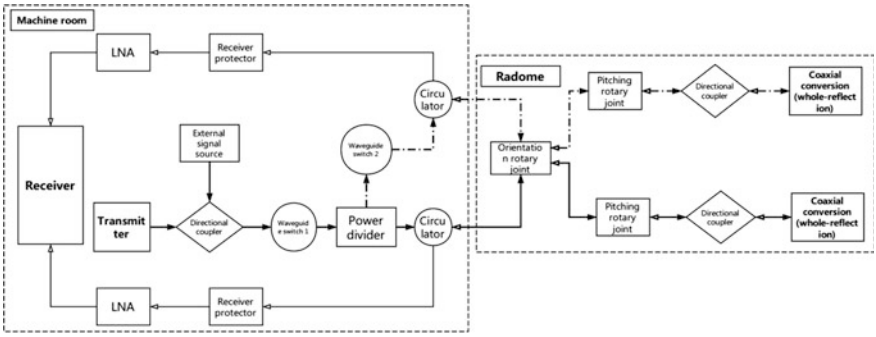



**Fig. 2.** Principle chart of whole-chain method for testing the impact of rotary joint on the amplitude and phase consistency

### 3.3 Amplitude and Phase Calibration Method

External signal source is used to detect the impact of rotary joint on $Z_{DR}$ and $\Phi_{DP}$ with the change of azimuth (PPI scanning) and elevation (RHI scanning). Baseline deviation curves of $Z_{DR}$ and $\Phi_{DP}$ are obtained according to the testing result, and the actual data of $Z_{DR}$ and $\Phi_{DP}$ are calibrated with baseline deviation curve. The specific method is to inject baseline testing signal and then control radar to operate in VCP21D scanning mode, collect and store base data. The stored base data are read; the amplitude and phase deviation of the fixed range bin are obtained according to azimuth and elevation and stored as baseline deviation curve of $Z_{DR}$ and $\Phi_{DP.}$ When radar operates normally, baseline deviation is subtracted from observation data, to eliminate systematic deviation, and to realize calibration of $Z_{DR}$ and $\Phi_{DP.}$ Baseline deviation curve is generated at calibration initialization. In addition, it also needs to consider the quality control of radar, and abnormal data processing, etc.

Figure 3 is the curve of baseline deviation of differential reflectivity ($\Delta Z_{n-m}$) and differential propagation phase ($\Delta\varphi_{n-m}$). PPI in each elevation layer in radar volume scan corresponds to a curve of baseline deviation. The correction formula of differential reflectivity and differential propagation phase is:

$$Z\prime_{DR(n-m)} = Z_{DR(n-m)} - \Delta Z_{n-m} \tag{13}$$

$$\varphi\prime_{DP(n-m)} = \varphi_{DP(n-m)} - \Delta\varphi_{n-m} \tag{14}$$

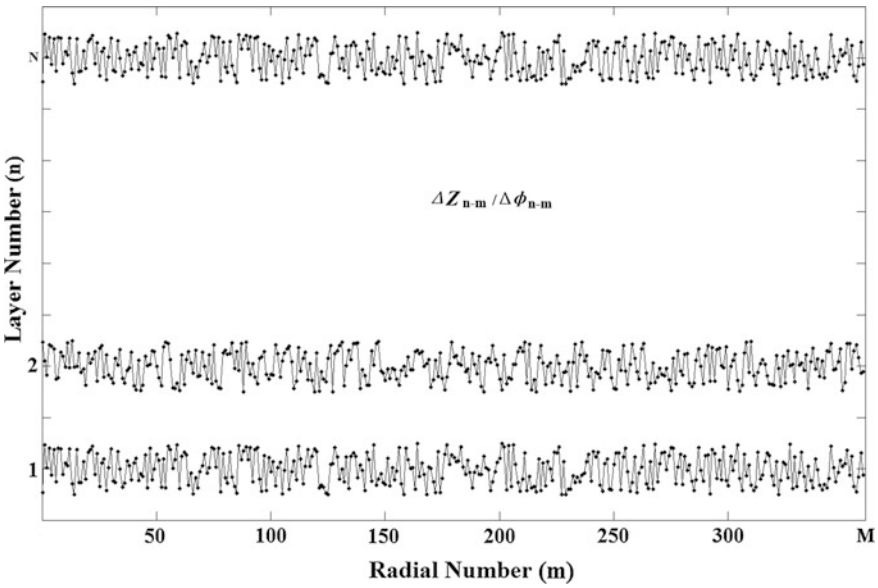



**Fig. 3.** Baseline deviation curves of differential reflectivity ($\Delta Z_{n-m}$) and differential propagation phase ($\Delta\varphi_{n-m}$)

In the formula, $Z'_{DR(n-m)}$, $\varphi'_{DP(n-m)}$ are calibrated $Z_{DR}$ and $\Phi_{DP}$, $Z_{DR(n-m)}$ and $\varphi_{DP(n-m)}$ are uncalibrated $Z_{DR}$ and $\Phi_{DP}$, and $\Delta Z_{n-m}$ and $\Delta\varphi_{n-m}$ are baseline deviation respectively. $N$ is the number of elevation layer, number is indexed from low-elevation angle to high-elevation angle starting from 1, that is, $n = 1, 2, \ldots, N$, the value $n$ is different in various scanning modes. $M$ is the number of azimuth index in different radial direction of every PPI. The 0° azimuth radial is numbered as 1 and number is indexed from low angle to high angle along the PPI circle, that is $m = 1, 2, \ldots, M$. The value of $m$ is different for different beam widths.

## 4 Impact of Rotary Joint on Amplitude and Phase Consistency and Calibration Test

### 4.1 Test and Calibration of Receiving Chain Method

As shown in Fig. 1, the method injects external signal with 0 dBm (signal-to-noise ratio entering LNA is about 67 dB) to dual-channel waveguide to test the change of $Z_{DR}$ and $\Phi_{DP}$ caused by rotary joint.

### 4.1.1 Change of $Z_{DR}$ and $\Phi_{DP}$ Caused by Azimuth Rotary Joint and Calibration

The change of $Z_{DR}$ and $\Phi_{DP}$ with azimuth before and after calibration is as shown in Fig. 4. Before the calibration, the maximum changing range of $Z_{DR}$ is 0.07 dB, and the standard deviation is 0.0108 dB; the maximum changing range of $\Phi_{DP}$ is 2.58°, and the standard deviation is 0.7254°. After the calibration, the maximum changing range of $Z_{DR}$ is 0.04 dB, and the standard deviation is 0.00464 dB; the maximum changing range of $\Phi_{DP}$ is 0.18°, and the standard deviation is 0.0187°.

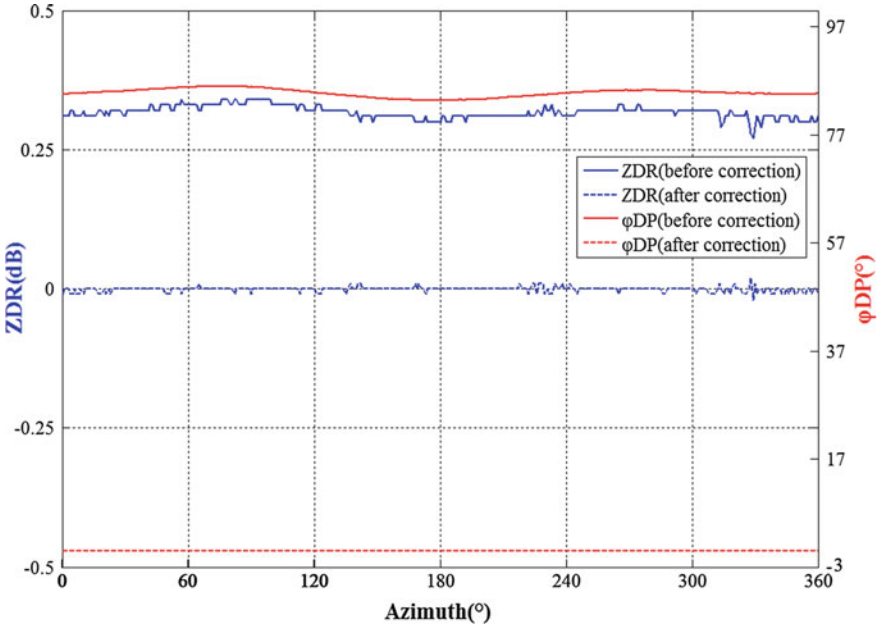



**Fig. 4.** Change of $Z_{DR}$ and $\Phi_{DP}$ with azimuth before and after calibration

### 4.1.2 Change of $Z_{DR}$ and $\Phi_{DP}$ Caused by Elevation Rotary Joint

The change of $Z_{DR}$ and ΦDP with elevation before and after calibration is as shown in Fig. 5. Before calibration, the maximum changing range of $Z_{DR}$ is 0.02 dB, and the standard deviation is 0.0051 dB; the maximum changing range of $\Phi_{DP}$ is 0.12°, and the standard deviation is 0.0266°. After the calibration, the maximum changing range of $Z_{DR}$ is 0.02 dB, and the standard deviation is 0.0025 dB; the maximum changing range of $\Phi_{DP}$ is 0.04°, and the standard deviation is 0.0145°.

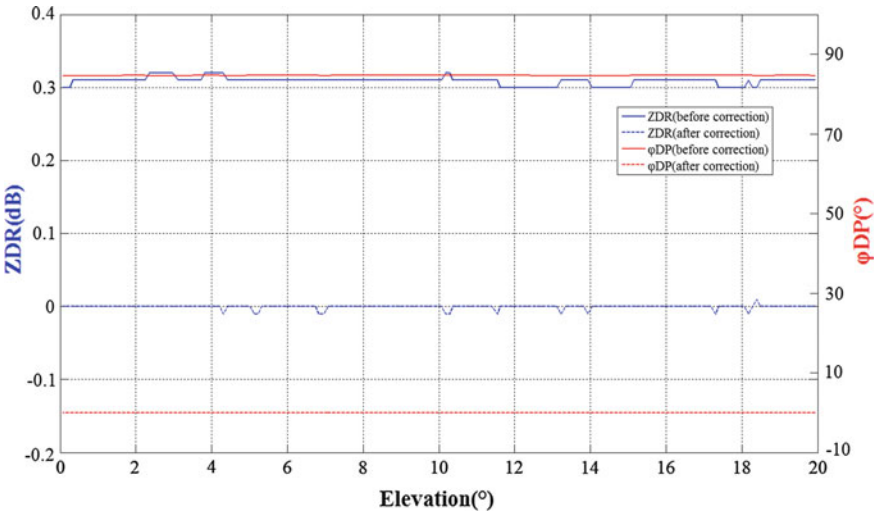



**Fig. 5.** Change of $Z_{DR}$ and $\Phi_{DP}$ with elevation before and after calibration

### 4.2 Whole-Chain Method Detection and Calibration

As shown in Fig. 2, a testing signal with 0 dBm is injected into the reverse power measurement port 1DC1 above the transmitter (the coupling ratio is about −36 dB) (signal-to-noise ratio entering LNA is about 70 dB), to detect the effect of the rotary joint on $Z_{DR}$ and $\Phi_{DP}$.

#### 4.2.1 Change of $Z_{DR}$ and $\Phi_{DP}$ Caused by Azimuth Rotary Joint and Calibration

The change of $Z_{DR}$ and $\Phi_{DP}$ with azimuth and angle before and after calibration is as shown in Fig. 6. Before the calibration, the maximum changing range of $Z_{DR}$ is 0.02 dB, and the standard deviation is 0.0071 dB; the maximum changing range of $\Phi_{DP}$ is 2.62°, and the standard deviation is 0.8656°. After the calibration, the maximum changing range of $Z_{DR}$ is 0.02 dB, and the standard deviation is 0.0011 dB; the maximum changing range of $Z_{DR}$ is 0.06°, and the standard deviation is 0.0121°.

#### 4.2.2 Change of $Z_{DR}$ and $\Phi_{DP}$ Caused by Elevation Rotary Joint

The change of $Z_{DR}$ and $\Phi_{DP}$ with elevation before and after calibration is as shown in Fig. 7. Before the calibration, the maximum changing range of $Z_{DR}$ is 0.01 dB, and the standard deviation is 0.0047 dB; maximum changing range of $\Phi_{DP}$ is 0.06°, and the standard deviation is 0.0165°. After calibration, the maximum changing range of $Z_{DR}$ is 0.01 dB, and the standard deviation is 0.0047 dB; the maximum changing range of $Z_{DR}$ is 0.04°, and the standard deviation is 0.0108°.

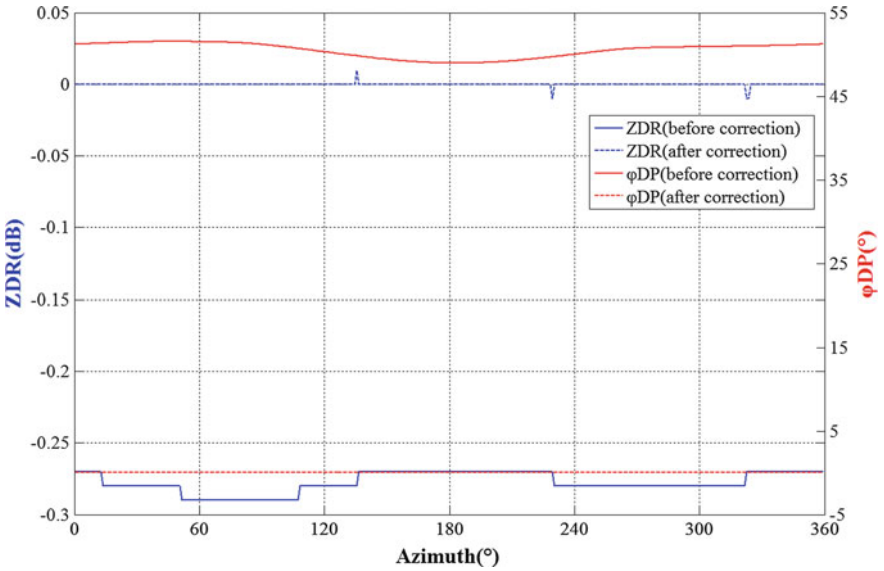



**Fig. 6.** Change of $Z_{DR}$ and $\Phi_{DP}$ with azimuth before and after calibration

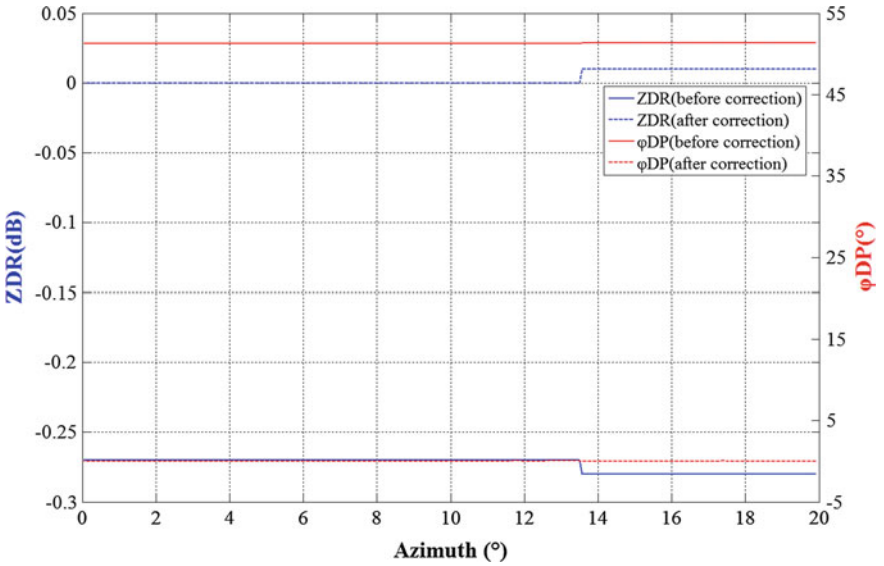



**Fig. 7.** Curve of the change of $Z_{DR}$ and $\Phi_{DP}$ with elevation before and after calibration

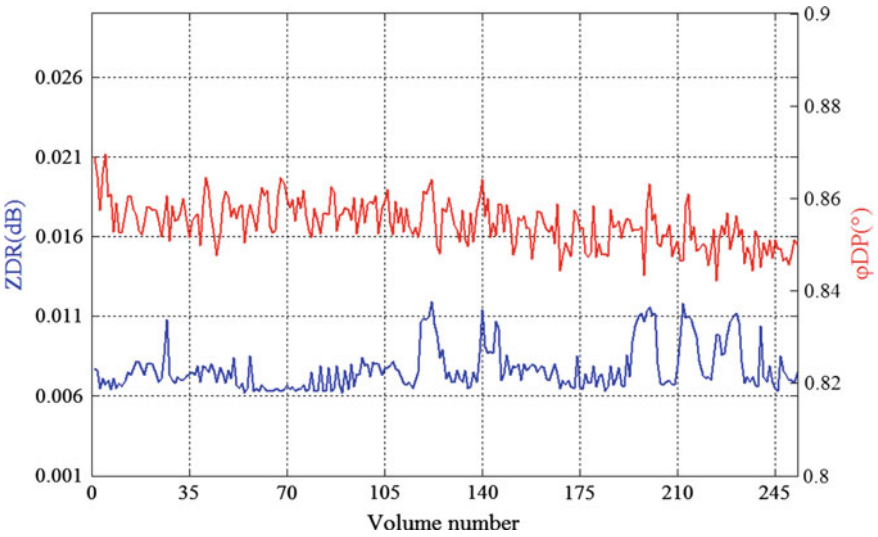



**Fig. 8.** Curve of the change of $Z_{DR}$ and $\Phi_{DP}$ within 24 h

## 5 Stability Analysis of Amplitude and Phase Consistency

The result shows that it can greatly reduce the deviation of amplitude and phase brought by rotary joint through calibration, and it can realize better correction to deviation of amplitude and phase caused by rotary joint in short term. After calibration, the effect of rotary joint can be neglected, but the stability of amplitude and phase consistency still needs to be tested for 24 h, and stability of standard deviation of $Z_{DR}$ and $\Phi_{DP}$ shall be analyzed through base data. Figure 8 is the testing result of $Z_{DR}$ and $\Phi_{DP}$ when radar operates for 24 h continuously, in which, the maximum changing range of the standard deviation of $Z_{DR}$ is 0.0057 dB, while the maximum changing range of the standard deviation of $\Phi_{DP}$ is 0.0274°.

## 6 Conclusion

This paper analyzes the impact of the rotary joint on the amplitude and phase consistency of dual-polarization weather radar, and it conducts detection and calibration to the deviation of $Z_{DR}$ and $\Phi_{DP}$ introduced by rotary joint through adopting receiving chain method and all chain method. The result shows that:

(a) Deviation of amplitude and phase introduced by azimuth and elevation rotary joints can satisfy requirements of no more than 0.2 dB and 3°, respectively.
(b) The impact of azimuth and elevation rotary joints on amplitude and phase has periodical characteristics.
(c) After calibration, the impact of azimuth and elevation rotary joints on amplitude and phase is very small, which can be neglected.

(d) Continuous testing result for 24 h indicates that the stability of amplitude and phase is good.
(e) Dual-polarization weather radar hasn't realized online detection of rotary joint with internal calibration signal. It is suggested to add the function to detect the operating status of rotary joint real time to avoid polarimetric data quality degradation caused by wearing of rotary joints due to long-term operation.

## References


1. Doviak RJ, Zrnic DS, Schotland RM. Doppler radar and weather observations. Appl Opt. 1994;33(21):4531.
2. Laroche S. Polarimetric doppler weather radar: principles and applications: V.N. Bringi and V. Chandrasekar. Cambridge University Press, 2001, 636 pp. Atmos Res. 2002;63(1–2):159–160.
3. Wang H, Ran Y, Deng Y, et al. Study on deep-learning-based identification of hydrometeors observed by dual polarization doppler weather radars. Eurasip J Wirel Commun Networking. 2017;2017(1):173.
4. Galati G, Pavan G. Estimation techniques for rainfall rate using differential phase shift in X-band weather radar. In: Geoscience and remote sensing symposium proceedings, 1998. IGARSS'98. 1998 IEEE International. IEEE; 1998, vol. 1. p. 138–40.
5. Liu H, Chandrasekar V. An adaptive neural network scheme for precipitation estimation from radar observations. In: Geoscience and remote sensing symposium proceedings, 1998. IGARSS '98. 1998 IEEE International. IEEE; 1998, vol. 4. p. 1895–7.
6. Schuur T, Ryzhkov A, Heinselman P, et al. Observations and classification of echoes with the polarimetric WSR-88D radar; 2003.
7. Bringi VN, Thurai M, Hannesen R. Dual-polarization weather radar handbook. AMS-Gematronik GmbH; 2007.
8. Liu H, Chandrasekar V. Classification of hydrometeors based on polarimetric radar measurements: development of fuzzy logic and neuro-fuzzy systems, and in situ verification. J Atmos Oceanic Technol. 2000;17(2):140–64.
9. Marks DA, Wolff DB, Carey LD, et al. Quality control and calibration of the dual-polarization radar at Kwajalein, RMI. J Atmos Oceanic Technol. 2011;28(28):181–96.
10. Kwon S, Lee GW, Kim G. Rainfall estimation from an operational S-band dual-polarization radar: effect of radar calibration. J Meteorol Soc Jpn. 2015;93(1):65–79.
11. Ryzhkov AV, Giangrande SE, Melnikov VM, et al. Calibration issues of dual-polarization radar measurements. J Atmos Oceanic Technol. 2005;22(8):1138–55.
12. Gorgucci E, Scarchilli G, Chandrasekar V. A procedure to calibrate multiparameter weather radar using properties of the rain medium. IEEE Trans Geosci Remote Sens. 1999;37(1):269–76.
13. Hubbert JC, Bringi VN, Brunkow D. Studies of the polarimetric covariance matrix. Part I: calibration methodology. J Atmos Oceanic Technol. 2003;20(5):696–706.
14. Williams ER, Cho JYN, Smalley DJ, et al. End-to-end calibration of NEXRAD differential reflectivity with metal spheres. In: Conference on radar meteorology. American Meteorological Society; 2013.
15. Scarchilli G, Gorgucci E, Chandrasekar V, et al. Self-consistency of polarization diversity measurement of rainfall. IEEE Trans Geosci Remote Sens. 1996;34(1):22–6.

16. Vivekanandan J, Zhang G, Ellis SM, et al. Radar reflectivity calibration using differential propagation phase measurement. Radio Sci. 2016;38(3):14-1-14-14.
17. Lim S, Chandrasekar V, Bringi VN. Hydrometeor classification system using dual-polarization radar measurements: model improvements and in situ verification. IEEE Trans Geosci Remote Sens. 2005;43(4):792–801.
18. Baldini L, Gorgucci E. Identification of the melting layer through dual-polarization radar measurements at vertical incidence. J Atmos Oceanic Technol. 2006;23(6):829–39.
19. Zrnić D, Doviak R, Zhang GF, et al. Bias in differential reflectivity due to cross coupling through the radiation patterns of polarimetric weather radars. J Atmos Oceanic Technol. 2009;27(10):1624–37.
20. Melnikov VM, Zrnic DS, et al. ZDR calibration issues in the WSR-88Ds. Report on 2013-MOU. p. 1–55.
21. Wang Y, Chandrasekar V. Polarization isolation requirements for linear dual-polarization weather Radar in simultaneous transmission mode of operation. IEEE Trans Geosci Remote Sens. 2006;44(8):2019–28.

# Research on Performance of Chaotic Direct Sequence Spread Spectrum UWB System Based on Chaotic Matrix

Peize Li, Bing Zhao(✉), and Zhifang Wang

School of Electronic Engineering, Heilongjiang University, Harbin 150080, China
zb0624@163.com

**Abstract.** Pulse ultra-wideband technology uses narrow pulse as the information carrier, occupies a very wide frequency bandwidth, and has many characteristics such as high transmission rate, large capacity and strong multipath resistance. Chaotic signals are stochastic and sensitive to initial values, so they are widely used in secure communication systems. In this paper, the chaotic sequence of quadratic chaotic map is generated by matrix method. The new digital chaotic sequence after binary quantization is applied to the direct sequence spread spectrum UWB system. It reduces the influence of quantization on the randomness and ameliorates the dispersion of UWB signals power spectrum, making the interference with other systems lower. The simulation shows that the chaotic direct sequence spread spectrum UWB system based on chaotic matrix quadratic mapping has better system coexistence.

**Keywords:** Direct sequence spread spectrum UWB system · Chaotic matrix · Power spectral density

## 1 Introduction

Ultra-wideband wireless communication is a typical high-speed broadband access technology, with extremely low power spectral density hidden in the noise, and coexists with communication system in an overlapping and sharing manner. As a result, scarce spectrum resources improved. Literature [1] studies the performance of direct sequences. Chaotic sequences contain plenty of advantages, such as a large number of sequences, strong confidentiality, excellent correlation characteristics and easy to emerge, making the design of chaotic spreading codes in DS-UWB systems more and more important. Literature [2, 3] analyzes the multiple access performance of DS-UWB based on chaotic spread sequences. In order to enhance the multi-access performance, an intermediate extraction truncation digital binary chaotic sequence method is proposed and used in DS-UWB [4]. The literature proposes a composite sequence of double logistic map [5].


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_100

## 2 Chaotic Matrix

In this paper, the chaotic matrix mainly studies a quadratic map based on Logistics and Chebyshev. First, $N$ iterations of the logistics are performed to get a chaotic sequence of length $N \times 1$. The sequences of the parameters satisfy the full map condition. Take this column as an initial value of Chebyshev for $M$ iterations. The last iteration of the matrix $X$ is shown in Eq. (2.1). Each row is Chebyshev chaotic sequence, but each column is a sequence generated by Logistics and Chebyshev. Take the $N$th column of $X$ $m_n$ for research, which $m_n = (x_{1n}, x_{2n}, x_{3n}, \ldots, x_{mn})$.

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \tag{2.1}$$

## 3 Chaotic Matrix Direct Sequence UWB System

The information sequence is repeatedly coded by the repeat encoding module, and the repeatedly coded sequence is encoded by the column vector $m_n$ emerged by the chaotic matrix. What needs to be noted is that the second chaotic sequence $m_n$ is real-valued and needs to be binarized before being applied to the chaotic coding. After the information encoded, it is sent to the PAM modulation module and former module for transmission. The transmission model is shown in Fig. 1. The resulting signal can be expressed as Eq. (3.1):

$$s(t) = \sum_{j=-\infty}^{+\infty} a_j m_j p(t - jT_s) \tag{3.1}$$

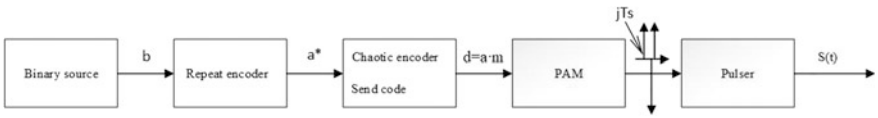



**Fig. 1.** DS chaotic matrix-UWB transmission model

## 4 Analysis of Simulation Results

### 4.1 Correlation Analysis

Before correlation performance analysis, binary quantization of the generated column vector is required. In the case of changing the parameter k, the Nth column vector of

the chaotic matrix $m_n$ is extracted and the correlation analysis is performed. The autocorrelation and cross-correlation functions are shown in Figs. 2 and 3. Then change the length of the sequence to record the balance as shown in Table 1. Due to its excellent autocorrelation and cross-correlation performance, the interference among users may be minimized by these chaotic sequences.

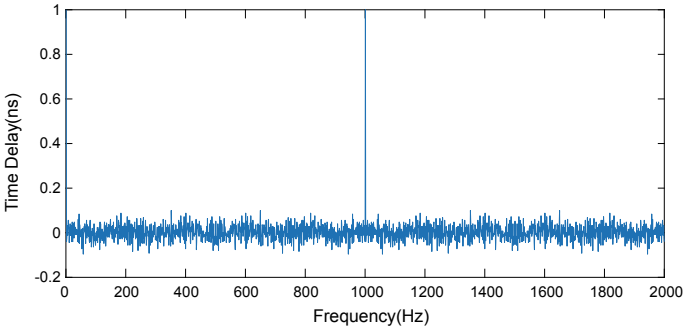



**Fig. 2.** Autocorrelation

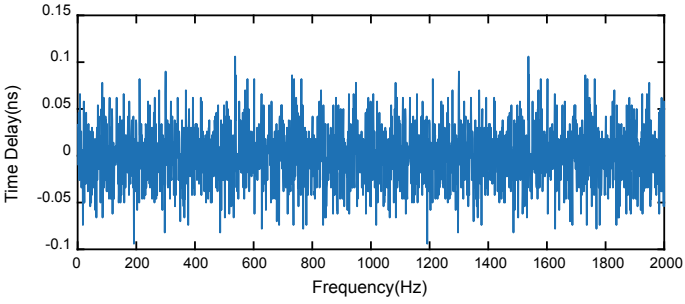



**Fig. 3.** Correlation

**Table 1.** Balance

| Length | Balance (%) |
|---|---|
| 2000 | 3.8 |
| 3000 | 0.367 |
| 4000 | 0.6 |
| 5000 | 1.72 |
| 6000 | 0.73 |
| 7000 | 0.014 |

Figure 2 shows that the value of the sequence except for the zero delay is basically 0 and is similar to an impulse function. In Fig. 3, the cross-correlation of the sequence is basically 0 when the delay is 0.5 ns, indicating that the sequence has better cross-correlation and autocorrelation. In Table 1, as the length of the sequence increases, the number of "0" and "1" of the sequence is gradually similar. When the sequence length is 7000, the number of "0" and "1" of the sequence is basically the same, and the balance performance is perfect.

### 4.2 Power Spectral Density Analysis

Set the parameter of sampling frequency 50 ns, average transmitted power −30 dBm, number of bits generated by the source 50 bits, pulse duration 2 ns, number of pulses per bit map 20, pulse duration $T_m$ 0.5 ns, the $\mu$, Logistics 4, the $k$ of Chebyshev 4, shaping factor for the pulse 0.25 ns, and the initial value of the Logistics sequence 0.8400.

Figure 4 shows the power spectral density of the DS-UWB transmitted signal for the second-mapped digital chaotic sequence. The envelopes of the graph have the shape of a basic pulse Fourier transform and the peaks of the graphs in Fig. 4 are low, and the spectral lines are concentrated, which also prove the concealment better.

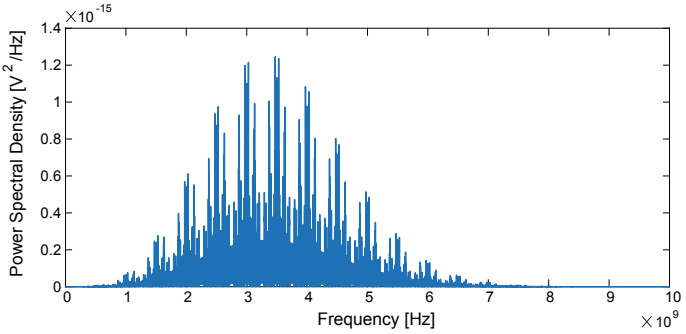



**Fig. 4.** Power spectral density of chaotic matrix DS-UWB

## 5 Conclusion

In this paper, the experimental results show that the digital chaotic sequence after secondary map has good autocorrelation and cross-correlation, which can be applied to multiple users. It has better confidentiality, easier generation, lower power spectral density and more in line with the standards of UWB low power spectrum.

**Acknowledgments.** This work is supported by Heilongjiang University Talent Innovation Special Project of Heilongjiang Province Universities Basic Scientific Research Business (No. HDRCCX-201605).

## References

1. Foerster JR. The performance of a direct-sequence spread ultrawideband system in the presence of multipath, narrowband interference, and multiuser interference. In: 2002 IEEE conference on ultra wideband systems and technologies (IEEE Cat. No.02EX580), Baltimore, MD, USA; 2002. p. 87–91.
2. Shanmugam SK, Leung H. Efficient chaotic spreading codes for DS-UWB communication system. In: 2006 IEEE international conference on acoustics speech and signal processing proceedings, Toulouse, France; 2006 May.
3. Long X, Dongdong W, Guoping J. Multi-access performance of ultra-wideband communication system based on chaotic spreading sequences. Southeast Univ J (Natural Science). 2005;35(2):22–5.
4. Honglei W, Yinhui Y, He S. Research on power spectrum of direct sequence UWB signal based on chaotic code. Jilin Univ J. 2008;26(4):338–42.
5. Guoqiang N, Xiaobo H, Bo H, Weidong Z. Research and application of a new chaotic spread spectrum sequence. 2009;46(05):1341–6.

# Compressed Sensing-Based Energy-Efficient Routing Algorithm in Underwater Sensor Networks

Qiuming Zhao, Hongjuan Yang, Bo Li(✉), and Chi Zhang

Harbin Institute of Technology (Weihai), Weihai, China
libol983@hit.edu.cn

**Abstract.** Due to the limited energy of nodes and the harsh working environment in underwater sensor networks, designing energy-efficient routing algorithms to achieve data acquisition is particularly important. Using the correlation of original signal in underwater sensor networks, in this paper, an uneven-layered, multi-hop routing based on distributed compressed sensing (DCS-ULM) is proposed to achieve data collection. The simulation results show that DCS-ULM can effectively prolong the lifetime of networks while ensuring the reconstruction accuracy of original data.

**Keywords:** Compressed sensing · Underwater wireless sensor network · Three-dimensional routing · Energy consumption

## 1 Introduction

At present, the design of routing protocol for underwater sensor networks is primarily focused on the research of data transmission mode, and it is necessary to propose energy-efficient data fusion measures to extend the network lifetime. The compression perception theory (CS) proposed by Candès, Romberg, Tao, and Donoho in 2004 has revolutionized the field of signal processing [1]. The application of CS technology effectively reduces the energy cost and data processing complexity of information acquisition, and provides a good opportunity for large-scale development and application of marine monitoring sensor network. If combined with the design of proper routing protocol, the energy consumption of nodes can be greatly reduced.

In the current research field, a large number of CS technologies have emerged in wireless sensor networks' (WSNs) [2–4] data collection method research. In this paper, CS technology is applied to the data acquisition process of three-dimensional underwater sensor network and combined with the design of routing protocol, which can effectively prolong the lifetime of the network while ensuring the accuracy of the original data reconstruction.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_101

## 2 System Model

Considering the node fixed in the three-dimensional static underwater sensor network with different depths, $n$ underwater nodes and a water surface sink node are distributed randomly, which can be used to realize real-time monitoring of underwater environment[5–7]. Assume that each node has the ability to locate and has a unique node ID. The node packet format adopted in this paper is shown in Fig. 1, and $N$ is the raw data quantity; the selection of projection coefficient $\phi_{ij}$ of nodes in this paper is as follows:

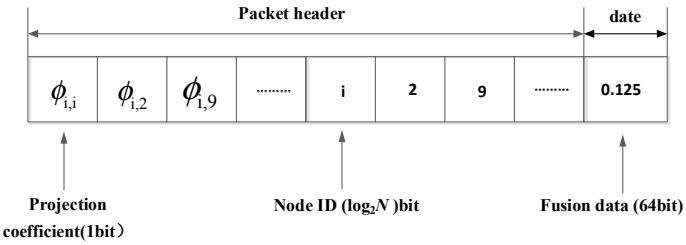



**Fig. 1.** Data packet format

$$\phi_{ij} = \begin{cases} +1, p = \dfrac{1}{2} \\ -1, p = \dfrac{1}{2} \end{cases} \tag{1}$$

As shown in Fig. 1, the data packet contains a node ID of $i$, 2, 9, and the fused data is,

$$y_i = \phi_{i,i} x_i + \phi_{i,2} x_2 + \phi_{i,9} x_9 \tag{2}$$

In addition, in this paper, introduces the mechanism of uneven stratification into the design of routing algorithm. We transform the layered mechanism based on depth into horizontal layering. Firstly, the three-dimensional underwater communication network is classified according to uneven-layered spacing $r$, $2r$, $3r$, $4r$, and so on, and $r$ represents the network base radius. Secondly, the node determines the rank according to its geographical location and completes the next-hop node selection.

## 3 Proposed Scheme

First of all, each node needs to judge the rank according to the location information, if the location information of the current node satisfies the formula,

$$\sum_{i=1}^{n_2-1} i \cdot r \leq D < \sum_{i=1}^{n_2} i \cdot r \tag{3}$$

$$D = \sqrt{(X_i - X_{\text{sink}})^2 + (Y_i - Y_{\text{sink}})^2} \tag{4}$$

where $X$ is the actual horizontal axis value of the node, and $Y$ is the actual ordinate value of the node, and the rank is $n_2$.

The process of DCS-ULM algorithm can be described as: (1) at the opening stage of each round data transmission, the underwater node completes the information collection and initializes the packet. (2) On the data transmission phase, each node generates $a$ random number $q$ between $(0, 1)$ by equipping an independent, uniformly distributed Bernoulli generator. $p$ is the access probability, if $q \leq p$, this node will be used as a routing open node and participates in data transmission. (3) The open node participating in data transmission is $i$, judged by its initialization rank, if the rank of node $i$ is 1, the packet would be forwarded directly to the surface sink node. (4) After receiving the broadcast packets, the node $j$ in the broadcast range compares its location information with the source node $i$ to determine whether it can be the next-hop node. (5) Source node $i$ completes information extraction based on received answer packets. (6) After the next hop node $j$ receives the data packet, according to its own data packet, determine whether the data fusion can be carried out. (7) If the number of nodes satisfying the conditions in the step 6 is $n_3 = 0$, then node $i$ selects the answer node with the rank $n_2$ as the next-hop node set, and then completes the next-hop node judgment based on the value of $W$ in the node set. (8) Finally, the surface sink node uses the received data packets to extract the observed value $\mathbf{y}$, and constructs the measurement matrix $\Phi$ according to the node ID and the random coding coefficient $\phi_{ij}$, and then realizes the effective reconstruction of the original data $\boldsymbol{x}$.

## 4 Simulation and Analysis

Random deployment of 2000 underwater nodes in an interval with longitude [238.8°E, 239.18°E], latitude [32.5°N, 32.88°N], and depth $[0, 800$ m] to complete collection and reconstruction the temperature data on September 3, 2012, in the Gulf of Southern California, USA, the original amount of data is $N$ = 8000.

It can be seen from the Fig. 2 that compared with DCS-HLM algorithm, DCS-ULM algorithm starts to appear death nodes when the number of rounds $r$ = 35, and the dead nodes appear the last time, but the death rate of the nodes is faster than that of DCS-HLM. Using the n-of-n criterion, DCS-ULM algorithm can effectively prolong the network life cycle.

It can be seen from the Fig. 3 that when the number of rounds of network execution is the same, the maximum energy consumption of nodes in DCS-ULM algorithm is the lowest. When the maximum energy consumption of the node is greater than $E_0$, the node's

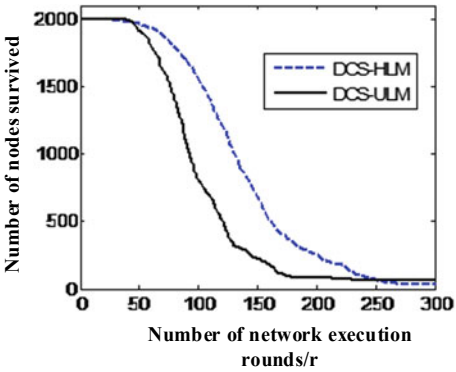



**Fig. 2.** Relationship between the number of nodes' survival and the number of rounds of network execution

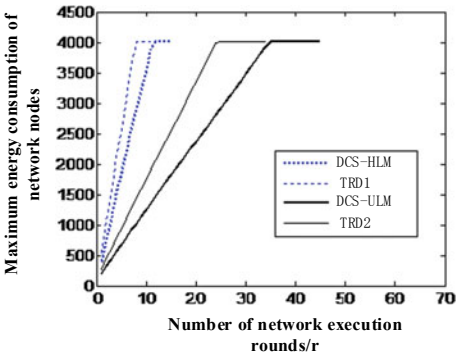



**Fig. 3.** Relationship between the maximum energy consumption of nodes and the number of rounds of network execution

maximum energy consumption of the network will not change due to the failure of the dead node. Therefore, DCS-ULM algorithm can effectively save network energy consumption.

## 5 Conclusions

For the three-dimensional underwater sensor networks with limited node energy, a data collection algorithm based on distributed compressed sensing (DCS-ULM) is proposed in this paper. The algorithm integrates data compression and transmission, and combines compressed sensing technology with the design of three-dimensional routing algorithm. Using real data as experimental processing object, the results show that DCS-ULM algorithm effectively solves the "hot spot" problem in the process of data

transmission, and prolongs the lifetime of three-dimensional underwater acoustic sensor network while ensuring the accuracy of original data reconstruction.

**Acknowledgments.** This work is supported in part by National Natural Science Foundation of China (No. 61401118, and No. 61671184), Natural Science Foundation of Shandong Province (No. ZR2018PF001 and No. ZR2014FP016), the Fundamental Research Funds for the Central Universities (No. HIT.NSRIF.201720 and HIT.NSRIF.2016100) and the Scientific Research Foundation of Harbin Institute of Technology at Weihai (No. HIT(WH)201409 and No. HIT (WH)201410).

## References

1. Liu Z, Guoliang X, Ying H. Timing signal subsection compression algorithm in WSNs based on compressed sensing theory. Chin J Sens Actuators. 2016;01:122–8.
2. Yuxiao C, Ying Q, Lei H. A kind of compressed sensing clustering algorithm for wireless sensor network. Microelectron Comput. 2015;11:59–63.
3. Wang X, Zhao Z, Xia Y, et al. Compressed sensing for efficient random routing in multi-hop wireless sensor networks. Int J Commun Netw Distrib Syst. 2011;7(3):275–92.
4. Bassi F, Liu C, Iwaza L. Compressive linear network coding for efficient data collection in wireless sensor networks. In: European signal processing conference; 2012. p. 714–8.
5. Coates RFW. Underwater acoustic systems. Halsted Press; 1989.
6. Zorzi M, Casari P, Baldo N, et al. Energy-efficient routing schemes for underwater acoustic networks. IEEE J Sel Areas Commun. 2008;26(9):1754–66
7. Liu G, Kang. W. Underwater sparse sensor network information acquisition technology based on compressed sensing. J Instrum Instrum. 2014;3(2):253–60.

# Performance Research of Multiuser Interference in Chaotic UWB System

Bing Zhao, Peize Li, and Jie Yang(✉)

Heilongjiang University, Harbin Xuefu Road. 74, Heilongjiang Province, China
guitaryang@126.com, 18846772820@163.com,
172175491@qq.com

**Abstract.** Ultra-broadband system uses extremely short pulse transmission information, occupies extremely wide bandwidth, and has very low power spectral density. Because of its low power consumption, low interception rate, and the characteristics of good benefit against multi path fading, UWB system is especially suitable for short-distance wireless communication. To a great extent, the development of ultra wide band system depends on the effective multiple access technology, mainly to solve how to let more users share limited spectrum resources. The maximum system capacity is achieved by reasonably allocating and using the given spectrum. In this paper, we mainly study the multiple access technology of direct sequence spread spectrum UWB system based on digital chaotic sequence, and compare it with the real random sequence direct sequence spread spectrum ultra wide band system. The simulation results show that the DS-WWB system with digital chaotic sequence spread spectrum multi-access distribution is close to the actual random sequence system in terms of user capacity and bit error rate (BER). The real random sequence cannot be copied, and the digital chaotic sequence is easy to be produced and copied, so it has stronger physical realizability.

**Keywords:** DS-UWB · Multiple-access technology · Digital chaotic sequence · Multiuser interference

Multiple access technology is a key technology in UWB communication system. Compared with the traditional wireless communication system, the UWB multi access system has the characteristics of small network coverage, fewer users, and overlap ofmultiple network coverage areas [1, 2]. The bandwidth of the UWB system is very wide, and if broadband is used to transmit all information data, the bandwidth utilization is very low. For UWB system, in order to improve spectrum utilization,it is necessary to transmit multi channel data information on the same band, that is, to improve multi address capability. Because UWB adopts direct pulse transmission technology, all users work within the same frequency range, so it is suitable to adopt the spread-spectrum technology.

The selection of spread-spectrum codes is a very important technical problem in spread-spectrum multiple-access communication system, whose characteristics directly affect the complexity, synchronization time, system performance, and system capacity of multiple-access communication systems. In all current candidate codes, the most


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_102

commonly used are linear sequences such as M sequence and Gold sequence. M sequence and Gold sequence as the linear sequences, their implementation are relatively simple, and have good auto correlation properties. But they have a limited number of spread spectrum codes, and some cross-correlation performance is limited. With the development and maturation of nonlinear chaos theory, chaotic sequences are used as address codes in UWB systems. As a sequence of spread spectrum, chaotic sequence has the advantages of high initial value sensitivity, goodcorrelation, large quantity, arbitrary length, good confidentiality, easy to produce and copy. Therefore, the sequence to spread-spectrum modulation technology has become the research hot spot of UWB system gradually.

## 1 Generation of Digital Chaotic Spread-Spectrum Sequences

The generation method of the digital chaotic sequence in this paper is shown in Fig. 1.

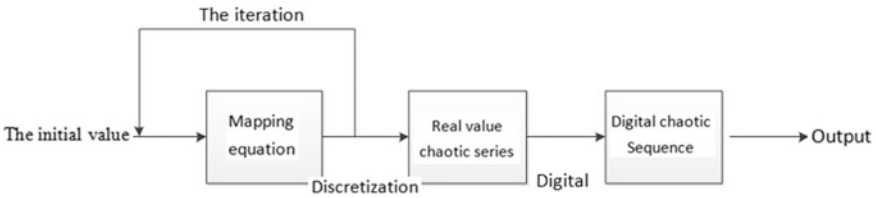



**Fig. 1.** Generation of digital chaotic sequences

The proper nonlinear equation was selected, and the system parameters were reasonably adjusted to keep the system in the chaotic state interval, and the real-valued chaotic sequence was generated by nonlinear iteration [3]. Logistic mapping was used in this paper.

Logistic mapping is defined as follows.

$$x_{k+1} = \mu x_k(1 - x_k) \quad x \in (0, 1) \tag{1}$$

The mapping is in a chaotic state when $3.57 < \mu \leq 4$.

Logistic mapping is distributed on the interval (0.1). The decimals in this region can be quantified in binary terms, which can be quantified as a binary dual-polarity sequence $X$.

$$X = \begin{cases} 1 & x > 0.5 \\ -1 & x \leq 0.5 \end{cases} \tag{2}$$

$X$ corresponds to the decimal $x$ one by one, and the actual calculation shows that the sequence produced by quantization still has chaotic characteristics.

## 2 Chaotic Direct-Sequence Spread-Spectrum UWB System

The generation process of chaotic direct-sequence spread-spectrum UWB signal is as follows. The information sequence of rate $R_b = 1/T_b$ is converted into a bipolar binary sequence after repeated $N_s$ times by repeated encoder. It is amplified by a digital chaotic sequence of length $N_p$, usually assuming that $N_p$ is equal to integer multiples of $N_s$. The amplified sequence passes through the PAM modulator to produce a unit pulse sequence with a rate of $R_P = N_s/T_b = 1/T_s$ (pulse/s). After forming a filter with impulse response $p_t$, the pulse sequence outputs a chaotic direct-sequence spread-spectrum ultra-broadband signal $s(t)$, where $p_t$ is an impulse with a duration far less than $T_s$.

The output signal $s(t)$ can be expressed as follows:

$$s(t) = \sum_{j=-\infty}^{+\infty} d_j p(t - jT_s) \tag{3}$$

where $d_j$ is the sequence after amplification.

The output waveform is a typical PAM waveform. Since there is no time shift, the pulse appears at regular intervals [4].

It is assumed that the useful signal $r_{\mathrm{u}}(t)$ of the receiver is the signal after the delay and attenuation of the transmitting signal $s(t)$, i.e.,

$$r_{\mathrm{u}}(t) = \alpha s(t - \tau) \tag{4}$$

The channel gain $\alpha$ and channel delay $\tau$ are all dependent on the transmission distance $D$ between transmitter and receiver [5].

The best receiver for AWGN channel consists of two parts: correlator and detector. The correlator changes the received signal into a set of decision variables $\{Z\}$, and the detector determines which signal waveform is sent based on the observed value. When sending bits are distributed independently, the average error rate $\mathrm{Pr}_b$ is:

$$\underset{b}{\mathrm{Pr}} = \frac{1}{2}\mathrm{erfc}\left[\sqrt{\frac{E_{RX}}{N_0}}\right] \tag{5}$$

where $E_{RX} = \alpha^2 E_{TX}$ is the receiving energy per pulse.

Assuming that the transmitter and receiver are fully synchronized, each pulse is limited to a specific time interval and the duration of the pulse itself is limited. Ideally, there is no interference (ISI) in a transmission signal, but after the actual channel transmission, due to the presence of multipath, the pulse may experience a variety of different time delays. Moreover, in the case of multiple users, pulses from other transmission links may collide with pulses sent from the reference transmission link, resulting in interference noise called multiuser (MUI).

## 3 Performance Analysis of Chaotic Direct-Sequence Spread-Spectrum UWB System with Multiuser Interference

SNR and BER are the important indexes of the data communication performance. In view of the chaotic direct sequence spread spectrum UWB users,this article describes the simulation content about multi-user interference and bit error rate under different SNR parameter. And the true random sequence ultra-wide band system theory is adopted to improve the contrast. This article assumes that the receiver and transmitter are completely synchronous.

System simulation parameters are as follows: The sampling frequency is 50 GHZ, the average transmission power is 30 dBm, the average pulse time is 2 ns, every bit repetitive pulse count is 10, the pulse duration is 0.5 ns, the spread-spectrum sequence length is 500, and there are four chaotic sequence classification parameters [6].

First, considering the case of a single user, Fig. 2 shows the error characteristic curve of the receiving end receiving the UWB signal when there is only one chaotic direct-sequence spread-spectrum UWB system user communication in the channel.

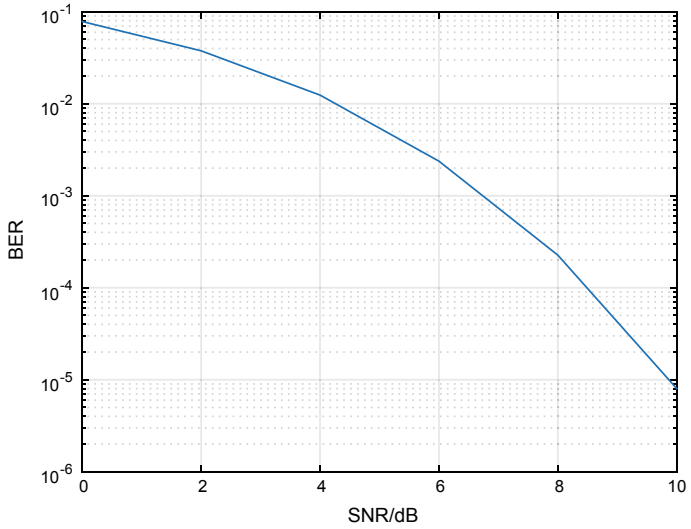



**Fig. 2.** Single-user error rate curve

When multiuser interference is not considered, it can be seen from the figure that the system error rate decreases with the increase of SNR. Through simulation, it can be seen that when SNR is greater than 16 dB, there will be no error rate.

When considering the multi-user interference, a channel exists in the 5 users of different delay and attenuation of different users. At the same time, the channel arrived at the receiving end. Under the condition of different SNR [7], the characteristics of UWB system error results are shown in Fig. 3.

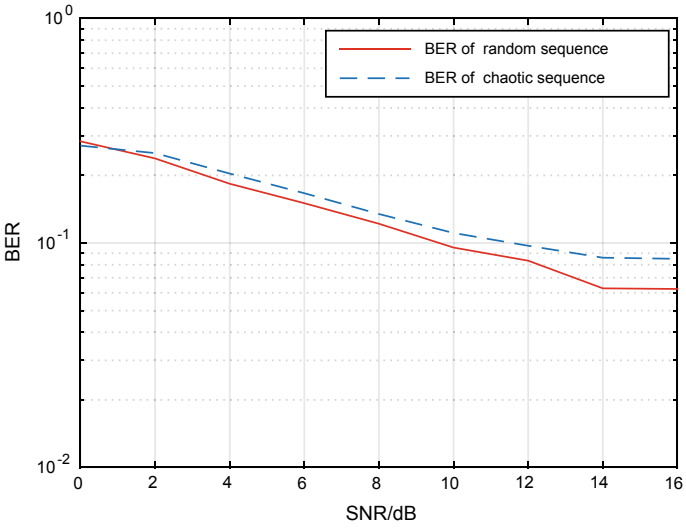



**Fig. 3.** Comparison of system error rate under multi-user interference

Figure 3 shows that the mutual correlation digital chaotic sequence is good, the system performance is close to the true random sequence ultra wide band system. But because of the true random sequence can't realize the replication, it is not conducive to produce. Therefore, it cannot be achieved in the real system [8]. Digital chaotic sequence in direct-sequence spread-spectrum UWB systems due to their good correlation and large number.

## 4 Conclusion

This article mainly for ultra wide band digital chaotic system of anti-jamming performance in a multiuser environment was studied, using the MATLAB software; the simulation experiment is based on the analysis of the BER curve that digital chaotic sequence in the direct-sequence spread-spectrum ultra wide band system has goodperformance.

**Acknowledgments.** This project is sponsored by the talent innovation project of Heilongjiang University and belongs to the basic scientific research business of Heilongjiang University. (No. HDRCCX-201605).

## References

1. Liu K, Shen X, Zhang R, et al. Performance analysis of distributed reservation protocol for UWB-based WPAN. IEEE Trans Veh Technol. 2009;58(2):907–14.

2. Scholtz R, Weaver R, Homier E, et al. UWB radio development challenges. In: The 11th IEEE international symposium on personal, indoor and mobile radio communications. London, UK; 2000. p. 621–6.
3. Caihong Li, Yibin Li, Lei Zhao, Yanhua Zhang, Xinhui Zhang. One-dimensional Logistic mapping chaotic pseudo—random sequence statistical characteristics (in Chinese). Comput Appl Res. 2014;31(5):1403–6.
4. Maggio GM, Laney D, Larson L. BER and error-floor calculation for multi-access PCTH. In: Global telecommunications conference, 2002. Globecom'02. IEEE, Taipei, Taiwan, vol. 2; 2002 Nov. p. 1324–8.
5. Hailian He, Guangyi Wang. Improvement on multiple access performance in DS-UWB by using hyperchaotic sequence (in Chinese). Commun Technol. 2009;42(01):36–8.
6. Laney DC, Maggio GM, Lehmann F, Larson L. Multiple access for UWB impulse radio with pseudo chaotic time hopping. IEEE J Sel Areas Commun. 2006;20(9):1692–700.
7. Maggio GM, Laney D, Lehmann F, Larson L. A multi-access scheme for UWB radio using pseudo-chaotic time hopping. In: 2002 IEEE conference on ultra wideband systems and technologies, Baltimore, MD, USA; 2002 May. p. 225–229.
8. Wei Zhang, Hongmei Xie, Baoping Wang. A novel segmented logistic turbidimetric spread spectrum communication algorithm (in Chinese). Comput Appl Res. 2013;40(1):59–62.

# Part IV
# Feature Selection

# Dual-Feature Spectrum Sensing Exploiting Eigenvalue and Eigenvector of the Sampled Covariance Matrix

Yanping Chen[1] and Yulong Gao[2(✉)]

[1] Harbin University of Commerce, Harbin 150028, Heilongjiang, China
yanping1009@163.com
[2] Harbin Institute of Technology, Harbin 150080, Heilongjiang, China
ylgao@hit.edu.cn

**Abstract.** The signal can be charactered by both eigenvalues and eigenvectors of covariance matrix. However, the existing detection methods only exploit the eigenvalue or eigenvector. In this paper, we utilize both eigenvalues and eigenvectors of the sampled covariance matrix to perform spectrum sensing for improving the detection performance. The features of eigenvalues and eigenvectors are considered integratedly, and the relationship between the false-alarm probability and the decision threshold is offered. To testify this method, some simulations are carried out. The results demonstrate that the method shows some advantages in the detection performance over the conventional method only adapting eigenvalues or eigenvectors.

**Keywords:** Dual-feature · Spectrum sensing · Cognitive radio · Eigenvalue and eigenvector

## 1 Introduction

At present, wireless spectrum resources are assigned statically, and the given spectrum bands are authorized to some communication systems. Simultaneously, other systems cannot occupy these spectrum bands. For the limited wireless resources, the static style of working results in a precipitous decline in available spectrum resources. However, some of these licensed frequency bands are underutilized [1–3]. To solve the unbalanced utilization of spectrum resources, cognitive radio is presented. Many institutions and scholars put more attentions on key issues of cognitive radio [4–7], such as spectrum sensing, spectrum sharing.

Nowadays, conventional spectrum sensing methods are composed of matched filter detection, energy detection, likelihood ratio test detection, and cyclostationary detection [8–11]. When the signal is corrupted by the Gaussian white


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_103

noise, matched filter detection has optimal detection performance. Unfortunately, some prior information about licensed users must be known. More difficultly, precise synchronization between licensed users and secondary users must also be achieved. It is mostly impossible for secondary users to acquire this prior knowledge. Energy detection is most commonly used in practice due to its simplicity and low computational complexity. No additional prior information about licensed users is required for energy detection. However, energy detection suffers from noise uncertainty and SNR wall, which degrade the detection probability and the false-alarm probability severely. Likelihood ratio test-based detection method has optimal detection performance under the Neyman–Pearson criterion. This method detects the signal by virtue of the difference of probability density function between the licensed user and noise. Obviously, the algorithm requires the corresponding prior knowledge. For cyclostationary detection method, the inherent features of the licensed user, which arise from modulation style, signal rate or other parameters, are exploited. This method can overcome the effect of fading and shadow and has a better performance than other detection methods in the low SNR region. But, the computational complexity is unaffordable for the practical application in most cases.

From the previous analysis, we can see that the above-mentioned methods have advantages and disadvantages. So the detection method based on random matrix was first introduced by Cardoso to deal with these problems. This method can cope with the noise uncertainty of energy detection and require no prior information about the licensed user. So this method attracts more attentions. In [12], the ratio of the maximum eigenvalue and the minimum eigenvalue is adopted as the test statistic to perform spectrum sensing, and the corresponding threshold is calculated according to MP law of random matrix. But the requirement of too many samples is unsatisfactory. Aiming to solve this problem, the authors in [13] exploit the Tracy–Widom distribution of maximum eigenvalue to modify the threshold and propose the maximum–minimum eigenvalue (MME) method. In light of these work, many modifications were made to improve the sensing performance.

In fact, it is well-known that eigenvectors also contain related information about the presence of signal. Compared to the method exploiting the eigenvalue, only a few works about the utilization of eigenvector are reported. In terms of principal component analysis, eigenvector-based feature template matching (FTM) method was proposed [14]. This method utilizes the eigenvector of sampled covariance matrix as the test statistic to carry out spectrum sensing. Additionally, the literature [15] combines kernel function of machine learning with FTM to propose the kernel feature template matching method. It is pointed that feature template matching (FTM) method is a special case of kernel feature template matching method. The authors of the literature also derived the false-alarm probability and decision threshold and analyzed the factors of affecting the detection performance.

Stimulated by the eigenvalue-based method and FTM algorithm, we combine the eigenvalue and eigenvector to carry out spectrum sensing.

In the proposed method, the product of main eigenvector and the ratio of maximum and minimum eigenvalue is utilized as the test statistic. In terms of the concept of random matrix, the false-alarm probability is derived by approximating some results. Finally, the decision threshold is calculated.

## 2 System Model and Algorithm Description

We assume that there are $K$ sensing nodes. For the $i$th node, the binary hypothesis test can be expressed as

$$y_i(n) = \begin{cases} \omega_i(n), & H_0 \\ h_i(n)x_i(n) + \omega_i(n), & H_1 \end{cases} \tag{1}$$

where $\omega_i(n)$, $x_i(n)$, and $h_i(n)$ denote the additive noise, the licensed user signal, and channel gain, respectively. Two different segments signals are exploited to construct the received signal matrix $\mathbf{Y}_1$ and $\mathbf{Y}_2$

$$\mathbf{Y}_1 = \begin{bmatrix} y_1(1)\, y_1(2) \ldots y_1(N) \\ y_2(1)\, y_2(2) \ldots y_2(N) \\ \vdots \quad \cdots \quad \vdots \\ y_K(1)\, y_K(2) \ldots y_K(N) \end{bmatrix} \tag{2}$$

$$\mathbf{Y}_2 = \begin{bmatrix} y_1(N+1)\, y_1(N+2) \ldots y_1(2N) \\ y_2(N+1)\, y_2(N+2) \ldots y_2(2N) \\ \vdots \quad \cdots \quad \vdots \\ y_K(N+1)\, y_K(N+2) \ldots y_K(2N) \end{bmatrix} \tag{3}$$

where $N$ is the number of signal samples. Their corresponding covariance matrix $\mathbf{R}_1$ and $\mathbf{R}_2$ are expressed as

$$\mathbf{R_1} = \frac{1}{N}\mathbf{Y}_1\mathbf{Y}_1{}^{\mathrm{T}}, \quad \mathbf{R_2} = \frac{1}{N}\mathbf{Y}_2\mathbf{Y}_2{}^{\mathrm{T}} \tag{4}$$

If only the noise exists, the maximum eigenvalue $\lambda_{\max}$ and minimum eigenvalue $\lambda_{\min}$ of the sampled covariance matrix have $\lambda_{\max} = \lambda_{\min} = \sigma^2$. But, when the signal is present, the maximum eigenvalue and minimum eigenvalue satisfy $\lambda_{\max} > \lambda_{\min} = \sigma^2$. Therefore, we have

$$\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)_{H_0} < \left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)_{H_1} \tag{5}$$

Analogously, for the main eigenvector of $H_0$ and $H_1$ case, the following relation holds

$$(|\langle \mathbf{a}_1, \mathbf{b}_1 \rangle|)_{H_0} < (|\langle \mathbf{a}_1, \mathbf{b}_1 \rangle|)_{H_1} \tag{6}$$

where $\mathbf{a}_1$ and $\mathbf{b}_1$ are the corresponding main eigenvectors of the sampled covariance matrix $\mathbf{R}_1$ and $\mathbf{R}_2$. Jointly considering (8) and (9), we can obtain

$$\left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)_{H_0} (|\langle \mathbf{a}_1, \mathbf{b}_1 \rangle|)_{H_0} < \left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)_{H_1} (|\langle \mathbf{a}_1, \mathbf{b}_1 \rangle|)_{H_1} \tag{7}$$

We can observe that the product of the ratio of the maximum eigenvalue and minimum eigenvalue and the correlation of main eigenvectors for the $H_0$ and $H_1$ case differ obviously. Compared to the method only employing eigenvalue or eigenvector, more obvious difference can be obtained for the product. Thus, we select the product as the test statistic of spectrum sensing. The proposed method is summarized as follows.

## 3 Solving the False-Alarm Probability and the Decision Threshold

From the definition of false-alarm probability, we have

$$P_{fa} = P(D_1|H_0) = P(T > \gamma|H_0) = P\left(\rho_{\,1}\rho_{\,2}\left|\langle \mathbf{a}_1, \mathbf{b}_1\rangle\right| > \gamma|H_0\right) \tag{8}$$

When only the noise exists, the covariance matrix of the received signal is Wishart matrix. From the MP law, the maximum eigenvalue and minimum eigenvalue in the $H_0$ case are

$$\lambda_{\max} = \sigma^2\left(1+\sqrt{c}\right)^2, \;\; \lambda_{\min} = \sigma^2\left(1-\sqrt{c}\right)^2 \tag{9}$$

where $c$ is the ratio of the number of nodes and the number of samples. The ratio of the maximum eigenvalue and the minimum eigenvalue is

$$\eta = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{\left(1+\sqrt{c}\right)^2}{\left(1-\sqrt{c}\right)^2} \tag{10}$$

For the matrix $\mathbf{R}_1$ and $\mathbf{R}_2$, we can calculate the corresponding $\eta_1$ and $\eta_2$.

It is pointed by the literature [16] that the eigenvalue and eigenvector of Wishart matrix are independent to each other. Let $\gamma = \varepsilon\eta^2$. Coupled with (10), the false-alarm probability can be expressed in the form

$$P_{fa} = P\left(\rho_{\,1}\rho_{\,2}\left|\langle \mathbf{a}_1, \mathbf{b}_1\rangle\right| > \gamma|H_0\right) \approx P\left(\left|\langle \mathbf{a}_1, \mathbf{b}_1\rangle\right| > \varepsilon|H_0\right) \tag{11}$$

To calculate the false-alarm probability, we derive the probability density function of $|\langle \mathbf{a}_1, \mathbf{b}_1\rangle|$. In terms of the concept of eigenvalue decomposition, we can get

$$\mathbf{R}_1 = \mathbf{A}\boldsymbol{\Lambda}_1\mathbf{A}^T, \;\; \mathbf{R}_2 = \mathbf{B}\boldsymbol{\Lambda}_2\mathbf{B}^T \tag{12}$$

where $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ are the diagonal matrix containing eigenvalues of $\mathbf{R}_1$ and $\mathbf{R}_2$. The columun vectors of and correspond to eigenvectors. And the main eigenvectors $\mathbf{a}_1$ and $\mathbf{b}_1$ of the sampled covariance matrix $\mathbf{R}_1$ and $\mathbf{R}_2$ are the first column of matrix $\mathbf{A}$ and $\mathbf{B}$. Because $\mathbf{A}$ and $\mathbf{B}$ are unitary matrix, we have $f\left(\mathbf{A}^T\mathbf{B}\right) = f\left(\mathbf{B}\right)$. We can say that the elements of $\mathbf{A}^T\mathbf{B}$ and $\mathbf{B}$ in the same position follow the same distribution, i.e., $f\left(\langle \mathbf{a}_1, \mathbf{b}_1\rangle\right) = f\left(b_{11}\right)$.

From the properties of unitary matrix, $b_{11}$ obeys the beta distribution with parameters $\alpha = \frac{1}{2}$ and $\beta = \frac{N-1}{2}$

$$f\left(x\right) = \frac{\left(1-x^2\right)^{(n-1)/2-1}}{B\left(\frac{1}{2}, \frac{N-1}{2}\right)} \tag{13}$$

Because the distribution are computationally complex, we approximate it with the Gaussian distribution for the large $N$ [15]. According to the result derived in the literature [15], $f(x)$ can be replaced by $h(x)$

$$h(x) = \frac{\sqrt{N}}{\sqrt{2\pi}} e^{-\frac{Nx^2}{2}} \tag{14}$$

Substituting (14) into (11) yields

$$P_{fa} \approx P\left(|\langle \mathbf{a}_1, \mathbf{b}_1 \rangle| > \varepsilon | H_0\right) = 2Q\left(\sqrt{N}\varepsilon\right) \tag{15}$$

When the false-alarm probability is preset, we get the parameter $\varepsilon$. Therefore, the corresponding threshold can be calculated as

$$\gamma = \varepsilon\eta^2 = \frac{(1+\sqrt{c})^4}{(1-\sqrt{c})^4} Q^{-1}\left(\frac{P_{fa}}{2}\right) \frac{1}{\sqrt{N}} \tag{16}$$

## 4 Numerical Simulation and Analysis

We first verify the rationality of the decision threshold. We offer the relationship among the test statistic of $H_0$ case, test statistic of H1 case, and the decision threshold in Fig. 1 when $K$=10 and $SNR = -10$ dB. It is demonstrated that the threshold varies dynamically with the number of samples. Additionally, we can also observe that the difference of the test statistic for H0 and H1 case is obvious for all the samples and the decision threshold is between the test statistics for

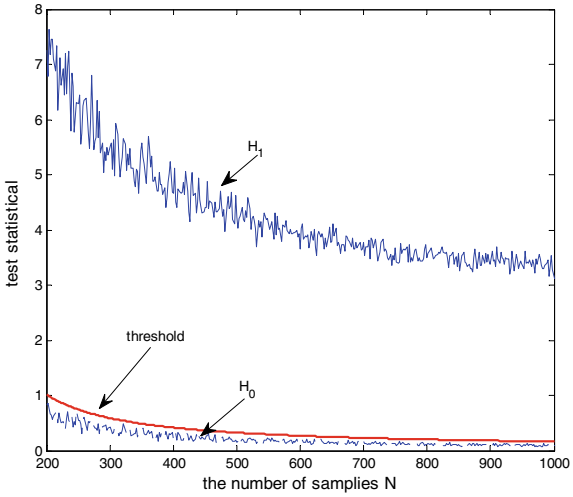



**Fig. 1.** Estimated pdf and theoretical result of the main eigenvector for $H_0$

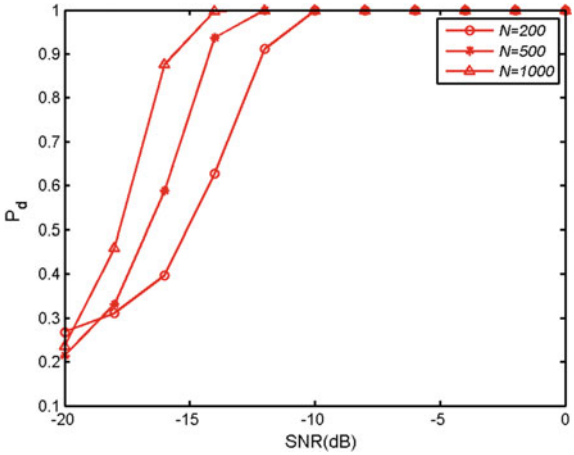



**Fig. 2.** Detection probability for the different samples and SNR

$H_0$ and $H_1$ case. All these results guarantee the proper threshold for obtaining satisfactory performance.

We now verify the detection performance of the proposed method. The simulated parameters are set as follows. $K = 10$, $N = 200$, 500, 1000. The number of Monte Carlo simulation is 2000. Figure 2 plots the detection probability for the different SNR. We can draw three conclusions. (1) The detection probability is improved with the increasing of samples for the low $SNR$ region. For example, the detection probability is 0.5 and 1 for $N = 200$ and 1000 when $SNR = -15$ dB. (2) When SNR is more than $-10$ dB, the detection probability remains a constant 1 for the different samples. (3) when the samples remain unchanged, the detection probability increases proportionately with $SNR$.

To testify the superiority of the proposed method over other algorithms, we show the detection probability of FFM, MME, and the proposed method in Fig. 3. The simulated parameters are set as follows. $K = 10$, $N = 500$. The number of Monte Carlo simulation is 2000. It is noted that the samples are divided into two segments for FTM and the proposed method, and each segment has $N/2$ samples. We can observe that the proposed method has good detection performance over other two algorithms.

## 5 Conclusion

Inspired by the FTM and eigenvalue-based spectrum sensing, we introduced the spectrum sensing algorithm based on dual-feature consisting of eigenvalues and eigenvector. We defined the product of the ratio of the maximum eigenvalue and the minimum eigenvalue and main eigenvector as the test statistic. This test statistic considers the eigenvalue and eigenvector comprehensively. Based on the test statistic, we derived the false-alarm probability and offer the closed

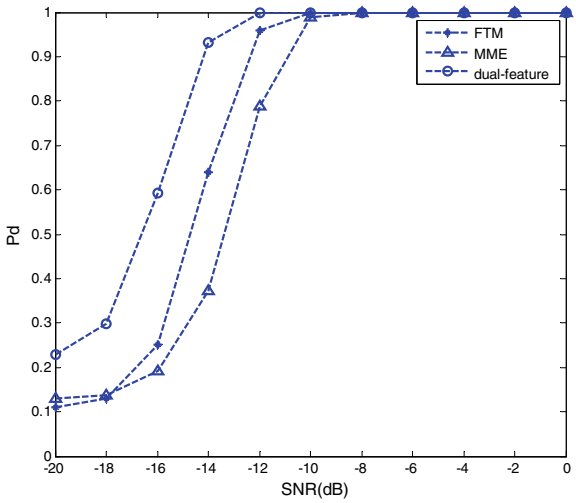



**Fig. 3.** Detection probability for the different samples and SNR

form of threshold under Neyman–Pearson criterion. In simulations, we testified the correctness of Gaussian approximation of probability density function and compared the proposed method with FTM and MME algorithms. The simulation results showed that the proposed method outperforms the FTM and MME algorithms with lower computational complexity.

**Acknowledgments.** This work is supported by National Natural Science Foundation of China (NSFC) (Grant No. 61671176).

## References

1. Abdelmohsen A, Hamouda W. Advances on spectrum sensing for cognitive radio networks: theory and applications. IEEE Commun Surv Tutorials. 2017;19(2):1277–304.
2. Guo H, Jiang W, Luo W. Linear soft combination for cooperative spectrum sensing in cognitive radio networks. IEEE Commun Lett. 2017;21(7):1573–6.
3. Wang B, Liu KJR. Advances in cognitive radio networks: a survey. IEEE J Sel Top Sig Process. 2011;5(1):5–23.
4. Mitola J, Maguire GQ. Cognitive radio: making software radios more personal. IEEE Pers Commun. 1999;6(4):13–8.
5. Mchenry M, Livsics E, Nguyen T, Majumdar N. XG dynamic spectrum access field test results [Topics in radio communications]. IEEE Commun Mag. 2007;45(6):51–7.
6. Liu Chang, Li Ming, Jin Ming-Lu. Blind energy-based detection for spatial spectrum sensing. IEEE Wirel Commun Lett. 2015;4(1):91–8 Feb.
7. Cabric D, Brodersen RW. Physical layer design issues unique to cognitive radio systems. Proc IEEE Pimrc. 2005;2:759–63.

8. Cabric D, Tkachenko A, Brodersen R. Spectrum sensing measurements of pilot, energy, and collaborative detection. In: IEEE Conference on Military Communications; 2006. p. 1–7
9. Sonnenschein A, Fishman PM. Radiometric detection of spread-spectrum signals in noise of uncertain power. IEEE Trans Aerosp Electron Syst. 1992;28(3):654–60.
10. Yucek T, Arslan H. A survey of spectrum sensing algorithms for cognitive radio applications. IEEE Commun Surv Tutorials. 2009;11(1):116–30.
11. An T, Kim D, Song I, et al. Cooperative spectrum sensing based on generalized likelihood ratio test under impulsive noise circumstances. In: 2012 IEEE military communications conference; 2012. p. 1–6
12. Cardoso LS, Debbah M, Bianchi P, et al. Cooperative spectrum sensing using random matrix theory. In: International symposium on wireless pervasive computing; 2008. p. 334–8
13. Zeng Y, Liang YC. Maximum-minimum eigenvalue detection for cognitive radio. In: IEEE international symposium on personal, indoor and mobile radio communications; 2007. p. 1–5
14. Zhang P, Qiu R, Guo N. Demonstration of spectrum sensing with blindly learned features. IEEE Commun Lett. 2011;15(5):548–50.
15. Hou S, Qiu RC. Kernel feature template matching for spectrum sensing. IEEE Trans Veh Technol. 2014;63(5):2258–71.
16. Eldar YC, Chan AM. On the asymptotic performance of the decorrelator. IEEE Trans Inf Theor. 2003;49(9):2309–13.

# Adaptive Scale Mean-Shift Tracking with Gradient Histogram

Changqing Xie, Wenjing Kang, and Gongliang Liu(✉)

School of Information and Electrical Engineering, Harbin Institute of Technology (Weihai), Weihai, China
15650131973@yeah.net, {kwjqq,liugl}@hit.edu.cn

**Abstract.** The mean-shift (MS) tracking is fast, is easy to implement, and performs well in many conditions especially for object with rotation and deformation. But the existing MS-like algorithms always have inferior performance for two reasons: the loss of pixel's neighborhood information and lack of template update and scale estimation. We present a new adaptive scale MS algorithm with gradient histogram to settle those problems. The gradient histogram is constructed by gradient features concatenated with color features which are quantized into the $16 \times 16 \times 16 \times 16$ bins. To deal with scale change, a scale robust algorithm is adopted which is called background ratio weighting (BRW) algorithm. In order to cope with appearance variation, when the Bhattacharyya coefficient is greater than a threshold the object template is updated and the threshold is set to avoid incorrect updates. The proposed tracker is compared with lots of tracking algorithms, and the experimental results show its effectiveness in both distance precision and overlap precision.

**Keywords:** Object tracking · Mean-shift · Scale estimation · Gradient

## 1 Introductions

Visual object tracking has always been a challenging work especially in the sequences with the deformation and rotation. The MS tracking has an outstanding performance and is easy to implement. It tracks by minimizing the Bhattacharyya distance between two probability density functions represented by a target and target candidate histograms. The histogram is a statistical feature that does not depend on the spatial structure within the search window. This makes it more robust than other algorithms. But it lacks the essential template update and pixel's neighborhood information leading to a worse accuracy.

The mean-shift algorithm is a nonparametric mode-seeking method for density functions proposed by Fukunaga and Hostetler [1]. Comaniciu et al. [2, 3] use it to track object. And Comaniciu et al. change the window size over multiple runs by a constant factor but produces little effect because the smaller windows usually have higher similarity. The image pyramids and an additional mean-shift algorithm for scale selection had been used after estimating the position to confirm the window size in Collins [4]. But its speed is lower than the conventional MS algorithm. A new histogram that exploits the object neighborhood has been proposed to help discriminate


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_104

the target which is called background ratio weighting (BRW) in Vojir et al. [5]. This approach is faster than others and has a superior effect in sequences with scale change but performs poorly for grayscale sequences.

Gradient information is crucial for appearance representation since it contains pixel's neighborhood information and is insensitive to illumination variation but it always be ignored. Based on this observation, we present a novel adaptive scale MS algorithm with gradient histogram. The gradient information is calculated by Canny edge detector which was developed by Canny [6].

Moreover, the BRW algorithm has been used to improve the performance of videos with scale change. The template of target will be updated by liner interpolation only if conditions are met to avoid addition of incorrect information. The template update also can cope with appearance variation of target. The proposed tracker is compared with lots of algorithms, and the experiment results show that it is more robust and accurate.

## 2 Canny Edge Detector

The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images. To remove the noise, a Gaussian filter is applied to the image; the Gaussian filter kernel of size $(2k + 1) \times (2k + 1)$ is given by:

$$H_{ij} = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(i-(k+1))^2+(j-(k+1))^2}{2\sigma^2}\right);\ 1 \le i,\ j \le (2k+1) \tag{1}$$

Let $I_o$ denote the original image and $H$ is a classic Gaussian filter matrix; we get the image without noise $I = H * I_o$. Then we extract intensity gradient of the image with the Sobel operator proposed by Sobel [7]. The gradient information in horizontal and vertical directions is $G_x$ and $G_y$. From this, the edge gradient can be determined by

$$G = \sqrt{G_x^2 + G_x^2} \tag{2}$$

where

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 2 \end{bmatrix} * I \text{ and } G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * I. \tag{3}$$

The edge extracted from the gradient value is still quite blurred after processing the $G$ with Gauss filter and Sobel operator. The non-maximum suppression and double-threshold joint should be applied to the processed image to improve effect of gradient. Then we can get the gradient image $I_g$ to track the target.

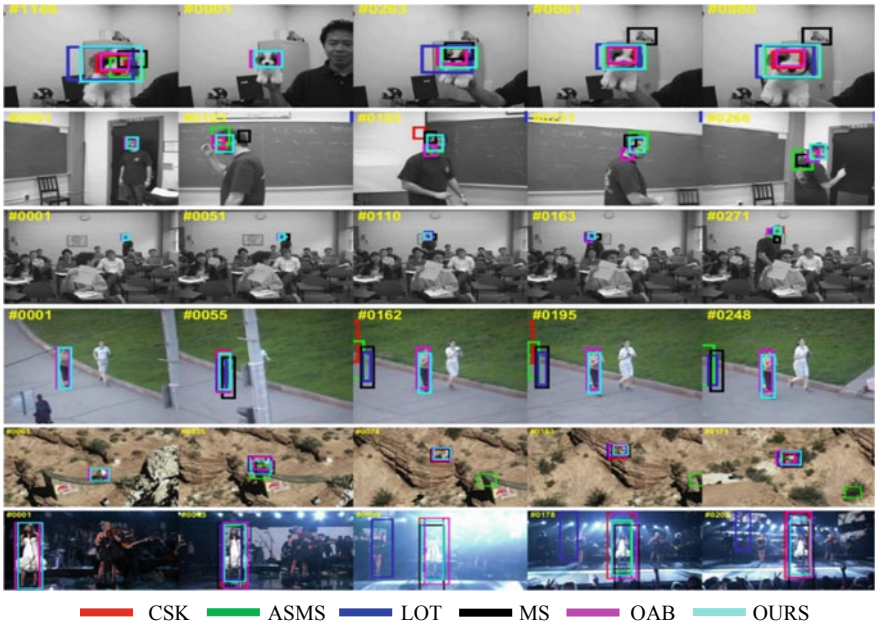



**Fig. 1.** Result of six trackers in six sequences. (The sequence names from top to bottom are dog1, freeman1, freeman3, jogging, mountainbike, and singer1)

## 3 The Tracking Algorithm

Different from the conventional MS tracking algorithm, we append the $I_g$ to original image to get $I_e$. After the combination of the images, we extract the histogram $\hat{\mathbf{q}}$ from $I_e$. To cope with the problem which caused by the size of target changes, we use the BRW-MS instead of conventional MS. We can get $\hat{\mathbf{q}}$ from:

$$\hat{q}_u = C\sum_{i=1}^{N} k\left(\frac{\left(x_i^{*1}\right)^2}{a^2} + \frac{\left(x_i^{*2}\right)^2}{b^2}\right)\delta\left[b\left(\boldsymbol{x}_i^*\right) - u\right] \tag{4}$$

and an ellipsoidal region is used to represent target $\frac{\left(x_i^{*1}\right)^2}{a^2} + \frac{\left(x_i^{*2}\right)^2}{b^2} < 1$ in current frame. The target candidate is given by

$$\hat{p}_{\boldsymbol{u}}(\boldsymbol{y}, h) = C_h\sum_{i=1}^{N} k\left(\frac{\left(y^1 - x_i^1\right)^2}{a^2h^2} + \frac{\left(y^2 - x_i^2\right)^2}{b^2h^2}\right)\delta[b(\boldsymbol{x}_i - u)] \tag{5}$$

where the $h$ is the scale factor. The location of the target is obtained by

$$\hat{y}_1^1 = \frac{1}{a^2} m_k^1(\hat{\boldsymbol{y}}_0, h_0) + \hat{y}_0^1,\ \hat{y}_1^2 = \frac{1}{b^2} m_k^2(\hat{\boldsymbol{y}}_0, h_0) + \hat{y}_0^2 \tag{6}$$

$$h_1 = \left[1 - \frac{A}{K}\right] h_0 + \frac{1}{h_0} \frac{B}{K} \tag{7}$$

where $A$ is

$$\sum_{i=1}^{N} w_i k\left(\frac{\left(y_0^1 - x_i^1\right)^2}{a^2 h_0^2} + \frac{\left(y_0^2 - x_i^2\right)^2}{b^2 h_0^2}\right), \tag{8}$$

$B$ is

$$\sum_{i=1}^{N} w_i \left(\frac{\left(y_0^1 - x_i^1\right)^2}{a^2} + \frac{\left(y_0^2 - x_i^2\right)^2}{b^2}\right) g\left(\frac{\left(y_0^1 - x_i^1\right)^2}{a^2 h_0^2} + \frac{\left(y_0^2 - x_i^2\right)^2}{b^2 h_0^2}\right). \tag{9}$$

And $g(x) = -k'(x)$ is the derivative of $k(x)$. $w_i$ can be obtained by

$$w_i = \max(0, W) \tag{10}$$

where W is

$$\mathrm{W} = \left(\sum_{u=1}^{m} \frac{1}{\hat{\rho}[\hat{\boldsymbol{p}}(\hat{\boldsymbol{y}}_0, h_0), \hat{\boldsymbol{q}}]} \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{\boldsymbol{y}}_0, h_0)}} - \frac{1}{\rho\left[\hat{\boldsymbol{p}}(\hat{\boldsymbol{y}}_0, h_0), 3\widehat{\boldsymbol{bg}}\right]} \sqrt{\frac{\widehat{bg_u}}{\hat{p}_u(\hat{y}, h_0)}}\right) \delta[b(\boldsymbol{x_i}) - u] \tag{11}$$

The $\widehat{\boldsymbol{bg}}$ is the histogram of background computed over the neighborhood of the target in the first frame. Let us denote

$$K = \sum_{i=1}^{N} w_i g\left(\frac{\left(y_0^1 - x_i^1\right)^2}{a^2 h_0^2} + \frac{\left(y_0^2 - x_i^2\right)^2}{b^2 h_0^2}\right) \tag{12}$$

and

$$\mathbf{m}_{\mathrm{k}}(\widehat{\boldsymbol{y}}, h_0) = \frac{\sum_{i=1}^{N} x_i w_i g\left(\frac{\left(y_0^1 - x_i^1\right)^2}{a^2 h_0^2} + \frac{\left(y_0^2 - x_i^2\right)^2}{b^2 h_0^2}\right)}{K} - \hat{\boldsymbol{y}}_0 \tag{13}$$

When the location was determined, the $\hat{q}$ is updated by

$$\hat{\mathbf{q}}_{\mathbf{new}} = \begin{cases} \hat{\boldsymbol{q}}_{\mathbf{old}} & \text{if } \boldsymbol{\rho}[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}_{\mathbf{old}}] \leq \alpha \\ (1-\lambda)\hat{\boldsymbol{q}}_{\mathbf{old}} + \lambda \hat{\boldsymbol{p}}_{\boldsymbol{u}}(\mathbf{y}) & \text{if } \boldsymbol{\rho}[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}_{\mathbf{old}}] > \alpha \end{cases} \tag{14}$$

## 4 Experiment

Experiments are conducted on sequences from Object Tracking Benchmark2013 (OTB2013) dataset [8]. The sequences in OTB2013 not only suffer the deformation but also have other change such as fast motion, background clutter, motion blur, and so on. So, we selected six sequences from OTB2013 dataset to show the results. We compared the proposed algorithm with conventional and state-of-the-art algorithms which are available as source code. They are conventional mean-shit algorithm [2], ASMS [5], OAB [9], LOT [10], and CSK [11]. The parameters for those algorithms are set default.

Figure 1 shows the result of six trackers in sequences. The score of distance precision (DP) rate, overlap success (OS) rate, and center location error (CLE) can be obtained from Table 1.

In general, the gradient histogram improves performance of MS algorithm for the gradient histogram. The data in Table 1 show that proposed algorithm has a higher score than others in DP, OS, and CLE which means our tracker is better than others. The sequence dog1 suffers the scale change, and results show standard MS failed to track target because it only uses the gray levels to calculate histogram and it lacks template update. But our tracker can deal well with this condition for the adopting of gradient histogram and RBW algorithm. The proposed tracker has a better performance than other trackers in sequence singer1 which suffers illumination variation since the addition of gradient histogram which does not dependent on the current pixel value but the difference of adjacent pixel. What is more we find that our tracker has a great improvement in gray image compared to the colorful image than the ASMS algorithm since the ASMS algorithm only can acquire information in the gray channel which is scanty.

**Table 1.** Scores of six trackers in six sequences

| Sequence name | | Dog1 | Freeman1 | Freeman3 | Jogging1 | Mountainbike | Singer1 |
|---|---|---|---|---|---|---|---|
| Our tracker | **OS** rate (%) | *1* | *0.975* | *0.998* | *0.974* | *0.996* | *0.806* |
| | **DP** rate (%) | *1* | *0.837* | *0.304* | *0.958* | *0.965* | *0.299* |
| | **CLE** (pixel) | *3.18* | *6.23* | *6.57* | *8.1* | *8.36* | 33.39 |
| MS | **OS** rate (%) | 0.224 | 0.423 | 0.598 | 0.15 | 0.75 | 0.14 |
| | **DP** rate (%) | 0.054 | 0.012 | 0.02 | 0.078 | 0.706 | 0.259 |
| | **CLE** (pixel) | 48.46 | 35.02 | 30.6 | 83.93 | 22.33 | 33.95 |
| ASMS | **OS** rate (%) | 0.842 | 0.696 | 0.696 | 0.228 | 0.132 | 0.792 |
| | **DP** rate (%) | 0.841 | 0.387 | 0.304 | 0.218 | 0.132 | 0.299 |
| | **CLE** (pixel) | 9.98 | 19.93 | 37.5 | 93.69 | 238.2 | 33.91 |
| CSK | **OS** rate (%) | 0.926 | 0.371 | 0.813 | 0.231 | 0.996 | 0.521 |
| | **DP** rate (%) | 0.59 | 0.224 | 0.165 | 0.208 | 0.925 | 0.296 |
| | **CLE** (pixel) | 8.91 | 182.64 | 33.94 | 112.91 | 8.49 | 17.45 |
| LOT | **OS** rate (%) | 0.507 | 0.11 | 0.593 | 0.192 | 0.715 | 0.165 |
| | **DP** rate (%) | 0.506 | 0.08 | 0.085 | 0.023 | 0.671 | 0.208 |
| | **CLE** (pixel) | 24.16 | 137.09 | 42.09 | 89.68 | 25.52 | 145.06 |
| OAB | **OS** rate (%) | 0.999 | 0.552 | 0.741 | 0.893 | 0.724 | 0.786 |
| | **DP** rate (%) | 0.583 | 0.236 | 0.122 | 0.827 | 0.684 | 0.248 |
| | **CLE** (pixel) | 7.36 | 18.09 | 51.7 | 13.9 | 13.85 | *15.46* |

(The best results are in italics)

## 5 Conclusion

In this paper, an adaptive scale mean-shift algorithm with gradient histogram has been proposed to improve the tracking performance of MS-like algorithms. The gradient histogram is constructed by color histogram and gradient feature calculated by Canny edge detector. To deal with the scale change, the RBW algorithm is adopted. Template update is used to cope with appearance variation when the Bhattacharyya coefficient between the current frame and the template is greater than the threshold. The setting of the threshold makes tracker more robust for incorrect information. The proposed tracker is compared with a lot of algorithms in OTB2013 dataset. The experiment results show the tracker's effectiveness in deformation, rotation, scale change, and illumination variation. Moreover, our tracker has a better preference in gray sequences than conventional MS-like algorithms.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (Grant No. 61501139) and the Natural Scientific Research Innovation Foundation in Harbin Institute of Technology (HIT.NSRIF.2013136).

## References

1. Fukunaga K, Hostetler L. The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Trans Inf Theor. 1975;21(1):32–40.
2. Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift. In: Proceedings IEEE conference on computer vision and pattern recognition, 2000. IEEE; 2000, vol. 2. p. 142–9.
3. Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking. IEEE Trans Pattern Anal Mach Intell. 2003;25(5):564–77.
4. Collins RT. Mean-shift blob tracking through scale space. In: Proceedings IEEE computer society conference on computer vision and pattern recognition, 2003. IEEE; 2003, vol. 2. p. II-234.
5. Vojir T, Noskova J, Matas J. Robust scale-adaptive mean-shift for tracking. Pattern Recogn Lett. 2014;49:250–8.
6. Canny J. A computational approach to edge detection. In: Readings in computer vision; 1987. p. 184–203.
7. Sobel I. An isotropic $3 \times 3$ image gradient operator. In: Machine vision for three-dimensional scenes; 1990. p. 376–9.
8. Wu Y, Lim J, Yang MH. Online object tracking: A benchmark. In: 2013 IEEE conference on computer vision and pattern recognition (CVPR). IEEE; 2013. p. 2411–8.
9. Grabner H, Grabner M, Bischof H. Real-time tracking via on-line boosting. Bmvc; 2006, vol. 1, no. 5. p. 6.
10. Oron S, Bar-Hillel A, Levi D, et al. Locally orderless tracking. Int J Comput Vision. 2015;111(2):213–28.
11. Henriques JF, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels. In: European conference on computer vision. Springer, Berlin, Heidelberg; 2012. p. 702–15.

# Improved Performance of CDL Algorithm Using DDELM-AE and AK-SVD

Xiulan Yu(✉), Junwei Mao, Chenquan Gan, and Zufan Zhang

Chongqing Key Labs of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
yuxl@cqupt.edu.cn, mjwlmt@163.com

**Abstract.** Due to the poor robustness and high complexity of the concentrated dictionary learning (CDL) algorithm, this paper addresses these issues using denoising deep extreme learning machine based on autoencoder (DDELM-AE) and approximate $k$ singular value decomposition (AK-SVD). Different from the CDL algorithm, on input, DDELM-AE is added for enhancing denoising ability and AK-SVD replaces K-SVD for improving running speed. Additionally, experimental results show that the improved algorithm is more efficient than the original CDL algorithm in terms of running time, denoising ability, and stability.

**Keywords:** Signal compression · CDL · Deep learning · DDELM-AE · AK-SVD

## 1 Introduction

With the advent of the era of big data, increasing the storage space or transmission path will bring a great burden on communication system. To meet these challenges, signal compression algorithm plays an important role in the field of signal processing.

Signal sparsity has gained considerable interests, especially since Aharon et al. [1,2] introduced the K-SVD algorithm (e.g., see Refs. [3–6]). In [3], an online dictionary learning algorithm was proposed and widely used in hyperspectral images. In [4], the inner product of signals in sparse domain was preserved by linear mapping so that the compression of original signal could be completed. The important conclusion drawn from [5] is that the features of low-dimensional space preserve the main features of original signal. The emergence of compressed sensing (CS) made that the original signal can be recovered from the low-frequency sampled signal [6]. However, most algorithms based on CS

Junwei Mao, M.S., Chongqing University of Posts and Telecommunications. His current main research interest includes: signal processing and deep learning.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_105

must obey the restricted isometry property (RIP) constraint, which poses a huge challenge for signal compression.

In 2016, Silong et al. [7] proposed a nonlinear dimension reduction algorithm based on K-SVD, namely the CDL algorithm, which preserves the geometric characteristics of the original signal and overcomes the limitations of RIP. However, when the data size is too large, the SVD in the iterative process will bring a great burden on the running speed of the algorithm itself. Thus, in this paper, the DDELM-AE neural network [8] and AK-SVD algorithm [9] are considered to reduce computational complexity and improve denoising ability of the CDL algorithm.

First, the original signal enters the DDELM-AE neural network, which can remove useless information and get a better denoising representation. Second, put the denoising representation as an input of the AK-SVD algorithm, then the denoising dictionary can be obtained. Finally, simulation results reveal that the improved algorithm not only improves the running speed, but also enhances the denoising ability and stability of the system.

The organization of the rest of this paper is as follows: Sect. 2 presents the improved CDL algorithm in detail. In Sect. 3, experimental results are explained. Finally, Sect. 4 summarizes our work.

## 2 Improved CDL Algorithm

The main idea of CDL algorithm is based on the conclusion drawn in [10,11] that for original signal $X$ and compressed signal $Y$, the high-dimensional dictionary $D$ and the low-dimensional dictionary $P$ share a sparse representation coefficient $\alpha$. Then, $\alpha$ becomes a bridge between $X$ and $Y$:

$$\begin{cases} X = D\alpha + \varepsilon_1, \\ Y = P\alpha + \varepsilon_2, \end{cases} \tag{1}$$

where $\varepsilon_1 \in R^m, \varepsilon_2 \in R^d$ are noise. Therefore, the derivation of $P$ becomes an important goal of this paper.

### 2.1 Improving the Denoising Ability of the CDL Algorithm.

Similar to [12], the number of each output layer in DDELM-AE network (see Fig. 1) equals to its input layer number. In this way, a higher level feature representation $X'$ of the signal can be obtained. Furthermore, take the $X'$ as AK-SVD input to generate a denoising dictionary, then the performance of AK-SVD algorithm can be improved.

In particular, when the hidden layer number of adjacent two nodes in the model is equal, orthogonal procrustes can be applied to solve the weight. The closed form solution can be obtained by finding the nearest orthogonal matrix

$$M = H^T X, \tag{2}$$

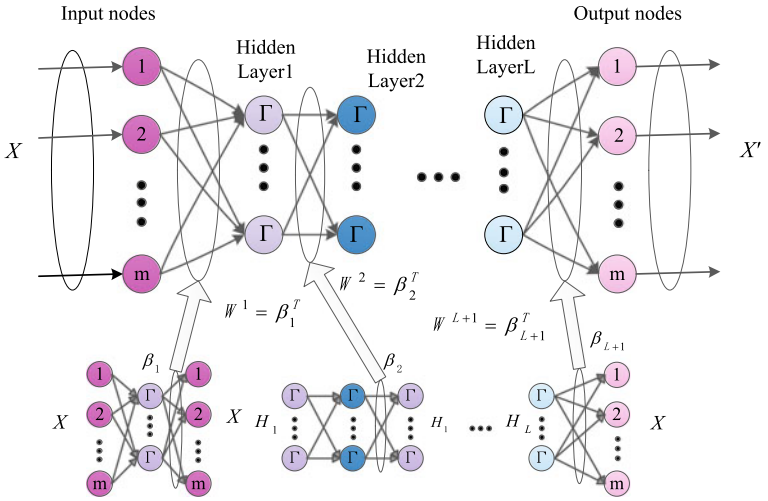



**Fig. 1.** A denoising deep ELM-AE model for the samples of training set $X$

where $H$ is hidden layer output matrix. To find the orthogonal weight matrix $\hat{\beta}$, the SVD of $M$:

$$M = USV^T \tag{3}$$

can be calculated. Thus, $\hat{\beta}$ can be computed by

$$\hat{\beta} = UV^T. \tag{4}$$

### 2.2 Updating Dictionary with Denoising Deep ELM-AE

Put the output signal $X'$ of the DDELM-AE as an input of AK-SVD algorithm and note that the sparse representation:

$$X' = DC, \tag{5}$$

and

$$D = X'C^T, \tag{6}$$

where $D = \{d_1, d_2, \ldots, d_k\}$ and $C = \{c_1, c_2, \ldots, c_n\}$ are sparse coefficient matrices. Additionally, in the K-SVD algorithm, the objective function:

$$\begin{aligned} \|X' - DC\|_F^2 &= \left\| X' - \sum_{i=1}^{n} d_{(:,i)} c_{(i,:)} \right\|_F^2 \\ &= \left\| R^{(l)} - d_{(:,l)} c_{(l,:)} \right\|_F^2 \end{aligned} \tag{7}$$

does not need an exact solution. Thus, the SVD process can be replaced by the formulas (8)–(10) in K-SVD algorithm [9], which greatly reduces the computational complexity and improves the running speed of the CDL algorithm.

$$d_l \leftarrow X' c_i^T - \left( \sum_{i \neq l} d_l c_i \right) c_i^T, \tag{8}$$

$$d_l \leftarrow d_l / \|d_l\|_2, \tag{9}$$

$$c_i \leftarrow X'^T d_l - \left( \sum_{i \neq l} d_l c_i \right)^T d_l. \tag{10}$$

The specific steps of the improved CDL algorithm are shown in Algorithm 1.

**Algorithm 1.** The improved CDL algorithm

**Input:** Signal $X = \{x_1, x_2, \ldots, x_n\}$, number of iterations $m$
**Output:** Trained dictionary $D$ and $P$

Init: Set $D = D_0$, $X$ as input to DDELM-AE, is trained to get $X'$
for $n = 1, 2, \ldots, m$ do
  $\forall i$: $\alpha_i \leftarrow \min \alpha_i \frac{1}{2} \|x'_i - D\alpha_i\|_2^2 + \lambda \|\alpha_i\|_1$
  $D^T D$ singular value decomposition: $D^T D = u \Lambda v^T$
  Updating dictionary $D$: $D' = \Gamma(D)$
  for $j = 1, 2, \ldots, L$ do
    $D_j = 0$
    $I =$ indices of the signals in $X'$ whose representations use $d_j$
    $g = \alpha_{j,I}^T$
    $d = X'_{Ig} - D\alpha_{Ig}$
    $d = d / \|d\|_2$
    $g = X_I'^T d - (D\alpha_I)^T d$
    $D_j = d$
    $\alpha_{j,I} = g^T$
  end if
end for
Dictionary $D$ singular value decomposition: $D = U \Theta V^T$
low-dimensional dictionary $P = U^T D$

## 3 Experimental Results and Analysis

### 3.1 Complexity Analysis

Under the same noise condition, taking Lena and Boat as the experimental objects, the difference between the original algorithm and the improved algorithm is test at the compression ratio (CR) of 16. The denoising ability of the


**Fig. 2.** Reconstructed images of Lena: **a** original image, **b** noisy image, **c** K-SVD, **d** DAK-SVD ($\sigma = 10, CR = 16$)

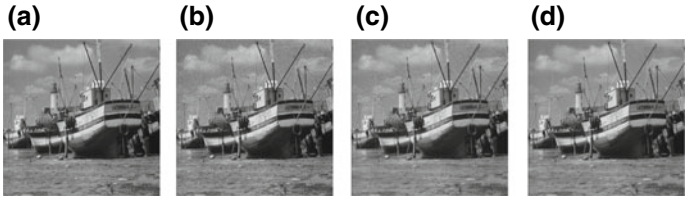



**Fig. 3.** Reconstructed images of Boat: **a** original image, **b** noisy image, **c** K-SVD, **d** DAK-SVD ($\sigma = 10, CR = 16$)

AK-SVD algorithm and K-SVD are used to compress and reconstruct the noisy image. The denoising effect is shown in Figs. 2 and 3.

Table 1 shows that the performance of the DAK-SVD algorithm is better than the K-SVD algorithm, according to the PSNR and TIME standard.

**Table 1.** Comparison of reconstructed images

| Name of the image | Parameters | K-SVD | Proposed system |
|---|---|---|---|
| Lena | PSNR | 31.5358 | 32.3233 |
| | TIME | 96.3186 | 73.0783 |
| Boat | PSNR | 29.7238 | 30.6521 |
| | TIME | 135.1286 | 103.3922 |

### 3.2 Stability Analysis

In the experiment, the original algorithm and the improved algorithm are tested under no noise condition and adding noise with standard deviations of 10, 20, 40 (Figs. 4, 5, 6 and 7) condition, respectively. Using PSNR to judge the quality of the decompressed image.

Table 2 reveals that in no noise condition, CDL compression performance is better than the proposed algorithm under different CR. However, after adding different noise, as shown from Figs. 5, 6, 7 and Table 3, the proposed algorithm

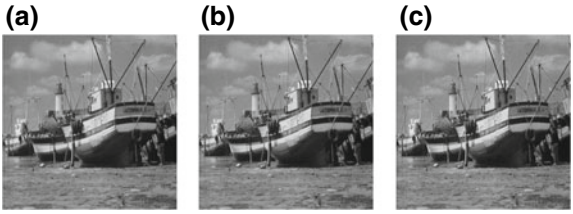



**Fig. 4.** Reconstructed images of Boat: **a** original image, **b** CDL, **c** proposed algorithm (no noise, $CR = 16$)

is more effective than the CDL algorithm in the same compression ratio. Due to the denoising processing of the deep learning model, the denoising ability of the original algorithm and the anti-interference ability are improved, which increases the system robustness.

**Table 2.** Comparison of reconstructed Boat (no noise)

| CR | | 4 | 16 | 32 | 64 |
|---|---|---|---|---|---|
| PSNR | CDL | 35.22 | 34.26 | 30.56 | 28.23 |
| | Proposed system | 34.46 | 33.85 | 28.63 | 25.98 |

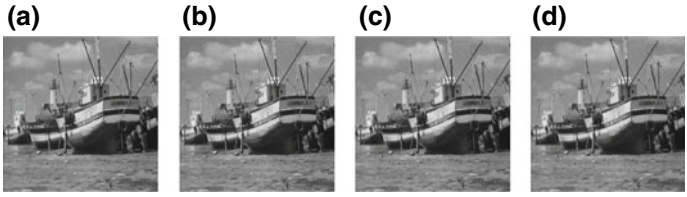



**Fig. 5.** Reconstructed images of Boat: **a** original image, **b** noisy image, **c** CDL, **d** proposed algorithm ($\sigma = 10, CR = 16$)

Table 3 and Fig. 8 indicate that, in the same compression ratio, the improved algorithm is more obvious advantage in the noise environment.

## 4 Conclusions

In this paper, the performance of CDL algorithm has been improved. The denoising ability has been enhanced by adding DDELM-AE on input, and the running speed has been improved by using AK-SVD. Additionally, experimental results have indicated that the improved algorithm is faster and more stable than the original algorithm under noisy environment.

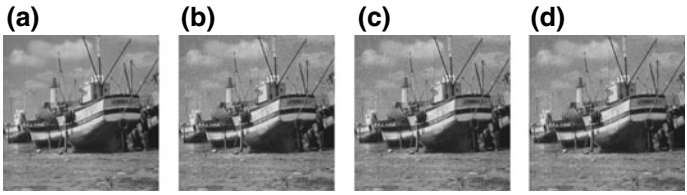



**Fig. 6.** Reconstructed images of Boat: **a** original image, **b** noisy image, **c** CDL, **d** proposed algorithm ($\sigma = 20, CR = 16$)

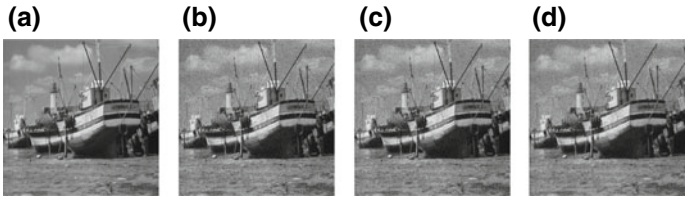



**Fig. 7.** Reconstructed images of Boat: **a** original image, **b** noisy image, **c** CDL, **d** proposed algorithm ($\sigma = 40, CR = 16$)

**Table 3.** Comparison of reconstructed Boat ($CR = 16$)

| Noise | | 10 | 20 | 40 |
|---|---|---|---|---|
| PSNR | CDL | 31.53 | 29.52 | 20.50 |
| | Proposed system | 32.75 | 31.44 | 25.88 |

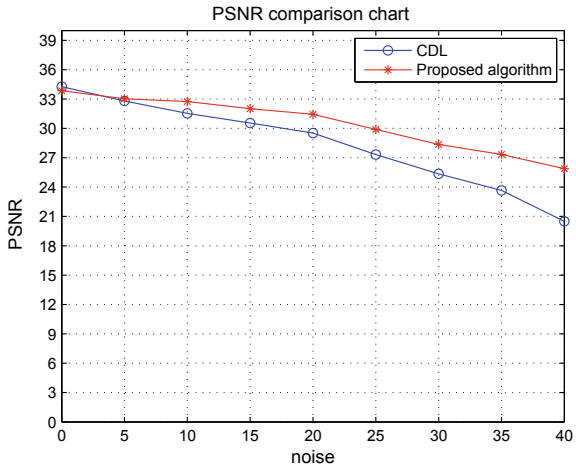



**Fig. 8.** Comparison of PSNR for CDL and proposed model

**Acknowledgments.** This work is supported by Natural Science Foundation of China (Grant No. 61702066), Scientific and Technological Research Program of Chongqing Municipal Education Commission (Grant No. KJ1704080), and Chongqing Research Program of Basic Research and Frontier Technology (Grant No. cstc2017jcyjAX0256).

## References

1. Aharon M, Elad M, Bruckstein A. K-SVD: design of dictionaries for sparse representation. In: Proceedings of the workshop on signal processing with adaptive sparse structured representations (SPARS05); 2005. p. 9–12.
2. Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Image Process. 2006;54(11):4311–22.
3. Wei F, Shutao L, Leyuan F, Benediktsson JA. Contextual online dictionary learning for hyperspectral image classification. IEEE Trans Geosci Remote Sens. 2017;1–12.
4. Gkioulekas IA, Zickler T. Dimensionality reduction using the sparse linear model. In: Proceedings of the 2011 advances in neural information processing systems; 2011. p. 271–9.
5. Calderbank R, Jafarpour S, Schapire R. Compressed learning: universal sparse dimensionality reduction and learning in the measurement domain. Princeton University, USA; 2009. Technical Report.
6. Donoho DL. Compressed sensing. IEEE Trans Inf Theor. 2006;52(4):1289–306.
7. SiLong Z, YuanXiang L, Xian W, XiShuai P. Nonlinear dimensionality reduction based on dictionary learning. In: ACTA AUTOMATICA SINICA; 2016. p. 1065–76.
8. Xiangyi C, Huaping L, Xinying X, Fuchun S. Denoising deep extreme learning machine for sparse representation. Memetic Comp. 2017;9(3):199–212.
9. Shuting C, Shaojia W, Binling L, Daolin H, Simin Y, Shuqiong X. A dictionary-learning algorithm based on method of optimal directions and approximate K-SVD. In: 2016 35th Chinese control conference (CCC); 2016. p. 6957–61.
10. Zeyde R, Elad M, Protter M. On single image scale-upusingsparse-representations. In: Proceedings of the 7th international conference on curves and surfaces; 2012. p. 711–30.
11. Jianchao Y, Wright J, Thoams H, Yi M. Image super resolution via sparse representation. IEEE Trans Image Process. 2010;19(11):2861–73.
12. Vincent P, Larochelle H, Lajoie I, Bengio Y. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J Mach Learn Res. 2010;11(6):3371–408.

# Body Gestures Recognition Based on CNN-ELM Using Wi-Fi Long Preamble

Xuan Xie, We Guo, and Ting Jiang(✉)

Key Laboratory of Universal Wireless Communication,
Beijing University of Posts and Telecommunications, Beijing, China
{xiexvan,tjiang}@bupt.edu.cn

**Abstract.** Recently, researchers around the world have been striving to develop human–computer interaction systems. Especially, neither special devices nor vision-based activity monitoring in home environment has become increasingly important,and has had the potential to support a broad array of applications. This paper presents a novel human dynamic gesture recognition system using Wi-Fi signals. Our system leverages wireless signals to enable activity identification at home. In this paper, we present a novel Wi-Fi-based body gestures recognition model by leveraging the fluctuation trends in the channel of Wi-Fi signals caused by human motions. We extract these effects by analyzing the long training symbols in communication system. USRP-N210s are leveraged to set up our test platform, and 802.11a protocol is adopted to implement body gestures recognition system. Besides, we design a novel and agile segmentation algorithm to reveal the specific pattern and detect the duration of the body motions. Considering the superiority of feature extraction, convolutional neutral networks (CNN) is adopted to extract gesture features, and extreme learning machine (ELM) is selected as classifier. This system is implemented and tested in ordinary home scenario. The result shows that our system can differentiate gestures with high accuracy.

**Keywords:** Body gesture recognition · CNN · ELM · Long training sequence · USRP · 802.11a · Wi-Fi

## 1 Introduction

Wearable devices basis of the sensor have gradually integrated into our daily life. Activity monitor in home environment like elder care and child safety most are implemented in video surveillance. These applications have been accepted by the masses as a part of life. While it is not pleasant to wear some sensors or be kept on an electronic eye each moment even though the surveillance is well

This work was supported by the National Natural Sciences Foundation of China (NSFC) (No. 61671075) and Major Program of National Natural Science Foundation of China (No. 61631003).


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_106

intentioned. Naturally, how to implement above function and make less adverse difference are worthy for research. Since wireless signals do not require line-of-sight and can traverse through walls, and some wireless signals such as Wi-Fi are widely used in home environment, suppose we can realize activity monitor or gesture recognition by wireless signals, user experience could be boosted greatly. Some wireless signals-based human activity recognition and localization systems have been proposed. WiSee [1] recognizes body gesture by extracting the micro-doppler shift from Wi-Fi signals. As we know, people makes an action, which results in a pattern of Doppler shifts at the wireless receiver, they proposed algorithms to extract Doppler shift from OFDM symbols. WiTrack [2] tracks the 3D motion of a user through the wall from the radio signals reflected off the body. They measures time it takes for its signal to travel from its transmit antenna to the reflecting body, and then back to each of its receive antennas. They leverage the geometric placement of its antennas to localize the moving body in 3D. WiHear [3] detects and analyzes fine-grained radio reflections from mouth movements by MIMO technology to recognize spoken words, they use PHY layer channel state information (CSI) on WLAN device to analyze micro-motion, then use multipath effect and discrete wavelet packet transformation to achieve lip reading with Wi-Fi. Besides, as we know, different gestures between transmitter and receiver have different effects on received signals, [4] extract the long preamble of Wi-Fi signals, and adopt the traditional improved classification arithmetic to classify and recognize different hands gestures.

This paper fully analyze the frame structure of IEEE 802.11a WLAN protocol. Without effecting communication quality, we recognize the human gestures by extracting and analyzing the long training sequence of Wi-Fi frame. In the process of communication, data is a kind of stochastic series, specific transmitting content have a decisive impact on waveforms. While, preamble sequence, short and long included, which are responsible for frame detection, frequency offset correction and symbol alignment, is fixed in different data frame. Inside, short training sequence, only two to three are available in signal detection [5], thus, the long training sequence is implemented to recognize the human motion. Long training sequence is used as frequency offset correction and symbol alignment in communication, which is composed of a fixed length pattern [6].

Our gesture recognition is omplemented in following steps. Firstly, long training sequence is extracted from wireless signals (IEEE 802.11a), the 802.11a system that [7] provided are adopted to extract long training symbols. Secondly, continuous periodic activities product, which make periodic influence on long preamble in communication system. A method is proposed to make sample extraction. Thirdly, features are extracted from the data by convolutional neutral networks (CNN). Finally, Extreme Learning Machine (ELM) is leveraged to classify and recognize different body gestures.

This paper was organized as follows. In Sect. 2, we introduce the features extraction algorithm and classifier. In Sect. 3, a brief introduction on experiment platform and scenario is provided, and the method that extracts long preamble is introduced. Then, we describe how to test fluctuation caused by body gestures.

In Sect. 4, we introduce our experiment and analyze the result. Finally, a conclusion is drawn in Sect. 5.

## 2 Recognition Approach

### 2.1 Convolutional Neutral Networks

Convolutional neutral networks (CNN) were first presented by Kunihiko Fushima; in recent years, CNNs fall in the category of deep learning techniques, which have been employed in many other fields such as speech recognition and images recognition [8,9]. In this paper, instead of learning an ad-hoc CNN for the body gestures recognition problem, CNN is used to choose the most informative features and the best combination of features. In our experiment, we can find that features extracted by CNN do better in our body gestures recognition than normal features selected manually.

CNN takes an input data, processes it with a series of different transformation layers, finally produce a prediction of the class. The transformation layers mainly contain three classes: input layer, hidden layer, and output layer, and convolutional layers are the main type of the hidden layers in CNNs. These are designed to perform a convolutional operation on the data. Besides, pooling layers are used between adjacent layers, which is average or maximum filters, they can reduce the impact of small variations and descend the dimension of previous layer output data. Dropout layers are used to fix overfitting, which could occur when the number of training data and CNN parameters are mismatching. When the CNNs are set, training data are used to train the net iteratively and the weight and bias of every neuron are adjusted until the loss function is fitted. The CNN network structure of this paper is Fig. 1.

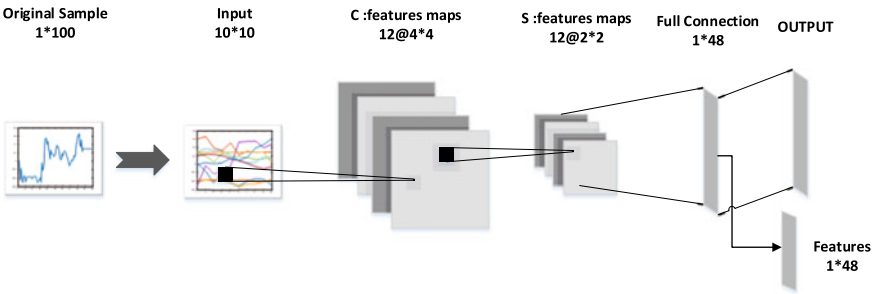



**Fig. 1.** CNN network structure.

**Extreme Learning Machine** Extreme learning machine (ELM) [10] is a kind of learning algorithm for the single-hidden-layer feed-forward neural networks. Comparing with the deep learning algorithm, it overcomes the slow training

speed and overfitting problems. And ELM also has milder optimization constraints than support vector machine (SVM). ELM is based on empirical risk minimization theory and its learning process needs only a single iteration. The algorithm avoids multiple iterations and local minimization. Especially in small dataset, due to better generalization ability and fast learning rate, it is used in various fields. We adopt the ELM in our recognition.

ELM is initially developed for a single-hidden-layer feed-forward neural networks (SLFNs). Given an input sample ($\mathrm{X}_i$,$\mathrm{y}_i$), where $\mathrm{X}_i = [\mathrm{x}_{i1}, \mathrm{x}_{i2}, \mathrm{x}_{i3}, \ldots, \mathrm{x}_{in}]^T \in R^n$ , $\mathrm{y}_i = [\mathrm{y}_{i1}, \mathrm{y}_{i2}, \mathrm{y}_{i3}, \ldots, \mathrm{y}_{im}]^T \in \mathrm{R}^m$. The output of ELM with L hidden nodes is designed as

$$\sum_{i=1}^{L} \beta_i g(\omega_i \cdot X_j + b_j) = o_j, \quad j = 1, \ldots, N \tag{1}$$

where $\beta_i$ is the hidden-layer bias, $\omega_i$ is the input weight of the ith neuron, g(x) is the activation function, and $\mathrm{o}_j$ is the output of hidden layer. $\omega_i \cdot \mathrm{X}_j$ means the inner product of $\omega_i$ and $\mathrm{X}_j$. In case of ELM approximates the data perfectly for all samples, we have

$$\sum_{j=1}^{L} \|o_j - y_j\| = 0 \tag{2}$$

that means existing $\beta_i$, $\omega_i$, and $\mathrm{b}_j$ such that

$$\sum_{i=1}^{L} \beta_i g(\omega_i \cdot X_j + b_j) = y_j, j = 1, \ldots, N \tag{3}$$

Denoted by matrix

$$H\beta = Y \tag{4}$$

where

$$\mathrm{Y} = (\mathrm{y}_1, \mathrm{y}_2, \mathrm{y}_3, \ldots, \mathrm{y}_N)^T$$
$$\beta = (\beta_1, \beta_2, \beta_3, \ldots, \beta_L)^T$$

and H is the hidden-layer output matrix:

$$A = \begin{bmatrix} g(\omega_1 \cdot X_1 + b_1) & \cdots & g(\omega_L \cdot X_1 + b_L) \\ \vdots & \cdots & \vdots \\ g(\omega_1 \cdot X_N + b_1) & \cdots & g(\omega_L \cdot X_N + b_L) \end{bmatrix} \tag{5}$$

In the process of training, ELM aims at $\widehat{\omega_i}$, $\widehat{b_i}$, and $\widehat{\beta_i}$, which minimum the training error:

$$\|H(\widehat{\omega_i}, \widehat{b_i})\widehat{\beta_i}\| = \max_{\omega,\beta,b} \|H(\omega_i, b_i) - Y\| \tag{6}$$

Which is equivalent to minimizing the loss function

$$E = \sum_{j=1}^{L} \left( \sum_{i=1}^{L} \beta_i g(\omega_i \cdot X_j + b_j) - y_j \right)^2 \tag{7}$$

According to ELM theory, input weights $\omega$ and b can be randomly assigned, and $\beta = H^*Y$ where $H^*$ is the Moore–Penrose of H.

## 3 Data Processing

**Experiment Scenario** We adopt a complete 802.11a communication system [7], which is implemented based on GNU Radio and fitted for operation on an Ettus USRP N210. USRP N210 is a kind of generic software defined radios (SDRs). We set the packet transmission rate to 50 packets per second to obtain sufficient information from body gestures.
We set up our system in a meeting room, as depicted in Fig. 2. The testing area is 6 m × 6 m. And the volunteer performs different gestures in room.
The distance between transmitter and receiver is 4.0 m. The center frequency of the wireless transmission is 5.3 GHz.

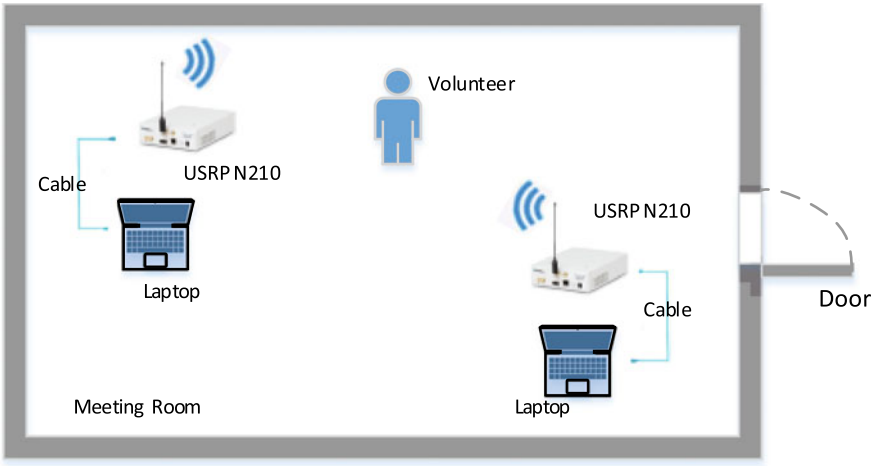



**Fig. 2.** Experiments environment and setting

**Long Symbols Extraction** The IEEE802.11a WLAN is a packet communication system [11]. Frame structure [6] is shown in Fig. 3, which consists of two parts, training sequence and OFDM Symbols; the training sequence contains designated short preamble and long preamble. Short preamble, a pattern that spans 16 samples and repeats ten times, which is used to detect the frame, and long preamble is used as frequency offset correction and symbol alignment, which is a pattern repeats 2.5 times, one of the purposes of the long preamble is symbol alignment, the intention of this process is detecting the start of signal. According to Fig. 3, when the start of signal is found, the long preamble can be extracted. The sample points of long preamble are related to the duration of long preamble and the sample rate:

$$N_{LP} = N_S \times T_S \tag{8}$$

where $N_{LP}$ is the sample points of one long preamble and $N_S$ is the sample rate; in this paper, we set $N_S$ to 20 MHz, $T_s$ is the length of long preamble patterns (two repeated sequence). Seen from Fig. 3, the length of long preamble is 8 μs, which contains a pattern that repeats 2.5 times. It is easy to get $T_s$, that equals to 6.4 μs. Each frame, we sample 128 points. Next, we talk about how to divide the continuous data into correct cycle while ignoring the minute error caused by participant in periodic action. In the process of data acquisition, we keep the normal communication,and transmit text. Meanwhile, people do periodic activities shown in Fig. 4. Obviously, periodic gestures cause cyclical effects on signals. Ideally, we can divide the serial data in fixed period. While, as we known, it is not able to avoid that some random difference or minute errors could be superposed in data. So proper method of period division produces vital influence on recognition performance.

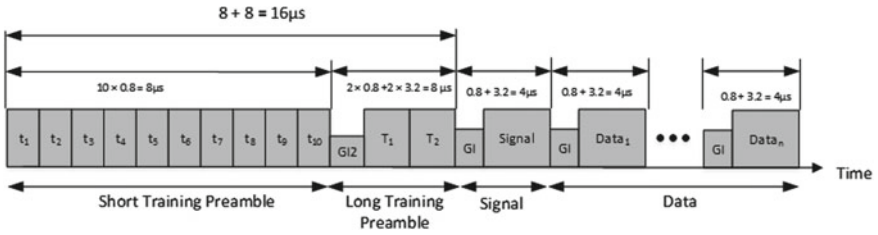



**Fig. 3.** OFDM training structure

**Data Preprocessing** Consider the long preamble $X_1, X_2, X_3, \ldots, X_i, \ldots, X_{n-2}, X_{n-1}, X_N$, where $X_i$ is a column vector of $1 \times N_{LP}$ that samples from long preamble of data frames, and N is the number of frame of one group body gestures (different gestures may have different N), we calculate the mean of $X_i$. Say $x_i$ denotes the mean power of the $X_i$, then, let $\bar{X}_i$ ($i = 1, 2, 3, \ldots, N-2, N-1, N$) be the new vector formed by $x_i$:

$$\bar{X}_i = [x_1, x_2, x_3, \ldots, x_i, x_{N-1}, x_N] \tag{9}$$

Figure 5 shows the amplitude spectrum of nine body gestures. Inside, the horizontal axis is the frame index, and the vertical axis is the mean power of $X_i$ (i.e. $\bar{X}_i$). The plots show the nine kinds of periodic body gestures defined in Fig. 4. It is shown that different gestures have completely different effects on the preamble periodically. The cycle of difference depends on the duration of body gesture, and the amplitude and the shape of wave are determined by the location and the range of the body.

**Endpoint Detection** A new method to detect the endpoint is proposed in this paper. Obviously, body gestures have some effects on preamble, so detecting the start and end point of gesture is vital. The steps of detection are in follows:

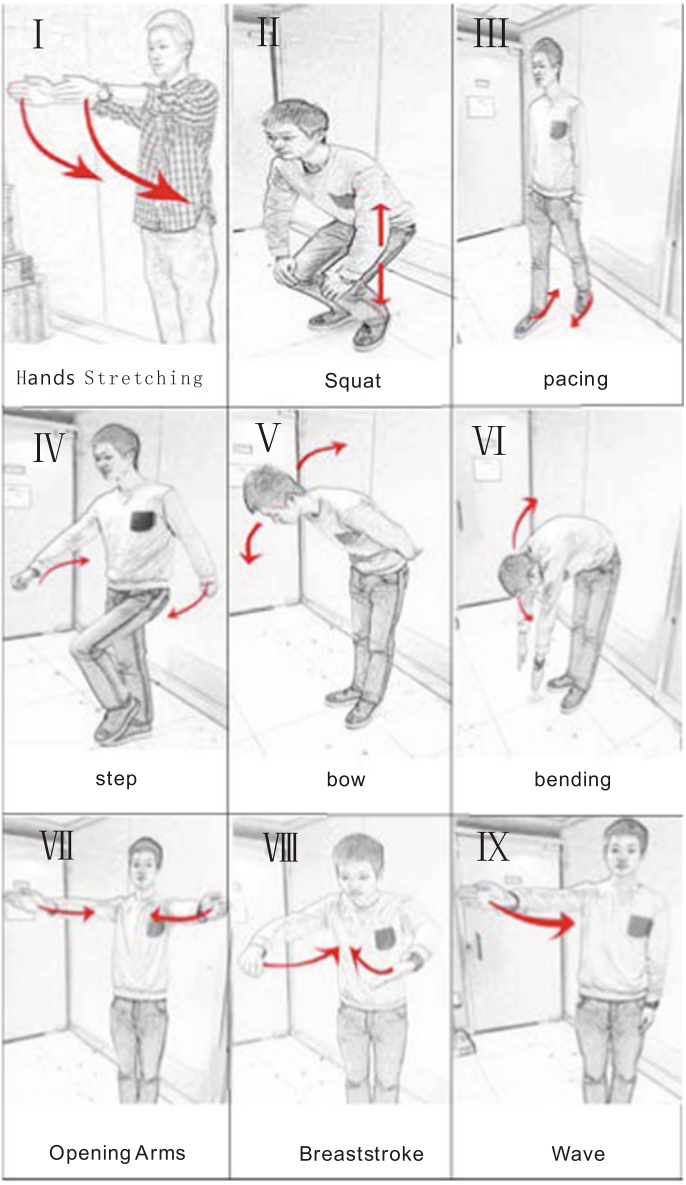



**Fig. 4.** Predefined dynamic body gestures

Step 1. Zero-mean need to be carried out on $\bar{X}_i$.

$$P_i = \bar{X}_i - \bar{X}_i^{mean} \tag{10}$$

where $\mathrm{P}_i(\mathrm{i} = 1, 2, 3, \ldots, \mathrm{N}-2, \mathrm{N}-1, \mathrm{N})$ denotes the zero-mean amplitude of $\bar{X}_i$, ${X_i}^{mean}$ is the mean of the $\bar{X}_i$.

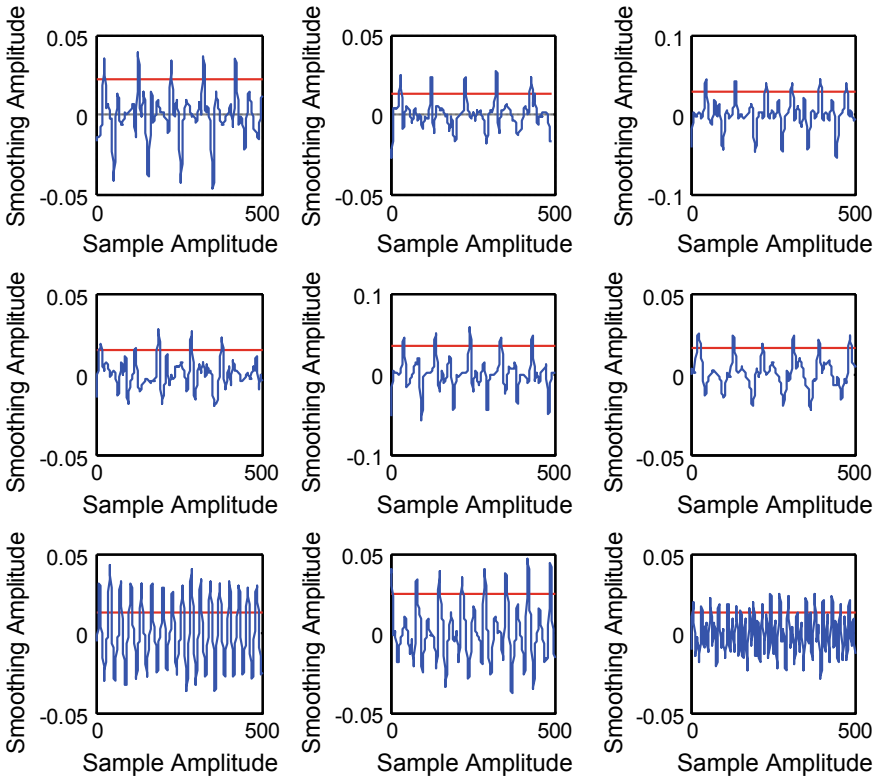



**Fig. 5.** The normalize amplitude spectrum of nine kinds of body gestures

The result is shown in Fig. 5.

Step 2. Make differential operation on $P_i$, and then, we let the differential operation result of the $P_i$ is the $Y_j$

$$Y_i = P_i - P_{i-1} \tag{11}$$

Next, we calculate the ${y_j}^{threshhold}$:

$${y_j}^{threshhold} = {y_j}^{mean} + \beta({y_j}^{max} - {y_j}^{mean}) \tag{12}$$

where j =1, 2, 3, ... ,7 ,8 ,9, which denotes nine kinds of body gestures. ${y_j}^{mean}$ is the mean of $Y_i$, ${y_j}^{max}$ is the max of the $Y_i$. And $\beta$ is the adjustable coefficient to make the cycle division reasonable.

Step 3. Calculate the mean of $y_{i-1}$, $y_i$ and $y_{i+1}$ then replace it with $z_i$, this operation needed to be done to eliminate some singular points (i.e.small peaks).

Step 4. Adjust $\beta$ and find i that meets the following conditions:

$$\begin{cases} z_i - z_{i-1} > 0, z_{i-1} - z_{i-2} > 0, \\ z_i - z_{i+1} > 0, z_{i+1} - z_{i+2} > 0, \\ z_i > {y_i}^{threshhold} \end{cases} \tag{13}$$

where we need adjust $\beta$ on the basis of the actual periods to find reasonable i. Later, we get a new vector $V_j$ consists of the special i that meets above conditions (i.e.,$v_1$ equals eligible i).
Say

$$V_j = [v_1, v_2, v_3, v_l, \ldots, v_{L-2}, v_{L-1}, v_L] \tag{14}$$

where L is the length of the $V_j$. The elements of this vector denotes the start and end point of one cycle. Using $V_j$, divide $\bar{X}_i$ to form a matrix $M_j$. The row of $M_j$ denotes the sample numbers of one body gestures. Acturally it is L − 1, the column of $M_j$ is the period of one body gestures, we set the column is:

$$\max_{1<l<L-1} \|v_{l+1} - v_l\| \tag{15}$$

Here, we get the whole samples of body gestures,which are shown in Fig. 6.

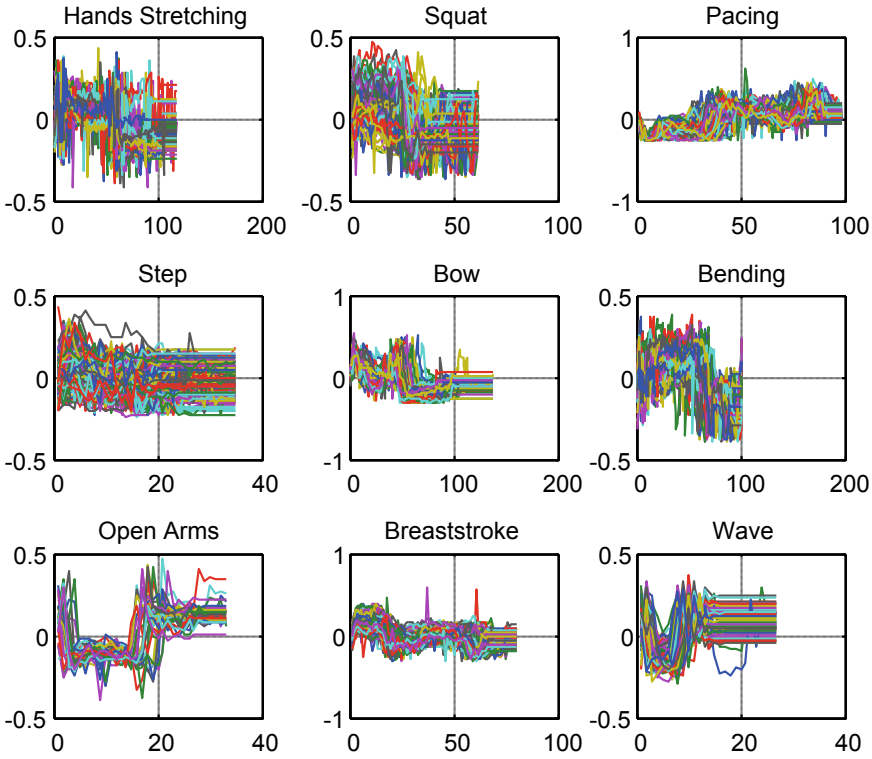



**Fig. 6.** Nine groups of samples, final samples of nine gestures

## 4 Experiment and Analysis

**Experiment** In Sect. 3, we get the human body gestures dataset, which consists of nine kinds of gestures shown in Fig. 4, every gesture contains 100 samples. Obviously, nine gestures have different cycle, considering facilitate subsequent process,we choose the median to make the dimension of all vectors become 100. That means the dataset is a matrix of 900 × 100, which denotes 900 samples. Besides, the label matrix is 900 × 9, is one-hot vector, which denotes the class of sample. The following processing, we use this dataset to train the CNN and ELM. In every process of training and testing, the train dataset is 540 × 100 and test dataset is 360 × 100.

In our experiment, the features produced by CNN in the last layers before the class assignment work are 48-dimensional, which are extracted from each 100-dimensional body gesture data. Considering the data size, our net selected one convolutional layer and one fully connected layer. The CNN was originally trained on nine kinds of body gestures with labels. Features are obtained by extracting activation values of the last hidden layer. The extracted features for each sample are then used as input to ELM for classification. The flow chart of the whole processing is shown in Fig. 7. In the process of training the CNNs, MSE loss function is used:

$$\min C(\widehat{Y}, Y) = \|\widehat{Y}^2 - Y^2\| = \sum_i (\widehat{y} - y)^2 \tag{16}$$

iteration result in train dataset is shown in Fig. 8.

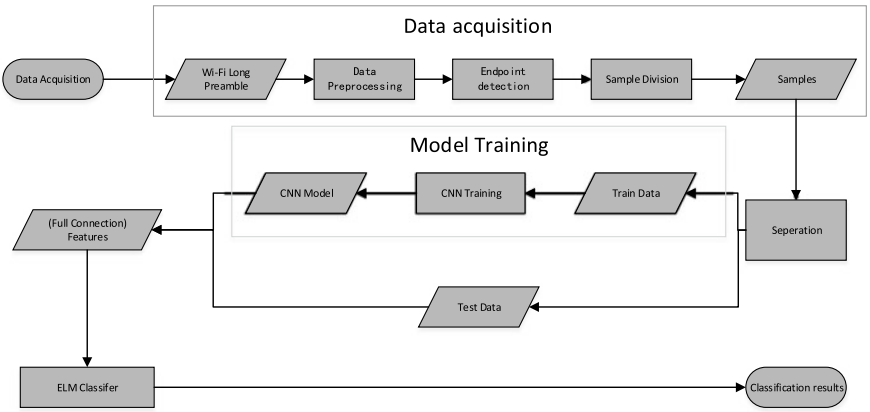



**Fig. 7.** Body gestures recognition framework

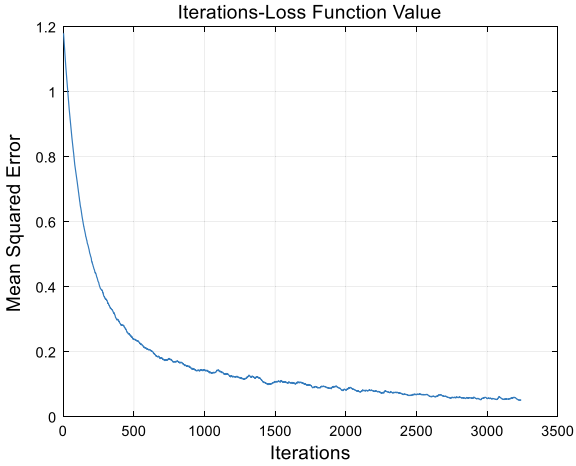



**Fig. 8.** World map

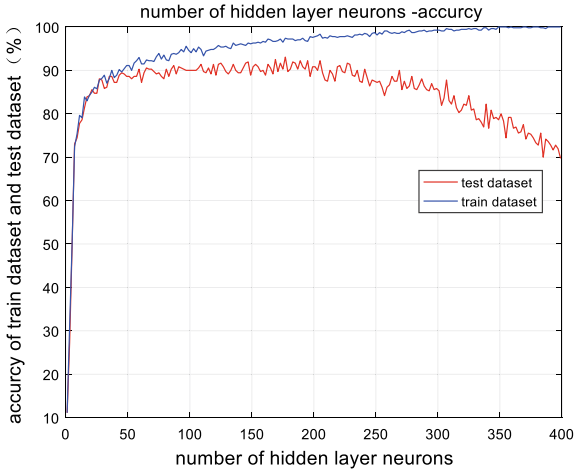



**Fig. 9.** Concrete and constructions

We choose randomly generated hidden nodes, sigmoid function is used as activation. The number of hidden neurons is determined through cross-validation. Figure 9 shows the process of it.

Compared with other classifiers and features extraction method, we also use them in our dataset. In the extraction features, we select ten kinds of statistics features manually [12]: power-delay-profile (PDP), mean excess delay (RMS), RMS delay spread, multipath components (MPCs) within 10 dB, number of MPCs that contain 60% of the received energy, max value of sample, mean of sample, standard deviation, kurtosis, and skewness. And support vector machine (SVM) is applied in our dataset, the result is shown in Table 1.

**Table 1.** Summary of classification accuracy (%) using different feature extraction methods and classifiers

| | Hands stretching | Squat | Pacing | Step | Bow | Bending | Opening arms | Breast stroke | Waves | Mean accuracy |
|---|---|---|---|---|---|---|---|---|---|---|
| Manual&SVM | 75.79 | 81.32 | 86.81 | 89.20 | 91.46 | 88.18 | 97.44 | 83.58 | 92.39 | 87.35 |
| CNN&SVM | 82.50 | 87.50 | 100.00 | 97.50 | 97.50 | 75.00 | 100.00 | 90.00 | 100.00 | 92.22 |
| CNN&ELM | 95.00 | 85.00 | 85.00 | 100.00 | 95.00 | 100.00 | 100.00 | 92.50 | 100.00 | 94.72 |

In Table 1, Manual&SVM means we selected features manually and SVM classifier is adopted. CNN&SVM stands for CNN is used to extract features and SVM classifier act as classifier. The last is what we proposed in this paper. The detailed recognition rates of the nine body gestures in Fig. 4 are shown in Table 2.

**Table 2.** Summary of classification accuracy (%) of different body gestures

| | i | ii | iii | iv | v | vi | vii | viii | ix |
|---|---|---|---|---|---|---|---|---|---|
| i | 95.00 | 0 | 0 | 0 | 0 | 2.50 | 0 | 2.50 | 0 |
| ii | 5.00 | 85.00 | 0 | 5.00 | 0 | 2.50 | 0 | 2.50 | 0 |
| iii | 2.50 | 0 | 85.00 | 0 | 0 | 0 | 10 | 2.50 | 0 |
| iv | 0 | 0 | 0 | 100.00 | 0 | 0 | 0 | 0 | 0 |
| v | 0 | 0 | 0 | 0 | 95.00 | 0 | 0 | 2.50 | 0 |
| vi | 0 | 0 | 0 | 0 | 0 | 100.00 | 0 | 0 | 0 |
| vii | 0 | 0 | 0 | 0 | 0 | 0 | 100.00 | 0 | 0 |
| viii | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 92.50 | 0 |
| ix | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100.00 |

## 5 Conclusion

In this paper, we treated the problem of human body gestures recognition with Wi-Fi signals. Instead of using CSI to analyze the motion, we adopt the long preamble of Wi-Fi signal through the analysis of the communication process. Compared with CSI, long preamble is more primitive, and more features of motion could be collected. Aimed at long preamble, in the data processing stage, we proposed new motion capture and endpoint detection algorithm. And inspired by speech processing and image processing, instead of extracting features manually, CNN is used in this task. In contrast, we select ten statistic features of samples, and use the same classifier SVM, the result shows that the accuracy rate of CNN is higher 4.87%. Besides, in the small dataset, ELM has milder optimization constraints and less computational complexity, we choose ELM as the classifier in the end, we can find from Table 1, when the same feature vector input ELM and SVM, ELM has the better accuracy. The mean accuracy of SVM

is 92.22%, while ELM is 94.75%. Meanwhile for SVM, the lowest recognition is 75%, while ELM is 85%. This means ELM is better than SVM in the body gestures recognition. Moreover, different gestures may have totally distinct effect on signals, which reflects on recognition rate. The nine body gestures we select consist of many parts of daily motions, this could promote the home care on the condition of respect of privacy.

**Acknowledgments.** We thank the anonymous for their thoughtful and constructive remarks that help me improve the quality of this paper.

## References

1. Qifan P, Gupta S, Gollakota S, Patel S. Whole-home gesture recognition using wireless signals. Comput Commun Rev. 2013;43(4):485–6.
2. Adib F, Kabelac Z, Katabi D, Miller RC. 3D tracking via body radio reflections; 2013. p. 317–29.
3. Wang G, Zou Y, Zhou Z, Wu K, Ni LM. We can hear you with wi-fi!. In: International conference on mobile computing and Networking; 2014. p. 593–604.
4. Zhou G, Jiang T, Liu Y, Liu W. Dynamic gesture recognition with wi-fi based on signal processing and machine learning. In: IEEE global conference on signal and information processing; 2015. p. 717–21.
5. Liu CH. On the design of OFDM signal detection algorithms for hardware implementation. In: Global telecommunications conference, 2003. GLOBECOM '03. IEEE; 2003, vol. 2. p. 596–9.
6. IEEE. Part 11: wireless LAN medium access control (mac) and physical layer (phy) specifications. In: the 5-GHz band, IEEE; 2007. p. C1–1184.
7. Bloessl B, Segata M, Sommer C, Dressler F. An IEEE 802.11 a/g/p OFDM receiver for GNU radio; 2013.
8. Sainath TN, Kingsbury B, Saon G, Soltau H, Mohamed AR, Dahl G, Ramabhadran B. Deep convolutional neural networks for large-scale speech tasks. Neural Netw Official J Int Neural Netw Soc. 2015;64:39–48.
9. Ciresan DC, Meier U, Masci J, Gambardella LM, Schmidhuber J. Flexible, high performance convolutional neural networks for image classification. In: IJCAI 2011, Proceedings of the international joint conference on artificial intelligence, Barcelona, Catalonia, Spain, July; 2011. p. 1237–42.
10. Huang GB, Zhu QY, Siew CK. Extreme learning machine: theory and applications. Neurocomputing. 2006;70(1–3):489–501.
11. Sourour E, El-Ghoroury H, Mcneill D. Frequency offset estimation and correction in the IEEE 802.11a wlan. In: Vehicular technology conference, 2004. Vtc2004-Fall. 2004 IEEE; 2005, vol. 7. p. 4923–7.
12. Zhai S, Jiang T. Target detection and classification by measuring and processing bistatic UWB radar signal. Measurement. 2014;47(1):547–57.

# Evaluation of Local Features Using Convolutional Neural Networks for Person Re-Identification

Shuang Liu[1,2(✉)], Xiaolong Hao[1,2], Zhong Zhang[1,2], and Mingzhu Shi[1,2]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
[2] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China
{shuangliu.tjnu,haoxiaolong17,zhong.zhang8848}@gmail.com
shimingzhu1@163.com

**Abstract.** In this paper, we mainly evaluate the influence of local features extracted by convolutional neural networks for person re-identification. Considering the variant body parts with different structural information, we divide the holistic person images into several parts and extract their features. Two kinds of aggregation methods are used to aggregate local features. Experiments on the challenging person re-identification database, Market-1501 database, show that the max aggregation is more effective for extracting the discriminative local features than the sum aggregation.

**Keywords:** Local features · Convolutional neural networks · Person re-identification

## 1 Introduction

Given a specified person image captured from one camera, person re-identification aims to spot the same person captured from other cameras [1]. Person re-identification plays a significant role in practical applications such as video surveillance, image retrieval, and so on [2,3]. Hence, person re-identification becomes gradually popular in the research field. Nevertheless, the performance of person re-identification is affected by many factors, such as different pose, various illumination, and so on [4]. In order to solve these problems, a large number of researchers have devoted themselves to the study of person re-identification.

Recently, the convolutional neural networks display potential for extracting the discriminative features for person re-identification [5,6] and gradually draw more attention in this domain. Up to now, the available methods mainly concentrate on extracting a discriminative features representation [7–9] and learning


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_107

an effective measure method [10–12]. As for the former, global and local features are the mainstream algorithms. Some approaches utilize the whole person image to train the network model in order to obtain global representations. For instance, the method named the domain guided dropout algorithm was proposed by Xiao et al. [13], which utilizes convolutional neural networks to obtain a robust global features from multiple domains. To acquire the high-level discriminative features, Wu et al. [7] used deep convolutional networks and extremely small filters to extract features. Combining the strengths of the identification model and verification model, Zheng et al. [14] proposed a siamese network model to learn discriminative global features using three loss functions. Although the global features extracted the convolutional neural networks are effective for person re-identification, it ignores the local structure information. Therefore, under the guidance of the deep learning methods, some researchers have focused on mining local structure information and have achieved promising results. In order to extract robust local features, Yi et al. [15] split the holistic person images into three partial areas and trained three convolutional neural networks, respectively. Considering the various structure information in different positions, the Spindle Net captured micro-body information and learnt the local features completely. Sun et al. [16] presented a network called part-based convolutional baseline (PCB), which utilizes a simple uniform partition strategy to obtain local features consisting of several part-level features. As for the measure method, some methods, such as the Euclidean distance [11,17], the Mahalanobis distance [18], the cosine distance [12], and so on, are exploited to compute the similarity between person images.

The above-mentioned local features can achieve good results for person re-identification [15,16,19]. In this paper, we mainly focus on evaluating the impact of local features with different aggregation methods for person re-identification. In order to learn local features using the convolutional neural networks, we first split the holistic person image into several parts and allocate a person re-identification label for each part. Then, the output of fully connected layer is regarded as the features vector of each part. After extracting the features for each part, we utilize two kinds of methods to aggregate these local features for each person image. Finally, we evaluate the performance with different aggregation methods on the Market-1501 database [20].

The remain of this article is organized as follows. The detailed implementation is given in Sect. 2. In Sect. 3, we show the experimental results on the Market-1501 database and we make a conclusion in Sect. 4.

## 2 Approach

In this section, we first introduce the framework of convolutional neural networks and then present implementation details. Finally, two kinds of aggregation methods are introduced.

### 2.1 The Framework of Convolutional Neural Networks

There are many convolutional neural networks models, such as CaffeNet [21], VGG16 [22], ResNet-50 [23], and so on. We select ResNet-50 as the convolutional neural networks to fine-tune on our database as shown in Fig. 1. In order to apply the ResNet-50 network on the Market-1501 database, the number of neurons in the original fully connected layer is changed to 751, and the features $f$ extracted from the fully connected layer are a 2048 dimensional vector.

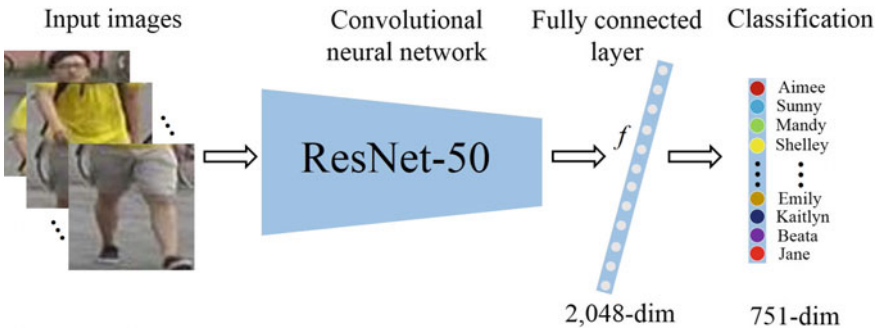



**Fig. 1.** The framework of convolutional neural networks.

### 2.2 Implementation Details

In order to extract the local features, we first divide each person image into three overlapped parts. Each part is then allocated a person re-identification label, which is the same as the original person re-identification label as shown in Fig. 2. Next, we employ these parts with the person re-identification labels to train the convolutional neural networks. After extracting the feature of each part from the fully connected layer, two aggregation approaches, i.e., sum aggregation and max aggregation, are used to aggregate the features of three parts, and the aggregated features are taken as the final features for each person image.

### 2.3 Two Kinds of Aggregation Methods

As shown in Fig. 3, since each person image is divided into three overlapped parts, we can obtain three partial features for each person image, i.e., $f_1$, $f_2$, and $f_3$. Afterward, two alternative methods are employed to aggregate the three partial features.

As for the sum aggregation, it aims to smooth the difference among all the parts. Figure 4 illustrates the fusion process. Using the sum aggregation, the local features can be aggregated into one global features. The aggregated features $f_{sum}$ can be formulated as:

$$f_{sum} = f_1 + f_2 + f_3 \tag{1}$$

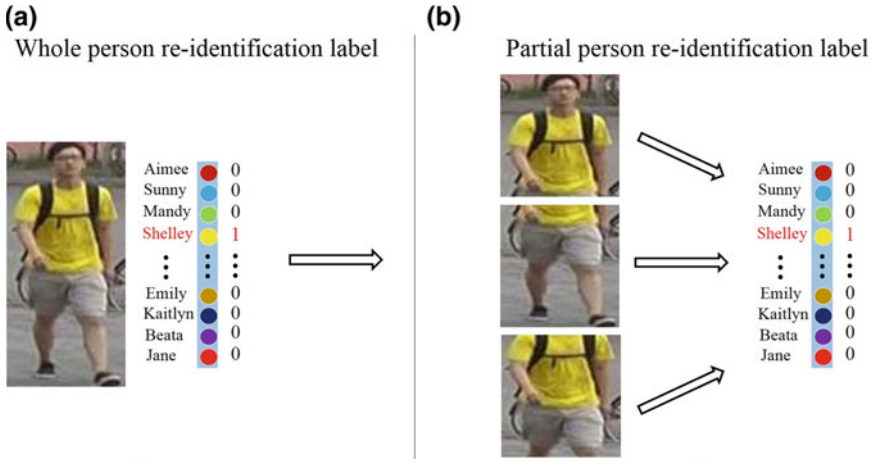



**Fig. 2.** **a** Is the whole person re-identification label, and **b** is the partial person re-identification label.

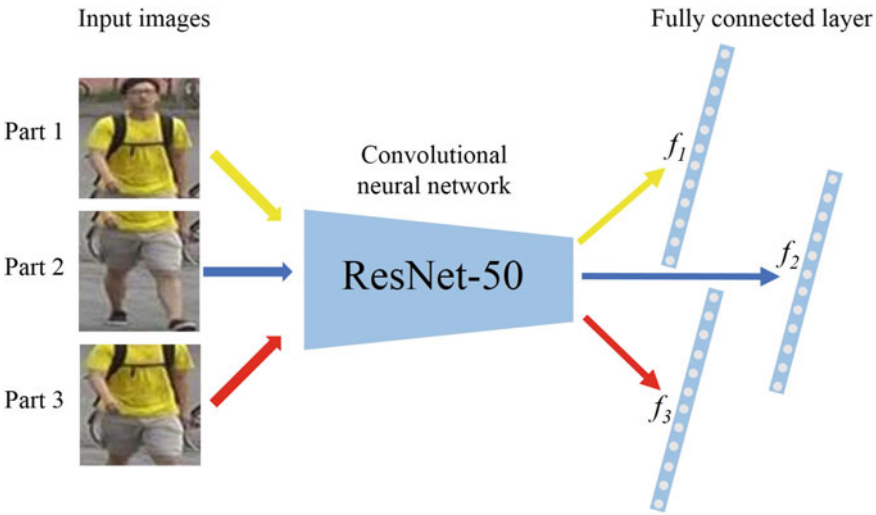



**Fig. 3.** Extracting three partial features for each person image.

where $f_1$, $f_2$, and $f_3$ are the partial features of a person image, respectively.

As shown in Fig. 5, the max aggregation method aims at maximizing the difference of the parts. In other words, the max aggregation method can maximize the detailed structure information for the person image. It can be formulated as:

$$f_{max} = \max(f_1, f_2, f_3) \tag{2}$$

where $\max(\cdot, \cdot, \cdot)$ indicates the maximum of three vectors in each dimension.

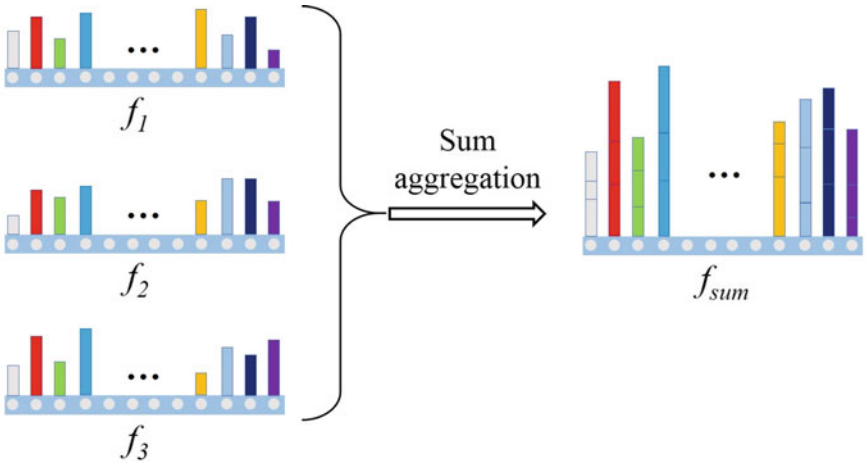



**Fig. 4.** Illustration of the sum aggregation.

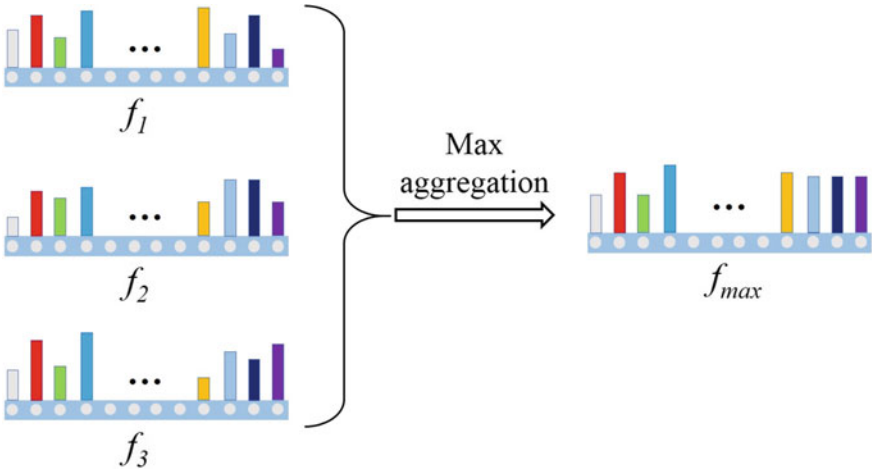



**Fig. 5.** Illustration of the max aggregation.

## 3 Experiments

In this section, we first briefly introduce the Market-1501 database [20] for person re-identification, and then evaluate the above-mentioned aggregation approaches on this database.

### 3.1 Market-1501 Database

Market-1501 database [20] is popular database for person re-identification. It utilizes six cameras to collect the person images in a campus and all the person

images are divided into three parts, i.e., training part, test part, and disturbance part. Training part includes 12,936 images of 751 persons and each person has an average of 17.2 images. Test part contains 19,732 images of 750 persons. This database utilizes the deformable part model (DPM) [24] to automatically detect the pedestrians, which is very close to the realistic setting. During the evaluation process, we utilize the rank-$i$ ($i = 1$, 5 and 10) accuracy and mAP algorithm to evaluate the experiment results.

### 3.2 Evaluation on Two Aggregation Methods

We evaluate two different aggregation approaches in this subsection. Table 1 lists the experimental results on the Market-1501 database. The max aggregation achieves the best results compared with the other method. Specifically, the max aggregation achieves 86.43% rank-1 accuracy and 67.78% mAP accuracy. Compared with the ResNet baseline, the max aggregation yields 3.74% rank-1 and 3.35% mAP improvements, respectively. The max aggregation achieves the best results due to two main reasons. On the one hand, the max aggregation is beneficial to excavation of the comprehensive structure information. On the other hand, utilizing the max aggregation, we can obtain the unique features of different persons. However, the sum aggregation is inferior to the ResNet baseline slightly. The reason is that the sum aggregation ignores some discriminative features, which decreases the performance for person re-identification.

**Table 1.** Comparison with two aggregation approaches on the Market-1501 database.

| Methods | Rank-1 | Rank-5 | Rank-10 | mAP |
|---|---|---|---|---|
| ResNet baseline | 82.69 | 92.90 | 95.63 | 64.43 |
| Sum aggregation | 80.73 | 91.35 | 94.15 | 58.89 |
| Max aggregation | **86.43** | **94.03** | **95.90** | **67.78** |

## 4 Conclusion

In this paper, we have evaluated the influence of local features extracted by convolutional neural networks on the Market-1501 database. Specifically, we first divide the person image into three parts, and then allocate a person re-identification label for each part, which is the same as the original person label. After extracting the local features, we utilize two kinds of approaches to aggregate the partial features, respectively. Experimental results show that the max aggregation can yield better performance than the sum aggregation.

**Acknowledgments.** This work was supported by National Natural Science Foundation of China under Grant No. 61501327, No. 61711530240 and No. 61501328, Natural Science Foundation of Tianjin under Grant No. 17JCZDJC30600 and No. 15JCQNJC01700, the Fund of Tianjin Normal University under Grant No.135202RC1703, the Open Projects Program of National Laboratory of Pattern Recognition under Grant No. 201700001 and No. 201800002, the China Scholarship Council No. 201708120039 and No. 201708120040, and the Tianjin Higher Education Creative Team Funds Program.

## References

1. Liao S, Hu Y, Zhu X, Li SZ. Person re-identification by local maximal occurrence representation and metric learning. In: IEEE conference on computer vision and pattern recognition; 2015. p. 2197–206.
2. Sathish PK, Balaji S. Person re-identification in surveillance videos using multi-part color descriptor. Int J Comput Appl. 2015;121(16):15–7.
3. Zhang R, Liang L, Zhang R, Wang M, Zhang L. Bit-scalable deep hashing with regularized similarity learning for image retrieval and person re-identification. IEEE Trans Image Process. 2015;24(12):4766–79.
4. Zhang Z, Wang C, Xiao B, Zhou W, Liu S. Action recognition using context-constrained linear coding. IEEE Sig Process Lett. 2012;19(7):439–42.
5. Zheng L, Zhang H, Sun S, Chandraker M, Yang Y, Tian Q. Person re-identification in the wild. In: IEEE conference on computer vision and pattern recognition; 2017. p. 1367–76.
6. Zheng L, Yang Y, Hauptmann AG. Person re-identification: past, present and future. arXiv preprint arXiv:1610.02984; 2016.
7. Wu L, Shen C, Hengel AVD. Personnet: person re-identification with deep convolutional neural networks. arXiv preprint arXiv:1601.07255; 2016.
8. Zhao H, Tian M, Sun S, Shao J, Yan J, Yi S, Wang X, Tang X. Spindle net: person re-identification with human body region guided feature decomposition and fusion. In: IEEE conference on computer vision and pattern recognition; 2017. p. 1077–85.
9. Zhang Z, Wang C, Xiao B, Zhou W, Liu S. Attribute regularization based human action recognition. IEEE Trans Inf Forensics Secur. 2013;8(10):1600–9.
10. Zhang Z, Wang C, Xiao B, Zhou W, Liu S, Shi C. Cross-view action recognition via a continuous virtual path. In: IEEE conference on computer vision and pattern recognition; 2013. p. 2690–7.
11. Ma B, Su Y, Jurie F. Local descriptors encoded by fisher vectors for person re-identification. In: European conference on computer vision; 2012. p. 413–22.
12. Zheng Z, Zheng L, Yang Y. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. In: International conference on computer vision; 2017. p. 3774–82.
13. Xiao T, Li H, Ouyang W, Wang X. Learning deep feature representations with domain guided dropout for person re-identification. In: IEEE conference on computer vision and pattern recognition; 2016. p. 1249–58.
14. Zheng Z, Zheng L, Yang Y. A discriminatively learned cnn embedding for person reidentification. ACM Trans Multimedia Comput Commun Appl. 2017;14(1):1–20.
15. Yi D, Lei Z, Liao S, Li SZ. Deep metric learning for practical person re-identification. In: International conference on pattern recognition; 2014. p. 34–9.

16. Sun Y, Zheng L, Yang Y, Tian Q, Wang S. Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline). arXiv preprint arXiv:1711.09349; 2018.
17. Dikmen M, Akbas E, Huang TS, Ahuja N. Pedestrian recognition with a learned metric. In: Asian conference on computer vision; 2010. p. 501–12.
18. Xiang S, Nie F, Zhang C. Learning a mahalanobis distance metric for data clustering and classification. Pattern Recognit. 2008;41(12):3600–12.
19. Cheng D, Gong Y, Zhou S, Wang J, Zheng N. Person re-identification by multi-channel parts-based cnn with improved triplet loss function. In: IEEE conference on computer vision and pattern recognition; 2016. p. 1335–44.
20. Zheng L, Shen L, Tian L, Wang S, Wang J, Tian Q. Scalable person re-identification: a benchmark. In: IEEE international conference on computer vision; 2015. p. 1116–24.
21. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Neural information processing systems; 2012. p. 1097–105.
22. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: International conference on learning representations; 2015.
23. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition; 2016. p. 770–8
24. Felzenszwalb PF, Girshick RB, Mcallester DA, Ramanan D. Object detection with discriminatively trained part-based models. IEEE Trans Pattern Anal Mach Intell. 2010;32(9):1627–45.

# A Modulation Recognition Method Based on Bispectrum and DNN

Jiang Yu, Zunwen He(✉), and Yan Zhang

School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, People's Republic of China
hezunwen@bit.edu.cn

**Abstract.** In this paper, we propose a new method for modulation recognition of received digital signals using bispectrum and AlexNet. The bispectrum analysis is used to generate the feature images, AlexNet, as a widely used deep neural network (DNN), is used as the classifier. It is able to classify six common digital communication signals, including 2ASK, 4ASK, 2FSK, 4FSK, 2PSK and 4PSK. Compared to the traditional decision-theoretic methods, the proposed method needs no prior information for the received signals. The numerical results indicate that this method is more robust and effective than the classical decision theory and its improved algorithm, particularly when the signal-to-noise ratio (SNR) is low. It is shown that the success rate of 90% can be achieved when the SNR is greater than or equal to 3 dB.

**Keywords:** AlexNet · Bispectrum · CNN · DNN · Modulation recognition

## 1 Introduction

With the rapid development of wireless communication technology in modern society, the communication environment has become more and more complicated. Under this complex environment, in order to ensure that the two parties can transmit information accurately and efficiently, the communication signal adopts different modulation formats. For the received signal, signal demodulation and other communication tasks can be completed only on the premise of correct recognition of signal modulation mode and parameters. Therefore, the recognition of modulated signals is very important. Modulation recognition refers to determining the modulation type of the received communication signal with unknown modulation information content [1].

The modulation recognition problem has been considered in many articles that are summarized in [2]. All modulation recognition algorithms considered, there are divided into two groups: one group deals with likelihood functions and the other is based on features extracted from received signals. The decision-theoretic method that deals with likelihood function uses the properties of the signal, and the mathematical formula is obtained after deriving the principle formula and distinguishing it from the appropriate threshold. Under the criterion of Bayesian minimum cost of misjudgment, this method ensures that the recognition result is optimal, but in practical applications, the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_108

decision-theoretic method has problems such as too many parameters required in the recognition process and the computational expressions of the likelihood ratio function are complicated, and the calculation amount is quite large [2].

Feature-based (FB) methods generally consist of feature extraction subsystem and the function of the pattern recognition subsystem. The function of the feature extraction subsystem is to extract features from the received data, and the features are predefined so that the dimensions represented by the patterns are reduced. The role of the second subsystem is to distinguish the type of modulation. As the feature extraction and classifier play a fatal role in the FB methods, much work has done to get a better result. Different artificial neural network (ANN) architectures are used for classification [3, 4]. As for the feature extraction, there are usually used spectral feature set (e.g., Instantaneous amplitude, frequency and phase) [3], statistical feature set (higher-order statistics) [5–7] and wavelet transform [8]. The advantage of FB methods is that the basic algorithm is simple, the selected parameters are sensitive, and it is very applicable in the field of complex signal recognition. The disadvantage is that noise has a great influence on its performance [2].

Recently, deep learning (DL) has become the core technology of artificial intelligence (AI). DL shines in domains such as image, speech and natural language processing, because it is hard to characterize real-world language or images with rigid mathematical models. By using a cascade of nonlinear processing unit layers with high dimensional input data, DL is good at feature extracting and transforming. Furthermore, unsupervised learning of the features and representations of the data is fundamental to DL [9]. As DNN will extract the features and use them by itself, it will reduce the demand of prior knowledge of the received signal.

In this paper, we propose a new method, which use the bispectrum analysis to get the bispectrum features that generate feature images, and use the convolutional neural network (CNN) as the classifier. By using this method, we can identify six modulation types such as 2FSK, 4FSK, 2PSK, 4PSK, 2ASK and 4ASK. Our experimental results show that this method performs well in recognition rate.

The remainder of the paper is organized as follows: Sect. 2 describes the system model, the theory of bispectrum analysis and the CNN classifier. Simulation result of the six modulation types is presented in Sect. 3. Finally, conclusions are made in Sect. 4.

## 2 Proposed Method Based on Bispectrum and DNN

### 2.1 System Model

The system model is shown in Fig. 1 as follows.

As shown in Fig. 1, our system model is consist of three parts. The digital receiver receives the signal through the antenna, and then down-converts the IQ two-way signal, and segments the complex signal according to the sampling rate. We perform non-parametric direct bispectrum estimation on each signal to get bispectrum features to generate 3D images, and then select appropriate view of angle to obtain 2D images. In order to improve recognition accuracy, discarding axis metrics and number symbols for

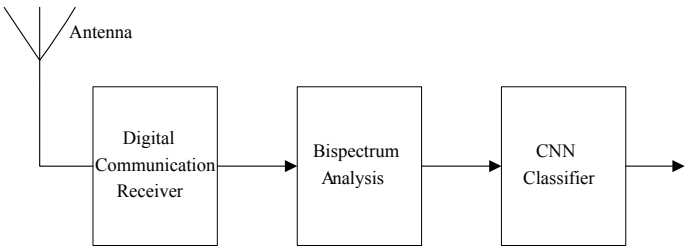



**Fig. 1.** System model

the generated images and selecting the white background as well. Set the image to 256 × 256 pixels, and use the folders as the image classification label, then put them into the CNN classifier to identify. The details are described as follows.

## 2.2 Signal Model

The initial phase of the carrier is assumed to zero, and the digital modulation signal with noise can be expressed as:

$$\widetilde{S}(t) = \overline{S}(t)\exp(j\omega_c t) + n(t) \tag{1}$$

where $\overline{S}(t)$ is the baseband signal, $\omega_c$ is the carrier frequency and $n(t)$ is the complex white Gaussian noise.

After down conversion, the ASK signal can be expressed as:

$$\overline{S}_{\mathrm{ASK}}(t) = \sum_{i=1}^{N} \sqrt{S_i} p(t - iT_b) + n(t) \tag{2}$$

where $T_b$ is the symbol cycle, $p(t)$ is the transmitted symbol waveform.

The FSK signal can be expressed as:

$$\overline{S}_{\mathrm{FSK}}(t) = \sqrt{S} \sum_{i=1}^{N} \mathrm{e}^{j(\omega_i t + \theta_i)} p(t - iT_b) + n(t) \tag{3}$$

where $\omega_i \in \{\omega_1, \omega_2, \omega_3, \ldots, \omega_M\}$, $S$ is the signal power.

The PSK signal can be expressed as:

$$\overline{S}_{\mathrm{PSK}}(t) = \sqrt{S} \sum_{i=1}^{N} \mathrm{e}^{j\Phi_i} p(t - iT_b) + n(t) \tag{4}$$

where $\Phi_i \in \left\{\frac{2\pi}{M}(m-1),\ m = 1, 2, \ldots, M\right\}$, $S$ is the signal power.

### 2.3 Bispectrum Analysis

Given higher-order cumulant $c_{kx}(\tau_1, \tau_2, \ldots, \tau_{k-1})$ of random sequence satisfy the next equation:

$$\sum_{\tau_1=-\infty}^{+\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{+\infty} |c_{kx}(\tau_1, \tau_2, \ldots, \tau_{k-1})| < +\infty \tag{5}$$

then a $k$-order spectrum is the dimensional DFT of $k-1$ order cumulant, i.e.:

$$S_{kx}(\omega_1, \omega_2, \ldots, \omega_{k-1}) = \sum_{\tau_1=-\infty}^{+\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{+\infty} |c_{kx}(\tau_1, \tau_2, \ldots, \tau_{k-1})| \mathrm{e}^{-j\sum_{i=1}^{k-1}\omega_i\tau_i} \tag{6}$$

where $|\omega_i| \le \pi, i = 1, 2, \ldots, k-1, \ |\omega_1 + \omega_2 + \cdots + \omega_{k-1}| \le \pi$.

Bispectrum is the third-order spectrum, defined as:

$$B_x(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{+\infty} \sum_{\tau_2=-\infty}^{+\infty} c_{3x}(\tau_1, \tau_2) \mathrm{e}^{-j(\omega_1\tau_1 + \omega_2\tau_2)} \tag{7}$$

where $|\omega_i| \le \pi, \ i = 1, 2, \ |\omega_1 + \omega_2| \le \pi$.

$$c_{3x}(\tau_1, \tau_2) = E\{x^*(n)x(n+\tau_1)x(n+\tau_2)\} \tag{8}$$

where * stands for the complex conjugate and $E\{\cdot\}$ denotes statistical expectation.

In general, two of the most popular conventional approaches to estimate the higher-order spectra are the indirect and direct methods, which may be seen as direct approximations of the definition of higher-order spectra [10], and in this paper we mainly adopt the latter bispectrum features of the 2ASK, 4ASK, 2FSK, 4FSK, 2PSK and 4PSK are illustrated in Fig. 2.

After that we have to select appropriate view of angle to obtain 2D images from 3D images as shown in Fig. 3. Due to the bispectrum symmetry, we set an observation angle of every 15° from $(-90°, 0°)$ to $(90°, 0°)$ with step of 2 dB from 0 to 8 dB according to the experimental setting in Sect. 3, generating 500 groups of samples per observation angle, of which 300 groups are used as training sets and 200 are used as validation sets, and use the same classifier as in Sect. 3. Experimental results show that the recognition rate of $(\pm 90°, 0°)$ and $(\pm 45°, 0°)$ is better than the other angles, and the recognition success rate between these four angles is almost the same, but $(-45°, 0°)$ requires a lower SNR when the recognition rate reaches 100%, so we use $(-45°, 0°)$ as the best viewing angle when generating 2D images from 3D images, and that may change when more signals are added for recognition. In order to improve recognition accuracy, discarding axis metrics and number symbols for the generated images and selecting the white background as well.

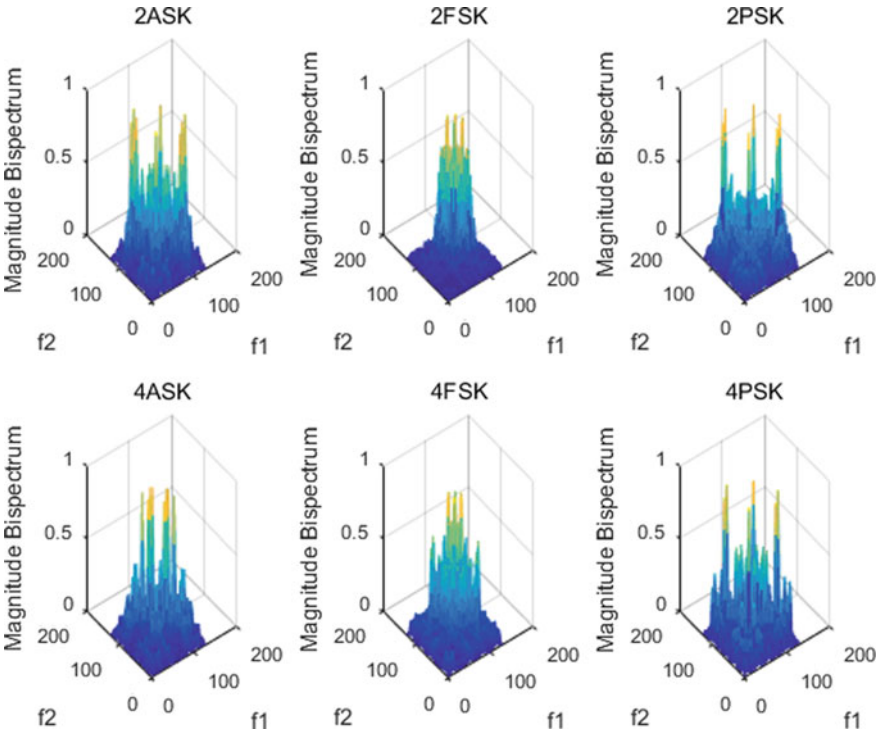



**Fig. 2.** Bispectrum features of the six modulated signals with SNR 5 dB

### 2.4 Deep Neural Network

Deep neural network (DNN) can express complex nonlinear relationships. There are some basic network architecture of DNN, such as fully connected feedforward neural network, convolutional neural network (CNN), recurrent neural network (RNN) and so on.

CNN is a famous model in architecture and it is developed from a fully connected feedforward network to prevent the rapid growth in the parameters when applied in image recognition. Generally, the basic architecture of CNN is to add pooling and convolutional layers before feeding into a fully connected network [11]. By taking most of the calculation, convolutional layer is the core part of CNN with the property of sparse connectivity and weight sharing. Convolutional layers pass the result to the next layer after applying a convolution operation to the input, and this operation involves the convolutional core. If the input image is fairly large, we would like to reduce the number of parameters with the pooling layer. The neurons in the pooling layer are grouped to compute for the mean value (average pooling) or maximum value (max pooling) to reduce the parameters before using the fully connected network. The fully connected layer is almost the same as in traditional neural networks that neurons have full connections to all activations in the previous layer.

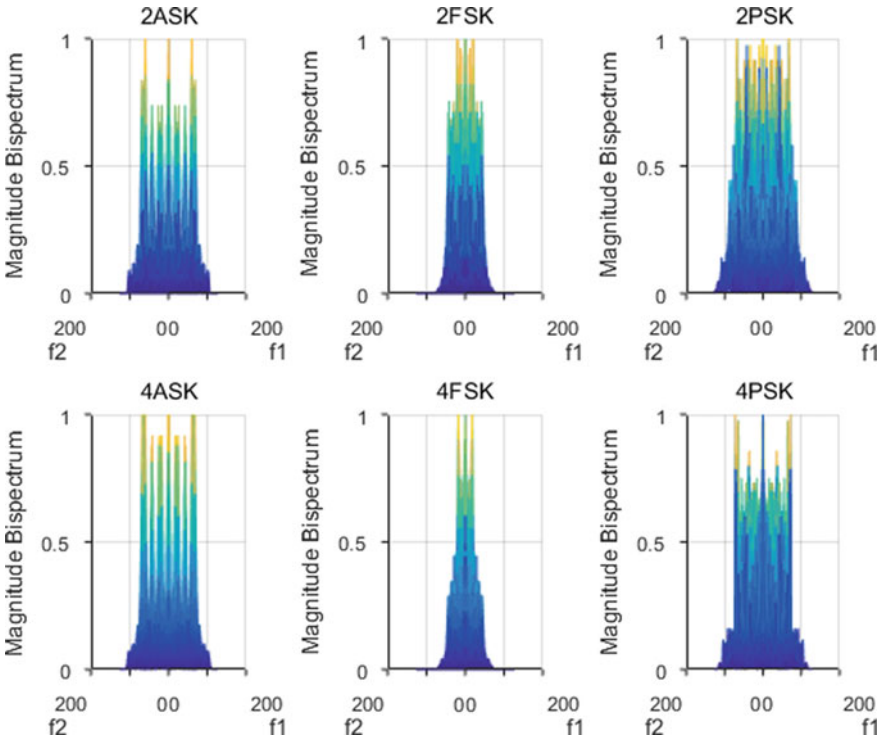



**Fig. 3.** Selected 2D images from 3D images as the input of the CNN classifier

AlexNet is consist of five convolutional layers and three fully connected layers with a 1000 ways softmax layer, and it proposes two primary methods that we can tackle with overfitting. One is "Dropout", the other is "Data Augmentation" [12]. To solve the vanishing gradient problem, AlexNet introduces a new activation function named rectified linear units (ReLU) to replace the classic sigmoid function.

As AlexNet performs well in image recognition, we choose it as the classifier, and the solver options of the AlexNet are set default while the base learning rate is 0.0001.

## 3 Performance Evaluation

### 3.1 Experimental Parameter Setup

In Sect. 2, we described how to generate classification images from received signals. In Sect. 3, we used simulation experiments to verify the validity of the proposed method. We use MATLAB 2016b communication signal toolbox to generate six kinds of simulation signals that are consistent with the parameters of [13], i.e., 2ASK, 4ASK, 2FSK, 4FSK, 2PSK and 4PSK.

The specific parameters of the simulation experiment were set as follows: the symbol rate was 500 bps, the carrier frequency was 2000 Hz, the sampling frequency

was 12,000 Hz, the signal frequency offset was equal to the symbol rate, the carrier amplitude was 1, the number of symbols was 500, the SNR range from 0 to 20 dB, the noise is Gaussian white noise. For each signal, 500 signal samples were collected every 1 dB in a SNR of 0 to 20 dB, of which 300 were used for training set and 200 were used for testing set.

### 3.2 Comparison and Analysis

The comparison of the proposed method named bis and [13] are shown in Table 1 and Fig. 4.

**Table 1.** Average recognition rate of this method and that of [13]

| SNR/dB | 2ASK/% | | 4ASK/% | | 2FSK/% | | 4FSK/% | | 2PSK/% | | 4PSK/% | | AVE/% | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | bis | [13] | bis | [13] | bis | [13] | bis | [13] | bis | [13] | bis | [13] | bis | [13] |
| 0 | 81.60 | 21.50 | 80.90 | 0.80 | 90.30 | 0.00 | 90.00 | 0.00 | 82.30 | 0.00 | 80.30 | 0.00 | 84.23 | 3.72 |
| 2 | 89.60 | 96.15 | 85.30 | 97.80 | 91.30 | 14.40 | 91.30 | 14.20 | 87.30 | 0.00 | 84.60 | 0.00 | 88.23 | 37.10 |
| 4 | 94.30 | 100.0 | 94.60 | 100.0 | 97.30 | 34.10 | 98.30 | 71.70 | 93.60 | 50.00 | 92.60 | 99.10 | 95.12 | 75.80 |
| 6 | 98.90 | 100.0 | 97.30 | 100.0 | 100.0 | 90.90 | 100.0 | 93.30 | 98.60 | 99.30 | 97.60 | 100.0 | 98.73 | 97.25 |
| 8 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

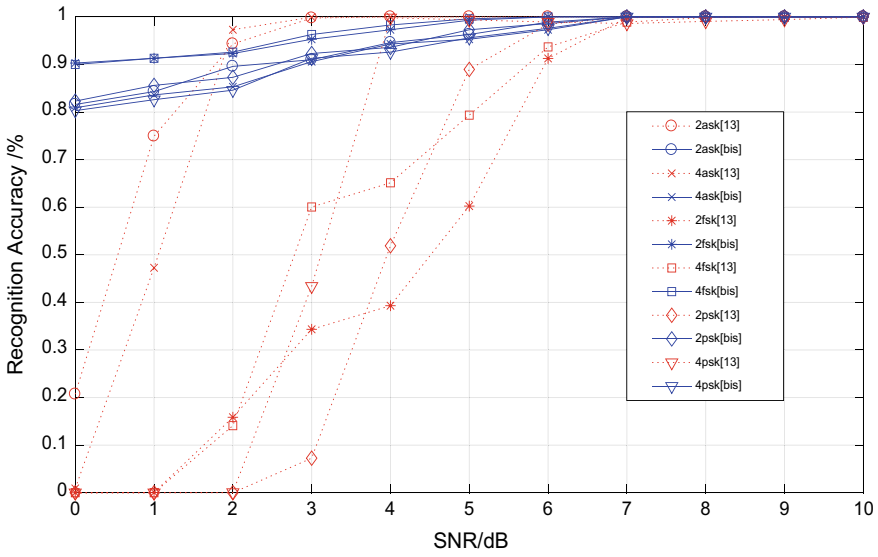



**Fig. 4.** Comparison of the recognition rate all six signals of the proposed method with [13]

Table 1 shows the comparison of the average recognition rate of the proposed method named bis and that of [13]. It can be seen from the table that the average recognition rate of this method is higher than that of [13] in 0–6 dB, which is more obvious at low SNR.

In Fig. 4, the blue solid line is the proposed method named bis and the red dotted line is the method of [13]. From Fig. 4, we can see that the proposed method generally performs better than that of [13] in distinguishing the 2PSK, 2FSK and 4FSK signals, especially in the low SNR. Furthermore, we can see that the proposed method can achieve a recognition success rate of approximately 100% when the SNR is greater than or equal to 7 dB, and success rate of 90% can be achieved when the SNR is greater than or equal to 3 dB, and success rate of 80% can be achieved when the SNR is no less than 0 dB. It performs much better than the algorithm in [4], which reaches 90% recognition rate when the SNR is 10 dB. Meanwhile, it has a better recognition effect compared with [13] when the SNR is low.

## 4 Conclusions

In this paper, we propose a new modulation recognition method based on bispectrum and DNN. The bispectrum analysis is used to generate the feature images, AlexNet, as a widely used DNN, is used as the classifier. The problem of image angle selection is analyzed and some useful conclusions are obtained. The numerical results show that compared with the classical decision theory and its improved algorithm, the proposed method performs better at low SNR, and success rate of 90% can be achieved when the SNR is greater than or equal to 3 dB, and success rate of 100% can be achieved when the SNR is no less than 7 dB.

**Acknowledgments.** This work was supported in part by National Nature Science Foundation of China under Grants No. 61201192 and the National High Technology Research and Development Program of China under Grants No. 2015AA01A708.

## References

1. Azzouz EE, Nandi AK. Algorithms for automatic modulation recognition of communication signals. IEEE Trans Commun. 1998;46(4):431–6.
2. Dobre OA, Abdi A, Bar-Ness Y, et al. A survey of automatic modulation classification techniques: classical approaches and new trends. IET Comm. 2007;1(2):137–56.
3. Wong MLD, Nandi AK. Automatic digital modulation recognition using artificial neural network and genetic algorithm. Sig Process. 2004;84(2):351–65.
4. Park CS, Kim DY. Modulation classification of analog and digital signals using neural network and support vector machine. In: 4th international symposium on neural networks, ICNN2007, Nanjing, Jun 2007.
5. Lopatka J, Pedzisz M. Automatic modulation classification using statistical moments and a fuzzy classifier. In: International conference on signal processing, vol. 3, 61(24); 2000. p. 1500–6.
6. Ebrahimzadeh A, Ardeshir GR. A new signal type classifier for fading environments. J Comp Info Technol. 2007;15(3):257–66.
7. Wong MDL, Nandi AK. Efficacies of selected blind modulation type detection methods for adaptive OFDM systems. In: Proceedings of the international conference on signal processing and communication systems, ISCPSC2007; Dec 2007.

8. Liu J, Luo Q. A novel modulation classification algorithm based on daubechies5 wavelet and fractional fourier transform in cognitive radio. In: IEEE ICCT; 2012. p. 115–20.
9. Li D, Yu D. Deep learning: methods and applications. Found Trends Sig Process. 2014;7(3–4):197–387.
10. Nikias CL, Mendel JM. Signal processing with higher-order spectra. IEEE Sig Process Mag. 1993;10(3):10–37.
11. Lin Y, Tu Y, Dou Z et al. The application of deep learning in communication signal modulation recognition. In: 2017 IEEE/CIC international conference on communications in China (ICCC), Qingdao; Oct 2017.
12. Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. In: International conference on neural information processing systems, p. 1097–105 (2012).
13. Zhang DM, Wang X. Improved method of digital signals modulation identification based on decision theory. J Comput Appl. 2009;29(12):3227–30.

# Image-to-Image Local Feature Translation Using Double Adversarial Networks Based on CycleGAN

Chen Wu(✉), Lei Li, Zhenzhen Yang, Peihong Yan, and Jiali Jiao

National Engineering Research Center of Communication and Network Technology School of Science, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
wc915495936@live.com

**Abstract.** Image-to-image translation is a hot field in the machine learning with the emergency of the generative adversarial networks. Most of the latest models easily lead to changes in the overall image and overfitting when they are used to local feature translation. To address these limitations, this article adds a suppressor and proposes a double adversarial CycleGAN. The suppressor is added to suppress the change of images, and the suppressor and generator form a new adversarial relationship. We hope it will achieve Nash equilibrium that is the change of image focus on the local feature. Finally, a contrast experiment was conducted. In the case of image local feature transfer, the change of image is focused on the local features and the overfitting phenomenon can be well resolved.

**Keywords:** Local feature translation · Generative adversarial networks · Double adversarial CycleGAN

## 1 Introduction

The image-to-image translation [1] was defined as the task of translating one possible representation of a scene into another. The ideal of image-to-image translation starts with the Image Analogies [2]. Lots of researches in computer version, image processing, computational photography, and graphics have been done to learn the mapping between images. The image-to-image translation contains lots of applications such as style transfer, object transfiguration, season transfer, and photograph enhancement. In this paper, we focus on the local feature transfer. Recently, as the emergence of the generative adversarial networks [3], a lot of methods based on GANs have been proposed to deal with the image-to-image translation. Some methods [4,5] were proposed to overcome the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_109

difficulty that it is hard to get a large quantity of paired images. The dual learning approaches [6–10] have been further exploited to the mapping between source image domain and target image domain. Based on these theories, the researchers proposed some image-to-image translation models such as pix2pix [11], CycleGAN [10], StarGAN [12], and GeneGAN [13]. Compared to other generative model, GAN can produce clearer images and the state-of-the-art models could learn the mapping of two domains through the unpaired data. However, when these models are used to image-to-image local feature translation by unaligned data, it will easily lead to overall image changes and overfitting. The aligned data means the local feature that we want to transfer is in the same position of all pictures. We try to use the CelebFaces Attributes Dataset for CycleGAN to train the *glass* ↔ *noglass* translation model (local feature) and fined that it changes the tint of face slightly (see Fig. 1). The CelebFaces Attributes Dataset that we used is the aligned dataset. However, when we try to use the unaligned *horse* ↔ *zebra* data for CycleGAN, the tint of the generated images is changed obviously (see Fig. 4). The CycleGAN does a good job in image-to-image style translation. However, when we use it in image-to-image local feature translation, it will lead to global changes and overfitting of the images. The other latest models rely on data alignment and they did not try to solve the potential problems in image-to-image local feature translation. It is a big step from paired data to unpaired data to reduce the difficulty of data preprocessing. The StarGAN, GeneGAN, and ELEGAN use the aligned CelebFaces Attributes Dataset and the


**Fig. 1.** Comparison chart: The first column (1): original people wearing glasses. The second column (2): *glass* → *noglass* using CycleGAN. The third column (3): original people who do not wear glasses. The fourth column (4): *noglass* → *glass* using CycleGAN

Radboud Faces Database to train the modules. Learning the mapping between domains by aligned data is easier than using the unaligned data. When we use the CycleGAN to learn the mapping $glass \leftrightarrow noglass$ by the aligned CelebFaces Attributes Dataset, the model works well and changes the tint of face a little bit. However, when we use the CycleGAN to learn the mapping $horse \leftrightarrow zebra$ by the unaligned data, the model changes the tint of the images obviously and it is easy to overfit. If we need to train new attributes, it will spend us lots of time aligning dataset. The purpose of this paper is to find a way to make the image-to-image translation focus on the local feature by the unaligned data. Our goal is to learn the mappings $G, F$ exactly between the domain X and domain Y($G : X \rightarrow Y, F : Y \rightarrow X$). The CycleGAN introduces an additional loss to inhibit the change of images. It was the technique of Taigman [14]

$$L_{identity}(G, F) = \mathbb{E}_{y \sim p_{data(y)}}[\|G(y) - y\|_1] + \mathbb{E}_{x \sim p_{data(x)}}[\|F(x) - x\|_1] \quad (1)$$

We try this method in local feature image-to-image translation, but this method is not good enough. The additional loss regularizes the generator directly; it means we hope the generator can change the images domain and inhibit the change of images at the same time. It is difficult for the generator to achieve the balance. In this paper, we propose a double adversarial network based on CycleGAN to improve the local feature image-to-image translation by the unaligned data. We let the $L_{identity}(G, F)$ as a suppressor instead of a regular. Therefore, the generator and the suppressor can train the $G, F$ alternately. We find that the confrontational relationship between generator and suppressor can accelerate convergence.

## 2 Related Work

### 2.1 Generative Adversarial Networks

The GAN is a competition between two networks. The generator network receives a source of noise and exports the corresponding fake data samples. The discriminator network receives the fake data samples and the real data samples, and then it distinguishes the fake data samples and the real data samples. The aim of the generator is to fool the discriminator, and the aim of the discriminator is to distinguish samples better. The minimax objective function of GAN:

$$\min_G \max_D \underset{x \sim p_r}{E}[\log(D(x))] + \underset{x \sim p_g}{E}[\log(1 - D(x))] \quad (2)$$

where the $p_r$ is the true data distribution and the $p_g$ is the output distribution of generator. Firstly, the discriminator is trained to optimality. Secondly, when minimizing the objective function to minimize the Jensen–Shannon divergence between $p_r$ and $p_g$, there are some problems with original GAN such as collapse problem, non-convergence problem, and vanishing gradient problem. Some methods were proposed to solve these problems such as DCGAN [15], LS-GAN [16], WGAN-gp [17]. CycleGAN uses the LS-GAN instead of the original GAN.

### 2.2 Cycle-Consistent Generative Adversarial Network

CycleGAN is a model algorithm based on the use of two generators and two discriminators to transform the distribution of two types of images based on generative adversarial network. The purpose of CycleGAN is to learn the mapping relationship between two data distributions $X$ and $Y$. Assuming only the generator and discriminator, the generator does not learn the mapping relationship between the two data distributions very well. It is possible that all the images of one data distribution are mapped to one picture in another data distribution. Therefore, the cycle-consistent loss is proposed. The forward cycle consistency is $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$, and the backward cycle consistency is $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$ (see Fig. 2). The full loss function of CycleGAN is:

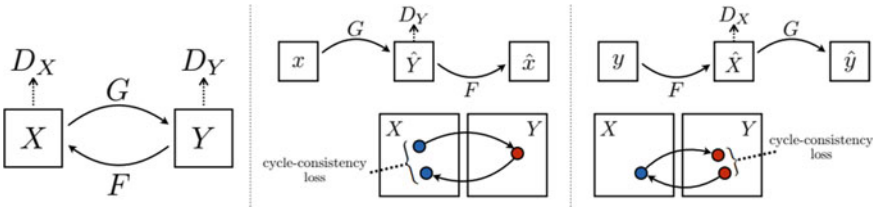



**Fig. 2.** CycleGAN tries to learn the mapping functions $G : X \rightarrow Y$ and $F : Y \rightarrow X$ by the loss function $L(G, F, D_{(X)}, D_{(Y)})$ combined by $L_{GAN}(G, D_Y, X, Y)$, $L_{GAN}(F, D_X, Y, X)$, and cycle-consistent loss.

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, Y, X) + \lambda L_{cyc}(G, F) \tag{3}$$

$$L_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{data(y)}}[\log(D_Y(y))] + \mathbb{E}_{x \sim p_{data(x)}}[\log((1 - D_Y(G(x))))] \tag{4}$$

$$L_{GAN}(F, D_X, Y, X) = \mathbb{E}_{x \sim p_{data(x)}}[\log(D_X(x))] + \mathbb{E}_{y \sim p_{data(y)}}[\log((1 - D_X(G(y))))] \tag{5}$$

$$L_{cyc}(F, D_X, Y, X) = \mathbb{E}_{x \sim p_{data(x)}}[\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{data(y)}}[\|F(G(y)) - y\|_1] \tag{6}$$

CycleGAN does a good job in unpaired image-to-image style translation such as season transfer and painting $\leftrightarrow$ photographs. However, the performance of CycleGAN in image-to-image local feature translation is not satisfied. It is easy to change the tint of the pictures, and it easily leads to the overfitting.

In the paper of CycleGAN, it adopts the technique of Taigman by a additional loss $L_{identity}(G, F) = \mathbb{E}_{y \sim p_{data(y)}}[\|G(y) - y\|_1] + \mathbb{E}_{x \sim p_{data(x)}}[\|F(x) - x\|_1]$. The purpose of the $L_{identity}(G, F)$ is to decrease the change between the image-to-image translation $G(x), F(y)$. It has a good performance in image-to-image style translation. However, when I try to use the new loss function in image-to-image local feature translation such as $horse \leftrightarrow horse$ or $glass \leftrightarrow noglass$, it does not work well.

## 3 Double Adversarial CycleGAN

When we try to transfer the $horses \rightarrow zebras$ by CycleGAN, the tint of the pictures is changed at the same time. The phenomenon of overfitting is occurred. Without the $L_{identity}(G, F)$, the CycleGAN is easy to lead to the overall change in the images. However, when we add the $L_{identity}(G, F)$ as a regular to the loss, we find that it is not easy for the generator to achieve the balance. It can even lead to the phenomenon that the local feature translation that should be changed is suppressed and the background area of the picture is changed. We can see that the translation of $horse \rightarrow zebra$ is more focused on local feature by the double adversarial CycleGAN (see Fig. 4). The experiment shows that the generated horse still has some zebra stripes (see Fig. 5). The generator is difficult to learn the mapping of $zebra \rightarrow horse$. When we add the regular term or the suppressor to the CycleGAN, it will lead to more zebra stripes. We have not yet solved this problem.

According to these problems, we take some methods to solve them. The previous method is to add the $L_{identity}(G, F)$ to the generator's loss function during the training process. It means that the $L_{identity}(G, F)$ loss is a regular term. When the generator tries to do image-to-image translation, it also limits the change of image-to-image translation at the same time. It is hard for the generator to learn the balance in one step. Inspired by the two-person zero-sum game, we let the $L_{identity}(G, F)$ as a suppressor and the generator can achieve the balance stably step by step (see Fig. 3). Therefore, the generator–discriminator and generator–suppressor constitute a double adversarial cycle-consistent loss. The objective function of suppressor:

$$min \quad S = \mathbb{E}_{y \sim p_{data(y)}}[\|G(y) - y\|_1] + \mathbb{E}_{x \sim p_{data(x)}}[\|F(x) - x\|_1]$$

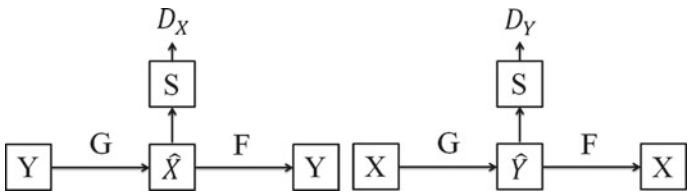



**Fig. 3.** The suppressor(S) is used to suppress the difference between the generated images and the original images.

## 4 Experiment

In this section, we compare these three methods: CycleGAN, CycleGAN + $L_{identity}(G, F)$, and double adversarial CycleGAN. As shown in Figs. 4 and 5, the CycleGAN model changes the tint of images obviously. The overall color of the image has become deeper, and generated ($horse \rightarrow zebra$) stripes are not clear. The CycleGAN + $L_{identity}(G, F)$ model improves the phenomenon of the tint change in image-to-image translation. However, the generated zebra ($horse \rightarrow zebra$) stripes are also not clear. From the fourth column of Fig. 4, we can see the generated zebra ($horse \rightarrow zebra$) stripes are clear and the tint of image is stable.

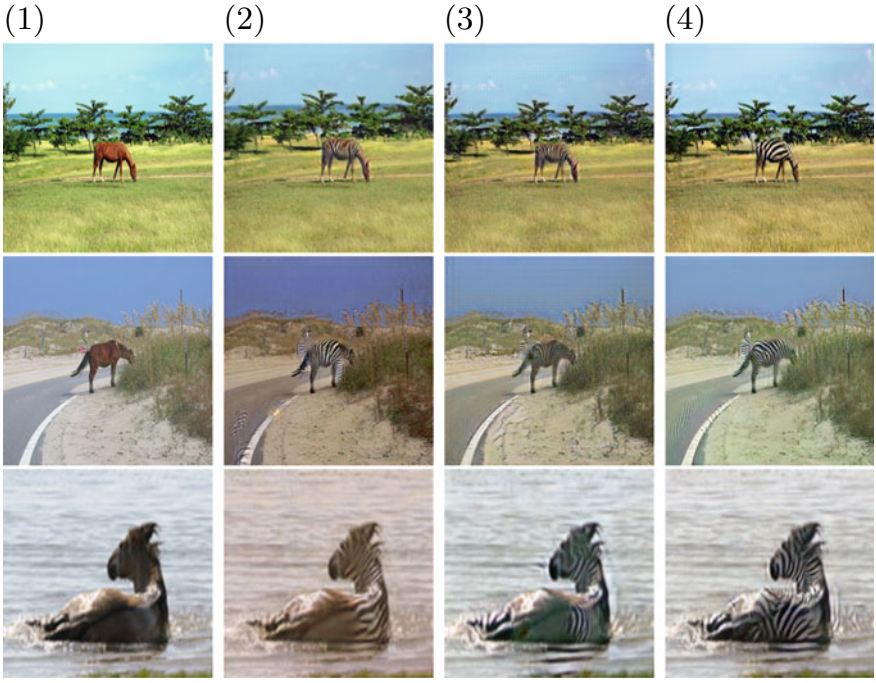



**Fig. 4.** Comparison chart ($horse \rightarrow zebra$): The first column (1): original images. The second column (2): $horse \rightarrow zebra$ using CycleGAN. The third column (3): $horse \rightarrow zebra$ using CycleGAN+$L_{identity}(G, F)$. The fourth column (4): $horse \rightarrow zebra$ using double adversarial CycleGAN

The Inception Score (IS) and Fréchet Inception Distance (FID) [18] are two ways to test the quality of generated pictures. We use the FID score to measure the distance between the real data and the generated data.

$$d^2((m, C), (m_\omega), C_\omega) = ||m - m_\omega||_2^2 + Tr(C + C_\omega - 2(CC_\omega)^{1/2}) \quad (7)$$

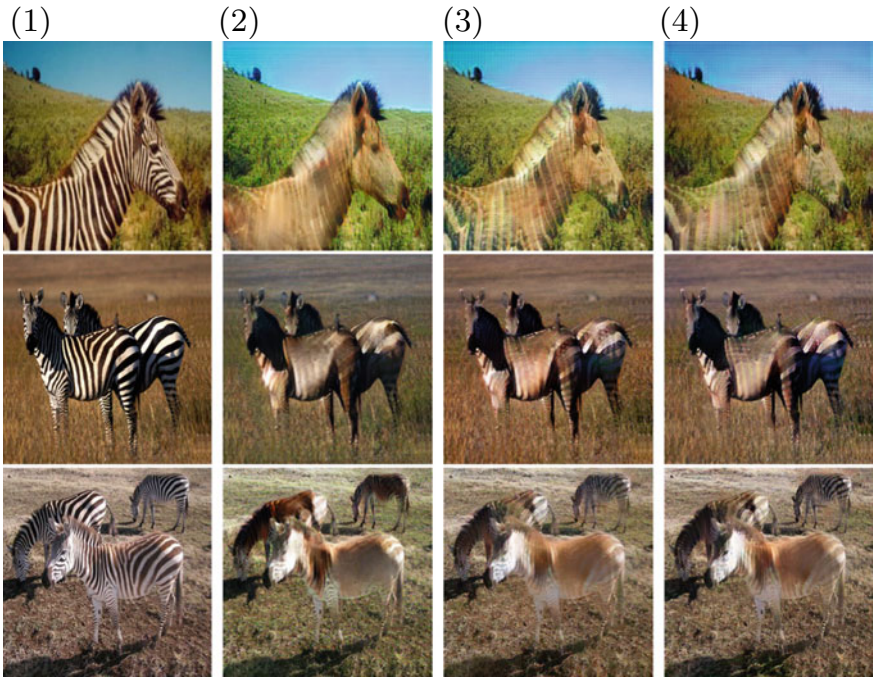



**Fig. 5.** Comparison chart ($zebra \rightarrow horse$): The first column (1): original images. The second column (2): $zebra \rightarrow horse$ using CycleGAN. The third column (3): $zebra \rightarrow horse$ using CycleGAN+$L_{identity}(G, F)$. The fourth column (4): $zebra \rightarrow horse$ using double adversarial CycleGAN

$(m, C)$ and $(m_\omega, C_\omega)$ are the means and covariance matrices of the real data distribution and fake data distribution. This table shows the FID of generated images. We can see the result of CycleGAN+$L_{identity}(G, F)$ is worse than CycleGAN. Double adversarial CycleGAN has a good performance in $horse \rightarrow zebra$, and it is not suitable for the $zebra \rightarrow horse$. However, both IS and FID are not suitable for our work. We need to measure the ratio of global and local changes in the image-to-image translation and the quality of the local features from generated pictures. We have not done this yet (Table 1).

**Table 1.** FID score

| Name | CycleGAN | CycleGAN $+L_{identity}(G, F)$ | Double adversarial CycleGAN |
|---|---|---|---|
| FID($zebra \rightarrow horse$) | 68.4 | 86.5 | 75.4 |
| FID($horse \rightarrow zebra$) | 69.0 | 65.6 | 51.9 |

This table shows the FID of generated images.

## 5 Conclusion

In this paper, we propose a double adversarial CycleGAN. Under the observation that only local part of the image should be modified in this task, we let the $L_{identity}(G, F)$ as a suppressor instead of regular term and it can accelerate convergence and achieve the balance more quickly. The confrontation between generator and suppressor can make the training more stable.

**Acknowledgments.** This work was supported by National Natural Science Foundation of China (Grant No. 61501251, 61373137, 61071167) and the Science Foundation of Nanjing University of Posts and Telecommunications Grant (NY214191).

## References

1. Liu M-Y, Breuel T, Kautz J. Unsupervised image-to-image translation networks. CoRR, abs/1703.00848, 2017.
2. Hertzmann A, Jacobs CE, Oliver N, Curless B, Salesin DH. Image analogies. In: Proceedings of the 28th annual conference on computer graphics and interactive techniques, SIGGRAPH '01. ACM: New York, NY, USA, 2001. p. 327–40.
3. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, editors. Advances in neural information processing systems 27. New York: Curran Associates, Inc.; 2014. p. 2672–80.
4. Perarnau G, van de Weijer J, Raducanu B, Álvarez JM. Invertible conditional gans for image editing. CoRR, abs/1611.06355, 2016.
5. Zhu J-Y, Krähenbühl P, Shechtman E, Efros AA. Generative visual manipulation on the natural image manifold. In: Proceedings of European conference on computer vision (ECCV), 2016.
6. Xia Y, He D, Qin T, Wang L, Yu N, Liu T-Y, Ma W-Y. Dual learning for machine translation. CoRR, abs/1611.00179, 2016.
7. Shen W, Liu R. Learning residual images for face attribute manipulation. CoRR, abs/1612.05363, 2016.
8. Kim T, Cha M, Kim H, Lee JK, Kim J. Learning to discover cross-domain relations with generative adversarial networks. CoRR, abs/1703.05192, 2017.
9. Yi Z, Zhang H, Tan P, Gong M. Dualgan: Unsupervised dual learning for image-to-image translation. CoRR, abs/1704.02510, 2017.
10. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. CoRR, abs/1703.10593, 2017.
11. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. CoRR, abs/1611.07004, 2016.
12. Choi Y, Choi M-J, Kim M, Ha J-W, Kim S, Choo J. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. CoRR, abs/1711.09020, 2017.
13. Zhou S, Xiao T, Yang Y, Feng D, He Q, He W. Genegan: Learning object transfiguration and attribute subspace from unpaired data. CoRR, abs/1705.04932, 2017.
14. Taigman Y, Polyak A, Wolf L. Unsupervised cross-domain image generation. CoRR, abs/1611.02200, 2016.
15. Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. CoRR, abs/1511.06434, 2015.

16. Qi G-J. Loss-sensitive generative adversarial networks on lipschitz densities. CoRR, abs/1701.06264, 2017.
17. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC. Improved training of wasserstein gans. CoRR, abs/1704.00028, 2017.
18. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Klambauer G, Hochreiter S. Gans trained by a two time-scale update rule converge to a nash equilibrium. CoRR, abs/1706.08500, 2017.

# Evaluation Embedding Features for Ground-Based Cloud Classification

Zhong Zhang[1,2(✉)], Donghong Li[1,2], and Shuang Liu[1,2]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
{zhong.zhang8848,donghongli1139,shuangliu.tjnu}@gmail.com
[2] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China

**Abstract.** Ground-based cloud classification plays a vital important role in meteorological research. However, the existing methods perform well confined to one weather station. In this paper, we present a detailed introduction of two representative embedding features for ground-based cloud classification in various weather stations. The features are learned from the metric learning and the convolutional neural network (CNN), respectively. The two kinds of features are evaluated on two weather stations.

**Keywords:** Ground-based cloud classification · Embedding features · Metric learning · Convolutional neural networks

## 1 Introduction

Ground-based cloud classification has an important impact on meteorological studies [1]. However, the professional human observers are the main conductors to implement this task, resulting in mass consumption of human efforts and material resources. In order to address this issue, many researchers have focused on developing automatic techniques for ground-based cloud classification in recent years. Isosalo et al. [2] classified ground-based cloud images by utilizing local binary patterns (LBPs) and local edge patterns (LEPs), and LBPs had better performances than LEPs. Calbó and Sabburg [3] represented cloud images using pattern features based on the Fourier transformation and statistical features. Liu et al. [4] extracted both texture and structure features from the color cloud images. Liu et al. [5] proposed illumination-invariant completed local ternary patterns (ICLTP), which can effectively handle the illumination variations. They soon proposed the salient LBP (SLBP) [6] to capture descriptive cloud information. The desirable property of SLBP is the robustness to noise.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_110

Recently, Ye et al. [7] first applied the convolutional neural network (CNN) to ground-based cloud classification, and they employed the Fisher Vector to encode the deep convolutional features. Shi et al. [8] concluded that the deep convolutional activation-based features outperformed traditional hand-craft features considerably for ground-based cloud classification. However, these methods are only suitable for one certain weather station. That means they can achieve desired classification accuracies on the certain dataset captured by the weather station, but obtain poor performances on other weather stations. In other words, these methods are not generalized to other weather stations, because the cloud images captured by them generally possess variations in illuminations, images resolutions, occlusions, camera settings, etc., as shown in Fig. 1. Therefore, we argue that training a classifier on one weather station, while test on another weather station.

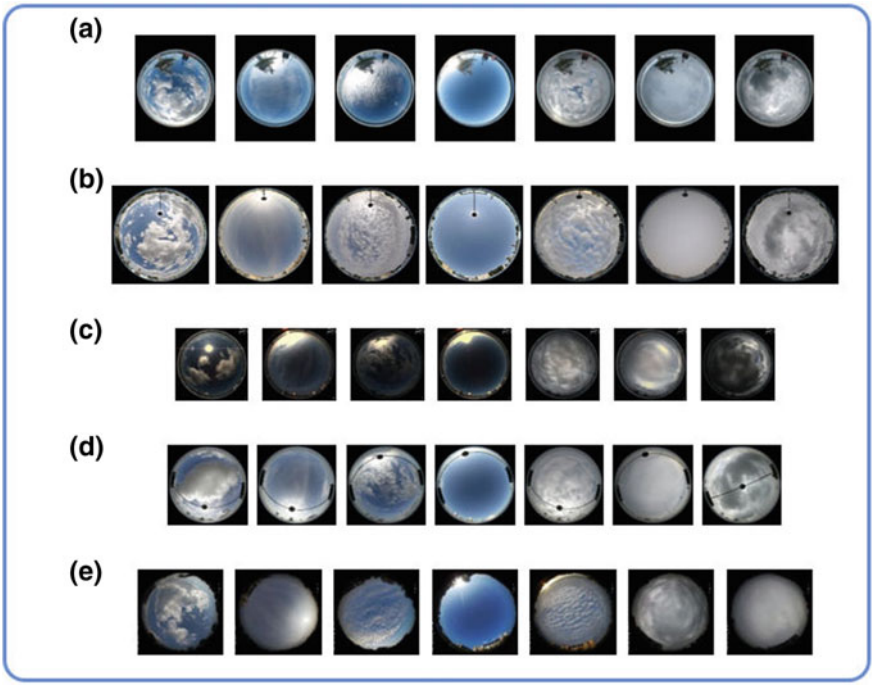



**Fig. 1.** Cloud samples are provided by various weather stations (**a**)–(**e**).

In this paper, we evaluate different cloud images provided by various weather stations based on two kinds of embedding features. The first one is to learn a subspace using metric learning methods. Metric learning has been successfully applied to the domain adaptation [9,10]. The metric learning is suitable for the situation that training on one weather station, while test on other weather station. The second one is the output of the fully connected (FC) layer of the CNN.

The CNN has achieved excellent performances in image classification [11–13], and it is a desirable choice to apply the CNN to ground-based cloud classification.

The remaining of the paper is organized as follows. Section 2 introduces the metric learning and CNN. Section 3 shows the experimental results of various weather stations based on the two embedding features. We conclude the paper in Sect. 4 the paper.

## 2 Approach

### 2.1 Metric Learning

First, the cloud images are divided into two sets of pairs where a pair consists of two cloud images. The first set $S$ includes similar pairs where the two cloud images of each similar pair belong to the same cloud category. The second set $D$ consists of dissimilar pairs where the two cloud images of each dissimilar pair belong to different cloud categories. The metric learning considers the relationships of pairs to learn a transformation matrix, which is used to project the cloud images into a subspace. Suppose that there is a pair $(i, j)$, where $i \in \mathbb{R}^{d\times 1}$ and $j \in \mathbb{R}^{d\times 1}$ are from different weather stations. We learn a transformation matrix $H \in \mathbb{R}^{d\times m}$ which parameterizes the (squared) Mahalanobis distance:

$$d_M(i, j) = (i - j)^T M(i - j), \tag{1}$$

where $M = HH^T (M \in \mathbb{R}^{d\times d})$. If $i$ and $j$ belong to a similar pair, there is:

$$(i, j) \in S, \tag{2}$$

in the same way, if $i$ and $j$ belong to a dissimilar pair, there is:

$$(i, j) \in D. \tag{3}$$

Based on the Mahalanobis distance, in order to learn $H$, Xing et al. [14] proposed a metric that minimizes the distance between all similar pairs:

$$\begin{aligned} \min_{M} \quad & \sum_{(i,j)\in S} (i - j)^T M(i - j) \\ s.t. \quad & \sum_{(i,j)\in D} \sqrt{(i - j)^T M(i - j)} \geq 1, \\ & M \geq 0. \end{aligned} \tag{4}$$

The first constraint ensures that $M$ does not collapse the dataset into a single point, and the second one ensures a valid metric.

Ghodsi et al. [15] proposed to minimize the squared distance between similar pairs, while maximize the squared distance between dissimilar pairs. Similarly, Globerson and Roweis [16] learned a metric by keeping images from the same category near each other and simultaneously far from the images from other categories. Weinberger et al. [17] utilize semidefinite programming to learn a metric for k-nearest neighbor (kNN) classification.

### 2.2 Convolutional Neural Network

The CNN model can learn the high nonlinear transformation for the input data. Thus, the features extracted from FC layer can be viewed as embedding features. The VGG-11 [13] model is representative, and therefore, we evaluate the VGG-11 model on the task of ground-based cloud classification. There are eight convolutional layers and three fully connected (FC) layers composing the VGG-11 model. The cloud images as the inputs of the VGG-11 model should be resized to $224 \times 224$ pixels. The filters of each convolutional layer are with a receptive field of $3 \times 3$ pixels, and the convolution stride is set to 1 pixel. Five max-pooling layers follow some of convolutional layers and are implemented by a sliding window of $2 \times 2$ pixels with stride 2. The first two FC layers both consist of 4096 neurons, and the last one includes 1000 neurons where each neuron denotes a category of the ImageNet dataset. However, when the VGG-11 model is fine-tuned on the ground-based cloud datasets, we change 1000 into the number of cloud categories, as illustrated in Fig. 2. We take the output of the second FC layer as features, which contain the overall spatial layout information. That means a cloud image is represented as 4096-dimensional vector.

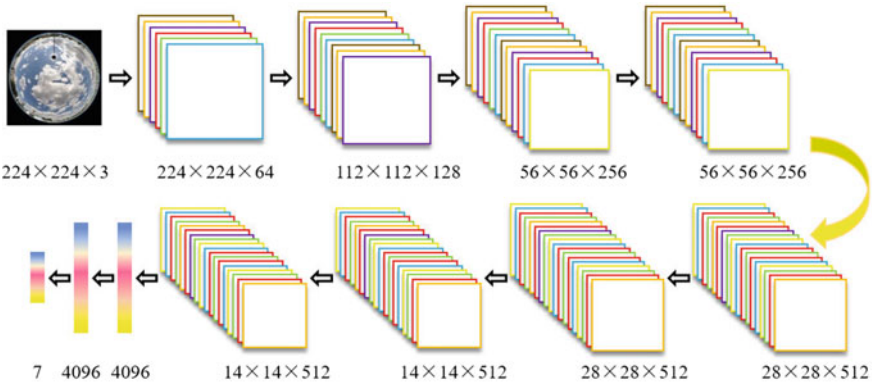



**Fig. 2.** Cloud image with the size of $224 \times 224 \times 3$ is fed into the VGG-11 model, which results in a series of feature maps with different sizes. The final output is 7 neurons corresponding to the number of cloud categories.

## 3 Experiments

### 3.1 Datasets and Experimental Setup

The are two cloud datasets serving for the following evaluations. The Institute of Atmospheric Physics, Chinese Academy of Sciences, provides the first cloud dataset, named IAP_e. There are seven categories and 3533 cloud images in this

dataset. The cloud images from the IAP_e dataset are $2272 \times 1704$ pixels. The second cloud dataset is named as CAMS_e with the number of 2491 cloud images, provided by Chinese Academy of Meteorological Sciences. Each cloud image has the resolution of $1392 \times 1040$. Since the two cloud datasets are captured from different weather stations, the IAP_e is quitely different from the CAMS_e in the aspects of camera settings, resolutions, occlusions, and illuminations, as shown in Fig. 3.

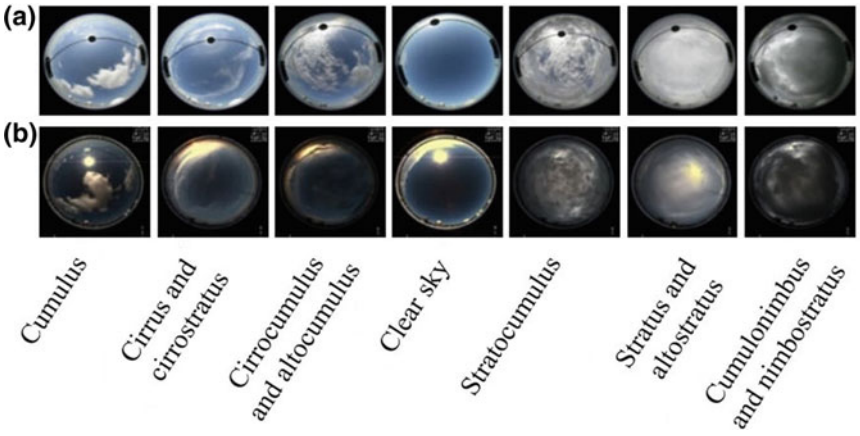



**Fig. 3. a** The IAP_e cloud dataset. **b** The CAMS_e cloud dataset.

All images from the two datasets are resized to $224 \times 224$ pixels. We extract LBP features with $(P, R)$ equal to (8, 1), (16, 2), and (24, 3), and then concatenate histograms of the three scales to form a feature vector with the size of $10 + 18 + 26 = 54$. The $M$ is finally multiplied by the resulting feature vector to form a transformed feature vector for each cloud image. We select $k = 150$. When we fine-tune the VGG-11 [13] model on our cloud datasets, we set the batch size and the number of training epochs to 32 and 60, respectively. The learning rate is initialized as 0.0001 and then set to 0.00006 for the final 20 epochs. A half of cloud images is randomly selected from each category and consists of the training set, and the remaining as the test set. The final results represent the recognition accuracy averaged over ten randomly selected training/test sets. The nearest neighborhood classifier is used to classify the cloud images.

Table 1 lists the classification results of the two embedding features on different weather stations. It is obvious that the high classification results are achieved when training and testing on the same dataset. However, we obtain poor classification performances when training and testing on different datasets. Moreover, the CNN performs much better than the metric learning. To conclude, it is essential to make a further study for the ground-based cloud classification in various weather stations.

**Table 1.** Classification results (%) using different embedding features.

| Datasets | | Methods | |
|---|---|---|---|
| | | Metric learning | CNN |
| IAP_e | IAP_e | 86.93 | 92.74 |
| CAMS_e | CAMS_e | 89.42 | 93.81 |
| IAP_e | CAMS_e | 38.65 | 53.84 |
| CAMS_e | IAP_e | 36.81 | 52.07 |

## 4 Conclusion

In this paper, we have introduced two kinds of representative embedding features which are learned from the metric learning and the CNN, respectively. Then, we evaluated the two features on two weather stations. The experimental results show that the metric learning and the CNN perform poorly when training on weather station, while test on another weather station. In other words, they are unsatisfied with the ground-based cloud classification in different weather stations, and it is still eager to develop effective methods on this issue.

**Acknowledgments.** This work was supported by National Natural Science Foundation of China under Grant No. 61501327 and No. 61711530240, Natural Science Foundation of Tianjin under Grant No. 17JCZDJC30600 and No. 15JCQNJC01700, the Fund of Tianjin Normal University under Grant No.135202RC1703, the Open Projects Program of National Laboratory of Pattern Recognition under Grant No. 201700001 and No. 201800002, the China Scholarship Council No. 201708120039 and No. 201708120040, and the Tianjin Higher Education Creative Team Funds Program.

## References

1. Chen Z, Zen D, Zhang Q. Sky model study using fuzzy mathematics. J Illum Eng Soc. 1994;23:52–8.
2. Isosalo A, Turtine M, Pietikäinen M. Cloud characterization using local texture information. In: Finnish signal processing symposium, Kuopio, Finland; 2007. p. 1–6.
3. Calbó J, Sabburg J. Feature extraction from whole-sky ground-based images for cloud-type recognition. J Atmos Ocean Tech. 2008;25:3–14.
4. Liu L, Sun X, Chen F, Zhao S, Gao T. Cloud classification based on structure features of infrared images. J Atmos Ocean Tech. 2011;28:410–7.
5. Liu S, Wang C, Xiao B, Zhang Z, Shao Y. Illumination-invariant completed LTP descriptor for cloud classification. In: International congress on image and signal processing, Chongqing, China; 2012. p. 449–53.
6. Liu S, Wang C, Xiao B, Zhang Z, Shao Y. Salient local binary pattern for ground-based cloud classification. Acta Meteorol Sin. 2013;27:211–20.
7. Ye L, Cao Z, Xiao Y. DeepCloud: ground-based cloud image categorization using deep convolutional features. IEEE Trans Geosci Remote. 2017;55:5729–40.

8. Shi C, Wang C, Wang Y, Xiao B. Deep convolutional activations-based features for ground-based cloud classification. IEEE Geosci Remote Sens Lett. 2017;14:816–20.
9. Ding Z, Fu Y. Robust transfer metric learning for image classification. IEEE Trans Image Process. 2017;26:660–70.
10. Saenko K, Kulis B, Fritz M, Darrell T. Adapting visual category models to new domains. In: European conference on computer vision, Heraklion, Crete, Greece; 2010. p. 213–26.
11. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, Lake Tahoe, Nevada, USA; 2012. p. 1097–105.
12. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition, Las Vegas, Nevada; 2016. p. 770–8.
13. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: International conference on learning representations, San Diego, California, USA; 2015. p. 1–14.
14. Xing EP, Jordan MI, Russell SJ, Ng AY. Distance metric learning with application to clustering with side-information. In: Advances in neural information processing systems, Vancouver and Whistler, British Columbia, Canada; 2003. p. 521–8.
15. Ghodsi A, Wilkinson DF, Southey F. Improving embeddings by flexible exploitation of side information. In: International joint conferences on artificial intelligence, Hyderabad, India; 2007. p. 810–6.
16. Globerson A, Roweis ST. Metric learning by collapsing classes. In: Advances in neural information processing systems, Vancouver and Whistler, British Columbia, Canada; 2006. p. 451–8.
17. Weinberger KQ, Blitzer J, Saul LK. Distance metric learning for large margin nearest neighbor classification. In: Advances in neural information processing systems, Vancouver and Whistler, British Columbia, Canada; 2006. p. 1473–80.

# A Gradient Invariant DCT-Based Image Watermarking Scheme for Object Detection

Xiaocheng Hu(✉), Bo Zhang, Huibo Li, Jing Guo, Yunxiang Yang, Yinan Jiang, and Ke Guo

China Academy of Electronic Information Technology, Beijing 100041, China
675342900@qq.com

**Abstract.** In this paper, we proposed a novel DCT-based watermarking scheme for grayscale images, which utilizes the connection between discrete cosine transform (DCT) and the Histogram of Oriented Gradient (HOG) feature extraction operation. We embed messages into the low-frequency band, and correspondingly, an effective coefficients pair selection scheme is constructed. The proposed scheme not only maintains the superiority of compression robustness but also keeps good visual quality for the watermarked image. Moreover, the proposed method is insensitive to HOG feature extraction, which makes the watermarked image more suitable for further objection detection and recognition scenarios.

**Keywords:** Watermarking · Robustness · HOG feature extraction

## 1 Introduction

With the rapid development of mobile devices and social networks, digital images are widely captured and shared throughout the Internet, which draws several urgent issues relating to copyright protection and authentication during transmission. Digital image watermarking is a kind of data hiding method that has been extensively studied in recent decades.

Most of the image watermarking schemes satisfy one or more of the five requirements: imperceptibility, non-detectability, security, robustness, and capacity. Blind image watermarking methods generally belong to the spatial or transform domain. Many recent advances related are largely inspired by the manipulation of the domain transform of multimedia objects. Watermarks embedded in the transform domain are usually more robust and less perceptible, while their computational requirements tend to be higher than those spatial domain ones. The major domains used for image watermarking include discrete cosine transform (DCT) [1–3], discrete Fourier transform (DFT) [4–6], discrete wavelet transform (DWT) [7–9], and singular value decomposition (SVD) [10,11].


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_111

DCT has proven particularly effective with regard to energy compaction and the ability to incorporate characteristics of human visual system (HVS) [12,13].

Recently in [14], Soumitra proposed a compromised scheme using the middle-frequency band DCT coefficients pairs, which are less vulnerable to modification and also the embedding capacity is guaranteed. Existing watermarking schemes mainly focus on robustness against lossy compression or geometric transformation attacks for message recovery, while before the message extraction procedure, kinds of image processing operations can be applied to the watermarked image. For example, the well-known Histogram of Oriented Gradients (HOG) feature descriptor is very popular for object detection and recognition in computer vision field. A modified watermarked image will affect the feature extraction and thus result in the decrease in detection and recognition performance. For some important applications such as pedestrian retrieval in a police or military system, slight accuracy decrease may involve fateful consequences. Therefore, watermarking scheme for images not interfering with further image processing is especially desired. In [15], Hou et al. proposed a recursive reversible data hiding method for color images, which keeps the gray version of the cover image unchanged, and thus, the feature extraction is not affected.

In this paper, we proposed a novel DCT-based watermarking scheme for grayscale images, which utilizes the connection between DCT transform and HOG feature extraction operation. We embed messages into the low-frequency band, and further, a novel coefficients pair selection scheme is constructed. The proposed scheme not only maintains the superiority of robustness but also keeps good visual quality for the watermarked image.

## 2 Previous Arts

In [14], Soumitra et al. proposed a block-based watermarking scheme using DCT coefficients pairs. First, the input $M \times N$ image $I$ is divided into non-overlapping blocks of size $k \times k$. Then, each block $I_b$ is transformed into the DCT domain through the following equation:

$$F_b(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{k-1} \sum_{y=0}^{k-1} I_b(x,y) \times cos\left[\frac{(2x+1)u\pi}{2k}\right] cos\left[\frac{(2y+1)v\pi}{2k}\right] \quad (1)$$

where

$$\alpha(u) = \begin{cases} \sqrt{1/k}, & u = 0 \\ \sqrt{2/k} & otherwise \end{cases} \quad and \quad \alpha(v) = \begin{cases} \sqrt{1/k}, & v = 0 \\ \sqrt{2/k} & otherwise \end{cases} \quad (2)$$

When computed, the DCT coefficients are collected according to the zigzag order. The method in [14] utilizes the size relation between the two coefficients in each pair. To be specific, while embedding message bit "1," if the first coefficient is not greater than the second, then swap their values, otherwise keep them unchanged; While for message bit "0," if the first coefficient is greater than the second coefficient, then swap there values, otherwise do nothing. After message embedding, the image block pixel values are reconstructed from the changed DCT coefficients through inverse DCT transform.

## 3 Proposed Scheme

### 3.1 Frequency Band Selection

Observing from (1), the computed DCT coefficients flow into three bands, namely the low-, middle-, and high-frequency bands. Generally, the energy of the image block mostly lies in the low-frequency band, and the middle- and high-frequency bands reflect the texture and edge details of the image block. DCT-based watermarking schemes usually utilize middle-band frequency for message embedding as it is less perceptible on modification, and the high-frequency band is rarely used because it is fragile to compression attacks.

The Histogram of Oriented Gradients (HOG) feature computes the horizontal and vertical gradients by simple high-pass filters and then, the direction of the gradient is resolved for each pixel, and later, the histogram of the gradient directions is counted for each block. HOG feature actually plays a high-pass filter role to the image block, such that modification on the high- or middle-frequency band will influence its computation. Seeing this connection, we propose to embed messages into the low-frequency band to minimize affection to the HOG features extraction. Moreover, we design a message size adaptive coefficients selection embedding scheme to decrease the perceptual distortion involved. The selected low-band DCT coefficients are depicted in Fig. 1, ranging from the second coefficient to the fifteenth coefficient according to the zigzag order.

### 3.2 Coefficient Pairs Selection

In [14], Soumitra et al. utilize the size relation between the two coefficients in each pair to embed message. For those pairs whose two DCT coefficients violate

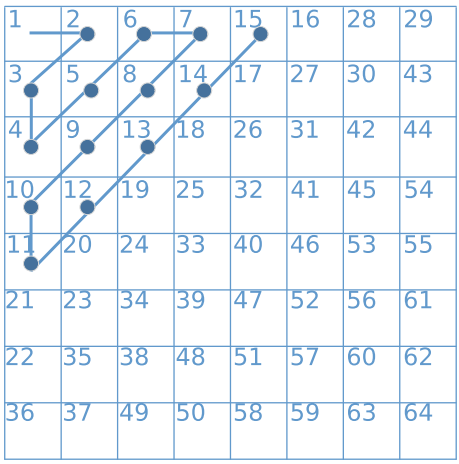



**Fig. 1.** Selected DCT coefficients of $8 \times 8$ image block following the zigzag order.

the order during embedding will swap their two coefficients, no matter how much they differ from each other.

As a matter of fact, whether the two coefficients in each pair swap their values or not, their absolute difference value stays invariant after message embedding. Utilizing this criterion, we can select coefficients pairs with small absolute difference to be embedded first to decrease the swapping distortion. Figure 2 shows the absolute difference value distribution between the two adjacent low-frequency band DCT coefficients in all the $8 \times 8$ image blocks for the Lena cover image. From Fig. 2, we can see that small absolute difference values between the two coefficients occurs more often than large absolute difference values, we can use a message length adaptive threshold to pick out close coefficient pairs for embedding. As a result, the coefficient pairs with large absolute difference values can be skipped, and the embedding distortion is better guaranteed.

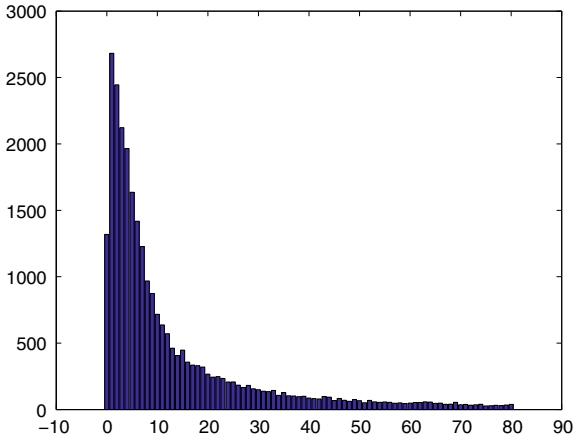



**Fig. 2.** Histogram of absolute difference value between the two adjacent DCT low-frequency band coefficients in all $8 \times 8$ image blocks for the Lena cover image.

### 3.3 Embedding and Extracting

To extract the message, we need to transfer the coefficient pairs filtering threshold to the receiver side. The overall message embedding and extraction procedure with adaptive coefficient pairs selection are described as follows.

**Embedding**

1. Divide the cover image into $8 \times 8$ blocks and apply discrete cosine transform(DCT) to each image block.
2. Define an initial threshold $\delta$ to filter all the $8 \times 8$ image blocks in the cover image, count all the adjacent DCT coefficient pairs whose absolute difference value is less equal than $\delta$. If the total count is less than the message payload length, increase the $\delta$ threshold by 1 and repeat the filtering procedure until enough DCT coefficient pairs are selected to embed the message.
3. For each DCT coefficients pair in each $8 \times 8$ image block, check where the absolute difference value between them is less equal than the threshold $\delta$, then a message bit can be embedded into this pair by:
   - While embedding message bit is "1," if the first coefficient is less than the second coefficient, then swap their values, else keep them unchanged.
   - While embedding message bit is "0," if the first coefficient is greater than the second coefficient, then swap their values, else keep them unchanged.
4. After embedding the message, apply the inverse discrete cosine transform to each image block to get the modified image block pixels, as a result, the whole watermarked image is obtained.
5. Actually, the threshold $\delta$ must be transmitted to the receiver side first, we simplify this through least significant bits (LSBs) replacement using the first several blocks in the cover image.

**Extracting**

1. Firstly, recover the threshold $\delta$ value from the LSBs in the first several blocks in the cover image.
2. Divide the cover image into $8 \times 8$ blocks and apply discrete cosine transform to each image block.
3. For each DCT coefficients pair in each $8 \times 8$ image block, check where the absolute difference value between them is less equal than the threshold $\delta$, then a message bit can be extracted from this pair by:
   - If the first coefficient is greater than the second coefficient, then extract message bit "1."
   - If the first coefficient is less than the second coefficient, then extract message bit "0."
4. After extracting each message bit, the whole message payload is recovered.

## 4 Experimental Results

To evaluate the embedding performance of the proposed DCT coefficients pair selection scheme, we conduct several kinds of simulations on different types of cover images. The cover image of size $512 \times 512$ is shown in Fig. 3, which varies from smooth ones to those full of textures and edge components. In our proposed scheme, low-frequency DCT coefficients from the second position toward the fifteenth position following the zigzag scanning order for each $8 \times 8$ block are

selected to embed messages. To demonstrate the effectiveness of the proposed scheme, comparisons between the proposed scheme and Soumitra et al.'s method in [14] are illustrated, in terms of visual imperceptibility, influence degree on HOG features, and also the compression robustness.

(a) Lena (b) baboon (c) barbara (d) boat

**Fig. 3.** Test cover images of size $512 \times 512$.

### 4.1 Imperceptibility Comparison

Peak signal-to-noise ratio (PSNR) is largely used to evaluate the similarity between the cover image and the watermarked image, which is defined as:

$$PSNR = 10 \times \log_{10} \frac{255 \times 255}{MSE} \tag{3}$$

where the mean square error(MSE) is given as $MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (x_{ij} - x'_{ij})^2$. Here, the $M$ and $N$ represent the width and height of the image, and $x_{ij}$ and $x'_{ij}$ mean the pixel values of the cover image and the watermarked image.

We compare our proposed DCT coefficients pair selection scheme against Soumitra et al.'s method in [14] with respect to PSNR criterion. Comparison results under different embedding rates are shown in Fig. 4, from which we can see that even though the proposed scheme gains distinctly better PSNR performance in comparison against method in [14].

### 4.2 HOG Feature Influence

The HOG feature is widely used for object detection, especially for pedestrian detection. We embed message into the cover images using both our proposed scheme and Soumitra et al.'s method in [14] and extract their HOG features. Later, the two modified HOG features are compared to the original HOG feature extracted from the cover image before embedding. The HOG feature similarity measurement(FSM) is computed by:

$$FSM = \sqrt{\sum_{i=1}^{L} (f_i - f'_i)^2} \tag{4}$$

where $f_i$ and $f_i'$ are the feature elements of the cover image and the watermarked image. Note, here we stitch all the block HOG features together to form a feature vector $f_i, i = 1, \ldots, L$, and $L$ is the overall vector length.

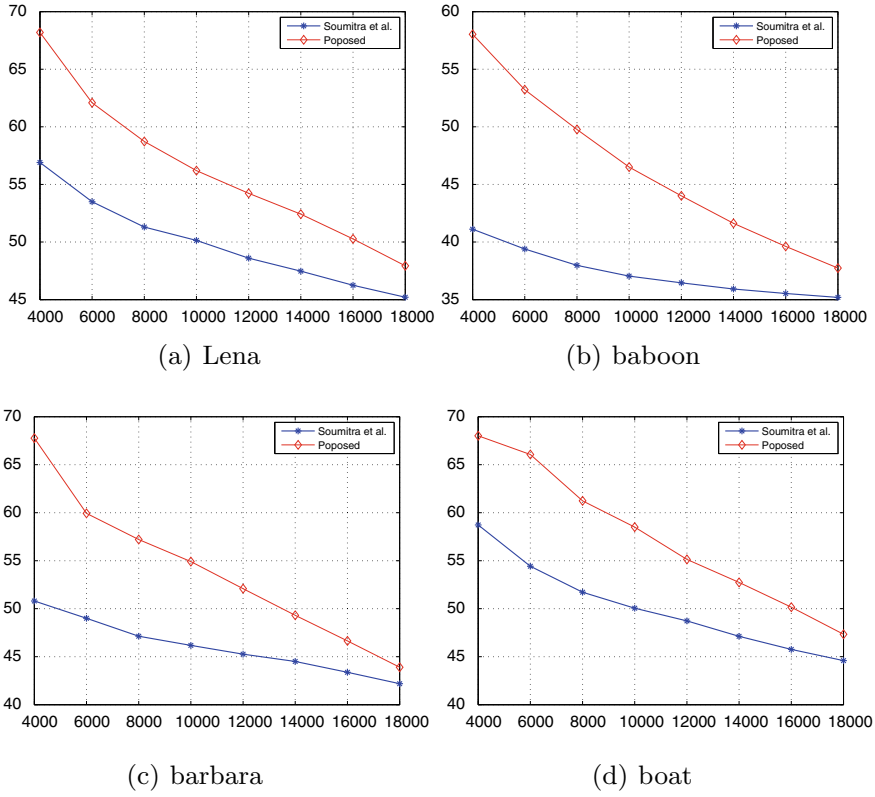



**Fig. 4.** Imperceptibility performance comparisons with Soumitra et al.'s method [14].

The HOG feature similarity measurement (FSM) comparison results are depicted in Table 1. The results exhibit our proposed method's superiority for HOG-like features' invariability after message embedding and thus make it more suitable for object detection and recognition scenarios.

### 4.3 Compression Robustness

JPEG compression is one of the most popular compression attacks during watermarked image transmission. As the low-frequency band coefficients are more robust to compression operations than those middle- and high-frequency band coefficients, our proposed embedding scheme will expect better robustness to JPEG compression attack theoretically, in comparison with a method in [14]. For the sake of simplicity, the experimental comparisons are omitted.

**Table 1.** HOG feature similarity measurement (FSM) comparisons.

| Image | Lena | Baboon | Barbara | Boat | Peppers | Airplane | Average |
|---|---|---|---|---|---|---|---|
| Soumitra et al. | 9.945 | 11.267 | 8.001 | 10.184 | 8.575 | 8.275 | 9.375 |
| Proposed | 9.231 | 9.446 | 7.350 | 8.907 | 8.417 | 8.079 | 8.572 |

## 5 Conclusion

In this paper, we proposed a novel DCT-based watermarking scheme for grayscale images, which utilizes the connection between discrete cosine transform (DCT) and the histogram of Oriented Gradient (HOG) feature extraction operation. We embed messages into the low-frequency band, and correspondingly, a novel coefficients pair selection scheme is constructed. The proposed scheme not only maintains the superiority of compression robustness but also keeps good visual quality for the watermarked image. Moreover, the proposed method is insensitive to HOG feature extraction, which makes the watermarked image more suitable for further objection detection and recognition scenarios.

**Acknowledgments.** This work is supported by Beijing NOVA Program (Z181100006218041) and National key R&D program of China (2017YFC0820106).

## References

1. Patra JC, Phua JE, Bornand C. A novel DCT domain CRT-based watermarking scheme for image authentication surviving JPEG compression. Digital Signal Process. 2010;20(6):1597–611.
2. Lin SD, Shie SC, Guo JY. Improving the robustness of DCT-based image watermarking against JPEG compression. Comput Stand Interfaces. 2010;32(1):54–60.
3. Hsu LY, Hu HT. Blind image watermarking via exploitation of inter-block prediction and visibility threshold in dct domain. J Vis Commun Image Represent. 2015;32(C):130–43.
4. Tao P, Eskicioglu AM. An adaptive method for image recovery in the DFT domain. J Multimedia. 2006;1(6):36–45.
5. Tsui TK, Zhang XP, Androutsos D. Color image watermarking using multidimensional fourier transforms. IEEE Trans Inf Forensics Secur. 2008;3(1):16–28.
6. Lang J, Zhang ZG. Blind digital watermarking method in the fractional fourier transform domain. Opt Lasers Eng. 2014;53(2):112–21.
7. Wang Y, Doherty JF, Dyck REV. A wavelet-based watermarking algorithm for ownership verification of digital images. IEEE Trans Image Process Publ IEEE Signal Process Soc. 2002;11(2):77–88.
8. Zhang G, Wang S, Wen Q. An adaptive block-based blind watermarking algorithm. In: International conference on signal processing, 2004. Proceedings. ICSP, Vol 3. New York: IEEE; 2004. p. 2294–7.
9. Liu N, Li H, Dai H, Guo D, Chen D. Robust blind image watermarking based on chaotic mixtures. Nonlinear Dyn. 2015;80(3):1329–55.

10. Chung K, Yang W, Huang Y, Wu S, Hsu Y. On SVD-based watermarking algorithm. Appl Math Comput. 2007;188(1):54–7.
11. Guo J, Zheng P, Huang J. Secure watermarking scheme against watermark attacks in the encrypted domain. J Vis Commun Image Represent. 2015;30:125–35.
12. Hernandez JR, Amado M, Perez-Gonzalez F. DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure. IEEE Trans Image Process Publ IEEE Signal Process Soc. 2000;9(1):55–68.
13. Agarwal C, Mishra A, Sharma A. Gray-scale image watermarking using GA-BPN hybrid network. J Vis Commun Image Represent. 2013;24(7):1135–46.
14. Roy S, Pal AK. A blind DCT based color watermarking algorithm for embedding multiple watermarks. AEU - Int J Electron Commun. 2017;72:149–61.
15. Hou D, Zhang W, Chen K, Lin SJ, Yu N. Reversible data hiding in color image with grayscale invariance. IEEE Trans Circuits Syst Video Technol. 2018;99:1.

# A Method for Under-Sampling Modulation Pattern Recognition in Satellite Communication

Tao Wen(✉) and Qi Chen

Battle Laboratory, Naval Command College, Nanjing 210016, China
1280691681@qq.com

**Abstract.** To solve the problem of reconnaissance and processing of broadband satellite communication signals, a kind of satellite communication signals BPSK/QPSK modulation pattern recognition method was put forward in this paper. This method deals with the satellite descending signal with BPSK/QPSK modulation in the under-sampling condition. Because the corrected spectrum of BPSK signal contains obvious crest, while QPSK signal does not contain this feature. The difference of the waveform characteristics is used to complete modulation pattern recognition. The simulation results show that this method can identify BPSK/QPSK modulation signals when SNR is greater than 1 dB. When the sampling points are reduced, the satellite communication signal under-sampling modulation pattern recognition method can still maintain good recognition performance.

**Keywords:** Satellite communication · Under-sampling · Modulation pattern recognition · Sparse reconstruction

## 1 Introduction

Satellite communication as a kind of high-frequency, large-broadband instant means of communication has the characteristics of wide coverage and signal transmission stability. It is an important means of battlefield communications at present. And noncooperative reception and processing of satellite communication signal are of great significance to grasp the trend of the space electromagnetic situation. The development of satellite communication anti-jamming technology, such as spectrum expansion, star processing, and limiting technology, presents a challenge to the noncooperative reception of satellite communication signals. However, the modulation pattern recognition process of satellite communication signals is not exactly the same as that of traditional communication signal modulation pattern recognition. First of all, the satellite communication signal has the characteristics of high frequency, large bandwidth, in noncooperative satellite communication signal processing, as a result of the signal sampling rate increase, the data quantity is big, have high requirements for ADC and data transmission. There are differences between satellite communication signal and traditional communication signal in transmission path. The transmission path of satellite communication has bigger longer noise influence , but also needs to take into


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_112

account the signal transmission in the atmosphere with the influence of the multipath effect, therefore, it is difficult to extract the characteristic parameters of the signal. In addition, it is an important characteristic parameter extraction method in the modulation pattern recognition of the satellite communication signal. The modulation types and parameters of satellite downlink communication signals can be obtained through modulation pattern recognition, and then useful information in satellite communication signals can be further obtained through demodulator.

Compressed sensing is an under-sampling acquisition and processing theory of signal, at the same time, the signal can be compressed by the acquisition system based on the compressed sensing theory framework [1]. Compressed sensing can retain the information in the original signal, the sampling rate is also far lower than that of the Nyquist sampling rate. Satellite communication signal has the characteristics of a large amount of data and difficult signal acquisition, and the problems need to be solved in under-sampling condition based on compressed sensing theory [2]. Therefore, this paper considers using under-sampling technology to get a small amount of sampling point satellite communication signal processing.

At present, there are two main types of signal modulated pattern recognition algorithms: likelihood ratio test and feature extraction algorithm. Literature 3 used in mixed likelihood ratio test (Hybrid likelihood ratio test , HLRT) method to complete BPSK/QPSK signal modulation recognition [3], the phase difference, and sequence of the signal is in the process of unknown parameters. This paper adopts maximum likelihood estimation, linear least squares estimation, and torque estimation method to estimate the unknown parameters. Furthermore, the whole process of HLRT recognition was completed. But the experimental environment was set in the additive white Gaussian noise (AWGN), which did not take into account the effect of the multipath effect, and the calculation was large. Feature extraction algorithms mainly include time–frequency method, wavelet transform method, and statistical extraction method [4]. Less prior information of the signal and low computational complexity are required in the feature extraction method, but it needs to solve the reliability problem of modulation style recognition performance. Literature 7 discussed several modulation pattern recognition algorithms based on feature extraction [5]. This method can not meet the requirements of satellite communication signal modulation recognition in sampling rate and signal-to-noise ratio.

BPSK/QPSK modulation is widely used in satellite communication due to its advantages of high-frequency utilization, strong anti-interference, and simple circuit implementation. A method for under-sampling modulation pattern recognition in satellite communication is proposed in this paper for the recognition of BPSK/QPSK modulation pattern recognition in satellite communication [6, 7]. This algorithm uses undersampling technology to reduce the sampling rate, and introduces the sparsity adaptive subspace tracking algorithm (SASP) in literature 9. As the reconstruction algorithm in the under-sampling technology, it improves the reconstruction accuracy and reduces the computational complexity at the same time [8]. Then, a method proposed in this paper for under-sampling modulation pattern recognition in satellite communication, using BPSK/QPSK signal differences of fixed frequency spectrum

waveform characteristic parameters, completed the recognition of BPSK and QPSK modulation signal.

## 2 Satellite Communication Signal Under-sampling Modulation Recognition Based on Modified Spectral Feature Extraction

The overall processing flow of BPSK/QPSK signal modulated pattern recognition technology based on under-sampling technology is shown in Fig. 1.

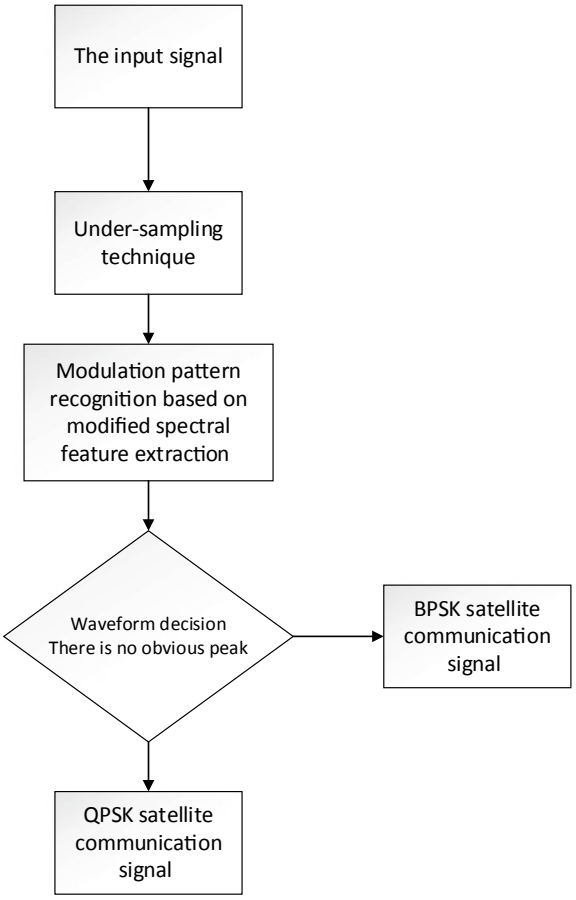



**Fig. 1.** BPSK/QPSK signal modulated pattern recognition technology based on under-sampling technology

As shown in the flow chart, this method is divided into three steps: first step, under-sampling technology processing, and the BPSK/QPSK satellite communication signal

with less sampling points. In the second step, the characteristic parameter extraction part, the reconstruction signal extracted from the first step is extracted and corrected. The third step is to identify the modulation pattern by comparing the modified spectral signature waveform.

### 2.1 Under-sampling Technology

The theory of under-sampling technology is based on the theory of compressive sensing: The sparse signal is not sampled in the sparse signal with the sparse signal in different domains. That is, the signal is compressed at the same time of sampling, and the process of under-sampling is realized by removing most of the redundancy in the signal. The main steps of under-sampling technique are sparse representation, random measurement, and reconstruction algorithm. Signals in nature can be represented by specific sparse $\{\Psi\}$ basis basically.

The received BPSK/QPSK signal $s$ can be regarded as a random process, which can be expressed as a linear combination of an orthonormal basis; signal $s$ can be expressed as:

$$s = \Psi x \tag{1}$$

where $x$ is the column vector of $N \times 1$, which is the projection coefficient $x = \langle s, \Psi_i \rangle$. Therefore, it can be seen that the projection coefficient $x$ and the signal $s$ are, respectively, the representation of the receiving signal in the $\Psi$ domain and the time domain, and the two are equivalent expressions.

In the process of compression sampling, $N$ visa number is expressed as $M$ visa, which naturally reduces the amount of data sampled. Construct the matrix $M \times N$ of $\Phi$, and the compression sampling value $y$ can be expressed as:

$$y = \Phi s \tag{2}$$

where the row vector of $\Phi$ is composed of $\{\Phi_i\}_{j=1}^{M} (M < N)$; $y_j = \langle s, \Phi_i \rangle$, $j = 1, \ldots, M$.

Figure 2 shows the schematic diagram of the basic principle of compression sampling.

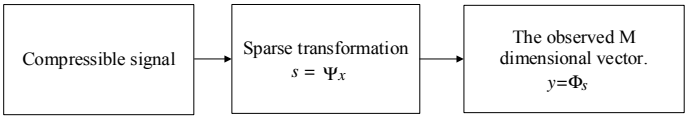



**Fig. 2.** Basic principle block diagram of compressed sampling

In addition, in order to ensure that the information in the original signal is not lost, the CS measurement matrix should satisfy conditions of the restricted isometry property (RIP). In general, random observation matrices composed of random elements can

meet the RIP conditions with a relatively large probability, for example, Gaussian distribution and Bernoulli distribution.

### 2.1.1 Sparse Representation

The observation matrix adopted in this paper can be equivalent to matrix form:

$$\Phi = HP \tag{3}$$

where $P$ is the orthogonal sparse matrix and $H$ is the random measurement matrix. In this paper, the discrete Fourier transform (DFT) is used as the orthogonal sparse matrix, and the positive transformation formula is as follows:

$$X(k) = \sum_{n=0}^{N=1} x(n) \mathrm{e}^{-j\frac{2\pi}{N}kn} = \sum_{n=0}^{N=1} x(n) W_N^{kn} \tag{4}$$

where $n,\ k = 0, 1, \ldots, N-1,\ W_N = \mathrm{e}^{-j2\pi/N}$.

This process can also be expressed through matrix operations:

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ \vdots \\ X(N-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N^1 & W_N^2 & \cdots & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N-1) \end{bmatrix} \tag{5}$$

After orthogonalization, the matrix form of discrete Fourier transform is obtained:

$$P = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N^1 & W_N^2 & \cdots & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix} \tag{6}$$

### 2.1.2 Random Measurement

The structure of the measurement matrix in the compressive sensing theory determines whether the information in the signal can be effectively collected. Normally, the measurement matrix should satisfy RIP measurement matrix properties to ensure that the sampling signal energy is unchanged before and after. At the same time, it need to be considered that the matching problem between the measurement matrix and measurement matrix reconstruction algorithm to ensure that the reconstruction effect optimization.

$H$ is the Gaussian random measurement matrix adopted in this paper, and an $M \times N$-size matrix $H$ is constructed, so that each element in $H$ is independent of the mean value of 0, and the variance is the Gaussian distribution, namely

$$H \sim N(1, \frac{1}{M}) \tag{7}$$

Gaussian measurement matrix is most widely used in the sparse reconstruction technology of measurement matrix. When meet the conditions of the $M \geq cK \log(N/K)$, the matrix which has strong randomness will meet great probability of RIP conditions [9].

### 2.1.3 Reconstruction Algorithm

Sparsity adaptive subspace tracking (SASP) uses a new sparse estimation method to obtain the initial estimate of sparsity and then updates the estimation by iteration. The weak matching principle is used in each iteration to select new atomic, again through the subspace tracking to improve result and reconstruction signals. The method will be sampling the sparse degree of adaptive subspace tracking to accomplish the signal reconstruction.

The sparse degree of adaptive subspace tracking (sparsity adaptive subspace pursuit, SASP) algorithm steps:

| |
|---|
| Input: $M$ dimension observation vector $y$, dimension measurement matrix |
| Output: reconstructed signal $\hat{x}$, $\hat{x}_{\Gamma^n} = \arg\min\|y - \Phi\Gamma^n \hat{x}_{\Gamma^n}\|_2^2$. The other elements in $\hat{x}$ are zero |
| (1) $g^0 = \Phi * y$, $K_0 = 1$ |
| (2) $\Gamma^0 = \{\|g_i^0\|\text{The maximum index value before } K_0\text{th}\}$ |
| (3) If $\|\Phi_{\Gamma^0} * y\|_2 < \frac{1-\delta_K}{\sqrt{1+\delta_K}}\|y\|_2$, then $K_0 = K_0 + 1$; repeat (2) |
| (4) $r^0 = \min\|y - F_{G^0} x_{G^0}\|_2^2$ |
| (5) $n = 1$ |
| (6) $g^n = \Phi * r^{n-1}$ |
| (7) $\hat{\Gamma} = \Gamma^{n-1} \cup \{\|g_i^n\|\text{The maximum index value before } \|\Gamma^{n-1}\|\text{th}\}$ |
| (8) $\hat{x}_{\hat{\Gamma}} = \arg\min\|y - \Phi_{\hat{\Gamma}} x_{\hat{\Gamma}}\|_2^2$ |
| (9) $\tilde{\Gamma} = \{\|x_{\hat{\Gamma}}\|\text{The maximum index value before } \|\Gamma^{n-1}\|\text{th}\}$ |
| (10) $\tilde{r} = \min\|y - \Phi_{\tilde{\Gamma}} x_{\tilde{\Gamma}}\|_2^2$ |
| (11) If $\|\tilde{r}\|_2^2 < \|r^{n-1}\|_2^2$, $\Gamma^n = \tilde{\Gamma}, r^n = \tilde{r}, n = n + 1$, repeat (6) |
| (12) When the correlation between atoms and residuals is less than a fixed value, otherwise $\hat{\Gamma} = \Gamma^{n-1} \cup \left\{ i : \|g_i^n\| \geq \alpha \max_{2 \leq j \leq N} \|g_j^n\| \right\}$, repeat (6) |

In the algorithm, the first four steps are estimated for the sparse degree, where $\Gamma^0$ is the initial estimation set and $r^0$ is the residual. As the main part of the iterative algorithm, steps 5–12, where $n$ is the number of iterations, the stopping condition of the

algorithm iteration is that the correlation between the atoms and residuals in step 12 is less than some fixed value.

At this point, it is possible to reconstruct signal $\tilde{x}(n)$, $n = 1, 2$, in less than the conditions of Nyquist sampling rate. Figures 3 and 4 are the simulation waveforms of sparse reconstruction process of BPSK/QPSK modulated signals when sampling points are 1000 and the sampling rate is 1/4. The results show that BPSK/QPSK signals can be sparse represented by discrete Fourier transform, and the original signal can be recovered by the sparse adaptive subspace tracking (SASP) reconstruction algorithm, and the reconstruction error is less than 10%.

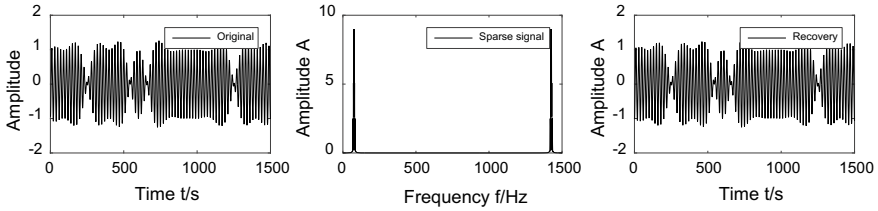



**Fig. 3.** Sparse reconstruction simulation waveform of BPSK signal

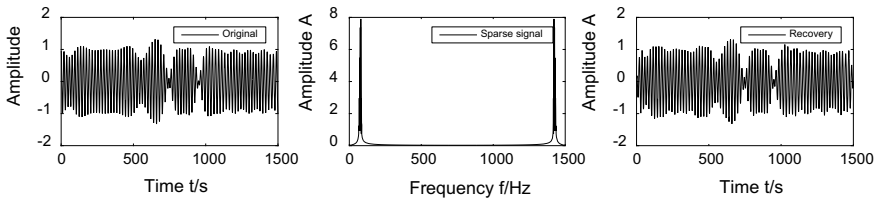



**Fig. 4.** Sparse reconstruction simulation waveform of QPSK signal

## 2.2 BPSK/QPSK Satellite Communication Signal Under-sampling Modulation Pattern Recognition

### 2.2.1 BPSK/QPSK Signal Model

The mathematical model of BPSK/QPSK modulation signal can be expressed as:

$$s(n) = A\exp(j[\phi(n)]),\ 0 \le n \le N-1 \tag{8}$$

where $s(n)$ represents BPSK/QPSK modulation signal, $A$ is the modulation signal amplitude, $\phi(n)$ is the phase function, and $N$ is the number of sampling points. For phase modulation, the phase function $\phi(n)$ reflects the different modulation styles.

$$\phi(n) = 2\pi f_0 n\Delta t + \pi d_i(n) + \theta \tag{9}$$

The multipath effect of signal in transmission process is an important problem in satellite communication and reconnaissance. The multipath effect is mainly caused by

multipath reflection in communication signal transmission path. The influence of synchronous timing error and incomplete matched filtering of receiver can also be equivalent to multipath model [10]. The discrete equivalent model of multipath fading channel can be expressed as:

$$h(n) = h_0 e^{j\theta_0}\delta(n) + \sum_{k=1}^{L} h_k e^{j\theta_k}\delta(n-k) \tag{10}$$

In the above equation, the number of channels in the channel is $L + 1$, the amplitude gain of $h_0$ main path, and the phase shift factor of $\theta_0$, while $h_k$ and $\theta_k$ are, respectively, the amplitude gain and phase shift factor of the KTH multipath.

This paper considers the simulation experiment scenario for mobile satellite communication in a wide area. The direct signal is basically unblocked, the direct wave envelope is a fixed value. The signals containing the direct waves and the diffuse components subject to Rayleigh distribution are subject to the rice distribution. At this point, the satellite channel is a linear time-varying channel, and the response varies with time. The effect of this interference on useful signals is multiplicative interference. Suppose the effective signal is $s(n)$, $w(n)$ as multiplicative interference, $e(n)$ as the additive white Gaussian noise, received signal $x(n)$ can be expressed as:

$$x(n) = s(n) \cdot w(n) + e(n) \tag{11}$$

### 2.2.2 Correction of Spectral Feature Extraction

Through the under-sampling technique, the reconstructed signal $\tilde{x}(n)$ of input signal $x(n)$ is obtained. The next step is to extract the characteristic parameters of the reconstructed signal.

First, the signal is squared, namely

$$y(n) = \tilde{x}^2(n) \tag{12}$$

If the observed signal is BPSK signal, then $y(n)$ is converted into $y(n)$ sine wave with phase $2\theta$ and carrier $2f_0$. If it is QPSK signal, it degenerates into BPSK signal [11, 12]. At this time, the modulation pattern recognition of BPSK/QPSK signal was translated into the recognition of the sinusoidal affected BPSK signal.

The frequency form of sine wave signal is concentrated in a certain frequency. BPSK signal spectral waveform is subject to sinc function distribution, bandwidth is affected by the symbol rate. Therefore, the former on the spectrum distribution is single frequency on a separate line, while the latter should be on the spectrum distribution peak shape distribution; therefore, this difference can be used for BPSK/QPSK signal modulation style recognition.

For the reconstructed signal $\tilde{x}(n)$, discrete Fourier transform (DFT) is taken and the model is taken, and $\widehat{X}(n)$ is defined as the modified spectrum:

$$\widehat{X}(n) = |\mathrm{DFT}[y(n)]| = \left| \sum_{n=0}^{N-1} W_N^{-nk} \right|, \ 0 \le k \le N-1 \tag{13}$$

where $W_N = \mathrm{e}^{j2\pi/N}$.

Figure 5 shows the modified spectra of BPSK/QPSK two signal squares, respectively, and the simulation conditions are: Gaussian noise intensity is 3 dB, the carrier frequency is 500 Hz, the symbol rate is 100 bps, element number is 100 bit, the modulation modes are BPSK and QPSK. Figure 5 shows that the BPSK signal is reduced to sine wave by square transformation. The QPSK signal is degraded to BPSK signal by square transformation. When the BPSK signal and QPSK signal are represented in the modified spectrum, it can be seen that the BPSK signal has a single spectral line component, and QPSK signal is without a single component of the spectrum. After removing the single line component, BPSK signal peak is not obvious, and QPSK signal has obvious peak value.

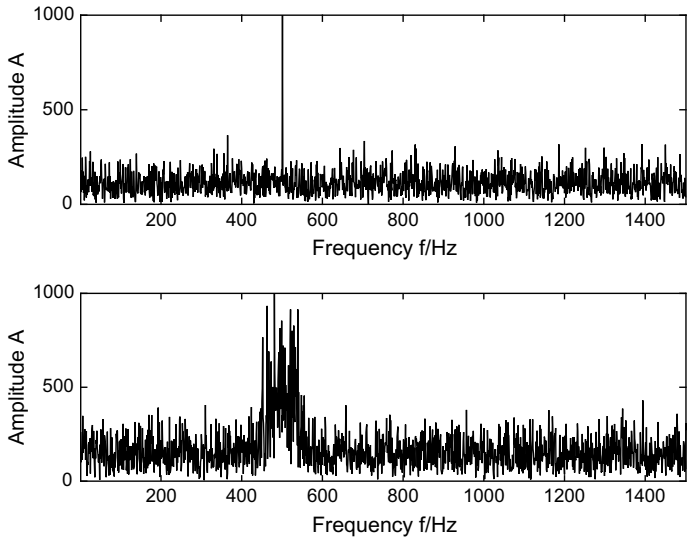



**Fig. 5.** Comparison of the corrected spectrums after the square of BPSK/QPSK

### 2.3 Modulation Pattern Recognition

BPSK signal transforms back a sine signal through the square transformation, QPSK signal through the square transforms back into Sinc function. Because of the effect of channel noise and multipath fading effect, after receiving the under-sampling of signal $x(n)$, the modified spectrum characteristic waveform $R(n)$ of the extraction $x(n)$ can be expressed as:

$$R(n) = \Omega(n) + N(n) \tag{14}$$

where $\Omega(n)$ is the effective characteristic waveform, and $N(n)$ is the noise waveform. When the input signal is BPSK signal, $R(n)$ contains a single spectral line, and the waveform is shown as a single spectral line distributed in the noise waveform. When the input is QPSK signal, $\Omega(n)$ still contains the single spectral line, but also contains the sinc function waveform. At this time, the modified spectrum characteristic waveform $R(n)$ shows a single spectral line and peak amplitude at the same time distributed in the noise waveform.

Therefore, in order to use the characteristics of the BPSK/QPSK signal waveform to realize the modulation pattern recognition, first remove the single line, the two kinds of characteristic waveforms due to the characteristics of the QPSK signal waveform contains obvious peak value, then sentence by setting the threshold for BPSK/QPSK signal modulation style recognition.

## 3 Simulation Analysis

Let's say the input signal $s(n)$ style for BPSK and QPSK modulation. After the transmission of the mobile satellite channel, the output signal $x(n)$ obeys the rice distribution and contains additive white Gaussian noise through under-sampling based on compression perception processing, and reconstruction signal $\tilde{x}(n)$ is obtained by less sampling point. The simulation experiment will analyze the performance of the satellite communication signal under-sampling modulation recognition method proposed in this paper under different conditions. The recognition rate of BPSK signal is $P(I_0)$, the recognition accuracy of QPSK signal for $P_t = (P(I_1) + P(I_2))/2$, has the performances evaluation for simulation of BPSK using 13 Barker code, QPSK with 14 Taylor code, number of simulation experiments/group is set to 100 times.

### 3.1 Under-sampling Modulation Pattern Recognition Simulation Experiment Under Different SNR Environment

In the process of satellite communication signal processing, the influence of damping and noise in the channel is a must to consider. Therefore, in order to test the performance of the method in satellite communication channel, a simulation experiment was carried out in conditions of different SNRs. The simulation condition is: carrier frequency is 75 MHz, element width is 800 ns, sampling rate $W = 1/2$, simulate the multipath effect, les channel attenuation factor intensity to 2 dB, gaussian noise intensity change interval [−8 dB, 8 dB], the step length is 1 dB.

The simulation results are shown in Fig. 6; under the influence of the multipath effect, under-sampling satellite communication signal modulation pattern recognition method in the signal-to-noise ratio of 3 dB complex still has 87% of the recognition accuracy in noisy environment; BPSK/QPSK modulation signal can be stabilized when SNR is greater than −1 dB.

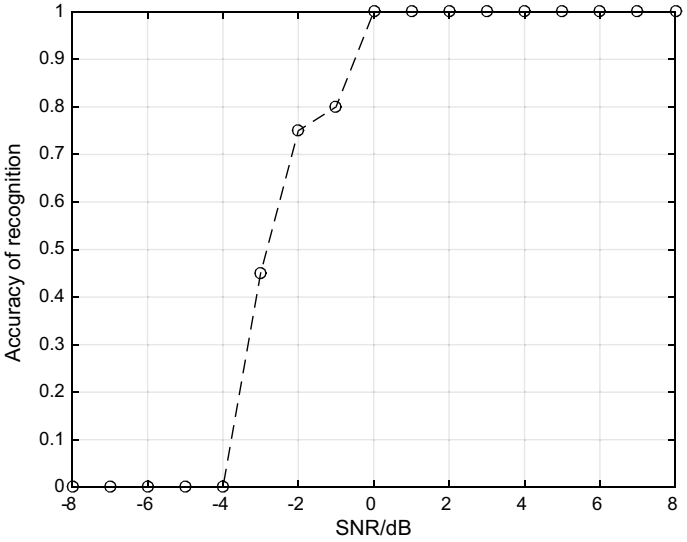



**Fig. 6.** Recognition rate of satellite communication signal under-sampling modulation pattern recognition method in different SNR environments.

### 3.2 Analysis of Under-sampling Modulation Pattern Recognition Performance

In order to verify the effect of the method in this paper under the condition of under-sampling. The performance analysis of under-sampling modulation pattern recognition was carried out. The simulation conditions is: the carrier frequency is 75 MHz; element width is 800 ns; les channel factor intensity is 2 dB; gaussian noise intensity is 2 dB.

Figure 7 shows the simulation results. The traditional sampling frequency conforms to the modulation pattern recognition method of Nyquist sampling theorem, and the sampling points are 1000, 2000, 3000, and 4000, respectively. Simulation results are shown in Fig. 7b, the modulation pattern recognition algorithm based on the technology of under-sampling recognition performance under different conditions of sample points, sampling points respectively in 1000, 2000, 3000, 2000, and other conditions of the same recognition performance experiment. The simulation results show that when the sampling points are reduced, the recognition accuracy of the unsampled modulation pattern recognition method can maintain the recognition performance better when the SNR changes.

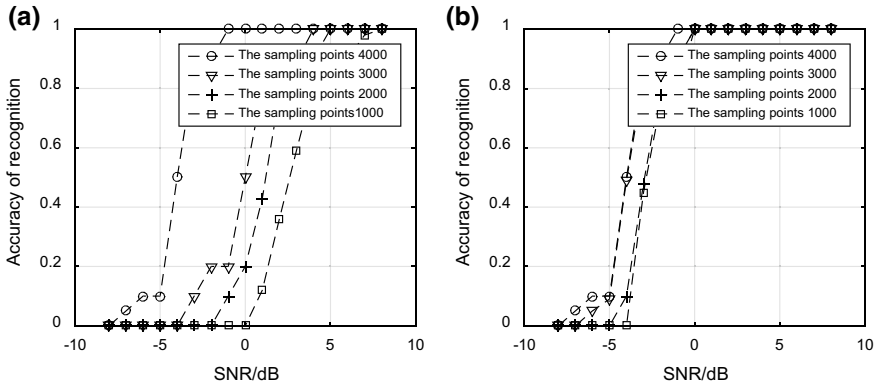



**Fig. 7.** Performance comparison of recognition rate under different sampling points

## 4 Conclusion

In this paper, a method is proposed to identify the under-sampling modulation based on feature parameter extraction. This method mainly aims at BPSK/QPSK modulation pattern recognition of communication signals. The methods in this paper considers the influence of the satellite communication channel multipath effect joins the additive gaussian noise at the same time to join the multiplicative noise, simulates the les channel attenuation phenomenon, in the process of signal source generated. In the signal sampling phase, sparse reconstruction of under-sampling technique is used to sample and reconstruct the signal with a lower sampling rate to solve the problem of a large sampling difficulty of satellite communication data. Finally, the spectrum characteristic waveform of the reconstructed signal was extracted, according to the characteristic difference of corrected spectral waveforms, and the BPSK/QPSK signal modulation pattern recognition was realized. The simulation results show that this method can be applied to the modulation pattern recognition in satellite communication channel environment and has a strong robustness.

**Fund Project.** National Natural Science Foundation of China (youth project): No. 61501484; Research and Development Fund of Naval Engineering University (Science and Technology [2016] No. 66), accounting subject: 425517K170; 425517K167.

## References

1. Xiang Q, Cheng B, Yu J. Hierarchical porous CdS nanosheet-assembled flowers with enhanced visible-light photocatalytic $H_2$-production performance. Appl Catal B Environ. 2013;138–139:299–303.
2. Huan X, Safta C, Sargsyan K, et al. Compressive sensing with cross-validation and stop-sampling for sparse polynomial chaos expansions. arXiv preprint arXiv:1707.09334. 2017.
3. Sejdić E, Orović I, Stanković S. Compressive sensing meets time–frequency: an overview of recent advances in time–frequency processing of sparse signals. Digital Sig Process. 2017.

4. Sur P, Chen Y, Candès EJ. The likelihood ratio test in high-dimensional logistic regression is asymptotically a rescaled chi-square. arXiv preprint arXiv:1706.01191. 2017; Xu JL, Su W, Zhou M. Likelihood-ratio approaches to automatic modulation classifigation. IEEE Trans Syst Man Cybern, Part C (Appl Rev). 2011;41(4):455–69.
5. Gong W, Huang K. An image-feature based method for feature extraction of intra-pulse modulated signals. Electron Opt Control. 2008;4:016.
6. Guo-Bing H, Yu L. Signal intrapulse modulation recognition algorithm based on sine wave extraction. Comput Eng. 2010;36(13):21–3.
7. Ming J, Guobing H. Intrapulse modulation recognition of radar signals based on statistical tests of the time-frequency curve. In: 2011 international conference on electronics and optoelectronics (ICEOE). IEEE; 2011, 1, p. V1-300–V1-304.
8. Yang L, Guobing H. An improved BPSK/QPSK signal modulation pattern recognition algorithm. Telecommun Eng, 2017;57(8).
9. Zefang XU, Shunlan LIU. Adaptive regularized subspace pursuit algorithm. Comput Eng Appl. 2015;51(3):208–11.
10. Mishali M, Eldar YC. From theory to practice: sub-Nyquist sampling of sparse wideband analog signals. IEEE J Sel Top Sig Process. 2010;4(2):375–91.
11. Simon MK, Alouini MS. Digital communication over fading channels. Hoboken: Wiley; 2005.
12. Guobing H, Xu J, Li Y, et al. Modulation radar signal analysis and processing technology of. Boston: Addison Wesley Publishing; 2014.

# Sequential Modeling for Polyps Identification from the Vocal Data

Fangqi Zhu[1], Qilian Liang[1], and Zhen Zhong[2(✉)]

[1] Department of Electrical Engineering, University of Texas at Arlington, 416 Yates St., Arlington, TX 76010, USA
fangqi.zhu@mavs.uta.edu, liang@uta.edu
[2] Department of Otolaryngology Head and Neck Surgery, Peking University First Hospital, Beijing 100034, China
{Zhong_zhen,Zhen_zhong}@sina.com

**Abstract.** Given the revival of neural networks and its recent impact in other disciplines and record-breaking performances in a variety of applications, in this paper, we employed a deep sequential model for polyps detection from the vocal data. Previous research of acoustic signal recognition (ASR) has focused on hand-crafted machine learning fashion, such as Mel-frequency cepstral coefficients with hidden Markov model and Gaussian mixture model. The deep model demonstrates its flexibility and potential to outperform the traditional methods, and we expand its scope on medical symptom identification. The mapping between the raw vocal signal and the symptom recognition is established, and we show that we can achieve a good recognition accuracy, which may appear to clinical diagnosis in the near future.

**Keywords:** Vocal features · Polyps · Sequential model · LSTM

## 1 Introduction

Classical model for time-series modeling is mainly based on statistical analysis, state-space description, and dynamical causality. From the statistical point of view, a Cox proportional hazard regression model for predicting stroke and heart disease was proposed [1]. The milestone work of Rabiner summaries the development of hidden Markov Model, which embedded the observation with hidden layer and used the transformation in the hidden states for inference [2]. A review of Wiener–Akaike–Granger–Schweder (WAGS) influence, state-space method, and dynamic casuality method for fMRI data is discussed [3]. Although all these three kinds of methods make a huge progress on some specific tasks, they rely on lots of hand-crafted details and assumptions, which may not always be satisfied in the real applications.

The emergence of deep learning has a great impact on the medical and biological fields. One of the classical but still challenging research branches is how to use


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_113

the neural network to achieve sequential modeling and inference. For example, a fuzzy logic-induced feed-forward neural network for sequential classification is proposed to handle the inherent uncertainty [4]. The DeepBind framework, which contains convolution and pooling layer for its neural network, achieves good performance on the DREAM5 evaluation for DNA and RNA sequence specification [5].

A more suitable network structure for time-series modeling is the recurrent neural network, especially the long short-term memory (LSTM) network [6–8]. It exhibits impressive performance in numerous sequence-based tasks such as speech recognition, acoustic modeling of speech, knowledge graph, and machine translation. LSTM is also a flexible and extensible structure, and different variations of the LSTM have been proposed in terms of different aspects such as saving the computation and memory cost, increasing the adaption for longer sequences, and tailored to fit more tasks [9]. The Granger causality analysis with $l1$-norm constraints on the weights of neighboring input periods leads to a lasso regularization on the LSTM network [10], which optimizes the structure of the LSTM to save computation and memory resources.

Some previous interdisciplinary work between RNN and pathophysiologic analysis has validated RNN's ubiquitous property. For instance, the gated recurrent unit network, one of the simplified versions of LSTM, has been employed for early detection of the heart failure onset [11]. In this paper, the sequential modeling for polyps identification from the vocal data is investigated. Twelve respondents, four of which have throat polyps and the rest are healthy (no throat polyps), are recorded for producing two vowels with high sampling rate. We use time-frequency analysis on these data and they are trained with sequential deep learning model. The main contributions of this paper are as follows:

- A concrete real vocal dataset for polyps identification is provided, and an easy-to-employ temporal frequency transform is utilized.
- Deep sequential learning model with dropout regularization is employed on the two specific temporal frequency features.
- Outstanding performance compared to previous work [12], which provides promising approach for clinical application.

The rest of the paper is arranged as follows: In Sect. 2, the vocal experiment and dataset description are demonstrated for formulating the problem. In Sect. 3, the deep sequential model is employed to learn the coefficients of time frequency analysis and achieve the polyps identification. In Sect. 4, the performance of the training will be demonstrated, and finally, we conclude in Sect. 5.

## 2 Vocal Experiment and Dataset Description

One of the typical symptoms among adults is the vocal cord polyps, which is the result of an acute injury (such as yielding and shouting really loud) and or several other causes (such as cigarette smoke). The original diagnosis of this problem is from the pathological viewpoint, and it is highly depended on the expert knowledge, which is not suited to the urgent requirement of self-diagnosis currently. In principle, sounds generated by human beings are filtered by the structure of the vocal tract which comprises of parts like tongue, teeth, etc., which is nonlinear processing comprising of both physical and aerodynamic effects. This kind of structure leads to what kind of sound will come out.

The recording of the voice for automatic diagnosis of the problem is achieved based on the speech analysis. We collect the data that are about the discrete vocal samples of recording twelve respondents, in which four of them have throat polyps and the rest are healthy (no throat polyps). In the study, the vowels /a:/ and /i:/ are collected and each of them lasts for approximately 10 seconds. The samples of vowels within the same scenarios are snipped off a longer vowel sample to choose the stable part of the whole sample for processing. The sampling rate of the vocal data is 192,000, and bits per sample is 32. The details of the vocal dataset are summarized in Table 1. The time and frequency of vocal data for four cases (/a:/ polyps, /a:/ non-polyps, /i:/ polyps and /i:/ non-polyps) are shown in Fig. 1.

**Table 1.** Vocal dataset summary

| Parameters | Values |
|---|---|
| No. of patients | 12 (4 with polyps, 8 without polyps) |
| Age range | 21–50 |
| Sample rate | 192,000 |
| Samples range | 1,157,200–1,344,000 |
| Bits per samples | 32 |
| Channels | 1 |
| Normalized data | FALSE |

As shown in Fig. 2, it is difficult to find features or correlation directly in the time domain, but we are able to distinguish the difference in the frequency domain. Therefore, we resort to display the latent features in the temporal frequency domain. The main components in the frequency domain lie within 0–2000 Hz, and we employ a rectangle window to filter out the rest of the part and adopt the short-time Fourier transform (STFT) to form a spectrogram. Such sampled window clipped STFT defined over the region $m \in [0, R-1]$ is given by [13]:

$$X_{STFT}[k, lL] = X_{STFT}\left(e^{j2\pi k/N}, lL\right) = \sum_{m=0}^{R-1} x[lL - m]w[m]e^{-j2\pi km/N}, \quad (1)$$

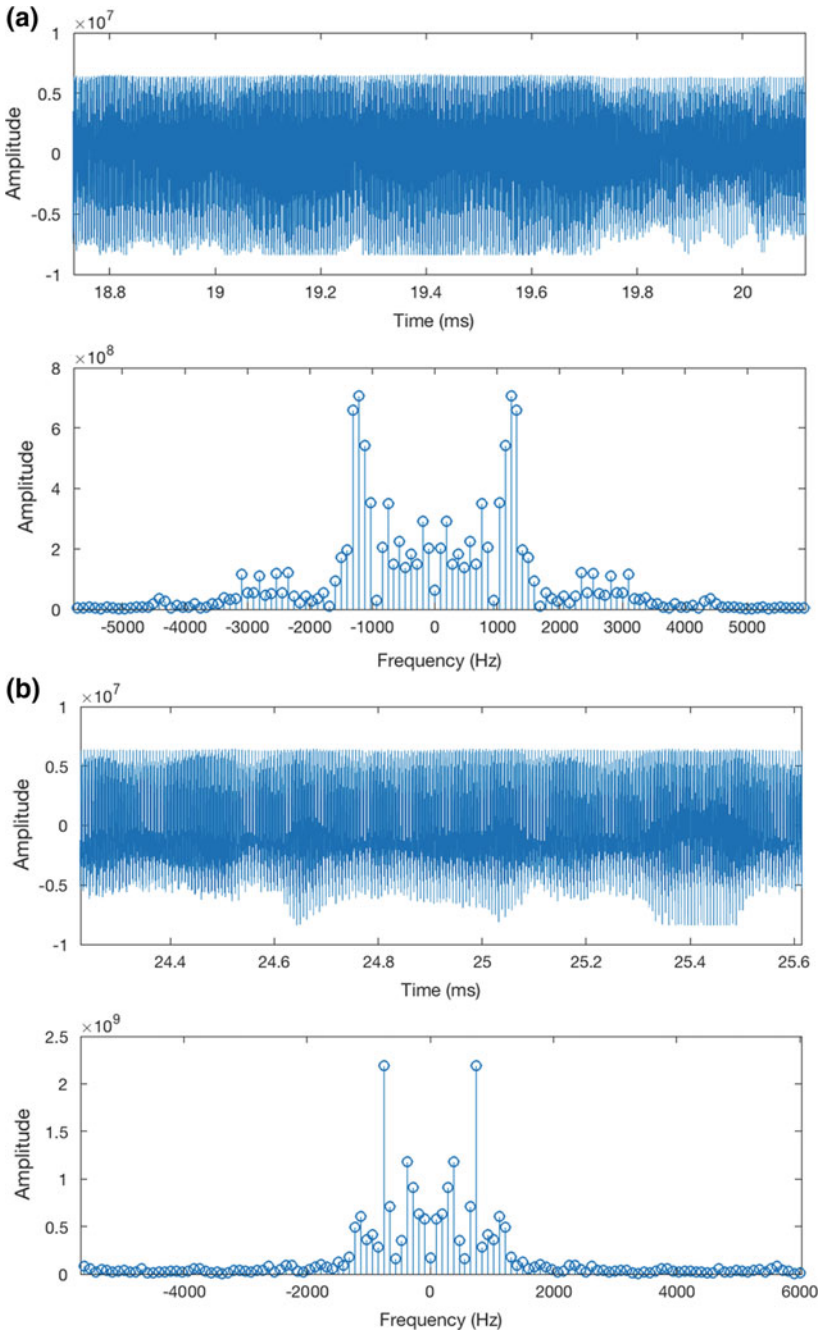



**Fig. 1.** Time and frequency of the vocal signal for vowels, **a** /a:/ normal; **b** /a:/ abnormal(polyps); **c** /i:/ normal; **d** /i:/ abnormal(polyps)

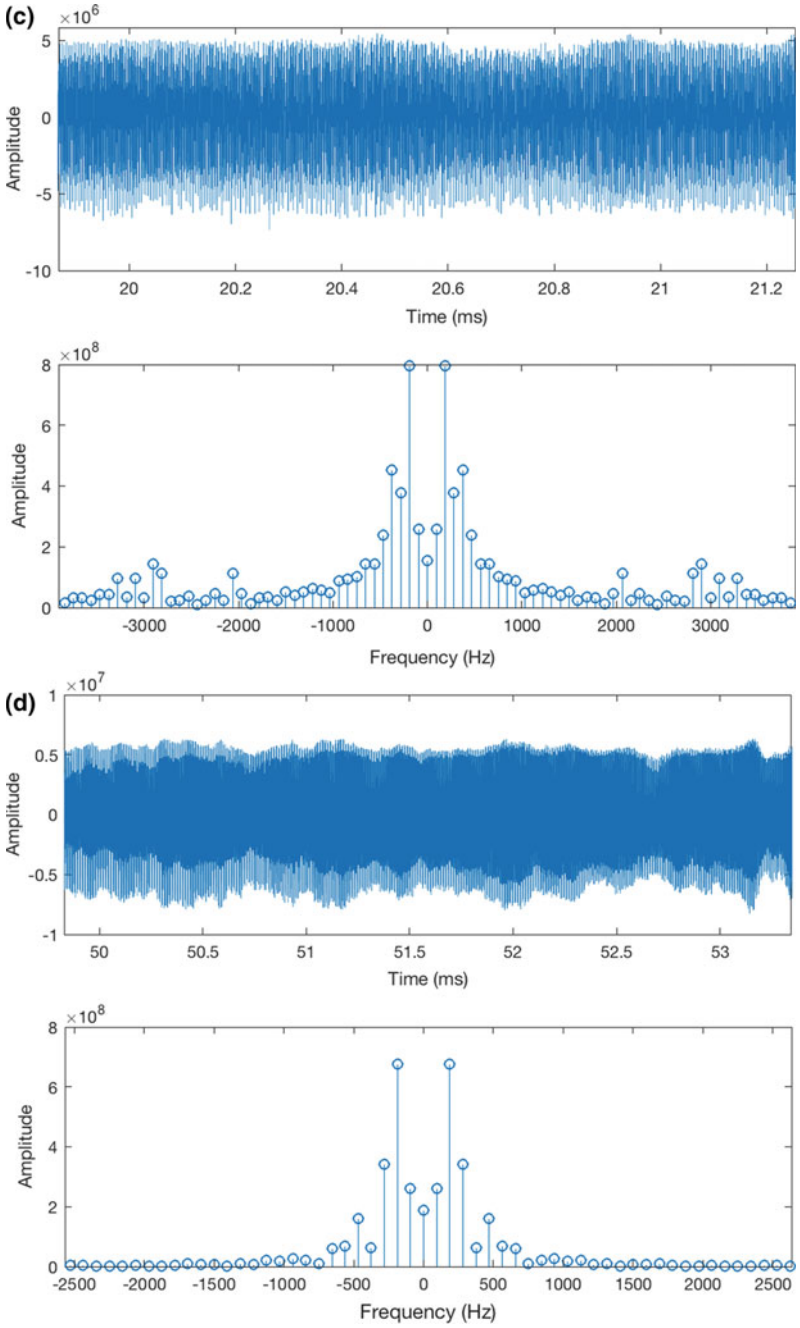



**Fig. 1.** (*continued*)

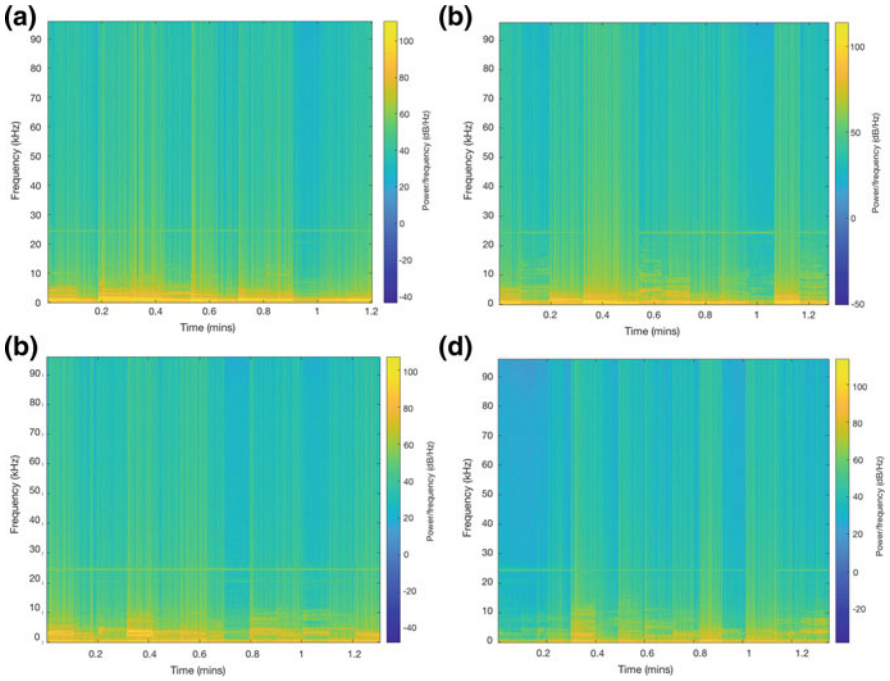



**Fig. 2.** Spectrogram (STFT) of the vocal signal for vowels, **a** /a:/ normal; **b** /a:/ abnormal(polyps); **c** /i:/ normal; **d** /i:/ abnormal(polyps)

Another widely adopted feature of ASR is the Mel-frequency cepstral coefficient (MFCC), which was proposed by Davis and Mermelstein in the 1980s.[1] The main steps are as follows:

- Apply the Mel filterbank to the power spectra, and sum the energy in each filter. $Y(i) = \sum_{k=0}^{N/2} \log \mid s(n) \mid H_i\left(k\frac{2\pi}{N}\right)$. ($S(n)$ is the DFT of the input signal, $H_i$ is the band filter, $N$ is the frame length, and $N'$ is the number of points in short-term DFT with zero padding)
- For each frame, calculate the periodogram estimate of the power spectrum.
- Frame the signal into short frames and take the Fourier transform
- Keep DCT coefficients 2–13, and discard the rest (can be adjusted accordingly).
- Take the logarithm of all filterbank energies, and take the DCT of the log filterbank energies.

## 3 Deep Sequential Modeling for Classification

LSTM is an efficient tool to selectively read, write, and forget the sequential input by leveraging the gated control mechanism. The LSTM framework that is

[1] We follow the implementation from https://github.com/jameslyons/python_speech_features.

adopted in our model contains the five modules. The input module is inspired by the requirement to choose what kind of information that we want to update, which can be expressed as follows:

Since for an RNN, it is difficult to own a very long memory and also it has constraints due to the problem of vanishing and exploding of the gradients, it is necessary to introduce a mechanism to selectivity lose some information which has less contribution to the learning task.

$$i_t = \sigma(W_i s_{t-1} + U_i x_t + b_i) \tag{2}$$
$$f_t = \sigma(W_f s_{t-1} + U_f x_t + b_f) \tag{3}$$
$$o_t = \sigma(W_o s_{t-1} + U_o x_t + b_o) \tag{4}$$
$$\tilde{s}_t = \phi(W(o_t \odot s_{t-1}) + U x_t + b) \tag{5}$$
$$s_t = f_t \odot s_{t-1} + i_t \odot \tilde{s}_t \tag{6}$$

where the $\sigma$ is the sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$ and $\phi$ is the hyperbolic tangent function ("tanh") $\phi(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. $W_*$ is the weight matrix from the input to the hidden layer, $U_*$ is the recurrent state transition matrix, $b_*$ is the bias vector, and $\odot$ stands for the element-wise product.

For the cost function, the well-known cross-entropy function is a proper choice for this target detection and data retrieval task, since we can model both the problems as classification-like problems. Given $M$ input–target pairs $\mathbf{x}^{(i)}, y^{(i)}, i = 1, \ldots, M$ and the estimation of the distribution for the $i$th sample $\hat{y}^{(i)}$, the discrete binary cross-entropy can be expressed as follows:

$$L(\mathbf{x}, y) = -\sum_{i}^{M} \left( y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)}) \right) \tag{7}$$

## 4 Performance Analysis

The datasets contain STFT coefficients and MFCC coefficients for both vowel /a:/ and vowel /i:/, for normal and abnormal situations. For both of the two data, they are separately imported into the LSTM networking and will be split as 60% : 20% : 20% for training, validation, and testing. The amount of data for STFT are larger than that of MFCC, and the performance of the two features is compared. The implementation of the whole program is accomplished by leveraging the Keras 2.0 with TensorFlow backend. The program is running on the Imac i7 4 GHz with 16 GB RAM.

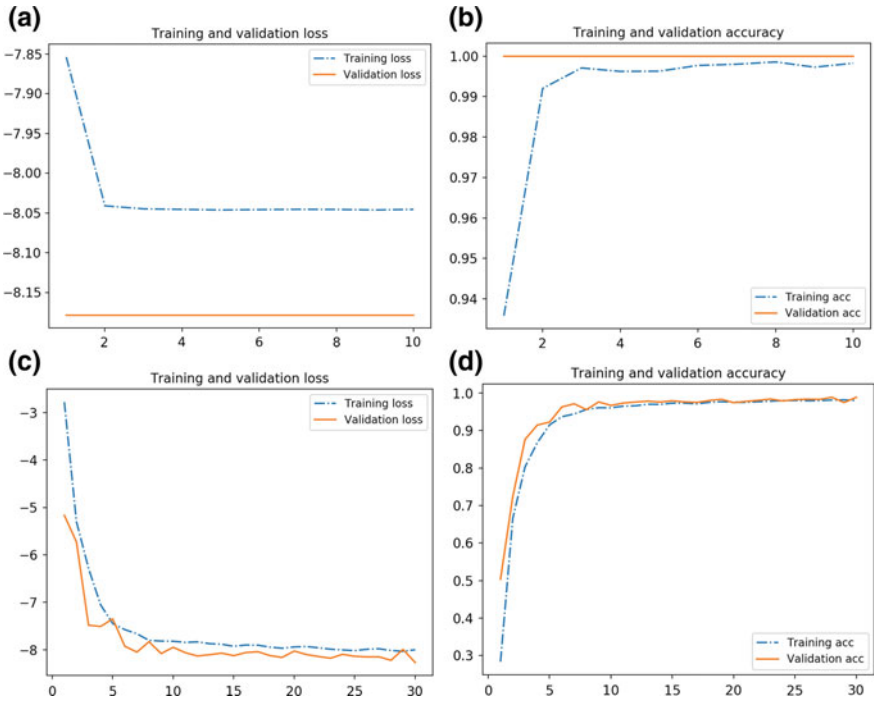



**Fig. 3.** Training and testing performance of the deep sequential model of **a** loss of STFT features; **b** accuracy of STFT features; **c** loss of MFCC features; **d** accuracy of MFCC features

A 32-unit LSTM is employed with dropout ratio of 0.4. A dense layer with hyperbolic tangent activation function is built on top of it. The Adam optimization method is adopted for iterations. The training process contains 10 epochs with minibatch size of 32. The random seed is set so that we do not need to run multiple times and average (it is still able to run average as well). The training performance and validation performance for both STFT and MFCC are shown in Fig. 3. The training accuracy approximates stable after 6 epochs and the loss approximates the lower bound after 2 epochs for STFT, while for MFCC, since the datasets are much smaller than STFT, it shows oscillation but eventually achieves a good performance, which shows the efficiency of the LSTM network.[2] The testing loss and accuracy are shown in Tables 2 and 3 for both the STFT features and MFCC features.

[2] The corresponding code will be uploaded to Github for open source purpose.

**Table 2.** Training, validation, and testing performance of the deep sequential model of STFT features

| Performance index | vowel /a:/ | vowel /i:/ |
|---|---|---|
| STFT training accuracy | 0.9983 | 0.9998 |
| STFT training loss | −8.0454 | −8.0449 |
| STFT validation accuracy | 1.0000 | 1.0000 |
| STFT validation loss | −8.1784 | −8.1784 |
| STFT test accuracy | 1.0000 | 1.0000 |
| STFT test loss | −7.9814 | −7.9784 |

**Table 3.** Training, validation, and testing performance of the deep sequential model of MFCC features

| Performance index | vowel /a:/ | vowel /i:/ |
|---|---|---|
| MFCC test accuracy | 0.9896 | −0.9778 |
| MFCC test loss | −8.1018 | −7.8884 |
| MFCC training accuracy | 0.9794 | 0.9774 |
| MFCC training loss | −8.0009 | −7.8938 |
| MFCC validation accuracy | 0.9886 | 0.9796 |
| MFCC validation loss | −8.2706 | −7.9664 |

## 5 Conclusion

In this paper, the polyps detection based on the vocal samples is solved by leveraging the deep sequential model. Two types of preprocessing methods, i.e., the STFT and MFCC, are adopted to project the raw data on new domain to demonstrate the inherent information. The problem is formulated as the binary classification task, and the LSTM network with Adam optimization is able to achieve good performance on the binary classification task, with high accuracy and fast convergence.

**Acknowledgments.** This work was supported in part by NSFC under Grant 61771342, 61731006, 61711530132, and Tianjin Higher Education Creative Team Funds Program.

## References

1. Wang TJ, Massaro JM, Levy D, et al. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the framingham heart study. J Am Med Assoc. 2003;290(8):1049–56.
2. Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE. 1989;77(2):257–86.
3. Roebroeck A, Seth AK, Valdes-Sosa P. Causal time series analysis of functional magnetic resonance imaging data. In: NIPS mini-symposium on causality in time series, 2011. p. 65–94.

4. Zhu F, Liang J. Soil moisture retrieval from UWB sensor data by leveraging fuzzy logic. IEEE Access, 2018. https://doi.org/10.1109/ACCESS.2018.2840159.
5. Alipanahi B, Delong A, Weirauch MT, et al. Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning. Nat Biotechnol. 2015;33(8):831–9.
6. Hochreiter S, Schmidhuber J. Long short-term memory. Neural comput. 1997;9(8):1735–80.
7. Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process Mag. 2012;29(6):82–97.
8. Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. Adv Neural Inf Process Syst, 2014, p. 3104–12.
9. van der Westhuizen J, Lasenby J. The unreasonable effectiveness of the forget gate. arXiv preprint arXiv:1804.04849, 2018.
10. Tank A, Cover I, Foti NJ, et al. An interpretable and sparse neural network model for nonlinear granger causality discovery. In: Accepted by NIPs time series workshop, 2017.
11. Choi E, Schuetz A, Stewart WF, et al. Using recurrent neural network models for early detection of heart failure onset. J Am Med Inf Assoc. 2016;24(2):361–70.
12. Zhong Z, Jiang T, Zhang W, et al. Analyzing speech of patients with vocal polyps based on channel parameters and fuzzy logic systems. Comput Math Appl. 2011;62(7):2834–42.
13. Mitra SK, Kuo Y. Digital signal processing: a computer-based approach. New York: McGraw-Hill Higher Education; 2006.

# Audio Tagging With Connectionist Temporal Classification Model Using Sequentially Labelled Data

Yuanbo Hou[1(✉)], Qiuqiang Kong[2], and Shengchen Li[1]

[1] Beijing University of Posts and Telecommunications, Beijing, China
hyb@bupt.edu.cn
[2] Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, UK

**Abstract.** Audio tagging aims to predict one or several labels in an audio clip. Many previous works use weakly labelled data (WLD) for audio tagging, where only presence or absence of sound events is known, but the order of sound events is unknown. To use the order information of sound events, we propose sequentially labelled data (SLD), where both the presence or absence and the order information of sound events are known. To utilize SLD in audio tagging, we propose a convolutional recurrent neural network followed by a connectionist temporal classification (CRNN-CTC) objective function to map from an audio clip spectrogram to SLD. Experiments show that CRNN-CTC obtains an area under curve (AUC) score of 0.986 in audio tagging, outperforming the baseline CRNN of 0.908 and 0.815 with max pooling and average pooling, respectively. In addition, we show CRNN-CTC has the ability to predict the order of sound events in an audio clip.

**Keywords:** Audio tagging · Sequentially labelled data (SLD) · Convolutional recurrent neural network (CRNN) · Connectionist temporal classification (CTC)

## 1 Introduction

Audio tagging aims to predict an audio clip with one or several tags. Audio clips are typically short segments such as 10 s of a long recording. Audio tagging has many applications in information retrieval [1], audio classification [2], acoustic scene recognition and industry sound recognition [3].

Many previous works of audio tagging rely on strongly or weakly labelled data. In strongly labelled data [4], each audio clip is labelled with both tags and onset and offset times of sound events. Labelling the strongly labelled data is time-consuming and labour expensive, so the size of strongly labelled dataset is often limited to minutes or a few hours [4]. Additionally, the onset and offset time of some sound events are ambiguous due to the fade-in and fade-out effect [5]. In fact, many audio datasets contain only tags, without the onset and offset times of sound events. This is referred to as weakly labelled data (WLD) [5]. In WLD, only the presence or absence of sound events is known, the occurrence sequence of sound events is not known. These weaknesses limit the use of strongly labelled data and WLD.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_114

To avoid the weakness of strongly labelled data and WLD and use order information of sound events, we propose sequentially labelled data (SLD). This idea is inspired by the label sequences in speech recognition [6]. In SLD, the tags and order of tags are known, without knowing occurrence time of tags. SLD reduces the workload of data annotation and indicates the order of tags in WLD. The order information of events will benefit tasks like acoustic scene analysis and environment recognition. Figure 1 shows an audio clip and its strong, sequential and weak tags.

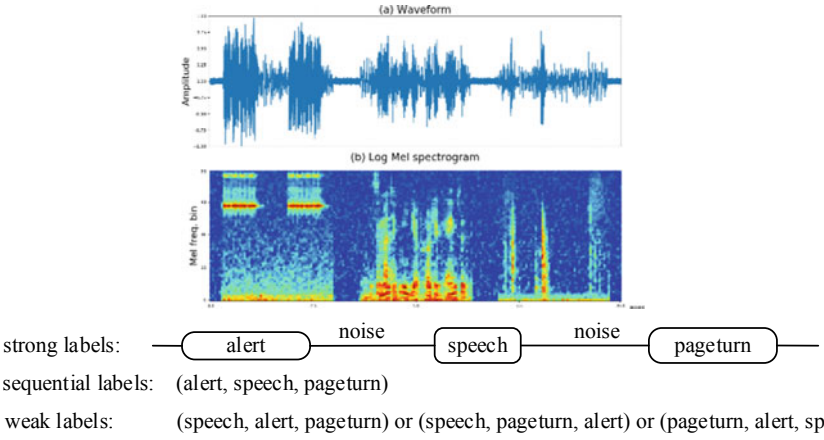



**Fig. 1** From top to bottom: **a** waveform of an audio clip containing three sound events: "alert", "speech" and "pageturn"; **b** log Mel spectrogram of (**a**); strong labels, sequential labels and weak labels of the audio clip

To utilize the SLD in audio tagging, we propose to use CTC technique to train a CRNN (CRNN-CTC). CTC is a learning technique for sequence labelling with RNN [7], which has achieved great success in speech recognition [6]. In fact, CTC is an objective function that allows RNN to be trained for sequence-to-sequence tasks, without requiring any prior alignment between the input and target sequences. In training, CTC computes the total probability of input sequences and sums over all possible alignments. CTC allows train an RNN without any prior alignment (i.e. the starting or ending times of each sound event), hence, even without strong labels, it is sufficient to do audio tagging with SLD based on CTC model.

There are two contributions in this paper. First, in audio tagging, we propose SLD, which reduces the workload and difficulties of data annotation in strong labels, and indicates the order of tags in weak labels. Second, to utilize SLD in audio tagging, we propose CRNN-CTC compare its performance with other common CRNN models in previous works. This paper is organized as follows, Sect. 2 introduces related works. Section 3 describes CRNN baseline. Section 4 describes CRNN-CTC with SLD. Section 5 describes dataset, experimental set-up and results. Section 6 gives conclusions.

## 2 Related Work

Audio classification and detection have obtained increasing attention in recent years. There are many challenges for audio detection and tagging such as DCASE 2013 [3], DCASE 2016 and DCASE 2017 [4].

In previous works in audio classification and tagging, Mel-frequency cepstrum coefficient (MFCC) and Guassian mixture model (GMM) is widely used in baseline system [3]. Recent methods include deep neural networks (DNNs), convolution neural networks (CNNs) and RNN [2], with inputs varying from Mel energy, spectrogram, MFCC to constant $Q$ transform (CQT).

Many methods described above rely on the bag of frames (BOF) model [5]. BOF is based on an assumption that tags occur in all frames, which is however not the case in practice. Some audio events like "gunshot" only happen a short time in audio clip. State-of-the-art audio tagging methods [8] transform waveform to the time-frequency (*T-F*) representation. *T-F* representation is treated as an image which is fed into CNNs. However, unlike image where the objects usually occupy a dominant part of the image, in an audio clip audio events only occur a short time. To solve this problem, some attention models [8] for audio tagging and classification are applied to attend to the audio events and ignore the background sounds.

## 3 CRNN Baseline in Audio Tagging

CRNN has been successfully used in audio tagging [8]. First, the waveforms of the audio are transformed to time-frequency (*T-F*) representation such as Mel spectrogram. Next convolutional layers are applied on the *T-F* representation to extract high-level features. Then, bidirectional gated recurrent units (BGRU) are adopted to capture the temporal context information. Finally, the output layer is a dense layer with the sigmoid activation function since it is a multi-class classification problem [5], the sigmoid activation function to predict probability of each sound event in the audio clip. Inspired by the good performance of CRNN in audio tagging [8], we use CRNN as our baseline system in this paper.

An audio clip from real life may contain more than one sound event, as environmental sound is often a mixture audio that comes from multiple sound sources simultaneously. Thus, the audio tagging task is a multi-label classification problem, and a binary decision is made for each class. In training phase, the binary cross-entropy loss is applied between the predicted probability of each tag and the ground-truth tag in an audio clip. The loss can be defined as:

$$E = -\sum_{n=1}^{N}(P_n \log Q_n + (1 - P_n)\log(1 - Q_n)) \tag{1}$$

where $E$ is the binary cross-entropy, $Q_n$ and $P_n$ denote the predicted and reference tags sequence of the $n$-th audio clip, respectively. $N$ is the batch size.

In CRNN baseline, clip level probability of tags can be obtained from the last layer. However, there is no frame level information of each event in it. To obtain the probability of each event at each frame, a dense layer with the number of event classes, following the BGRU layer, as shown in Fig. 2. These frame level predictions can be used for sound event detection. To map the frame level tags to clip level tags, pooling layer was used. In training, the clip level predictions are compared against the weak labels of the audio clip to compute the loss function of model.

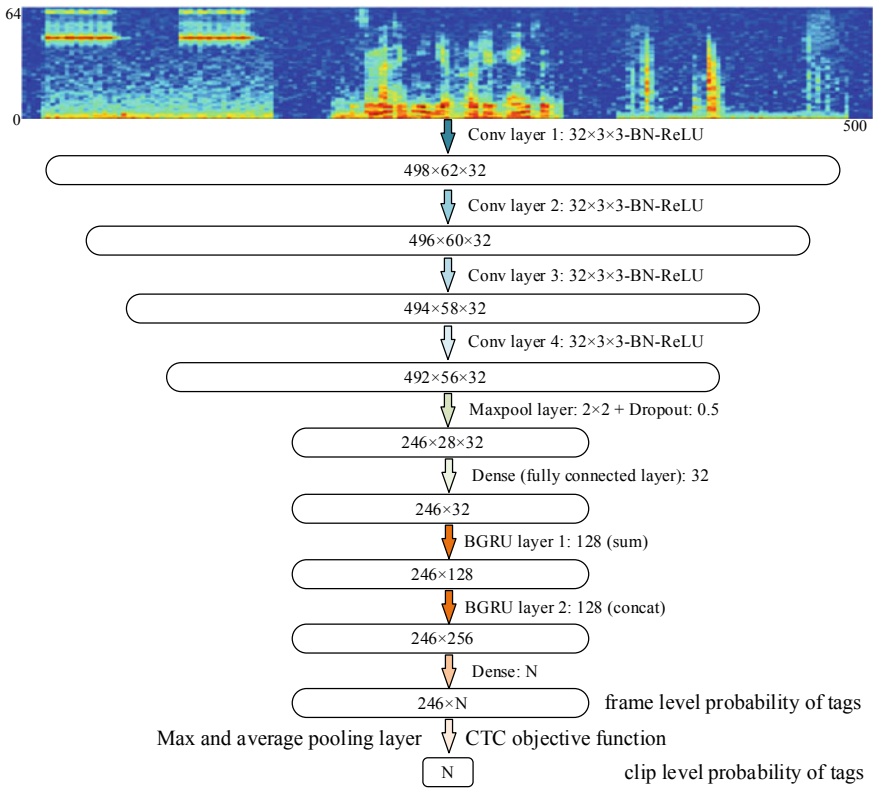



**Fig. 2** Model structure. BN: Batch normalization. ReLU: Rectified linear unit. For baseline, CRMP and CRAP, $N = 16$. For CRNN-CCT, $N = 17$ (16 + 1), the extra "1" indicates the blank label

There are two pooling operations in Fig. 2, max pooling and average pooling. For CRNN with max pooling (CRMP) and CRNN with average pooling (CRAP), pooling performs down-sampling along time axis and transforms the frame level probability of tags to clip level tags, respectively. Max pooling and average pooling as way of aggregation have been successfully used [8].

## 4 CRNN-CTC in Audio Tagging

As discussed before, strongly and weakly labelled data have their own drawbacks in audio tagging, so we propose sequentially labelled data (SLD) and use CRNN-CTC to detect presence or absence of several sound events in SLD.

### 4.1 Sequentially Labelled Data

Let $\mathcal{D}$ be a set of training examples drawn from audio dataset. Input space $\mathcal{X} = (\mathbb{R}^n)$ is the set of all sequences of $n$ dimensional vectors. Target space $\mathcal{Z} = L$ is the set of all sequences of labels over audio events. In general, we refer to elements of $L$ as label sequences or labelling [7]. Each example in $\mathcal{D}$ consists of a pair of sequences $(\boldsymbol{x}, \boldsymbol{z})$. The target sequence $\boldsymbol{z} = (z_1, z_2, \ldots, z_Q)$ is at most as long as input sequence $\boldsymbol{x} = (x_1, x_2, \ldots, x_T)$, i.e. $Q \leq T$. Since, the input and target sequences are not generally the same length, there is no a priori way of aligning them [7]. In the label sequence $\boldsymbol{z}$, the tags of the audio clip and sequence of tags are known, without knowing their occurrence time, that is, there are no starting/ending times of sound events. We refer to audio data labelled by label sequence as sequentially labelled data (SLD).

In essence, SLD is a weakly labelled data with events sequence information. In audio tagging using SLD, we can use the model like CRNN described in Sect. 3. However, there is no order information of sound events in predictions of baseline, CRMP and CRAP. So we propose CRNN-CTC in audio tagging using SLD.

### 4.2 CRNN-CTC in Audio Tagging Using SLD

CTC has achieved great success in speech recognition [6, 7]. In this section, we will show how to use CTC technique to train a CRNN in audio tagging using SLD.

CTC is a learning technique for sequence labelling, it shows a way for training RNN with label unsegment sequences. CTC redefines the loss function of RNN [7] and allows RNN to be trained for sequence-to-sequence tasks without requiring any prior alignment (i.e. starting or ending time of sound events) between the input and target sequences. Thus, it is sufficient to train a CRNN using SLD with CTC. Given $y_t(k)$ is probability of observing label $k$ at time $t$ output by the last recurrent layer in CRNN, and $z_t$ is the ground-truth label, conventional loss function of RNN for a sequence $X$ of length $T$ is $L = -\sum_{t=1}^{T} \log y_t(z_t)$, which is the negative logarithm of the joint probability of desired label sequence and its alignment. In audio tagging, we are only interested in label sequence, not the ground-truth alignment.

CTC gives a solution to how to marginalize out the alignment. First, CTC adds an extra "blank" label (denoted by "–") to original label set $L$ [7]. Then, it defines a many-to-one mapping $\beta$ that transforms the alignment (i.e. the sequence of output labels at each time step, also called a path [7]) to label sequence. The mapping $\beta$ removes repeated labels from the path to a single one and then removes the "blank" labels. For example, $\beta(C - AT-) = \beta(-CC - -ATT) = CAT$, that is, path "$C - AT-$" and "$-CC - -ATT$" both map to the label sequence "$CAT$".

The CTC objective function is defined as the negative logarithm of total probability of all paths that map to the ground-truth label sequence. The total probability can be found using dynamic programming algorithm [7] on the trellis shown in Fig. 3. On the $x$-axis is time steps, on the $y$-axis is "modified label sequence", target label sequence with blank labels added to the beginning and end and inserted between every pair of labels. Given the length of modified label sequence is $L$ and $l_i$ denotes $i$-th label. An effective path may start at either $l_1$ or $l_2$ and end at $l_{L-1}$ or $l_L$. At each time step, the path may (i) stay at the same label; (ii) move to the next label; (iii) move to the label after the next if it is not a blank label different from the current label. Let $\alpha_t(s)$ be the total probability of $l_{1:s}$ at time $t$. Assuming conditional independence between $y_t(k)$ (i.e. probability of observing label $k$ at time $t$) across time steps, the $\alpha_t(s)$ can be calculated as follows:

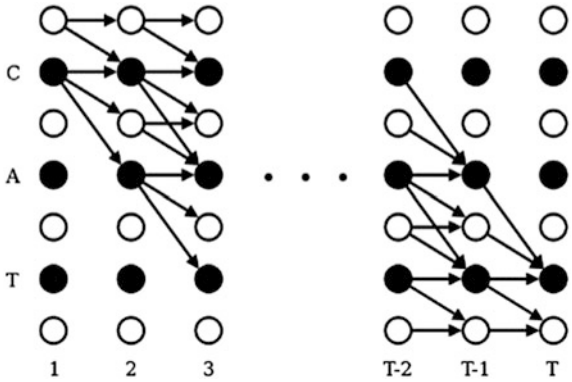



**Fig. 3** Trellis for computing CTC objective function [7] applied to the example labelling "CAT". Black circles represent labels, white circles represent blanks. Arrows signify allowed transitions

$$\alpha_1(s) = \begin{cases} y_1(l_s) & s \leq 2 \\ 0 & s > 2 \end{cases} \tag{2}$$

$$\alpha_t(s) = [\alpha_{t-1}(s) + \alpha_{t-1}(s-1) + \delta_s \alpha_{t-1}(s-2)] y_t(l_s), t > 1 \tag{3}$$

where $\delta_s = 1$ if $l_s \neq l_{s-2}$, and terms that go past the start of the modified label sequence are zero. The sum of total probability of paths that map to original label sequence is $\alpha_T(L-1) + \alpha_T(L)$, and its negative logarithm is CTC loss function.

To decode the CTC output, we use the simple best path decoding [7] in this paper. This method is to select the label with the maximum probability at each frame, reduce adjacent repeating labels to a single one and remove the blank labels.

The output of CTC model is directly a label sequence corresponding the audio clip. The detailed structure of CRNN-CTC was shown before in Fig. 2.

## 5 Experiments and Results

We use the audio events in DCASE 2013 [3] to make SLD and evaluate the proposed method. There are 16 kinds of sound events in DCASE 2013 includes: *alert, clear-throat, cough, doorslam, drawer, keyboard, keys, knock, laughter, mouse, pageturn, pendrop, phone, printer, speech* and *switch*. We remixed these sound events to 10 s audio clips totalling 7.1 h, where each audio clip contains no overlapped three or several sound events mixed with noise background.

For experimental set-up, fourfold cross-validation was used for model selection and parameter tuning. Dropout, batch normalization and early stopping criteria are used in training phase to prevent over-fitting. The model is trained for maximum 1000 epochs with Adam optimizer with learning rate of 0.001.

The results are evaluated by precision, recall and $F$-score [9]. To calculate these metrics, we need to count the number of: true positive (TP), false negative (FN) and false positive (FP). Precision ($P$), recall ($R$) and $F$-score are defined as:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \quad R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \quad F = \frac{2\mathrm{P} \cdot \mathrm{R}}{\mathrm{P} + \mathrm{R}}. \tag{4}$$

To evaluate the true positive rate (TPR) versus false positive rate (FPR), the receiver operating characteristic (ROC) curve was used [9]. AUC score is the area under the ROC curve which summarizes the ROC curve to a single number. Larger $P$, $R$, $F$-score and AUC indicates better performance. The AUC score of audio tagging is shown in Table 1; CRAP, CRMP and CRNN-CTC outperform baseline system. CRNN-CTC achieves an averaged AUC of 0.986.

Table 2 shows the averaged statistic including precision, recall, $F$-score and AUC over 16 kinds of sound events, and CRNN-CTC performs better than other models. Figure 4 shows the frame level predictions of models on example audio clip. In Fig. 4, CRNN-CTC predicts the tag sequence of audio clip, typically as a series of spikes [7]. Although the spikes align well with the actual position of sound events in audio clip, there is no time span information about these events.

In Fig. 4, CRMP produces wide peaks, indicating the onset/offset times of each event. That shows max pooling has ability to locate audio events, while average pooling seems to fail. The reason may be max pooling encourages the response for a single location to be high, for similar audio events which can obtain similar features. While average pooling in CRAP encourages all response to be high, the difference features of each event make it difficult to locate audio events.

**Table 1** AUC of audio tagging

| | *alert* | *clearthroat* | *cough* | *doorslam* | *drawer* | *keyboard* | *keys* | *knock* | *laughter* | *mouse* | *pageturn* | *pendrop* | *phone* | *printer* | *speech* | *switch* | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Baseline | 0.609 | 0.627 | 0.674 | 0.691 | 0.690 | 0.569 | 0.702 | 0.816 | 0.617 | 0.668 | 0.693 | 0.662 | 0.654 | 0.862 | 0.550 | 0.625 | 0.669 |
| CRAP | 0.737 | 0.948 | 0.792 | 0.804 | 0.895 | 0.811 | 0.864 | 0.971 | 0.783 | 0.587 | 0.759 | 0.809 | 0.715 | 0.910 | 0.800 | 0.850 | 0.815 |
| CRMP | 0.959 | 0.970 | 0.915 | 0.875 | 0.953 | 0.735 | 0.918 | 0.973 | 0.883 | 0.835 | 0.892 | 0.936 | 0.892 | 0.985 | 0.887 | 0.922 | 0.908 |
| CRNN-CTC | 0.968 | 1.0 | 1.0 | 0.977 | 1.0 | 0.959 | 0.972 | 1.0 | 1.0 | 0.995 | 0.990 | 0.972 | 1.0 | 0.995 | 0.990 | 0.965 | 0.986 |

**Table 2** Averaged stats of audio tagging

| | Precision | Recall | *F*-score | AUC |
|---|---|---|---|---|
| Baseline | 0.687 | 0.371 | 0.482 | 0.669 |
| CRAP | 0.847 | 0.647 | 0.733 | 0.815 |
| CRMP | 0.933 | 0.827 | 0.877 | 0.908 |
| CRNN-CTC | 0.983 | 0.975 | 0.98 | 0.986 |

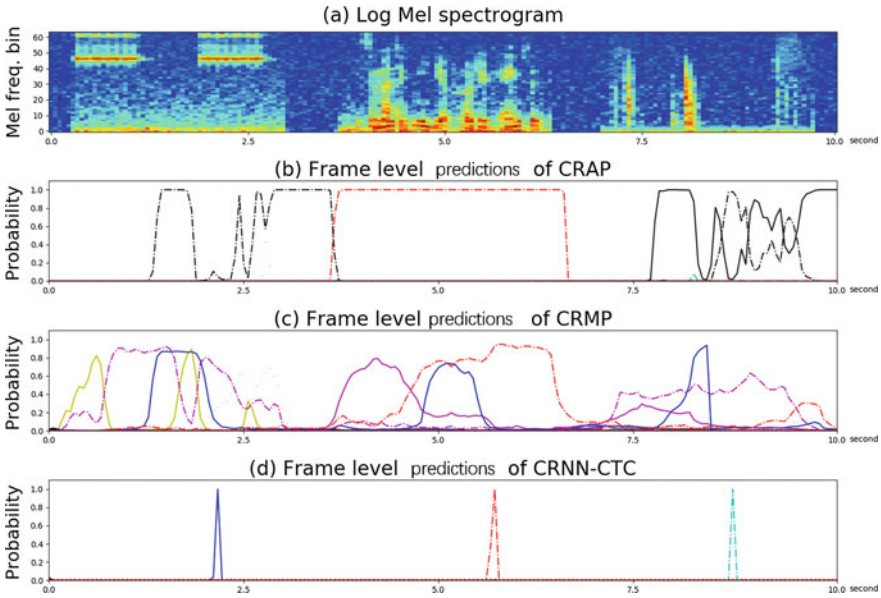



**Fig. 4.** A frame level predictions of CRAP (**b**), CRMP (**c**) and CRNN-CTC (**d**). The ground-truth tag is "alert, speech, pageturn". Peaks are annotated with corresponding tag

## 6 Conclusion

In this paper, we analyse the weakness of strongly and weakly labelled data, then propose SLD. To utilize SLD in audio tagging, we propose CRNN-CTC. In CRNN-CTC, CTC layer maps frame level tags to clip level tags, similar to the pooling layer. So we compare them. Experiments show CRNN-CTC outperforms CRAP, CRMP and baseline. The frame level predictions of models in Fig. 4 show CRNN-CTC predicts the presence/absence and tag sequence of events in the audio clip well.

## References

1. Guo, G, and S. Z. Li. Content-based audio classification and retrieval by support vector machines. IEEE Press, 2003.
2. Xu Y, Kong Q, Wang W and Plumbley MD. "Largescale weakly supervised audio classification using gated convolutional neural network," arXiv preprint arXiv:1710.00343, 2017.
3. Stowell D, Giannoulis D, Benetos E, Lagrange M, and Plumbley MD. "Detection and classification of acoustic scenes and events," IEEE Transactions on Multimedia. 17(10): 1733–1746, 2015.
4. Mesaros A, Heittola T, et al. Dcase 2017 challenge setup: Tasks, datasets and baseline system, in Workshop on DCASE 2017, Munich, Germany, 2017.
5. Kong Q, Xu Y, Wang W and Plumbley MD. A joint separation-classification model for sound event detection of weakly labelled data, in IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Apr. 2018.
6. Graves A and Jaitly N. Towards end-to-end speech recognition with recurrent neural networks, in Proc. of ICML, 2014.
7. Graves A and Gomez F, Connectionist temporal classification:labelling unsegmented sequence data with recurrent neural networks, in ICML, 2006, pp. 369–376.
8. Xu Y, Kong Q, Huang Q, Wang W and Plumbley MD. Attention and localization based on a deep convolutional recurrent model for weakly supervised audio tagging, in INTERSPEECH, 207, pp. 3083–3087.
9. Bhavna K, Jain K and Sharma SK. Estimation of Area under Receiver Operating Characteristic Curve for Bi-Pareto and Bi-Two Parameter Exponential Models. Open Journal of Statistics 4.1(2014):1–10.

# Implementation of AdaBoost Face Detection Using Vivado HLS

Sanshuai Liu, Kejun Tan(✉), and Bo Yang

Information Science and Technology College,
Dalian Maritime University, Dalian, Liaoning Province, China
tankejun@dlmu.edu.cn

**Abstract.** For the problem that Adaptive Boosting (AdaBoost) face detection algorithm is slowly implemented on the embedded platform by software, this paper adopts the method of the full hardware acceleration. The intellectual property (IP) core of AdaBoost algorithm is designed by Vivado high-level synthesis (HLS), which may reduce the development difficulty and shorten the development cycle. The design adopts the serial–parallel structure to accelerate face detection and uses several methods of optimizing hardware resource. The face detection algorithm is implemented on the Zedboard platform and achieves the purpose of real-time detection.

**Keywords:** AdaBoost · Face detection · HLS · Real-time detection

## 1 Introduction

Face detection refers to the process of determining whether there are faces in the image, and their positions, and sizes. The face detection problem is originally proposed as the positioning link of the automatic face recognition system. Due to its application value in the areas of security access control and visual inspection, it has become an independent topic that has received widespread attention [1].

As face detection becoming more and more important in many fields, researchers have tried various methods to achieve face detection. Software implementation implemented on the embedded platform is slow [2]. Therefore, some scholars have studied the hardware acceleration of AdaBoost algorithm; for example, Shi Yuehua used Verilog HDL to implement the face detection algorithm [3]. In this paper, AdaBoost algorithm intellectual property (IP) core is designed by HLS and implemented on the Zedboard platform. HLS could convert C, C++, and other codes into hardware description language [4], and the generated IP cores can be directly connected with register-transfer-level (RTL) design in Vivado. Compared to software implementations, this method has a faster detection speed. Compared to the hardware description language implementations, the algorithm is easier to implement and the development cycle is shorter.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_115

## 2 Face Detection Algorithm

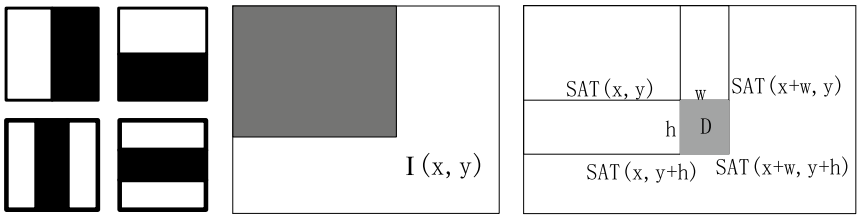
### 2.1 Haar Features and Integral Image

The Haar features reflect the gray pixel subtraction module in an image and are mainly used to describe human face feature, and the Haar features used in this paper are shown in Fig. 1a. For facilitating the calculation of Haar feature, Viola and Jones introduced the concept of integral image [5]. The definition of the integral image is:

$$SAT(x, y) = \sum_{x' < x, y' < y} I(x', y') \tag{1}$$

The value of $I(x', y')$ represents the pixel value of the point $(x', y')$, and $SAT(x, y)$ is the pixel sum of all pixels in the upper-left corner of the point $(x, y)$, as shown in Fig. 1b. Figure 1c shows the incremental calculation algorithm which is used to quickly calculate the Haar features. The sum of pixels in the rectangular shaded region could be obtained by adding and subtracting the values of the integral image of the four vertex positions. The calculation formula is as follow:

$$Sum(D) = SAT(x, y) + SAT(x+w, y+h) - SAT(x+w, h) - SAT(x, y+h) \tag{2}$$



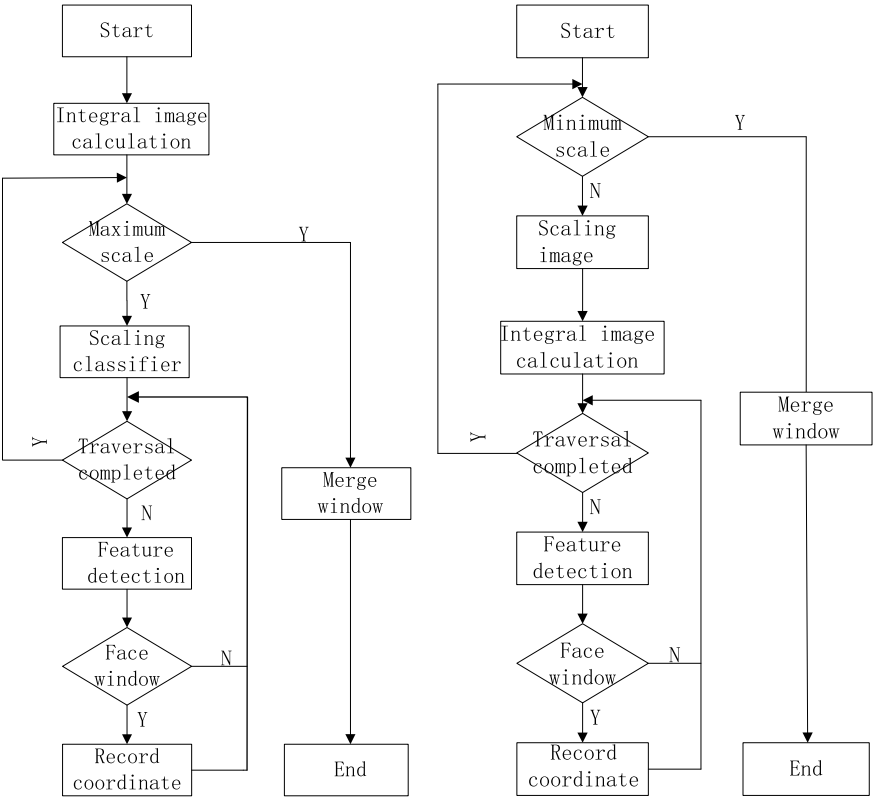

(a) Haar features. (b) Integral image. (c) Integral image increment calculation.

**Fig. 1.** Haar features and integral image.

### 2.2 Face Detection Process

Scaling the image and scaling the classifiers are two methods of face detection [6]. The method of scaling the classifiers only needs to calculate the integral image once, and it needs to cache the integral image. It takes less time, but it consumes a large amount of memory resources. The detection process is shown in Fig. 2a. The method of scaling the image slides a fixed size detection window in the

integral image, while the size of detection window in the method of scaling the classifiers is changed. The memory resources consumption of scaling the image is less, because it requires to cache the raw image, but each time the image is scaled the integral image needs to calculate again. Therefore, scaling the image method is used in this paper. The specific detection process is shown in Fig. 2b.




(a) The processing flow of scaling the classifiers.

(b) The processing flow of scaling the image.

**Fig. 2.** Processing flow of AdaBoost.

The process of scaling the image face detection firstly judges whether the image is scaled to the minimum scale that is the size of training sample and, if not, continues to scale the image. The square integral image and the integral image are used to obtain the image variance which could correct the illumination of the image. The detection window traverses in the integral image to find face window. If a face window is detected, the coordinate and the scaling ratio of face window are recorded. When the integral image traversal is over, it is judged whether the image is scaled to the smallest scale that means the end of the detection. The face windows are merged to obtain the correct coordinate position after the detection is completed.

# 3 Hardware Design and Optimization

## 3.1 Face Detection Module

Feature detection module is the core module in the face detection. The classifier used in the module is the front face detection classifier provided by the Open Source Computer Vision (OpenCV) library. The Haar features in the classifier we used have only two rectangular features and three rectangular features, and we encoded them. It can be determined what kind of feature is used by identifying the encoding, which could save memory resources compared to all default to three rectangular features (the third rectangular data in two rectangular features is filled with zeros). Firstly, the module of feature detection reads the stored rectangular data of the Haar features of the classifier, including the rectangular starting coordinates, widths, lengths, and weights. The sum of pixels in the rectangular region can be calculated according to formula (2). The sum of pixels in each rectangular area is multiplied by itself related weight and then added. The sum is compared with the product of the corresponding weak classifier threshold and the variance within the window. If it is greater than the product, it takes the right value, and if it is less than the product, it takes the left value. The results of the weak classifiers accumulated compared with the threshold of the strong classifier; if it is greater than the threshold, it means that the strong classifier is passed and the next strong classifier will be detected; otherwise, it is not a face window, and the next candidate window will be detected.

## 3.2 Serial and Parallel Structure Design

The classifier has 22 strong classifiers, each of which could filter out a certain amount of candidate windows. The filtered results of top 5 strong classifiers are tested in 50 different face pictures as shown in Fig. 3. It could be seen that the filtered increase of the top 3 strong classifier is larger, and the filtered increase in the fourth stage and the fifth stage gradually decreases. Since the top 3 strong

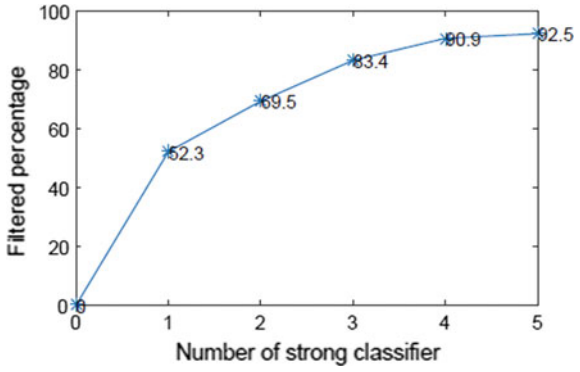



**Fig. 3.** Filtered rate of strong classifier.

classifiers could filter out 83.4% of non-face windows, taking into account the relationship between resources and speed, the top 3 strong classifiers use a parallel structure to design, and the latter 19 strong classifiers use a serial structure to design, as shown in Fig. 4. The weak classifiers of the top 3 strong classifiers are all expanded, and a parallel structure could be obtained. If any one of top 3 strong classifiers detects failed, the next window is directly detected. If the detection of the parallel module is successful, the detection of the serial module will be continued. In the serial module, only the strong classifier at current could continue to detect the next strong classifier. Otherwise, it will detect the next window. If the candidate window passes all strong classifiers, the window is a face window, the coordinate of which will be recorded, and the next window continues to be detected.

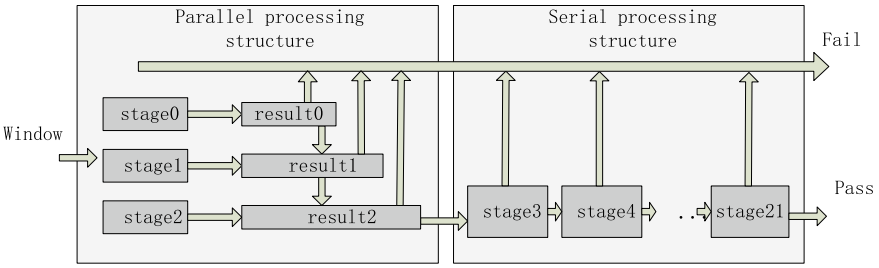



**Fig. 4.** Serial and parallel structure.

### 3.3 Data Bit Qualification and Type Conversion

The calculation of the integral image could be easily obtained according to the incremental calculation method of the integral image, but the integral image needs to cache when the window traverses. Limiting the bit width of data can save block random access memory (BRAM) resource consumption. For example, the maximum of the $320 \times 240$ image's integral image is $320 \times 240 \times 255 = 19,584,000$, so at least 25 bits are needed to store the data of the integral image. The maximum of the squared integral image is the maximum of the integral image moved left 8 bits, so storage requires at least 33 bits.

Floating-point calculation is used in the bilinear interpolation scaling image and the scaling calculation. Fixed-point data is defined in Vivado HLS, which could set the length of integer part and decimal part according to the required accuracy. As shown in Table 1, solution 1 is the resource consumption of the fixed-point data, the bit width of the integer part is 10, the bit width of the decimal part is 16, and solution 2 is the resource consumption report of the floating-point-type data. Solution 1 is compared with solution 2, the look-up table (LUT) resource is reduced about 15%, and other hardware resources are also reduced. Since the floating-point-type data consumes lots of resources, the

thresholds, the left-value parameters, the right-value parameters of the weak classifiers, and the thresholds of the strong classifiers are expanded 4096 times to become integer data.

**Table 1.** Comparison of resource consumption

| | Solution 1 | Solution 2 |
|---|---|---|
| BRAM_18K | 197 | 197 |
| DSP48E | 54 | 62 |
| FF | 44424 | 48715 |
| LUT | 30978 | 39069 |

### 3.4 Multiplication Optimization

Multiplication calculation takes much time and occupancies digital signal processing (DSP) resources. Multiplication calculation can be split into shift operations which take less time, so the multiplication calculation in the code could be replaced with shift operations. It not only saves time, but also reduces the use of DSP resources.

## 4 Experiment Result

The experiment is implemented on the Zedboard platform, which includes a dual-core ARM Cortex A9 and more than 5 million programmable logic elements. In the experiment, the image size is $320 \times 240$, and the scaling is 1.2. The method of scaling the image is bilinear interpolation, and the system operating frequency is 100 MHz.

The experiment selected 100 images which contain face and no face in different backgrounds. The rate of false detection is connected to the background, we may reduce the false detection rate by adjusting the algorithm of window selection, and the detection rate will decrease. Comprehensively, adjusting the window screening algorithm that when two or more scales are continuously passed detection, the candidate face windows are determined to be face windows, and the accuracy rate of the experimental detection is 95%. Software implementation of the face detection algorithm is on the Zedboard platform processing system (PS), and it takes about 3377 ms in the case of having face in the image and 3265 ms in the case of having no face. The hardware implementation is 93 and 82 ms, respectively, and the hardware acceleration effect is more than 30 times, achieving the purpose of real-time detection. In some specific applications, the distance between the person and the sensor is often relatively fixed, such as driver fatigue detection. We may fix scale in several scaling sizes, which could greatly reduce the traversal window, so the speed can reach about 10 ms.

## References

1. Liang L, Ai H, Zhang B. A survey of human face detection. Chin J Comput. 2002;25(5):449–58.
2. Zhu M, Lu X, Lu H. Transplantation and optimization of AdaBoost face detection algorithm on DSP. Comput Eng Appl. 2014;50(20):197–201.
3. Zhang Z, Feng Z, Wei S. Implementation of face detection system based on AdaBoost algorithm on embedded system. Inf Technol. 2008;07:167–70.
4. Xilinx Inc. High-level-synthesis v17.1 [EB/OL]. URL https://www.xilinx.com/support/documentation/sw_manuals/xilinx2018_1/ug902-vivado-high-level-synthesis.pdf.
5. Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: Computer society conference on computer vision and pattern recognitionl, vol. 1; 2001. p. I-511–8.
6. Acasandrei L, Barriga A. Accelerating Viola-Jones face detection for embedded and SoC environments. In: ACM/IEEE international conference on distributed smart cameras, Ghent; 2011. p. 1–6.

# Research on Rolling Bearing On-Line Fault Diagnosis Based on Multi-dimensional Feature Extraction

Tianwen Zhang(✉)

College of Information and Communication Engineering,
Harbin Engineering University, Harbin, China
{287133770,1325069385}@qq.com

**Abstract.** In the paper, a novel rolling bearing fault diagnostic method was proposed to fulfill the requirements for effective assessment of different fault types and severities with real-time computational performance. Firstly, multi-dimensional feature extraction is discussed. And secondly, a gray relation algorithm was used to acquire basic belief assignments. Finally, the basic belief assignments were fused through Yager algorithm. The related experimental study has illustrated the proposed method can effectively and efficiently recognize various fault types and severities.

**Keywords:** Rolling element bearing · Pattern recognition · Gray relation algorithm · Yager algorithm

## 1 Introduction

Rolling bearing as an important part is widely used in almost all types of rotating machinery, such as gas turbine, steam turbine and diesel engine. Rolling bearing failure is one of the main causes of failure and damage of rotating machinery and leads to huge economic losses [1–3]. Among the many fault diagnosis approaches for bearings, vibration-based diagnostic methods have received much attention in the past few decades [4, 5]. Bearing vibration signal contains a wealth of information on mechanical health status. This also makes it possible to extract the dominant features that characterize the mechanical health state from vibration signals through signal processing techniques [6]. Currently, many signal processing techniques have been applied to bearing off-line fault diagnosis. However, due to many nonlinear factors (e.g., stiffness, friction, clearance, etc.), bearing vibration signals (especially in a faulted condition) will exhibit nonlinear and unsteady character [7]. In addition, the measured vibration signal contains not only information about the operating conditions associated with the bearing itself, but also information about a large number of other rotating components and structures in the plant equipment [8]. Due to the usually large background noise, slight bearing fault information is easily submerged in background noise and difficult to extract. Therefore, conventional time-domain and frequency-domain methods do not


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_116

allow for an accurate assessment of the bearing health status [9]. In this paper, a novel rolling bearing fault diagnostic method was proposed to fulfill the requirements for effective assessment of different fault types and severities with real-time computational performance.

## 2 Methodology

### 2.1 Multi-dimensional Feature Extraction

Firstly, multi-dimensional feature extraction on the basis of Entropy characteristics, Holder coefficient characteristics were proposed for extracting health status feature vectors from the bearing vibration signals, respectively.

#### 2.1.1 Entropy Characteristics

Entropy is a crucial concept in information theory and is a measure of information uncertainty of the signal distribution and the complexity of the signal [10].

Suppose the bearing vibration signal is $f$. The signal is first sampled and discretized into a discrete signal sequence $f(i), i = 1, 2, 3, \ldots, n$, where $n$ is the total number of discrete signals. Perform FFT transform as follows:

$$F(k) = \sum_{i=0}^{n-1} f(i) \exp\left(-j\frac{2\pi}{n}ik\right), \quad k = 0, 1, \ldots, n-1 \tag{1}$$

After obtaining the signal spectrum, calculate the energy of each point:

$$E_k = |F(k)|^2 \tag{2}$$

Calculate the total energy value of each point:

$$E = \sum_{k=0}^{n-1} E_k \tag{3}$$

Shannon entropy $E_1$ and exponential entropy $E_2$ can be defined as follows:

$$E_1 = -\sum_{i=1}^{n} P_i \log_e P_i \tag{4}$$

$$E_2 = \sum_{i=1}^{n} P_i \mathrm{e}^{1-P_i} \tag{5}$$

The entropy characteristics $[E_1, E_2]$ are taken as dominant feature vectors for rolling element bearing fault pattern recognition.

### 2.1.2 Holder Coefficient Characteristics

The Holder coefficient algorithm evolves from the Holder inequality [11, 12]. Holder coefficient can be used to measure the similar degree of two sequences.

Holder coefficient of these two discrete signals is obtained as follows:

$$H_c = \frac{\sum_{i=1}^{n} f_1(i) f_2(i)}{\left(\sum_{i=1}^{n} f_1^p(i)\right)^{1/p} \cdot \left(\sum_{i=1}^{n} f_2^q(i)\right)^{1/q}} \tag{6}$$

where $0 \leq H_c \leq 1$.

Calculate the Holder coefficient value $H_1$ between the bearing vibration signal sequence $f(i)$ and the rectangular signal sequence $s_1(i)$.

$$H_1 = \frac{\sum_{i=1}^{n} f(i) s_1(i)}{\left(\sum_{i=1}^{n} f^p(i)\right)^{1/p} \cdot \left(\sum_{i=1}^{n} s_1^q(i)\right)^{1/q}} \tag{7}$$

In the same way, we obtain the Holder coefficient value $H_2$ between the vibration signal sequence $f(i)$ and the triangular signal sequence $s_2(i)$.

$$H_2 = \frac{\sum_{i=1}^{n} f(i) s_2(i)}{\left(\sum_{i=1}^{n} f^p(i)\right)^{1/p} \cdot \left(\sum_{i=1}^{n} s_2^q(i)\right)^{1/q}} \tag{8}$$

The Holder coefficient characteristics $[H_1, H_2]$ are taken as dominant feature vectors for rolling element bearing fault pattern recognition.

## 2.2 Diagnostic Procedure

Totally, the process of the proposed method for rolling bearing on-line fault diagnosis is as follows.

Step 1: The vibration signals from the object bearing are sampled under different healthy status, including normal operating condition and conditions with different fault types and severities, to establish the knowledge base (i.e., the recognition template).

Step 2: The health status feature vectors are extracted from the sample knowledge base through the multi-dimensional feature extraction based on entropy characteristics $[E_1, E_2]$, Holder coefficient characteristics $[H_1, H_2]$, respectively.

Step 3: The sample knowledge base is established based on the fault symptom (i.e., the extracted fault feature vectors) and the fault pattern.

Step 4: The health status feature vectors extracted based on bearing vibration signals to be identified are input into the gray relation algorithm (GRA) to acquire basic belief assignments (i.e., BBA1, BBA2, BBA3) and fuse the basic belief assignments through Yager method and output the diagnostic results (i.e., fault types and severities).

## 3 Experimental Validation

In the paper, the rolling bearing vibration signals for testing are from Case Western Reserve University Bearing Data Center [13]. The fault types contain outer race fault, the inner race fault, and the ball fault, and the fault diameters, i.e., fault severities, contain 28 mils, 21 mils, 14 mils and 7 mils. An accelerometer is installed on the motor drive end housing with a bandwidth of up to 5000 Hz, and the vibration data for the test bearing under different fault patterns is collected by a recorder. The bearing vibration data used for analysis is obtained under the motor speed of 1797 r/min and load of 0 hp (Table 1).

**Table 1** Description of experimental data set

| Health status condition | Fault diameter (mils) | The number of base samples | The number of testing samples | Label of classification |
|---|---|---|---|---|
| Normal | 0 | 10 | 40 | 1 |
| Inner race fault | 7 | 10 | 40 | 2 |
| | 14 | 10 | 40 | 3 |
| | 21 | 10 | 40 | 4 |
| | 28 | 10 | 40 | 5 |
| Ball fault | 7 | 10 | 40 | 6 |
| | 14 | 10 | 40 | 7 |
| | 28 | 10 | 40 | 8 |
| Outer race fault | 7 | 10 | 40 | 9 |
| | 14 | 10 | 40 | 10 |
| | 21 | 10 | 40 | 11 |

Where the abscissa axis $E_1$ represents the Shannon entropy, and ordinate axis $E_2$ represents the exponential entropy.

Where the abscissa axis $H_1$ represents the Holder coefficient with the rectangular sequence as the reference sequence, and ordinate axis $H_2$ represents the Holder coefficient with the triangular sequence as the reference sequence.

From Figs. 1 and 4, it is interesting to see that the dominant fault feature vectors extracted from the bearing vibration signals with different fault types and severities through the multi-dimensional feature extraction based on entropy characteristics, Holder coefficient characteristics show apparent differences, respectively. The knowledge base (i.e., the recognition template) for GRA is established based on the fault symptom (i.e., the extracted feature vectors) and the fault pattern (i.e., the known fault types and severities). The fault feature vectors extracted based on the testing rolling bearing vibration signals to be recognized input to GRA, and the diagnostic results (i.e., fault types and severities) are output after the fusion of BBAs, shown in Table 2 (Figs. 2 and 3).

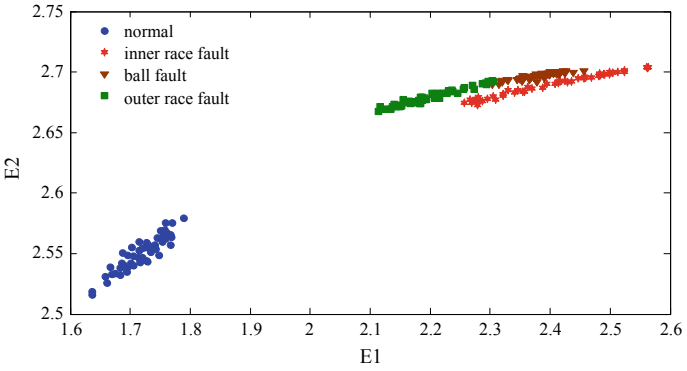



**Fig. 1** Entropy characteristics of a random selected sample from normal operating condition and various fault conditions with fault diameter 7 mils

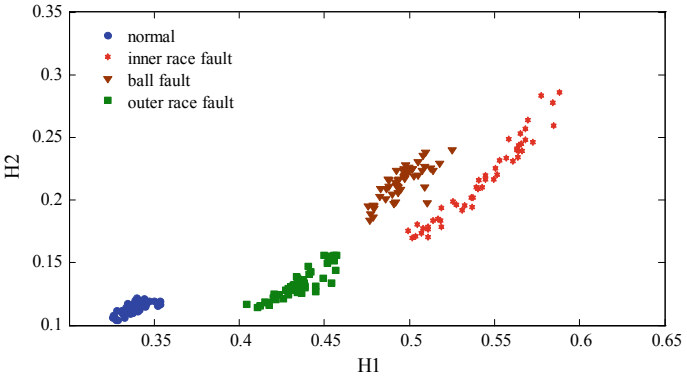



**Fig. 2** Holder coefficient characteristics of a random selected sample from normal operating condition and various fault conditions with fault diameter 7 mils

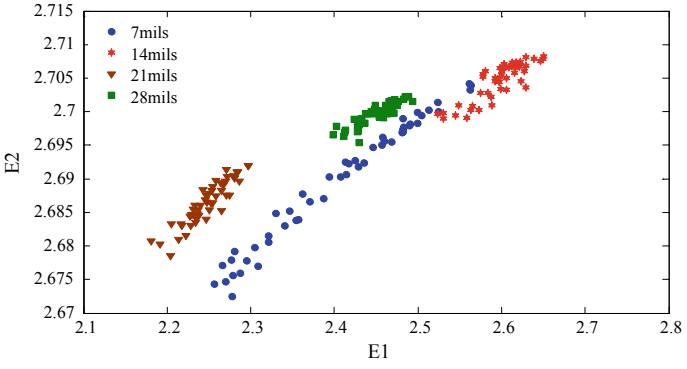



**Fig. 3** Entropy characteristics of a random selected sample from inner race fault condition with various severities

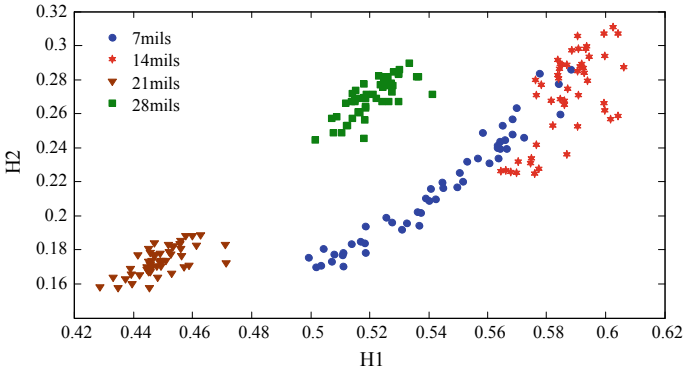



**Fig. 4** Holder coefficient characteristics of a random selected sample from inner race fault condition with various severities

The diagnostic results from Table 2 show that the detecting success rate for bearing faulty conditions can reach 100%, with the total fault pattern recognition success rate almost 99.09%, which shows a certain improvement in diagnostic accuracy compared with the methods from references [14]. The time cost of the proposed method through a laptop computer with a 4.0 GHz dual processor for one test case is only 0.016 s.

**Table 2** Diagnostic results by the proposed method compared with results from Ref. [14]

| Label of classification | The number of testing samples | The number of misclassified samples | | Testing accuracy (%) | |
|---|---|---|---|---|---|
| | | [14] | Proposed | [14] | Proposed |
| 1 | 40 | 0 | 0 | 100 | 100 |
| 2 | 40 | 0 | 0 | 100 | 100 |
| 3 | 40 | 0 | 2 | 100 | 95 |
| 4 | 40 | 3 | 0 | 92.5 | 100 |
| 5 | 40 | 0 | 0 | 100 | 100 |
| 6 | 40 | 2 | 0 | 95 | 100 |
| 7 | 40 | 3 | 2 | 92.5 | 95 |
| 8 | 40 | 3 | 0 | 92.5 | 100 |
| 9 | 40 | 0 | 0 | 100 | 100 |
| 10 | 40 | 0 | 0 | 100 | 100 |
| 11 | 40 | 4 | 0 | 90 | 100 |
| In total | 440 | 15 | 4 | 96.59 | 99.09 |

## 4 Conclusion

In the paper, a novel rolling bearing fault diagnostic method was proposed to fulfill the requirements for accurate assessment of different fault types and severities with real-time computational performance. The related experimental study has illustrated the following conclusions: The diagnostic results by the proposed approach show that the diagnostic success rate for bearing faulty conditions can reach 100%, with the total diagnostic success rate almost 99.09%; the proposed approach can improve the fault diagnostic accuracy compared with the existing expert systems, and it is very suitable for on-line bearing fault diagnosis.

**Acknowledgments.** This work is supported by the National Natural Science Foundation of China (61771154) and funding of State Key Laboratory of CEMEE (CEMEE2018K0104A).

Meantime, all the authors declare that there is no conflict of interests regarding the publication of this article.

We gratefully thank of very useful discussions of reviewers.

## References

1. Hecke BV, Qu Y, He D. Bearing fault diagnosis based on a new acoustic emission sensor technique. J Risk Reliab. 2015;229(2):105–18.
2. Jiang L, Shi T, Xuan J. Fault diagnosis of rolling bearings based on marginal fisher analysis. J Vib Control. 2014;20(3):470–80.
3. Xu J, Tong S, Cong F, Zhang Y. The application of time-frequency reconstruction and correlation matching for rolling bearing fault diagnosis. ARCHIVE Proc Inst Mech Eng Part C J Mech Eng Sci 1989–1996, vols. 203–210. 2015;229(17):3291.
4. Lin Y, Wang C, Ma C, Dou Z, Ma X. A new combination method for multisensor conflict information. J Supercomputing. 2016;72(7):2874–90.
5. Zhang X, Hu N, Hu L, Chen L, Cheng Z. A bearing fault diagnosis method based on the low-dimensional compressed vibration signal. Adv Mech Eng. (2015);7(7).
6. Zhang DD. Bearing fault diagnosis based on the dimension–temporal information. ARCHIVE Proc Inst Mech Eng Part J. J Eng Tribol 1994–1996, vols. 208–210. 2011;225 (8):806–13.
7. Vakharia V, Gupta VK, Kankar PK. A multiscale permutation entropy based approach to select wavelet for fault diagnosis of ball bearings. J Vib Control. 2014;21(16):3123.
8. Tiwari R, Gupta VK, Kankar PK. Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier. J Vib Control. 2015;21(3):461–7.
9. Sun W, Yang GA, Chen Q, Palazoglu A, Feng K. Fault diagnosis of rolling bearing based on wavelet transform and envelope spectrum correlation. J Vib Control. 2013; 19(6):924–41.
10. Li J, Guo J (2015) A new feature extraction algorithm based on entropy cloud characteristics of communication signals. In: Mathematical problems in engineering. p. 1–8.
11. Li J. A new robust signal recognition approach based on holder cloud features under varying SNR environment. KSII Trans Internet Inf Syst. 2015;9(12):4934–49.

12. Li J. A novel recognition algorithm based on holder coefficient theory and interval gray relation classifier. KSII Trans Internet Inf Syst. 2015;9(11):4573–84.
13. The Case Western Reserve University Bearing Data Center, http://csegroups.case.edu/bearingdatacenter/pages/download-data-file. Accessed 11 Oct 2015.
14. Li J, Cao Y, Ying Y, Li S. A rolling element bearing fault diagnosis approach based on multifractal theory and gray relation theory. PLoS ONE. 2016;11(12):1–16.

# Multi-pose Face Recognition Based on Contour Symmetric Constraint-Generative Adversarial Network

Ning Ouyang[1,2], Liyuan Liu[2], and Leping Lin[1,2](✉)

[1] Key Laboratory of Cognitive Radio and Information Processing, Ministry of Education, Guilin University of Electronic Technology, Guilin 541004, Guangxi, China
ouyangning@guet.edu.cn, lin_leping@163.com

[2] School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, Guangxi, China
342520927@qq.com

**Abstract.** In order to address the impact of large-angle posture changes on face recognition performance, we propose a contour symmetric constraint-generative adversarial network (CSC-GAN) for the multi-pose face recognition. The method employs the convolutional network as the generator for face pose recovery, which introduces the global information of the constrained pose recovery of positive face contour histogram. Meanwhile, the original positive face is used as the discriminator, and the symmetric loss function is added to optimize the learning ability of the network. The positive face with gesture recovery is obtained by striking the balance between training of the generator and discriminator. Then we employed the nearest neighbor classifier to identify. The experimental results show that CSC-GAN obtained good posture reconstruction texture information on the multi-pose face reconstruction. Compared with the traditional deep learning method and 3D method, it also achieves higher recognition rate.

**Keywords:** Generative adversarial network · Pose recovery · Face contour · Symmetric loss

## 1 Introduction

In recent years, face recognition [1–4] technology is a hot topic in machine vision research. However, the performance of face recognition depends largely on the influence of various factors such as lighting, posture, and time span. Currently, the posture change is the most challenging face recognition problem. Asthana et al. [5] used the view-based active appearance model (VBAAM) to project images of non-positive faces into an aligned 3D face model for normalizing the pose. With the rapid development of deep learning in recent years, face posture reconstruction based on deep learning has also made great progress. Kan et al. [6] proposed stacked progressive auto-encoders (SPAEs) for face changes caused by different poses. It uses shallow stepping auto-encoding to gradually reconstruct a non-positive face image into a positive face image,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_117

but the partial details of the positive face image recovered are not clear. Later, Zhu et al. [7] proposed a multi-view perceptron (MVP) deep network, which employed multi-layer perceptions to obtain better reconstruction images. In 2014, Goodfellow et al. [8] proposed a generative adversarial network (GAN). This model consists of a generator and a discriminator, where the generator generated new sample data, and the discriminator evaluated the authenticity of the generated data. Generative adversarial network is an unsupervised learning model, which effectively solves a series of data generation problems. Meanwhile, it has more self-learning ability in feature extraction, which makes the image output from the generator more realistic.

Various factors such as lighting, posture, and occlusion will have a certain impact on the recognition performance of realistic faces. For large-angle pose reconstruction issues, a method based on contour symmetric constraint-generative adversarial network (CSC-GAN) is proposed to reconstruct and recognize a multi-face poses. This method competes each other by generated network and adversarial network, optimizes, and trains iteratively until the output images are close to real people's faces. Experimental results show that the proposed method can not only eliminate the attitude error in the large-angle attitude reconstruction, but also enrich the detail of the reconstructed positive face. Compared with other face recognition methods, the recognition rate is also greatly improved.

## 2 Contour Symmetric Constraint-Generative Adversarial Network Model

The network model uses a convolution neural network as a generator in a generative adversarial network, the positive face contour histogram was introduced to constrain the restoration of the positive face to ensure the global contour quality of the reconstructed image, and the data output from this section serves as the generator part of the overall network. The adversarial network uses the real face data as a discriminator to distinguish the authenticity and filter the face data reconstructed in the resolution generator, in order to achieve a minimized synthesis error to introduce a symmetrical loss function. When the network is continuously trained against the generator and discriminator, the entire network is balanced when no real data or data can be generated; the output data is a real image that can be confused. Figure 1 shows the convolution generation adversarial network model. SF is the composite picture output by the generator, and GT is the true face data of the sample.

### 2.1 Generated Network

In the generated network, the image of any pose under normal lighting is processed through the convolution pool, and the output positive face image is used as the output of the generated network. It can be seen from Fig. 1 that the face picture input in the generation network is $w \times h$, and any gesture under normal lighting is $x^0$. The generated network consists of two $5 \times 5$ convolution layers and two $3 \times 3$ pooled layers. $W_i^1$ is the characteristic map produced by the first layer convolution, and $W_j^2$ is the characteristic map produced by the second layer convolution. $V^1$ and $V^2$ are the first

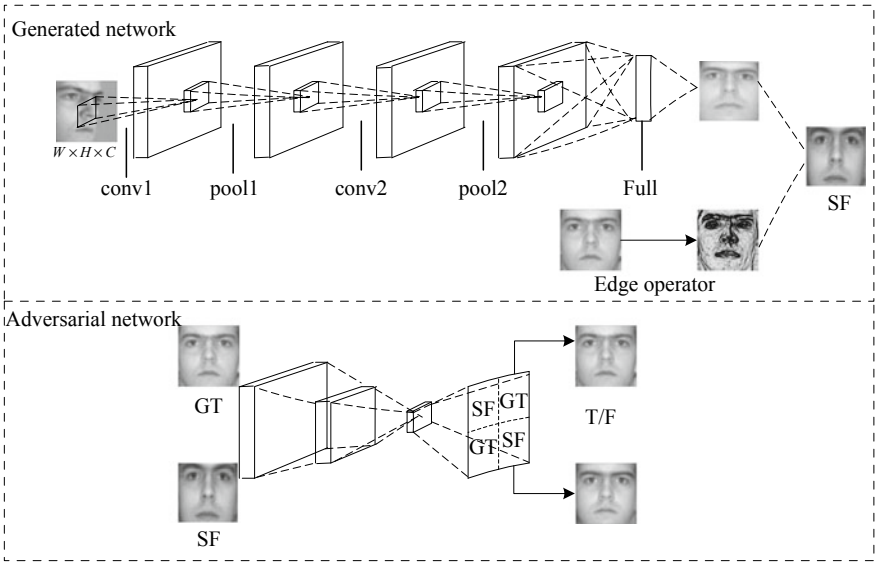



**Fig. 1.** Contour symmetric constraint-generative adversarial network

and second levels of matrix pooling, so the feature map formula of image $x^0$ through two-layer convolution network is as follows:

$$x_i^2 = \sum_{j=1} V^2 \sigma\left(W_j^2 V^1 \sigma\left(W_i^1 x^0\right)\right) \tag{1}$$

where $W$ is the feature map of the weight matrix map, $\sigma$ indicates activation function; select RELU function here. Positive face image output normal lighting is reconstructed positive face $y_1$. The extraction of the feature map $x_i^2$ is performed through the $W$ weight matrix for $W^3$ reconstruction as follows:

$$y_1 = \sigma\left(W^3 x_i^2\right) \tag{2}$$

Generated network is based on convolution network-based reconstruction of the positive face, using gradient descent to continuously update the network parameters' back-propagation; $\Delta$ represent intermediate variables, and $\gamma$ learning rate is 0.0001. The updated parameters $W_{k+1}^i$:

$$\Delta_{k+1} = 0.9\Delta_k - \gamma \frac{\partial E}{\partial W_k^i} \tag{3}$$

$$W_{k+1}^i = \Delta_{k+1} + W_k^i \tag{4}$$

In the process of face reconstruction with a large angle, the restored positive face often results in the discontinuity of the gray value of the edge due to the partial loss of

the overall information feature, and the recognition effect is not good. The edge information histogram can reflect the local discontinuity of the image, along the edge of the pixel intensity to ensure the integrity of the overall contour, and better noise resistance can effectively highlight the edge direction information. In order to ensure that the faces reconstructed at large angles are more realistic and the features extracted are more abundant, after generating a positive face image recovered from the network, a positive face contour histogram is added to constrain the quality of the global features. The central position of the image is represented by any pixel $(i,j)$ of the positive face image $f(x,y)$, and the gradient of the center pixel of the given window in the $m$ and $n$ directions is calculated. Obtain the edge points of the image coordinates $S_{m,n}$ through $S$:

$$S = \arctan\left(\frac{S_m}{S_n}\right)\sqrt{S_m^2 + S_n^2} \tag{5}$$

Convolution of the image output from $\theta$ and convolution networks to obtain a reconstructed positive face image:

$$y_i = \sigma\left(W^3 S \otimes x_i^2\right) \tag{6}$$

Finally, obtain the cost function of convolution network after positive face contour constraint, $X^0$ represents the input image set $\left\{x_i^0\right\}$, $\overline{Y} = \{\overline{y_i}\}$ is the original image, and $Y = \{y_i\}$ is a reconstructed image:

$$E(X^0; W) = \left\|Y_1 - \overline{Y}\right\|^2 \tag{7}$$

## 2.2 Adversarial Network

The adversarial network is similar to a comparative analysis, and the authenticity of the reconstructed image $I^P$ output from the network output is compared with that of the real data $I^F$ until an image similar to the real data is identified. That is, the true and false samples are distinguished, and the two-dimensional matrix values that are opposed to the network output are converted into a probability value. The network is trained by using a separate alternate iterative update method. Remember that the data generated by the input against the network is $G_{\tau_G}(I^P)$ and the real data is $I^P$. First, the generated network $G$ is fixed and trained adversarial network $D$ to maximize the accuracy of the confrontation discrimination. Secondly, the fixed adversarial network $D$ is trained to generated network $G$, so that the accuracy of confrontation discrimination is minimized. This results in the final destination function, which is 1 when the output value is true and 0 when the output value is false. Therefore, when training the network, the adversarial network $D$ is updated twice and the parameters of the generated network $G$ are updated once:

$$\min_{\tau_G}\max_{\tau_D} = E_F\left(X^0; W\right)\log D_{\tau_D}\left(I^F\right) + E_P\left(X^0; W\right)\log\left(1 - D_{\tau_D}\left(G_{\tau_G}\left(I^P\right)\right)\right) \tag{8}$$

## 3 Minimize Loss Function

In the generation network, in order to ensure the naturalness of the generated face, a parameter g is proposed to update the loss function to minimize the synthesis error. Reconstruct the image as the dataset $I_n^P = Y\{y_1, y_2, y_3, \ldots\}$. The real data $I_n^F = \overline{Y}\{\overline{y_i}\}$ is the face positive face image. $L_{\text{syn}}$ is the total loss function. The parameter $\hat{\tau}_G$ is updated by:

$$\hat{\tau}_G = \frac{1}{N} \arg\min \sum_{n=1}^{N} \left\{ L_{\text{syn}}(G_{\tau_G}(I_n^P), I_n^F) \right\} \tag{9}$$

The total loss function $L_{\text{syn}}$ that minimizes the synthesis error is composed of real data discriminating losses and generating symmetry losses of the reconstructed image.$I_{x,y}^P$ is the reconstructed sample pixel level, $I_{x,y}^{gt}$ is the real image pixel level, and $L_r$ is the error loss of training real sample reconstruction:

$$L_r = \frac{1}{w \times h} \sum_{x=1}^{w} \sum_{y=1}^{h} \left| I_{x,y}^P - I_{x,y}^{gt} \right| \tag{10}$$

In the traditional generation adversarial network training learning, the problem of degeneration of generator samples may occur, the same sample points are always generated, and the sample face tends to shrink as the number of iterations increases. According to the characteristics of face symmetry, the width of all input images is covered by half, and the feature points are gradually described by reconstructing the absolute value of the subtracted sample image, which ensures the symmetry of the reconstruction of the visible part and the masked part. Especially for large-angle reconstructions, the asymmetry of the local feature reconstruction is very obvious. Introducing the symmetric loss $L_{\text{sl}}$ can effectively constrain the symmetry of the positive face after reconstruction of different poses, and it also has a face detail that is reconstructed and better richness:

$$L_{\text{sl}} = \frac{1}{w/2 \times h} \sum_{x=1}^{w/2} \sum_{y=1}^{h} \left| I_{x,y}^P - I_{w-(x-1),y}^P \right| \tag{11}$$

$L_{\text{cee}}$ is the cross-entropy loss function used to limit hidden activation functions. $\lambda_1$ and $\lambda_2$ are coefficients of balanced penalty terms, and take the empirical values $\lambda_1 = 0.001$, $\lambda_2 = 0.03$. The final loss function is a weighted loss function whose purpose is to achieve the minimum synthesis error:

$$L_{\text{syn}} = L_r + \lambda_1 L_{\text{sl}} + \lambda_2 L_{\text{cee}} \tag{12}$$

## 4 Experimental Results and Analysis

The paper verifies the effectiveness of the proposed algorithm on the Multi-PIE [9] face image library. The database contains 337 people with a total of 754,204 face images in different poses, lighting, and expressions; each person contains 15 postures, and 20 different lighting are included in the same posture. One of the subsets of the Multi-PIE database was selected, which included 11 poses in the range of −75° to +75°, and 15° between poses. The selected face data is normal expressions under normal lighting. The image size of alignment clip is set to 64 × 64, the batch size is 10, and the learning rate is $10^{-4}$.

Most of the previous multi-pose face–face reconstruction models aim to solve the problem of attitude reconstruction in the ±45° range. In recent years, it is generally believed that it is difficult for a model to recover postures larger than ±60°. However, this paper provided a solution for recovering postures with large deflection angles. In the experiment, Local Gabor binary pattern histogram sequence (LGBP) [10], Stacked progressive auto-encoders (SPAE) [6], Convolutional restricted boltzmann machines (CRBM) [11], Fully automatic Ensemble Gabor Fisher Classifier (FA-EGFC) [12], Deep Convolutional Neural Networks (DCNN) [1] et al. Act recognition rate comparison experiment. From Fig. 2, it can be seen intuitively that the positive face image reconstructed by the CSC-GAN method can restore the texture information of the original image better. Lines 1, 3, and 5 are different gestures for different faces in −75° to +75°. Lines 2, 4, and 6 are face–face images reconstructed by this method when corresponding to different angles. It can be seen from the above that the positive face of the face reconstructed by this method under the large angle of +75° still maintains the symmetry and visual clarity of the original image.

Figure 3 shows the comparison between the method and GAN, DCNN, MVP, and other deep learning methods when the face pose angle is +75°. From the visual observation, it can be seen intuitively that when CSC-GAN reconstructs a large-angle gesture face, facial symmetry is better than other methods, making facial information more natural. In particular, the comparison with the common GAN network also verifies the importance of adding the symmetric loss function in this paper. In addition, facial features such as eyebrows, eyelids, and beards are more comprehensive than DCNN and MVP models, and face reconstruction is more abundant, ensuring more face features.

Table 1 compares the recognition rate of different face reconstruction methods on Multi-PIE face database. It can be seen that when the attitude angle is very small, such as ±15°, the recognition rate of each method is not significantly different. When the angle reaches ±45°, the method of this paper is slightly lower than that of DCNN. However, with the gradual increase of the angle, when the angle is greater than ±60° CSC-GAN compared with other methods, the recognition rate has a greater improvement.

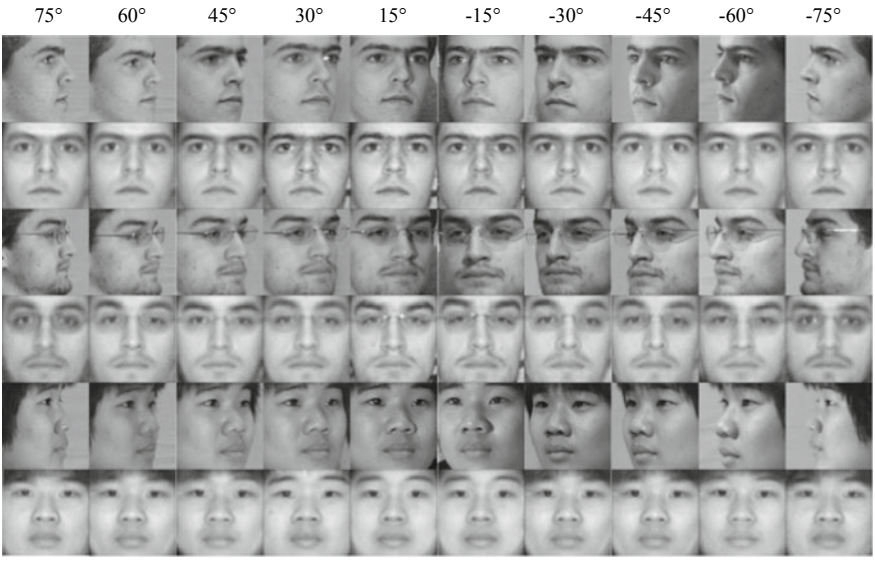



**Fig. 2.** Face reconstructed by CSC-GAN at various angles 75° to −75°

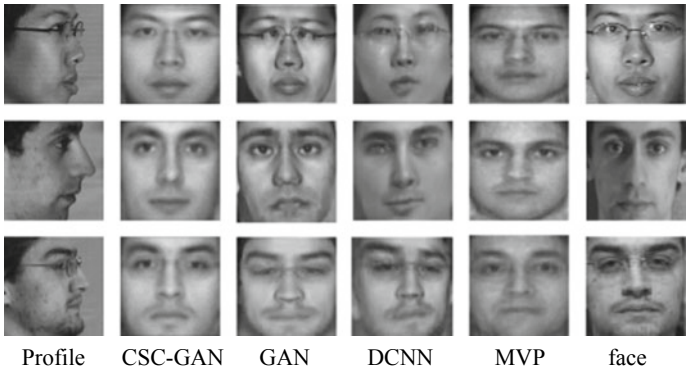



**Fig. 3.** Comparison of CSC-GAN and other reconstruction models at 75°

**Table 1.** Comparison of recognition rates of different face reconstruction methods on Multi-PIE face database

| Methods | −75° | −60° | −45° | −30° | −15° | 15° | 30° | 45° | 60° | 75° |
|---|---|---|---|---|---|---|---|---|---|---|
| LGBP | 10.2 | 20.2 | 37.7 | 62.5 | 77.0 | 83.0 | 59.2 | 36.1 | 22.6 | 11.3 |
| VBAAM | 30.2 | 52.3 | 74.1 | 91.0 | 95.1 | 95.7 | 89.5 | 74.8 | 54.0 | 32.4 |
| CRBM | 40.1 | 69.4 | 80.3 | 90.5 | 94.9 | 96.4 | 88.3 | 75.2 | 65.8 | 44.5 |
| SPAE | 60.4 | 72.2 | 84.9 | 92.6 | 96.3 | 95.7 | 94.3 | 84.4 | 77.6 | 68.2 |
| FA-EGFC | 70.5 | 80.6 | 87.4 | 95.0 | 99.3 | 99.0 | 92.9 | 85.2 | 81.8 | 70.8 |
| MVP + LDA | 72.4 | 80.2 | 89.6 | 96.5 | 100 | 99.0 | 94.5 | 90.8 | 79.8 | 70.2 |
| DCNN + LDA | 74.8 | 82.6 | 95.6 | 98.8 | 100 | 99.3 | 98.2 | 97.8 | 80.4 | 73.2 |
| SC-GAN | **82.6** | **86.2** | 95.2 | **99.4** | 100 | **99.8** | **99.4** | 97.4 | **85.4** | **82.4** |

## 5 Conclusions

We present a method based on contour symmetric constraint-generative adversarial network (CSC-GAN) to reconstruct and recognize a variety of face poses; this method uses a convolution network as a generator to introduce a positive face contour histogram to constrain the global quality and to confront the real data positive face pose. Symmetrical loss is added during the attitude reconstruction to achieve a better posture recovery effect. Experiments show that this method not only eliminates the effect of pose on face recognition performance, but also improves the quality of face reconstruction to a certain extent. It also maintains a good recognition rate after large-angle pose reconstruction and achieves experimental expectations. Future research work will take into account various factors such as lighting and expression to reconstruct the human face and integrate the influence in the actual environment to further improve the recognition rate.

**Acknowledgments.** This work is partially supported by the following foundations: the National Natural Science Foundation of China (61661017); the China Postdoctoral Science Fund Project (2016M602923XB); the Natural Science Foundation of Guangxi province (2017GXNSFBA198212, 2016GXNSFAA38014); the Key Laboratory Fund of Cognitive Radio and Information Processing (CRKL160104, CRKL150103, 2011KF11); Innovation Project of GUET Graduate Education (2016YJCXB02); the Scientific and Technological Innovation Ability and Condition Construction Plans of Guangxi (159802521); the Scientific and Technological Bureau of Guilin (20150103-6).

## References

1. Zhu Z, Luo P, Wang X, et al. Deep learning identity-preserving face space. In: Proceedings of the IEEE international conference on computer vision. Darling Harbour, Sydney; 2013.
2. Taigman Y, Yang M, Ranzato MA, Wolf L. Closing the gap to human-level performance in face verification. In: CVPR, Colombia; 2014.
3. Sun Y, Wang X, Tang X. Deep learning face representation from predicting 10,000 classes. In: CVPR, Colombia; 2014.
4. Taigman Y, Yang M, Ranzato MA, Wolf L. Web-scale training for face identification. arXiv:1406.5266; 2014.
5. Asthana A, Marks TK, Jones MJ, Tieu KH, Rohith M. Fully automatic pose-invariant face recognition via 3D pose normalization. In: ICCV, Barcelona, Spain; 2011.
6. Kan M, Shan S, Chang H, et al. Stacked progressive auto-encoders (spae) for face recognition across poses. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2014. p. 1883–90.
7. Zhu Z, Luo P, Wang X, et al. Multi-view perceptron: a deep model for learning face identity and view representations. In: Advances in Neural Information Processing Systems; 2014. p. 217–25.
8. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, et al. Generative adversarial nets. In: NIPS, Montreal, Canada; 2014.
9. Gross R, Matthews I, Cohn J, et al. The CMU multi-pose, illumination, and expression (multi-PIE) face database. CMU Robotics Institute. TR-07-08, Tech. Rep; 2007.

10. Zhang W, Shan S, Gao W, Chen X, Zhang H. Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition. In: ICCV, Beijing, China; 2005.
11. Huang GB, Lee H, Learned-Miller E. Learning hierarchical representations for face verification with convolutional deep belief networks. In: CVPR, Rhode Island, America; 2012.
12. Li S, Liu X, Chai X, Zhang H, Lao S, Shan S. Morphable displacement field based image matching for face recognition across pose. In: ECCV, Florence, Italy; 2012.

# Flight Target Recognition via Neural Networks and Information Fusion

Yang Zhang(✉), Zhenzhen Duan, Jian Zhang, and Jing Liang

University of Electronic Science and Technology of China
Information and Communication Engineering, Chengdu, China
xixrui@yeah.net

**Abstract.** The purpose of this research is to increase the target recognition rate by means of neural networks and feature fusion. We analyze the performance of different recognition methods (Bayesian classifier, support vector machine (SVM), and neural networks) based on high-resolution range profile (HRRP). The result shows the superiority of neural networks to Bayesian classifier and SVM in classification. We apply multi-source feature fusion to target recognition based on neural networks. The results show that, in certain cases, the target recognition ratio using fusion feature is higher than that of HRRP only.

**Keywords:** Target recognition · Neural networks · Information fusion

## 1 Introduction

High-resolution range profile (HRRP) is the coherent summations of complex echoes from target scatters in each range cell [1]. It represents the projection of target scattering centers on radar line of sight. Since HRRP is easy to be obtained and contains much structure information of targets, it has attracted intensive attentions from the radar automatic target recognition (RATR) community [2,3].

After a long term of research, the US Army Missile Command thought that target recognition based on HRRPs is the most promising identification method in the existing technology [4]. This viewpoint made a rapid development of this method. Zyweck and Bogner [5] proposed a method to solve the problem of aspect sensitivity of HRRP. It uses multi-template matching to carry out target recognition with HRRPs of different aspect. Du et al. [6] proposed a method for calculating the Euclidean distance in higher-order spectra feature space. It avoids calculating the higher-order spectra directly. Botha proposed an identification method based on neural network in 1996 [7]. He thought that, compared with other recognition mehtods, neural networks can extract more unknown features from HRRP and provide a very high target correct recognition rate if the amount of training data is large enough. However, the neural networks he used are simple


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_118

perceptrons. More complex and useful neural networks are not applied because of the lack of theory or hardware performance.

Now, with the development of hardware technology, deep neural networks have applied in various fields. In 2012, Geoffrey Hinton [8] won the competition of ImageNet image recognition by using deep neural network. He reduced the error rate of "top 5" to 15.315%. The same year, Graves A. proposed improved long short-term memory (LSTM) networks [9], which can effectively process the timing issues. For HRRP, deep neural networks can extract more features from it for identification and LSTM networks can process the dependence between sampling points. We believe the target recognition ratio based on HRRP can be improved via deep neural networks.

However, the information of target provided by HRRP is limited. Although the deep neural networks can extract more features, the correlation of these features is too high to improve the recognition rate further. We need to use more other features to do that. Nowadays, the reconnaissance system is mostly composed of multi-source sensor networks. The whole system can work effectively with the fused information obtained by different sensors (visible light sensors, infrared sensors, and radar sensors). It is believed that the target recognition ratio can be improved via the fusion feature of different sensors.

In this paper, we analyze the performance of different recognition methods based on HRRP. The result shows the superiority of neural networks to Bayesian classifier and SVM in classification. Meanwhile, we apply multi-source feature fusion to target recognition based on neural networks. We establish a multilayer perceptron to carry out target recognition with fusion feature obtained by color images, grayscale images, and HRRPs. The results show that, in sunny day, the target recognition ratio using fusion feature is higher than using HRRP only.

The structure of this paper is organized as follows. In Sect. 2, we propose the methods of getting fusion feature of images and HRRPs. Section 3 introduces the neural networks. In Sect. 4, we conduct the training of neural network and the simulation of target recognition. Finally, Sect. 5 summarizes our investigation.

## 2 Feature Fusion

First of all, we need to extract some certain features from the images and HRRP of target to form the fusion feature.

### 2.1 HRRP

HRRP cannot be used for target recognition or feature extraction directly. We must do some processing about it. The amplitude of HRRP obtained by radar sensor is defined as:

$$R = [r(1),\ r(2), \ldots, r(N)] \tag{1}$$

$N$ is the number of the distance units.

The processing of HRRP is power transformation and normalization. Botha et al. [7] think these processing can improve the target recognition rate. Power transformation is defined as:

$$H(n) = r(n)^{\alpha} \quad 0 < \alpha < 1 \tag{2}$$

$H(n)$ is the amplitude of HRRP's sample after power operation.

The next step is normalization. After that, we can calculate the length and the change rate according to [10]. Then, we can obtain a feature vector:

$$F_{radar} = [L,\ t] \tag{3}$$

$L$ is the length of target, and $t$ is the change rate.

### 2.2 Image Feature Extraction

We can obtain color images and grayscale images of target from visible light sensor and infrared sensor, respectively. It is difficult to use the entire images for fusion. So, we need to extract some typical features from these images.

For color image, we extract its color moments [11] for target recognition. The mathematical definitions of three color moments are:

$$u_i = \frac{1}{N}\sum_{j=1}^{N} p_{i,j} \tag{4}$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}(p_{i,j} - u_i)^2\right)^{\frac{1}{2}} \tag{5}$$

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{N}(p_{i,j} - u_i)^3\right)^{\frac{1}{3}} \tag{6}$$

$p_{i,j}$ represents the probability of the appearance of pixel with a gray level of $j$ in color channel $i$. $N$ represents the number of pixels in the image.

According to three color channels of RGB image, we can obtain a feature vector of color moments:

$$F_{color} = [u_R,\ \sigma_R,\ s_R,\ \mu_G,\ \sigma_G,\ s_G,\ \mu_B,\ \sigma_B,\ s_B] \tag{7}$$

For grayscale image, we can extract several features based on gray histogram as follows:

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \tag{8}$$

$$\sigma = \sqrt{\sum_{i=0}^{L-1}(z_i - m)^2 p(z_i)} \tag{9}$$

$$e = -\sum_{i=0}^{L-1} p(z_i) \log_2 (p(z_i)) \tag{10}$$

$m$, $\sigma$, and $e$ represent the average brightness of texture, the average contrast of texture, and the entropy of gray histogram, respectively. $L$ is the number of the gradation, $z_i$ represents the $ith$ gradation and $p(z_i)$ represents the probability of the appearance of pixel with a gradation of $z_i$.

Then, we can obtain another feature vector for fusion:

$$F_{gray} = [m,\ \sigma,\ e] \tag{11}$$

Finally, we can get the fusion vector:

$$F_{fusion} = [F_{gray},\ F_{radar},\ F_{color}] \tag{12}$$

## 3 Neural Network

After getting the features, we begin to introduce the neural network. Here, we only briefly introduce two types of neural networks used in the research.

### 3.1 Multilayer Perceptron

Multilayer perceptron is the most common neural network. It is composed of input layer, hidden layers, and output layer. The structure of multilayer perceptron is shown in Fig. 1.

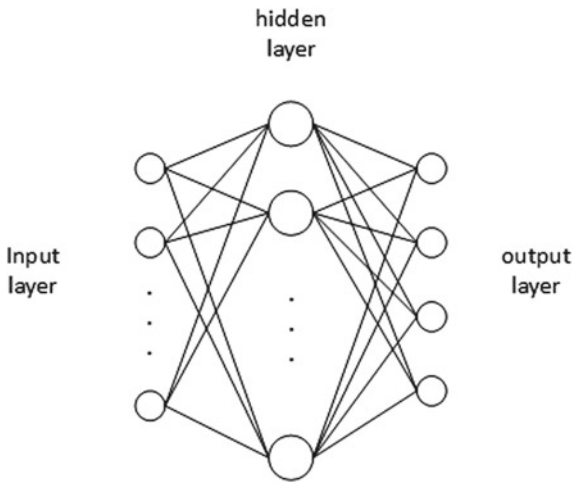



**Fig. 1.** The structure of multilayers perceptron

For multilayer perceptron, we make the entire HRRP or feature vectors as input data. Output is set as the types of targets. We can change the number of features to be extracted by changing the number of hidden layers or the nodes in each layer.

### 3.2 LSTM Network

LSTM network is a variant of recurrent neural network. It can be used for learning long-term dependence information [9]. Each LSTM unit is composed of cell, input gate, output gate, and forget gate. The structure is shown in Fig. 2.

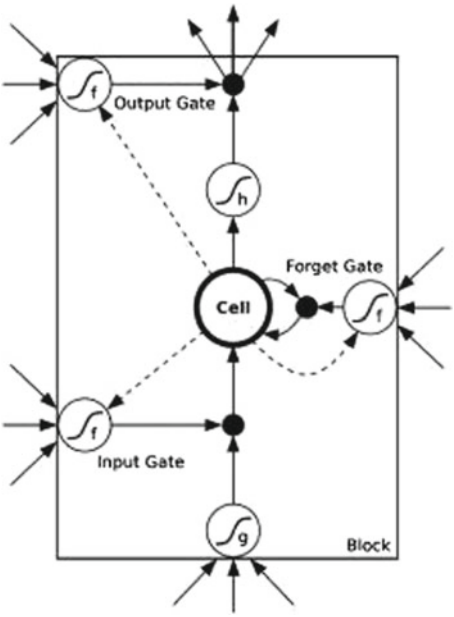



**Fig. 2.** The structure of LSTM networks

In LSTM unit, the first step is to compute what information should be thrown away from the cell state according to the output of last state and current input. The next step is to compute the new information which should be stored in the cell state. After that, updating the old cell state. Finally, computing the output based on current input, current cell state, and the output of last state.

HRRP is a variant of radar echo. And it can be considered as a time series vector. We believe it is reasonable to use LSTM networks to conduct recognition based on HRRP.

## 4 Simulation

The echo data in the simulation obtained by the radar simulation software BSS. In this research, F-15, Tu-16, and AH-64 are selected as targets to be recognized. We get 300 sets of HRRPs, color images, and grayscale images for each kind of aircraft. The attitude angle is transformed from 3 to 30° with the step size of 3°. The number of sampling points of each HRRP is 320.

After processing the data, we can acquire HRRP and $F_{fusion}$ of targets. Then, the rest work is divided into two parts:

1. Performance analysis of different recognition methods based on HRRP (simple perceptron, multilayer perceptron, LSTM networks, Bayesian classifier, and SVM).
2. Performance analysis of recognition method based on fusion feature.

### 4.1 Performance Analysis of Different Recognition Method

This part is to analyze the performance of five different recognition methods. First, we build three different neural networks (LSTM networks, multilayer perceptron, and simple perceptron). The parameters of them are shown in Tables 1, 2, and 3.

**Table 1.** Parameters of LSTM networks

| Time step | Input size | Hidden size | Hidden layer | Output size |
|---|---|---|---|---|
| 16 | 20 | 64 | 1 | 3 |

**Table 2.** Parameters of multilayer perceptron

| Input size | $W_1$ | $W_2$ | $W_3$ | $W_{out}$ |
|---|---|---|---|---|
| 320 | $320 \times 160$ | $160 \times 64$ | $64 \times 16$ | $16 \times 3$ |

**Table 3.** Parameters of simple perceptron

| Input size | $W_1$ | $W_{out}$ |
|---|---|---|
| 320 | $320 \times 160$ | $160 \times 3$ |

In Tables 2 and 3, $W_i$ is the shape of weight matrix of hidden layer $i$. And $W_{out}$ is the shape of weight matrix of output layer.

HRRPs are used as training data of three neural networks. For LSTM networks, We need to reshape the HRRP to (16,20) to fit the input size and time step. The number of training times for all neural networks is 100,000, batch size is 20, the loss function is cross entropy and the optimizer is Adam optimizer.

Then, we add a Bayesian classifier [10] and a SVM to compare the performance of different recognition methods based on HRRP.

In testing, each time we change the SNR, we will conduct 600 tests and calculate the average recognition rate of three aircrafts. The curves of SNR and recognition rate are shown in Fig. 3.

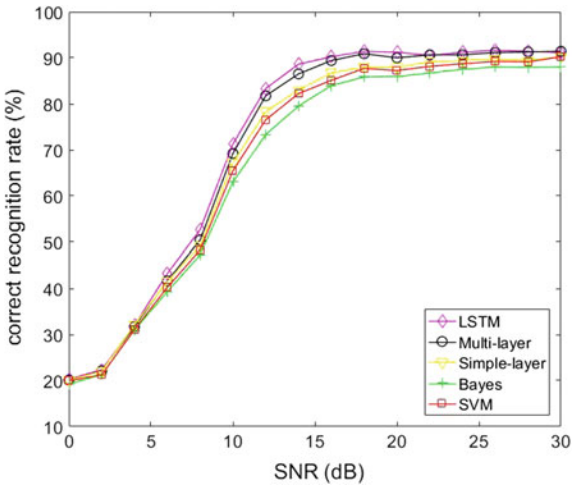



**Fig. 3.** The recognition rate of different methods based on HRRP

Figure 3 shows the performance of five different classifiers based on HRRP. It is obvious that neural networks are superior to other classifiers. Because they can extract more features from HRRP to conduct target recognition. From neural networks, simple perceptron has the worst performance because of the fewer nodes, fewer layers, and fewer features. So, we believe that the performance can be improved by adding the number of layers and nodes. We can see that the recognition rate based on LSTM networks is very close to that of multilayer perceptron. Although LSTM networks have advantages in timing processing, multilayer perceptron have more layers or nodes and can extract more features from the target. So, for target recognition rate, they almost have the same performance.

### 4.2 Performance Analysis of Recognition Method Based on Fusion

It can be learned from the above that neural networks can provide higher target correct recognition rate than that of SVM and Bayesian classifier. In this part,

we establish another multilayer perceptron to conduct target recognition based on fusion vectors $F_{fusion}$. The parameters are shown in Table 4.

**Table 4.** Parameters of multilayer perceptron based on fusion feature

| Input size | $W_1$ | $W_2$ | $W_3$ | $W_{out}$ |
|---|---|---|---|---|
| 14 | $14 \times 64$ | $64 \times 64$ | $64 \times 16$ | $16 \times 3$ |

In testing, we simulate a rainy scene by adding Gaussian blur filter to the parts of images. And we make the rest of the images as test data on sunny day. Then, we can obtain two sets of $F_{fusion}$ to conduct target recognition. Meanwhile, we add two neural networks based on HRRP (LSTM networks and multilayer perceptron) to analyze the performance. The testing results are shown in Figs. 4 and 5.

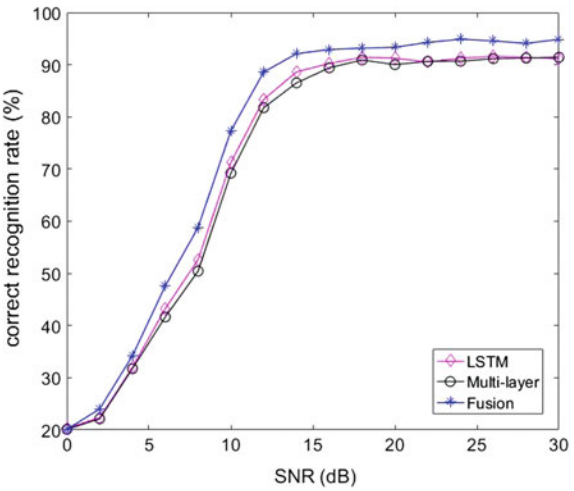



**Fig. 4.** The recognition rate of different methods in sunny day

From Fig. 4, we can see that the correct recognition rate based on fusion is better than that of HRRP only when the SNR changes from 8 to 13 dB. It is believed that, on sunny day, color images and grayscale image can provide important information of targets. Using fusion feature is conducive to the recognition.

Figure 5 shows the performance of method based on fusion is worse than the others on rain day. Because in this case, the resolution of image is too low to contribute to the recognition. The features extracted from images are useless and the effective features are only the length and the change rate extracted from

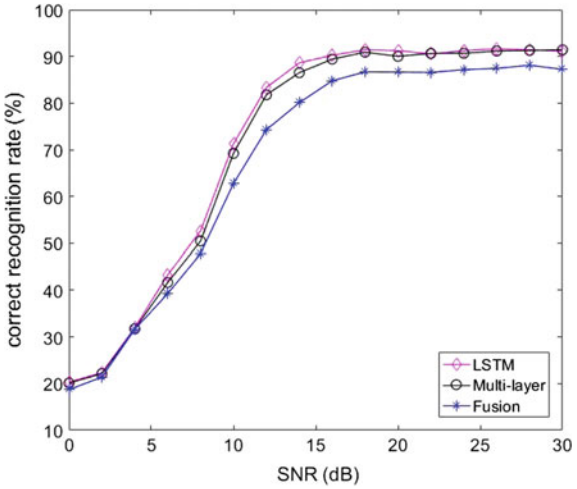



**Fig. 5.** The recognition rate of different methods in rainy day

HRRP. It is difficult to get a high recognition rate based on only two features. So, it is proved that the strategy of recognition should change with the environment.

## 5 Conclusion

In this paper, we analyzed the performance of different recognition methods based on HRRP. The result shows the superiority of neural networks to SVM and Bayesian classifier. The next, we established a multilayer perceptron to carry out target recognition with fusion feature obtained by color images, grayscale images, and HRRPs. The results show that, in sunny day, when the resolution of the images is high, the target recognition ratio using fusion feature features is higher than using HRRP only.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61671138, 61731006), and was partly supported by the 111 Project No. B17008.

## References

1. Smith CR, Goggans PM. Radar target identification. IEEE Antennas Propag Mag. 1993;35(2):2738.
2. Zhang XD, Shi Y, Bao Z. A new feature vector using selected bispectra for signal classification with application in radar target recognition. IEEE Trans Sig Proces. 2001;49(9):1875–85.
3. Kim DH, Seo DK, Kim HT. Efficient radar target recognition using the music algorithm and invariant features. IEEE Trans Antennas Propag. 2002;50(3):325337.

4. Webb AR. Gamma mixture models for target recognition. Pattern Recogn. 2000;33(12):2045–54.
5. Zyweck A, Bogner RE. Radar target classification of commercial aircraft. IEEE Trans Aerosp Electron Syst. 1996;32(2):598–606.
6. Du L, Liu H, Bao Z, Xing M. Radar HRRP target recognition based on higher order spectra. IEEE Trans Sig Process. 2005;53(7):2359–68.
7. Botha EC, Barnard E, Barnard CJ. Feature-based classification of aerospace radar targets using neural networks. Amsterdam: Elsevier Science Ltd.; 1996.
8. Montavon G, et al. Explaining nonlinear classification decisions with Deep Taylor decomposition. Pattern Recogn. 2016;65:211–22.
9. Graves A. Long short-term memory. Supervised sequence labelling with recurrent neural networks. Berlin Heidelberg: Springer; 2012. p. 1735–80.
10. Yang Z, et al. Flight recognition via HRRP using fusion schemes; 2016.
11. Stricker MA, Orengo M. Similarity of color images. In: Storage and retrieval for image and video databases III; 1995. p. 381–92.

# Specific Emitter Identification Based on Feature Selection

Yingsen Xu[1(✉)], Shilian Wang[1], and Luxi Lu[2]

[1] College of Electronic Science and Engineering, National University of Defense Technology, Changsha, China
591297993@qq.com, wangsl@nudt.edu.cn
[2] National Key Laboratory of Science and Technology on Blind Signal Processing, Chengdu, China

**Abstract.** For the high dimension of fingerprint feature set in the process of specific emitter identification (SEI), feature selection method is utilized to reduce the feature dimension and improve individual recognition rate. This paper adopted the filter feature selection in four ways: MIFS, mRMR, CMIM, and JMIM fingerprint feature set of high-dimensional feature selection and combined with PCA dimensionality reduction algorithm to minimize the feature dimension. The simulation results show that feature selection is feasible in individual recognition of the radiation source and can be effectively combined with dimension reduction algorithm.

**Keywords:** Specific emitter identification · Feature selection · Dimension reduction

## 1 Introduction

Due to design tolerance, the manufacture error, equipment aging and the influence of environmental factors, there must be some differences in hardware equipment between different radiation sources, and these differences are shown in the signal that there are subtle features that do not affect the detectability of information transmission. Through fingerprint identification, a certain amount of information can be obtained if the captured enemy signal cannot be decrypted. The technology of communication radiation source identification includes the following steps fine feature extraction, feature processing, and feature classification.

Xu [1] uses the square integral bispectra (SIB) method to extract the radiation source fingerprint characteristics. Literature [2] combined empirical mode decomposition (EMD) and Hilbert spectrum analysis to analyze signal characteristics at time-scale and frequency scale; Bertoncini et al. [3] established fingerprint characteristic quantity by using multi-scale wavelet transform and statistical feature quantity and


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_119

conducted individual fingerprint identification experiments for different RF tags; some scholars used intrinsic time-scale decomposition (ITD) [4] to decompose signal extraction characteristics. The subtle feature extraction method mentioned above based on the signal itself, to achieve high resolution at the same time increased the feature dimension. The existence of feature redundancy may result in a large amount of classification calculation, and at the same time, there may be interference problems between various features. These questions, using the feature selection process, characteristics on the one hand, reduce the feature dimension, on the other hand, reduce the characteristics of redundancy, is helpful to improve the efficiency of emitter individual fingerprint identification.

Feature selection is a work of selecting the most effective features of a set of features to reduce the dimension of feature space [5]. Dash [6] describes the basic framework of feature selection in the literature, as shown in Fig. 1. Feature selection based on evaluation criteria can be divided into filter, wrapper, and embedded [7]. Filter method uses the evaluation criteria to ensure the relationship between the characteristics and classification and reduce the correlation between features and characteristics. Wrapper method evaluates each of the selected feature subsets, to choose the optimal subset. Filter method using the statistical characteristic of all the learning samples, fast at the same time, may be subjected to the imbalance of the small sample and the sample problem, the wrapper method based on the follow-up study to evaluate feature subset, classification effect is best, but is not suitable for large data sets. Because the feature dimension extracted by the feature extraction method mentioned above is sharp, the filter method is chosen.

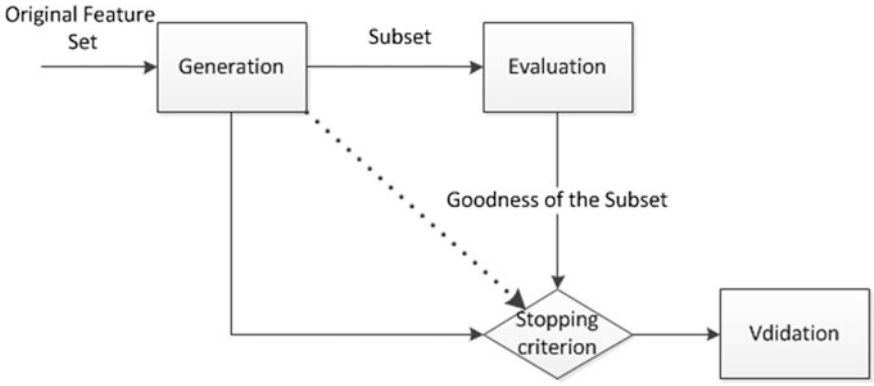



**Fig. 1.** Basic framework of feature selection

## 2 Feature Selection Methods Based on Filter

The main idea of the filter method is to evaluate each characteristic property, distribution of weight, and then choose the feature subset. The method needs to follow the principle to maximize the principle of classification effect and minimize redundancy. According to the evaluation content, it can be divided into four categories: distance measurement, information measurement, dependent measurement, and consistent measurement. In this paper, we choose evaluation feature based on information measurement which usually uses the information gain (IG) mutual information (MI). Information gain is defined as a priori uncertainty and expectations of the difference between a posteriori uncertainty, it can effectively select key characteristics, eliminate irrelevant features. Mutual information describes the interdependence between two random variables.

### 2.1 Related Theory

The method used in this paper is based on the measurement of mutual information. Suppose there are two random variables, one of which provides information on another random variable, that is, mutual information. This section mainly introduces some theoretical knowledge related to mutual information and subsequent methods.

The entropy of a discrete random variable $f = \{x_1, x_2, \ldots, x_N\}$ is denoted by $H(f)$ defined as:

$$H(f) = -\sum_{i=1}^{N} p(x_i) \log p(x_i) \tag{1}$$

where $p(x_i)$ is the probability mass function. For any two variables $f_1$ and $f_2$, the conditional entropy of the variable $f_1$ given $f_2$ is defined as:

$$H(f_1|f_2) = -\sum_{f_1} \sum_{f_2} p(f_1|f_2) \log p(f_1|f_2) \tag{2}$$

And the MI is the amount of uncertainty in $f_1$ due to the knowledge of $f_2$, it is defined as:

$$I(f_1;f_2) = \sum_{f_1} \sum_{f_2} p(f_1,f_2) \log \frac{p(f_1,f_2)}{p(f_1)p(f_2)} \tag{3}$$

MI represents the interdependence between the variables. If the mutual information is larger, it means that the correlation between two random variables is larger, and vice versa. We can also say:

$$I(f_1;f_2) = H(f_1) - H(f_1|f_2) = H(f_1) + H(f_2) - H(f_1f_2) \tag{4}$$

## 2.2 Process of Feature Selection

The feature selection method used in this paper adopts the heuristic search strategy. Specifically, the search in the state space evaluates the location of each search, gets the best location, and then searches from this location to the target. In this way, we can omit a large number of unnecessary search paths and improve the efficiency. The disadvantage is that the most complete subset cannot be obtained.

The filter feature selection method based on mutual information measurement focuses on selecting the appropriate information measurement function. There are many different forms of functions, but the purpose is the same, that is, the selected feature subset has the greatest correlation with the category, and the selected features have the least correlation with each other. The general form of measurement function is [8]:

$$J(f_i) = \alpha \cdot g(f_i, S, C) - \delta \tag{5}$$

where $C$ is the category set, $f_i$ is candidate feature, and $S$ is a subset of selected features.

The basic process of the filter methods is shown in Table 1.

**Table 1.** Basic process of the filter methods

| Initialize: Set $F = \{f_1, f_2, \ldots, f_D\}$, $S = \{\}$ |
|---|
| Step 1: Compute $I(f_i; C) i = 1, 2, \ldots, D$<br>Step 2: Select $f_i = \arg\max(I(f_i; C))$, set $F \leftarrow F \backslash \{f_i\}$ and $S \leftarrow \{f_i\}$<br>Step 3: Select the next feature $f_i$ according to the $J(f)$, set $F \leftarrow F \backslash \{f_i\}$ and $S \leftarrow S \cup \{f_i\}$<br>Step 4: Repeat the third step until $\lvert S \rvert = k$ |
| Output: $S = \{f_1, f_2, \ldots, f_k\}$ |

## 2.3 MIFS and mRMR

MIFS (mutual information feature selection) and mRMR (minimal redundancy and maximal relevance) are selected as the selection methods of fingerprint features of radiation sources, mainly considering that their calculation is simple, rapid, and easy to realize [9]. The information measurement function of the MIFS algorithm considers the candidate feature's contribution to the classification and takes the MI between the candidate feature and the selected feature as the penalty factor. The MIFS evaluation function is:

$$J(f_i) = I(f_i : C) - \beta \sum_{f_k \in S} I(f_i; f_k) \tag{6}$$

where $\beta$ is regulation constant in the interval [0.5, 1].

Compared with MIFS method, mRMR method is more cautious about redundant computation between the features. The idea is that when the quantity of selected features increases, the regulating coefficient in the evaluation function should be relatively reduced. The mRMR evaluation function is:

$$J(f_i) = \frac{I(f_i; C)}{\frac{1}{|S|^2} \sum_{f_k \in S} I(f_i; f_k)} \tag{7}$$

### 2.4 CMIM and JMIM

CMIM (conditional mutual information maximization) determines the importance of candidate features through the dependence degree of candidate features and categories under the condition that the selected feature set is known. The evaluation function of CMIM is:

$$J(f_i) = \arg\max_{f_i \in F} I(f_i; C|S) \tag{8}$$

JMIM (joint mutual information maximization) is similar to CMIM, but there are different ways to define the importance of candidate features. Specifically, there are different ways to select features that can represent the selected subset. The evaluation function of JMIM is:

$$J(f_i) = \arg\max_{f_i \in F} \left( \min_{f_k \in S} (I(f_i, f_k; C)) \right) \tag{9}$$

Venn diagram is used to illustrate the ways in which these two methods can be used to select features that represent selected subsets. $I(f_i, f_k; C)$ is the union of areas 1, 2, and 3; $I(f_i; C|f_k)$ is area 3 in Fig. 2. Because the calculation cost of MI is high and the number of features is large, a single selected feature which maximizes the $I(f_i; C|f_k)$ is selected to replace the entire selected subset in the CMIM method. Formula (8) can be written as:

$$J(f_i) = \arg\max_{f_i \in F} \left( \max_{f_k \in S} (I(f_i; C|S)) \right) \tag{10}$$

Unlike the CMIM method, the JMIM method first finds the $f_k$ that minimizes $I(f_i, f_k; C)$ and then selects the $f_i$.

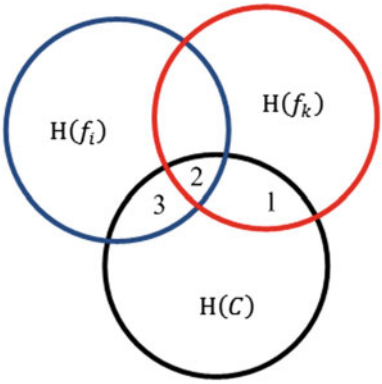



**Fig. 2.** Venn diagram illustrating the relation between feature and class

## 3 Experiments and Results

This article extracted different fingerprint based on the same set of simulation signals used the above four feature selection methods to reduct the dimension of characteristic fingerprint in order to validate the feasibility of the emitter signal features of fingerprint feature selection.

We consider the simulation based on the steady-state characteristics of the radiation source signal and divide the radiation source transmitter into several modules [10]; each module will affect the signal to generate different distortion. In this paper, the *I*/*Q* orthogonal modulator module and power amplifier module are considered.

We add the gain imbalance and orthogonal error of the orthogonal modulator into the simulation model and use Taylor series to model the power amplifier module. Set the Taylor series to the third order: $y(n) = \varepsilon_1 x(n) + \varepsilon_2 x(n)^2 + \varepsilon_3 x(n)^3$.

Five types of transmitters are constructed. Each type of transmitter generates 600 sample signals, including 300 training samples and 300 test samples. Each sample contains 200 randomly generated QPSK modulation symbols. The transmitter simulation coefficient is shown in Table 2.

**Table 2.** Parameter caused signal distortion

| | *T*1 | *T*2 | *T*3 | *T*4 | *T*5 |
|---|---|---|---|---|---|
| Gain imbalance | 0.1 | 0.13 | 0.15 | 0.17 | 0.19 |
| Orthogonal error (angle) | 4.2 | 3.3 | 3.9 | 3.6 | 3 |
| $\varepsilon_1$ | 1.521 | 1.181 | 0.820 | 1.000 | 0.987 |
| $\varepsilon_2$ | −0.104 | −0.001 | −0.001 | 0.109 | 0.014 |
| $\varepsilon_3$ | 0.0021 | 0.0015 | 0.0011 | −0.0019 | 0–0.002 |

In a noise environment where the signal-to-noise ratio is 10 dB, we took 2000 consecutive points from each sample to run the experiments. In the first experiment, we used EMD and ITD to decompose signals, respectively [11], and get five layers from each component signal, we extract each of the signal component samples of permutation entropy, approximate entropy, and fuzzy entropy as the fingerprint characteristics, forming a 40 story feature set. The feature set of 60 layers was extracted by rectangular double spectrometry in the second experiment. On the basis of experiment 1, PCA dimensionality reduction was added. All three experiments were identified by SVM classifier.

In the first experiment, the recognition probability of radiation source is 82% when feature selection is not carried out. Figure 3 shows the change of recognition probability along with the quantity change of selected feature. As can be seen from the figure, the curves of the four methods are basically similar and achieve the best recognition rate when the value of $k$ is in the interval [20, 25]. Meanwhile, the curves of CMIM method and JMIM method are slightly better than those of mRMR method and MIFS method.

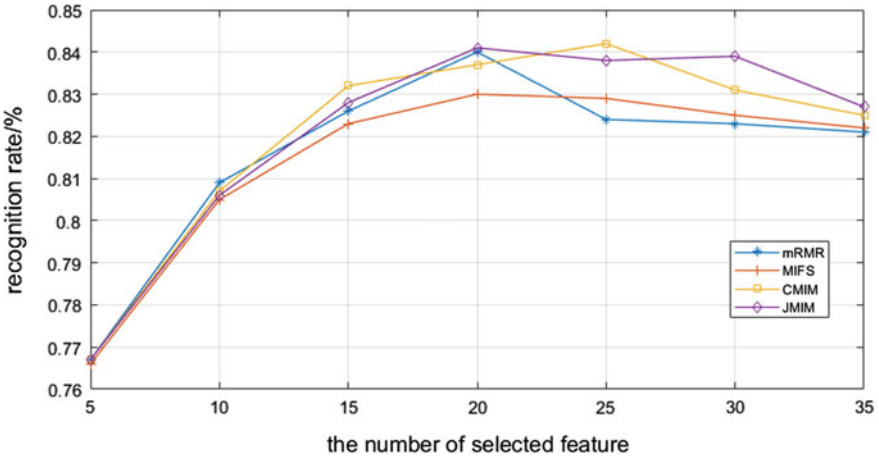



**Fig. 3.** Recognition rate in the first experiment

Figure 4 shows the recognition probability in experiment 2, along with the change of feature selection number. It can be seen that CMIM method and JMIM method have obvious advantages. From the analysis on evaluation function, mRMR method and MIFS method may ignore some features which have high values. Figure 5 shows the relationship between the $f_i$ and $f_k$ in some particular case. When the value of the penalty factor (area of $I(f_i;f_k)$) is close to the value of $I(f_i;C)$, those two feature selection methods cannot select the candidate feature.

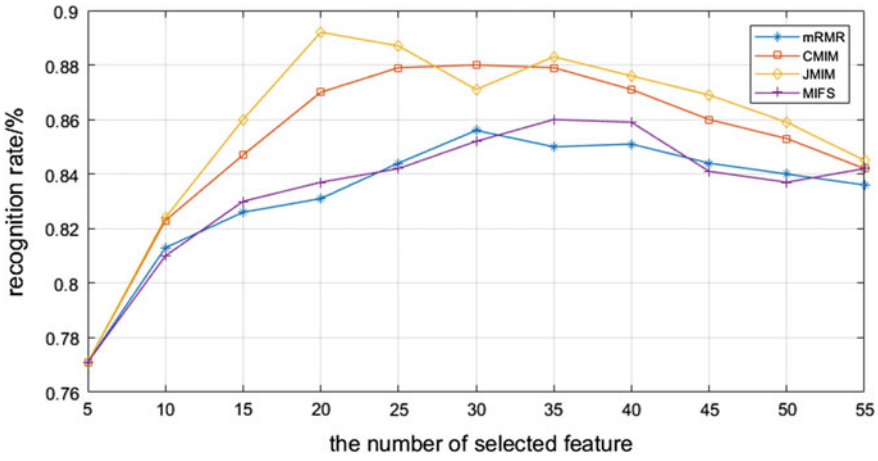



**Fig. 4.** Recognition rate in the second experiment

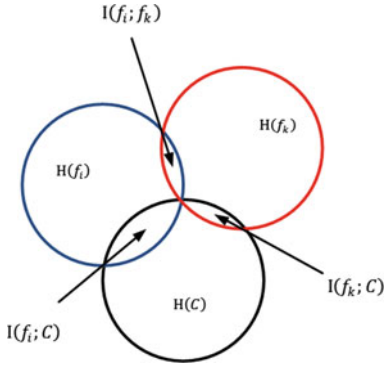



**Fig. 5.** Venn diagram illustrating the relation between the $f_i$ and $f_k$ in some particular case

The simulation results in the third experiment are shown in Fig. 6. The feature subset which includes 20 feature selected by JMIM method was sent into the PCA dimension reduction function. The experimental results show that the feature dimension reduction can be effectively combined with feature selection and bring the advantages that improve emitter individual identification probability and further reduce the subset dimension.

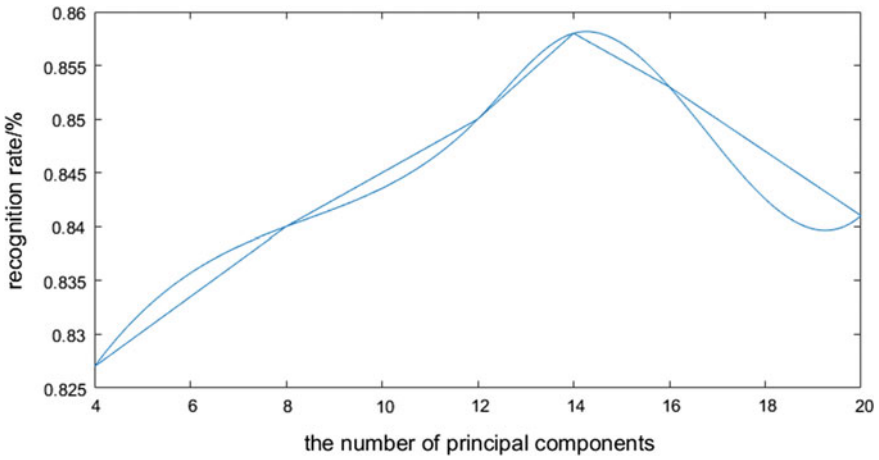



**Fig. 6.** Recognition rate in the third experiment

## 4 Conclusion

Considering the problem of large feature dimension would be met in the process of SEI, feature selection is helpful in reducing the feature dimension and improving individual recognition rate. Moreover, the feature selection can be effectively combined with feature dimension reduce function and bring a better performance.

## References

1. Xu S, Huang B, Xu Z, et al. A new feature vector using local surrounding-line integral bispectra for identifying radio transmitters. In: International symposium on signal processing and its applications. IEEE; 2007. p. 1–4.
2. Yuan Y, Huang Z, Wu H, et al. Specific emitter identification based on Hilbert-Huang transform-based time-frequency-energy distribution features. IET Commun. 2014;8(13):2404–12.
3. Bertoncini C, Rudd K, Nousain B, et al. Wavelet fingerprinting of radio-frequency identification (RFID) tags. IEEE Trans Industr Electron. 2012;59(12):4843–50.
4. Frei MG, Osorio I. Intrinsic time-scale decomposition: time-frequency-energy analysis and real-time filtering of non-stationary signals. Proc Math Phys Eng Sci. 2007;463(2078):321–42.
5. 边肇祺, 张学工. 模式识别(第二版). 清华大学出版社; 2000.
6. Dash M, Liu H. Feature selection for classification. Intell Data Anal. 1997;1(1–4):131–56.
7. Blum AL, Langley P. Selection of relevant features and examples in machine learning. Artif Intell. 1997;97(1–2):245–71.
8. 刘华文. 基于信息熵的特征选择算法研究. 吉林大学; 2010.
9. Somol P, Haindl M, Pudil P. Conditional mutual information based feature selection for classification task. In: Congress on pattern recognition, Iberoamerican conference on

progress in pattern recognition, image analysis and applications. Berlin: Springer; 2007. p. 417–26.
10. 黄渊凌, 郑辉. 通信辐射源指纹产生机理及其仿真. 电信技术研究. 2012;1:1–12.
11. 谢阳, 王世练, 张尔扬,等. 基于差分近似熵和EMD的辐射源个体识别技术研究. 全国信号和智能信息处理与应用学术会议专刊; 2016.

# Nonlinear Dynamical System Analysis for Continuous Gesture Recognition

Wenjun Hou and Guangyu Feng(✉)

School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications, Beijing 100876, China
kinsney@bupt.edu.cn

**Abstract.** Extracting applicable features from continuous gesture is uneasy since it shows up as a nonlinear dynamic system with a spatial–temporal pattern. This paper introduces a continuous gesture recognition framework that analyzes, models, and classifies the nonlinear dynamics of gestures based on chaotic theory. In this system, the trajectories of finger joints are captured as the discrete observations of nonlinear dynamic system, which defines the feature matrix of gestures by reconstructing a phase space through employing a delay-embedding scheme, the properties of the reconstructed phase space are captured in terms of dynamic and metric invariants that include Lyapunov exponent, correlation integral, and fractal dimension. Finally, we extract a feature matrix for training several classifiers with relatively few samples and get best accuracy of around 96.6% to prove our assumption that the nonlinear dynamics of continuous gesture can be approximated by a particular type of dynamical system for classification.

**Keywords:** Continuous gesture recognition · Human computer interaction · Feature extraction · Chaotic theory

## 1 Introduction

Mid-air gesture interaction, as a common way of communication in the real world, has advantages of being potentially natural, efficient and multidimensional in the field of human–computer interaction. This interactive model contains not only the mouse-based panning operations of a desktop coordinate system, but also other unlimited operations with different postures. The requirement of natural interaction is met from the aspects of both multiple dimensions and intuition. Especially in the virtual reality (VR) environment, the mid-air gesture interaction is able to free users from the wearable devices, so that they can enjoy a more immersive experience. Recent years have witnessed the progress of the graphics


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_120

processing algorithms, many hardwares have the ability to capture the trajectories of the skeletons with built-in units. Specifically, LeapMotion and Kinect can identify human's bare hands and obtain the position of finger joints in real time. There have been more and more attention to the mid-air gesture interaction; the priority lies in the semantic processing of a large amount of dynamic spatial–temporal information that is mostly metaphoric, ambiguous, and personalized. The performance of conventional continuous gestures recognition algorithms are mainly affected by such four factors:

1. A large number of degrees of freedom (DOF); The human hand is estimated to possess 27 DOF, with each pose and motion modeled by many more state variables.
2. Nonlinearity. The movement of gestures is inherently nonlinear, which is hard to extract accurate features for describing the motion in mathematical terms.
3. Noise. The appearance of dynamical gesture seems irregular due to different scales, shelter, orientations, luminosity, and other factors.
4. Different input dimensions. Gestures with same meaning are usually captured as time series of different length, while traditional machine learning methods require unified input dimensions.

In the early time, many researchers have borrowed models from speech recognition such as HMM [1], conditional random fields (CRFs) [2] and dynamic time scheduling (DTW) [3]. Among them, the HMM and CRF mainly employ a probability distribution model to establish a adaptive standard of target gestures and unintended gestures, but their temporal modeling capabilities are not good enough for dynamic spatial information. Although DTW methods have the advantage of low computational complexity, their sensitivity to the noise and outliers will reduce the robustness. Besides, descriptive variables such as tangential angular change [4] and coordinate of center point [5] are mostly used in spatial feature extraction, their abilities to characterize the movement are rather weak due to unknown governing equations of the system. There are also researchers focusing on gesture strokes and proposing a combination of new features such as left and right sector trajectory features, but they are too dependent on orientation, distance and other spatial features of strokes, and the differences between some similar gestures with confusing strokes such as five with S are hard to tell [6]. Recently, there have been growing interests in convolutional neural network (CNN) [7] and recurrent neural network (RNN) [8]. The hardware revolution has greatly increased the speed of computers and helped the original complex algorithms of deep learning to get a finer result than in the past. According to the results and analysis of the 2016 ChaLearn [9], the best recognition accuracy of isolated gesture recognition has improved from 56.90 to 67.71%, the Mean Jaccard Index (MJI) of continuous gesture recognition has improved from 0.2869 to 0.6103. However, their massive computing power does not pay off in the practical system, this reveals the fact that there is still a lack of effective modeling methods in feature engineering.

The aim of this paper is to provide a modeling method of continuous gesture recognition based on chaotic theory. The hypothesis is made that the irregularly

sampled data of continuous gestures can be approximated with a particular type of dynamic system, and the characterization of this nonlinear dynamics will help with the establishment of feature vectors. The remainder of this paper is organized as follows: Sect. 2 presents the methodology of main proposals and its solutions. The proposed algorithm is validated experimentally on a set of mid-air alphabetic gestures in Sect. 3. Future work and conclusion are given in Sect. 4.

## 2 Chaotic Analysis of Continuous Gesture Recognition

When performing a particular alphabetical gesture, the trajectory we draw with bare hand usually contains not only the stroke itself, but also those unintended transition movement. When and how participants begin and end their gestures depend on their own habits, which make the pattern recognition more difficult. Besides, the underlying mechanism is hard to characterize in mathematical terms, descriptive statistics such as position of average coordinate points or orientation angle are not sufficient for characterizing since they make no physical sense without a detailed mathematical knowledge of the underlying dynamics.

The novelty of this paper stems from assuming a particular type of dynamical system to approximate the nonlinear dynamics of gesture and describe the features by discussing the structure and stability of this system. In other words, rather than letting data speak for itself about numbers of independent variables and unknown parameters, we try to make assumptions about it and fit the experimental data to the model by finding the parameters that best explains the patterns.

### 2.1 Strange Attractor

Assuming that there is a determinism present in the seemingly stochastic dynamics of gestures, which can be used to extract rich information for identification and classification. In a mathematical field of dynamical systems, the attractor is used to denote the set of numerical values toward which system tends to evolve, when the trajectory in an attractor with similar initial conditions tend to move apart with increasing time, there is a strange attractor characterizing this chaotic system with sensitive dependence on initial conditions. The next issue to be considered is the invariants of system's attractor, which also decides the properties of the constructed phase space for further classification. There are three major features in discussion—correlation integral, fractal dimension, and maximal Lyapunov exponents. The correlation integral is the mean probability that the states at two different times are close, while the fractal dimension measures the change in the density of phase space with respect to the neighborhood radius and characterizes the geometric structure of a strange attractor. The Lyapunov exponents describe how trajectories on the attractor move under the evolution of the system dynamics and the uncertainty about the future state. A chaotic

process is generated by a nonlinear deterministic system with at least one positive Lyapunov exponent, which as a sufficient condition for stability also helps us filter the training data. As these exponents quantify the exponential rate at which nearby trajectories separate from each other while moving on the attractor, the maximal Lyapunov exponent is chosen as a dynamical invariant of the attractor to measure the exponential divergence of the nearby trajectories in the phase space.

### 2.2 Dynamic Reconstruction

Dynamic reconstruction describes a state of mapping from one-dimensional signal to an n-dimensional signal, which provides the dynamical modeling of a time series for an unknown n-dimensional system to capture the underlying dynamics. A fundamental method in dynamic reconstruction theory is a geometric theorem called the delay-embedding proposed by Takens [10], which reconstructs a phase space equivalent to the original phase space by delaying coordinate values. To form a matrix with each row representing a point in the reconstructed phase space (RPS), delayed copies of the original time series are used to slide a window of length $m$ through the time series, and stack the m-dimensional vectors into the matrix with $m$ noting embedding dimension and $t$ noting delay parameter [11].

$$X = \begin{pmatrix} x_0 & x_t & x_{2t} & \cdots & x_{(m-1)t} \\ x_1 & x_{t+1} & x_{2t+1} & \cdots & x_{1+(m-1)t} \\ x_2 & x_{t+2} & x_{2t+2} & \cdots & x_{2+(m-1)t} \\ x_3 & x_{t+3} & x_{2t+3} & \cdots & x_{3+(m-1)t} \end{pmatrix} \tag{1}$$

There are two main methods to estimate a proper embedding delay: sequence correlation method and phase space geometry. The autocorrelation is a kind of sequence correlation method, which reduces the correlation between the reconstructed time series after selecting the delay time by the autocorrelation function to make the sequence dynamics less lost as much as possible. However, this method is essentially a linear concept, suitable for judging linear correlation rather than a nonlinear system. Based on this, Fraser and Swinney suggest using the first local minimum of the mutual information between the delayed and non-delayed time series, effectively identifying a value of delaying time for which they share the least information [12]. The false nearest neighbor method is used for finding the optimal embedding dimension $m$ [13]. From the geometric point of view, the chaotic time series is the projection of the high-dimensional phase space on one-dimensional space with distorted trajectory, while the false neighbor points are those independent points which turn out be adjacent after projection in high-dimensional space and make chaotic time series irregular. With the increase of the embedded dimension $m$, the false neighbor points will be gradually removed and the trajectory of the chaotic motion will be restored to regular one. This reveals the fact that if points are sufficiently close in a reconstructed phase space, then they should remain close during a forward iteration, a

phase space point that does not fulfill this criterion has a false neighbor. Figure 1 shows the three-dimensional projection of the reconstructed phase space for the chosen values of $t$ and $m$.

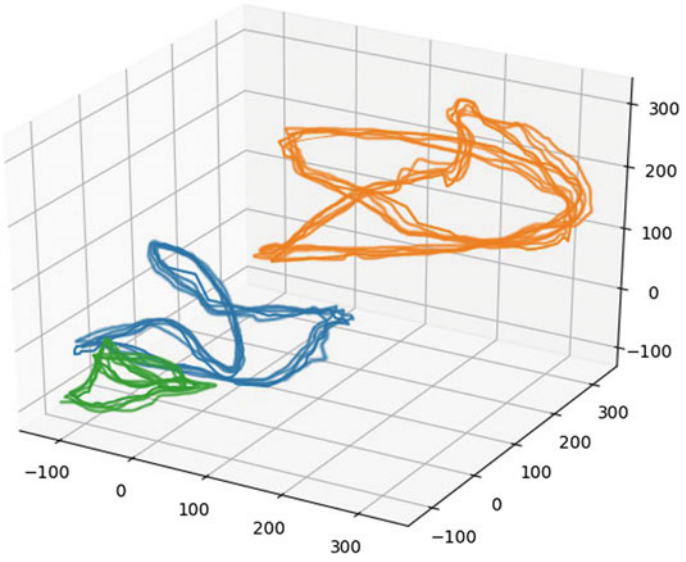



**Fig. 1.** Three-dimensional projection of reconstructed phase space

### 2.3 Correlation Factors

The correlation factors include correlation integral and fractal dimensions, they two characterize the metric structure of the attractor. Of them, the correlation integral measures the number of points within a neighborhood of radius, averaged over the entire attractor as

$$C(\varepsilon) = \frac{2}{N \times (N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} H(\varepsilon - \|\boldsymbol{x_i} - \boldsymbol{x_j}\|), \tag{2}$$

where $\boldsymbol{x}_i$ are points on the attractor, $H(x)$ is the Heaviside function and $\varepsilon$ is the value of radius. When set a specific value of $\varepsilon$ to compute $C(\varepsilon)$, we can get the correlation integral as a feature vector. In our training, we go through several gesture samples to observe the stability of data with varying radius, and finally set 50 for a distinguishing distribution. The correlation function of the attractor, denoted as $C(q, r)$, describes the probability that any two points on the attractor are separated by a distance $r$ for some integer $q$, indicated as

$$C(q, r) = r^{(q-1)D_q}. \tag{3}$$

For $q = 2$ and taking the logarithm of both sides, the resulting dimension

$$D_q = lnC(\varepsilon)/ln\varepsilon \tag{4}$$

is correlation dimension we need and bounds the degrees of freedom required to describe the system.

## 3 Experiment Result

### 3.1 Tracking Results of Gestures

In our experiment, we choose eleven continuous alphabetic gestures for classification viz 'A', 'B', 'E', 'F', 'H', 'I', 'K', 'P', 'R', 'X', and 'Z'. With the initial coordinate data captured from LeapMotion, we extract each nonlinear trajectory, respectively, in x, y, z dimensions from finger joints movement, and draw the approximate coordinate function after smoothing (shown in Fig. 2).

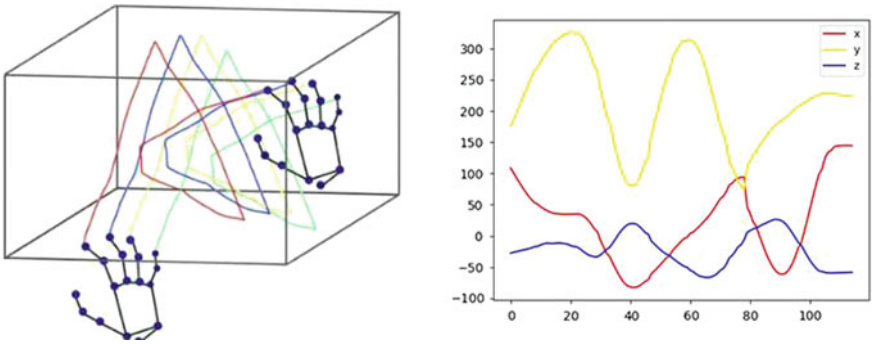



**Fig. 2.** The trajectory of mid-air gesture

### 3.2 Classification Results

The performances of the feature extraction are evaluated using three classifiers viz k-NN, SVM, and random Forest. In our experiment, fivefold cross-validation is conducted to analyze the performance of each classier with different parameters, the dataset is equally divided into five subsets, among them four subsets are used for training, the other is for testing. The framework we proposed put a lot of concentration on feature engineering, which improves the performance of the model and even achieves a accurate result from a simple dataset. Our dataset has 200 gestures of which 160 are used for training and remaining 40 are used for testing. Finally, the average accuracy of cross-validation is calculated and shown in Table 1. The three-dimensional projection of feature vector is shown in Fig. 3 with good clustering performance.

**Table 1.** The average recognition rate of each alphabetic gesture in three classifiers

| Gesture | K-NN (%) | Random forest (%) | SVM (%) | N/G (%) |
|---|---|---|---|---|
| A | 95 | 97.5 | 97.5 | 2 |
| B | 87.5 | 92.5 | 80 | 1 |
| E | 97.5 | 100 | 95 | 1.5 |
| F | 97.5 | 100 | 95 | 2.5 |
| H | 92.5 | 92.5 | 92.5 | 3 |
| I | 95 | 100 | 97.5 | 3.9 |
| K | 100 | 100 | 100 | 2.5 |
| P | 90 | 92.5 | 80 | 3 |
| R | 85 | 90 | 80 | 4.5 |
| X | 85 | 100 | 95 | 9.5 |
| Z | 92.5 | 100 | 92.5 | 7.8 |
| Overall | 91.8 | 96.6 | 91.6 | 3.7 |

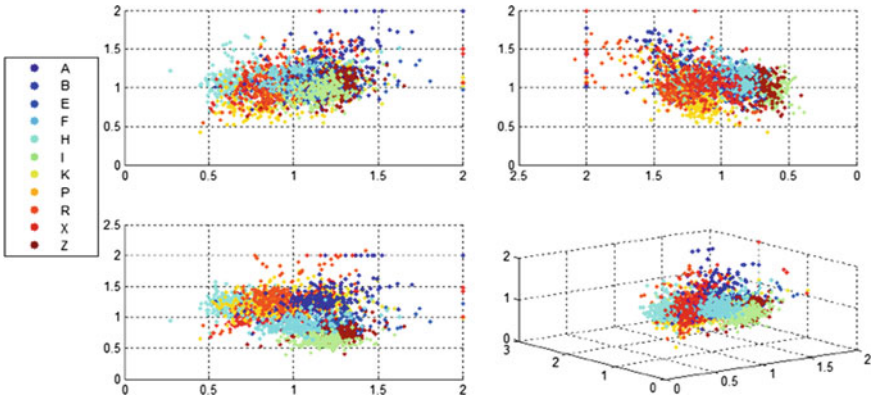



**Fig. 3.** Distribution of chaotic factors in feature matrix

From the average recognition classification accuracy of several alphabets' gesture, we can draw a conclusion that the algorithm we proposed for feature extraction applies to these three popular classification algorithm efficiently. Specifically in case of individual's gesture, the alphabets' gesture 'F', 'H', 'I', 'K', and 'Z' achieve the best performance of about 100%, they all share a common character of presenting as sequential movement, while for alphabets' gesture 'B', 'P', and 'R' with certain curves, it is uneasy for LeapMotion to capture the real trajectory without compensation, which put some constrains and effect on recognition, this will help with our further design of some interactive gestures. The experimental results suggest that our model provides pretty good recognition rate for most of the gestures; however, there is still reduction in overall accuracy due to mis-

classification of few similar gestures, such as D with P, E with F, and I with L, which is also in line with our expectations and indirectly illustrates the effective characterization of chaotic variables.

## 4 Conclusion

There are very few studies on the application of chaotic theory in dynamical gesture recognition. In this paper, we have attempted to provide some insights into feature extraction to recognize continuous alphabetical gestures, the recognition rate of around 93% in three popular classifiers have proven our assumption that these dynamical variables can characterize gesture motions in part and help us with the further study from a new perspective of chaotic theory. However, these three indicators we extract through phase space construction can only characterize the approximate chaotic movement tendency to some extent, which means our classifiers still cannot tell the difference between similar motions, there are still many valid feature variables to be explored. So our future work will include finding more indicators that characterize the gesture motion precisely to improve the recognition performance and design easy-to-identify interactive gesture considering its chaotic structure.

## References

1. Yang Z, Narayanan SS. Modeling dynamics of expressive body gestures in dyadic interactions. IEEE Trans Affect Comput. 2017;8(3):369–81.
2. Yang HD, Sclaroff S, Lee SW. Sign language spotting with a threshold model based on conditional random fields. IEEE Trans Pattern Anal Mach Intell. 2009;31(7):1264–77.
3. Celebi S, Aydin AS, Temiz TT, Arici T. Gesture recognition using skeleton data with weighted dynamic time warping; 2013. p. 620–5.
4. Vo DH, Huynh, HH, Nguyen TN, Meunier J. Automatic hand gesture segmentation for recognition of Vietnamese sign language; 2016. p. 368–73.
5. Lu W, Tong Z, Chu J. Dynamic hand gesture recognition with leap motion controller. IEEE Sig Process Lett. 2016;23(9):1188–92.
6. Singha J, Misra S, Laskar RH. Effect of variation in gesticulation pattern in dynamic hand gesture recognition system. Neurocomputing. 2016;208:269–80.
7. Wang P, Li W, Liu S, Gao Z, Tang C, Ogunbona P. Large-scale isolated gesture recognition using convolutional neural networks. In: ArXiv e-prints; 2017.
8. Chai X, Liu Z, Yin F, Liu Z, Chen X. Two streams recurrent neural networks for large-scale continuous gesture recognition; 2016. p. 31–6.
9. Wan J, Escalera S, Escalante HJ, Baro X, Guyon I, Allik J, Lin C, et al. Results and analysis of ChaLearn LAP multi-modal isolated and continuous gesture recognition, and real versus fake expressed emotions challenges; 2017.
10. Takens F. Detecting strange attractors in turbulence. Berlin, Heidelberg: Springer; 1981. p. 366–81.
11. Oselio B, Hero A. Dynamic reconstruction of influence graphs with adaptive directed information; 2017.

12. Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. Phys Rev A. 1986;33(2):1134–40.
13. Brown L, Abarbanel J, Brown D, Kennel J, Kennel C. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys Rev A. 1992;45(6):3402–4311.

# Feature Wave Recognition-Based Signal Processing Method for Transit-Time Ultrasonic Flowmeter

Yanping Mei[1(✉)], Chunling Zhang[1], Mingjun Zhang[1], and Shen Wang[2]

[1] City Institute, Dalian University of Technology, Dalian, China
meiyanp@dlut.edu.cn
[2] Dalian Hui Ming Instrument Co., Ltd., Dalian, China

**Abstract.** In order to improve the measuring precision and stability of transit-time ultrasonic flowmeter as well as the locating accuracy of datum point for ultrasonic received signal, a feature wave recognition-based signal processing method is proposed in this study, which derives from analyzing the cause of errors in conventional threshold approach. By introducing a phase-shifted pulse into the ultrasonic excitation one, a feature wave with different period and phase is consequently produced in the ultrasonic received signal and recognized using a high-precision TDC chip according to the period of the received signal at first. Then the datum point of the received signal is accurately located with regard to the relationship between the position of feature wave and the initial position of the received signal so that the transit time of ultrasonic signal is finally measured. The following experiments focusing on a real-world problem demonstrate that the proposed method can effectively reduce the measurement errors caused by the amplitude change of the received signal. Such an approach is greatly beneficial for improving the precision of measurement along with the stability of the ultrasonic flowmeter.

**Keywords:** Ultrasonic flowmeter · Transit time · Signal processing · Feature wave recognition · Measuring precision

## 1 Introduction

Ultrasonic flowmeter (UF), with its non-contact characteristic and superiority on measuring accuracy, has become one of the most widely applied flow measurement instruments [1]. And the transit-time ultrasonic flowmeter (TTUF) is always the first alternative for solving practical problems regarding the fast response, easy installation, and low maintenance costs.

Considering the relationship between the transit-time difference caused by upstream and downstream and the velocity when ultrasonic pulse transmits through the flowing medium, the TTUF can calculate the flow rate, in which the transit time of ultrasonic signal plays a pivotal role for the flow rate measurement [2–4]. Such a parameter is mainly affected by the measurement precision of transit time and the detection accuracy of ultrasonic received signal. As for the former, it can be solved by utilizing time-to-


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_121

digital converter (TDC), of which the measuring precision is typically up to picosecond level nowadays [5]. Bearing this in mind, the accurate detection of ultrasonic received signal becomes the most important factor affecting for the flow rate measurement. Conventional threshold method is commonly deployed for detecting the ultrasonic received signal, whereas measurement error is always produced by signal attenuation or noise disturbance. As such, various solutions are proposed to solve this problem. For instance, a digital signal processing algorithm using wavelet threshold for de-noising is proposed by Meng et al. [6]. A variable delay-time filter method is proposed by Zhao et al. [7]. A variable threshold-based zero-crossing detection signal processing method to determine a feature wave as well as the zero-crossing point and the transit time is presented by Wang et al. [8]. A dual-threshold method involving the determination on the abnormality of the signal is reported by Chen et al. [9].

In this paper, aiming at the measurement error and poor stability of the conventional threshold method caused by the noise in favor of practical application, considering the fact that the amplitude of the received signal could be changed easily while its period and phase are not changed significantly by the noises, a feature wave recognition-based signal processing method is proposed in this study. By means of introducing a phase-shifted pulse into the specific position of ultrasonic excitation one, a received signal with feature wave is consequently produced, of which the feature wave can be recognized by a high-precision TDC chip according to its period. Based on the position relationship with the feature wave, the datum point of the received signal is accurately located which leads to an accurate detection of ultrasonic received signal. Such an approach essentially reduces the measurement errors comparing with the one of threshold-based methods caused by the amplitude change of the received signal. The following experimental results show that the proposed method exhibits superior performance over the conventional methods, and the real-world application further demonstrates its practicability.

## 2 Analysis of Signal Processing

### 2.1 Error Analysis of Conventional Threshold Method

The principle of conventional threshold method is shown in Fig. 1. If the amplitude of the received signal reaches the preset threshold voltage, the detection circuit will trigger the zero-crossing detection and produce the stop pulses at the same time, then the ultrasonic transit time will be measured. However, when the amplitude of the received signal attenuates by disturbance, the stop pulse will have at least one periodic error because the datum point is changed. As such, the measurement error of transit time will be produced.

It can be depicted that the main problem for conventional threshold method is the detection error of the received signal. Therefore, it is of great significance to improve the accuracy on locating the datum point of ultrasonic received signal regardless of the varied amplitude caused by the disturbance.

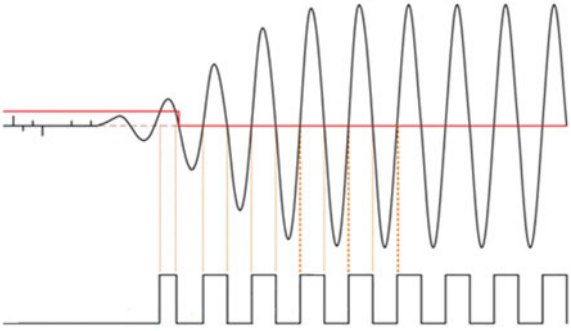


**Fig. 1.** Conventional threshold method

### 2.2 Feature Wave Recognition-Based Signal Processing Method

According to the basic principles of underwater acoustics [10], the vibration produced on its surface when the transducer receives the pulse signal can be described as

$$F = F_m \sin(\omega t + \varphi) \tag{1}$$

where $F_m$, $\omega$, and $\varphi$ denotes the amplitude, angular frequency, and phase angle of vibration. The received signal shares the same period and phase with the ones of (1), but its amplitude is related to the velocity of the vibration which can be calculated as

$$u_2 = \frac{2P_1 S}{Z_M} \tag{2}$$

where $P_1$ denotes the acoustic pressure of incident wave, $S$ and $Z_M$ refers to the section area and mechanical impedance of receiver transducer, respectively. When the acoustic waves are interrupted by disturbance during propagation, the $S$ in (2) is correspondingly reduced which will cause the amplitude attenuation of the received signal, whereas the related period and phase will not be influenced. Taking such characteristic into account, the feature wave recognition-based signal processing method is proposed. The detailed elaboration as well as analysis is as follows.

A normal excitation pulse and its received signal are shown in Fig. 2. If a 180° phase-shifted pulse is introduced into the excitation one, its received signal will be changed as shown in Fig. 3. It can be depicted by comparing the two cases that the period and phase of the received signal in specific position are varied by a 180° phase-shifted excitation pulse, where a feature wave is introduced into the specific position of the received signal. Considering the favorable recognition of feature wave and the sensitivity and vibration speed of the utilized transducer with a resonant frequency of 1 MHz, an ideal received signal with higher initial amplitude and distinct feature wave can be obtained if the 180° phase-shifted pulse is inserted after the third excitation pulse as shown in Fig. 3.

Take a transducer with a resonant frequency of 1 MHz as an example, its received signal with feature wave is shown in Fig. 4. After the zero-crossing detection, the

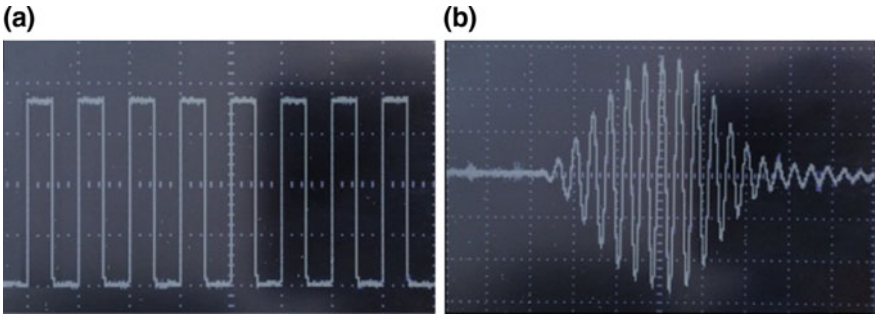



**Fig. 2.** **a** Normal excitation pulse and **b** received signal

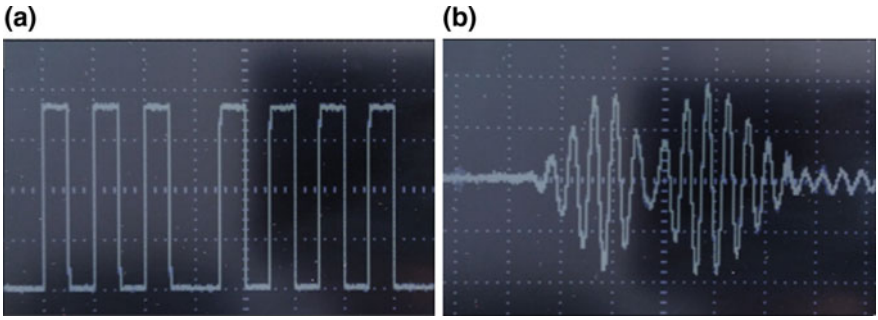



**Fig. 3.** **a** Excitation pulse with 180° phase-shifted and **b** received signal with feature wave

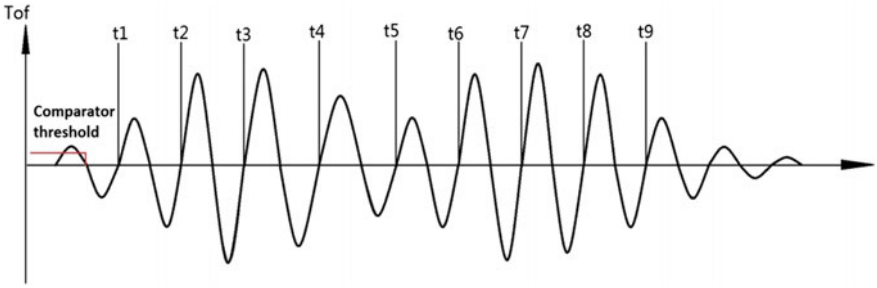



**Fig. 4.** Received signal with feature wave

periods of the received signal can be measured, i.e., $t2 - t1$, $t3 - t2$, $t4 - t3$, $t5 - t4$, $t6 - t5$,…, and the period change rate is defined as,

$$S\_{\mathrm{p}} = \frac{T_n}{T_{\mathrm{s}}} \quad (n = 1, 2, 3, \ldots) \tag{3}$$

where $T_n$ is the period of the $n$th received signal after zero-crossing detection; $T_s$ is the standard resonant period of the transducer used.

A threshold value of the period change rate $S\_{std}$ is preset according to the characteristics of the transducer and the practical experience. If $S\_{p} < S\_{std}$, it can be depicted there are no changes in the periods of the received signal. Else if $S\_{p} \geq S\_{std}$, then the periods of the received signal are considered change, and the first change received signal wave is identified as the feature wave. The position relationship between the feature wave and the initial wave of the received signal can be determined by the position of phase-shifted pulse signal in the excitation pulse. Thus, the datum point of the received signal can be located accurately.

## 3 System Tests and Comparative Study

### 3.1 System Hardware Structure

Figure 5 gives the system hardware structure of TTUF method on the basis of the feature wave recognition. EFM32G880F128 is used to be the controller of the system. The high-precision TDC chip MAX35101 is used to transmit and receive the ultrasonic signal as well as measure the transit time of ultrasonic signal. The phase-shifted pulse drive circuit is responsible for generating an excitation pulse with fixed sequence. The dual-channel SPDT analog switch TS5A23159 is adopted to the transfer switch for the ultrasonic transmit circuits and the receive circuits.

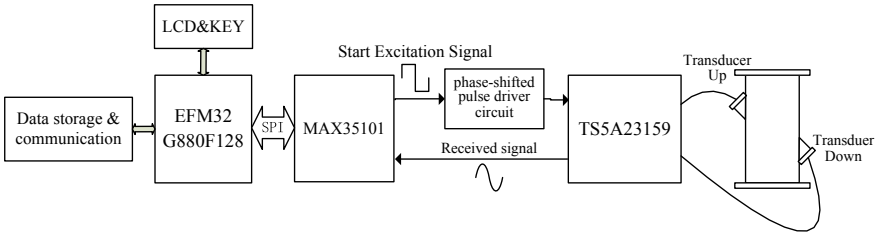



**Fig. 5.** System hardware structure diagram

### 3.2 Transit-Time Measurement

The accurate measurement of transit time requires the requirement on TDC chip is not only on the high-precision, but also having the capability of recording multiple times stop pulse for supporting the feature wave recognition method. The measurement precision of MAX35101 is up to 20 ps, and it can continuously record six stop pulses at one time, i.e., five periods of the received signal can be measured continuously, which the recognition of feature wave can be completed within one-time measurement scan.

When the measurement stars, MAX35101 emits the ultrasonic excitation pulses with no phase-shifted pulse in order to measure the $T_s$ in (3) firstly, then it reemits one excitation pulse to the phase-shifted pulse drive circuit as well as generate a start pulse. The phase-shifted pulse drive circuit immediately generates a pulse sequence containing ten excitation pulses in which the first three pulses have 180° phase difference with the latter seven ones. When the signal containing feature wave is received, six stop pulse signals are produced by MAX35101 with which five consecutive periods of the received signal and the $S\_{p}$ are calculated. Comparing the $S\_{p}$ and the $S\_{std}$ presented, the corresponding feature wave is identified.

According to the position of the phase-shifted pulse introduced in the excitation pulse, it can be determined that the third wave before the feature wave is the initial received signal wave. Consequently, the datum point of the received signal is located and the transit time of ultrasonic is measured.

### 3.3 Comparative Study

In order to verify the effectiveness of the feature wave recognition method presented in this paper, two comparison experiments are conducted to the TTUF based on feature wave recognition method (TTUF-FWR) and the TTUF based on conventional threshold method (TTUF-CT): the transit-time measurement experiment on simulated condition and the flow rate measurement experiment on flow test platform.

#### 3.3.1 Transit-Time Measurement Experiment

This experiment is conducted to test the influence for the precision and stability of transit-time measurement when the amplitude of the received signal changes. The transducers with the resonant frequency of 1 MHz are installed on the TTUF with DN50 in the 16 °C static clean water, and the voltage of excitation pulse is 3.0 V, the threshold value of comparator is set to 30 mV. The amplitude attenuation of the received signal is simulated by adjusting the resistance of resonant resistor which series with the transducers. Table 1 shows the results of measurement.

**Table 1.** Results of transit time measurement with different methods

| Maximum amplitude of received signal (mV) | TTUF-FWR | | TTUF-CT | |
|---|---|---|---|---|
| | Downstream transit time (μs) | Upstream transit time (μs) | Downstream transit time (μs) | Upstream transit time (μs) |
| 400 | 48.275–48.276 | 48.275–48.276 | 48.628–48.629 | 48.628–48.629 |
| 300 | 48.281–48.282 | 48.281–48.282 | 49.633–49.634 | 49.633–49.634 |
| 200 | 48.286–48.287 | 48.286–48.287 | 50.631–50.633 | 50.631–50.633 |
| <150 | Low jitter, measurable | | Large jitter, more than 2 μs | |

The results show that the transit times measured by TTUF-FWR basically keep the same regardless the amplitude changes of the received signal, but the ones measured by TTUF-CT produce about 2 μs errors. When the amplitude of the received signal is lower than 150 mV, a data jitter occurs in the both flowmeters, which is smaller and measurable for the TTUF-FWR and greater than 2 μs for the TTUF-CT which cannot work normally. It can be seen that the TTUF-FWR has higher precision and stability of measurement than the TTUF-CT when the amplitude of the received signal changes.

### 3.3.2 Flow Rate Measurement Experiment on Flow Test Platform in Real Working Condition

In order to comparing the practical effect of flow rate measurement, this experiment is conducted on the flow test platform of Dalian Hui Ming Instrument Co., Ltd. Figure 6 shows the experimental layout. The excitation pulse voltage of the transducers with 1 MHz resonance frequency is 3.0 V, and the threshold value of comparator is set to 30 mV. The water in the pipe is added amount of sediment in order to reflect the practical situation that the bubbles and silt mixed in the water will cause the amplitude change of the received signal when the pump runs. The photograph of the experiment is shown in Fig. 7.

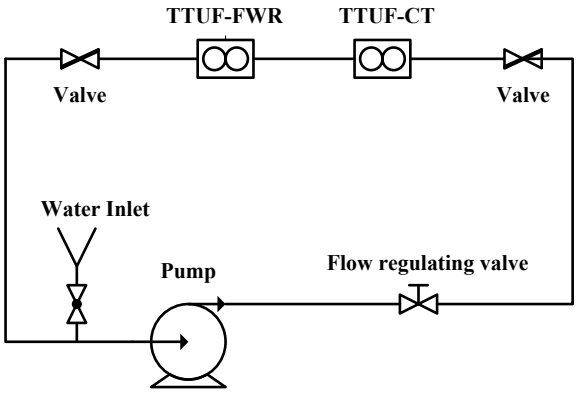



**Fig. 6.** Layout of experimental devices

The flow rate measurement experiments are conducted at three different flow rates of 1.5, 15, and 45 $m^3$/h, and the measurements are done 10 times every 5 s at each flow rate. The experimental results including the repeatability calculated are shown in Table 2.

Table 2 shows that the measurement repeatability's of TTUF-FWR are all better than TTUF-CT at every flow rate, which indicates the more effectiveness of the feature wave recognition method. Because the accurate flow rate cannot be calibrated in this experiment, the minimum flow rate of 10 sets of measured data at 45 $m^3$/h is, respectively, selected as the standard flow rates to calculate the measurement errors of the two flowmeters as shown in Table 3.

**Fig. 7.** Experimental photograph

**Table 2.** Measurement of flow rate

| Group number | TTUF-FWR | | | TTUF-CT | | |
|---|---|---|---|---|---|---|
| | Flow rate 1.5 $m^3/h$ | Flow rate 15 $m^3/h$ | Flow rate 45 $m^3/h$ | Flow rate 1.5 $m^3/h$ | Flow rate 15 $m^3/h$ | Flow rate 45 $m^3/h$ |
| 1 | 1.523 | 15.619 | 45.561 | 1.492 | 15.961 | 44.784 |
| 2 | 1.562 | 15.895 | 46.024 | 1.593 | 16.334 | 45.436 |
| 3 | 1.422 | 15.724 | 46.352 | 1.651 | 15.805 | 45.825 |
| 4 | 1.475 | 15.636 | 46.684 | 1.454 | 15.529 | 46.391 |
| 5 | 1.395 | 15.525 | 46.236 | 1.328 | 15.035 | 46.632 |
| 6 | 1.613 | 15.149 | 45.853 | 1.466 | 15.297 | 45.457 |
| 7 | 1.453 | 15.266 | 45.352 | 1.354 | 15.536 | 45.203 |
| 8 | 1.537 | 15.749 | 45.213 | 1.473 | 15.154 | 44.565 |
| 9 | 1.569 | 15.817 | 45.942 | 1.629 | 14.897 | 44.239 |
| 10 | 1.637 | 15.725 | 46.204 | 1.626 | 14.662 | 44.354 |
| Repeatability | 0.081 | 0.238 | 0.461 | 0.115 | 0.550 | 0.825 |

According to Table 3, the maximum measurement error of TTUF-FWR is 3.253%, whereas the one of TTUF-CT is 5.409%. Although the measurement error mentioned above cannot accurately reflect the real measurement error, it proves that the feature wave recognition method has a better performance in accuracy.

The two comparing experiments show that the TTUF-FWR is superior to the TTUR-CT in the accuracy of measurement, the repeatability, the stability, and the measurement precision because of decreasing the measurement error of transit time caused by the amplitude changed of the received signal.

**Table 3.** Measurement error statistics

| Group number | TTUF-FWR | | TTUF-CT | |
|---|---|---|---|---|
| | Flow rate 45 $m^3/h$ | Measurement error (%) | Flow rate 45 $m^3/h$ | Measurement error (%) |
| 1 | 45.561 | 0.770 | 44.784 | 1.232 |
| 2 | 46.024 | 1.794 | 45.436 | 2.706 |
| 3 | 46.352 | 2.519 | 45.825 | 3.585 |
| 4 | 46.684 | 3.253 | 46.391 | 4.864 |
| 5 | 46.236 | 2.263 | 46.632 | 5.409 |
| 6 | 45.853 | 1.416 | 45.457 | 2.753 |
| 7 | 45.352 | 0.307 | 45.203 | 2.179 |
| 8 | 45.213 | 0 | 44.565 | 0.737 |
| 9 | 45.942 | 1.612 | 44.239 | 0 |
| 10 | 46.204 | 2.192 | 44.354 | 0.260 |

## 4 Conclusions

Aiming at the measurement precision and stability of TTUF, a feature wave recognition-based signal processing method is proposed in this study. By measuring the period of feature wave which has a specific position in the received signal, the datum point of ultrasonic received signal is accurately located. The zero-crossing detection error of threshold comparator caused by the amplitude change of the received signal is effectively reduced so that the measurement accuracy of transit time is further improved. The following experimental study demonstrates that the proposed method performs better on both precision and stability comparing with the conventional approach, which has a high feasibility and practical value for real application.

## References

1. Yu Y, Zong GH, Ding FL. Comparison of flow rate calculation method for ultrasonic flow measurement. J Beijing Univ Aeronaut Astronaut. 2013;39(1):37–41.
2. Svilainis L, Dumbrava V. The time-of-flight estimation accuracy versus digitization parameters. Ultrasound. 2008;63(1):12–7.
3. Wanderson ES, Edson DCB. Development and signal processing of ultrasonic flowmeters based on transit time. In: IEEE international conference on industry applications; 2016. p. 1–7.
4. Wang XF. Research on the key technologies of transit time ultrasonic gas flowmeters. Doctoral Dissertation. Dalian: Dalian University of Technology; 2011.
5. Rajita G, Nirupama M. Review on transit time ultrasonic flowmeter. In: 2016 2nd international conference on control, instrumentation, energy & communication; 2016. p. 88–92.
6. Meng H, Wang H, Li MW. High-precision flow measurement for an ultrasonic transit time flowmeter. In: International conference on intelligent system design and engineering application; 2010. P. 823–6.

7. Zhao WG, Jiang YF, Huang CC. A new ultrasonic flowmeter with low power consumption for small pipeline applications. In: IEEE international instrumentation and measurement technology conference; 2016. p. 1–6.
8. Wang W, Xu KJ, Fang M, Zhu WJ, Shen ZW, Wang G, Wang B. Study of a signal processing method for gas ultrasonic flowmeter. J Electron Measur Instrum. 2015;29 (9):1365–73.
9. Chen J, Yu SS, Li B, Fan CY. Signal processing based on dual-threshold of ultrasonic flow meter. J Electron Measur Instrum. 2013;27(11):1024–33.
10. Gu JH, Ye XQ. Foundation of underwater acoustics. Beijing: Defence Industry Press; 1981.

# Realization of Unmanned Cruise Boat for Water Quality

Zhongxing Huo(✉), Yongjie Yang, and Yuelan Ji

School of Electronics and Information, Nantong University, Nantong 226019, China
1254588282@qq.com, yang.yj@ntu.edu.cn

**Abstract.** In order to solve the problems of difficult wiring, poor flexibility, and high cost in aquaculture water quality monitoring, an unmanned water quality monitoring cruise ship with water quality monitoring device was built. The ship navigated automatically on the surface of the water according to the set course and collected water quality data during the voyage, which saved a lot of resources in this innovative way. There are two modes of operation for cruise ships: manual mode and autopilot mode. In the manual mode, the user can realize the manual operation of the ship through the remote controller and can set the autopilot path of the cruise ship. In the autopilot mode, the cruise ship moves automatically according to the preset path.

**Keywords:** Cruise boat · Water quality detection · Data acquisition · Automatic navigation

## 1 Introduction

Aquaculture water is the environment that fishes depend on for survival, so the good and bad culture environment is directly related to the growth and development of cultured fish, which determines the yield and quality of aquatic products [1–3]. The normal online water quality monitoring system, which usually replaces manual water quality monitoring system, adopts the scheme of arranging a number of nodes on the surface of the water. The system has the advantages of good real-time performance and simple operation, but the price of each set of sensors is very expensive, which is beyond the economic affordability of some farmers.

This paper designed an intelligent cruise boat for water quality monitoring, which only needs to carry a set of water quality detection sensors, saving a lot of equipment costs. At present, most of the researches on the water quality monitoring boat are manually controlled by the shore personnel, so it is impossible for the boat to run automatically for a long time.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_122

## 2 Overall Design of System

The system block diagram of the scheme is shown in Fig. 1. The cruise boat is connected to the satellite positioning system, an electronic compass, the remote controller, and the PC. The satellite positioning system is used to obtain the longitude and latitude of the position information of cruise boat, and the electronic compass is used to get boat's course.

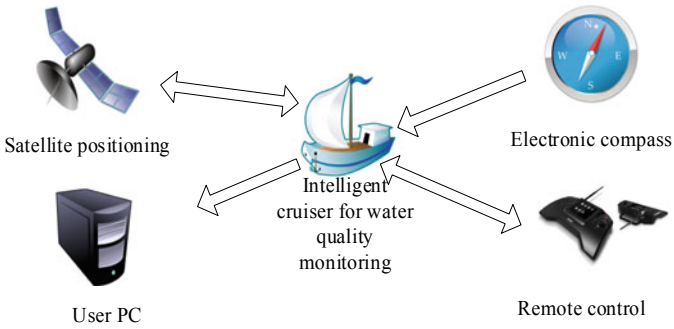



**Fig. 1.** Intelligent cruise boat system for water quality monitoring

The remote controller is used to communicate with the cruise boat. Through it, users can manually control the cruiser to move. They can also set up automatic navigation routes for cruisers, which need to avoid obstacles on the surface of the water. After completion, the remote control is required to set up the cruise boat into autopilot mode, and then the cruise boat continuously collects data on water quality parameters during the voyage. During autopilot, the user triggers a forward, left, right, or backward operation, the cruise boat exits autopilot mode.

## 3 Hardware Design of Cruise Boat for Water Quality Monitoring

In this paper, the circuit is divided into two parts: boat cruise and remote control.

### 3.1 Hardware Design of Cruise Boat

In order to realize intelligent cruising, the cruise boat carries the Beidou module, the electronic compass module, the steering gear, the motor drive circuit, the 433 MHz communication module, and the display module. The circuit diagram is shown in Fig. 2 [4]. The cruiser adopts a 32-bit STM32F407 single-chip microcomputer with high performance and high-performance price ratio. The frequency of operation is as high as 72 MHz, and the code execution efficiency is high.

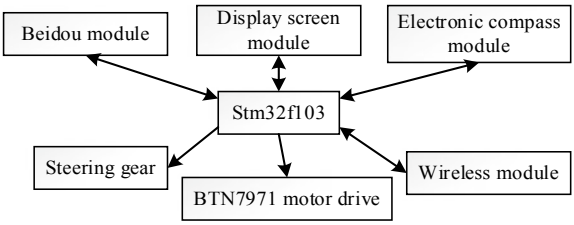



**Fig. 2.** Hardware block diagram of the front-end device

The geographic position of the cruise boat is indicated by latitude and longitude, which is obtained by the Beidou module UM220-III [5]. The positioning accuracy of this module is about 2.5 m CEP. The module uses serial port to send data to STM32.

The electronic compass adopts the GY-273 module, the measuring range is ±1.3–8 Gauss, and the MCU can communicate with the module through the IIC protocol to obtain the angle between the prow of cruising boat and the geomagnetic North Pole. Combined with the output information of the above two modules, the motion state of the cruise boat can be determined.

The wireless module uses the SX1278 module under the 433 MHz frequency to communicate with the wireless remote control provided by the cruise boat [6].

The OLED carried on the cruise boat is used to display the boat's information, mainly including the current longitude, latitude, heading angle and so on [7].

### 3.2 The Hardware Design of Remote Controller

The remote controller is used to interact with the user and the interactive interface is displayed on the display screen. Finally, the wireless module under the frequency of 433 MHz transmits the data to the cruise boat. The hardware block diagram is shown in Fig. 3.

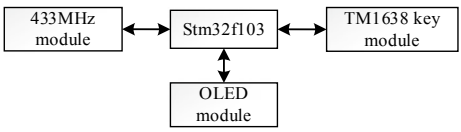



**Fig. 3.** Hardware block diagram of remote controller

The key module uses TM1638 chip to reduce the input/output occupation of microprocessor and simplify the complexity of program. In the design of the remote controller, the keyboard is used to receive user information and LED displays the state of the button pressed by the user. The digital communication of the module is designed by serial port, which only occupies three input/output ports of microprocessor, and the circuit connection is simple.

# 4 Software Design of Intelligent Cruise Boat for Water Quality Monitoring

## 4.1 Programming of Remote Control

The function realization of cruise boat needs to cooperate with the remote control, shown in Fig. 4, which needs to issue manual command, set path command, and autopilot command to cruise boat. The program first initializes the required hardware by using the function Board_Init(), including the 433 MHz wireless module, the key module and the display module. The Input_deal() function then enters to detect the input state of the user, assembles different data packets according to the different keys pressed, sends them out through the Send_packet() function, and displays the operation carried out by the current user on the display screen.

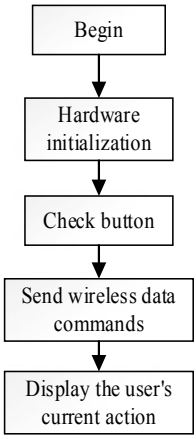



**Fig. 4.** Flowchart of remote control software

The remote control command packet format includes the following: “forward” button corresponding to “forward” is the cruise boat’s advance command; “left turn” button corresponding to “left” is the cruise boat assembly command; “right turn” button corresponding to “right” for cruiser right-turn command; the “back” button corresponds to “backward” for the cruise boat’s back command. These are manual driving orders. The “setting place” button corresponds to “Set1” as the first point command to set the path, and the “clear all” button corresponds to “clear” to clear the set location command, and the “autopilot” button corresponds to the “auto” as the autopilot command. When the user sets the path, the current set point mark is displayed back on the display screen, and the settings are automatically added 1. The schematic diagram of the remote control keys is shown in Fig. 5.

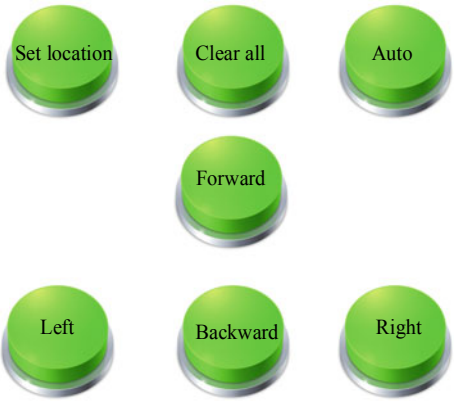



**Fig. 5.** Schematic diagram of remote control keys

### 4.2 Design of Autopilot Software for Cruise Boat

The embedded program is written by KEIL software and C language, and the flow of autopilot software is shown in Fig. 6 [8–10]. After the software is started, the wireless module, the Beidou module, the electronic compass module, the display screen module, the steering gear, the motor drive circuit and the stm32 internal peripherals are initialized by Board_Init() function.

The wireless module receives command data from the cruise boat's remote controller, which is divided into three types: the "manual pilot" command, the "set the path" command, and the "autopilot" command. The Get_lc(&hello) function performs different programs according to different commands, and if it is a "manual pilot" command, the cruiser can move back and forth from side to side.

If it is the "set path" command, the program first reads the coordinate data of the cruiser's location, then the stm32flash_write() function records it to the inner flash of the STM32 and displays the coordinate point label on the display screen.

If the "autopilot" command, then immediately enter autopilot mode, the program first read the electronic compass and Beidou satellite data, and then determine whether the cruiser's current coordinates are near the coordinates of the destination. According to the precision of the Beidou module, if the distance is less than 2.5 m, it means to reach the destination. If the cruise boat reaches its destination, the next stored coordinate is used as the destination. Then, according to the above information, the angle of the cruise boat need to be adjusted is calculated. Finally, the calculation results are input into the PID algorithm, and the output of the PID algorithm is used to control the cruiser movement [11–13]. Through the continuous cycle of the above process, the trajectory of the cruise boat is continuously adjusted, and then the cruise boat intelligent cruise is realized.

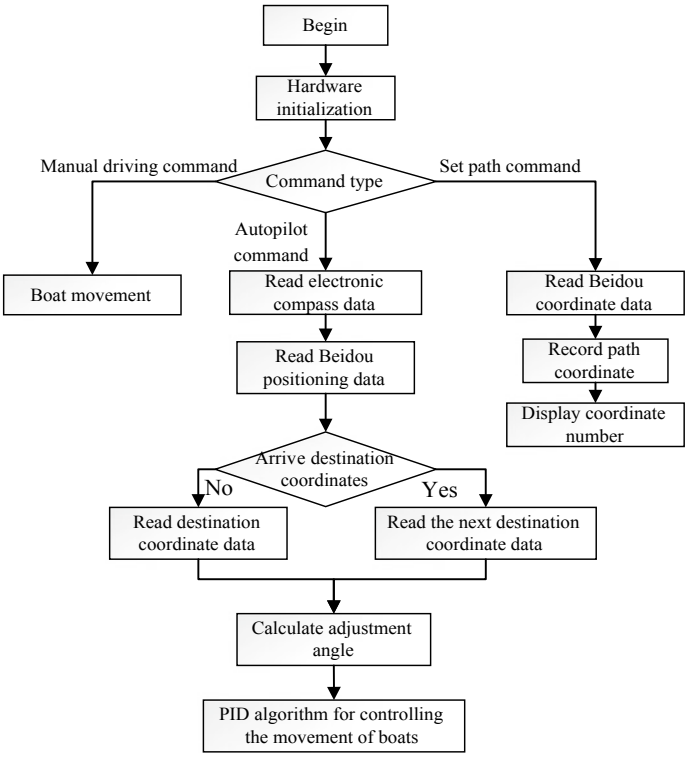



**Fig. 6.** Flowchart of cruise boat autopilot software

### 4.3 Software Design for Environmental Information Collection

When a cruise boat controlled by autopilot software is navigating on the surface, the data of water quality parameters collected by the cruise boat will be continuously uploaded, and the software flowchart is shown in Fig. 7. The program first initializes the 485 interface and the GPRS module by using Board_Init(), in which the 485 interface uses the baud rate of the 9600 bps, and each device has a unique address for each device on the 485 bus, according to which each sensor can be distinguished [14]. And then the program reads the data from each sensor. After initialization, when the cruiser is sailing on the water, the program sends the collected data to the user's PC through Gprs_send() function.

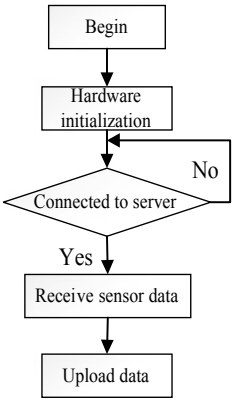



**Fig. 7.** Flowchart of environment information acquisition software

## 5 System Test

The test uses three target points to form a triangle, and each point is 10 m apart. After the target point is set, the cruiser is controlled into autopilot mode by the remote controller. The actual path is shown in Fig. 8.


**Fig. 8.** Schematic diagram of cruise boat operation path

In the course of the experiment, the distance of each cruiser from the target point was recorded four times, and the data was sent to the user PC by using the GPRS of the data acquisition part. The result is shown in Table 1.

The cruiser can reach every target point correctly, and the distance error is less than 2.5 m. This precision can meet the requirements on the surface of the water, can be used, and has a good application prospect.

**Table 1.** Cruise boat navigation record (unit: m)

| Serial number | Distance from point 1 | Distance from point 2 | Distance from point 3 |
|---|---|---|---|
| 1 | 1.0 | 1.6 | 1.3 |
| 2 | 1.2 | 2.1 | 1.3 |
| 3 | 1.9 | 0.8 | 1.1 |
| 4 | 1.4 | 0.8 | 1.9 |

## 6 Conclusions

In this paper, a kind of intelligent cruise boat with water quality parameters is realized, and the water quality parameters information of the whole water area is collected in real time by a novel "multi-point one boat" method, which saves a lot of material resources and can meet the needs of aquaculture for water quality measurement.

**Acknowledgments.** This work was supported by Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX17-1921) and First phase project of Jiangsu University Brand Specialty Construction Project (PPZY2015B135). In addition, it was completed under the support of Nantong University-Nantong Intelligent Information Technology Joint Research Center Open Topic (KFKT2017B05).

## References

1. Shah MR, Lutzu GA, Alam A, et al. Microalgae in aquafeeds for a sustainable aquaculture industry. J Appl Phycol. 2017;1:1–17.
2. Mo WY, Man YB, Wong MH. Use of food waste, fish waste and food processing waste for China's aquaculture industry: needs and challenge. Sci Total Environ. 2017;635:613–4.
3. Tai H, Ding Q, Li D, et al. Design of an intelligent PH sensor for aquaculture industry. IFIP Adv Inf Commun Technol. 2017;347:642–9.
4. Yang B. Design and implementation of intelligent home wireless gateway based on STM32. In: International conference on information science and control engineering. IEEE computer society; 2017. p. 258–60.
5. Goncharova I, Lindenmeier S. Compact satellite antenna module for GPS, Galileo, GLONASS, BeiDou and SDARS in automotive application. IET Microwaves Antennas Propag. 2018;12(4):445–51.
6. Zhou L, Sun S, Zhang Y, et al. Long-distance running test system based on 433 MHz wireless module. In: IEEE international conference on communication technology. IEEE; 2016. p. 339–43.
7. Kurban M, Gündüz B. Physical and optical properties of DCJTB dye for OLED display applications: experimental and theoretical investigation. J Mol Struct. 2017;1137:403–11.
8. Mustakerov I, Borissova D. A framework for development of e-learning system for computer programming: application in the C programming language. J e-Learn Knowl Soc. 2017;13(2):89–101.
9. Kim JH, Whang IH. Augmented three-loop autopilot structure based on mixed-sensitivity $H_\infty$ optimization. J Guidance Control Dyn. 2017;4:1–6.

10. Graham DM. An AutoPilot platform for high-resolution light-sheet microscopy. Lab Anim. 2017;46(2):25.
11. Kong H, Fang Y. Neural network PID algorithm for a class of discrete-time nonlinear systems. Int J Online Eng. 2018;14(2):103.
12. Dideriksen JL, Feeney DF, Almuklass AM, et al. Control of force during rapid visuomotor force-matching tasks can be described by discrete time PID control algorithms. Exp Brain Res. 2017;235(2):1–13.
13. Pradhan PC, Sahu RK, Panda S. Firefly algorithm optimized fuzzy PID controller for AGC of multi-area multi-source power systems with UPFC and SMES. Eng Sci Technol Int J. 2016;19(1):338–54.
14. Xie Y, Yu M, Fu J, et al. A hazmat transportation monitoring system based on global positioning system/beidou navigation satellite system and RS485 bus. In: International congress on image and signal processing, biomedical engineering and informatics. IEEE; 2017. p. 1059–63.

# Improved K-Means Clustering for Target Activity Regular Pattern Extraction with Big Data Mining

Guo Yan[1(✉)], Lu Yaobin[1], Ning Lijiang[1], and Wang Jing[2]

[1] Nanjing Institute of Electronic Technology, Guorui Road 8#, Nanjing, China
9821078@qq.com
[2] Troop, PLA, Beijing 66132, China

**Abstract.** The traditional target activity regular pattern extraction methods replay previous target tracks, activities of the specified target are manually analyzed by checking all the tracks on map. This paper adopts big data mining technology to solve the problem of automatically extracting target classic tracks and converts the original pure manual map analysis into system automatic track extraction. This method greatly reduces the operation intervention of classic track extraction, which can reduce the 3–4 manual days to 3–4 h.

**Keywords:** Big data mining · K-means clustering · Target activity regular pattern

## 1 Introduction

The traditional method for target activity regular pattern extraction analyzes the activity of a specified target in a manual manner by replaying previous trajectory data. It is necessary to manually analysis the target trajectory according to time, region, country, mission to form a classic trajectory, and then sum up the target activity pattern. Therefore, there is a lack of effective methods for automatic extraction of target activity patterns.

Big data mining technology [1] uses non-traditional machine learning tools to process massive amounts of structured and unstructured data, to extract data relationships that are unknown, such as group data analysis (cluster analysis), unusual data monitoring (abnormal monitoring), and relationship mining. Using big data mining technology, intelligent processing similar to human intelligence can be performed on the previous target trajectory data to extract target activity patterns.

Target activity regular pattern extraction with big data mining technology, whose input time, region, and other conditions can be specified by requirements, utilize the K-means clustering previous big target trajectory mining algorithm to extract classic tracks, utilize geographic grid to merge the classic tracks into classic track patterns. This technology solves the problem of automatically extracting interested target from big amount of track data and converts the classic manually selection method to automatic extraction with little human participation. This change for track regular pattern extraction is fundamental, which greatly reduces the workload of operator.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_123

## 2 Extraction of Target Activity Patterns Based on Big Data Mining

### 2.1 Target Classic Track Data Mining Method

This method utilizes big data mining technology to analyze the intrinsic correlation among the large amount of target track data. According to the machine number, time, area information, machine learning algorithms, such as data cleaning, data clustering are used to excavate track pattern, then the classic tracks are shown to combat personnel for combat target rapid analysis. The flowchart for target classic track data mining method is shown in Fig. 1, including data cleaning, track extraction, track clustering, track merging, classic track generation, and classic track display. The core of this flowchart is to track clustering.

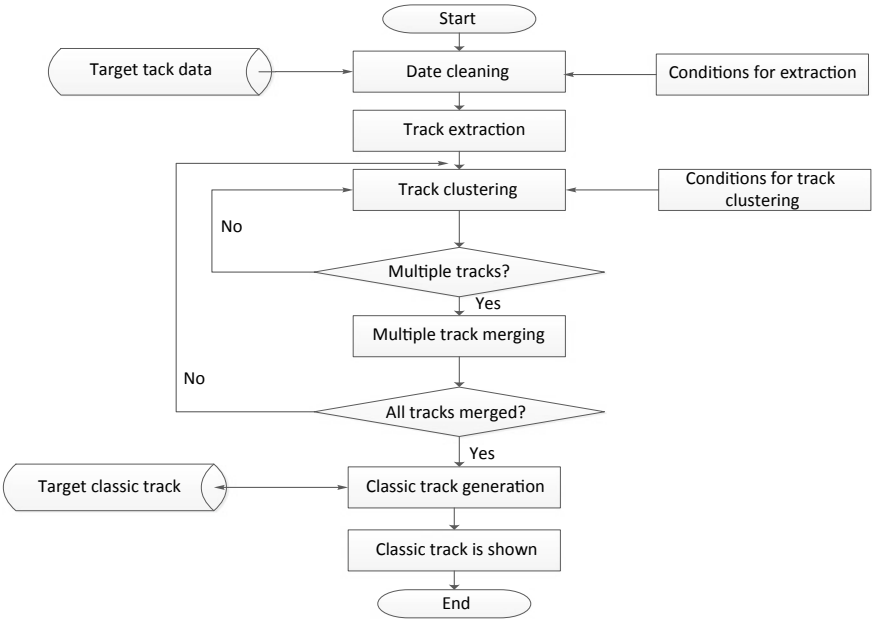



**Fig. 1.** Target activity pattern extraction with big data mining

(1) Data cleaning

The data for data mining is complex track data, which cover planes, ships, missiles, and satellites. The data sources include ground-based air defense radars, airborne radars, shipborne radars, and space-based radars, the sensor types include passive radar and optical equipment. Due to the different data sources, detection mechanisms, transmission mechanisms, the consistency and validity of the data should be considered. Data cleaning is necessary for data anomalies eliminating, normalization processing, and data formatting.

The track data is stored in the big data platform with near-line relational database and off-line distributed file system, where the track data within three months is stored in database, and the track data more than three months is stored with specified format in file system. During data cleaning, the data range is determined according to the conditions of track extraction, then the related track data is read into memory of big data platform, and the track data is cleaned.

(2) Track extraction

Using human–computer interaction window, the track extraction conditions, such as type of plane, start and stop time of specified target, area range (up-left and down-right corner of rectangular area, custom polygon area), country information, are properly set.

When the track is extracted, the big data platform acquires data components on demand. According to the different track extraction conditions set by warfighters, the related track data is parallel and quickly extracted into big data platform, and sorted by types, start and stop time, area range, and country.

(3) Track clustering

Track clustering is the core of the target activity pattern with big data mining. It mainly refines the analysis and processing of the large amount of unordered track data and classifies the different tracks with different shapes.

In this process, Euclidean position distance constraint [2] is improved as multiple track feature distance constraint for K-means clustering [3]. Euclidean position distance constraint performs the nearest neighbor calculation based on Euclidean distance of track point position. The computational complexity is low, and it is suitable for large-scale track clustering. Multiple track feature distance constraint performs the track feature Hausdorff distance [4], which includes track point position, velocity, attributes, course, and so on. The multidimensional features are used, and the clustering accuracy is better than the classic Euclidean position constraint.

The steps of K-means clustering algorithm is as follow:

| Input: track data $C_1, C_2, \ldots, C_n$, clustering initial value $k$<br>Output: classic track clustering results $x_1, x_2, \ldots, x_k$ | |
|---|---|
| Algorithm steps | Step 1: Initialize $k$ classic track seeds for track clustering. Extract $n$ track data $C_1, C_2, \ldots, C_n$ from big data memory database, select random $k$ tracks as the initial clustering seeds of the classic track clustering<br>Step 2: Distribute track point data samples to the nearest center vector and construct disjoint clusters. For the remaining objects, according to their similarity with these cluster centers, they are assigned to their most similar clusters: $\sum_{i=1}^{n} \min_{j\in\{1,2,\ldots,k\}} (x_i - p_i)^2$<br>Step 3: Use the sample mean in each cluster as the new cluster center (average of all objects in the cluster): $x = \frac{1}{\lvert C_i \rvert} \sum_{x\in C_i}^{n} x$<br>Step 4: Repeat steps 2 and 3 until the cluster center no longer changes. Repeating this process until the standard measure function $E$ to start to |

(*continued*)

(*continued*)

| Input: track data $C_1, C_2, \ldots, C_n$, clustering initial value $k$<br>Output: classic track clustering results $x_1, x_2, \ldots, x_k$ | |
|---|---|
| | converge, and then get $k$ clusters, making the clusters themselves as compact as possible, and separating the clusters as much as possible. The standard deviation function is used here as the standard measure function ad follow: $E = \sum_{i=1}^{k} \sum_{x \in C_i} \lvert x - x_i \rvert^2$ |

The above K-means track clustering method based on Euclidean distance can achieve rapid clustering processing with massive target track data, but the accuracy of clustering is limited. In order to improve the clustering accuracy, K-means clustering algorithm based on multidimensional track features can be considered in this process. The algorithm steps are the same as the above method, but the distance calculation is calculated using Hausdorff as follows:

$$\delta_H(\mathrm{TR_A}, \mathrm{TR_B}) = \max_{P_\mathrm{A} \in \mathrm{TR_A}} \left\{ \min_{P_\mathrm{A} \in \mathrm{TR_B}} \{\mathrm{dist}(P_\mathrm{A}, P_\mathrm{B})\} \right\} \tag{1}$$

In the above equation, $\mathrm{dist}(P_\mathrm{A}, P_\mathrm{B})$ is the Euclidean distance considering the position between the two points, $\delta_H(\mathrm{TR_A}, \mathrm{TR_B})$ is the maximum distance between $\mathrm{TR_A}$ and $\mathrm{TR_B}$. Where from the middle point to the middle point and represents the degree of similarity with. Where $\mathrm{TR}_i$ is multidimensional track point sequence as follows:

$$\mathrm{TR}_i = \left\{ P_{i1}, P_{i2}, \ldots, P_{ij}, \ldots, P_{im} \right\} \tag{2}$$

$P_{ij}$ is multidimensional track point vector as follow:

$$P_{ij} = [\,\mathrm{label} \quad \mathrm{longitude} \quad \mathrm{latitude} \quad \mathrm{altitude} \quad \mathrm{attribute} \quad \mathrm{velocity} \quad \ldots \quad \mathrm{course}\,] \tag{3}$$

Which contains the target number, longitude, latitude, altitude, attributes, speed, velocity, etc. of the track point $P_{ij}$.

(4) Multiple tracks merging

The results of track clustering are merged according to the type of target. For aircraft and ship targets, tracks are sorted by model, number, time, area, and country.

(5) Classic track generation

For the collection of target tracks obtained by the multiple tracks merging, the classical track is extracted using the geographic grid technology. For the target tracks belonging to the same category, the track points are divided by a suitable grid width, and multiple tracks falling within a grid are replaced by one classic track point. The connectivity of the grid connects the classic track points to form a classic track.

(6) The classic track displaying

According to the model, time, area, country, etc. that the warfighter pay attention to, the classic tracks are displayed on the map.

### 2.2 General Machine Learning Algorithm Model

Target activity regular pattern extraction needs big data platform to implement data cleaning, clustering analysis. Therefore, a general machine learning hardware and software platform is necessary.

This paper develops a universal machine learning management tool to realize multisource data training and real-time data prediction. The tool encapsulates seven types of algorithms such as classification, regression, integration, and clustering, among which anomaly detection and clustering are applied to target activity regular pattern extraction. The universal machine learning algorithm model library is shown in Table 1.

**Table 1.** Universal machine learning algorithm model library

| Algorithm classification | Algorithm name | Algorithm applicability |
|---|---|---|
| Anomaly detection | Support vector machine | Detecting unusual or abnormal data |
| Clustering | Enhanced K-means, orthogonal partition clustering | Finding natural grouping |
| Classification | Logistic regression, Bayes classification, support vector machines, decision trees | Predicting specific results |
| Regression | Ridge regression, lasso regression, elastic network | Prediction for continuous numerical results |
| Integration | Binary decision tree, gradient promotion | Random forest data |
| Feature extraction | Non-negative matrix factorization | Generate new attributes as a linear combination of existing attributes |

## 3 Verification Results

### 3.1 Lightweight Big Data Platform

To build a lightweight platform for the target activity regular pattern extraction, a big data hardware platform with blade server I9000 + disk array, core software components such as Hive, Spark, Yarn, and Zookeeper is constructed. The hardware and software configurations are shown in Table 2.

**Table 2.** Big data platform hardware and software environment

| | |
|---|---|
| Server | There are 4 servers, 3 for constructing a fleet, and 1 for building a stand-alone test environment. Blade server I9000 + disk array |
| Operation system | Linux |
| Distributed storage system | HDFS |
| Distributed computing framework | Spark computing framework |
| Distributed resource management system | YARN |
| Distributed collaborative component | Zookeeper |
| Distributed message queue | Kafka |
| Data migration tools | Sqoop, kettle |
| Data processing algorithm library | Mahout, Mlib |
| Cluster monitoring deployment tool | CM |

### 3.2 Target Classic Track Clustering Mining Results

In order to verify the validity of the target classic track extraction, the track data of an airport in 2013 was used, and the data of one day is selected for cluster mining. Using the above process, multiple tracks of similar shapes are grouped into clusters from a large number of unrelated tracks, and multiple groups of tracks of different shapes are divided into clusters. Figure 2a shows the five track clusters of interest, and Fig. 2b, c shows the two of the five tracks.

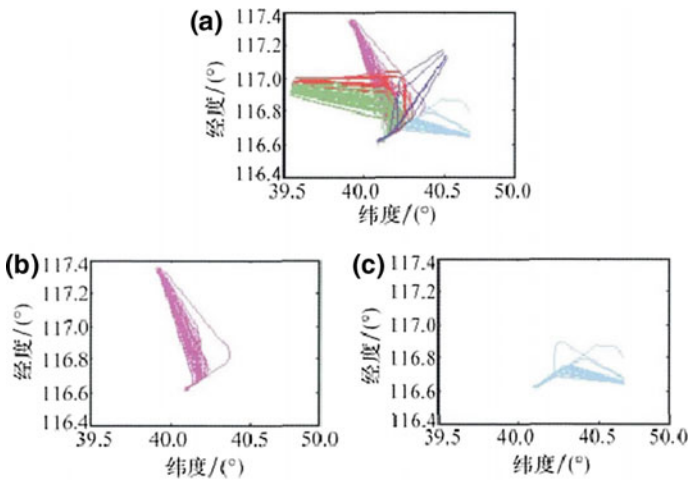



**Fig. 2.** Target classic track clustering results

## 4 Conclusion

Target activity regular pattern extraction technology, according to the time, area, and country constraint, utilizes the improved track-feature-based K-means clustering method for big data mining in large amount of previous tracks. This technology greatly reduces the operation intervention of classic track extraction, which can reduce the 3–4 manual days to 3–4 h. This technology has a wide range of application, such as radar netting system and edge sea defense system.

## References

1. Liang JY, Qian YH, Li DY, et al. Theory and method of granular computing for big data mining. China Sci Inf Sci. 2015;45(11):1355–69.
2. Xu T, Li YX, Lv ZP. Track clustering based on distance from the track point metod. Syst Eng Electron Technol. 2015;37(9):2198–204.
3. Wang ZF, Pan Q, Lang L, et al. Dynamic track clustering algorithm based on subtractive clustering. Syst Simul J. 2009;21(16):5240–3.
4. Chen H, Zhang BY, Chen Y. Research on multi-hypothesis tracking with adaptive depth adaptation. Syst Eng Electron Technol. 2016;38(9):2000–7.

# PSO-RBF Small Target Detection in Sea Clutter Background

ZhuXi Li, ZhenDong Yin(✉), and Jia Shi

School of Electronics and Information Engineering, Harbing Institute of Technology, Harbin 150006, China
51280684@qq.com, yinzhendong@hit.edu.cn, 13182169819@163.com

**Abstract.** Target detection in the background of sea clutter is an important part of sea surface radar signal processing. The traditional detection of weak targets in sea clutter is based on the statistical characteristics of sea clutter, which does not reflect the intrinsic dynamics of sea clutter. Therefore, the detection results are not ideal. Based on the chaotic characteristics of sea clutter, this dissertation reconstructs the space structure of the sea clutter and proposes an improved particle swarm optimization (PSO) algorithm based on adaptive time-varying weights and local search operators. This method was applied to the optimization learning of the parameters of the radial basis function (RBF) neural network kernel function. The method was validated by using McIX University in Canada to measure the sea clutter data with the target in the Dartmouth area using IPIX radar. The results showed that the PSO-RBF algorithm in the background of chaotic sea clutter has good predictability. Compared with the general radial basis neural network, the improved algorithm not only has fast convergence speed but also has high error accuracy.

**Keywords:** Particle swarm optimization (PSO) · Radial basis function (RBF) · Target detection · Neural network · Sea clutter

## 1 Introduction

In recent years, the small target detection technology in the sea has attracted more and more attention. Sea clutter is an echo signal that radar shines on the sea surface and is very vulnerable to external environmental factors such as waves, wind speed, and wind power. Since sea clutter can seriously affect the detection of small targets on the sea (including low-altitude flying planes, small warships, floating ice on the sea, navigation buoys, etc.), the study of sea clutter characteristics is of great significance.

The research of sea clutter based on chaos theory can reflect the inherent nonlinear dynamics of sea clutter [2]. The introduction of nonlinear theory and neural network technology has improved the detection performance of small sea targets. Many scholars have done a lot of related works. Optimization of RBFNN center points, variances, and number of hidden layer nodes by genetic algorithm [3]. An algorithm based on dual-constraint least square support vector machine is introduced to detect the weak signal in the background of chaotic sea clutter [4]. Based on the study of normalized RBF neural


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_124

network and least squares support vector machine (LSSVM) nonlinear prediction of sea clutter, a prediction method of LSSVM-coupled image lattice is proposed [5].

In this dissertation, the sea clutter is reconstructed by phase space, and an improved PSO algorithm based on adaptive time-varying weights and local search operators is proposed. This algorithm is applied to the optimal learning of kernel function parameters in RBF neural network. Algorithm flowchart is shown in Fig. 1. The simulation results show that the PSO-RBF algorithm has good predictability in the chaotic sea clutter background. Compared with the general radial basis neural network, the improved algorithm not only has fast convergence speed but also has high error precision.

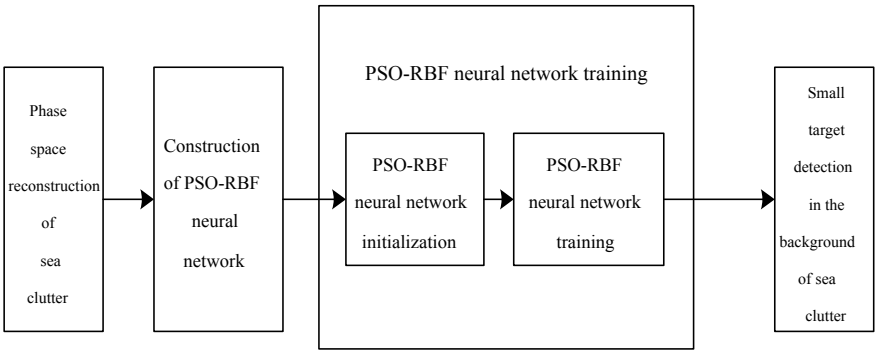



**Fig. 1.** PSO-RBF target detection flowchart

## 2 Phase-Space Reconstruction Model

Phase space is a multidimensional space that can reflect the inherent dynamic structural characteristics of the system. For a nonlinear system, the internal structure is usually more complex, and the phase space has a higher dimension. Therefore, it is difficult to determine the dimension. In practical applications, for a given one-dimensional time series, it is usually extended to a three-dimensional or even higher-dimensional space in order to extract the information contained in the time series. This process is called phase-space reconstruction. Its essence is to reconstruct a one-dimensional time series into an $m$-dimensional phase-space vector:

$$y(t+m\tau) = F(y(t), y(t+\tau), \ldots, y(t+(m-1)\tau)) \tag{1}$$

where $m$ is the embedding dimension, $\tau$ is the delay time, and $F$ is the nonlinear function. Phase-space reconstruction's main steps are as follows:

Step 1: Calculate the embedding dimension $m$. The GP algorithm is a simple and relatively mature method for calculating the embedding dimension.

Step 2: Calculate the delay time $\tau$. The autocorrelation function method is a commonly used calculation method that is relatively simple and mature.

## 3 PSO-RBF Neural Network Target Detection Model

### 3.1 The Basic Principle of Particle Swarm Optimization

Particle swarm optimization is a simulation of bird flocking behavior. The global search is achieved through the collective cooperation and competition between the particles. All particles have a fitness determined by the function being optimized, and a speed determines the direction and distance of their movement. During the iteration of the PSO algorithm, the particle updates its speed and position by tracking two extreme values (individual extremum $p_{\text{best}}$ and global extremum $g_{\text{best}}$). When these two optimal values are found, the particles update their speed and position according to the following formula:

$$v_{id}^{k+1} = v_{id}^{k} + c_1 \text{rand}() \left(p_i^k - x_{id}^k\right) + c_2 \text{rand}() \left(p_{gd}^k - x_{id}^k\right) \tag{2}$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \tag{3}$$

where $c_1$ and $c_2$ are learning factors; $v_{id}$ is the speed of the particle, $v_{id} \in [-v_{\max}, v_{\max}]$, $v_{\max}$ is a constant, set by the user to limit the speed of the particle; rand() is a random number between [0, 1].

The improved PSO algorithm proposed in this paper: In iterative formula (2) adding self-knowing time-varying weights $w_i$, and in iterative formula (3) adding $\alpha$. Its mathematical description is as follows:

$$\begin{aligned} v_{id}^{k+1} &= v_{id}^{k} + c_1 \text{rand}() \left(p_i^k - x_{id}^k\right) + c_2 \text{rand}() \left(p_{gd}^k - x_{id}^k\right) \\ &= \left\{ w_{\text{end}} + (w_{\text{start}} - w_{\text{end}}) \exp\left[ -15 \left( \frac{k}{k_{\max}} \right)^3 \right] \right\} v_{id}^k \\ &\quad + c_1 \text{rand}() \left(p_i^k - x_{id}^k\right) + c_2 \text{rand}() \left(p_{gd}^k - x_{id}^k\right) \end{aligned} \tag{4}$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \alpha = x_{id}^{k} + v_{id}^{k+1} \text{rand}() \theta \tag{5}$$

where $w_i$ is in the range of $(w_{\min}, w_{\max})$; $k_{\max}$ is the maximum number of iterations set by the algorithm; rand() is a random number between $[0, 1]$; $\theta = [\text{rand}() + 0.5]$ (where [] is a rounding operation).

### 3.2 RBF Neural Network

RBF neural network is a kind of feed-forward back-propagation network constructed based on function approximation theory. It has strong ability to classify and approximate arbitrary nonlinear continuous functions. The structure of RBF neural network is shown in Fig. 2.

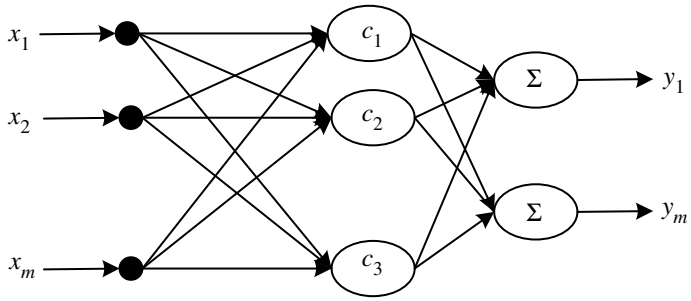



**Fig. 2.** RBF neural network structure

### 3.3 PSO-RBF Algorithm Flowchart

PSO-RBF algorithm flowchart is shown in Fig. 3.

## 4 Simulation

### 4.1 Simulated Sea Clutter Prediction

Assuming that in the case of a turbid background signal, the grazing echoes for the presence of sea clutter and noise are described by the following mathematical model:

$$\begin{cases} H_0 : x(n) = c(n), & 1 \le n \le N_T \\ H_1 : x(n) = s(n) + c(n), & 1 \le n \le N_T \end{cases} \tag{6}$$

where $x(n)$ is the observation time series, $s(n)$ is the target signal, $c(n)$ is the chaotic background signal, $N_T$ is the total number of points of the observation sequence signal, $H_0$ represents the signal that does not contain the target, and $H_1$ represents the signal that has the target.

This paper uses Lorentz turbidity time series to simulate sea clutter. System parameters $\sigma = 16, b = 4, r = 45.92$. The initial value $x = -1, y = 0, z = 1$. And using the 4th-order Runge–Kutta method with an integration step of 0.01 to solve the Lorenz equation.

Set a single periodic target signal $s(n)$ as:

$$s(n) = 0.01 \sin(2\tau f), \quad f = 0.3 \tag{7}$$

Figure 4 shows the spectrum of a single periodic signal. SNR $= -75.98$ dB. Phase-space reconstruction of the mixed time series, using PSO-RBF neural network to predict the mixed time series with rectangular target signal. Forecast results are as shown in Fig. 5, and forecast error is as shown in Fig. 6.

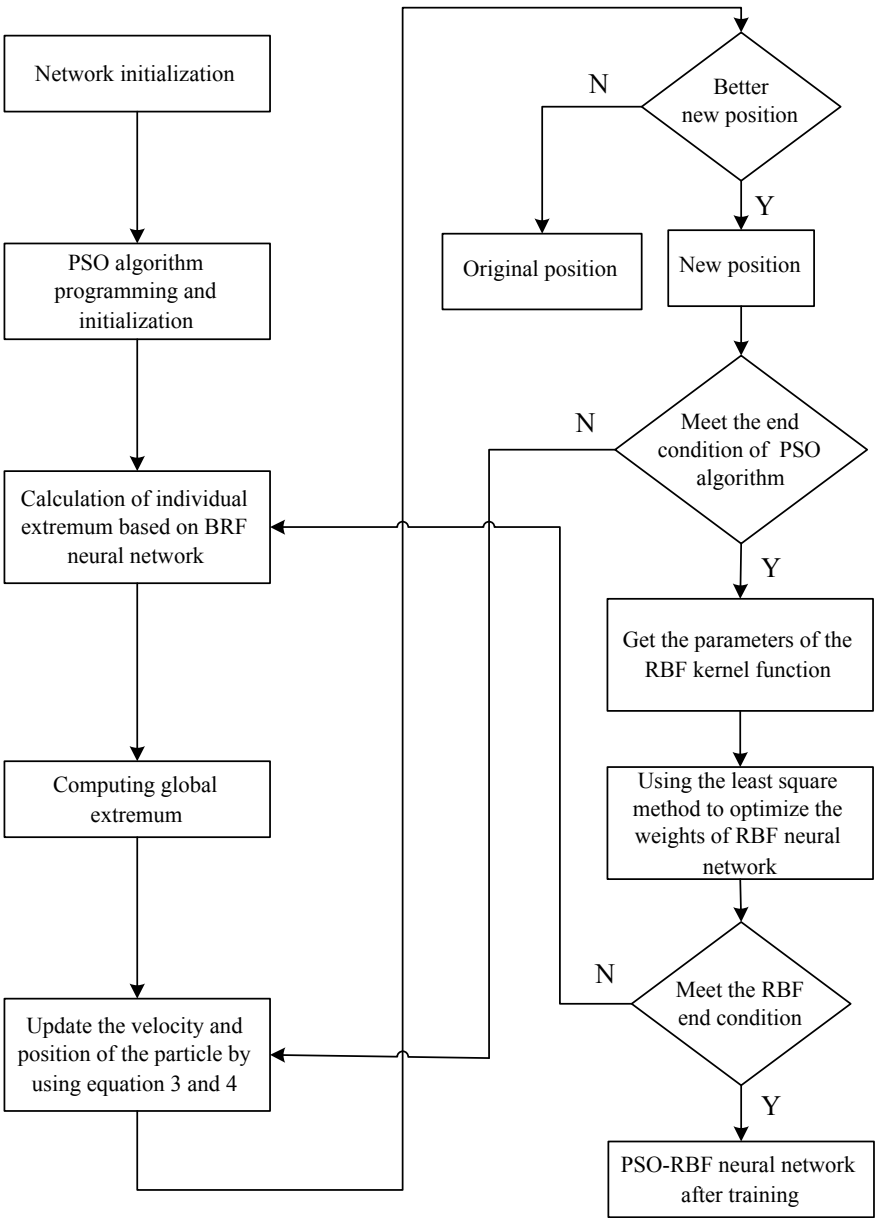



**Fig. 3.** PSO-RBF train algorithm flowchart

From Fig. 6, it can be seen that the frequency spectrum peaks at $f = 0.3$, thus enabling the detection of periodic signals.

From Table 1, it can be seen that the improved PSO-RBF can achieve faster convergence and better accuracy.

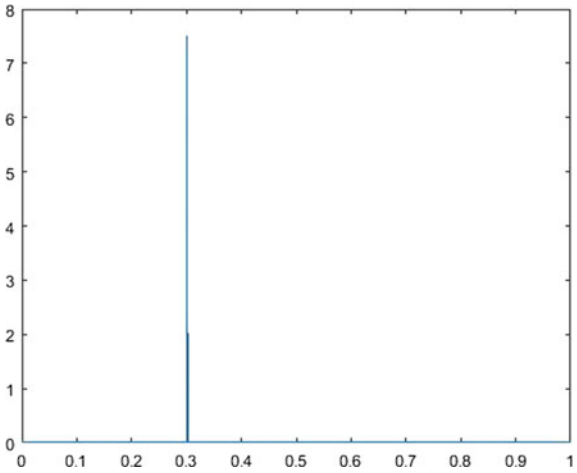



**Fig. 4.** Spectrum of a single periodic signal

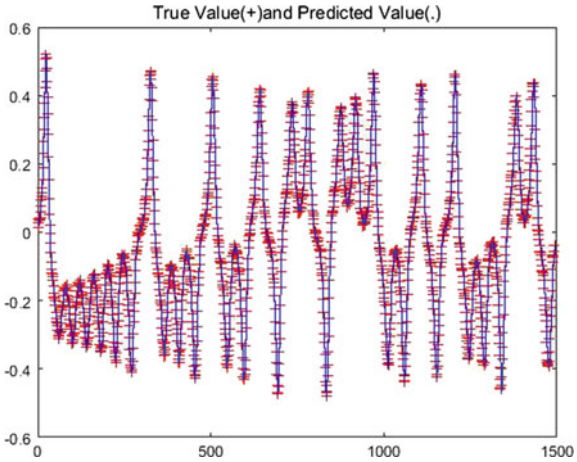



**Fig. 5.** Prediction value and real value of chaotic time series with a single periodic signal

### 4.2 IPIX Radar-Measured Sea Clutter Prediction

Select the #18 sea clutter data of the IPIX radar data set and select the sampling points of the 10,000 distance sea clutter data as the test sample. Using the RBF neural network trained to achieve prediction accuracy, the results are shown in Fig. 7.

It can be seen from Fig. 7 that there is a target at sampling point 1800-2200 and sampling point 8700-8800, which is consistent with the actual results (Table 2).

The experimental simulation shows that for the measured sea clutter data without weak targets, it can use RBF neural network and PSO-RBF neural network to perform single-step prediction because of its chaotic characteristics. For the measured sea clutter

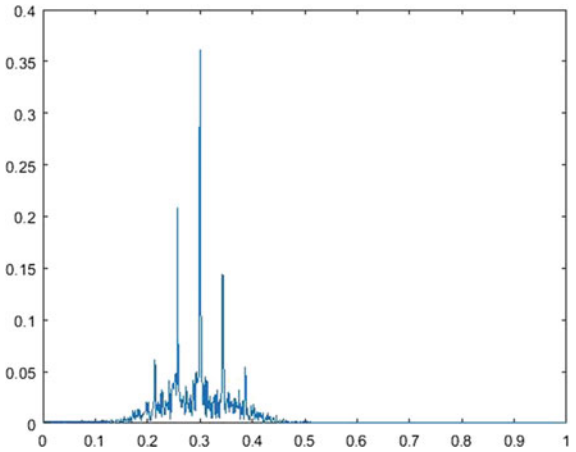



**Fig. 6.** Spectrum of prediction error of a single periodic signal

**Table 1.** Performance comparison of three algorithms

| Algorithm name | RBF | Classic PSO-RBF | Improved PSO-RBF |
|---|---|---|---|
| Train time (s) | 883.6 | 678.2 | 258.5 |
| RMSE | 0.032 | 0.183 | 0.000323 |

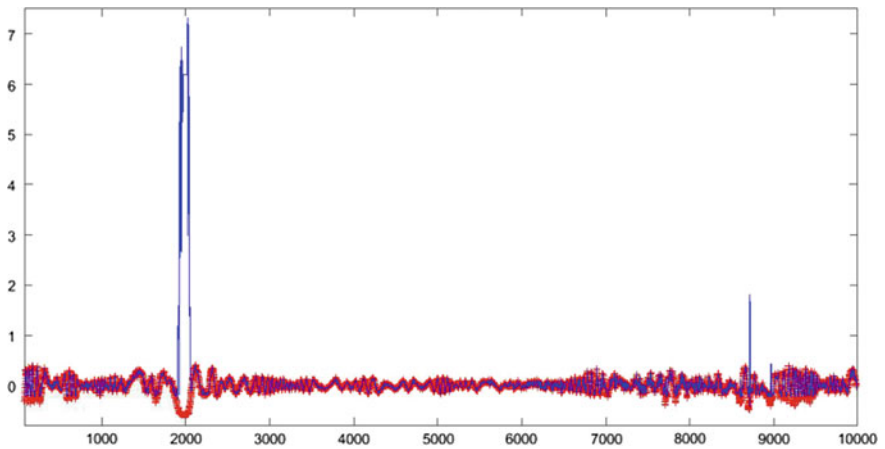



**Fig. 7.** IPIX radar-measured true value and predicted value sea clutter data

**Table 2.** Mean square error comparison of three algorithms

| Algorithm name | RBF | Classic PSO-RBF | Improved PSO-RBF |
|---|---|---|---|
| RMSE | 0.047 | 0.0143 | 0.00234 |

data with weak targets, the presence of the weak target destroys its inherent chaotic characteristics, which causes the prediction error to be larger than that of the pure sea clutter data. Therefore, we can effectively detect the weak target of the sea by comparing the turbidity characteristics of sea clutter by comparing the size of the prediction error. According to the comparison of prediction errors, it can be concluded that the improved PSO-RBF neural network detection is better.

## 5 Conclusion

In this paper, based on the analysis of particle swarm optimization, an improved particle swarm optimization algorithm is proposed for the PSO algorithm to fall into local minimum defects. The improved algorithm is applied to the hidden layer node parameters of the RBF neural network, namely the data center and the width. In the selection through simulation and IPIX radar data tests, it is verified that the improved PSO-RBF neural network has a better effect on weak target detection under sea clutter background.

## References

1. Roy LP, Kumar RVR. Accurate K-distributed clutter model for scanning radar application. IET Radar Sonar Navigation. 2010;4(2):158–67.
2. Hennsssey G, Leung H, Yip PC. Sea-clutter modeling using a radar-basis-function neural network. IEEE J Oceanic Eng. 2001;28(3):358–72.
3. Leung H, Dubash N, Xie N. Detection of small objects in clutter using a GA-RBF neural network. IEEE Trans Aerosp Electron Syst. 2002;38(1):98–118.
4. Hongyan JT. Weak signal detection based on dual constraint least squares support vector machine in Chaotic Sea clutter background. Acta Physica Sinica. 2010.
5. Si W, Tong N, Wang Q. Small target detection algorithm research in sea clutter background. Sig Process. 2014.
6. Wang F. Sea clutter fractal characteristic analysis, modeling and small target detection. Doctoral Dissertation. Harbin Engineering University; 2009.
7. Si W, Tong N, Wang Q. Small target detection algorithm in the sea clutter background. Sea clutter processing and target detection based on nonlinear analysis. Doctoral Dissertation. Dalian Maritime University; 2008.
8. Li D. Characteristic detection method for small and medium targets in sea clutter. Master Thesis. Xidian University; 2016.
9. Xing H, Xu Y. Neural network based on the detection of weak signals in chaotic background; 2007.

# An Identity Identification Method for Multi-biometrics Fusion

Yingli Wang, Yan Liu, Hongbin Ma(✉), Xin Luo, and Danyang Qin

Electronic Engineering College, Heilongjiang University,
XueFu Road. 74, Harbin 150080, China
mahongbin@hlju.edu.cn

**Abstract.** In order to improve the reliability and security of biometric-based identity authentication system and reduce the risk of unauthorized access caused by forgery feature attacks, this paper proposes a method for identifying the identity of visitors. The method is based on D-S evidence theory. The palm print and palm vein are used as authentication features. Firstly, the same collection device is used to collect palm print and palm vein images under different wavelengths of light source and extract the HOG features of the image; then, use the one-vs-one multi-classification method of SVM to classify different individuals, and finally, using the D-S fusion strategy at the decision-making level to improve the security and accuracy of the identity authentication system. Through many experiments, the recognition rate of decision-making layer fusion is above 98%, which confirms the effectiveness of the proposed method.

**Keywords:** Multiple biometrics · D-S fusion · Identification

## 1 Introduction

In the existing data protection system, identity authentication is the first line of defense. At present, the mainstream identity authentication method is biometric authentication technology, including fingerprint, palm print, iris, vein, and so on. In recent years, bionic technology has developed rapidly. The emergence of items such as fingerprint caps makes identity identification schemes based on a single biometric no longer secure. However, it is difficult for criminals to obtain raw data of multiple biometrics at the same time, which provides new ideas for improving the reliability and security of existing biometric-based identification systems. In the existing feature fusion scheme, fusion is mainly performed at the following four levels: original data layer fusion, feature layer fusion, matching layer fusion, and decision layer fusion. It is very important to choose reasonable biometric features and fusion strategies. In a complex interactive environment, it is worthwhile to improve the ability of biometric-based identity authentication


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_125

systems to defend against counterfeit feature attacks while ensuring recognition rates. Combined the above issue, this paper proposes a method for identifying the identity of visitors. The method is based on D-S evidence theory, and the palm print and palm vein are used as authentication features. Firstly, the same collection device is used to collect palm print and palm vein images under different wavelengths of light source and extract the HOG features of the image; then, use the one-vs-one multi-classification method of SVM to classify different individuals, and finally, using the D-S fusion strategy at the decision-making level to improve the security and accuracy of the identity authentication system.

## 2 Related Knowledge

Multi-feature fusion is the application of information fusion technology in the field of pattern recognition. The fusion of the decision-making layer belongs to the top-level fusion method. According to the result of single feature recognition, the fusion decision is made according to a certain distribution principle, and finally, the identity authentication will be output [1].

In the process of multi-feature fusion, there are some uncertain factors. The D-S evidence theory has good fuzzy inference performance for the uncertainty evidence. Xu and Yu [2] pointed out that the D-S method provides strong theoretical basis for the expression and synthesis of uncertain information. Chang et al. [3] used the D-S evidence theory to fuse the features of the image such as color and texture. Experiments have shown that it has a good recognition effect. In the field of identity authentication, researches of using D-S evidence theory for multi-feature fusion are rare.

In order to ensure the identification efficiency of the identity authentication system in the actual complex environment, this paper applies the D-S evidence theory fusion algorithm to the decision-making layer. Firstly, the HOG algorithm be used to extract the eigenvalues of a single biological image on the training data set. Then, the recognition subsystem be used to train. Next, the parameter correction set be used to test the recognition accuracy rate of the feature, which will be considered as the BPA of D-S evidence theory, and the different BPA is used for decision-level fusion recognition. Finally, the recognition result and rate of the system after feature fusion will be obtained.

The decision output $f(x)$ of the standard SVM is a hard output, which means that only $x$ is or does not belong to a certain class, and does not provide the probability of belonging to this class. The Dempster synthesis rule first needs to determine the basic probability setting of each piece of evidence. The recognition accuracy rate of each biometric in this paper is a piece of evidences BPA. Therefore, the SVM posterior probability proposed by Platt will be used as the BPA for each evidence.

After learning the training data set of a single feature, in order to ensure the objectivity of the BPA assignment, a parameter-corrected data set will be used to obtain more reliable position parameters, scale parameters, and recognition accuracy $\alpha$ of the feature-corresponding classifier. The $C_i$ refers to the user's

category number, and $i$ is the number of categories. The definition of the palm print feature recognition result is the first proof body, the palm vein feature recognition result is the second proof body, and the BPA function corresponding to the $j$th proof body can be described in the following form:

$$m_j(C_1, C_2, \ldots, C_i, \ldots, C_N) = \begin{pmatrix} p_{11}a_1, p_{12}a_1, \ldots, p_{1N}a_1 \\ \cdots \\ p_{j1}a_j, p_{j2}a_j, \ldots, p_{jN}a_j \end{pmatrix}, j = \{1, 2, \ldots\} \quad (1)$$

The degree of uncertainty of the $j$th evidence body is described in the following form:

$$m_j(\Theta) = 1 - \alpha_j \quad (2)$$

According to the Dempster synthesis rules and calculating the belief function and the plausibility function of the two evidence bodies, the reliability of the evidence to identify all the propositions of the framework and the uncertainty of the evidence are obtained.

## 3 Experiments and Analysis

### 3.1 The Source of Image Database

In order to verify the effectiveness of the proposed method, it is necessary to do the simulation experiment. This paper uses the database of palm print and palm vein images provided by the Hong Kong Polytechnic University as experimental data.

In this paper, 40 people were randomly selected from the above image library as the internal staff of a small company. The palm print image and the palm vein image were combined one by one. One person has eight palm print images and eight palm vein images. Four palms were randomly selected from one feature. The four of palm print images and the four of palm vein images are used as training data set to train the model proposed in this paper. The next two of them are selected as parameter correction data set to correct the probability of the single feature SVM classifier output and obtain the classification accuracy rate. The last two of them are used as test data set to obtain the accuracy of the fusion method proposed in this paper. The evaluation index of the proposed method is correct recognition rate (CRR).

### 3.2 Results and Analysis

In order to verify the recognition rate of individual features, experiments were performed on test data sets used to test palm print and palm vein. The construction of BPA function of palm print and palm vein evidence body depends on the classification accuracy rate. According to the experiment, the recognition rate of palm print is 85%, and the palm vein recognition rate is 90%. They are

**Table 1.** Comparison of the D-S fusion performance

| The name of experiment | The numbers of test images | The right numbers of the results | CRR in (%) | The max time of one sample in(s) |
|---|---|---|---|---|
| Palm print | 80 | 68 | 85 | 1.091947 |
| Palm vein | 80 | 72 | 90 | 1.533456 |
| D-S fusion | 80 | 79 | 98.75 | 5.365985 |

brought into the D-S fusion rule. The method proposed in this paper was tested using the test data set. Table 1 shows the recognition results of the D-S fusion method based on the palm print and palm vein proposed in this paper.

By analyzing the above experiment, we can draw the following conclusions:

(1) The traditional HOG algorithm can extract palm print features with a recognition rate of 85%, but it can be easily forged. Therefore, using only the palm print feature does not guarantee that the system is safe.
(2) Since we are not sure whether it is a fake palm print, we can combine the palm print and vein features to improve the safety of the system. The experimental results show that the proposed D-S fusion strategy has good performance for joint authentication of two features. A higher recognition rate demonstrates the effectiveness of the method.

## 4 Conclusion

In this paper, the D-S evidence theory is applied to the authentication method based on palm print and palm vein. The experimental results show that the method is simple and effective. It can not only use the authentication based on the palm vein to resist the forgery feature attack, but also realize the high recognition rate authentication system and provide an effective scheme with certain engineering significance. However, there are some deficiencies in this paper. For example, the number of test samples is not enough, and the performance of the method under the large-scale sample set is not verified. The recognition performance of palm vein needs to be further improved. The next step is to extract the texture features of the palm vein using other feature extraction methods (such as LBP) to improve the recognition rate of the palm vein.

## References

1. Zhang L, Liang T. Adaptive multi-modality biometric fusion based on classification distance scores. J Comput Res Develop. 2018;1:151–62.
2. Xu C, Yu W. A review of theory and application of dempster-shafer's evidence reasoning method. Pattern Recogn Artif Intell. 1999;12(4):424–30.
3. Chang C, Xiaoyang Y, Guang Y. The research of image retrieval based on multi feature ds evidence theory fusion. Int J Sig Process Image Process Pattern Recogn. 2016;9(1):51–62.

# Iced Line Position Detection Based on Least Squares

Yanwei Wang[1(✉)] and Jiaqi Zhen[2]

[1] College of Mechanical Engineering, Harbin Institute of Petroleum, Harbin 150080, China
xianxinyue@163.com
[2] College of Electronic Engineering, Heilongjiang University, Harbin 150080, China

**Abstract.** Ice and snow have a detrimental effect on the transmission and distribution lines. In the calculation of the ice thickness of the ice-covered transmission lines by the image method, the center position of the target transmission lines is crucial for the detection result. The detection of ice-covered transmission lines is affected by external factors, resulting in images that are mostly inclined ice-covered images. Therefore, it is necessary first to straighten the wire and then by fitting the image rotation technique.

**Keywords:** Ice-covering transmission lines · Least squares · Position detection

## 1 Introduction

In the ice detection system, the thickness of the ice coating is required to be detected due to different environmental conditions [1, 2]. In calculating the ice thickness of the target transmission line, we use the pixel difference in the vertical direction of the edge of the transmission line instead of the thickness of the ice coating. However, due to problems such as the angle of shooting, the power lines obtained after dividing the image are often inclined. So before calculating the thickness of the ice coating, we need to straighten the power line by image rotation technology [3, 4]. First, measure the tilt angle of the transmission line. Here, use the straight line fitting technique to obtain the tilt angle of the wire ice coating image, and then use the image rotation technique to straighten the inclined wire.

This work was supported by the National Natural Science Foundation of China under Grant No. 61501176, Heilongjiang Province Natural Science Foundation (F2018025), University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (UNPYSCT-2016017), The postdoctoral scientific research developmental fund of Heilongjiang Province in 2017 (LBH-Q17149).


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_126

## 2 Fitting Method of Ice Coating Line

If the number of sample points is a set of data, through a set of experimental data, according to the principle of the smallest sum of squared residuals, look for an analytic function, which is the least squares method, for a given set of data $(x_i y_i)(i = 0, 1, \ldots, m)$. It is required to find a function $y = s * (x)$ in function class $\Phi = \{\Phi_0, \Phi_1, \ldots, \Phi_n\}$, so that the sum of squared errors

$$\|\delta\|^2 = \sum_{i=0}^{m} \delta_i^2 = \sum_{i=0}^{m} [s * (x_i) - y_i]^2 \tag{2.1}$$

tends to be the smallest, ie:

$$\|\delta\|^2 = \min_{s(x)\in\Phi_i} \sum_{i=0}^{m} [s * (x_i) - y_i]^2 \tag{2.2}$$

polynomial in (2.1):

$$s(x) = a_0 * \Phi_0(x) + a_1 * \Phi_1(x) + \cdots + a_n * \Phi_n(x)\,(n < m) \tag{2.3}$$

is the least squares method of curve fitting.

The coefficients $a_1$, $a_2$, … , $a_n$ in the Eq. (2.2) can be obtained by solving the problem of the minimum value of (2.1). The final result $s(x)$ found is generally a high degree polynomial.

## 3 Detection of Ice-Covered Wires by Least Squares Straight Line Fitting Method

Combine the approximation parameter method to give an angle arbitrarily (according to our experimental experience, recommended as 70) give a maximum error according to the accuracy requirement, and then take the first two points according to the measured experimental data. Use the least squares method (for the first case, it is actually two points to determine a straight line) to fit its model (straight line), calculate the coordinates of its starting point and end point, and the connection between the next point and the starting point. The angle $\alpha$ of the fitted line: Assume that the starting point of the fitted line is $S(X_S, Y_S)$, the ending point is $E(X_E, Y_E)$, and the next point is $N(X_N, Y_N)$, then Lack, the distance between points $S_E$ is

$$|S_E| = \sqrt{(Y_S - Y_E)^2 + (X_S - X_E)^2} \tag{2.4}$$

the distance between points $S$ and $N$ is

$$|S_N| = \sqrt{(Y_S - Y_N)^2 + (X_S - X_N)^2} \tag{2.5}$$

the distance between points $E$ and $N$ is

$$|E_N| = \sqrt{(Y_E - Y_N)^2 + (X_E - X_N)^2} \tag{2.6}$$

the angle obtained by the cosine theorem is

$$\alpha = \cos^{-1} \frac{|S_E|^2 + |S_N - E_N|^2}{2|S_E S_N|} \tag{6.7}$$

It is judged whether the angle $\alpha$ is smaller than a given angle. If $\alpha$ is less than a given value, then this point $N$ is added to the previous point column to redo the least squares line fit until $\alpha$ is greater than the given value, and the resulting model is the model of the first straight line. Repeat the above-mentioned fitting and judgment with the last point on the straight line and the next point until all the points are fitted and obtain the models of all the segmented lines. Then use these models to calculate the intersection of adjacent lines. Get their intersections, which are the end points (start or end) of the fitted straight line segment. Once the model has been obtained, the error analysis of the model can be performed, that is, whether the obtained model satisfies the error requirement (Figs. 1, 2, and 3; Table 1).

**Fig. 1** Linear fitting of the first conductor

The experimental results show that the classification accuracy of color-based image feature extraction is between 80 and 90, and even 95% accuracy can be achieved. The more the number of samples, the higher the accuracy.

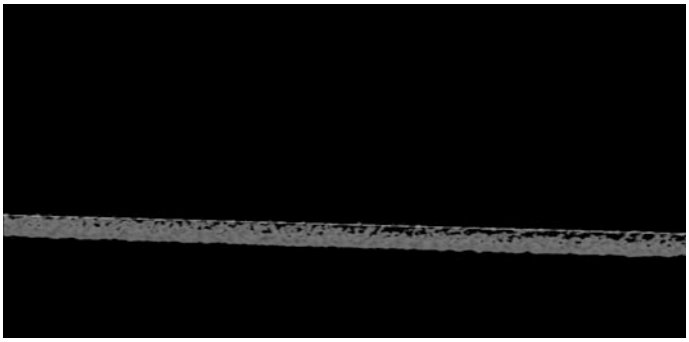


**Fig. 2** Linear fitting of the second conductor

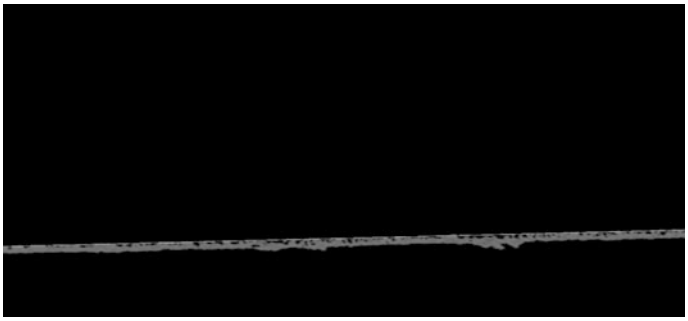


**Fig. 3** Linear fitting of the third conductor

**Table 1** Analysis

| Wire | Slope |
|---|---|
| Ice coated wire 1 | −0.5512 |
| Ice coated wire 2 | 4.8472 |
| Ice coated wire 3 | −3.9271 |

## 4 Conclusions

This paper proposes a method for realizing the position detection of ice-covered wires by means of least squares fitting. The method has the advantages of simple and rapid adjustment of adjusting the position of the ice-covered wire, measuring the inclination angle of the wire by the image method, and adjusting the inclination angle. The test result shows that the method can effectively collect the vertical direction of the wire edge in the ice-covered image. The pixel difference provides a reliable basis for calculating the thickness of the ice.

## References

1. Yano K. Studies of lcing and Ice-snow Accretion on Abies Mariesii. Bull Yamagata Univ Nat Sci. 1986;11:227–47.
2. Dai L, Huang H, Chen Z. Ternary sequential analytic optimization algorithm for SVM classifier design. Asian J Inf Technol. 2005;3:2–8.
3. Liao FM. Analysis of image processing and recognition technology of transmission line wire icing. South Agric Mach. 2017;48:94–5.
4. Gao J, Shen L, Xu YS, et al. Study on the detection of thickness and shortage diameter of ice coating based on sequence alignment method. Yunnan Electr Power. 2017;45:17–9.

# Sequentially Distributed Detection and Data Fusion with Two Sensors

LI Cheng(✉)

College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, China
lichengongl@l63.com

**Abstract.** The relationship of decision rule of sensor for each other is relevant to data fusion, so different topological networks of sensors usually results in different performances. This paper considers the sequential network fusion with two sensors in some detail and compares its performance with that of single detection and fusion. In this paper, the detection model is specified for binary hypotheses testing problem. In particular, this paper supposes that Bayesian risk cost of different decisions and the prior probability distribution of two hypotheses are known. Finally, this paper simulates the probabilities of error and Bayesian risk by these fusion rules with corresponding to different values of prior probabilities of two hypotheses by these fusion methods. And compared to single detection and fusion, the performance of sequential detection and fusion is better.

**Keywords:** Distributed detection · Single detection and fusion · Optimal fusion · Sequential network fusion

## 1 Introduction

Multi-sensor distributed detection and data fusion is an emerging technology that is applied in a wide range of areas such as automated target recognition, field surveillance, and so on [1, 2]. Contrast to the multi-sensor centralized detection which needs to collect the whole data received from all sensors to make decision, multi-sensor distributed detection refers to that each sensor could make temporary decision to be referred to the fusion center to make final decision, so multi-sensor data fusion refers to synergistic combination of information inferred by sensors for a better decision usually corresponding to a less probabilities of error [3].

## 2 Single Detection Fusion and Sequential Fusion

The relationship of decision rule of sensor for each other is relevant to data fusion, so different topological networks of sensors usually results in different performance. This paper considers sequential topological data fusion for two sensors on binary hypothesis testing problem in some detail. And then, this paper discusses its difference from the fusion of single detection and sequential fusion.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_127

### 2.1 Distributed Detection with Two Sensors

In a binary hypothesis testing problem, each hypothesis represents either the absence of a target or the presence of a target, and it can also represent symbol zero or one. Usually, the two hypotheses are respectively denoted by $H_0$ or $H_1$. For Bayesian detection method, the prior probabilities of the two hypotheses are known and denoted by $P_0$ or $P_1$. And what will be declared yields to a certain conditional probability under each hypothesis. This paper denotes the average cost or Bayesian risk function by $R$.

$$R = \sum_{i=0}^{1} \sum_{j=0}^{1} C_{ji} P_i P\big(\text{declaring } H_j | H_i \text{ is present}\big) \tag{1}$$

When it comes to two sensors, the conditional probability is related to the two declarations $u_1$ and $u_2$, respectively, from two sensors. $u_1$ is 1 if sensor 1 declares $H_1$ or else is 0, likewise for $u_2$. So, the conditional probability above can be expressed by

$$P\big(H_j | H_i\big) = \sum_{u_2=0}^{1} \sum_{u_1=0}^{1} P\big(\text{declaring } H_j | u_1, u_2\big) P(u_1, u_2 | H_i \text{ is present}) \tag{2}$$

And the conditional probability $P(u_2, u_1 | H_i \text{ is present})$ is determined by observations of sensors, $y_2$ and $y_1$. For analog observations, $y$ is stochastically sampled in continuous space and yields to a certain conditional probability distribution. And hence,

$$P(u_2, u_1 | H_i \text{ is present}) = \int_{Z_1(u_1)} \int_{Z_2(u_2)} p(u_2, u_1, y_2, y_1 | H_1) \mathrm{d}y_2 \mathrm{d}y_1 \tag{3}$$

where $Z_1(u_1)$ expresses the region corresponding to declaration $u_1$ for observation $y_1$, likewise, it is the same with $Z_2(u_2)$.

In above equation, the factor $P(u_2, u_1 | H_i \text{ is present})$ usually means distributed detection process and factor $P\big(\text{declaring } H_j | u_2, u_1\big)$ does data fusion process. However, not in all distributed detection and data fusion architecture, the two processes can be clearly separated. As to sequential network, the two processes of distributed detection and data fusion are combined.

### 2.2 Single Detection and Fusion with Two Sensors

For single detection, performances of sensors are conditionally independent,

$$P(u_1, u_2 | H_i) = P(u_1 | H_i) P(u_2 | H_i) \tag{4}$$

And in the case of two sensors' detection, there are $2^{2^2}$ fusion rules, however, only six of them are monotonic fusion rules [4]. The two of the six are simply all-one and all-zero, and their probabilities of error are the prior probabilities. And another two are

the decision of each sensor selves, totally disregarding the other. So, the probabilities of error of fusion are that of corresponding sensor. And the other two fusion rules are AND rule and OR rule, which comply with logical functions as follow,

$$\text{AND rule}: \ u_0 = \begin{cases} 1, & \text{if } u_1 = 1 \text{ and } u_2 = 1 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

$$\text{OR rule}: u_0 = \begin{cases} 1, & \text{if } u_1 = 1 \text{ or } u_2 = 1 \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

According to logical function and conditional independence, by any OR rule or AND rule, we can obtain the probability of detection, false alarm and missing which are denoted by $P_{\mathrm{D}}$, $P_{\mathrm{F}}$ and $P_{\mathrm{M}}$, respectively.

$$\text{AND rule}: \begin{array}{l} P_{\mathrm{D}} = P_{\mathrm{D1}} P_{\mathrm{D2}} \\ P_{\mathrm{F}} = P_{\mathrm{F1}} P_{\mathrm{F2}} \\ P_{\mathrm{M}} = 1 - P_{\mathrm{D}} \end{array} \tag{7}$$

$$\text{OR rule}: \begin{array}{l} P_{\mathrm{M}} = (1 - P_{\mathrm{D1}})(1 - P_{\mathrm{D2}}) \\ P_{\mathrm{F}} = 1 - (1 - P_{\mathrm{F1}})(1 - P_{\mathrm{F2}}) \\ P_{\mathrm{D}} = 1 - P_{\mathrm{M}} \end{array} \tag{8}$$

And, the probability of error is given by

$$P(\text{error}) = P(H_0) P_{\mathrm{F}} + P(H_1) P_{\mathrm{M}} \stackrel{\text{def}}{=} P_{\mathrm{E}} \tag{9}$$

### 2.3 Sequential Network Detection and Fusion

The architecture of sequential network as followed picture, two sensors observe the same source, and their observations are conditionally independent. The first sensor will transmit its decision to the second. And the second sensor will integrate the decision of the first sensor and its observation. The combination mode of decision is serial, that is to say, the second has two different thresholds corresponding to the decision of sensor 1. And the decision of second is not only the second decision but also the fusion result (Fig. 1).

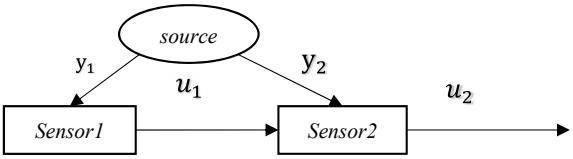



**Fig. 1.** Sequentially distributed detection and data fusion structure

The Bayesian risk function for this architecture is given by

$$R = \sum_{i,j,k} \iint_{y_1,y_2} C_{jk} P_k p(u_2|u_1, y_1, y_2, H_k) p(u_1, y_1, y_2|H_k) \mathrm{d}y_1 \mathrm{d}y_2 \tag{10}$$

where $i, j, k = 0, 1$, and $i$ is the decision $u_1$ and $j$ is decision $u_2$. And the decision $u_2$ does not depend on $y_1$, therefore,

$$R = \sum_{i,j,k} \iint_{y_1,y_2} C_{jk} P_k p(u_2|u_1, y_2) p(u_1, y_1|H_k) p(y_2|H_k) \mathrm{d}y_1 \mathrm{d}y_2 \tag{11}$$

Explicitly summing over $u_2$ and considering $p(u_2 = 1|u_1, y_2) + p(u_2 = 0|u_1, y_2) = 1$.

$$\begin{aligned} \mathrm{R} = & \sum_{i,k} \iint_{y_1,y_2} C_{jk} P_k p(u_1, y_1|H_k) p(y_2|H_k) \mathrm{d}y_1 \mathrm{d}y_2 \\ & + \sum_i \int_{y_2} p(u_2 = 0|u_1, y_2) \sum_k \int_{y_1} (C_{0k} - C_{1k}) \\ & \times P_k p(u_1, y_1|H_k) p(y_2|H_k) \mathrm{d}y_1 \mathrm{d}y_2 \end{aligned} \tag{12}$$

In order to minimize the Bayesian risk, the decision $u_2 = 1$ is under the condition that $\sum_k (C_{0k} - C_{1k}) P_k p(u_1|H_k) p(y_2|H_k) > 0$, then the decision rule of sensor 2 is,

$$\frac{p(y_2|H_1)}{p(y_2|H_0)} \begin{matrix} u_2 = 1 \\ > \\ < \\ u_2 = 0 \end{matrix} \frac{P_0(C_{10} - C_{00}) p(u_1|H_0)}{P_1(C_{01} - C_{11}) p(u_1|H_1)} \tag{13}$$

Therefore, $u_1$ is 0 or 1, representing different values of $\frac{p(u_1|H_0)}{p(u_1|H_1)}$. And, we denote the two thresholds by $t_2^1$ and $t_2^0$. They are expressed by

$$t_2^1 = \frac{P_0(C_{10} - C_{00}) p(u_1 = 1|H_0)}{P_1(C_{01} - C_{11}) p(u_1 = 1|H_1)}, \qquad t_2^0 = \frac{P_0(C_{10} - C_{00}) p(u_1 = 0|H_0)}{P_1(C_{01} - C_{11}) p(u_1 = 0|H_1)} \tag{14}$$

Next, we come back to risk function Eq. (11), and rewrite it as follows

$$R = \sum_{i,j,k} \iint_{y_1,y_2} C_{jk} P_k p(u_2|u_1, y_2) p(u_1, y_1|H_k) p(y_2|H_k) \mathrm{d}y_1 \mathrm{d}y_2 \tag{15}$$

Expanding over $u_1$ and also considering $p(u_1 = 1|y_1) + p(u_1 = 0|y_1) = 1$

$$R = \sum_{j,k} \iint_{y_1,y_2} C_{jk} P_k p(u_2|u_1 = 1, y_2) p(y_1|H_k) p(y_2|H_k) \mathrm{d}y_1 \mathrm{d}y_2 + \int_{y_1} p(u_1 = 0|y_1)$$
$$\sum_{j,k} \int_{y_2} C_{jk} P_k p(y_1|H_k) p(y_2|H_k) [p(u_2|u_1 = 0, y_2) - p(u_2|u_1 = 1, y_2)] \mathrm{d}y_1 \mathrm{d}y_2 \tag{16}$$

In order to minimize the Bayesian risk, the decision $u_1 = 1$ is under the condition that $\sum_{j,k} C_{jk} P_k p(y_1|H_k)[\mathrm{p}(u_2|u_1 = 0, H_k) - p(u_2|u_1 = 1, H_k)] > 0$ [5]. Expanding over $j$ and $k$, then considering $C_{10} > C_{00}$ and $C_{01} - C_{11}$, we can obtain the decision rule of sensor 1 with considering $t_2^1$ and $t_2^0$,

$$\frac{p(y_1|H_1)}{p(y_1|H_0)} \underset{u_1 = 0}{\overset{u_1 = 1}{\gtrless}} \frac{P_0(C_{10} - C_{00})\left[P_{\mathrm{F2}}\left(t_2^1\right) - P_{\mathrm{F2}}\left(t_2^0\right)\right]}{P_1(C_{01} - C_{11})\left[P_{\mathrm{D2}}\left(t_2^1\right) - P_{\mathrm{D2}}\left(t_2^0\right)\right]} \overset{\mathrm{def}}{=} t_1 \tag{17}$$

And, the probability of false alarm and missing are, respectively, following

$$P_{\mathrm{F}} = P_{\mathrm{F1}} P_{\mathrm{F2}}\left(t_2^1\right) + (1 - P_{\mathrm{F1}}) P_{\mathrm{F2}}\left(t_2^0\right) \tag{18}$$

$$P_{\mathrm{M}} = P_{\mathrm{D1}} \cdot \left(1 - P_{\mathrm{D2}}\left(t_2^1\right)\right) + (1 - P_{\mathrm{D1}}) \cdot \left(1 - P_{\mathrm{D2}}\left(t_2^0\right)\right) \tag{19}$$

The probability of error is the same as Eq. (9).

## 3 Simulation Result

### 3.1 Description of Two Hypotheses

Suppose the noise and target observations yield to unit Gaussian distribution for both two sensors. The distribution of observation and Bayesian cost factors is shown in Tables 1 and 2. In the simulation, the average cost is expressed by minus number, so the lower is the average cost, the bigger is Bayesian risk.

**Model**

**Table 1.** Likelihood function at two sensors

| Sensor 1 | $p(y_1\|H_0) \sim N(0,1)$<br>$\mathrm{p}(y_1\|H_1) \sim N(2,9)$ ’ | Sensor 2 | $p(y_2\|H_0) \sim N(0,1)$<br>$\mathrm{p}(y_2\|H_1) \sim N(1,1)$ |
|---|---|---|---|

**Table 2.** Bayesian cost factors

| Cost factors | Declaring $H_0$ | Declaring $H_1$ |
|---|---|---|
| Presenting $H_0$ | 0 | −1 |
| Presenting $H_1$ | −1 | 0 |

### 3.2 Single Sensor Detection and Fusion

First, according to Bayesian rules for single sensor detection, we check the probability of error and Bayesian risk of both two sensors. And the performances of each sensor are shown in Fig. 2. We find that the performance of sensor 1 is better than that of sensor 2, because the difference of observation means under both hypotheses is larger at sensor 1 than that at sensor 2. And, the maximal probability of error of sensor 1 is 0.24, while that of sensor 2 is 0.31. And then, we check the performances by the fusion of AND rule and OR rule as shown in Fig. 3, the maximal probability of error by the fusion of OR rule is 0.25, lower than that by the fusion of AND rule, 0.29, and a bit higher than that of single detection at sensor 1, 0.24. That is because, in this case, the probability of missing is reduced by fusion of OR rule and the probability of false alarm, however, is increased a bit.

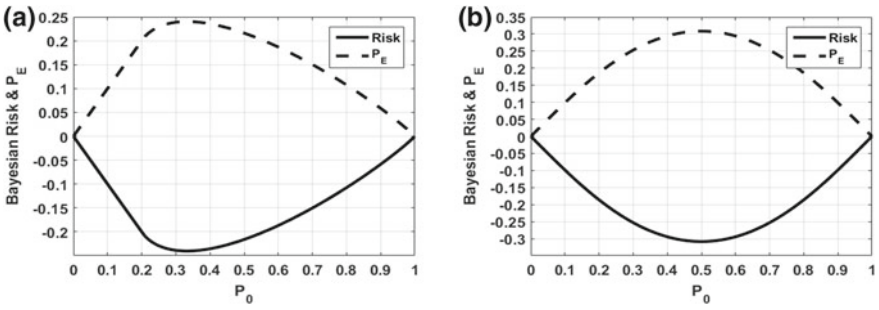



**Fig. 2.** **a** Performance of sensor 1 **b** Performance of sensor 2

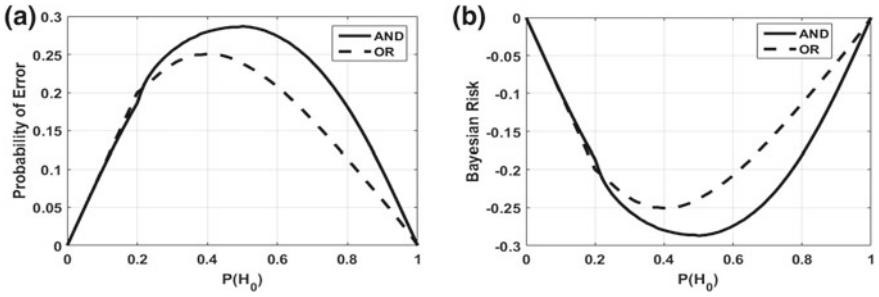



**Fig. 3.** **a** Probability of error by the fusion of AND rule and OR rule **b** Bayesian risk by the fusion of AND rule and OR rule

### 3.3 Sequential Fusion Result

For two sensors' detection, the sequential fusion has two states corresponding different sequences of sensors. And according to Eqs. (14) and (17), the decision of the latter sensor depends on the former. And we know by Fig. 2 that sensor 1 is better than sensor 2. So, the thresholds and performances in two states may be different. We first simulate the sequence in which sensor 2 is placed latter to fuse decisions, and then the other sequence. And the simulation result of sequential fusion network is showed in Fig. 4a, and the worst Bayesian Risk is −0.21, better than that of both OR rule fusion, −0.25, and single detection at sensor 1, −0.24. Because we adapt numerical methods to solve the thresholds. Figure 4b shows the logarithm of three thresholds of two sensors versus the prior probability of hypothesis $H_0$. When sensor 1 is placed latter to fuse decisions, the simulation result is showed in Fig. 5a. And its worst Bayesian risk is −0.21, similar to that of above sequence. But by comparing Bayesian risks of both sequences respect to $P_0$ as 0.7, 0.8, and 0.9, the Bayesian Risks of current sequence are better (Fig. 4).

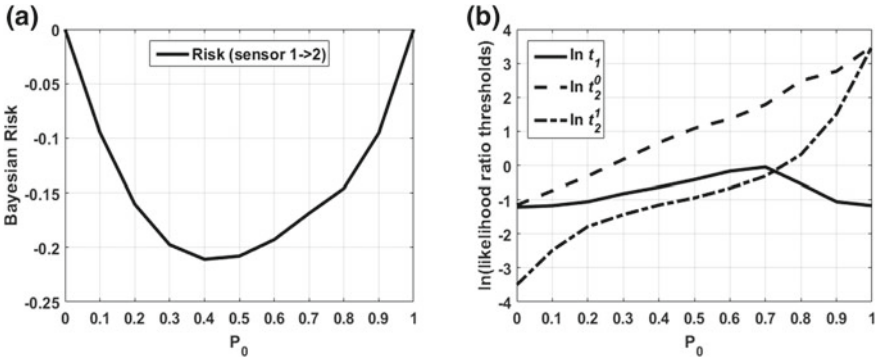



**Fig. 4.** **a** Performance of sequential network with sensor 2 as the latter to fuse decisions. **b** Logarithm of three thresholds of two sensors versus prior probability of hypothesis $H_0$

Figures 4b and 5b show the three logarithmic thresholds versus the prior probability of $H_0$. When we consider the thresholds rationally at the point that prior probability of $H_0$ is approaching zero, the logarithmic thresholds of sensor 2 should tend to be negative infinite so that the probability of detection is 1, and they should tend to be infinite when the prior probability of $H_0$ is approaching one. In fact, in Fig. 6a and b, the logarithmic thresholds $ln t_2^1$ and $ln t_2^0$ haven't converged yet, but the Bayesian risk and prob-ability of error has been convergent to zero along with iterations.

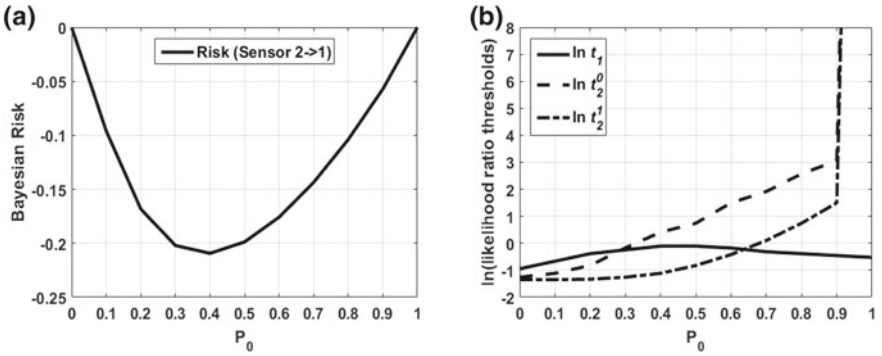



**Fig. 5.** **a** Performance of sequential network with sensor 1 as the latter to fuse decisions. **b** Logarithm of three thresholds of two sensors versus prior probability of hypothesis $H_0$

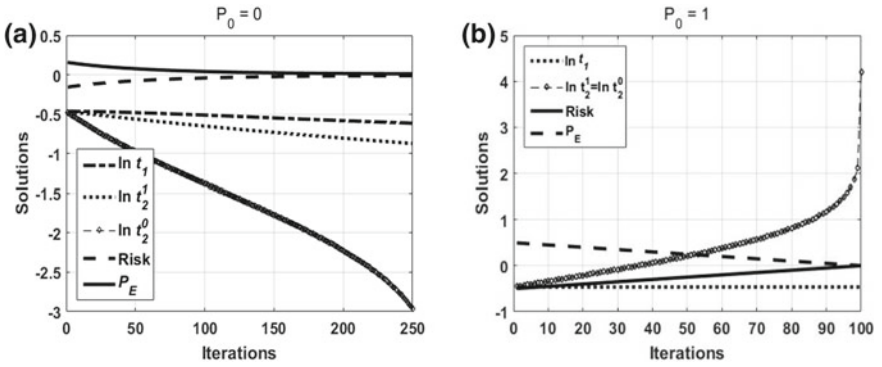



**Fig. 6.** **a** Iterative solutions of performance and thresholds when prior probability of $H_0$ is 0. **b** Iterative solutions of performance and thresholds when prior probability of $H_0$ is 1

## 4 Conclusion and Discussion

This paper discusses the fusion rule of two sensors' detection, and the observations at two sensors are conditionally independent. For single detection and fusion with two sensors, the two decisions are not conditionally independent, and there are six monotonic fusion rules. Two of the six fusion rules are all-one and all-zero which probability of error are probability of hypotheses selves, so maybe not referred in most occasions. This paper, particularly, compares the other four fusion rules with sequentially distributed detection data fusion. For sequential detection, the decisions of two sensors are not conditionally independent any more in sequential fusion, and they are coupled serially. And the thresholds of the latter sensor are dependent on the declaration of the former sensor, so the sequencing is relevant. The latter the better sensor is placed, the better is the fusion performance.

## References

1. Chair Z, Varshney PK. Optimal data fusion in multiple sensor detection systems. IEEE Trans Aerosp Elect Syst. 1986;AES-22(1):98–101.
2. Llinas J, Hall DL. Introduction to multi-sensor data fusion. Proc. IEEE. 1998;(6). 537–540 vol.6. 10.1109/ISCAS.1998.705329.
3. Blum RS. Quantization in multisensor random signal detection. IEEE Trans On Info Theory. 1995;41(1):204–215.
4. Chair Z, Varshney PK. Distributed bayesian hypothesis testing with distributed data fusion. IEEE Trans Syst Man Cybern. 1988;SMC-18(5):695–9.
5. Varshney PK, Burrus CS. Distributed detection and data fusion. New York: Springer; 1997.

# Part V
# Localization and Navigation

# Particle Filter with Correction of Initial State for Direction of Arrival Tracking

Huang Wang, Qiyun Xuan, Yulong Gao(✉), and Xu Bai

Harbin Institute of Technology, Harbin 150028, China
16S105138@stu.hit.edu.cn, hitxuanqiyun@126.com,
{ylgao,x_bai}@hit.edu.cn

**Abstract.** Generally, particle filter is used in the single snapshot situation and the initial state is assumed to be known. To make the measurement interval be small enough, we construct a multiple measurement vectors model for DOA tracking since it usually outperforms the single measurement vector model. And we take the initial state into consideration. The initial tracking error of the particle filter becomes very large when the initial state is unknown. Thus, we modify the initial state according to the likelihood of the generated random samples. The method is numerically evaluated using a uniform linear array in simulations. The results show that the proposed algorithm has higher tracking accuracy.

**Keywords:** DOA tracking · Particle filter · Multiple measurement vectors

## 1 Introduction

Decades of research have given rise to many algorithms that solve the direction of arrival (DOA) estimation problem which has important applications in fields like radar, sensor networks, and communications. The high-resolution DOA estimation is realized by algorithms like multiple signal classifications (MUSIC) [1] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [2]. But these methods fail or suffer performance degradations when the DOA is variant during the observation period. However, in many scenarios, the sources are moving. Therefore, it is necessary to study the DOA tracking algorithm.

Some DOA tracking algorithms are based on subspace tracking [3, 4]. The subspace tracking algorithm, such as projection approximation subspace tracking (PAST) [5], updates the subspace by minimizing the objective function and avoids the heavy eigenvalue decomposition operation. In general, the performance of subspace-based algorithms degrades with signal correlation. Another DOA tracking algorithms are based on state filtering, including extended Kalman filter (EKF) [6], unscented Kalman filter (UKF) [7], and particle filter (PF) [8–11]. EKF transforms the nonlinear model into a linear model by using the Taylor series expansion and ignoring the higher

This work is supported by National Natural Science Foundation of China (NSFC) (Grant No. 61671176).


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_128

order terms. UKF uses unscented transform to approximate the nonlinear distribution by sampling. Although these two methods are used to deal with nonlinear systems, they can not be guaranteed to be convergent. Particle filter is a recursive algorithm which is widely used in nonlinear/non-Gaussian system. A filtering idea based on Bayesian sequential importance sampling is first put forward by Gordon, Salmond, and Smith, which establishes the theoretical foundation of particle filter algorithm. Then, a variety of improved algorithms has been developed. In [9], PF is introduced to DOA tracking problem. In [10, 11], DOA tracking algorithms based on PF are proposed to deal with two-dimensional DOA and impulsive noise, respectively.

In this paper, we implement DOA tracking in the framework of sampling importance resampling (SIR) filter and improve the algorithm. First, we construct a dynamic state-space model and use multiple measurement vectors in the observation equation. We make the measurement interval very small so that the source is regarded to be stationary. Through the use of multiple snapshot data, the estimation accuracy is improved. Then, DOA tracking is carried out by sequential importance sampling and resampling. In addition, we consider the case that the initial state is unknown. In this case, the initial tracking accuracy is very low. To solve this problem, we modify the initial state using the observation data. It follows that higher accuracy is obtained with the correction of initial state.

## 2 System Mode

Suppose that $K$ narrowband far-field signals impinge onto an array of $M$ omnidirectional sensors from unknown time-varying directions $\boldsymbol{\theta}(t) = [\theta_1(t), \theta_2(t), \ldots, \theta_K(t)]^{\mathrm{T}}$. The wavelength of signals is $\lambda$. The space of the adjacent antennas is $d$. The sequential sampling approach we adopt admits a first-order state-space hidden Markov model. The time-varying direction vector is modeled as

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t-1) + \mathbf{v}(t) \tag{1}$$

where $\mathbf{v}(t)$ is normally distributed with zero mean and the covariance matrix $\sigma_v^2\mathbf{I}$. Here, $\sigma_v^2$ represents the noise spectral parameter, and $\mathbf{I}$ is the identity matrix.

For a collection of observed output date of $M$ sensors in the array, $\mathbf{x}(t) = [x_1(t), \ldots, x_M(t)]^{\mathrm{T}}$, the matrix formulation of outputs of the sensors at time $t$ is obtained as follows

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}(t))\mathbf{s}(t) + \mathbf{n}(t) \tag{2}$$

where $\mathbf{s}(t) = [s_1(t), \ldots, s_K(t)]^{\mathrm{T}}$ is the signal waveform, $\mathbf{n}(t) = [n_1(t), n_2(t), \ldots, n_M(t)]^{\mathrm{T}}$ is the complex independent Gaussian noise with zero mean and variance $\sigma_n^2$, $\mathbf{A}(\boldsymbol{\theta}(t)) = [\mathbf{a}(\theta_1(t)), \ldots, \mathbf{a}(\theta_K(t))]$ is the array manifold matrix, and $\mathbf{a}(\theta_k(t)) = [1, \mathrm{e}^{-j2\pi d \sin\theta_k(t)/\lambda}, \ldots, \mathrm{e}^{-j2\pi(M-1)d \sin\theta_k(t)/\lambda}]^{\mathrm{T}}$ is the steering vector.

We assume that the signal is observed very $T$ seconds. $L$ measurements are taken for each increment $T$, and the time interval over which these are gathered is sufficiently

small proportion of $T$ for us to approximate the DOA as stationary over the measurement interval. That is, $\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t+\tau) = \cdots = \boldsymbol{\theta}(t+(L-1)\tau)$, where $\tau$ is the time between measurements and $L\tau \ll T$. Then, the observation model with multiple measurement vectors is expressed as

$$\mathbf{X}(t) = \mathbf{A}(\boldsymbol{\theta}(t))\mathbf{S}(t) + \mathbf{N}(t) \tag{3}$$

where $\mathbf{X}(t) = [\mathbf{x}(t), \mathbf{x}(t+\tau), \ldots, \mathbf{x}(t+(L-1)\tau)]$, $\mathbf{S}(t)$ and $\mathbf{N}(t)$ are defined similarly.

## 3 Description of the Proposed Method

### 3.1 Particle Filter

From a Bayesian perspective, the tracking problem is to recursively calculate some degree of belief in the state $\boldsymbol{\theta}_t$ at time $t$, taking different values, given the data $\mathbf{X}_{1:t}$ up to time $t$. In principle, the posterior probability density function $p(\boldsymbol{\theta}_t|\mathbf{X}_{1:t})$ may be obtained recursively in two stages: prediction and update. Suppose that $p(\boldsymbol{\theta}_{t-1}|\mathbf{X}_{1:t-1})$ at time $t$-1 is available, the prediction stage is as follows

$$p(\boldsymbol{\theta}_t|\mathbf{X}_{1:t-1}) = \int p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) p(\boldsymbol{\theta}_{t-1}|\mathbf{X}_{1:t-1}) \mathrm{d}\boldsymbol{\theta}_{t-1} \tag{4}$$

At time step $t$, the measurement $\mathbf{X}_t$ becomes available, the update stage is implemented via Bayes rule

$$p(\boldsymbol{\theta}_t|\mathbf{X}_{1:t}) \propto p(\mathbf{X}_t|\boldsymbol{\theta}_t) p(\boldsymbol{\theta}_t|\mathbf{X}_{1:t-1}) \tag{5}$$

The basic idea behind the particle filter is that we represent the posterior distribution of interest with a set of weighted particles, each of which forms an independent hypothesis of the state at a given time. The general particle filter consists of three stages. First, we use the set of particles from the previous time step to propose a new set for the new time step, according to the importance function $q(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \mathbf{X}_t)$. Then, the weights of particles are calculated by

$$w_t^{(i)} \propto \tilde{w}_{t-1}^{(i)} \frac{p\left(\mathbf{X}_t \middle| \boldsymbol{\theta}_t^{(i)}\right) p\left(\boldsymbol{\theta}_t^{(i)} \middle| \boldsymbol{\theta}_{t-1}^{(i)}\right)}{q\left(\boldsymbol{\theta}_t^{(i)} \middle| \boldsymbol{\theta}_{t-1}^{(i)}, \mathbf{X}_t\right)} \tag{6}$$

where $i = 1, \ldots, Ns$, $Ns$ is the number of particles. And, the normalized weights are given by

$$\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{i=1}^{Ns} w_t^{(i)}} \tag{7}$$

Finally, the weighted particles may be resampled to convert them into an unweighted set without changing the distribution they represent. We use SIR algorithm to implement DOA tracking which chooses the prior $p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1})$ as the importance function so that the particle weights are then given by the likelihood only.

### 3.2 Correction of the Initial State

Considering the case that the initial state is unknown, we improve the algorithm under the framework of SIR. First, we produce a set of random samples $\left\{\boldsymbol{\theta}_0^{(i)}\right\}_{i=1}^{Ns}$ which obey the distribution $U[-90^\circ, 90^\circ]$, where $U[\cdot]$ denotes a uniform distribution. Then, the likelihood function of each particle is calculated according to the observation data at the initial time, which is followed as

$$p\left(\mathbf{X}_0\middle|\boldsymbol{\theta}_0^{(i)}\right)=\frac{\exp\left(-tr\left(\left(\mathbf{X}_0-\mathbf{A}\left(\boldsymbol{\theta}_0^{(i)}\right)\hat{\mathbf{S}}_0\right)^H\left(\sigma_n^2\mathbf{I}_M\right)^{-1}\left(\mathbf{X}_0-\mathbf{A}\left(\boldsymbol{\theta}_0^{(i)}\right)\hat{\mathbf{S}}_0\right)\right)\right)}{\left(\pi^M\det\left(\sigma_n^2\mathbf{I}_M\right)\right)^L} \tag{8}$$

where $tr(\cdot)$ denotes the trace, $\det(\cdot)$ denotes the determinant. $\hat{\mathbf{S}}_0$ is the least squares estimate of $\mathbf{S}_0$, which is given by

$$\hat{\mathbf{S}}_0=\left(\mathbf{A}\left(\boldsymbol{\theta}_0^{(i)}\right)^H\mathbf{A}\left(\boldsymbol{\theta}_0^{(i)}\right)\right)^{-1}\mathbf{A}\left(\boldsymbol{\theta}_0^{(i)}\right)^H\mathbf{X}_0 \tag{9}$$

Each particle is weighted by the likelihood function, that is $w_0^{(i)}=p\left(\mathbf{X}_0\middle|\boldsymbol{\theta}_0^{(i)}\right)$. Then, we get the normalized weights $\tilde{w}_0^{(i)}$ and the corrected initial state can be obtained by

$$\tilde{\theta}_0=\sum_{i=1}^{Ns}\tilde{w}_0^{(i)}\boldsymbol{\theta}_0^{(i)} \tag{10}$$

To eliminate particles that have small weights and to concentrate particles with large weights, we generate a new set $\left\{\boldsymbol{\theta}_0^{(i*)}\right\}_{i=1}^{Ns}$ by resampling which satisfies

$$P\left(\boldsymbol{\theta}_0^{(i*)}=\boldsymbol{\theta}_0^{(j)}\right)=\tilde{w}_0^{(j)} \tag{11}$$

Then, the weights are reset to $\tilde{w}_0^{(i)}=\frac{1}{Ns}$. Thus, we can treat this new set as the initial particle set in the SIR algorithm and perform SIR to realize DOA tracking in the following time.

A summary of the proposed algorithm is as follows

- Correct the initial state
  - Produce a set of random samples $\left\{\boldsymbol{\theta}_0^{(i)}\right\}_{i=1}^{Ns}$ from $U[-90^\circ, 90^\circ]$

  - Calculate the likelihood function of each sample by (8)
  - Obtain the corrected initial state by (10)
  - Determine the initial particle set by resampling
- Carry out SIR filter
  - Draw $\boldsymbol{\theta}_t^{(i)} \sim p\left(\boldsymbol{\theta}_t \middle| \boldsymbol{\theta}_{t-1}^{(i)}\right)$, $i = 1, \ldots, Ns$
  - Calculate the particle weights $w_t^{(i)} = p\left(\mathbf{X}_t \middle| \boldsymbol{\theta}_t^{(i)}\right)$ and normalize them by (7)
  - Resampling
- Output the state estimation

## 4 Numerical Simulations

Several simulations are performed to verify the feasibility of the proposed method. In the simulations, $K = 1$, $M = 10$, $d = \lambda/2$, $\sigma_v^2 = 1$, $\sigma_n^2 = 1$, $T = 1$, $Ns = 100$ and SNR = 5 dB.

We first analyze the DOA tracking error of SIR for the various numbers of snapshots and the root mean square error (RMSE) is given by $\mathrm{RMSE}(t) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}^i(t) - \theta^i(t)\right)^2$, where $N$ is the times of Monte Carlo trials. In this simulation, the times of Monte Carlo are 100, and the initial state is assumed to be known. The results are shown in Fig. 1. It can be seen that the tracking error reduces with the increasing of the number of snapshots, which proves that the multiple measurement vectors model has better performance than the single measurement vector model.

Then, we compare the DOA tracking results of SIR algorithm with the known initial state and the unknown initial state. In this simulation, we utilize the multiple

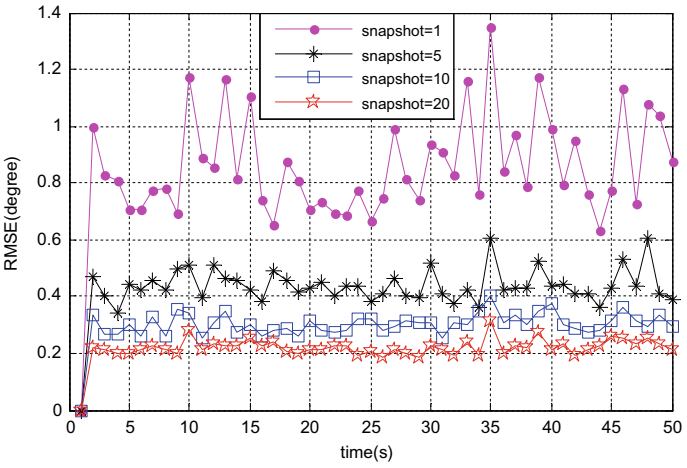



**Fig. 1.** Tracking error of SIR with different numbers of snapshots

measurement vectors model and set $L = 5$. The simulation results are illustrated in Fig. 2. It is indicated that the information of the initial state will affect the performance of the algorithm. The performance is obviously degraded when the initial state is unknown. As shown in the figure, the tracking error is stayed small in the known initial state case while it is very large in the first few seconds in the unknown initial state case. At time $t = 5$, the estimated DOA with unknown initial state starts to get close to the true DOA

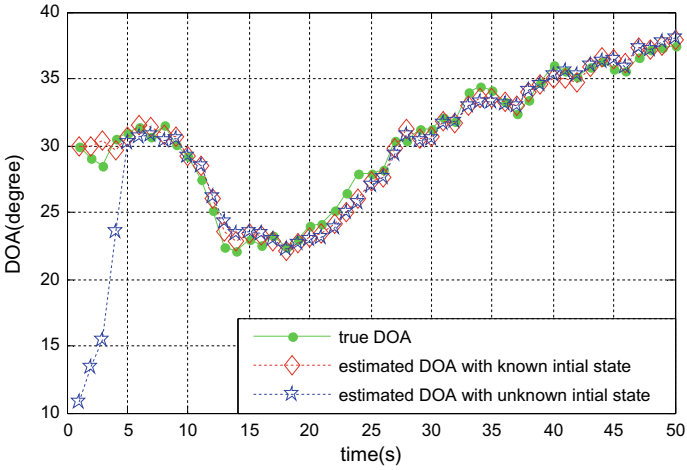



**Fig. 2.** DOA tracking with known initial state and unknown initial state

To reduce the initial tracking error in the case of the unknown initial state, we add the correction step. In the following, we evaluate the performance of the proposed method. In this simulation, we set $L = 5$ and carry out 1000 Monte Carlo trials. The results are shown in Fig. 3. The RMSE of SIR with unknown initial state is higher than 30 degree while that of the proposed method is lower than 3 degree. It is illustrated that the proposed method can improve the initial tracking performance.

## 5 Conclusions

In this paper, a particle filter algorithm with correction of the initial state for DOA tracking is proposed. A dynamic state-space model is constructed to describe the state variation tendency. In this model, we utilize multiple measurement vectors in the observation equation in order to take advantage of more data information. To solve the problem that the initial tracking performance is degraded under the situation of unknown initial state, we correct the initial state by weighting the random particles

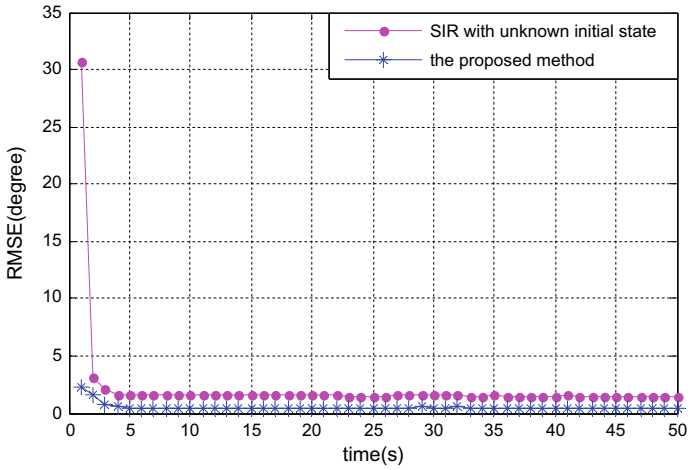



**Fig. 3.** Performance comparison between the proposed method and SIR with unknown initial state

based on likelihood function. The proposed algorithm performs better than SIR filter algorithm.

## References

1. Schmidt RO. Multiple emitter location and signal parameter estimation. IEEE Trans Antennas Propag. 1986;34(3):276–80.
2. Roy R, Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques. IEEE Trans Acoust Speech Signal Process. 1989;37(7):984–95.
3. Badeau R, David B, Richard G. Fast approximated power iteration subspace tracking. IEEE Trans Signal Process. 2005;53(8):2931–41.
4. Liao B, Zhang ZG, Chan SC. DOA estimation and tracking of ULAs with mutual coupling. IEEE Trans Aerosp Electron Syst. 2012;48(1):891–905.
5. Yang B. Projection approximation subspace tracking. IEEE Trans Signal Process. 1995;43 (1):95–107.
6. Kong D, Chun J. A fast DOA tracking algorithm based on the extended Kalman filter. In: National aerospace and electronics conference, 2000. Naecon 2000. Proceedings of the IEEE. IEEE Xplore; 2000. p. 235–8.
7. Kumar CV, Rajagopal R, Kiran R. An optimal integrated tracking (ITS) for passive DOA tracking using unscented Kalman filter. In: Information, decision and control, 2002. Final Program and. IEEE; 2002. p. 253–8.
8. Arulampalam MS, Maskell S, Gordon N, et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Trans Signal Process. 2002;50(2):174–88.
9. Orton M, Fitzgerald W. A Bayesian approach to tracking multiple targets using sensor arrays and particle filters. IEEE Trans Signal Process. 2002;50(2):216–23.
10. Zhong X, Premkumar AB, Madhukumar AS. Particle filtering and posterior cramér-rao bound for 2-D direction of arrival tracking using an acoustic vector sensor. IEEE Sens J. 2012;12(2):363–77.
11. Zhong X, Premkumar AB, Madhukumar AS. Particle filtering for acoustic source tracking in impulsive noise with alpha-stable process. IEEE Sens J. 2013;13(2):589–600.

# Localization of a Mobile Node Using Fingerprinting in an Indoor Environment

Sohaib Bin Altaf Khattak[1], Min Jia[1(✉)], Mir Yasir Umair[2], and Attiq Ahmed[2]

[1] School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China
jiamin@hit.edu.cn

[2] National University of Sciences and Technology, Islamabad, Pakistan

**Abstract.** Localization is an important requirement in today's world, and numerous modern applications require location tracking. An indoor localization of a mobile node using the range-free fingerprinting technique in WLAN environment is presented. The work focuses on improvement in the accuracy of localization using some additional parameters in the fingerprint, along with the conventional received signal strength (RSS). ToA has been used to enrich the fingerprint data for more unique fingerprints. The impact of AP placement on localization accuracy is also addressed. In this paper, a technique is proposed that is not complex to implement using existing infrastructure and is also easy to understand. Significant improvement has been achieved from about 20 to 40% in different scenarios including line of sight and non-light of sight scenarios, small and large areas.

**Keywords:** Indoor localization · WLAN AP placement · LBS for IoT

## 1 Introduction

Localization is the process of finding the position of a specific target, based on some observable phenomenon. It has numerous applications like robot movement, navigating self-driving cars, automatic object location detection. Recently, location-based services (LBS) have a great impact on our everyday lives, it is also related to IoT. The poor performance of GPS indoors and the popularity of WLANs have opened up many opportunities for LBS. Numerous approaches have been proposed that find user locations without the aid of GPS [1]. Among these approaches, Wi-Fi-based localization has gained immense attention, it saves extra infrastructural costs, because the Wi-Fi access points are widely deployed in every building nowadays. Indoor localization method using Wi-Fi can generally be divided into two categories: range based (an approach based on triangulation/ trilateration) and range free (fingerprinting approach).

Location fingerprinting is a technique that matches the fingerprint of certain characteristics of a signal as a function of the position. It has two phases: offline


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_129

and online. In the offline stage, the location coordinates and RSS from APs are collected. In the online phase, a positioning technique uses the currently observed RSS and previously collected information to figure out an online query by the user.

RADAR [2] and Horus [6] systems will be used for comparison with the proposed technique, as these are accepted and famous algorithms in the area, also widely used for comparison. RADAR system is a deterministic algorithm based on nearest neighbor algorithm, and Horus is probabilistic based on maximum likelihood and Bayes' rule. We propose a system which uses a combination of Wi-Fi RSSI values and ToA values for localization based on fingerprinting framework. In the literature, multiple techniques have been used along with location fingerprinting [3,4]. TOA is previously used along with RSS but as range-based technique [3], whereas in this paper, it is taken in range-free method and combined it with the RSS in the fingerprint. AP placement techniques [5] are also used to further improve the accuracy of the system. The paper is divided into the following sections. Section 1 is the introduction , Sect. 2 is the proposed technique, and Sect. 3 is simulation and results. The paper is concluded in Sect. 4. At last, bibliography is given.

## 2 Proposed Technique

This section describes the system model. The system is simulated, and the proposed technique is tested along both deterministic and probabilistic for comparison in different scenarios.

### 2.1 System Model

A 2D indoor environment is considered covered by a certain number of WLAN APs, which are visible throughout the area of interest. To build a radio map, area of interest is divided into grids of 2 m by 2 m. In the offline phase, the readings from Wi-Fi APs are taken at the RPs and stored in the FP database. The point can be represented as (x, y) 2D coordinates on the floor. As shown in Fig. 1, this figure will be further used to test localization algorithm. The notations used are given as follows, and the parameters used in simulation are given in Table 1.

| | |
|---|---|
| $\chi$ = two-dimensional space | $(x, y)$ = coordinates in 2D |
| $\Gamma$ = vector having RSS reading | $\tau$ = vector having ToA readings |
| $N_{ap}$ = number of APs | $Pixelvalue = 2$ |
| $\psi$ = samples of RSS from S | $\xi$ = samples of ToA from T |
| $L_x$ = dimensions of area along x-axis | $L_y$ = dimensions of area along y-axis |
| $Dif_x = [0.5, 2.5, \ldots, L_x]$ | $Dif_y = [0.5, 2.5, \ldots, L_y]$ |

### 2.2 Offline Phase

RSSI can be used to estimate distance because the power of electromagnetic waves is inversely proportional to the distance [7]. Practically, the RSS along with distance depends on other environmental factors also. So the following pathloss model is used:

$$P(d) = P(d_0) + 10.n.\log(d/d_0) + \zeta + W \tag{1}$$

Here, P(d) is the RSS at distance d, $P(d_0)$ is the RSS at the reference distance $d_0$ . 'd' is the distance of the RP from the AP. 'n' is the path loss coefficient, $\zeta$ is the random noise element, $W$ is the wall attenuation factor. Multiple number of samples for RSS had to be taken, as values of RSS change at every instant. All the values of RSS for a certain RP from a certain AP are averaged and are stored in the fingerprint database.

$$RSS_{ij}(avg) = \frac{\sum_{n=1}^{N} RSS_{ij}}{N} \tag{2}$$

Time of arrival (TOA) of each AP is also calculated at all RPs in the proposed technique and stored in the database along with the RSS values to enrich the fingerprint data. The values for TOA in the simulation are found by using the formula $s = v.t$, where $s$ is the distance, $v$ is the velocity, and $t$ is the time. This relation has been used as $t = d/c$, here $d$ is the distance between the AP and RP,$c$ is the propagation velocity of the signal, and $t$ is the TOA for that specific RP. Due to the scattering and multipath effect 'd' is not accurate, so a random noise is added as $d = d_i + d_n$. Using the values of RSS and TOA together in fingerprint, certain weight($W_n$) has to be assigned to TOA values because there is a large difference among the values of these parameters which can lead to wrong estimation during matching. So, the ToA calculation can be mathematically represented by the following equation:

$$ToA = \frac{\sum_{i=1}^{N} ToA(n)}{N} * W_n \tag{3}$$

The system detects the walls in the channel between the AP and RP and includes the effect in the propagation model for RSS calculation.

The training data in the offline phase is stored and then database of radio map will have the form:

$$\lambda = \begin{bmatrix} x_1, y_1 & \Gamma_{1,1} = (\psi_1, .., \psi_{Nap}) & \tau_{1,1} = (\xi_1, .., \xi_{Nap}) \\ : & ... & : \\ x_{L_x}, y_{L_y} & \Gamma_{Lx,Ly} = (\psi_1, \ldots, \psi_{Nap}) & \tau_{Lx,Ly} = (\xi_1, \ldots, \xi_{Nap}) \end{bmatrix}$$

Order of the matrix is $\alpha$ * $\beta$ , where $\alpha$ is the number of data points and $\beta$ is dimensionality.

### 2.3 Online Phase

A query is put to the system having the values (RSS and TOA), and the system provides the estimated location by using matching technique, the most similar match to the query is given as the estimated location and coordinate points are returned. One thousand random test points are created in the simulation, where the matrix (R, S) has x, y coordinates of the generated test points. The query matrix is also saved in the form:

$$Q = [(R_i S_i), \Gamma_i = (\psi_1, \ldots, \psi_{Nap}), \tau_i = (\xi_1, \ldots, \xi_{Nap})] \tag{4}$$

The two matrices $\lambda$ and Q will be matched for location determination. For the matching purpose, the K-nearest neighbor (KNN) algorithm is used. The KNN algorithm basically calculates the distance of the query from the fingerprinting database, selects the K-nearest neighbors, and gives the averaged value.

Both RSS and ToA are used in the proposed technique, so the combined vectors will be used,

$$\sqrt{(Query_t - DataBase_f)^2} \tag{5}$$

So, $(Query_t)$ is the query, by the user that wants to find its location, and $(DataBase_f)$ is the vector from fingerprint database. K-nearest neighbors are selected on the basis of minimum distance, their coordinate points are then averaged, and the position is estimated.

$$(x, y) = \frac{1}{k} \sum_{i=1}^{K} (x_i, y_i) \tag{6}$$

(x, y) is the estimated position, K is number of nearest neighbors, $(x_i, y_i)$ are coordinate points of the K-nearest neighbors. The root mean square error (RMSE) is calculated to check the overall performance of the system,

$$RMSE = \sqrt{\sum (P_R - P_{Es})^2} \tag{7}$$

$P_R$ is the real position coordinates, and $P_{Es}$ are the estimated position coordinates.

In the simulation for KNN implementation and location determination, the following procedure was adopted.

KNN ($\lambda$ , Q)

$$Q - (R_i S_i) = [\Gamma_i = (\psi_1, \ldots, \psi_{Nap}), \tau_i = (\xi_1, \ldots, \xi_{Nap})] = \omega \tag{8}$$

$$(R_i, S_i) = Actualtestlocations = L \tag{9}$$

$$D = \sqrt{\sum (\omega_i - \lambda_i)^2} \tag{10}$$

$$D_{sort} = argmin \sum_{i=1}^{1000} \sqrt{\sum (\omega_i - \lambda_i)^2} \tag{11}$$

$$D_{sort} = [d_{min}, \ldots, d_{max}] \tag{12}$$

$$Sort = Indexes(D_{sort}) \tag{13}$$

$$EstimatedPosition = \frac{\sum P(x), P(y)}{K} \tag{14}$$

$$EstimatedPositions(i = 1:1000) = E = P(x_i), P(y_i) \tag{15}$$

$$Error = \varepsilon = \sqrt{\sum (E_i - L_i)^2} \tag{16}$$

After the results from KNN and the position estimation tests performed for 1000 random queries, their cumulative density function (CDF) of Error was taken as a metric for precision of the system.

The values of parameters used in simulation are shown in Table 1.

**Table 1.** Parameters used in simultaion

| Parameter | Value |
|---|---|
| Path loss exponent $n$ | 3 |
| Wall attenuation factor $WAF$ | 2 |
| Signal propagation speed $c$ | 299792458 |
| Reference distance $d_0$ | 1 |
| Power at d0 $P(d_0)$ | $-30$ |
| Transmission power $Ptx$ | 10 |
| $K$ in KNN | 5 |
| Grid size | 2*2 |
| No. of position queries | 1000 |

## 3 Simulations and Results

In this section, the simulation scenario is explained and discussed with the obtained results. The RSS-based localization techniques and the proposed technique are tested in different indoor scenarios, and results are compared. The number and placement of AP are also addressed.

### 3.1 Case A and Case B

The first scenario is mapped upon an area of $50 \times 50$ (m) in which there is a LOS distance between the APs and the RPs and there are four APs placed at the four corners of the simulation area. See Fig. 1. The results were very encouraging, the proposed technique achieved 21% improvement. The CDF plot is shown in Fig. 2 and the detailed results in Table 2.

Second scenario considers an area of $100 \times 100$ m. The number of APs is 5. See Fig. 3. The proposed technique gave better results as compared to the simple RSS-based localization techniques. CDF plot is shown in Fig. 4 and results in Table. 2.

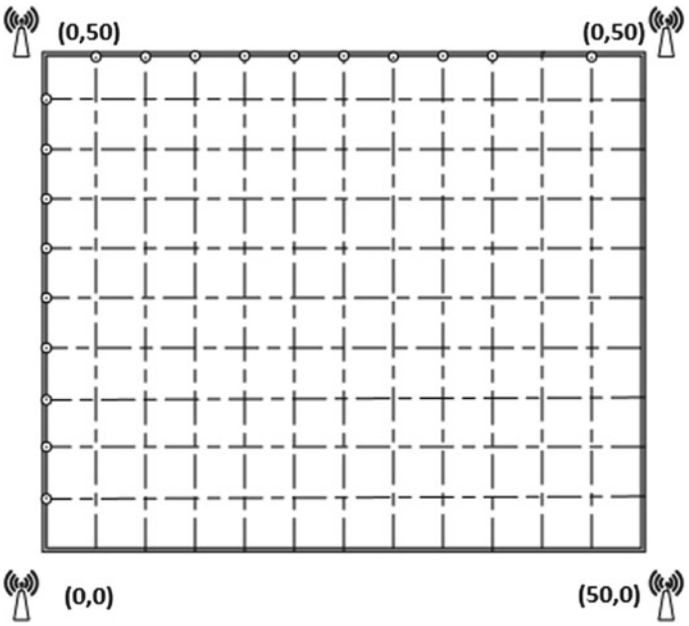



**Fig. 1.** Case A: 4 APs 50 m by 50 m area

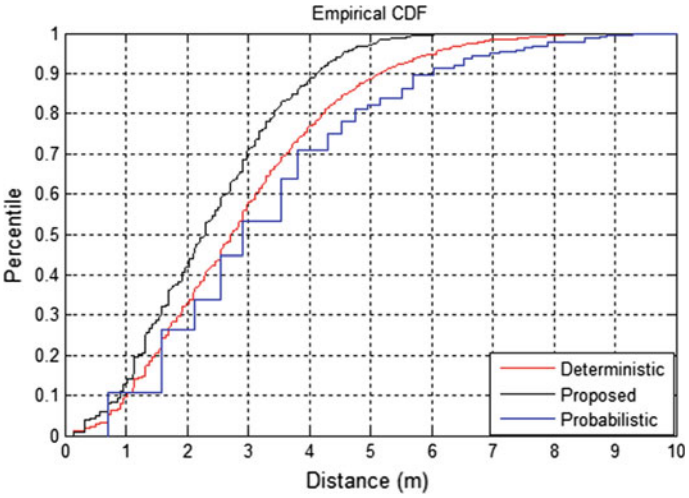



**Fig. 2.** Simulation results for case A

**Table 2.** CDF plot case A and B

| Case | Technique | CDF (error) at 0.7 (m) |
|---|---|---|
| A | Probabilistic | 3.8 |
| A | Determinisric | 3.6 |
| A | Proposed | 2.9 |
| B | Probabilistic | 6 |
| B | Deterministic | 5.4 |
| B | Proposed | 3.3 |

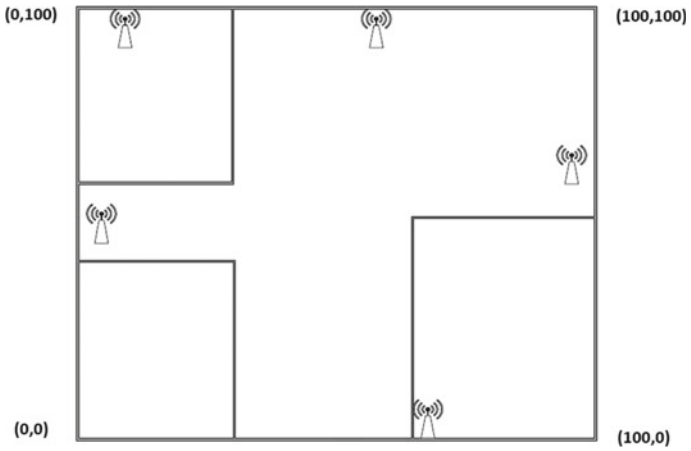



**Fig. 3.** Coordinates of 5 APs deployed (Case B 5 APs 100 m by 100 m Area)

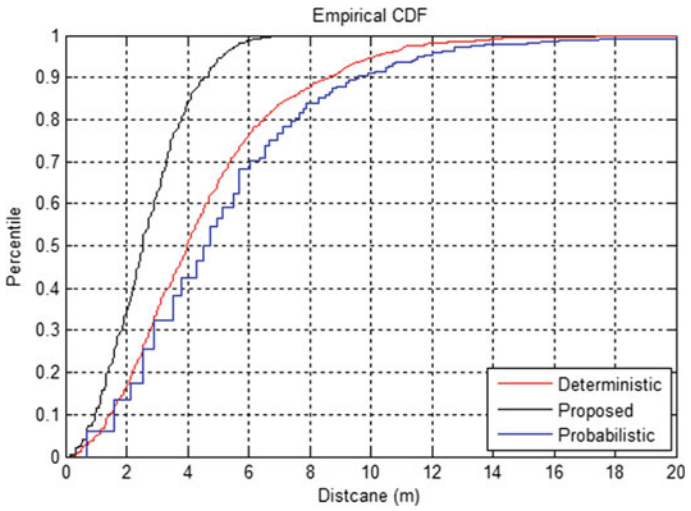



**Fig. 4.** Simulation results of case B (5 APs 100 m by 100 m area)

### 3.2 Impact of Number of APs and Placement

Increasing the number of APs to 8, the impact of number of APs and their placement on accuracy will be checked. Eight APs are deployed in the same model used in Case B as shown in Fig. 3. The APs are deployed first randomly, then two times in different symmetries. The coordinates of the APs can be seen in Table 3, and detailed results in Table 4. The CDF plots of the three AP deployment configurations are shown in Figs. 5, 6, and 7, respectively.

It can be seen from the results that placement of APs has an impact on the accuracy of the system, and it can also be observed that increasing the number of APs can have a positive impact on the accuracy. But one thing has to be noted that the proposed technique is not much prone to changes due to APs placements as compared to the RSS-based system and the results are consistent over all scenarios.

**Table 3.** APs coordinates

| AP number | Random | Symmetry 1 | Symmetry 2 |
|---|---|---|---|
| 1 | 0, 0 | 0, 0 | 13, 25 |
| 2 | 0, 50 | 0, 100 | 13, 75 |
| 3 | 31, 100 | 100, 100 | 37, 25 |
| 4 | 100, 0 | 100, 0 | 37, 75 |
| 5 | 50, 50 | 35, 20 | 63, 25 |
| 6 | 75, 80 | 35, 80 | 63, 75 |
| 7 | 100, 20 | 75, 20 | 87, 25 |
| 8 | 50, 0 | 75, 80 | 87, 75 |

**Table 4.** Error comparison and results

| AP placement strategy | Technique | CDF (error) at 0.7 |
|---|---|---|
| Random APs placement | Probabilisitc | 6.3 |
| | Deterministic | 5.3 |
| | Proposed | 3.6 |
| APs placement symmetry 1 | Probabilisitc | 5.7 |
| | Deterministic | 5.1 |
| | Proposed | 3.2 |
| APs placement symmetry 2 | Probabilisitc | 3.8 |
| | Deterministic | 3.5 |
| | Proposed | 2.8 |

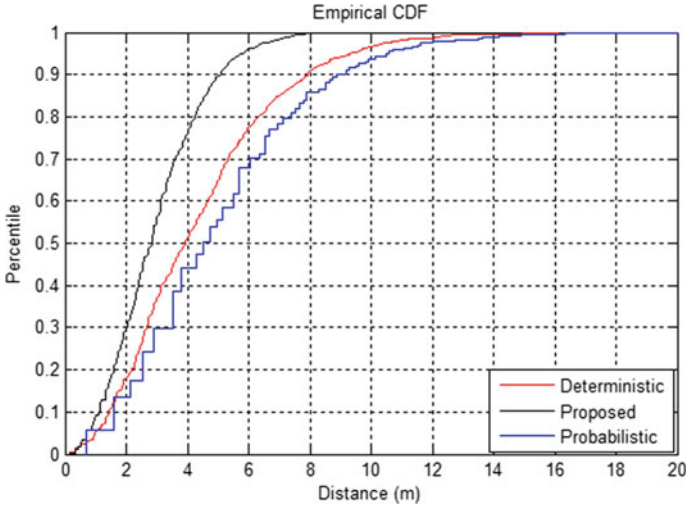



**Fig. 5.** CDF plot (8 APs randomly deployed)

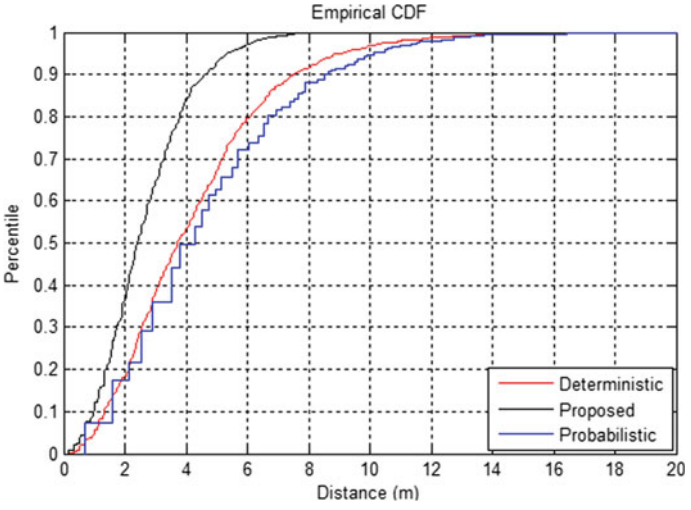



**Fig. 6.** CDF plot (8 APs deployed in symmetry 1)

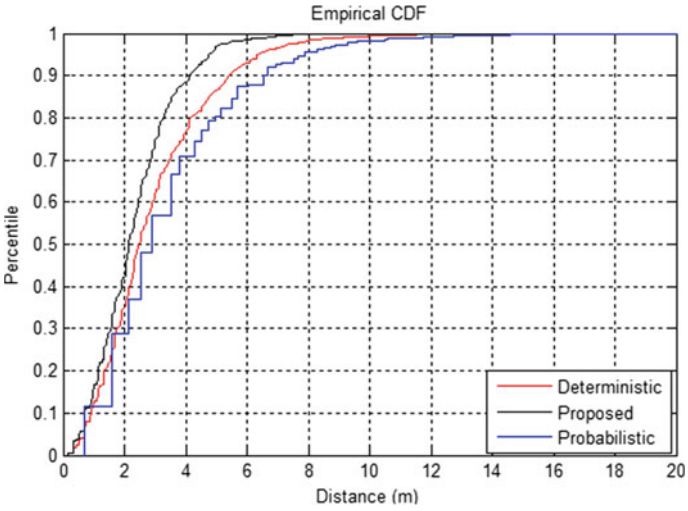



**Fig. 7.** CDF plot (8 APs deployed in symmetry 2)

## 4 Conclusion

The proposed technique that uses an additional parameter of ToA in fingerprint shows significant improvement in overall results. It can be concluded that utilizing other location dependent parameters in WLAN or any other wireless network can improve the localization accuracy. The impact of AP placement is also addressed and the results show changes due to placement. The proposed technique is consistent overall and provides much better results. The idea can be taken further by investigating other location dependent parameters for fingerprinting and AP placement can be also investigated, relations depending on the characteristics of the signal and geometry of the area can provide optimal AP locations.

**Acknowledgments.** This work was supported by the National Science Foundations of China (No.61671183 and 61771163).

## References

1. Liu H, Darabi H, Banerjee P, Liu J. Survey of wireless indoor positioning techniques and systems. Ieee Trans Syst Man Cybern Part C: Appl Rev. 2007;37(6):6.
2. Bahl P, Padmanabhan VN. RADAR: an in-building RF-based user location and tracking system. In: IEEE INFOCOM 2000, 2000.
3. Torteeka P, Chundi X, Dongkai Y. Hybrid Technique for indoor positioning system based on wi-fi received signal strength indication. In: 2014 international conference on indoor positioning and indoor navigation, 27th–30th October, 2014.
4. Machaj J, Brida P. Using of GSM and wi-fi signals for indoor positioning based on fingerprinting algorithms. Inf Commun Technol Serv. 2015;13(3).

5. Chen Y, Francisco J-A, Trappe W, Martin RP. A Practical approach to landmark deployment for indoor localization IEEE SECON 2006 proceedings.
6. Youssef M, Agrawala A. The horus WLAN location determination system. In: MobiSys05: proceedings of the 3rd international conference on Mobile systems, applications and servicesl, USA; 2005. p. 205–18.
7. Rappaport TS. Wireless communications: principles and practice. New Delhi: Prentice-Hall of India; 2003.

# An Improved State Coherence Transform Algorithm for the Location of Dual Microphone with Multiple Sources

Shan Qin and Ting Jiang(✉)

Beijing University of Posts and Telecommunications, Haidian District, Beijing 100000, China
tjiang@bupt.edu.cn

**Abstract.** This paper proposes a new kernel function in state coherence transform to perform multiple time difference of arrival estimation in order to increase the resolution of location in frequency-domain blind source separation. The state coherence transform associated with each source generalizes the GCC for multiple sources and generates envelopes with clear peaks corresponding to the maximum-likelihood TDOAs. However, the weight allocation of the kernel function is unreasonable for small spacing microphones. We propose an improved kernel function to enhance the resolution of small values, which means that a larger weight allocated to smaller values. Experimental results show that the proposed approach allows to separate four speakers, using very short utterances, in highly reverberant environment even with small-spaced microphones of 2 cm.

**Keywords:** Nonlinear weighting compensation · State coherence transformation · Blind source separation

## 1 Introduction

During the last thirty years, a huge number of methods of blind source separation have been proposed but separation in real life with small microphone pairs is still a challenging problem [1, 2]. The estimation of time difference of arrival (TDOA) is an essential step in several approaches for blind source separation [3]. In the indoor reverberation environment, frequency-domain approaches outperform the time-domain methods in computing complexity performance and convergence property, which is the most investigated like independent component analysis [4] and nonnegative matrix factorization [5]. However, the removal performance of internal permutation ambiguity directly affects the separation effect. The separation perform on small-spaced microphone pair is still lag, since the accuracy of time delay estimation is limited for microphone space, resulting in permutation problem. Among the most promising ones, a robust way to solve the permutation problem is to apply state coherence transform to increase the inter-source TDOA resolution [6]. However, this approach has a poor performance for the case of separation with small-spaced microphone pairs when reverberation and spatial aliasing existing.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_130

In this work, we present a new kernel function to improve SCT solving the permutation problem on small-spaced microphone pairs, which is robust both to spatial aliasing and to reverberation even for high $T_{60}$. The principle of time delay estimation is analyzed in Sect. 2, and the real data are used for experimental verification. Finally, it confirms that the proposed improved method performs effectively allowing the separation with microphone spaced at 2 cm in highly reverberant environments.

## 2 Physical Interpretation of SCT

To simplify the understanding of the proposed improved approach, we first give a simple physical interpretation of the TDOA and SCT. For simplicity, we analyze a number of source N of 1 and the number of the microphones M of 2. A schematic diagram is shown in Fig. 1.

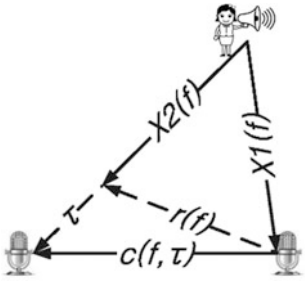



**Fig. 1.** Wave propagation model of source

It is known that $X_1$ and $X_2$ are signals received by microphones, and the microphone spacing is represented by $c(f, \tau)$. The TDOA of the source relative to the microphones is represented by $\tau$. $r(f)$ represents the phase difference of two observed acoustic propagation from the source to the microphone pair $(a, b)$. Under ideal conditions, the model of the inter-microphone delay for a given source can be represented by $c(f, \tau)$. According to the mathematical relation in the graph, TDOA is expressed as

$$\vec{\tau} = \overrightarrow{c(f,\tau)} - \overrightarrow{r(f)} \tag{2.1}$$

Furthermore, as long as the acoustic waves related to the direct propagation paths are relatively strong compared with the reflection paths, a likelihood measure of the TDOAs for multiple sources can be estimated. TDOAs estimation associated with each source can be performed by minimizing the following quantity [7]:

$$\overrightarrow{\tau_k} = \operatorname{argmin} \sum_{N} \sum_{M} \| c(f,\tau) - r(f) \| \tag{2.2}$$

SCT will be maximized for values of t that for each frequency minimize the sum of the above distances. In practice, given the coherence of the states across the frequency, the state coherence transform is formulated as follows

$$\mathrm{SCT}(\tau) = \max \sum_{N} \sum_{M} \left[ 1 - g\left( \frac{\|c(f,\tau) - r(f)\|_2}{2} \right) \right] \tag{2.3}$$

Since $\vec{\tau} \leq \overrightarrow{c(f,\tau)}$, if we do not perform any nonlinear transformation of the Euclidean distance with the decreasing of microphone space, the inter-source TDOAs will be too small to be distinguished with decreasing resolution, which makes the peak selection more sensitive to errors. $g(\cdot)$ is a generic nonlinear function which operates as a weighting function and allocates more weight to the small value τ that corresponds to the maximum likelihood of propagation parameters of each source. The common nonlinear monotonic function has been described in [8] as

$$g(\cdot) = \tan\, h(\alpha x) \tag{2.4}$$

Although the SCT is effective in normal situations, it lacks reliability if speech separation is analyzed with small microphone spacing at 2 cm or smaller. $g(\cdot)$ function has not enough resolution to distinguish multiple peaks in small space condition, which may cause the separation performance of small microphone array decline.

## 3 Proposed Kernel Function

Although the SCT in estimation of TDOAs is more accurate than GCC, the decreasing resolution reduces the accuracy of estimation [7]. When the distance between microphones decreases, the reduction in resolution may be occurred. Improving the ability of $g(\cdot)$ in separating multiple local peaks may enhance the resolution.

In this work, we propose an improved function with reasonable performance by using a locally confined kernel function that enhances the resolution on $t \approx 0$, which implicitly resolves the permutation problem and implements the estimation of TDOAs of multiple sources with small spacing microphone pair.

$$g(\cdot) = 1 - \tan\, h(\alpha x) \tag{3.1}$$

Figure 2 shows the graph of different kernel functions with $\alpha = 8$. Abscissa represents time delay, and ordinate represents the weight assigned to distance. Compared from the diagram, it can be inferred that SCT does have a resolution magnification for the distance, but the allocation of weights under the smaller size is not reasonable. The smaller space requires greater weights to enhance the resolution of the delay estimation. The higher weights are allocated to $\tau$ when $\tau$ has small values when the kernel function $g(\cdot) = 1 - \tan\, h(\alpha x)$ with better performance than $g(\cdot) = \tan\, h(\alpha x)$, and vice versa. The removal performance of internal permutation ambiguity outperforms the SCT. When $N = 2$, the improved SCT formula is shown as follows:

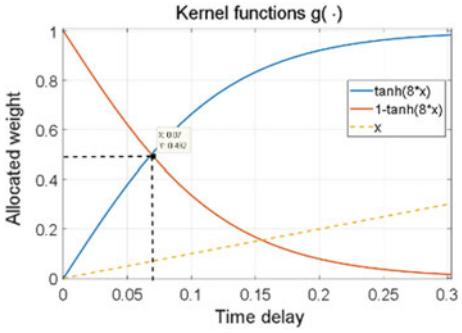



**Fig. 2.** Values of different kernel functions with $\alpha = 8$

$$\mathrm{ISCT} = \left(1 - \left(1 - \tan h\left(\frac{\alpha\|c(f,\tau) - r_1(f)\|_2}{2}\right)\right)\right) + \left(1 - \left(1 - \tan h\left(\frac{\alpha\|c(f,\tau) - r_2(f)\|_2}{2}\right)\right)\right) \tag{3.2}$$

And as known that $\|c(f,\tau) - r(f)\|_2 = \sqrt{[c(f,\tau) - r_2(f)][c(f,\tau) - r_2(f)^*]}$, so the above formulator can be expressed as:

$$\mathrm{ISCT} = \max \sum_N \sum_M \tan h\left(\alpha\sqrt{1 - \mathrm{Re}[c(f,\tau) * r(f)}\right) \tag{3.3}$$

In previous research, it was proved that $\mathrm{GCCPHAT}(\tau) = \mathrm{Re}[c(f,\tau) * r(f)]$ [6]. The (3.4) describes the relationship between ISCT and GCCPHAT, which indicates that the essence of SCT is to assign different weights to the value of GCC and enhance the resolution of location to estimate the TDOA more accurately:

$$\max \mathrm{ISCT} = \tan h\left(\alpha\sqrt{1 - GCC}\right) \tag{3.4}$$

Regarding the values of parameter $\alpha$, it is decided by doing experiments or experiences of researchers. In this paper, we propose a constraint function to restrict $\alpha$:

$$\alpha(d) = \begin{cases} \frac{1}{5d} & (0 < d \leq 0.2) \\ 1 & (0.2 < d \leq 1) \end{cases} \tag{3.5}$$

where $d$ is the microphone space in meters.

Inverse proportional sequence has a good convergence property. Monotonic bounded sequence must converge in the defined domain. Therefore, the (3.4) can obtain the local extreme values which represent the estimation of the TDOAs of multiple sources.

## 4 Experiments and Results

In this work, we solve the removal permutation problem according to the estimated TDOA and the optimization in formula (3.4). The parameter $\alpha$ is set according to the (3.5). The nonnegative matrix factorization algorithm has been applied to separate three sources with distances as shown in Fig. 3. Speakers place about 1.2 meter away from the array with the microphones spacing of $d = 2$ cm and the room impulse response with $T_{60} = 600$ ms. The separation performance is evaluated by waveform. The STFT was implemented with hamming window at the length of 1024 samples and 50% overlap. The speech collected by the microphone array is considered as experimental data. The voice acquisition module is shown in Fig. 4.

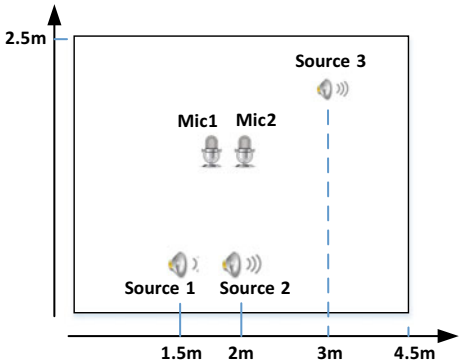



**Fig. 3.** Recording scene of the speech data

**Fig. 4.** Circuit board diagram of the microphone array

In order to demonstrate the ability of the proposed ISCT methods, we used the GCC-NMF algorithm to obtain the estimation of the TDOAs and separated sources. The voice acquisition module consists of six microphones spaced of 2 cm to each one. Here, we only keep the voice data of two adjacent microphones a and b. We compare our proposed algorithm ISCT with SCT.

Figure 5 describes the results estimation of TDOA of SCT, and Fig. 6 describes the result of ISCT. The abscissa is the angular spectrum, and the ordinate is the sum of the SCT in all frequencies. The solid line indicates the value of SCT and the dotted lines are the local maximum value, that is, the time difference we estimate. Compared with the two figures, the location resolution of ISCT algorithm is higher than SCT algorithm.

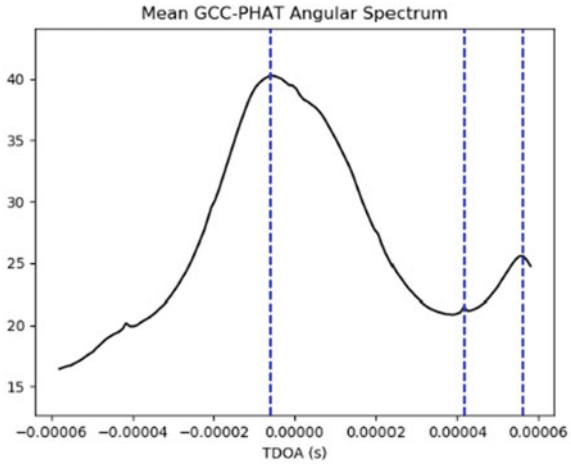



**Fig. 5.** SCT angular spectrum

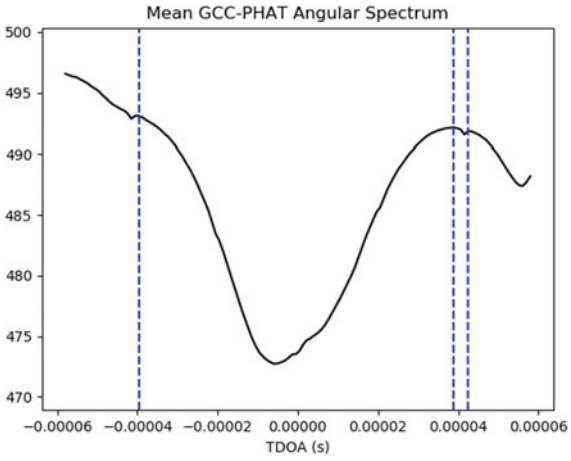



**Fig. 6.** ISCT angular spectrum

Since our experimental data are actually measured data, there is no pure speech contrast to calculate SIR. Comparing the waveform of the actual speech with the waveform of the separated speech is selected to evaluate the performance. The performance of the two algorithms is shown in the following waveform diagrams.

Figure 7 is the waveform of the separation results of ISCT. Figure 8 is the waveform of the separation results of SCT. The separation effect based on ISCT

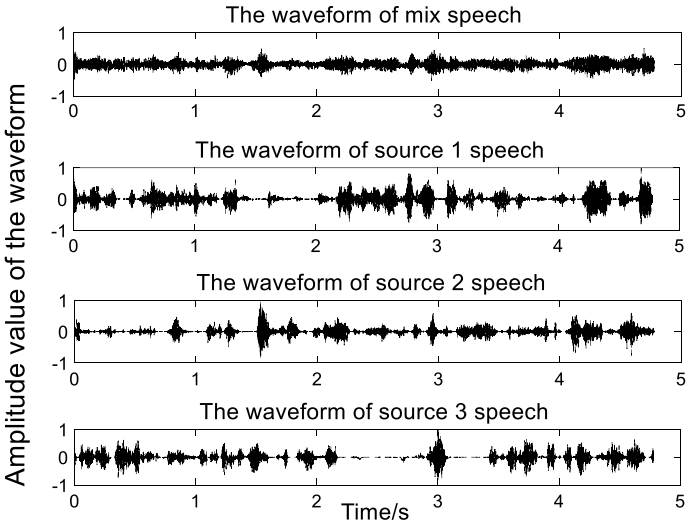



**Fig. 7.** Waveform of separated speech of ISCT

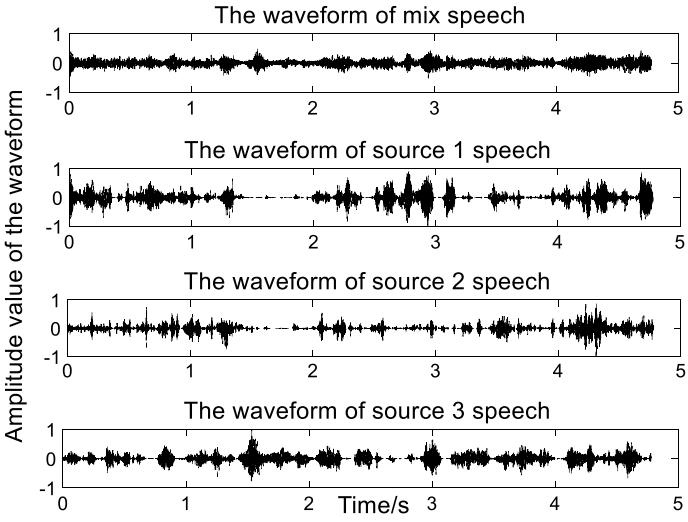



**Fig. 8.** Waveform of separated speech of SCT

outperforms the SCT. Compared with the separation results in Figs. 7 and 8 shows only two separation sources, which indicates that the separation effect based on ISCT outperforms the SCT.

ISCT performs better when the size of the microphone array is very small about 2 cm or smaller. If the size is changed to 5 cm or bigger, the effect of SCT is obvious. Therefore, different nonlinear functions can be selected for compensation according to the size of different microphone arrays as follows:

$$\text{maxSCT}(\tau) = \begin{cases} \tan h\left[\alpha(d)\sqrt{1 - GCC(\tau)}\right] & 0 < d \leq 0.04 \\ 1 - \tan h\left[\alpha(d)\sqrt{1 - GCC(\tau)}\right] & 0.05 \leq d < 1 \end{cases} \tag{4.1}$$

## 5 Conclusion

In this paper, we analyze the principle of estimation of TDOA. Spatial resolution decreases with decreasing size of microphone array if no nonlinear compensation is made. Compared with the nonlinear function of SCT, the new kernel function proposed in this paper will allocate a larger weight to smaller values. Therefore, there will be a higher spatial resolution in the case of small microphone spacing to obtain more accurate positioning accuracy, solve the problem of arrangement, and get the purpose of blind source separation. We also propose constraint functions for the parameters $\alpha$ according to the microphone space. Experimental results show that the proposed approach allows to separate four speakers, using very short utterances, in highly reverberant environment even with small-spaced microphones of 2 cm.

**Acknowledgments.** This work was supported by National Natural Science Foundation of China (NSFC) (No.61671075) and Major Program of National Natural Science Foundation of China (No.61631003).

## References

1. Hosseini MS, Rezaie A, Zanjireh Y. Time difference of arrival estimation of sound source using cross correlation and modified maximum likelihood weighting function. Sci Iran. 2017;24(6).
2. Jia RS, Gong Y, Peng YJ, Sun HM, Zhang XL, Lu XM. Time difference of arrival estimation of microseismic signals based on alpha-stable distribution. 2017; p. 1–17.
3. Zhu H, Li Z, Cheng Q. Sound source localization through optimal peak association in reverberant environments. 2017; p. 1–6).
4. Mirzal A. NMF versus ICA for blind source separation. Adv Data Anal Classif. 2017;11 (1):25–48.
5. Wood SU, Rouat J. Real-time speech enhancement with GCC-NMF. In: INTERSPEECH. 2017
6. Nesta F, Omologo M. Generalized state coherence transform for multidimensional TDOA estimation of multiple sources. IEEE Trans Audio Speech Lang Process. 2012;20(1):246–60.

7. Azadi M, Abutalebi HR. Modified state coherence transform to reduce spatial aliasing in TDOA estimation of multiple sound sources. In: International symposium on telecommunications. IEEE; 2015. p. 492–6
8. Nesta F, Omologo M. Generalized state coherence transform for multidimensional localization of multiple sources. In: Applications of signal processing to audio and acoustics, 2009. WASPAA '09. IEEE workshop on Vol.4. IEEE; 2009. p. 2360–71

# Route Navigation System with A-Star Algorithm in Underground Garage Based on Visible Light Communication

Ying Yu, Jinpeng Wang(✉), Xinpeng Xue, and Nianyu Zou

School of Information Science and Engineering, Dalian Polytechnic University, No. 1st Qinggongyuan, Ganjingzi, Dalian 116034, Liaoning, China
wangjp@dlpu.edu.cn

**Abstract.** In order to solve the problem of parking lot difficulty, communication security and low efficiency of garage, with the garage using LED lighting, designed an underground garage navigation system based on visible light communication, which uses the lighting system in the garage to realize the real-time monitoring of the parking space and the navigation of the vehicle. This paper designed the navigation system of underground garage based on the principle of visible light communication in the garage with LED lighting. Result shows that the system achieves signal transmission in the range of 4 m, which can meet the needs of vehicle navigation.

**Keywords:** Visible light communication · Underground garage · A* algorithm · LED

## 1 Introduction

In big cities, particularly in megacities, serials of problems, including the city traffic jam, environment pollution, and easy-angry drivers, etc., are caused by the parking problem. In the traditional way of parking, most people choose the parking space blindly, and some parking lots have to use manual guidance to enter the parking space aiming at the problem of parking [1]. However, the manual guidance has caused the waste of human resources. Some underground parking lots use the local sound control [2] method to guide the vehicle to stop, which is considering that the garage cannot receive the outside signal, and the ordinary navigation cannot be used in the underground garage. However, for the open environment of underground garage, when multiple vehicles enter the parking garage, the way of voice control can easily cause confusion for the garage [3]. For an underground garage, it is difficult for the owner to quickly find the best parking space without good guidance, thus reducing the efficiency of the garage.

This paper uses visible light communication technology (VLC) [4, 5], sensing technology, and the A*(A-Star) [6–8] search algorithm, to give a new underground garage parking mode. The photoelectric detector, consisted of LED lamps above the parking lots and photosensitive resistors on the ground, is used to detect the condition of parking lots; when vehicles enter the garage, the A* algorithm is selected to


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_131

calculate the optimal path according to the vacant condition of the parking space. The system will use A* algorithm to find optimal path giving to the drivers according to the condition of vacant parking space. The path, which can be used to guide, could be expressed as square waves with different frequency generated via single chip. When the system works, a modem demodulates signals to LED, then receivers transform optical signals to different frequency electrical signals, and demodulates the electrical signal to the LCD screen through the modem [9]. After processing, LCD displays the path information, which can prompt the driver's driving direction and improve the efficiency of the garage.

## 2 The Plane Model of the Underground Parking Lot

The system has to be abstracted the entity appearance of the garage into a plane model, so it can determine the vehicle's running path. The A* algorithm, most effective direct search algorithm, is suitable for path planning. This paper used the A* search algorithm to find the optimal path, and the plane model of the garage is shown in Fig. 1.

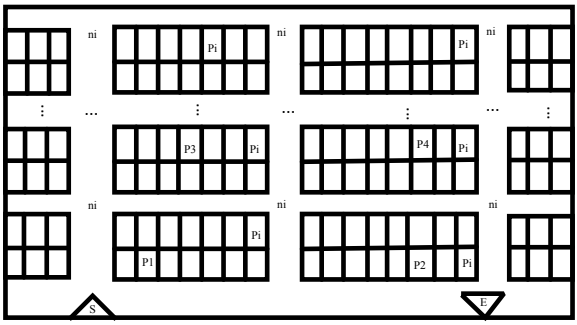



**Fig. 1.** Underground garage plane model

The principle of A* search algorithm is to design an evaluation function: $f(n) = h(n) + g(n)$ In this function:

$f(n)$ the evaluation function from the initial point s to the target point $P$ through the node $n$;

$h(n)$ the real cost of from the starting node '$s$' to the current node '$n$' in space;

$g(n)$ the estimated cost from the current node '$n$' to the destined node '$p$'. The evaluation function is used to evaluate each point that can be reached at the next step of the current position, and the system can find the smallest point of the value $f(n)$ as the path node by searching each step. The algorithm must restart when the path is forwarded one step.

The specific steps of the A* algorithm to calculate the path:

1. Create a 'start' table and a 'route' table. The two tables are initiated to empty, and the starting point 's' is placed in the 'start' table;
2. Search for nodes in 'start' table. If 'start' table is empty, it means that no path is found, and fail to search;
3. If the 'start' table is not empty, this system will select a node with the minimum value 'f' to be the optimal node, which is denoted as 'm' and put into the 'route' table;
4. Determine whether node 'm' is the target node 'e,' and if node 'm' is the target node, a path is found successfully;
5. If the node 'm' is not the target node, then it is extended to generate the sub-node $m_1$, $m_2$ … For each child node, the following procedure is performed (with the example of child node '$m_1$'):
   - If '$m_1$' is already in the 'start' table, calculate '$g(m_1)$,' and the original node '$m_1$' in the 'start' table is called node 'oa,' compared with '$g(m_1)$' and 'g(oa).' If $g(m_1) < g(oa)$, the parent pointer of 'oa' is modified to 'm,' and the 'g(oa)' value is corrected, and the smaller value '$g(m_1)$' is assigned to 'g(oa),' and the corresponding update 'f(oa)' value; If '$g(m_1)$' is greater than 'g(oa)', then the extension node is stopped;
   - If '$m_1$' is already in the 'route' table, then this system skips this node and continues to extend the other nodes;
   - If '$m_1$' is not in the 'start' table and not in the 'route' table at the same time, then put it in the 'start' table, and add a pointer to '$m^1$' to its parent node 'm,' calculate $g(m_1)$;
6. Turn to step (2) and continue circulation until the solution is found or no solution is left.

The block diagram of the algorithm is shown in Fig. 2.

## 3 Experimental Model of Communication System

The VLC system [10] needs the following steps to accomplish: Firstly, the information that needs to be transferred is converted into an electrical signal. Then, the signal is processed by the transmitter circuit, the signal is loaded to the LED through the drive circuit, and the LED sends out the light wave with information after the modulation. Finally, the modulated optical signal is converted to electrical signal through photoelectric detector. After receiving the terminal circuit, the electrical signal is restored to the data information. The principle block diagram of the system work is shown in Fig. 3.

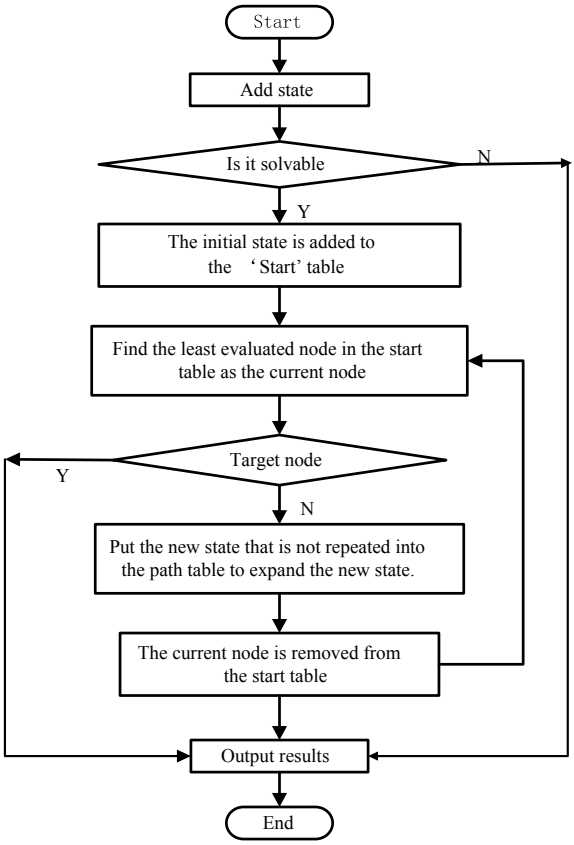



**Fig. 2.** Algorithm flowchart

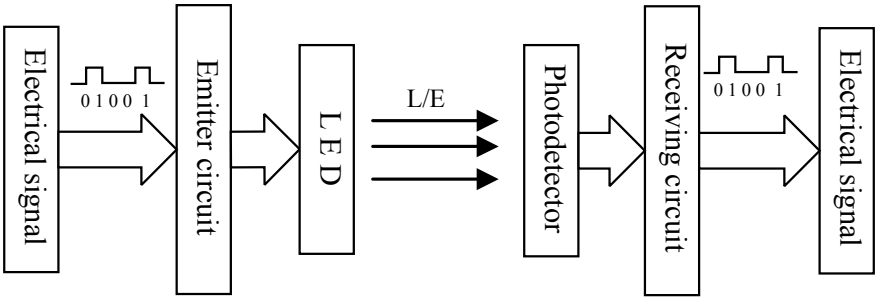



**Fig. 3.** System block diagram

### 3.1 Parking Information Monitoring

MSP430 control chip is used for the acquisition of parking information, and the monitoring of the parking lot is realized by the circuit of LED light source,

photosensitive resistor, power supply, and voltage comparator. The photosensitive resistor is placed in the center of the parking space. The LED light source is placed above the photosensitive resistance, when the car is parked in the parking space, the light signal is blocked, the resistance value of the photosensitive resistor increases, and the voltage on both ends will increase. The signals produced by the amplifier are transferred to the voltage comparator and converted into high and low-level signals that can be identified by the single-chip microcomputer. Display on LCD screen after processing by single-chip microcomputer, the red block indicates that the car is parking, and the green block indicates no parking, in order to monitor the parking space.

### 3.2 The Signal Transmitter

After calculating by the A* algorithm, the system is used in expressing the path of vehicle planning by means of different frequency square wave signals produced by MSP430 and modulates the generated square wave signal to the LED drive circuit. Using the modulation characteristics of the high-frequency flicker of the LED light source, the direction information of the vehicle is loaded into the light wave and radiated in the free space, in which the structure diagram of the transmitter is shown in Fig. 4.

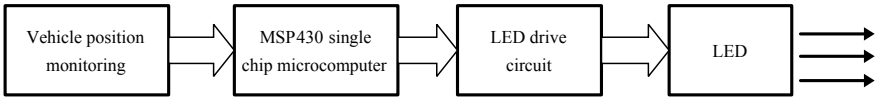



**Fig. 4.** Structure of the transmitter

The LED drive circuit is the main part of the transmitter. Combined with the actual garage situation, the power of LED lighting lamps is very large. In order to ensure the voltage constant at both ends of the lamp, the current drive LED method is adopted to ensure the normal work of the LED. The driving module takes the LM324 operational amplifier as the core driver and converts the voltage change of the front-stage circuit into the current change of the LED through the V/I conversion, in order to provide the normal work of the LED light source, and the LED light source drive circuit is shown in Fig. 5. At the same time, the regulation of variable resistance Rf can change output, and the circuit response speed is fast.

### 3.3 The Signal Receiver

The signal receiver is mainly composed of photodetector, second-order active preamplifier, filter circuit, and MSP430 single-chip microcomputer. In the optical receiving terminal changes the optical signals into electrical signals, which is amplified by the preamplifier circuit and sent to the single chip by serial port. Finally, the driving information is displayed on the LCD screen, and Fig. 6 shows the structure of the receiver.

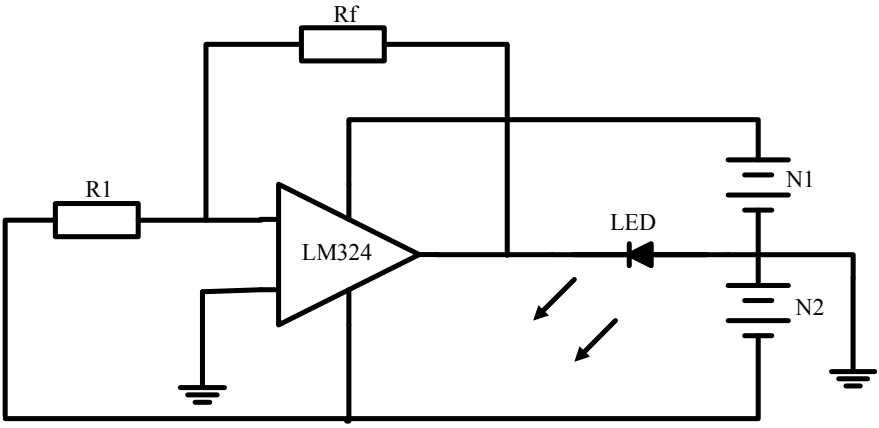



**Fig. 5.** LED light source drive circuit

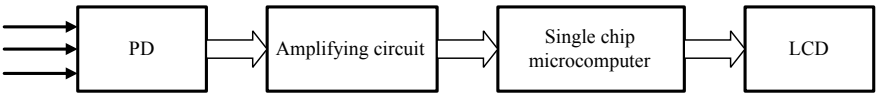



**Fig. 6.** Structure of the receiver

The preamplifier circuit of the receiver is the key part of the recovery of the optical signal and the sensitivity of the butt end has a direct influence, and it will also affect the transmission quality of the signal. The preamplifier circuit is based on the OP37 device, which has the advantages of extremely low input bias current and voltage noise. It is an ideal choice for high-speed trans-impedance amplifier and high impedance sensing amplifier. Figure 7 shows the trans-impedance preamplifier circuit and the two-order active filter circuit. When the receiver receives light at different frequencies, there will be a current Ir in the silicon photocell. Through the preamplifier circuit, the weak current signal will be amplified and converted to the voltage signal. The signal generated by the filter circuit is processed into a standard voltage signal and input to the single-chip microcomputer.

### 3.4 Modulation and Coding

In visible light communication system, there are mainly OOK modulation, PPM modulation, DMT modulation, OFDM modulation, and so on. In terms of the complexity of circuit design, the advantages of OOK modulation and PPM modulation are obvious, while DMT modulation and OFDM modulation have better spectrum utilization and overcoming multipath performance. In this system, because the circuit design is not complex, the OOK modulation mode is adopted. Through the '0' and '1' high and low pulse coding, we can control the brightness and darkness of LED lights and send data.

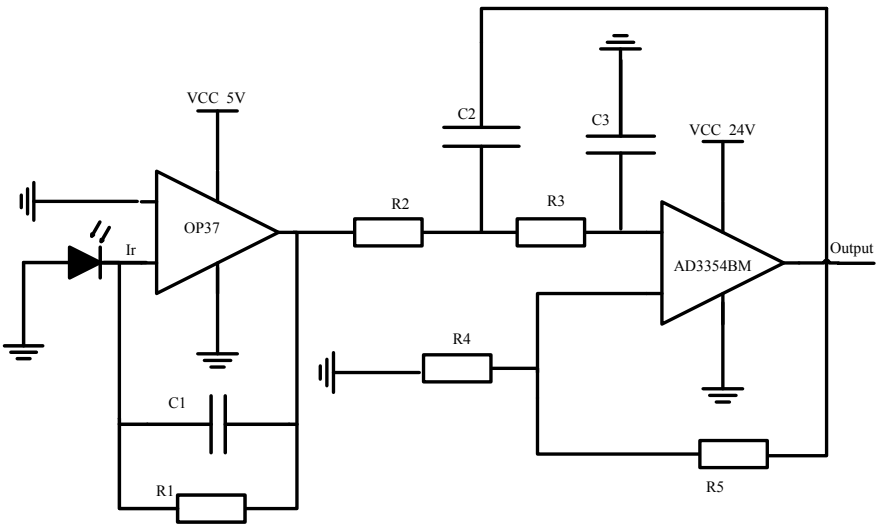



**Fig. 7.** Trans-impedance preamplifier circuit and active filter circuit

As the transmitter of the visible light communication, the LED as the transmission carrier can transmit a long series of 0 and 1 sequence signals with the brightness of LED lights which will not interfere with the radio frequency signal. Usually, when the frequency of LED in 50–60 Hz, because of the persistence of vision of the human eye, does not pay attention to the LED high-speed flash lamp; however, if the signal '0' and '1' are longer, it is possible that people will perceive the brightness change of LED. In order to meet the lighting requirements of the parking lot, the LED can still be used as a lighting appliance after transmitting the corresponding digital signal, so it is necessary to make the LED work at a high level after the signal transmission. In this paper, the non-return-to-zero (NRZ) coding of the signal can avoid the obvious brightness change of the LED to a certain extent, and it will not affect the normal work of the LED. The encoding of NRZ is shown in Fig. 8.

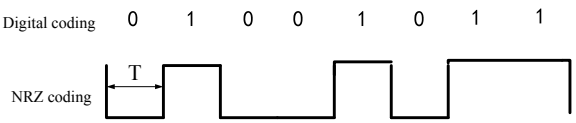



**Fig. 8.** NRZ encoding

## 4 Path Navigation

Motion sensors are used to determine the location of the moving vehicle, and the system acquirement of vehicle's path by algorithm, the LED at each node in the path indicates the driving direction of the vehicle in that position by emitting different '0'

and '1' combination signals. When the motion sensor [11] detects the movement of the vehicle, the LED on the node emits the frequency signal, and the light detector on the car receives the signal and displays the signal on the LCD, and converts the path information that the driver can identify. When the vehicle leaves the node, the point can continue executing the operation instructions of the system for subsequent vehicles; similarly, when the vehicle is away from the parking location, the sensor detects the change of the vehicle, and the system plans the outgoing route for the vehicle. The program schematic diagram for path navigation is shown in Fig. 9.

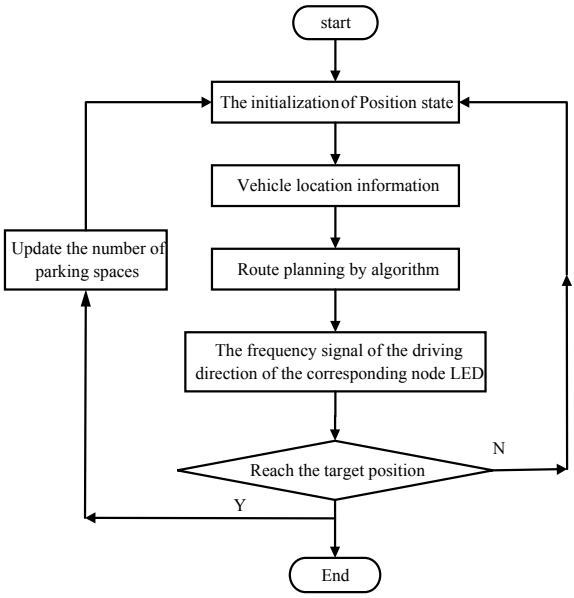



**Fig. 9.** Path navigation flowchart

## 5 Experiments and Discussions

Refer to the design standard of JGJ100-98 'specifications for building design of garage,' as shown in Table 1. In order to ensure that the system can be used for minivans, small cars, light cars, buses, and truck garages, the distance between the photodetectors of the LED light source vehicle is set to 4 m for testing.

**Table 1.** Minimum height of the car garage

| Models | Minimum height (m) |
|---|---|
| Small cars | 2.20 |
| Light vehicle | 2.80 |
| Medium, large, coach | 3.40 |
| Medium, large freight car | 4.20 |

The signal transmitting terminal of LED and the signal receiving terminal of photodetector are connected to the oscilloscope. We observe whether there are two similar waveforms on the oscilloscope. According to the design of each module circuit, the physical experiment model generated is shown in Fig. 10.

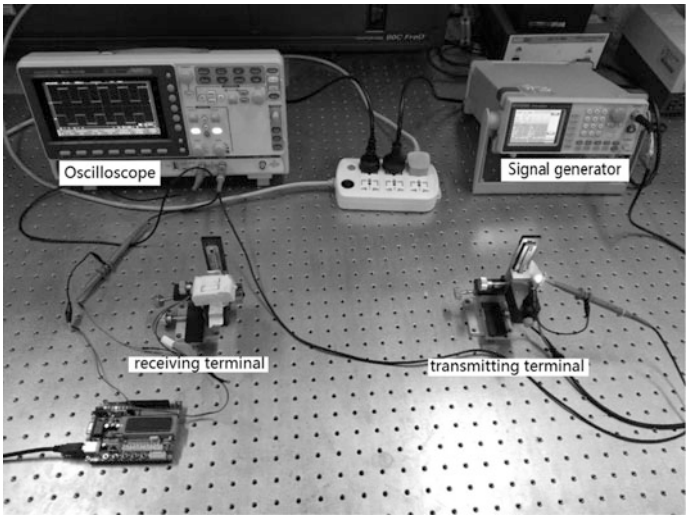



**Fig. 10.** Physical experiment model

Figure 11 shows the transmitting terminal and the receiving terminal signals detected by oscilloscope during signal transmission. The 4 m distance set in the experiment is acceptable to the signal, and the communication performance is good,

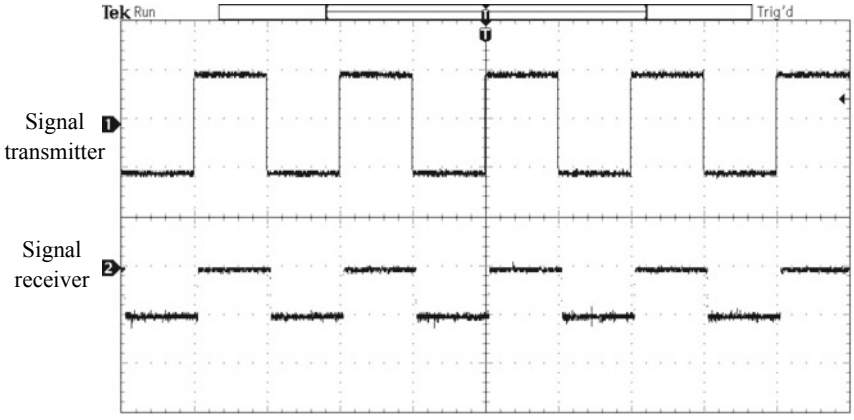



**Fig. 11.** Intercepted send and receive waveforms

and the received signal waveform is almost identical with the emitted signal waveform. Therefore, it can be seen that the system has certain practical application significance in underground garage.

## 6 Conclusion

This paper designed the navigation system of underground garage based on the principle of visible light communication in the garage with LED lighting. The system effectively reduces time to parking, avoids the blindness of people to find parking, by the LED lighting of the garage and implements the underground garage navigation based on visible light communication, is again for energy utilization. The system should be further studied and developed, which can be combined with the secondary searching system, and the influence of the moving speed of the vehicle on the reception signal should be considered in the practical application. The system has great market development value.

**Acknowledgments.** This research was financially supported by Project of the National Natural Science Foundation of China (61402069), 'the Fundamental Research Funds for the Central Universities' (3132016317), 2017 Project of the Natural Science Foundation of Liaoning province (20170540059), and General Project of Liaoning Education Department in 2016 (2016J205).

## References

1. Bao X, Dai J, Zhu X. Visible light communications heterogeneous network (VLC-HetNet): new model and protocols for mobile scenario. Wirel. Netw. 2017;23(1):299–309.
2. Beijing Institute of construction and technology, editor in chief. JGJ100–98 garage design code Design Code for Garage. Beijing: China Construction Industry Press. August 2002.
3. Wang J, Zou N, Zhang Y, Li P. Study on downlink performance of multiple access algorithm based on antenna diversity. ICICExpress Lett. 2015;9(4):1221–5.
4. Zhao J, Li Y, Zhang Y, Zou N, Wang J. A supplementary lighting system for plant growth with lighting-emitting diode based on DT TS&IC. 2016;61(7):548–51.
5. Pan G, Ye J, Ding Z. Secure hybrid VLC-RF systems with light energy harvesting. IEEE Trans. Commun. 2017;65(10):4348–59.
6. Li L. A* algorithm analysis and research. Sci. Wealth. 2016;(9):3–3. https://doi.org/10.3969/j.issn.1671-2226.2016.09.003.
7. Iskander M, Aboumoussa W, Gouvin P. Instrumentation and monitoring of a distressed multistory underground parking garage. J. Perform. Constr. Facil. 2001;15(3):115–23.
8. Onan A, Bulut H, Korukoglu S. An improved ant algorithm with LDA-based representation for text document clustering. J Inf Sci. 2017;43(2):275–292.
9. Yao Q, Mou X, Jia Y, Zhao G, Zhang H. Underground garage intelligent lighting control system. Sci. Technol. Eng. 2014;(14):239–243.

10. Rahman MS, Kim B-Y, Bang M-S, Park, Y-I, Kim, K-D. Color space mapping and medium access control techniques in visible light communication. J Inst Internet Broadcast. Commun. 2009;9(4):99–107.
11. Rabadan J, Guerra V, Rodríguez R, et al. Hybrid visible light and ultrasound-based sensor for distance estimation. Sensors (Basel). 2017;17(2):330.

# The Research of Fast Acquisition Algorithms in GNSS

Xizheng Song(✉)

Department of Electronic Engineering, Dalian Neusoft University of Information, Dalian 116023, Liaoning, China
songxizheng@neusoft.edu.cn

**Abstract.** Recently, GNSS has been applied in various domains deeply and widely. In some of applications such as carbon canyon, GNSS signals degrade severely. The conventional receivers have no ability to deal with such weak signals. The sensitivity performance has already been one of the most important features in modern receivers. Consider the advantages of easy implement and high efficiency, we choose it with coherent integration and differential coherent integration to acquire signals which of power is −145 dBm.

**Keywords:** High-Sensitivity · Weak signal · PMF

## 1 Introduction

Signal acquisition, also known as coarse synchronization, is roughly obtaining the code phase of the satellite signal and carrier frequency, which is the first step for the GPS receiver to process the baseband signal. However, the signal strength of GPS signals in an indoor environment or a sheltered environment is 15–30 dB lower than that in outdoor open areas. Therefore, how to capture such a weak signal has become the first hard problem that needs to be solved for high-sensitivity receivers.

The traditional pseudo-code acquisition is achieved through correlation and energy detection. When the correlator outputs an energy peak and that exceeds the threshold, it indicates that the input signal pseudo-code phase is consistent with the local pseudo-code phase. However, this algorithm search is very slow, which is more obvious especially during the high-speed movement of the carrier.

Furthermore, a faster speed of acquisitions is required to improve the capturing sensitivity; therefore, it has become a necessity to take the fast capturing technology.

In this paper, the structure of the matched filter +FFT is adopted, and according to the design parameters, the signal with the signal strength of −145 dBm can be captured.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_132

## 2 The Partial Matching Filtering and the FFT Fast Acquisition Algorithm

### 2.1 The Integral Forms

Integration is necessary to improve the acquisition sensitivity. When processing the communication information, there are three integral forms, namely, coherent integration, non-coherent integration and differential coherent integration. Since the integrator is a low-pass filter, the high-frequency signal components and noise in the signal are filtered out, and therefore, the output signal has a high signal-to-noise ratio. The gain obtained is called the coherent integration gain $G_{ci}$ [1]. Coherent integration is very effective for improving receiver sensitivity. However, the length of integration time is affected by various factors, which limits the further improvement of the coherent integration gain. Therefore, in order to have the ability to capture lower energy signals, the receiver generally performs N times of incoherent integration or differential coherent integration on the basis of coherent integration. Assuming that the total time for capturing points is $T_{acq}$, the coherent integration time is $T_{COH}$, then the incoherence times are:

$$N_{NC} = \frac{T_{acq}}{T_{COH}} \tag{1}$$

Since a differential coherency requires the use of the two before and after coherent integration sampling points, the number of differential coherence then becomes:

$$N_{DC} = \frac{T_{acq}}{T_{COH}} - 1 \tag{2}$$

For the same capture time $T_{acq}$, there are different combinations like $T_{COH}$ and $N_{NC}$, and $T_{COH}$ and $N_{DC}$. That is, if the coherent integration time increases, the number of incoherent coherence or differential coherence decreases; if the coherent integration time decreases, incoherent coherence or differential coherence will increase. Theoretically, the coherent integration has the best gain, while the non-coherent integration and the differential coherence integration have different degrees of loss. Therefore, the coherent integration time should be maximized to reduce the number of the non-coherent integration or the differential coherence integration. To compare the signal capturing capability, Monte Carlo experiments were performed on the acquisition of 1000 different 2 ms input signals with SNR = −20 dB, and the peak-to-noise gain was measured using the peak-to-average ratio. In the case of short bit-hopping, coherent integration is the best, and the coherence integration is better than non-coherent integration (as shown in Fig. 1). That is, in the same capture time, coherent integration time should be increased as much as possible. However, in the hardware receiver, the acquisition module is implemented by hardware, and longer coherent integration time means more sampling points and that more storage units are required.

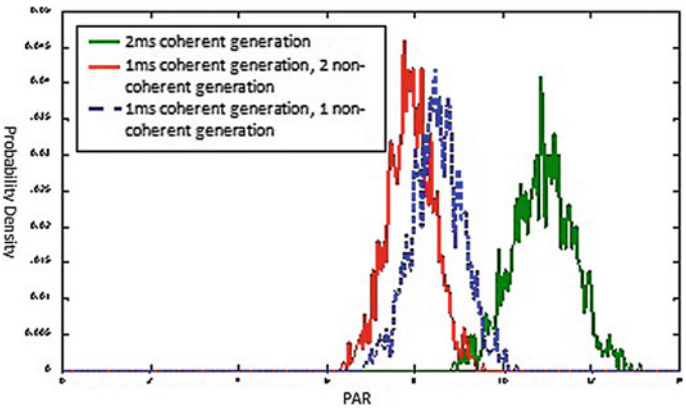



**Fig. 1.** Gain comparison of three kinds of integral form

## 2.2 The Basic Principle of Partially Matched Filter and FFT Fast Acquisition Algorithm

The fast acquisition algorithm with the combination of partially matched filter and Fourierism transform is a combination of time-frequency domain method [2], in which the matched filter to complete the serial search of code dimension, and the Fourierism transform to complete Doppler compensation to achieve the parallel Doppler dimensional search algorithm. Compared with the parallel frequency-domain-based processing technique, this algorithm is simpler, and more flexible, and it reduces the number of FFT kernels greatly, and saves resources, and it is more suitable for capturing under low signal-noise ratio and high dynamic environment. The matching filter is implemented in the form of piecewise matched filtering. The specific schematic diagram is shown in Fig. 2.

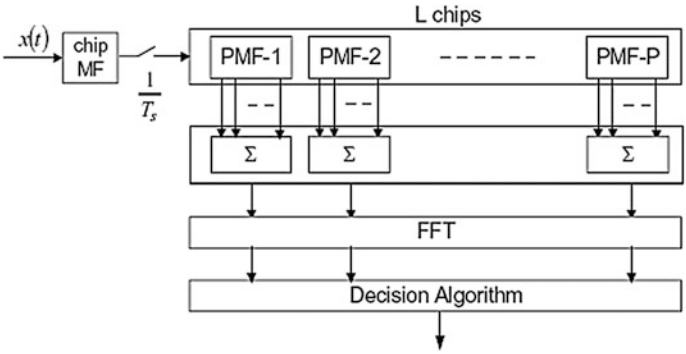


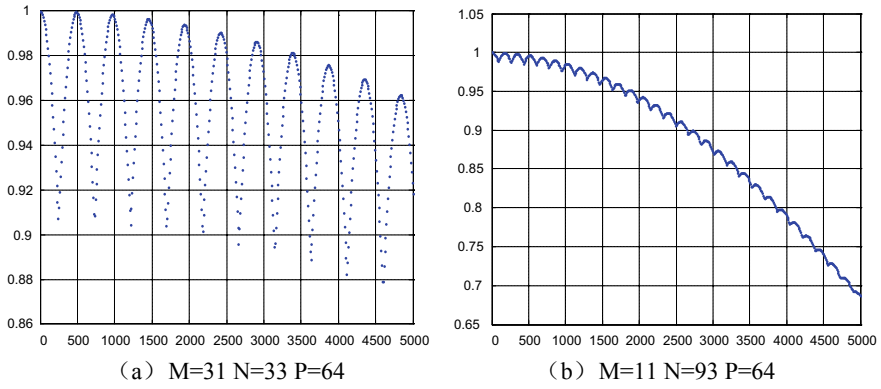

**Fig. 2.** Diagram of the piecewise matched filtering and FFT algorithm

Matching filter implementation is the core of the algorithm. For a communication system, the detection probability of a signal is directly related to the signal-to-noise ratio: The greater the signal-to-noise ratio, the greater the detection probability and the smaller the error probability; on the contrary, the smaller the signal-to-noise ratio, the smaller the probability of detection and the greater the probability of error. Therefore, in order to improve the detection probability of the communication system or to suppress the error probability, a preprocessing of the received signal should be performed so that we can get a greater output signal-to-noise ratio after the preprocessing. This principle is called the maximum signal-to-noise ratio criterion.

### 2.3 The Impact of the Partial Correlation Length and FFT Points on the Acquisition Performance

Under normal circumstances, when the spread spectrum gain is selected (i.e., the number of coherently added points is a fixed value), the number of segment points, the number of chips per segment and the number of FFT points will lead to a decrease of the accumulated peak attenuation coefficient [3]. Figure 3 shows the normalized curve of the system peak when different segments $M$ and FFT value $P$, when there is no noise involved.

As can be seen from Fig. 3, the larger the correlation length $N$, the smaller the normalized correlation peak of the FFT output. Therefore, the selection of the partial correlation length $M$ and the FFT value $P$ is determined by the bandwidth of the frequency difference, the frequency resolution and the system attenuation function, and depending on the different types of application, it can be selected flexibly.




（a） M=31 N=33 P=64　　（b） M=11 N=93 P=64

**Fig. 3.** Impact of the partial correlation length and FFT points on the acquisition performance

### 2.4 The Structure of Matching Filter

Matched filter is the core of the whole design of the fast acquisition module, which is equivalent to a FIR filter [4], shown in Fig. 4. $R$ stands for the register, and assuming

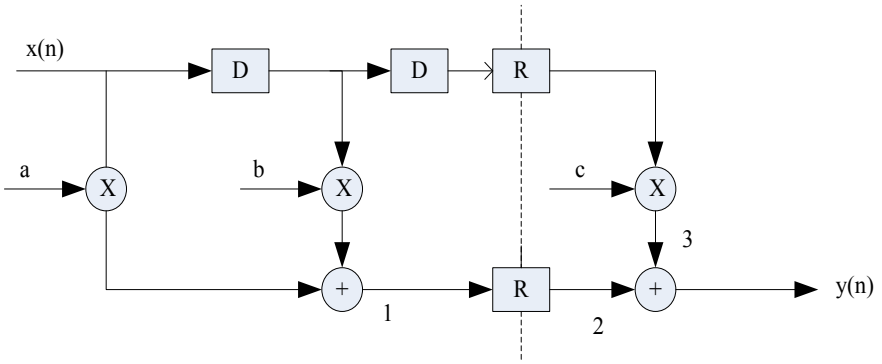



**Fig. 4.** diagram of the pipeline matching filter structure

there is no register case, $T_M$ is the multiplier delay and $T_A$ is the adder delay, then the critical path delay is $T_M + 2T_A$, so the data cycle of the filter output $T_{out}$ needs to meet

$$T_{out} = T_M + 2T_A \tag{3}$$

The pipelined approach is used in the actual design. Two registers were introduced in the figure, where x(n) is equivalent to the local C/A code sequence, and a, b, c is equivalent to the received data after AD.

By using the pipeline, the critical path delay is reduced from $T_M + 2T_A$ to $T_M + T_A$, which has effectively improved the data processing speed. Table 1 describes the processing timing of the water matched filter.

**Table 1.** Processing timing of the pipeline matching filter structure signal

| Clk | Input | Node 1 | Node 2 | Node 3 | Output |
|---|---|---|---|---|---|
| 0 | x(0) | ax(0) + bx(−1) | – | – | – |
| 1 | x(1) | ax(1) + bx(0) | ax(0) + bx(–1) | cx(–2) | y(0) = ax(0) + bx(–1) + cx(–2) |
| 2 | x(2) | ax(2) + bx(1) | ax(1) + bx(0) | cx(–1) | y(1) = ax(1) + bx(0) + cx(–1) |
| 3 | x(3) | ax(3) + bx(2) | ax(2) + bx(1) | cx(0) | y(2) = ax(2) + bx(1) + cx(0) |

As can be seen from the table, in the case when there is a large amount of data, theoretically, each clock will produce an output, and because of the match filtering hierarchy, the logic delay of each stage decreases, and the clock speed increases accordingly.

## 3 Partially Matched Filter and FFT Fast Acquisition Scheme

According to the actual requirements, a compromise was made between the acquisition time and the sensitivity. Therefore, in this scheme, the coherent integration length of the matched filter is 10 ms, which is divided into 128 segments with integral time of (10/128) ms, and then 128 points of FFT were used to estimate Doppler. Through the simulation analysis of MATLAB, the acquisition sensitivity gets to 141 dBm after using 10 ms coherent integration. Considering the quantification loss, noise correlation and various control loss (2 dB) of the limited data bits, the acquisition sensitivity can reach −139 dBm. For weak signals, non-coherent integration capture is started by software setting. Non-coherent acquisition is performed seven times, and in order to solve the message bit reversal, two consecutive 10 ms data segments are captured, of which a bit must not be bit-flipped. The specific capture strategy is to pre-store 140 ms of data, and then for every 10 ms process a fast acquisition, and then do the accumulation of the odd times of 10 ms and even times of 10 ms processed FFT matrix, with a total of seven cumulative to complete the non-coherent processing. The capture results are shown in Fig. 5. As can be seen from the figure, the capture peak is obvious compared with other clutter peaks, with a capture sensitivity of −145 dBm (remove the loss of 2 dB).

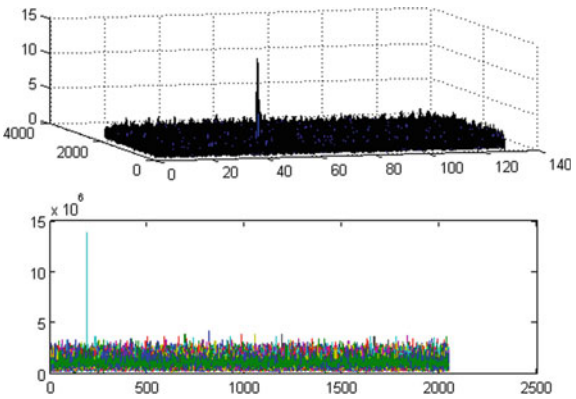
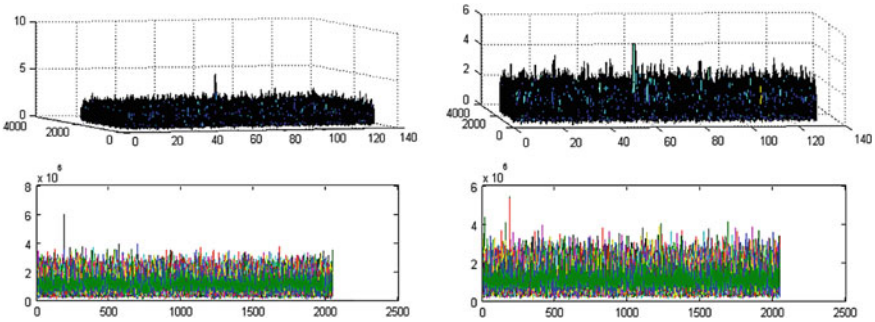



**Fig. 5.** Simulation results when SNR = −36 dB (−147 dBm), fd = 3800 Hz

According to the theoretical analysis, when using the above parameters, the theoretical Doppler search range reaches −12.8 to 12.8 kHz, but when the Gaussian white noise is added in the simulation, the Doppler search range is significantly reduced. In addition, since at this time the Doppler accuracy is 100 Hz, when the Doppler frequency is exactly the valley of the scallop loss, the acquisition peak is no longer noticeable. Figure 6 illustrates the impact that both cases have on the acquisition results.

As can be seen from Fig. 6, when Doppler reaches 4.8 kHz, the acquisition peak is nearly submerged in the noise, while for a typical receiver, −4.8 to 4.8 kHz basically can cover the normal Doppler range, but for the receivers that have higher dynamically demands, the search range cannot cover all Doppler frequencies. Therefore, the local carrier frequency must be changed to search for other Doppler frequency ranges.



(a) fd=4850Hz (b) fd=5050Hz

**Fig. 6.** Impact of different doppler on the acquisition results when SNR = −36 dB

## 4 Conclusion

The first step of processing baseband signal in the acquisition of satellite navigation signals, the improvement of the acquisition sensitivity and the improvement of the acquisition speed has always been the primary goals of satellite navigation signal acquisition. In this paper, by analyzing the matched filter and several integral forms, by using the method of the combination of the partial matched filter and FFT fast acquisition, the acquisition speed is increased, and at the same time the acquisition sensitivity is improved. Moreover, under the general dynamic conditions, this method generally covers the general Doppler range, and the experimental simulation and actual hardware implementation tests both show that the method has a great effect of the signal acquisition.

## References

1. yi J, Shufang Z, Qing H, Xiaowen S. A new FFT-based acquisition algorithm for GPS signals. In: 2008 education technology training and 2008 international workshop geoscience remote sensing. ETT and GRS 2008 international workshop on, vol. 2, 2008.
2. Tian S, Pi Y. Research of weak GPS signal acquisition algorithm. In: International conference on communications circuits and systems 2008 (ICCCAS 2008). 2008; p. 793–6.
3. Akopian D, Fast FFT based GPS satellite acquisition methods. IEE Proceeding-Radar Sonar Navig. 2005;152(4).
4. Spillard CL, Spangenberg SM, Povey GJR. A serial-parallel FFT correlator for PN code acquisition from LEO Satellites. In: Proceedings of ISSSTA'98, 1998.

# Research on BDS/GPS Combined Positioning Algorithm

Hong-Fang He(✉), Xin-Yue Fan, and Kang-Ning An

Key Laboratory of Optical Communication and Networks, School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
hehfang@hotmail.com, fanxy@cqupt.edu.cn, ankangning@163.com

**Abstract.** When single satellite navigation system is used for positioning, there exist the following problems: accuracy of positioning is low and reliability of positioning is also low. This paper investigates the combined positioning algorithm of BDS and GPS for static and dynamic observation point. For static observation point positioning, the weight coefficient integrated positioning algorithm of BDS and GPS is proposed. For dynamic observation point positioning, the Kalman filter can achieve smoothing of the movement trajectory which is the once combined observation point. The experimental results show that the combination of BDS and GPS is more accurate and reliable than any single system for static observation point positioning. At the same time, for the consideration of the weight coefficient, the system has good adjustability and practicality, and the Kalman filter can better modify the dynamic combined observation point.

**Keywords:** Combined positioning · Static point positioning · Dynamic point positioning · Weight coefficient

## 1 Introduction

Beidou Navigation Satellite System (BDS) is one of the four satellite navigation systems independently developed by China. Since the laboratory system was established in 2000, 15 orbital satellites have been developed to cover the entire Asia-Pacific region by 2012, and it will provide services around the world by 2020. It includes geostationary (GEO) satellites, inclined geosynchronous satellite orbit (IGSO) satellites, and medium earth orbit (MEO) satellites distributed over 21,000 km and over 35,000 km [1]. Global Positioning System (GPS) is developed by the US Land and Air Force. It was researched quite early and had achieved a global coverage of 98% in 1994. There are MEO satellites distributed over 20,200 km [2]. Under the research of all parties, the accuracy of single point positioning of BDS and GPS can reach a level of 10 m and an elevation of 10 m [3, 4].


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_133

With the gradual improvement of the four major satellite navigation systems, the future air constellation satellites will have more than 120 satellites, providing more stable, reliable, and accurate services of navigation and positioning for global users [5]. However, facing many satellite navigation information resources, its body of the application has gradually evolved from single navigation positioning to a new era of multi-information integration, such as collective positioning and the Internet. Global navigation satellite system (GNSS) has also entered the era of multi-system compatible and cooperative development [6]. The development time of BDS is short, and the technology is not yet perfect in this filed. Therefore, the combination of mature GPS and BDS can further improve accuracy and reliability of navigation positioning [7,8].

The satellites of BDS and GPS have different orbital altitudes, which has created some obstacles for the study of tight combinations [9]. Currently, BDS and GPS are capable of doing independently navigation positioning, the accuracy of single point positioning is not accurate enough, and researching the combined positioning of observation points under single point positioning is rare [10]. This paper designs the combined observation point positioning algorithm of BDS and GPS. For once positioning and multiple positioning of static observation points, a corresponding weight coefficient integrated positioning algorithm is put forward. At the same time, based on once positioning of static observation points, the Kalman filter can smooth the movement trajectory for the dynamic observation point. This design not only has adjustability, but also has strong practicality and confidentiality.

## 2 Static Point Positioning

When the static observation point is located, once and multiple positioning can be performed on the same observation point by using BDS and GPS simultaneously. Therefore, we can get different positioning coordinates. A more accurate coordinate can be obtained by combined method ultimately.

### 2.1 Once Positioning

At a fixed point, once positioning is used to do the same static observation point by using BDS and GPS simultaneously. Thus, two positioning coordinates are obtained to perform the combined positioning. In the longitude and latitude direction of the observation point, assume that $(x_B, y_B)$ and $(x_G, y_G)$, respectively, are predictive coordinates of BDS and GPS, and $(x, y)$ is the combined coordinate of BDS and GPS. The positioning accuracy of specific receiver determines observed gain errors of BDS and GPS. $\varepsilon_B = 5\,\mathrm{m}$ and $\varepsilon_G = 2\,\mathrm{m}$, respectively, are the gain error counted from the receiver in the paper. When the positioning coordinates of two-systems are the same, $(x, y) = (x_B, y_B) = (x_G, y_G)$ is the combined coordinate of BDS and GPS. However, the positioning accuracy of two systems is different in practical operation. Assume that the combined

coordinate and range, respectively, are calculated as follows:

$$(x, y) = \frac{1}{2}(x_B, y_B) + \frac{1}{2}(x_G, y_G)$$
$$r_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} \ (i = B, G) \tag{1}$$

When $r_i \leq 2\varepsilon_i$ is established, the weight coefficient as follows:

$$\delta_i = 1 - \frac{r_i}{r_B + r_G} \ (i = B, G) \tag{2}$$

When $r_i > 2\varepsilon_i$ is established, if the range is large, the weight coefficient will be 0, and if the range is small, the weight coefficient will be 1, which means that if one system suffers from strong interference, it will not participate in positioning. Therefore, the combined coordinate of BDS and GPS is calculated as follows:

$$(x, y) = \delta_B (x_B, y_B) + \delta_G (x_G, y_G) \tag{3}$$

### 2.2 Multiple Positioning

At a fixed point, multiple positioning is used to do the same static observation point by using BDS and GPS simultaneously. Thus, multiple discrete coordinates are obtained to perform combined positioning. The idea of combined positioning is to leave dense coordinates and remove discrete coordinates. When there are enough sample points, it is evenly distributed among a certain neighborhood of the precise coordinates, and offset errors can cancel each other out.

In the longitude and latitude direction of the observation point, assume that $\left(x_i^B, y_i^B\right)$ and $\left(x_i^G, y_i^G\right)$, respectively, are predictive multiple discrete coordinates of BDS and GPS, and $(x, y)$ is accurate combined coordinate of BDS and GPS. Let us take BDS as an example, the specific method is as follows:

Step 1: Calculating the desired point coordinate in all discrete coordinates is as follows:

$$\left(x_0^B, y_0^B\right) = E\left\{\left(x_1^B, y_1^B\right), \left(x_2^B, y_2^B\right), \ldots, \left(x_n^B, y_n^B\right)\right\} \tag{4}$$

Step 2: Respectively, taking the discrete point $\left(x_i^B, y_i^B\right) (i = 0, 1, \ldots, n)$ as the center of circle $O_{Bi}$ and the error gain $\varepsilon_B$ as the radius. The number of all coordinates contained in the circle $O_{Bi}$ is $n_{Bi}$. Therefore, their probability is $p_i = \frac{n_{Bi}}{n}$.

Step 3: When $p_i \geq p_0$ is established, corresponding positioning coordinates are remarked as $\left(x_j^B, y_j^B\right) (j = 0, 1, 2, \ldots, m_B)$. If the number of observed effective positioning coordinates is $m_B \geq 1$, the collection of available positioning positions is $\theta_B = \bigcap_{j=1}^{m_B} O_{Bj}$, and if the number of observed effective coordinates is $m_B = 0$, the collection of available positioning positions is $\theta_B = O_{B0}$.

As above, we can derive the number of effective coordinates $m_G$ and the collection of available positioning positions $\theta_G$ observed by GPS. Then, $(x_B, y_B)$

and $(x_G, y_G)$, respectively, are barycentric coordinates of the collection of available positioning positions. Finally, the precise combined coordinate of BDS and GPS is shown in formula (3), where the weight coefficient as follows:

$$\delta_i = \frac{m_i}{m_B + m_G} (i = B, G) \tag{5}$$

## 3 Dynamic Point Positioning

When the dynamic observation point is located, data is gathered using BDS and GPS simultaneously in a very short period of time, in which combined positioning of observation points is carried out. The specific method is the same as the static observation point. In fact, the movement trajectory of the BDS/GPS combined observation point is determined, and two ways can be used generally. The first is to collect many discrete coordinates as combined positioning coordinates in very short time intervals, which fit it directly to the movement trajectory. The second is to collect once coordinates as combined positioning coordinates in a very short time interval, which smooth the movement trajectory using Kalman filter. The first operation is less difficult, and the second effect is good.

This paper determines the movement trajectory of BDS/GPS combined observation point by the second method. The Kalman filter has a correlation with each epoch. The observations of the previous moment have an effect on the latter moment, and it has good smoothness. The study, which is about the movement trajectory of combined observation points of BDS and GPS based on Kalman filter, is assumed to be performed under linear conditions. Now only considering the uniform tracking position of the observation point in the two-dimensional plane, the selected state vector of the observation point is as follows:

$$x_k = \left[ x_t(k) \; \dot{x}_t(k) \; y_t(k) \; \dot{y}_t(k) \right]^T \tag{6}$$

where the coordinate of observation point under each epoch is $(x_t(k), y_t(k))$, and the speed of observation point under each epoch is $(\dot{x}_t(k), \dot{y}_t(k))$.

Owing to lack of control variable, the available system model and measurement model of the observation point are as follows:

$$\begin{aligned} x_{k+1} &= Fx_k + B_w w_k \\ z_{k+1} &= Hx_k + v_k \end{aligned} \tag{7}$$

where the state transition matrix $F, B_w$ and the measurement matrix $H$ are as follows:

$$F = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \; B_w = \begin{bmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{bmatrix} \; H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{8}$$

The sampling interval is set to $T$. $w_k$ and $v_k$, respectively, are system state noise and measurement noise of Gaussian white noise sequence. $D_{w_k}$ and $D_{v_k}$,

respectively, are its covariance matrix. The equations are as follows:

$$\begin{cases} E(w_k) = 0 \\ E(v_k) = 0 \\ \mathrm{cov}(w_k, w_j) = D_w(k)\delta_{kj} \\ \mathrm{cov}(v_k, v_j) = D_v(k)\delta_{kj} \\ \mathrm{cov}(w_k, v_j) = 0 \end{cases} \tag{9}$$

$\delta_{kj}$ is the Kronecker function, and $E(x_0) = u_x(0)$ $\mathrm{var}(x_0) = D_x(0)$ $\mathrm{cov}(x_k, v_k) = 0$, $\mathrm{cov}(x_k, w_k) = 0$ denotes the initial state of the system respectively.

## 4 Experiment Analysis

We choose the outdoor playground in Chongqing University of Posts and Telecommunications as our experiment area. Experimental data were collected by GPS/BDS two-system compatible receiver, which respectively collected 600 static data and 100 dynamic data by GPS and BDS. The measurement interval was set to 1s. The following is a test of static positioning error using multiple positioning, and the dynamic positioning error using the Kalman filter. Respectively, we analysis the positioning error in the longitude direction of the x-axis and latitude direction of the y-axis.

### 4.1 Static Point Positioning Analysis

To evaluate the combined positioning performance of static observation point, we compare the positioning error of single-system and two-system combined observation point after multiple positioning. In the two-dimensional coordinate system of longitude and latitude, standard deviation and root mean square error are defined as follows:

$$\sigma_{2D} = \sqrt{\frac{\sum_{i=1}^{n} \left((x_i - u_x)^2 + (y_i - u_y)^2\right)}{n}} \tag{10}$$

$$\delta_{2D} = \sqrt{\frac{\sum_{i=1}^{n} \left((x_i - x_0)^2 + (y_i - y_0)^2\right)}{n}}$$

where $n$ is the number of epochs, and $(x_i, y_i)$ is the positioning coordinates of the epoch data. $u_x$ and $u_y$, respectively, are the expected values of their corresponding positioning point. The true coordinate of the test location is $(x_0, y_0)$.

Table 1 shows the STD and RMS values obtained from the statistics of static observation points after multiple positioning.

As can be seen from Table 1, in the longitude and latitude direction of the observation point, the positioning effect of static observation point by GPS is better in the single system, and the positioning effect of BDS is relatively poor. At the same time, compared with the positioning of single system, the positioning

error of the combined observation point of BDS and GPS is the smallest, and the positioning error of the combined observation point is relatively minimal in the two-dimensional coordinate system of longitude and latitude. Therefore, the combined positioning of static observation point of BDS and GPS is more accurate after multiple positioning.

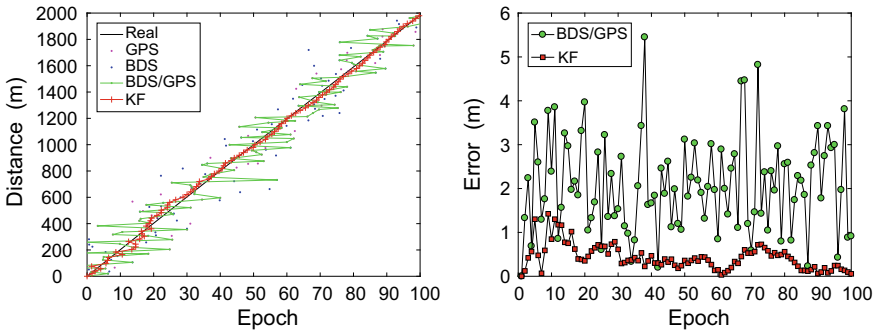
**Table 1.** STDs and RMSs of single-system and two-system combined observation point

| Scheme | STD | | | RMS | | |
|---|---|---|---|---|---|---|
| | $\sigma_x$ | $\sigma_y$ | $\sigma_{2D}$ | $\delta_x$ | $\delta_y$ | $\delta_{2D}$ |
| BDS | 1.4441 | 2.0834 | 2.9590 | 1.5391 | 2.3390 | 3.0588 |
| GPS | 0.8524 | 1.4384 | 2.0506 | 1.3493 | 1.8708 | 2.5992 |
| BDS/GPS | 0.7614 | 0.9062 | 1.7318 | 0.8063 | 0.9329 | 1.8331 |

### 4.2 Dynamic Movement Trajectory Analysis

To evaluate the combined positioning performance of dynamic observation point, we compare the movement trajectory of single-system and two-system combined observation point after once positioning. At the same time, we compare the positioning error of directly fitting combined observation points and the Kalman filter smoothing observation points. Figure 1a shows the movement trajectory of single-system and two-system combined observation point and the Kalman filter smoothing observation point. Figure 1b shows the positioning errors of BDS/GPS combined and the Kalman filter smoothing observation points.

It can be seen from Fig. 1a that the corrected movement trajectory using the Kalman filter is closer to the true movement trajectory. Tracking points of




(a) Movement trajectory (b) Positioning error

**Fig. 1.** Movement trajectory and positioning error

BDS and GPS are distributed on both sides for real movement trajectory. The fitted movement trajectory did not achieve very good effect after the combined observation points of BDS and GPS, but the movement trajectory is smoother using the Kalman filter.

It can be seen from Fig. 1b that the positioning error using the Kalman filter is less than the combination of BDS and GPS, and the positioning error is stable within 1.5 m, which means that the Kalman filter has a good corrected effect on the processing of dynamic observation points.

## 5 Conclusion

In this paper, we investigate the combined positioning of static and dynamic observation points. Once positioning and multiple positioning of static observation point are separately analyzed, two situations that can be used for the movement trajectory of the dynamic combined observation points have been introduced. Finally, for static combined observation point positioning, multiple positioning is chosen to determine a more accurate combined positioning coordinate. For the movement trajectory of the dynamic combined observation points, the Kalman filter has a good corrected effect after once combined positioning.

At the same time, the proposed weight coefficient integrated algorithm is adjustable, which means that not only an accuracy positioning can get by BDS and GPS, but also the positioning does not depend on BDS or GPS. The algorithm can be used in the military field, because they are practical and confidential.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61471077).

## References

1. Wang B, Lou Y, Liu J, Zhao Q, Su X. Analysis of BDS satellite clocks in orbit. GPS Solut. 2015;20(4):1–12.
2. Yang D, Yang J, Li G, Zhou Y, Tang CP. Globalization highlight: orbit determination using BeiDou inter-satellite ranging measure-ments. GPS Solut. 2017;21(3):1395–404.
3. Chen H, Xu C, Gao J, Song X, Yuan L. Precision analysis of pseudorange single point positioning by BDS. GPS and combined BDS/GPS: J Shandong Univ Sci Technol; 2015.
4. Man X, Sun F, Liu S, Li H, Ding H. Analysis of positioning performance on combined BDS/GPS/GLONASS. Berlin, Heidelberg: Springer; 2015.
5. Wei E, Liu X, Liu J. Accuracy evaluation and analysis of single point positioning with BeiDou and GPS. Bulletion Surv Mapp. 2017.
6. Yuanxi Y. Progress, contribution and challenges of Compass/Beidou satellite navigation system. Acta Geod Cartogr Sin. 2010;39(1):1–6.
7. Dai F, Mao X. BDS/GPS dual systems positioning based on kalman filter in urban canyon environments, In: IEEE International conference on intelligent transportation systems; 2014. p. 1882–3

8. Zhang K, Hao J. Research on BDS/GPS combined single-epoch attitude determination performance (2017).
9. Zeng A, Yang Y, Ming F, Jing Y. BDS-GPS inter-system bias of code observation and its preliminary analysis. GPS Solut. 2017;21(2):1–9.
10. Liu R, Gao X. The integrated positioning algorithm of "Beidou". GPS and GLONASS Satellite navigation system: Basic Sci J Text Univ; 2017.

# Indoor Positioning with Sensors in a Smartphone and a Fabricated High-Precision Gyroscope

Dianzhong Chen(✉), Wenbin Zhang, and Zhongzhao Zhang

Communication Research Center, Harbin Institute of Technology, Harbin 150001, China
dc2el2@163.com, zwbgxyl973@hit.edu.cn, zzzhang@hope.hit.edu.cn

**Abstract.** In the paper, an indoor positioning scheme combining pedestrian dead reckoning (PDR) and magnetic strength matching (MSM) is proposed. PDR is conducted by sensing acceleration and angular speed through the 3-axis accelerometer in iphone7 and a fabricated high-precision rotational gyroscope. Low bias stability (0.5°/h) of the gyroscope contributes to a small accumulative error in heading angle estimation. Through data analysis to outputs of the accelerometer and the gyroscope, human motion, such as walking a step, walking upstairs or downstairs, turning left or right, is recognized and walking path is reckoned with motion information. Magnetic strength is measured by the magnetometer in iphone7 and MSM positioning result is used to reduce error of reckoned heading angle. The error rate of downstairs/upstairs step count is low and after heading angle correction by MSM, a satisfactory indoor positioning result is obtained.

**Keywords:** Indoor positioning · Pedestrian dead reckoning (PDR) · Magnetic strength matching (MSM) · Modified dynamic time warping (DTW) algorithm

## 1 Introduction

Service requirements of indoor positioning have been growing rapidly over time. Though global navigation satellite system (GNSS) for outdoor navigation has been mature and widely utilized, indoor positioning is still a task for the facts that: (1) GNSS signal in indoor environment is unavailable or seriously attenuated [1]; (2) indoor environment is complex with the phenomenon of not line of sight (NLoS) [2] and interference from humans [3]; (3) human path is not restricted to specific route as vehicles [4], however, with unpredictable turns [5]. Existing indoor positioning technology includes Wi-Fi positioning, Bluetooth Low Energy (BLE) positioning, radio frequency identification (RFID) positioning, dead reckoning (DR), MSM, and so on. Wi-Fi positioning, with accuracy of 10–20 m, is widely used for mature IEEE 802.11 standard and universality of Wi-Fi signal receiver, such as smartphones [6]. However, this technology has problems of high reliability on availability and distribution of Wi-Fi signals [7], attenuation [8], and multipath effect [9]. BLE positioning depends on dense arrangement of signal sources and RFID positioning requires special signal


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_134

receivers, which restrict their widespread application. DR, based on measured linear acceleration signal and angular speed signal from inertial sensors (accelerometer, gyroscope) to calculate motion trail without receiving any external measured signals, is a positioning method different from others. Thus, DR is immune to outside interference or change of environment. However, for frequently used inertial MEMS sensors in watches or smartphones, inherent problems such as temperature sensitivity [10, 11], non-ideal repeatability of input–output characteristic for different measurements, low bias stability, cause measurement errors. Moreover, heading angles and linear speeds are the accumulation of measured angular speeds and linear accelerations, which will accumulate measurement errors [12]. DR with MEMS sensors is often used in short-distance path reckoning, and the other method is needed to modify the path. MSM can be a supplementary method to modify positioning error by PDR, as introduced in [13, 14]. Positioning performance of MSM relies highly on distribution condition of magnetic strength [15, 16]. Higher stability in time and larger variations in space of magnetic strength distribution contribute to better positioning precision by MSM.

In this paper, an indoor positioning scheme combining PDR and MSM is proposed. The three-axis accelerometer in iphone7 and a fabricated high-precision gyroscope (with the bias stability of 0.5°/h) with a ball-disk shaped rotor [17, 18] compose the inertial measurement unit (IMU). Construction of IMU is introduced in Sect. 2 with the structure of the fabricated gyroscope illustrated. In Sect. 3, PDR based on human motions through analysis to IMU signals is described in detail. In Sect. 4, the principle of MSM positioning and procedure of modified DTW algorithm are introduced and an example of MSM positioning with the modified DTW algorithm is given. An experiment of indoor positioning is conducted and the precision of the proposed scheme is analyzed in Sect. 5. Section 6 concludes the significance of the proposed indoor positioning scheme.

## 2 IMU Construction

The bias stability of the MEMS rate-grade gyroscope in smartphones is in the range of 10–1000°/h, which will lead to a large accumulative error in indoor positioning. A tactical-grade gyroscope with the bias stability of 0.5°/h is fabricated, constructing the IMU with the 3-axis accelerometer in iphone7. The fabricated gyroscope is a rotational gyroscope with a ball-disk shaped rotor supported by a water-film bearing, based on magnetic self-restoring effect to balance Coriolis torque, as shown in Fig. 1. Measurement range of the fabricated gyroscope is −30°/s to 30°/s, which is restricted by the linearity of differential capacitance detection [17, 18]. Nevertheless, for positioning application with memory module and processing module, it is feasible to construct a lookup table with corresponding differential capacitances and input angular speeds. To further increase the measurement range, the distance between the detection electrode and the rotor disk is increased from 100 μm to 500 μm. And the rotor deflection angle range increases from 1° (restricted by linearity) to 6°. And correspondingly, maximum measurable input angular speed increases from 30°/s to 180°/s (3.14 rad/s), enough for walking action recognition. But as a trade-off, sensitivity decreases, however, enough for PDR application. Low bias stability value of the

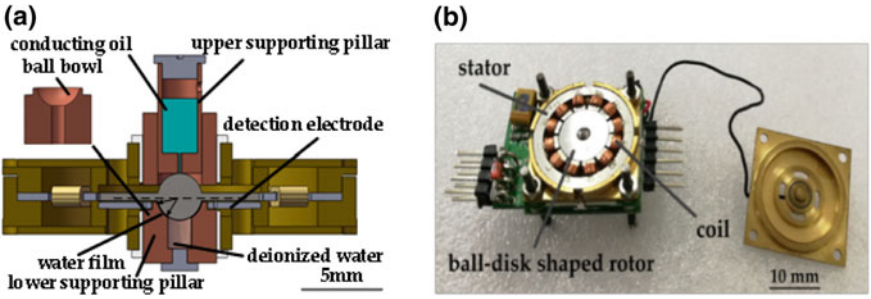



**Fig. 1.** **a** Engineering diagram of the gyroscope. **b** Photograph of the gyroscope

fabricated gyroscope (0.5°/h) contributes to a small accumulative error of heading angle estimation. The gyroscope is fixed on its side upon the iphone7 screen and put in a box, named as positioning box. There is a belt on the box to be fastened on the waist of the human for motion tests.

## 3 PDR Based on Human Motion

PDR is conducted by detecting typical human motions of walking a step on the flat floor, walking upstairs/downstairs, turning right/left through inertial sensors. I wear the positioning box on the waist, collect data from sensors during these actions, analyze their features, select those can be utilized for motion detection. In order to express it more clearly, the forward direction (*Y*-axis direction of the accelerometer) and the vertical direction (*Z*-axis direction of the accelerometer) during walking are defined as shown in Fig. 2. Motion of walking on the flat floor is divided into two stages. In stage one, from the time when heel of the left (right) foot touches the floor to the right time before tiptoes of the right (left) foot leave the floor, acceleration in forward direction is increased, as curve A–B in Fig. 3a. In stage two, from the time when tiptoes of the right (left) foot leave the floor to the time when heel of the right (left) foot touches the floor, acceleration in forward direction is decreased, as curve B–C in Fig. 3a. Thus, motion of walking with two stages can be detected by change of acceleration in forward direction.

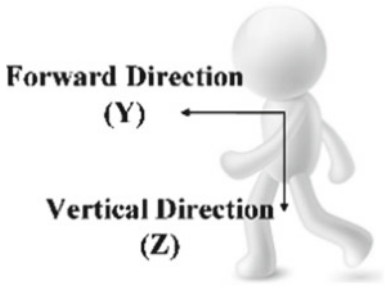



**Fig. 2.** Schematic diagram of directions during walking motion

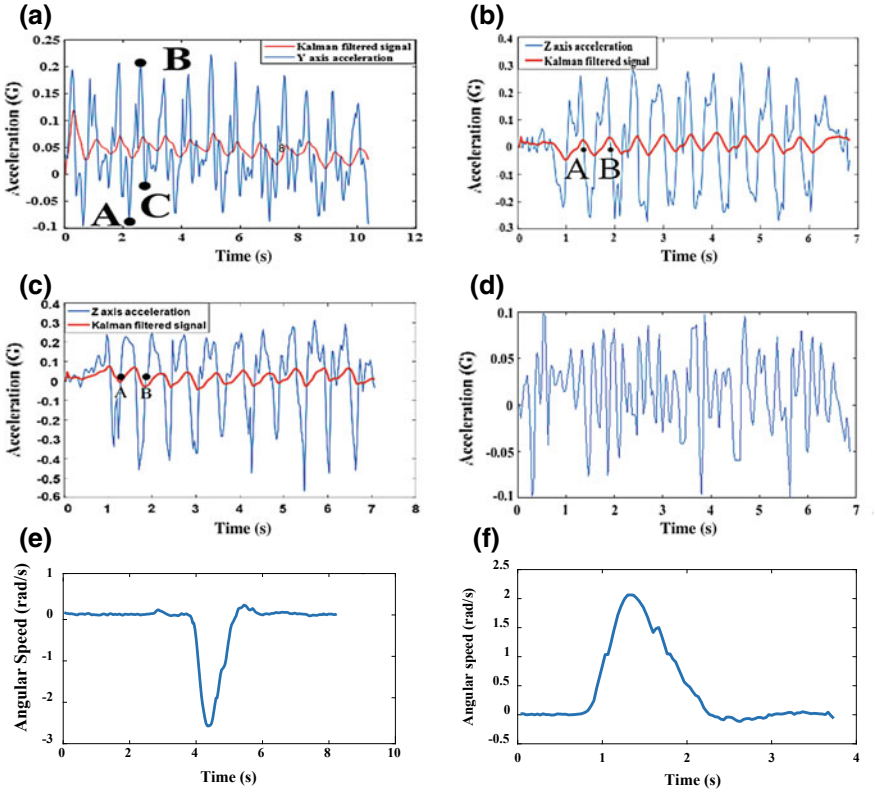



**Fig. 3.** **a** *Y*-axis acceleration while walking on the flat floor. **b** *Z*-axis acceleration of walking upstairs of 10 stairs. **c** *Z*-axis acceleration of walking downstairs of 10 stairs. **d** *Z*-axis acceleration of walking on the flat floor. **e** Angular speed of turn right. **f** Angular speed of turn left

Kalman filtering an algorithm jointing newly measurement and prediction based on last output to decrease measurement error is adopted to filter *Y*-axis acceleration. Filtered signal is noted as the red line as shown in Fig. 3a and peak number of the filtered signal is equal to sum of positive and negative peaks of the blue curve. Thus, through Kalman filtering, walking steps can be obtained by the function findpeaks in Matlab. Motion of walking upstairs and downstairs of 10 steps is detected by *Z*-axis acceleration (Fig. 3b, c), and the curve for motion of one step walk is marked between points A, B. As with walking on the flat floor, curves are Kalman filtered and the number of steps is equal to positive (negative) peaks of the filtered red curve. Figure 3d represents the curve of *Z*-axis acceleration in motion of walking on the flat floor. Compare Fig. 3b–d, curve difference of *Z*-axis accelerations is obvious, which can be a distinguish standard to differentiate three motions before step counting. Heading angle change is sensed by the fabricated gyroscope in IMU. Gyroscope can sense the angular speed and the turning angle is the integral of angular speed to time. Gyroscope outputs for experiments of turning right and turning left are as shown in Fig. 3e, f. Turning direction can be easily

distinguished by angular speed direction. Turning angle is calculated by $\theta = \sum_i \frac{1}{f}\omega_i$, where $f$ is the sampling frequency (30 Hz) and $\omega_i$ is the angular speed at the sampling point. Calculated turning angle of turning right and turning left is −1.62 and 1.54 rad, respectively, absolute value of which are approaching to $\pi/2$ rad.

## 4 Magnetic Strength Matching (MSM) Indoor Positioning

Magnetic strength matching works under the precondition of stable magnetic strength distribution in time and existence of magnetic strength variation in space. Indoor positioning by MSM is conducted through two stages: database construction by mobile matching and positioning. This scheme increases the dimension of information from 3 to the length of magnetic strength sequence. Database stores magnetic strength sequences of continuous RPs and positions of the first and the last RP as the form of:

$R_k = \{m_{k1}, m_{k2}, m_{k3}, \ldots m_{kn}, \text{pos}_{k1}, \text{pos}_{k2}\}$, where $m_{k1}, m_{k2}, m_{k3}, \ldots m_{kn}$ represent magnetic field strength of n RPs in route $R_k$. When the measured continuous magnetic strength data sequence matches $R_k$ in shape, the walking path is positioned to be between $\text{pos}_{k1}$ and $\text{pos}_{k2}$.

For a different height and walking speed of humans, curves of magnetic strength sequences through the same path will have similar overall component shapes but with different average amplitudes and sequence lengths as graph of Fig. 4a [19]. To identify similarity of signal 1 and signal 2, a time domain signal processing method of modified DTW is adopted. DTW algorithm stretches or compresses signal curves for dissimilarity analysis as shown in Fig. 4b. After signal 2 stretches to the same length as signal 1, signal similarity is evaluated by sum of squared euclidean metric of corresponding points. Modification to DTW algorithm is to eliminate influence of difference in average amplitude of magnetic strength. Procedure of modified DTW is as below: let data sequences $A(a_1, a_2,\ldots a_n)$, $B(b_1,b_2,\ldots b_m)$ be the reference data sequence and measured data sequence, respectively. For magnetic strength change trend matching, measured data sequence $B$ is modified ($B'$ $(b_1 + d, b_2 + d,\ldots, b_m + d)$) to be with the same average as the reference data sequence A, initially. Then, a $n \times m$ matrix grid is constructed with one element at each grid $(i, j)$ representing squared euclidean distance between $a_i$ and $b_j$. A route from (0, 0) to $(m, n)$ is searched for with the smallest sum of elements in grids. A chosen grid $(g, h)$ means $a_g$, $b_h$ are corresponding data in

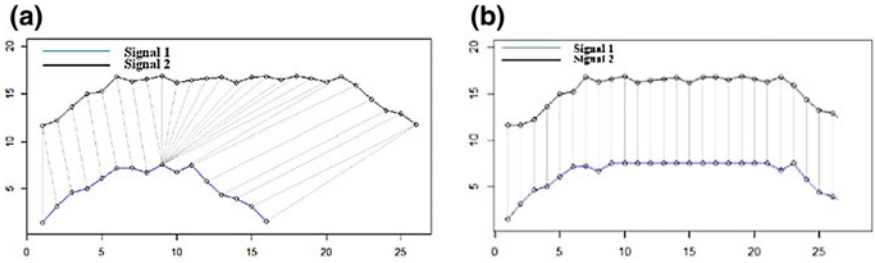



**Fig. 4.** Example of signal alignment by DTW

sequences A, B and the smallest sum is the similarity expressed in value between A and B, noted as parameter SUM. The smaller the SUM is, the better sequence A matches sequence B. The algorithm is implemented in Matlab.

To evaluate the functionality of MSM with modified DTW algorithm, a path is walked through twice with measured data sequence 1 (D1), 2 (D2) and a path nearby is walked through with measured data sequence 3 (D3). D2 is set as the reference data sequence and data pair of (D1, D2), (D3, D2) are input to the Matlab algorithm of modified DTW. (D1, D2), (D3, D2) with modified sequences MD1, MD3 are shown as Fig. 5a, b. SUM of (D1, D2), (MD1, D2), (MD3, D2) after DTW processing is 59.4, 30.1, 498.1, respectively. The result reveals that average modification can improve similarity between data sequences and discriminability of magnetic matching positioning with modified DTW algorithm is satisfactory.

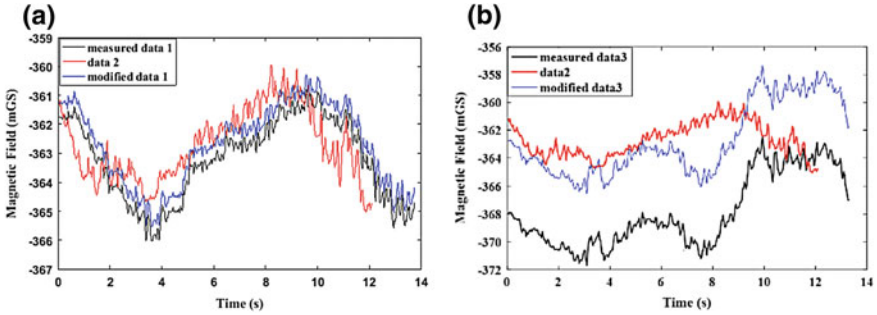



**Fig. 5.** Magnetic field strength signal alignment by modified DTW of **a** two walks through the same path **b** two walks through two nearby paths

## 5 Experiment and Discussion

To examine the functionality and the precision of the proposed indoor positioning scheme with the positioning box worn, I walk from the first floor to the second floor of a building in my campus. Layout of two floors is approximately the same, as shown in Fig. 6, with position of stairs marked. After constructing database of MSM by mobile mapping, I walked as the path of the red line in Floor 1, up two flights of stairs (20 steps), and as the path of green line of Floor 2 (Fig. 6). An additional experiment of walking downstairs is conducted, afterward. During the walking, a number of steps in Floor 1 and Floor 2 are counted. Magnetic matching corrects obvious route error by PDR which crosses areas with large magnetic field strength change. PDR path and magnetic matching correcting (MMC) path are marked in blue, yellow lines (Fig. 6). PDR path can roughly indicate the walking path and MMC path decreases errors, especially those caused by heading angle estimation (accumulative error by the gyroscope). Curves of PDR error, MMC error, and error of heading angle at corners (Fig. 7) reflect precision of positioning. It can be seen that maximum PDR error and PDR error at ending point are 3.5 m and 3 m, respectively. Heading angle error is

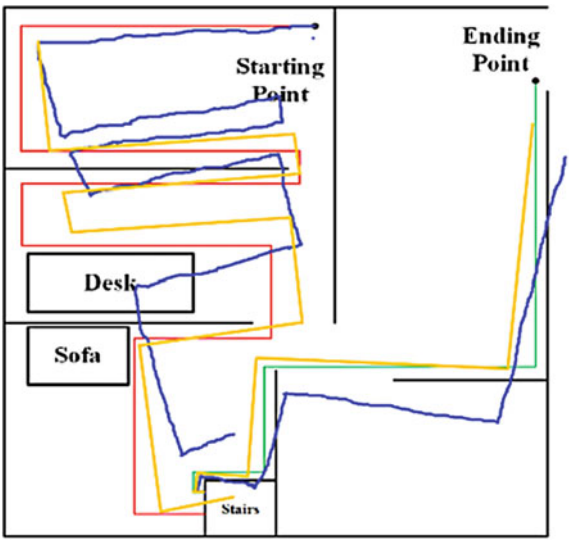



**Fig. 6.** Similar layout of Floor 1 and Floor 2

within 20°. After correcting of heading angle error (MMC path), positioning error at ending point is decreased to 1.5 m. Heading angle correction can effectively find out PDR route errors of crossing walls or other barriers. Figure 7b reveals that heading angle error accumulates with successive turn of the same direction, which derives from scale factor error of the gyroscope. Counted upstairs and downstairs numbers are 19, 20 respectively, with 1 step upstairs count error. The accuracy rate of stair count is high enough.

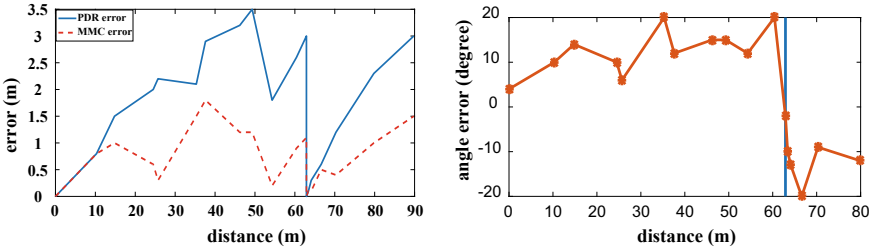



**Fig. 7.** **a** PDR and MMC error **b** heading angle error (blue line represents position of stairs)

## 6 Conclusion

An indoor positioning scheme combining PDR and MSM is proposed in the paper. The 3-axis accelerometer and the magnetometer in iphone7 with a fabricated high-precision gyroscope construct the hardware system of the positioning box. Through Kalman filtering to outputs from accelerator and gyroscope, typical human motions during

walking are recognized with low error rate. MSM positioning as a supplementary method to decrease accumulative errors is conducted through data matching between measured magnetic strength sequence and sequences in database by modified DTW algorithm. The proposed positioning scheme is applied in an experiment of walking path tracking from floor 1 to floor 2 in a building of my campus. PDR error caused by heading angle estimation error is obvious. After MSM correction, positioning error is lowered. Further work will focus on developing a real-time data processing module to process MSM data as the walking action and correct heading angle error in real time to decrease accumulative effect of PDR positioning by inertial sensors.

## References


1. He Z. High-sensitivity GNSS Doppler and velocity estimation for indoor navigation. Engineering Engineering–AerospaceEngineering–Electronics and Electrical, 2013.
2. Xiao Z, Wen H, Markham A, Trigoni N, Blunsom P, Frolik J. Non-line-of-sight identification and mitigation using received signal strength. IEEE Trans Wirel Commun. 2015;14:1689–702.
3. Schmitt S, Adler S, Kyas M. The effects of human body shadowing in RF-based indoor localization. In: 30th international conference on indoor positioning and indoor navigation, 2014.
4. Saeedi S. Context-aware personal navigation services using multi-level sensor fusion algorithm. Ph.D., University of Calgary, 2013.
5. Morrison A, Renaudin V, Bancroft JB, Lachapelle G. Design and testing of a multi-sensor pedestrian location and navigation platform. Sensors. 2012;12:3720–38.
6. Talvite J, Renfors M, Lohan ES. Distance-based interpolation and extrapolation methods for RSS-based localization with indoor wireless signals. IEEE Trans Veh Technol. 2015;64 (4):1340–53.
7. Cheng Y, Wang X, Morelande M, Moran B. Information geometry of target tracking sensor networks. Inf Fusion. 2013;14:311–26.
8. Torres-solis J, Falk TH, Chau T. A review of indoor localization technologies: toward navigational assistance for topographical disorientation. Ambient Intell. 2010;51–84.
9. Bose A, Foh CH. A practical path loss model for indoor WiFi positioning enhancement. In: 6th international conference on information, communication & signal processing. IEEE; 2007. p. 1–5.
10. Niu X, Li Y, Zhang H, Wang Q, Ban Y. Fast thermal calibration of low-grade inertial sensors and inertial measurement units. Sensors. 2013;13:12192–217.
11. Wang Q, Li Y, Niu X. Thermal calibration procedure and thermal characterisation of low-cost inertial measurement units. J Navig. 2015;1–18.
12. Akeila E, Salcic Z, Swain A. Reducing low-cost INS error accumulation in distance estimation using self-resetting. Trans Instrum Meas IEEE. 2014;63:177–84.
13. Xie H, Gu T, Tao X, Ye H, Lv J. Maloc. A practical magnetic fingerprinting approach to indoor localization using smartphones. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing. ACM; 2014. p. 243–53.
14. Zhang C, Subbu K, Luo J, Wu J. GROPING: geomagnetism and crowdsensing powered indoor navigation. IEEE Trans Mob Comput. 2015;14(2):387–400.
15. Pritt N. Indoor navigation with use of geomagnetic anomalies. In: Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International; IEEE. 2014. p. 1859–62.

16. Li Y, Zhuang Y, Lan H, Zhang P, Niu X, El-sheimy N. WiFi-aided magnetic matching for indoor navigation with consumer portable devices. Micromachines. 2015;6:747–64.
17. Chen D, Liu X, Zhang H, Li H, Weng R, Li L, Rong W, Zhang Z. A rotational gyroscope with a water-film bearing based on magnetic self-restoring effect. Sensors. 2018;18(2).
18. Chen D, Liu X, Zhang H, Li H, Weng R, Li L, Rong W, Zhang Z. Friction reduction for a rotational gyroscope with mechanical support by fabrication of a biomimetic superhydrophobic surface on a ball-disk shaped rotor and the fabrication of a water film bearing. Micromachines. 2017;8(7):223.
19. Zhen D, Zhao H, Gu F, Ball A. Phase-compensation-based dynamic time warping for fault diagnosis using the motor current signal. Meas Sci Technol. 2012;23:055601.

# Design and Verification of Anti-radiation SPI Interface in Dual Mode Satellite Navigation Receiver Baseband Chip

Yi Ran Yin(✉) and Xiao Lin Zhang

Electronic and Information Engineering, BeiHang University,
XueYuan Road no. 37, Haidian District, Beijing, China
728660626@qq.com

**Abstract.** This paper designs the SPI interface module in the baseband chip of the dual mode satellite navigation receiver based on the AMBA bus, and this paper is based on the first edition of the baseband chip accepted by the Science and Industry Corp. This design provides the IP core of SPI for the SoC using the LEON series processor and uses the TMR on register transfer level and evaluates the final results after the reinforcement. The results of verification and evaluation show that the SPI master designed by this paper can communicate normally based on the AMBA bus, and the SPI master after reinforcement can resist SEU to a certain extent and improve its stability in the space radiation environment.

**Keywords:** Anti-radiation · Baseband chip · AMBA bus

## 1 Introduction

The digital baseband chip is an important part of the navigation receiver compatible with the GPS and BD-2 systems and is the core unit of the dual system compatible receiver. This paper designs the SPI module as a part of a major special project. From a global point of view, there are many kinds of on-chip buses, and many semiconductor manufacturers have developed their own bus standards. The literature [1] shows that the wishbone bus has been widely used in the market with its diverse topological structures and most of the IP cores of SPI are currently designed based on the wishbone bus. The literature [2] shows that the AMBA bus designed by ARM has been widely recognized in the market with excellent architecture. The literature [3] shows that the failure of the space radiation in the satellite accounts for 55% of the total number of faults, so the stability of the integrated circuit is very important. So this paper describes the top-down design process for SPI based on the AMBA bus in detail and verifies the design with the LEON3 platform.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_135

## 2 Design of SPI Bus Control Module

There are two main methods for designing IP cores. One is top-down, and the other is from bottom to top. Because the structure of SPI master is relatively simple and there are less modules, it is relatively easy to grasp from the whole. Therefore, the top-down approach is used to design.

The main function of the SPI master is to define and describe the related registers, so the APB bus can control the SPI master to select the SPI slave and to receive data from the SPI slave and to send data to SPI slave. The architecture of the SPI master section is shown in Fig. 1.

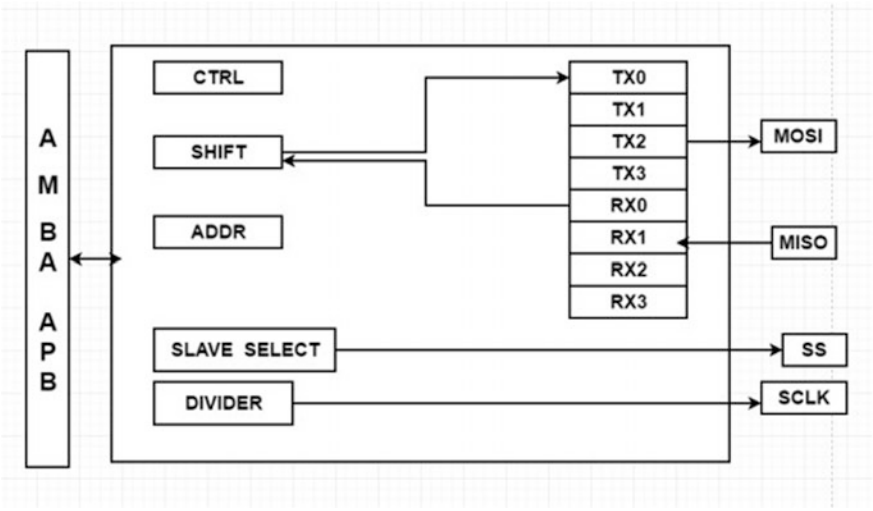



**Fig. 1.** Architecture of SPI master

The modules involved in Fig. 1 include: control and state register module, clock generation module, shift register module, address decoding module, etc.

### 2.1 Clock Generation Module

In order to match the high-speed clock of the APB bus and the low-speed clock of the external device, the clock division module 'spi_clgen' is needed in the design of the SPI master. The main function of the module is to divide the clock frequency PCLK from the APB bus and generate the clock signal SCLK to control the sending and receiving data.

This module divides the frequency from the APB bus and generates the internal clock signal. The frequency division function is realized by the reciprocal counter (CNT). The initial value of the counter (CNT) is stored in the register named 'divider.' The idle state of the 'clk_out' is low, and the value of the counter (CNT) is reduced by 1 at the coming of the rising edge of each PCLK. When the value of counter (CNT) is

reduced to 0, the value of the signal named 'clk_out' is turned to high level, and the high level lasts for a PCLK period, and then, it returns to low level. The signal passes through the transceiver logic module and generates a signal which is half of the signal named 'clk_out.' The signal generated is the main clock signal of the SPI communication.

### 2.2 Address Decoding Module

When it needs to write data or read data, the module will compare the address from the APB bus with the defined address constants one by one. If the APB bus address is consistent with the defined address, it performs the operations of reading and writing. In this judgment, the conditions should be met; that is, only when the PENBALE is enabled, PSELx chooses the corresponding SPI slave. The address signal named PADDR decodes the address and configures corresponding registers when it receives data.

### 2.3 Shift Register Module

The main function of this module is to transform parallel data from APB bus to serial data and deliver the serial data to the SPI master, transform serial data from SPI master to parallel data and deliver the parallel data to APB bus.

The main function of this module is completing the conversion of serial and parallel data through the shift register. The specific implementation process is as follows: The serial data named 's_out' from SPI slave stores in the receiving register RX in a certain order and then the data in the RX is passed to the register named 'data.' When the APB bus sends the corresponding task to the SPI master, the SPI master sends the data stored in the register named 'data' in parallel to the APB bus. In the same way, when the APB bus sends the corresponding command to the SPI master, SPI stores the parallel data named 'p_out' from the APB bus in the shift register named 'data,' and the data in the shift register is then transferred to the register TX in a certain order. Then, the data transfers to serial data and delivers to the SPI slave one by one.

### 2.4 Control and State Register Module

The main function of the control and state register is to control the transmission state and transmission mode of the SPI master. The transfer mode can be controlled by setting the related bits of the register. The control and state register has 32 bits, and the main internal signals are as follows: ASS, this bit is the twelfth bit of the register. If this bit is set, SS signals are generated automatically. IE, this bit is the eleventh bit of the register. If this bit is set, the interrupt output is set active after a transfer is finished. The interrupt signal is deasserted after reading from any registers or writing to any registers. TX_NEG, this bit is the second bit of the register. If this bit is set, the MOSI signal is changed on the falling edge of a SCLK clock signal, or otherwise, the MOSI signal is changed on the rising edge of SCLK.

## 3 Simulation and Verification of SPI Master Interface

This paper verifies the designed SPI master in two aspects: One part is writing testbench to do the preliminary verification, and the other is linking the design which has passed the initial verification to the LEON3 platform provided by ESA.

### 3.1 Write a Testbench for Verification

Before connecting the designed SPI master to the LEON3 SoC platform, we first need to write a testbench to verify the designed code preliminarily, and test whether the function of the designed circuit is consistent with the expected target. The structure diagram of testbench is shown in Fig. 2.

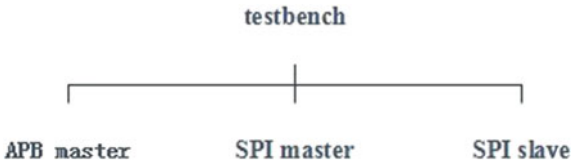



**Fig. 2.** Architecture of testbench

This paper uses Verilog to design the basic functions of the APB bus to verify the normal work of the designed SPI master. In this section, the writing task, reading task, and data comparison task of the APB bus are defined. In the testbench, this paper describes the corresponding WADDR and the corresponding WDATA, and the corresponding RADDR and the corresponding RDATA in detail. Through the COMPARE TASK this paper defined, we can judge whether the APB bus can write data to the corresponding address and read data from the corresponding address. The result of the test is shown in Fig. 3.

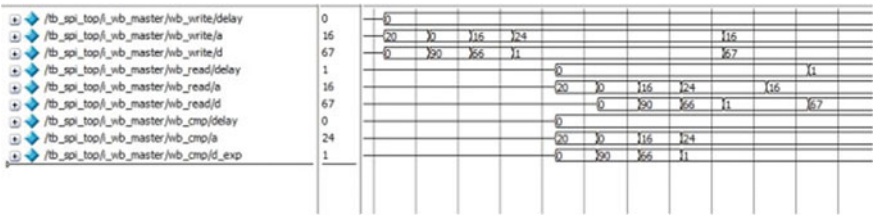



**Fig. 3.** Test of communication between APB and SPI master

In order to improve the reliability of the design results, this paper verifies the function of the designed SPI master fully. The SPI master also reads and writes 1000 times randomly to the related registers and compares the correctness of the written data and the data which is read from the related registers. The result of the test is shown in Fig. 4.

```
# the result is right, rand is 00000000000000000000011010000101,prdata is 00000000000000000000011010000101
# the result is right, rand is 11111111111111111111101000010111101,prdata is 11111111111111111111101000010111101
# the result is right, rand is 00000000000000000000010000010111,prdata is 00000000000000000000010000010111
# the result is right, rand is 00000000000000000000111001101010101,prdata is 00000000000000000000111001101010101
# the result is right, rand is 00000000000000000000011111001111101,prdata is 00000000000000000000011111001111101
# the result is right, rand is 11111111111111111101101101110111,prdata is 11111111111111111101101101110111
# the result is right, rand is 00000000000000000000010110101001100,prdata is 00000000000000000000010110101001100
# the result is right, rand is 11111111111111111111010000000000111,prdata is 11111111111111111111010000000000111
# the result is right, rand is 11111111111111111111110010110111,prdata is 11111111111111111111110010110111
# the result is right, rand is 11111111111111111101101001101001,prdata is 11111111111111111101101001101001
# the result is right, rand is 11111111111111111101100101110001,prdata is 11111111111111111101100101110001
# the result is right, rand is 00000000000000000000010010100111011,prdata is 00000000000000000000010010100111011
# the result is right, rand is 00000000000000000000010100111011001,prdata is 00000000000000000000010100111011001
# the result is right, rand is 00000000000000000000010010000100001,prdata is 00000000000000000000010010000100001
# the result is right, rand is 11111111111111111111101010100001,prdata is 11111111111111111111101010100001
# the result is right, rand is 00000000000000000000010110000101010,prdata is 00000000000000000000010110000101010
# the result is right, rand is 00000000000000000000010010111010,prdata is 00000000000000000000010010111010
# the result is right, rand is 11111111111111111110010110010101,prdata is 11111111111111111110010110010101
# the result is right, rand is 00000000000000000000011011010011,prdata is 00000000000000000000011011010011
# the result is right, rand is 11111111111111111111101000001100,prdata is 11111111111111111111101000001100
# the result is right, rand is 11111111111111111101101010010010,prdata is 11111111111111111101101010010010
# the result is right, rand is 11111111111111111101111010111110,prdata is 11111111111111111101111010111110
# the result is right, rand is 11111111111111111101001010111000,prdata is 11111111111111111101001010111000
# the result is right, rand is 11111111111111111101111001011111,prdata is 11111111111111111101111001011111
# the result is right, rand is 00000000000000000001000101100110,prdata is 00000000000000000001000101100110
# the result is right, rand is 00000000000000000001000001010000,prdata is 00000000000000000001000001010000
# the result is right, rand is 00000000000000000011000101110000,prdata is 00000000000000000011000101110000
# the result is right, rand is 00000000000000000000010101000000,prdata is 00000000000000000000010101000000
```

**Fig. 4.** Final test of communication between APB and SPI master

### 3.2 Connect to the LEON3 Platform for Verification

LEON3 uses the 7-level pipeline structure, and it connects the INSTRUCTION REGISTER with the MEMORY CONTROL REGISTER through the high-speed AMBA-AHB bus. The DATA CACHE REGISTER is connected with the high-speed external interface to transmit data. Low-speed AMBA-APB bus connects with low-power peripherals, such as timers, serial ports, and network interfaces [4].

To connect to the LEON3 platform, it is needed to understand the plug-and-play mechanism of APB bus, configure the designed SPI address, and connect the SPI master to the LEON3 platform. This paper uses the Verilog language for the design of SPI, while the open source code of LEON3 uses the VHDL language, so this paper needs to write a wrapper to complete the conversion of two languages.

For this paper, the function of verifying the designed SPI is to write to the SPI master and read data from the SPI master through the APB bus. The APB bus writes data to the related registers to transmit the related data and configure the related registers.

The basic function of this code is to write data $0 \times 5a$ to the TX register of the SPI master. By comparing the data read out with the written data, we can judge whether the SPI master can be used to transmit data with the APB bus.

This code is written in/grlib/software/leon3/apbuart.c. This paper writes the test function int spitest(int addr) and invokes the function in systest.c. When this paper invokes the function, the APB bus must know the address of the related registers to be written.

According to the above method, the APB bus can get the address of the related registers. So, the APB bus can write data into the related registers and read data from the related registers. The final results are shown in Fig. 5.

As is shown in Fig. 5, first, 000005A is written to the DIVIDER REGISTER of the SPI master. The address of the DIVIDER REGISTER is 0000084, and then, the APB


**Fig. 5.** Final result of test

bus can read 000005A from the DIVIDER REGISTER. In the same way, 000000 is written to the SS REGISTER and the APB bus can read the same data from the SS REGISTER. The address of the SS REGISTER is 00000818.

It can be concluded that the SPI master has been successfully connected to the APB bus and can be written and read out through the APB bus.

## 4 TMR of the SPI Master

In the actual aerospace circuit, to meet the reliability requirements of SPI, we usually use the fault-tolerant technology to ensure the function of the SPI master. The most common fault-tolerant technology is TMR. The main parts of TMR are three identical modules and a voter [5]. The output of the voter is in accordance with the majority of the three inputs. The precondition of the voter is that there are two or more than two modules working normally in the circuit, and the overall circuit function is normal, so that the fault of the single circuit module is shielded.

TMR technology is considered the most reliable anti-radiation method. To achieve the best effect of anti-SEU, we can instantiate all the modules into three identical modules. While this method will bring about an increase in resources and cost, we should consider the importance of the module and sensitivity to SEU. This paper only reinforces the main registers of the SPI master. There are the data receive and transmit registers named RX0, RX1, RX2, RX3; slave select register named SS; divider register named DIVIDER; control and state register named CTRL.

The main implementation process is that this paper instantiates three identical modules and assumes that one of the three registers has SEU that means the data in this register has changed. Then, this paper observes the output of the register whether has changed; if there is no change in the output, we can consider that the SPI master designed in this paper has a certain ability of anti-radiation. Because the probability of happening SEU at one time is very low and the probability of happening SEU is lower,

it can be regarded as impossible. In this paper, we only discuss the situation of only one register has SEU [6].

To judge whether the design achieves the purpose of reinforcement, this paper verifies the final results after TMR. The results of TMR are shown in Fig. 6.

This paper also reinforces other registers of the SPI master in the same way. The total test result after TMR is shown in Fig. 7.

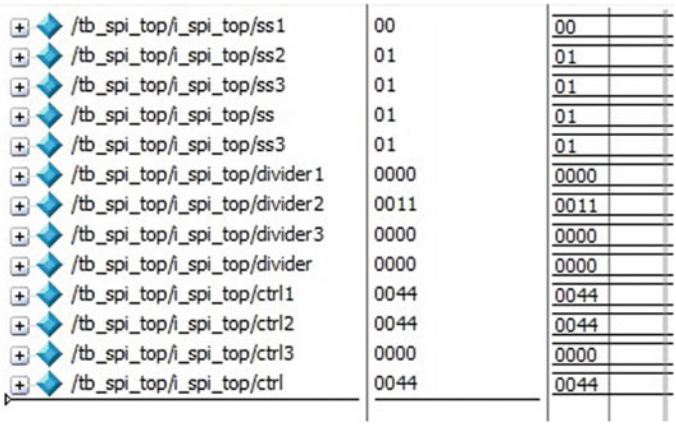



**Fig. 6.** Test result of TMR

```
# status:                    0 Testbench started
#
#
# status:                19500 done reset
# status:                28600 programmed registers
# status:                36600 verified registers
# status:                38600 generate transfer:  8 bit, msb first, tx posedge, rx negedge
# status:                61600 transfer completed: ok
# status:                67600 generate transfer:  8 bit, msb first, tx negedge, rx posedge
# status:                90600 transfer completed: ok
# status:                96600 generate transfer: 16 bit, lsb first, tx negedge, rx posedge
# status:               134600 transfer completed: ok
# status:               142600 generate transfer: 64 bit, lsb first, tx posedge, rx negedge
run
# status:               279600 transfer completed: ok
# status:               291600 generate transfer: 128 bit, msb first, tx posedge, rx negedge
run
# status:               563600 transfer completed: ok
# status:               569600 generate transfer: 32 bit, msb first, tx negedge, rx posedge, ie
VSIM 31> run
# status:               639500 transfer completed: ok
# status:               645600 generate transfer: 32 bit, msb first, tx posedge, rx negedge, ie, ass
# status:               715500 transfer completed: ok
# status:               721600 generate transfer: 1 bit, msb first, tx posedge, rx negedge, ie, ass
# status:               729500 transfer completed: ok
#
#
# status:               729500 Testbench done
```

**Fig. 7.** Final test result of TMR

## 5 Concluding Remarks

This paper completes the design and verification of the SPI interface module in the baseband chip of the dual mode satellite navigation receiver, reinforces the SPI interface modules through the TMR method, and completes simple evaluation after reinforcement. In summary, this paper improves the radiation resistance and improves the stability of its work.

**Acknowledgments.** This work was partially supported by the National Natural Science Foundation of China under Grant No.61601295.

This work was financially supported by Xi'an Aisheng innovation and Development Foundation (ASN-IF2015-1405).

## References

1. Tian Ze, Zhang YH, Yu DS. The Summary of SoC OCB. Semicond. T Echnology. 2003;11:11–5.
2. ARM Ltd. AMBA Specification Revision 2. 1999.
3. Hui X, Wei T, Yuan HY, Liang Z. Research of radiation harden based on triple modular redundancy for ASIC. Microprocessors. 2015;10:1–4.
4. Lin L, Lin ZX, Xi Y. Building and testing of SoC platform based on LEON open source soft core. Microcontrollers Embed. Syst. 2007;01:32–5.
5. Hua GJ, Run X, Lin Z. The research of radiation harden for ASIC. J. CAEIT. 2013;12: 644–645.
6. Xing K, Yang J. Study on radiation harden method for SRAM based FPGA in space. Chin. J. Electron Devices. 2007;30(1).

# Indoor Localization Algorithm Based on Particle Filter Optimization in NLOS Environment

Weiwei Liu(✉), Tingting Liu, and Lei Tang

Nanjing Institute of Technology, Hongjing Road 1, Nanjing, China
466346830@qq.com, lwwlll7@njit.edu.cn

**Abstract.** The performance of indoor localization algorithm is limited by non-line-of-sight (NLOS) error, a positioning system includes Bluetooth module, Bluetooth gateway and cloud monitoring center based on particle filter is presented to enhance positioning accuracy. Our experimental results indicate that the proposed localization scheme leads to higher localization accuracy and lower power consumption.

**Keywords:** Localization · Non-line-of-sight · Particle filter

## 1 Introduction

Over the past decades, localization service has been applied to a variety of different applications, especially in civil class positioning navigation. While Global Positioning System (GPS) localization does not work well in certain indoor regions. To tackle the problems with GPS, many researchers have proposed a series of alternative indoor localization schemes. Ultrasound positioning has high accuracy and simple structure, but it is greatly affected by multipath effect and non-line-of-sight propagation. Wi-Fi is applied to small-scale indoor positioning with low cost; meanwhile, its accuracy is easy to be affected by other signals. ZigBee location determines the location of the object by calculating the distance between the object and the reference nodes with known locations, which is not common in practical applications. Based on Bluetooth positioning network, it can span a large physical space, increases the coverage of monitoring area, reduces blind area, and be suitable for remote communication process. That is to say, the Bluetooth LAN with intelligent terminals and several Bluetooth beacons, which transmits data between Bluetooth gateway and cloud monitoring center [1–9].

At present, the Bluetooth location method mainly is based on triangular location algorithm. Other auxiliary methods such as Bayesian-based [10] sequential Monte Carlo [11–13] particle filter (PF), Kalman filter algorithm, and Gauss filter algorithm.

Many improved methods have been proposed [14–16]. Xie et al. [14] came up with an unscented particle filter (UPF). The main idea of UPF is to use unscented Kalman filter method integrating the latest observation information to get a good proposal distribution. Gaussian particle filter (GPF) was also proposed to solve degeneracy problem [16] with a low error when the environment is highly non-Gaussian, especially in line-of-sight (NLOS) environment.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_136

In our research, we focus on the outdoor car localization schemes that can be adopted by smartphones. In this paper, we present the original of car location detection algorithm in the database of cloud data center maps the one-to-one mapping relationship between car location and intelligent LED positioning lamp. The mobile phone submits the query request service through the beacon protocol communication. The nearest Bluetooth lamp receives the request information and broadcasts the request information to the LED wireless ad hoc MESH network for broadcast traversal to search of car bits information, which reduces the cost of hardware compared to traditional video parking detection methods.

The rest of this paper is organized as follows. In Sect. 1, the related work on existing localization schemes is presented. The problem formulation and describes the proposed hybrid localization scheme are included in Sect. 2. Section 3 includes the details of our experimental results. Finally, our conclusions are presented in Sect. 4.

## 2 The Problem Formulation and Describes the Proposed Hybrid Localization Scheme

Particle filter is to generate a set of random samples (particles). The weights and positions of the particles are adjusted according to the measurement information, and the initial empirical condition distribution is modified by the adjusted particle information. Observation model is given as

$$z_k = o_k(x_k, v_k, u_k) \tag{1}$$

where $z_k$ is observation position at time $k$, $o_k$ is Observation function, $v_k$ and $u_k$ are observation noise and test noise, respectively.

### 2.1 The Model Structure

The model structure of the localization in NLOS environment, which is shown in Fig. 1. In order to realize the stochastic characteristic analysis of the NLOS environment in the region, the dynamic process of observation noise and test noise should be modeled first in the region.

### 2.2 The Proposed Hybrid Localization Scheme

The position determines its position based on the LED data received with non-line-of-sight (NLOS) serving for this investigation. The optical power received from a single LED without overlap is given as

$$x_k^i \sim q(x_k | x_{1:k-1}^i, z_{1:k}) \tag{2}$$

where $k$ is the time indicator, $x_k = \left\{x_k^1, x_k^2, \cdots x_k^N\right\}$ is the system state vector consisting of all states at the time $k$, $z_{1:k} = \{z_1, z_2, \cdots z_k\}$ is a vector set of all measurement, $q(\bullet)$ is the prior probability function, $i \in [1, N]$, $N$ is the number of generated samples.

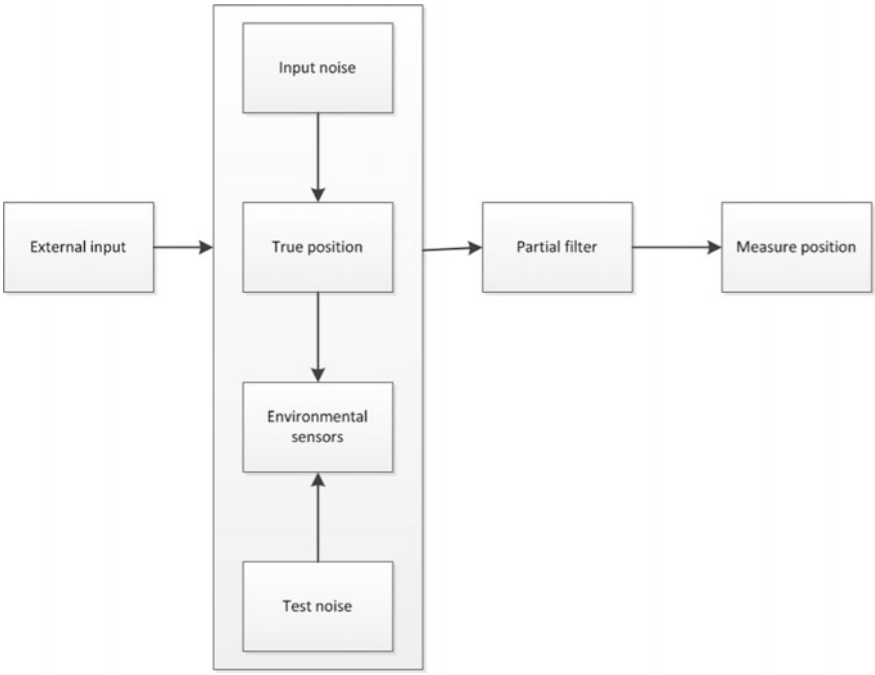



**Fig. 1.** Model structure

Suppose

$$w_k^i = w_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{0:k-1}^i, z_{1:k})} \tag{3}$$

Denote their corresponding weights.
Thus, we have

$$p(x_k|z_{1:k}) = \sum_1^N w_k^i \delta(x_k - x_k^i) \tag{4}$$

as random quantity to represent the posteriori probability density function. Here, the weight satisfies the regular condition. Where $\delta$ is the Dirac delta function.

After updating the state of particles, we weighted and estimated the location of particles.

$$\widehat{x}_k = \sum_1^N w_k^i x_k^i \tag{5}$$

## 3 Simulation Result

In this paper, we propose a position mechanism, and the simulation results of particle filtering are shown as shown. Then, we discuss the effects of car position data.

### 3.1 Localization Accuracy

In our research, we implemented a testbed around the Academic Building at Nanjing Institute of Technology, China. In this testbed, an android system running the proposed localization scheme was used to achieve localization.

Figure 2 shows the evolution of the entire system. On the one hand, by using the original measurement equations and the random disturbances, measurements are obtained in the forward process. On the other hand, by using the partial filter process of the system, measurement function and the partial observation of the system state, we are able to realize a dynamic backtracking from the measurement data to the system state. We compare the effect of the improved state evolution process with the process without the direct observation for partial system states. And this dimensionality of the system is tested to have the best performance.

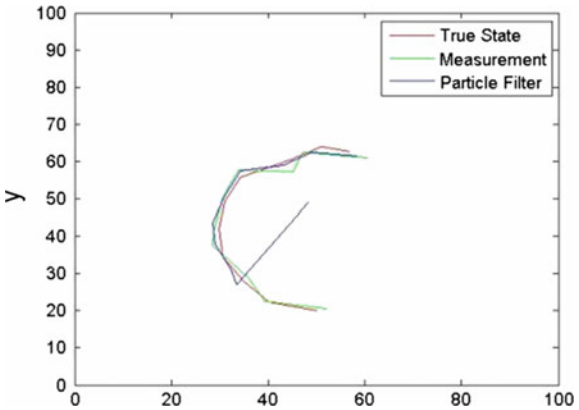



**Fig. 2.** Evolution of the entire system

In the following section, we present a numerical case to the effect of the indoor navigation observation for partial system.

### 3.2 The Effective of the Proposed System

We plotted these results into indoor car position as Fig. 3 shows.

The car owners can match the real-time parking demand and parking lot of accurate data through the mobile phone APP intelligent parking. The intelligent parking owners completely and efficiently realize the parking lot to maximize the use of resources in Fig. 3a. The parking users who are needed to parking can get parking navigation

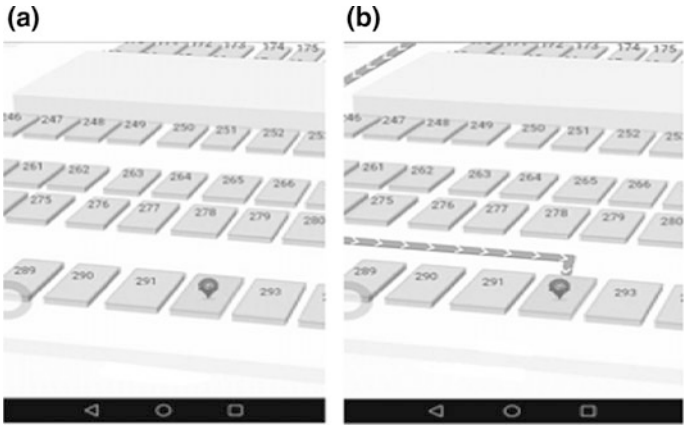



**Fig. 3.** **a** Three-dimensional map distribution of empty car position. **b** Empty car position navigation

services in Fig. 3b. The parking lot operators can master the parking data in real time through the system. The analysis of parking big data can provide strong decision-making basis for operators to improve the operation efficiency of parking lots, and support the unified management of multiple parking lots and the unified management of multiple massive devices.

## 4 Conclusions

In this paper, we present an indoor car localization scheme, which utilizes smartphones. With the partial filter mechanism, this scheme is expected to be more feasible than the existing indoor localization methods. Our experimental results indicate that the proposed localization scheme leads to higher localization accuracy and lower power consumption.

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation of China under Grant 61702258, in part by the China Postdoctoral Science Foundation under grant 2016M591852, in part by Postdoctoral research funding program of Jiangsu Province under grant 1601257C, in part by the China Scholarship Council Grant 201708320001 and the NJIT Foundation(Grant No. YKJ201419).

## References

1. Chen R, Guinness R. Geospatial computing in mobile devices. Artech House. 2014.
2. Zhou B, Li Q, Mao Q. Activity sequence-based indoor pedestrian localization using smartphones. IEEE Trans Hum-Mach Syst. 2015;45(5):562–74.
3. Groves PD. principles of GNSS inertial, and multisensor integrated navigation systems, 2nd edn. Aerosp. Electron. Syst. Mag. IEEE. 2015;30(2):26–27.

4. Shi K, Ma Z, Zhang R. Support vector regression based indoor location in IEEE 802.11 Environments. Mob Inf Syst. 2015.
5. Harle R. A survey of indoor inertial positioning systems for pedestrians. Commun Surv Tutor, IEEE. 2013;15(3):1281–93.
6. Cheng J, Yang L, Li Y. Seamless outdoor/indoor navigation with WIFI/GPS aided low cost inertial navigation system. Phys Commun. 2014;13:31–43.
7. Chen Z, Zou H, Jiang H. Fusion of WiFi, smartphone sensors and landmarks using the Kalman filter for indoor localization. Sensors. 2015;15(1):715–32.
8. Jiao J, Courtade TA, Venkat K. Justification of logarithmic loss via the benefit of side information. Inf Theory IEEE Trans On. 2015;61(10):5357–65.
9. Han B, Joo SW, Davis LS. Adaptive resource management for sensor fusion in visual tracking. Theory and applications of smart cameras. Springer Netherlands; 2016. p. 187–213.
10. Srkk S. Bayesian filtering and smoothing. Inst Math Stat Textb. 2013;40(4):S407.
11. Lu X,Dong Y, Wang XA. Monte carlo localization algorithm for 2-D indoor self-localization based on magnetic field. International ICST conference on communications and networking in China. IEEE; 2013. p. 563–8.
12. Smith A. Sequential monte carlo methods in practice. Springer Science & Business Media; 2013.
13. Klaas M, Freitas ND, Doucet A. Toward practical N2 monte carlo: the marginal particle filter. J Neurophysiol. 2012;94(3):2045–52.
14. Xie H, Gu T, Tao X. A reliability-augmented particle filter for magnetic fingerprinting based indoor localization on smartphone. IEEE Trans Mob Comput. 2016;15(8):1877–92.
15. Yin S, Zhu X. Intelligent particle filter and its application to fault detection of nonlinear system. Ind Electron, IEEE Trans On. 2015;62(6):3852–61.
16. Ades M, Van Leeuwen PJ. An exploration of the equivalent weights particle filter. Q J R Meteorol Soc. 2013;139(672):820–40.

# An Improved Sensor Selection for TDOA-Based Localization with Correlated Measurement Noise

Yue Zhao[1(✉)], Zan Li[1], Feifei Gao[2], Jia Shi[1], Benjian Hao[1,3], and Chenxi Li[1]

[1] State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China
{yuezhao, chenxili}@stu.xidian.edu.cn
{zanli, jiashi, bjhao}@xidian.edu.cn
[2] National Laboratory for Information Science and Technology, Tsinghua University, Beijing, China
feifeigao@ieee.org
[3] Collaborative Innovation Center of Information Sensing and Understanding, Xi'an, China

**Abstract.** This paper focuses on the problem of sensor selection in time-difference-of-arrival (TDOA) localization scenario with correlated measurement noise. The challenge lies in how to select the reference sensor and ordinary sensors simultaneously when the TDOA measurement noises are correlated. Specifically, the optimal sensor subset is found by introducing two independent Boolean selection vectors and formulating a nonconvex optimization problem, which motivates to minimize the localization error in the presence of correlated noise and energy constraints. Upon transforming the original nonconvex problem to the semidefinite program (SDP), the randomization method is leveraged to tackle the problem, and thereby proposing the novel algorithm for sensor selection. Simulations are included to validate the performance of proposed algorithm by comparing with the exhaustive search method.

**Keywords:** Sensor selection · Time-difference-of-arrival · Source localization · Convex optimization

## 1 Introduction

Source localization has been of considerable interest for many applications in radar [1], wireless communications [2], and spectrum monitoring [3]. In the wireless sensor network, source localization can be boosted by time-of-arrival (TOA), time-difference-of-arrival (TDOA), frequency-difference-of-arrival (FDOA), angle-of-arrival (AOA) measurements, or a combination of them [4,5]. This paper focuses on the TDOA localization for a stationary source.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_137

In practice, sensors are always deployed in the harsh conditions and powered by batteries [6]. Although localization with a great number of sensors can improve the accuracy [7], the system energy consumption and complexity are also enhanced [8]. To balance the energy consumption, the number of sensors, with the localization accuracy [7], the sensor selection problem arises and aims at invoking only a subset of sensors used for localization.

The straightforward approach to obtain the optimal sensor subset is enumerating all the possible sensor combinations and selecting one which has the optimal performance. However, this solver is computationally intensive and cannot be applied to the actual system. A heuristic approach has been proposed in [9] to solve the sensor selection problem by convex optimization. The authors of [10] have proposed a method that selects the minimum number of sensors in the premise of guaranteeing a certain estimation performance. All the above literatures deal with sensor selection problems by convex relaxation that relaxes the original nonconvex problem as a semidefinite program (SDP). In addition, several papers addressed the sensor selection problem by greedy approaches [11,12].

In the existing literatures, whether the measurement model is linear or nonlinear, only single-sensor selection vector is used to determine whether a sensor is selected. However, in TDOA localization scenario, both reference sensor and other ordinary sensors have a great influence on localization performance [13]. Besides, ubiquitous TDOA correlated noise makes the sensor selection problem more complicated since the contribution of each sensor to the Fisher information matrix (FIM) is no longer in an additive manner [14]. Therefore, development of sensor selection scheme in TDOA localization scenario in the presence of correlated noise, which can select the optimal reference sensor and ordinary sensors simultaneously, is crucial.

This paper develops an optimization framework for sensor selection in TDOA localization based on two Boolean vectors, which motivates to minimize the localization error in the presence of correlated noise and energy constraints. Though not focusing on a specific estimate algorithm, the Cramer-Rao Lower Bound (CRLB) is used as the performance metric [10], which results in the lowest possible variance that an unbiased linear estimator can achieve [2] and the trace being the minimum achievable localization mean square error (MSE) [15].

The rest of this paper is organized as follows. Section 2 introduces the measurement model and CRLB. Section 3 presents the problem formulation. Sensor selection approach is derived in Sect. 4. Section 5 presents the simulation results. Finally, conclusions will be drawn in Sect. 6.

## 2 Measurement Model and CRLB

We consider a wireless radio-based sensor localization network scenario where $M$ sensors are used to determine the source position $\boldsymbol{u}^o$ employing TDOAs. Let us denote the true position of the source as $\boldsymbol{u}^o = [x^o, y^o, z^o]^T$ and the position of sensor $i$ as $\boldsymbol{s}_i = [x_i, y_i, z_i]^T$, $i = 1, 2, \ldots, M$, where $(*)^T$ stands for transpose operation. The true distance between the source and sensor $i$ is

$$r_i^o = \|\boldsymbol{u}^o - \boldsymbol{s}_i\| = \sqrt{(\boldsymbol{u}^o - \boldsymbol{s}_i)^T(\boldsymbol{u}^o - \boldsymbol{s}_i)}, \tag{1}$$

where $\| * \|$ denotes the Euclidean distance. Assume that each sensor in the network can act as the reference sensor, and sensor $\boldsymbol{s}_1$ is regarded as the reference sensor in our analysis without loss of generality. The true TDOA of a signal received by the sensor pair $i$ and 1 is

$$t_{i1}^o = \frac{1}{c} r_{i1}^o = \frac{1}{c}(r_i^o - r_1^o), \tag{2}$$

where $c$ is the signal propagation speed.

In practice, the observed TDOA measurements are

$$\boldsymbol{t} = \boldsymbol{t}^o + \Delta\boldsymbol{t}, \tag{3}$$

where $\boldsymbol{t} = [t_{21}, \ldots, t_{(M-1)1}]^T$ is the collection of the TDOA measurements, $\Delta\boldsymbol{t} = [\Delta t_{21}, \ldots, \Delta t_{(M-1)1}]^T$ is the additive zero mean Gaussian noise with covariance matrix $\boldsymbol{Q}_t$. $\Delta t_{i1}$ is the difference of TOA noise $\Delta t_i$ and $\Delta t_1$, where $\Delta t_i$ is zero mean Gaussian noise with standard deviation $\sigma_i$ with defining $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_M)$. In this paper, we assume that TDOA noises between different sensor pairs are correlated due to the common reference, so the covariance matrix $\boldsymbol{Q}_t$ is not diagonal [16].

The CRLB of $\boldsymbol{u}$ is equal to the inverse of the Fisher information matrix [17]

$$CRLB(\boldsymbol{u}) = (\boldsymbol{J}_u)^{-1}, \tag{4}$$

where FIM is defined as

$$\boldsymbol{J}_u = \boldsymbol{H}\boldsymbol{Q}_t^{-1}\boldsymbol{H}^T = \left(\frac{\partial \boldsymbol{t}}{\partial \boldsymbol{u}^T}\right)^T \boldsymbol{Q}_t^{-1} \frac{\partial \boldsymbol{t}}{\partial \boldsymbol{u}^T}, \tag{5}$$

$$\frac{\partial \boldsymbol{t}}{\partial \boldsymbol{u}^T} = \frac{1}{c}\begin{bmatrix} \frac{(\boldsymbol{u}-\boldsymbol{s}_2)^T}{r_2} - \frac{(\boldsymbol{u}-\boldsymbol{s}_1)^T}{r_1} \\ \vdots \\ \frac{(\boldsymbol{u}-\boldsymbol{s}_M)^T}{r_M} - \frac{(\boldsymbol{u}-\boldsymbol{s}_1)^T}{r_1} \end{bmatrix}, \boldsymbol{Q}_t = \begin{pmatrix} \sigma_2^2 + \sigma_1^2 \ldots & \sigma_1^2 \\ \vdots \quad \ddots & \vdots \\ \sigma_1^2 \quad \ldots \sigma_M^2 + \sigma_1^2 \end{pmatrix}. \tag{6}$$

## 3 Problem Formulation

In this section, it discusses the general theory for the sensor selection problem in TDOA localization scenario when the noises are correlated, and the optimization problem is formulated and analyzed. In particular, we are motivated to select the best sensor subset with $K$ sensors of $M(M > K)$ available sensors that minimizing the localization error, which is subject to a constraint on the number of active sensors.

We introduce two Boolean vectors

$$\boldsymbol{q} = [q_1, q_2, \ldots, q_M]^T, \ q_i \in \{0, 1\} \tag{7}$$

$$\boldsymbol{p} = [p_1, p_2, \ldots, p_M]^T, \ p_j \in \{0, 1\} \tag{8}$$

to select the reference sensor and other ordinary sensors, respectively. The $i$th element of $\boldsymbol{q}$ describes whether or not the $i$th sensor is selected as the reference sensor and the $j$th element of $\boldsymbol{p}$ indicates whether or not the $j$th sensor is selected into the ordinary sensor subset.

We define two matrices $\boldsymbol{\Phi}_p$ and $\boldsymbol{\Phi}_q$, which are related to Boolean selection vectors and can formulate the FIM matrix and TDOA noise covariance matrix for selected sensors by picking out some columns or rows from information matrix. $\boldsymbol{\Phi}_p$ is a submatrix of $diag(\boldsymbol{p})$ after all columns corresponding to the unselected sensors are removed and $\boldsymbol{\Phi}_q$ is expanded by $\boldsymbol{q}$ as $\boldsymbol{\Phi}_q = [\boldsymbol{q}, \boldsymbol{q}, \ldots, \boldsymbol{q}]$.

Since two Boolean vectors are orthogonal, some properties are presented:

1. The products of two matrices:

$$\boldsymbol{\Phi}_p^T \boldsymbol{\Phi}_q = \mathbf{0}_{(K-1)}, \quad \boldsymbol{\Phi}_p \boldsymbol{\Phi}_q^T = \boldsymbol{p}\boldsymbol{q}^T, \quad \boldsymbol{\Phi}_p^T \boldsymbol{\Phi}_p = \boldsymbol{I}_{(K-1)}, \quad \boldsymbol{\Phi}_p \boldsymbol{\Phi}_p^T = diag\{\boldsymbol{p}\},$$
$$\boldsymbol{\Phi}_q \boldsymbol{\Phi}_q^T = (K-1) diag\{\boldsymbol{q}\}, \quad \boldsymbol{\Phi}_q^T \boldsymbol{\Phi}_q = \mathbf{1}_{(K-1)\times 1} \mathbf{1}_{1\times(K-1)},$$

   where $\boldsymbol{I}_{(K-1)}$ is an identity matrix with size $(K-1)$, $\mathbf{0}_{(K-1)}$ is a zero matrix with size $(K-1)$, $\mathbf{1}_{(K-1)\times 1}$ is a vector composed of '1' with size $(K-1) \times 1$.
2. We have

$$(\boldsymbol{\Phi}_p + \boldsymbol{\Phi}_q)^T \boldsymbol{E} (\boldsymbol{\Phi}_p + \boldsymbol{\Phi}_q) = (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T \boldsymbol{E} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q),$$

   where $\boldsymbol{E}$ can be any diagonal matrix. This property plays an important role in the derivation of optimization problem.

In our analysis, since the source position $\boldsymbol{u}$ is the only unknown parameter, the FIM for selected sensors can be expressed as

$$\boldsymbol{J}_u^{sel} = (\boldsymbol{\Gamma}_t \boldsymbol{\Phi}_p - \boldsymbol{\Gamma}_t \boldsymbol{\Phi}_q) \left( \boldsymbol{\Phi}_p^T \boldsymbol{\Gamma}_q \boldsymbol{\Phi}_p + \boldsymbol{\Phi}_q^T \boldsymbol{\Gamma}_q \boldsymbol{\Phi}_q \right)^{-1} (\boldsymbol{\Gamma}_t \boldsymbol{\Phi}_p - \boldsymbol{\Gamma}_t \boldsymbol{\Phi}_q)^T \tag{9}$$

where

$$\boldsymbol{\Gamma}_t = \frac{1}{c} \left[ \tfrac{(\boldsymbol{u}-\boldsymbol{s}_1)}{r_1}, \ \tfrac{(\boldsymbol{u}-\boldsymbol{s}_2)}{r_2}, \ \ldots, \ \tfrac{(\boldsymbol{u}-\boldsymbol{s}_M)}{r_M} \right], \tag{10}$$

$$\boldsymbol{\Gamma}_q = diag\{\Delta t_1^2, \Delta t_2^2, \ldots, \Delta t_M^2\}. \tag{11}$$

**Theorem 1.** *The closed form of the FIM for selected sensors $\boldsymbol{J}_u^{sel}$ with respect to $\boldsymbol{p}$ and $\boldsymbol{q}$ is*

$$\boldsymbol{J}_u^{sel} = \boldsymbol{\Gamma}_t \left( \boldsymbol{\Gamma}_o^{-1} - \boldsymbol{\Gamma}_o^{-1} \left( \boldsymbol{\Gamma}_o^{-1} + \boldsymbol{C} \right)^{-1} \boldsymbol{\Gamma}_o^{-1} \right) \boldsymbol{\Gamma}_t^T, \tag{12}$$

*where*

$$\boldsymbol{C} = \alpha^{-1} \left( diag\{\boldsymbol{p}\} + diag\{\boldsymbol{q}\} - \frac{(\boldsymbol{p}+\boldsymbol{q})(\boldsymbol{p}+\boldsymbol{q})^T}{K} \right)$$

*is the dependence between these two Boolean vectors and FIM, positive definite matrix $\boldsymbol{\Gamma}_o$ and scalar $\alpha$ are related to the TOA noise covariance matrix $\boldsymbol{\Gamma}_q$.*

*Proof.* To present the closed form of the FIM with respect to $\boldsymbol{p}$ and $\boldsymbol{q}$, the TOA noise covariance matrix $\boldsymbol{\Gamma}_q$ is decomposed and the TDOA noise covariance matrix for selected sensors is simplified after some algebraic manipulations.

The TOA noise covariance matrix $\boldsymbol{\Gamma}_q$ can be decomposed as

$$\boldsymbol{\Gamma}_q = \boldsymbol{\Gamma}_o + \alpha \boldsymbol{I} \tag{13}$$

where $\alpha$ is a positive scalar, $\boldsymbol{I}$ is the identity matrix, and $\boldsymbol{\Gamma}_o$ is a positive definite matrix [14]. This decomposition can be obtained by setting $\alpha$ to be half of the any eigenvalue, and it can help simplify the inverse operation in (9).

The TDOA noise covariance matrix for selected sensors is expressed as

$$\boldsymbol{\Gamma}_q^{sel} = \boldsymbol{\Phi}_p^T \boldsymbol{\Gamma}_q \boldsymbol{\Phi}_p + \boldsymbol{\Phi}_q^T \boldsymbol{\Gamma}_q \boldsymbol{\Phi}_q.$$

According to the properties of $\boldsymbol{\Phi}_p$ and $\boldsymbol{\Phi}_q$, it derives

$$\begin{aligned} \boldsymbol{\Phi}_p^T \boldsymbol{\Gamma}_q \boldsymbol{\Phi}_p + \boldsymbol{\Phi}_q^T \boldsymbol{\Gamma}_q \boldsymbol{\Phi}_q &= (\boldsymbol{\Phi}_p + \boldsymbol{\Phi}_q)^T \boldsymbol{\Gamma}_q (\boldsymbol{\Phi}_p + \boldsymbol{\Phi}_q) \\ &= (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T \boldsymbol{\Gamma}_q (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q) \\ &= \boldsymbol{\Lambda} + (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T \boldsymbol{\Gamma}_o (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q) \end{aligned} \tag{14}$$

where $\boldsymbol{\Lambda} = \alpha (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T \boldsymbol{I} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)$.

Substituting (14) into $(\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q) \left( \boldsymbol{\Gamma}_q^{sel} \right)^{-1} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T$ and employing the matrix inversion lemma [14], it derives

$$\begin{aligned} &(\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q) \left( \boldsymbol{\Gamma}_q^{sel} \right)^{-1} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T = \boldsymbol{\Gamma}_o^{-1} - \\ &\boldsymbol{\Gamma}_o^{-1} \left( \boldsymbol{\Gamma}_o^{-1} + (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)(\alpha (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T \boldsymbol{I} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q))^{-1} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T \right)^{-1} \boldsymbol{\Gamma}_o^{-1} \end{aligned} \tag{15}$$

where $\boldsymbol{\Gamma}_o$ is determined by the matrix decomposition in (13) and has no effect on the analysis.

Hence, the major challenge is to simplify the element in the inverse operation

$$\begin{aligned} \boldsymbol{C} &= (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q) \left( \alpha (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T \boldsymbol{I} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q) \right)^{-1} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T \\ &= \alpha^{-1} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q) \left( \boldsymbol{I}_{K-1} + \mathbf{1}_{(K-1)\times 1} \mathbf{1}_{(K-1)\times 1}^T \right)^{-1} (\boldsymbol{\Phi}_p - \boldsymbol{\Phi}_q)^T. \end{aligned} \tag{16}$$

Considering the relationship

$$\left( \boldsymbol{I}_{K-1} + \mathbf{1}_{(K-1)\times 1} \mathbf{1}_{(K-1)\times 1}^T \right)^{-1} = \left( \boldsymbol{I}_{K-1} - \frac{1}{K} \mathbf{1}_{(K-1)\times 1} \mathbf{1}_{(K-1)\times 1}^T \right),$$

we have

$$\boldsymbol{C} = \alpha^{-1} \left( diag\{\boldsymbol{p}\} + diag\{\boldsymbol{q}\} - \frac{(\boldsymbol{p}+\boldsymbol{q})(\boldsymbol{p}+\boldsymbol{q})^T}{K} \right).$$

Substituting (15) into (9), it derives

$$\boldsymbol{J}_u^{sel} = \boldsymbol{\Gamma}_t \left( \boldsymbol{\Gamma}_o^{-1} - \boldsymbol{\Gamma}_o^{-1} \left( \boldsymbol{\Gamma}_o^{-1} + \boldsymbol{C} \right)^{-1} \boldsymbol{\Gamma}_o^{-1} \right) \boldsymbol{\Gamma}_t^T.$$

□

Hence, the optimization problem can be formulated as

$$\begin{aligned} &\underset{\boldsymbol{p},\boldsymbol{q}}{\text{minimize}} \ \text{trace}\ (\boldsymbol{J}_u^{sel})^{-1} \\ &\text{subject to} \quad \mathbf{1}^T(\boldsymbol{p}+\boldsymbol{q}) = K \\ &\qquad\qquad\quad \boldsymbol{p}^T\boldsymbol{q} = 0 \\ &\qquad\qquad\quad \boldsymbol{p},\boldsymbol{q} \in \{0,1\}^M. \end{aligned} \tag{P1}$$

## 4 Sensor Selection Approach

Since the presence of Boolean constraint and the inner product of vectors, the optimization problem (P1) is nonconvex and is not easy to address [10]. We present the sensor selection method based on convex relaxation techniques and the major operation is relaxing these nonconvex constraints.

Substituting (12) into (P1), it derives

$$\begin{aligned} &\min_{\boldsymbol{p},\boldsymbol{q}} \ \text{tr}\ \left(\boldsymbol{B} - \boldsymbol{D}\left(\boldsymbol{\Gamma}_o^{-1} + \boldsymbol{C}\right)^{-1}\boldsymbol{D}^T\right)^{-1} \\ &s.t. \quad \mathbf{1}^T(\boldsymbol{p}+\boldsymbol{q}) = K \\ &\qquad\quad \boldsymbol{p}^T\boldsymbol{q} = 0 \\ &\qquad\quad \boldsymbol{p},\boldsymbol{q} \in \{0,1\}^M \end{aligned} \tag{17}$$

where $\boldsymbol{B} = \boldsymbol{\Gamma}_t\boldsymbol{\Gamma}_o^{-1}\boldsymbol{\Gamma}_t^T$ and $\boldsymbol{D} = \boldsymbol{\Gamma}_t\boldsymbol{\Gamma}_o^{-1}$.

Two auxiliary matrices $\boldsymbol{Z}$ and $\boldsymbol{V}$ are introduced to successively satisfy

$$\boldsymbol{B} - \boldsymbol{D}\left(\boldsymbol{\Gamma}_o^{-1} + \boldsymbol{C}\right)^{-1}\boldsymbol{D}^T \succeq \boldsymbol{Z}^{-1}, \tag{18}$$

$$\boldsymbol{V} - \boldsymbol{D}\left(\boldsymbol{\Gamma}_o^{-1} + \boldsymbol{C}\right)^{-1}\boldsymbol{D}^T \succeq \mathbf{0}, \tag{19}$$

where notation '$\succeq$' stands for an operation that $\boldsymbol{A} \succeq \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite. Besides, defining a vector $\boldsymbol{\iota} = \boldsymbol{p}+\boldsymbol{q}$ and a matrix $\boldsymbol{R}_\iota = \boldsymbol{\iota}\boldsymbol{\iota}^T$, $\boldsymbol{C}$ can be rewritten as

$$\boldsymbol{C} = \alpha^{-1}\left(diag\{\boldsymbol{p}\} + diag\{\boldsymbol{q}\} - \frac{1}{K}\boldsymbol{R}_\iota\right). \tag{20}$$

We can use the Schur complement to transform (18)–(20) into the following linear matrix inequalities (LMIs)

$$\begin{bmatrix} \boldsymbol{B}-\boldsymbol{V} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{Z} \end{bmatrix} \succeq \mathbf{0}, \quad \begin{bmatrix} \boldsymbol{V} & \boldsymbol{D}^T \\ \boldsymbol{D} & \boldsymbol{\Gamma}_o^{-1}+\boldsymbol{C} \end{bmatrix} \succeq \mathbf{0}, \quad \begin{bmatrix} \boldsymbol{R}_\iota & \boldsymbol{\iota} \\ \boldsymbol{\iota}^T & 1 \end{bmatrix} \succeq \mathbf{0}. \tag{21}$$

As for (17), $\boldsymbol{p} \in [0,1]^M$ and $\boldsymbol{q} \in [0,1]^M$ are used to relax the last constraint and the optimal solution may be fractional. The equation $\boldsymbol{p}^T\boldsymbol{q} = 0$ means that the values of corresponding elements in $\boldsymbol{p}$ and $\boldsymbol{q}$ cannot be '1' at the same time. This constraint is abandoned in the algorithm and is considered when the fractional vector is judged to be a Boolean vector.

Substituting (21) into (17) and considering the above relaxation operations, the problem can be transformed to a SDP

$$\begin{aligned} &\min_{\boldsymbol{p},\boldsymbol{q},\boldsymbol{Z},\boldsymbol{V},\boldsymbol{R}_\iota} \ \mathrm{tr}\,(\boldsymbol{Z}) \\ &s.t. \quad \text{LMIs in (21)} \\ &\quad \mathbf{1}^T(\boldsymbol{p}+\boldsymbol{q}) = K \\ &\quad 0 \le p_i \le 1,\ i = 1,2,\ldots,M \\ &\quad 0 \le q_i \le 1,\ i = 1,2,\ldots,M. \end{aligned} \tag{22}$$

The SDP problem can be solved efficiently by using interior-point approach [9,18]. However, since the relaxation operation, the optimal selection vectors for (22) are fractional and must be transformed to the Boolean vectors by judgement to select the suboptimal sensor subset from all the sensors. The simplest method is to select the maximal $(K-1)$ weight values from $\boldsymbol{p}$ and one maximal value from $\boldsymbol{q}$ [19]. Besides, we can also use the randomization method [14,20], generating a set of random fractional vectors related to the SDP solution and giving a set of random Boolean selection vectors, to select the minimum value of objective function when each random Boolean selection vector pair is substituted.

## 5 Simulation Results

This section presents the simulation results of the proposed sensor selection method and compares its performance with the exhaustive search and closest selection method. The performance metric is CRLB when the true position of source is substituting.

In the simulation, the sensor network contains 30 sensors in which 8 of them are selected. Their positions are randomly obtained and the coordinate components are Gaussian parameters when standard deviation is 3000. The position of source is also generated randomly in a large cube at $[0,0,0]^T$ with a length of 1000. TOA noise covariance matrix $\boldsymbol{\Gamma}_q$ is a diagonal matrix with diagonal elements as $\boldsymbol{\sigma}^2$, where $\boldsymbol{\sigma}$ is a random vector. In the SDP with randomization method, we set 50 random sensor combinations to calculate. Figure 1 shows the localization accuracy comparison when the range of arrival (ROA) noise strength varies from $-5$ to 10 dB. The circular symbol denotes the SDP method, the cross symbol represents the SDP with randomization method, and the inverted triangle represents the exhaustive search method. It is evident from the figure that the performance of the proposed SDP with randomization method is able to reach the exhaustive search method and is close to the scene when all sensors participate in localization. In particular, when five sensors of network are selected and ROA noise strength is fixed at 0.1 m, the selected reference sensor and ordinary sensors indices comparison of different methods are shown in Fig. 2. The pink and blue lines denote the estimation fractional vector of SDP and SDP with randomization method, and the solid symbols represent selected sensors.

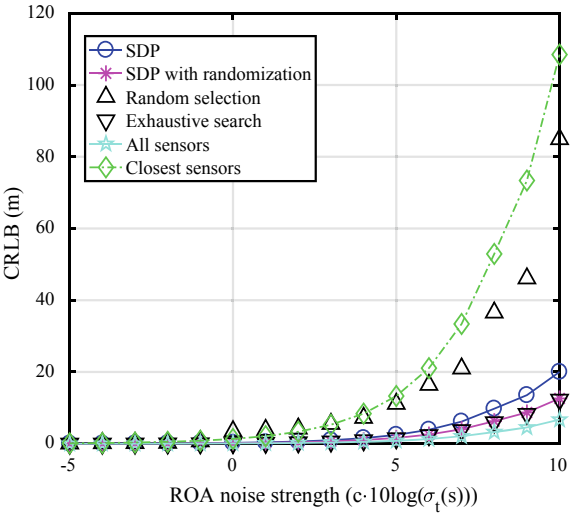



**Fig. 1.** CRLB versus ROA noise strength

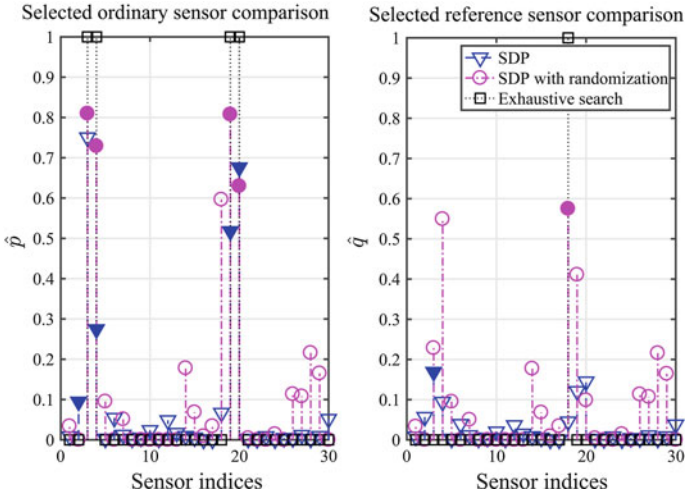



**Fig. 2.** Selected sensor indices comparison

Besides, for ease of comparison, we set the selection vector of exhaustive search method to Boolean. It is seen that the selected sensors of proposed SDP with randomization method are similar to the exhaustive search method including the reference sensor and other ordinary sensors.

## 6 Conclusion

This paper has proposed a sensor selection method to select the reference sensor and ordinary sensors simultaneously in TDOA localization scenario when the measurement noises are correlated. Two Boolean vectors are introduced to formulate the nonconvex optimization problem, which motivates to minimize the localization accuracy in the constraint of energy constraint. The problem has been solved by convex relaxation technique and randomization method. Simulation results have validated that the localization performance and selected sensor indices of proposed SDP with randomization method are similar to the exhaustive search method.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China under Grant 61631015, 61401323, 61471395 and 61501356, by the Key Scientific and Technological Innovation Team Plan (2016KCT-01), the Fundamental Research Funds of the Ministry of Education (7215433803 and XJS16063).

## References

1. Yin J, Wan Q, Yang S, Ho KC. A simple and accurate TDOA-AOA localization method using two stations. IEEE Signal Process Lett. 2016;23:144–8.
2. Ho KC, Xu W. An accurate algebraic solution for moving source location using TDOA and FDOA measurements. IEEE Trans Signal Process. 2004;52:2453–63.
3. Li Z, Guan L, Li C, Radwan A. A secure intelligent spectrum control strategy for future THz mobile heterogeneous networks. IEEE Commun Mag. 2018.
4. Wang Y, Ho KC. An asymptotically efficient estimator in closed-form for 3-D AOA localization using a sensor network. IEEE Trans Wireless Commun. 2015;14:6524–35.
5. Ho KC, Lu X, Kovavisaruch L. Source localization Using TDOA and FDOA measurements in the presence of receiver location errors: analysis and solution. IEEE Trans Signal Process. 2007;55:684–96.
6. Calwo-Fullana M, Matamoros J, Anton-Haro C. Sensor selection and power allocation strategies for energy harvesting wireless sensor networks. IEEE J Sel Areas Commun. 2016;34:3685–95.
7. Kaplan LM. Global node selection for localization in a distributed sensor network. IEEE Trans Aerosp Electron Syst. 2006;42:113–35.
8. Hao B, Zhao Y, Li Z, Wan P. A sensor selection method for TDOA and AOA localization in the presence of sensor errors. In: Proceedings of IEEE/CIC International Conference on Communications in China, Qingdao; 2017. p. 1–6
9. Boyd S, Vandenberghe L. Convex optimization. New York, NY, USA: Cambridge Univ. Press; 2004.
10. Chepuri SP, Leus G. Sparsity-promoting sensor selection for non-linear measurement models. IEEE Trans Signal Process. 2015;63:684–98.
11. Rao S, Chepuri S P, Leus G. Greedy sensor selection for non-linear models. In: Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing. 2015. p. 241–4
12. Shamaiah M, Banerjee S, Vikalo H. Greedy sensor selection: Leveraging submodularity. In: Proceedings of IEEE Conference on Decision and Control. 2010. 2572–7

13. Rene J E, Ortiz D, Venegas P, Vidal J. Selection of the reference anchor node by using SNR in TDOA-based positioning. In: Proceedings of IEEE Ecuador Technical Chapters Meeting. 2016. p. 1-4
14. Liu S, et al. Sensor selection for estimation with correlated measurement noise. IEEE Trans Signal Process. 2016;64:3509–22.
15. Hao B, Li Z, Ren Y, Yin W. On the cramer-rao bound of multiple sources localization using RDOAs and GROAs in the presence of sensor location uncertainties. In: Proceedings of IEEE Wireless Communications and Networking Conference. 2012. 3117–22
16. Qi Y, Kobayashi H, Suda H. Analysis of wireless geolocation in a non-line-of-sight environment. IEEE Trans Wireless Commun. 2006;5:672–81.
17. Kay SM. Fundamentals of statistical signal processing, estimation theory. Englewood Cliffs, NJ: Prentice-Hall; 1993.
18. Shen X, Liu S, Varshney PK. Sensor selection for nonlinear systems in large sensor networks. IEEE Trans Aerosp Electron Syst. 2014;50:2664–78.
19. Joshi S, Boyd S. Sensor selection via convex optimization. IEEE Trans Signal Process. 2009;57:451–62.
20. Luo Z, et al. Semidefinite relaxation of quadratic optimization problems. IEEE Signal Process Mag. 2010;27:20–34.

# Location Precision Analysis of Constellation Drift Influence in TDOA Location System

Liu Shuai[1], Song Yang[1], Guo Pei[2], Meng Jing[1], and Wu Mingxuan[1(✉)]

[1] China Academy of Space Technology, NO.104, You Yi Road, Haidian District, Beijing, China
lsshr@163.com, wumingxuan@sina.com

[2] Institute of Manned Space System Engineering, NO.104, You Yi Road, Haidian District, Beijing, China
stillif@sina.com

**Abstract.** In the TDOA location system,the constellation drift can affect the location precision. Aiming at this problem, based on the analysis of the three satellites constellation drifting in a period of time, we proposed the location precision results of constellation drift influence in TDOA location system. Considering the constraint of location accuracy and satellite resources, a feasible recommendation for location maintenance of three satellites system is put forward, which provides the theoretical support for the location maintenance strategy of three satellites TDOA location.

**Keywords:** Three satellites constellation · Constellation drift influence · Location precision

## 1 Introduction

Three satellites TDOA(time difference of arrival) location system is a kind of passive electronic reconnaissance location mode, which uses three formation flying satellites to determine two equal time difference surface by measuring the time arrival difference of radar emitter signal. Combined with the surface constraint equations, real-time and high precise location on electronic signal is achieved [1].

However, when the satellite is in orbit, its position will drift, which can affect the configuration of the satellite constellation than can affect the location precision of TDOA location system [2, 3]. In the paper [4], based on the best configuration mathematical expressions of the minimum positioning error variance, Li Wenhua demonstrated the effect of three satellites location precision with design of the constellation configuration, but did not explain the effect of the constellation changes on the location precision. In the paper [5], based on the TDOA location error model under the constraint of the earth, Gu Liming et al. demonstrated the effect of location performance by satellite constellation configuration, but did not make the relationship between constellation configuration and time changes.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_138

To solve the above problems, based on the three satellites TDOA system, we analyze the location precision results of constellation drift influence in a period of time, put forward the recommendation of location maintenance and improve the location precision.

## 2 Location Mechanism of TDOA System

By measuring the arrival time of the electronic signal pulse, the TDOA location system realizes the high precision location of the electronic signal radiation source. The location mechanism is shown in Fig. 1, *S*0, *S*1 and *S*2 are the obtained two independent pulse arrival times, which can determine the location points of a hyperboloid. Three satellites can make two hyperboloids as *A*1 and *A*2, which can form two curves *L*1 and *L*2 by intersection with the earth's surface respectively, and the intersection point of two curves is the location of radiation source.

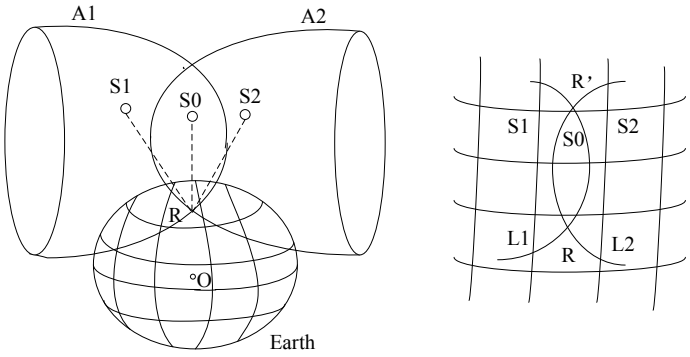



**Fig. 1.** Mechanism of three satellites TDOA location

Generally, there is only one location point *R* of the intersection between *A*1 and *A*2 and the earth's surface. When the distribution of three satellites is unreasonable, there will be two location points *R* and *R′*. The true radiation source is *R* and the false radiation source location is *R′*. With the motion of the satellites, *R′* is divergent, *R* is stable convergent [6, 7].

## 3 Analysis of the Constellation Drift

### 3.1 Analysis of Constellation Configuration Drift

Select a time as the starting point, the satellite constellation position of the current time is the initial position [8] and then the latitude and longitude location of the three satellites constellation is simulated at an interval of one month. In Fig. 2, the far right is

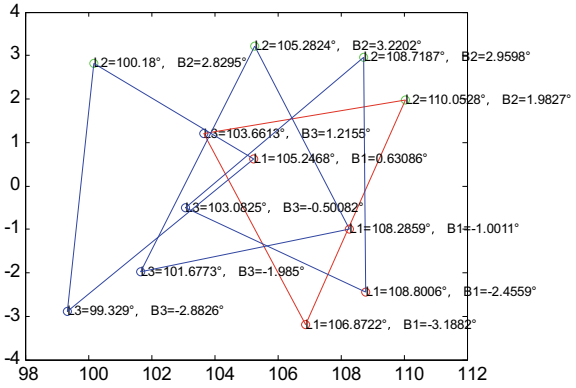



**Fig. 2.** Three months constellation change

the three satellites position at the starting point, and the left constellation lines are three satellites position after one month, two months and three months.

Table 1 is the latitude and longitude positions of the three satellites with three months change.

**Table 1.** Longitude and latitude drift with an interval of three months

| | Initial position (°) | Three months later (°) | Drifting degree (°) |
|---|---|---|---|
| Satellite longitude (A) | 106.8722 | 105.2468 | −1.62533 |
| Satellite latitude (A) | −3.18817 | 0.630862 | 3.819034 |
| Satellite longitude (B) | 110.0528 | 100.18 | −9.87278 |
| Satellite latitude (B) | 1.98275 | 2.829539 | 0.84679 |
| Satellite longitude (C) | 103.6613 | 99.32903 | −4.33232 |
| Satellite latitude (C) | 1.215527 | −2.88258 | −4.09811 |

As we can see, the maximum constellation drifting degree of the east and west direction (longitude) is −9.87°, and the maximum constellation drifting degree of the north and south direction (latitude) is −4.10°.

### 3.2 Analysis of Constellation Angle

The changed angle of constellation lines also illustrated the changes of constellation. In the initial design, the three satellites are an equilateral triangle and the angles are 60°. In the observation of a month, the baseline angle changes within the range of 5°, which will infect the location precision. The angles of the constellation are changed as shown in Fig. 3.

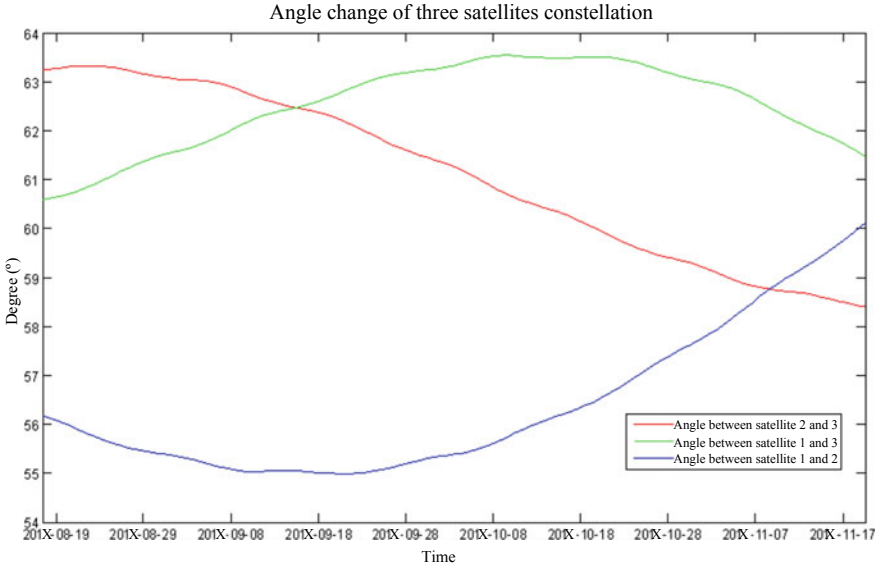



**Fig. 3.** Baseline angles change of satellite constellation

### 3.3 Analysis of the Coverage Area of Constellation

The coverage area of the three satellites constellation is also the main factor affecting the location precision. In order to compare the area changes with the initial constellation during different periods visually, we simulated the constellation changes each month, the results are shown in Fig. 4.

From Fig. 4, we can see that there is no overlap between the changed constellation and the initial constellation since October 16. As time goes on, the distance between the constellations is bigger and bigger, and the satellite coverage target area decreases. Table 2 listed the coverage area and coverage changing trends per ten days from August 17 to December 16.

As shown in Table 2, in the period of three months from August 17, the coverage area of three satellites is decreased. The coverage rate is decreased to 96.60% one month later and decreased to 57.01% three months later, which will affect the location precision of three satellites largely.

## 4 Simulation Analysis of Location Precision Influence

Choosing key parameters such as the constellation position error, time error and satellites velocity measurement error [9], then we can obtain the system location precision results, as shown in Fig. 5.

Select the specific single point in the coverage area, then we can analyze the influence of location precision of the three satellites in the drift state. Choosing the ground electronic signal emitter with the distances from the location below satellites

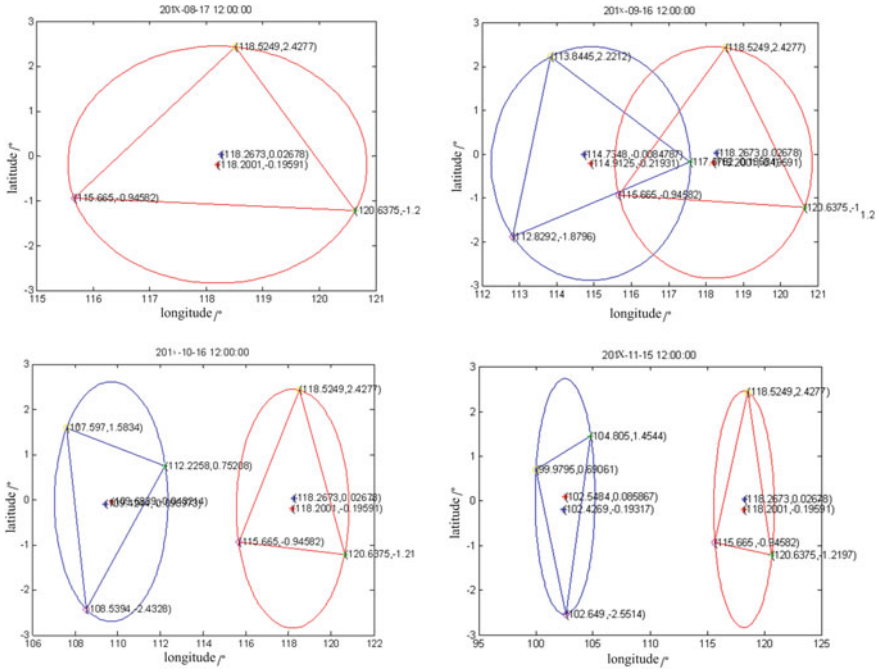



**Fig. 4.** Constellation changes each month

**Table 2.** Changing trends of coverage area

| Time (12 h) | Initial area ($km^2$) | Coverage area ($km^2$) | Coverage rate (%) |
|---|---|---|---|
| 8.17 | 19,595,000 | 19,595,000 | 100 |
| 8.27 | | 19,576,000 | 99.90 |
| 9.6 | | 19,356,000 | 98.78 |
| 9.16 | | 18,929,000 | 96.60 |
| 9.26 | | 18,250,000 | 93.14 |
| 10.6 | | 17,113,000 | 87.34 |
| 10.16 | | 15,787,000 | 80.57 |
| 10.26 | | 14,341,000 | 73.19 |
| 11.5 | | 12,804,000 | 65.34 |
| 11.15 | | 11,171,000 | 57.01 |

center is 2000, 3000, 4000 and 5000 km respectively, then we can obtain the location precision of the three satellites(GDOP)with the period of three months, as shown in Fig. 6. In Fig. 6, select the worst value (the maximum value of GDOP) of the location precision in the all equal distance target points at any time.

Simulation results show that the farther of the distance below the satellites center point, the longer of the three satellites constellation drift time, the worse of the location

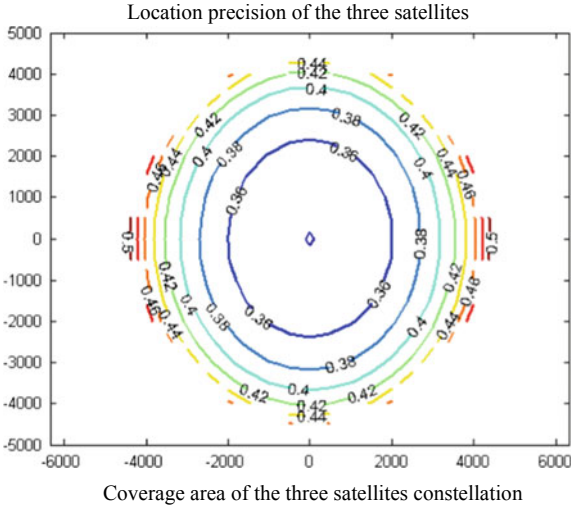



**Fig. 5.** Location precision of the three satellites

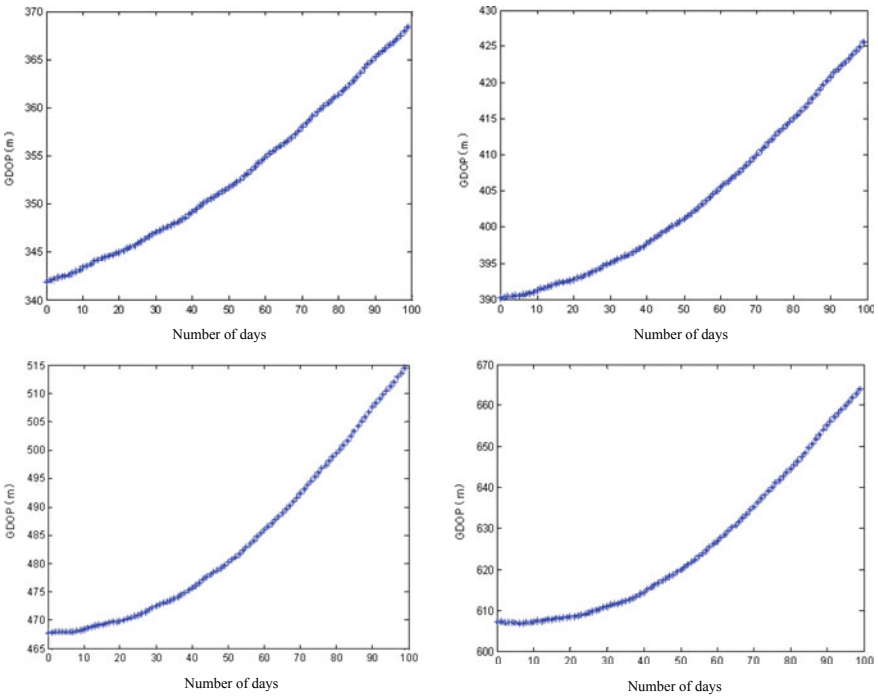



**Fig. 6.** Location precision for distances of 2000, 3000, 4000 and 5000 km

precision, as shown in Table 3. Generally, considering the location precision results and the satellite resources, we recommend that the three satellites TDOA system carried out a location maintenance for about 100 days.

**Table 3.** Summary of location precision at different distances

| Distance below the center point of satellites (km) | GDOP (m) |
|---|---|
| 2000 | 342–368 |
| 3000 | 390–426 |
| 4000 | 468–515 |
| 5000 | 608–665 |

## 5 Summary

In the three satellites TDOA system, the three satellites will drift, which affects the location precision of system. In this paper, the constellation change of the three satellites over a period of time is demonstrated, and the influence of constellation drift on location precision is also analyzed. Considering the constraints of location precision and satellite resources, we proposed feasible location maintenance recommendations for the three satellites, which provided a theoretical support for the location protection strategy of three satellites TDOA constellation.

## References

1. Zhu W, Huang P, Ma Q,Lu X. Emitter location with multi-station using TDOA/FDoA measurements. J Data Acquis Process. 2010; 05:307–12.
2. Yu M, Xu J, Xia Y, Lu BK. A Study of optimal configuration of three-satellite constellation for localizing the object on earth. ACTA Astron Sin. 2003;08:302–9.
3. Chen Y, Li C, Li X. A precision analyzing&reckoning model in tri-station TDOA location. ACTA Electron Sin..2004;32(9):1452–5.
4. Li W. Research on configuration of tri-satellites and location precision of TDOA. J Astronaut. 2010;03:701–6.
5. Gu L, Zhao Y, Zhu J, Miao S. Constellation configuration and its performance based on TDOA with four satellites. Telecommun Eng. 2017; 01:33–8.
6. Fu W. DTOA Location precision of target on the ground.Electron Inf Warf Technol. 2008;23 (6):17–20.
7. Rao H. Research on key technologies of TDOA with multi-satellites. Nan jing: Nanjing University of Science and Technology; 2009.
8. Du L, Zhang Z, Li X, Wang R, Liu L, Guo R. Station-keeping maneuver monitoring and moving-window ground track fitting of GEO satellites. Act Geod Cartogr Sin. 2014;03: 233–9.

9. Ping Z, Zhen H, Jianhua L. Passive stationary target positioning using adaptive particle filter with TDOA and FDOA Measurements, vol. 4. Issue 12. 9th Asia-peck Conference about Communications. 2010. p. 224–235.

# A Weighted and Improved Indoor Positioning Algorithm Based on Wi-Fi Signal Intensity

Guanghua Zhang and Xue Sun(✉)

Northeast Petroleum University Youth Fund: XN2014111, Daqing, China
1054351628@qq.com

**Abstract.** In order to solve the problem of the influence of signal strength fluctuation on indoor positioning accuracy, an improved indoor positioning algorithm based on WiFi signal strength is proposed in the paper. Based on the *K*-nearest neighbor location algorithm, the weight of the signal strength is further increased, the characteristics of the received signal and the fingerprint database are optimized, the interference of the weak signal on the positioning is reduced, and the accurate indoor positioning is achieved. The calculation in this work suggests that the positioning error can be reduced on the original basis, the accuracy of the algorithm is improved. On the basis of the original algorithm, the error range is narrowed and the positioning accuracy is improved. The average error of the improved algorithm is controlled at about 1.87 m.

**Keywords:** Wi-Fi · Indoor positioning · Weighted signal strength

## 1 Introduction

Indoor positioning technology refers to determining the position of a receiving terminal in a certain reference coordinate system at a certain time in an indoor environment. The "indoor" here not only refers to the general interior of a typical building, but also includes underground mines and dense high-rise buildings. Some fully enclosed or half enclosed space are also called indoors, such as buildings or woods [1]. These enclosed or semi-enclosed spaces are collectively referred to as interior spaces. With the development of science and technology, there are more and more indoor positioning technologies, such as wireless network technology, Bluetooth, infrared, ZigBee technology, and RFID. The principle of different positioning methods is also different; there are positioning methods based on signal arrival time or angle of arrival, or signal strength, long time, etc. For different indoor environments, the most suitable positioning method should be used to reduce the positioning error as much as possible. At present, several commonly used positioning software, such as Baidu Maps and Gaode maps, have a good positioning effect in the outdoors and also bring great convenience to people's lives. However, the positioning in the room is severely degraded due to the blockage of the GPS signal by the obstacles and cannot be applied indoors. The massive use of mobile phones has led to an increasing demand for indoor positioning. In addition, the coverage of Wi-Fi has been continuously expanding, which has led to


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_139

the emergence of indoor Wi-Fi positioning technology. This article mainly introduces the commonly used indoor positioning algorithms. Based on this, it proposes a method to increase the weights and improves the positioning accuracy through database filtering and error analysis.

## 2 Traditional Indoor Positioning Algorithm

### 2.1 KNN Algorithm

KNN algorithm is the *K*-nearest neighbor classification algorithm. The core principle of the algorithm is that if the most neighboring samples of a sample in the feature space mostly belong to a certain category, then the sample is also said to belong to this category, and has the basic characteristics of the sample of this category [2, 3]. In the KNN algorithm, the most important affair is to correctly select the value of K. When the value of *K* is small, the number of nearby points to be selected is small, and the category of the reference point cannot be accurately determined, thereby increasing the error. When the *K* value is large, the farther point will be selected, resulting in inaccurate results. Therefore, multiple measurements are required determining the value of parameter *K* during the experiment. Then selecting a certain number of collection points in the measurement area, and collecting the position information and signal intensity values of each point. When the user issues a positioning request, the coordinate information of the point is fed back to the user through the RSS of the signal.

### 2.2 Triangle Positioning Algorithm

Triangle principle of distance measurement: First we must determine the coordinate information of at least 3 Wi-Fi stores, by calculating the signal strength received by the mobile receiver, according to the formula to calculate the distance from the measuring point to each AP $P_r(d)$ =

$$\frac{P_t G_t G_r \lambda^2}{4\pi d^2 l} \tag{1}$$

$P_t$ is the transmission power of the AP, $G_t$ is the gain of the signal transmitting end, $G_r$ is the gain of the mobile receiving end, $\lambda$ is the wavelength of the electric wave, $P_r(d)$ is the signal power received by the mobile receiving end, d is the distance between AP and the mobile terminal, l is the loss factor [3]. Only d in the formula is unknown. Solve the distance from each AP to the point to be measured. Examples are A, B, C APs with known coordinates. D is the point to be measured, the three transmitting ends are centered, the distance between A and D is d1, the distance between B and D is $d_2$, and the distance between C and D is $d_3$. The distance $d_3$ is a radius, and the intersection of the three circles is the point to be measured. Let D coordinate be $(x, y)$ A, B, and C coordinates $(x_1,y_1)$, $(x_2, y_2)$, $(x_3, y_3)$, and calculate the known Wi-Fi

hot spot to the point D to be determined by formula For distances $d_1$, $d_2$, and $d_3$, $(x - x_i)^2 + (y - y_i)^2 = d_i^2$, where $i = 1, 2$, and 3. The values of $x$ and $y$ can be obtained by solving a system of linear equations.

### 2.3 Disadvantages and Demerits of Traditional Location Algorithms

The *K*-nearest neighbor method determines that the user equipment is located near the sensor position when the signal of the user equipment is detected by a sensor of a fixed position. In this way, the location of the sensor can be used to characterize the spatial location of the user which is a relative location data. And the initial investment is not suitable for a wide range of applications. Triangulation, in the data acquisition process, has a high requirement for signal stability and is affected by complex indoor environments and equipment. Signals can fluctuate and distort due to occlusion of buildings or the flow of people.

## 3 Improved Weight Measurement

The main idea based on the weight measurement method is to compare the RSSI collected in real time with the feature database, and select the centroid of the nearest point as the target's estimated position. In order to complete the matching, a feature database should be established first, including: location identification, location coordinates, signal strength, and intensity change range. The location identifier is an identifier for each AP in the collected data in the indoor area. Positioning is a constant parameter used [4, 5].

### 3.1 Based on Weight Measurement

Position coordinates refer to the *X* and *Y* coordinates of a signal strength set, which is used to calculate and determine the target position. The signal strength refers to a list of signal strengths ($AP_n$, $RSSI_n$) after a part of the location-specific reference AP has averaged the signal strengths several times, which is the number of selected reference points in the target area.

The signal strength and intensity variation range in the feature database are the features used to determine the position. The current position is determined by matching the feature library with the signal strength acquired in real time at the current position. Current Wi-Fi signal strength ($AP_1$, $RSSI_1$) ($AP_2$, $RSSI_2$) … ($AP_n$, $RSSI_n$), n is the number of currently scanned APs. The first name from the list starts with matching in the signature database, finds the AP with the same name, and then judges its signal strength. When the current signal strength is within the variation range of the signal strength of the gather, the weight of the position is increased by 1, so repeated will get a certain number of weight lists, in different positions, the maximum weight $q_{\max}$ is selected as a reference, and the weight and the position of the phase difference within e are used as reference points, where the value of e varies with the environment,

and $q_{\max} - q_i < e$. $q_i$ is the first $i$ weights. The weights of the reference points are arranged from the largest to the smallest and are recorded as $Q_1 = (q_1, x_1, y_1)$ to $Q_n = (q_n, x_n, y_n)$ according to the formula as below.

$$x = \frac{\sum_{i=1}^{n} q_i x_i}{\sum_{i=2}^{n} q_i} \tag{2}$$

$$y = \frac{\sum_{i=1}^{n} q_i y_i}{\sum_{i=2}^{n} q_i} \tag{3}$$

### 3.2 The Main Influencing Factors of Indoor Positioning

In the indoor positioning process, due to the different environments, the influence on the positioning accuracy is also different. The main influencing factors on the positioning accuracy are: the impact of obstacles on the strength of Wi-Fi signals, different obstacles such as walls, furniture, and people. These obstacles reflect or absorb the Wi-Fi signal. If these influencing factors are brought into the calculation, it will increase the positioning error. In order to reduce the error, an improvement factor is introduced [6]. We denote the correction of the influence of obstacles on the signal as a. The transmission distance will also affect the signal strength. As the distance increases, the signal strength will continue to attenuate. When the distance exceeds a certain distance, the signal will reach the receiving end and the intensity will be weak or even absent. In this paper, the distance will affect the signal. Correction is marked b. There are also multipath effects in the propagation of Wi-Fi signals. Because the signal is transmitted to the receiving process, it will be blocked by different interference and obstructions, which will cause multiple paths during the propagation process. In this way, the arrival time of the signal will be different, and the positioning longitude will be attenuated. In this paper, the correction of the effect of multipath effects on the signal is denoted as c.

### 3.3 The Realization of Improved Algorithm

In this article, improvements are made based on signal strength acquisition and impact factors. Through a large number of data acquisitions, it is found that the intensity advantage of Wi-Fi signals collected at two locations with a certain distance will be greatly different. If an AP's signal is lost several times within the same time, it cannot be scanned. This is not the case at another point that is a certain distance away from this point. Through comparison, when the signal strength at a certain point is lower than a certain value, the AP signal has a large fluctuation at this point. When the characteristic database is established, stocks can be excluded from this type of point and can be to a certain extent. Improve positioning accuracy.

After positioning based on the weights, the values of x and y can be obtained, and then a judgment is made on the result of the last accurate positioning based on the coordinates. Then, the difference between the two sets of coordinates is calculated, and if the difference is within the required range of the positioning accuracy, the positioning

is considered to be true. Otherwise, relocate. This method is used in the positioning stage. After the results are obtained by the positioning algorithm, firstly, the position coordinates obtained by the positioning algorithm are obtained, the distance $d_1$ between the coordinates and the coordinates of the last correct positioning is calculated, and the accelerometer and timing on the communication device are used. Calculate the speed and time from the first coordinate to the second coordinate by using the application of the acceleration sensor and the timer on the communication device, and then calculate the distance $d_2$ that the target moves during this time, and record the positioning accuracy of the system. For d, when $|d_1 - d_2| < d$, it means the positioning result is correct, otherwise the result is wrong.

However, the influencing factors in the signal propagation process cannot be ignored. To improve the positioning accuracy, the corresponding $x$ and $y$ values can be multiplied by the correction factor of the corresponding influencing factor. Then the formulas for calculating $x$ and $y$ are as follows: $x' = x * a * b * c, y' = y * a * b * c$, where $x'$ and $y'$ are the results of the last positioning. The value of $a$, $b$, $c$, is obtained from the review of the reference or the measurement of a large amount of data. For different positioning environments, the three values are not the same.

## 4 The Main Performance of the Positioning System

### 4.1 Data Collection and Filtering

In the initial period of database establishment, a large amount of data must be collected, and the strength of the Wi-Fi signal has great volatility. If it can be filtered based on changeable range of the intensity, the positioning accuracy will be greatly improved. In this paper, based on the RSS signal strength filtering method, the propagation of wireless signals in space is more complex, and the RSS values measured at the same point are also not the same, and its measured values are affected in many ways [7, 8], for example, the multipath of signal propagation, non-line-of-sight propagation, interference of other signals and obstacles. This result in different RSS worth cases, and because of the non-linear time-varying nature of the signal, the single-measured RSS value will also be different. Provided that only one RSS value is used to measure, the error is large, so the processing of the RSS value is necessary. In this paper, the average value filtering method is used to process the RSS value. The main idea of the mean filter is to perform multiple measurements at the same location to obtain different RSS values and then average these values so that the effect of a single intensity value can be avoided. The formula is as follows.

$$\mathrm{RSS}' = \frac{1}{N}\sum_{i=1}^{N}\mathrm{RSS}_i \tag{4}$$

$\mathrm{RSS}_i$ represents the value obtained by the $i$-th measurement, to avoid the error caused by a single value.

### 4.2 Positioning Function

The positioning function is the main function of the positioning system. The main workflow of the positioning system is as follows. First, the positioning module needs to complete the data acquisition in real time. Then, the acquired data is stored, and the obtained data and data in the database are completed through data calculation. Matching and calculating the current position coordinates. Since there is a certain error in the result of the positioning, an inaccurate judgment is made on this result. If the error is within the required positioning accuracy, the result is correct; otherwise, the result is discarded and repositioned [9, 10].

## 5 Experimental Results and Error Analysis

After the data is collected, the system diagram is simulated with MATLAB to obtain the positioning result. According to the positioning results, the positioning error is calculated, a chart depicting the distance of the error of the *x*-axis and *y*-axis is plotted to calculate the cumulative probability. From Fig. 1, the cumulative probability that the error distance of the *x*-axis is within 0.5 m is approximately 50% within 1 m. The cumulative probability is about 72%, and the cumulative probability within 2 m can reach more than 90%.

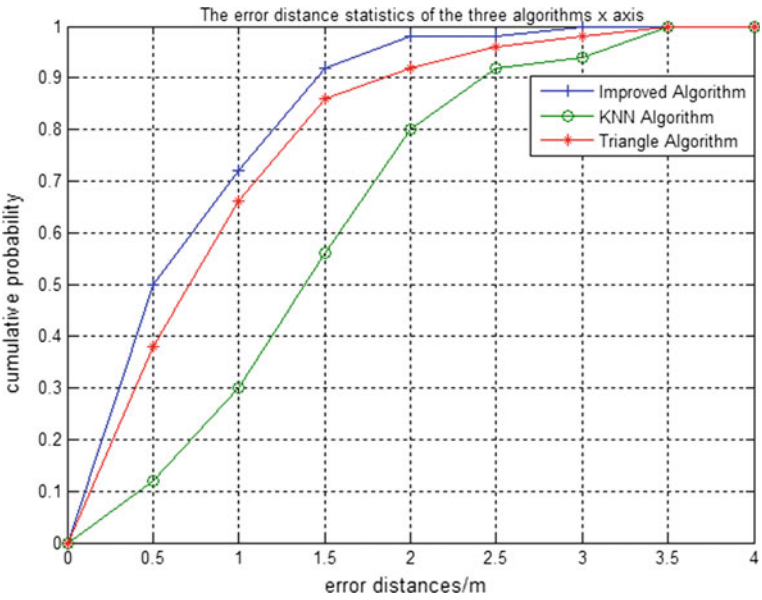



**Fig. 1.** Three algorithms *x*-axis error distance chart

From Fig. 2 we can see that the cumulative probability of *y*-axis error within 0.5 m is 38%, the cumulative probability of error within 1 m is about 70%, and the cumulative probability within 2 m is 98%.

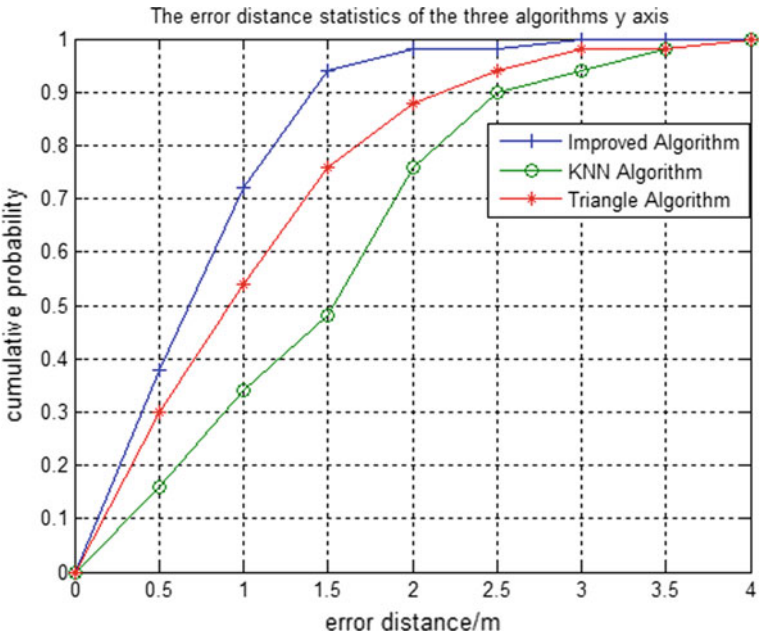



**Fig. 2.** Three algorithms *y*-axis error distance chart

According to the Fig. 3, the error histogram and average error histogram are plotted and compared from the error maximum, minimum, and average error. It can be seen that the improved positioning algorithm is smaller than the other two algorithms in numerical value. The positioning error can be shortened to within 2 m.

According to the Fig. 4, the paper uses CEP to calculate accuracy. It can be calculated that CEP = 1.87 m. From the figure, we can see that just 50% of the points fall within the circle.

## 6 Conclusion

This paper improves the positioning algorithm based on the weights and improves the positioning accuracy through database filtering, error analysis, and correction of impact factors. By comparing the advantages and disadvantages of the KNN algorithm and the triangle positioning algorithm, an improved positioning method based on weights is proposed to lay a foundation for the future development of indoor positioning technology. Wi-Fi-based positioning technology can be used in vehicle positioning systems, large shopping malls, libraries, etc. to ensure the accuracy of positioning through

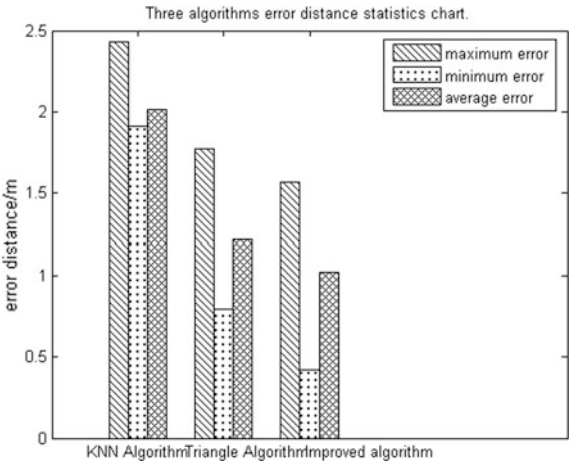



**Fig. 3.** Three kinds of algorithm error comparison chart

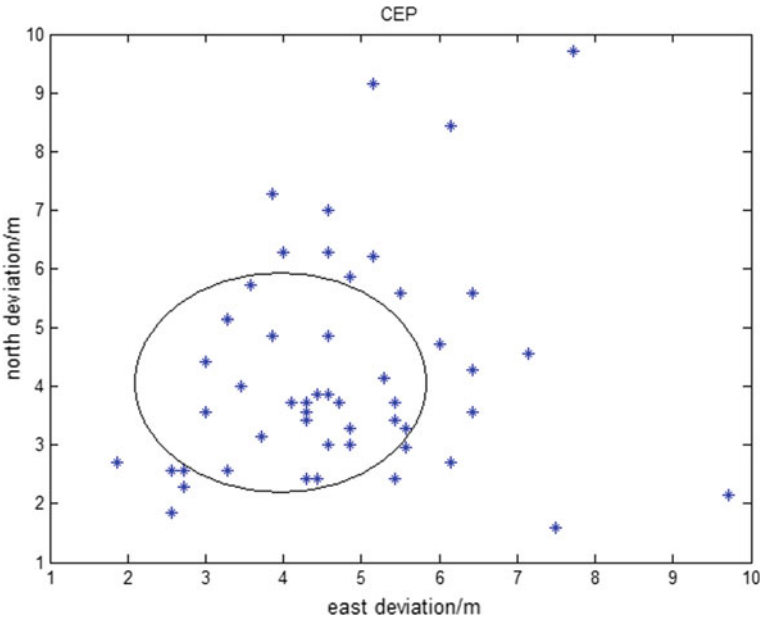



**Fig. 4.** CEP

real-time updates to the database. In the past two years, the application and research in China have become more and more widespread. Many hot spots have been installed in tourist attractions, restaurants, or subways. You can use mobile devices such as mobile phones and tablet computers to connect to the Internet, and various Wi-Fi-based applications are also used. The research on Wi-Fi positioning technology has become

more and more in-depth, and there are more and more positioning algorithms. People hope to solve the limitations of the indoor environment through the advantages of this technology and achieve indoor precise positioning. With the deepening of research, it is constantly improving. The accuracy of positioning brings greater convenience to people's lives.

## References

1. Qing W. Design and implementation of WiFi indoor positioning system. Beijing: Beijing Jiao tong University; 2014.
2. Bi-Chao Y. Research on indoor location technology based on WiFi. Chengdu: University of Electronic Science and Technology; 2017.
3. Yang P. A Weighted value selection and weighted localization algorithm based on RSSI. Inf Electron Eng. 2012;148–151.
4. Jin C, Qiu D. Research on indoor positioning technology based on WiFi signal. Bull Surv Mapp. 2017;21–25.
5. Lin H. Location-fingerprint indoor positioning algorithm based on Wi-Fi. Shanghai, Nanjing: East China Normal University; 2016.
6. Yan J. Research on indoor localization technology based on Wi-Fi. Guangzhou: South China University of Technology; 2013.
7. Rui M, Qiang G. An improved WiFi indoor positioning algorithm by Weighted Fusion. Sensors. 2015;21824–21843.
8. Hung-Huan L, Wei-Hsiang L. A WiFi-based weighted screening method for indoor positioning systems. Wireless Pers Commun. 2014;611–627.
9. Yang C, Shao H-R. WiFi-based indoor positioning. IEEE Commun Mag. 2015;150–157.
10. Huang H. WiFi Indoor positioning system design. J Guangxi Ademc Sci. 2016;59–61.

# Evaluation Distance Metrics for Pedestrian Retrieval

Zhong Zhang[1,2(✉)], Meiyan Huang[1,2], Shuang Liu[1,2], and Tariq S. Durrani[3]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
{zhong.zhang8848,meiyanhuang7295}@gmail.com
[2] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China
shuangliu.tjnu@gmail.com
[3] Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow Scotland, UK
t.durrani@strath.ac.uk

**Abstract.** Pedestrian retrieval is an important technique of searching for a specific pedestrian from a large gallery. In this paper, we introduce three types of distance metrics for pedestrian retrieval, including learning-free distance metric methods, metric learning methods, and convolution neural network (CNN) methods, and evaluate the performance of different distance metrics using the Market-1501 database. The experiment shows that the CNN methods achieve the best results.

**Keywords:** Pedestrian retrieval · Learning-free distance metric methods · Metric learning methods · CNN methods

## 1 Introduction

Pedestrian retrieval mainly refers to a task of matching persons from different cameras. It is widely applied to pedestrian tracking, image retrieval, human behavior analysis, and so on [1–5]. However, the challenges of low resolution, viewpoint change, occlusion and illumination difference make pedestrian retrieval a non-trivial problem.

In addition to extracting robust feature representation to deal with the above-mentioned challenges, distance metric plays an essential role in pedestrian retrieval. An effective distance metric could improve the matching accuracy for pedestrian retrieval. The existing distance metrics are mainly divided into three categories, i.e., learning-free distance metric methods, metric learning methods, and convolution neural network (CNN) methods. The learning-free distance metric methods exploit pre-defined distance functions, including Euclidean distance [6], Cosine distance [7], Bhattacharyya distance [8], and so


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_140

on, to calculate the similarity between two pedestrian images in the Euclidean space as shown in Fig. 1a. The metric learning methods are based on learning a transformation matrix $M$ using the positive and negative sample pairs. The learned matrix $M$ can transfer the feature representation from the Euclidean space to a new space as shown in Fig. 1b, where the positive samples are close to each other and the negative samples are relatively far away. The CNN methods employ the end-to-end CNN architecture to obtain the classification results directly as shown in Fig. 2.

In this paper, we review the Euclidean distance [6] and Cosine distance [7] to analyze the learning-free distance metric methods and enumerate several typical algorithms to describe the metric learning methods concretely. The CNN methods are introduced based on two types of CNNs that are the verification network and the identification network, respectively. We aim to evaluate the impact of different distance metrics on pedestrian retrieval.

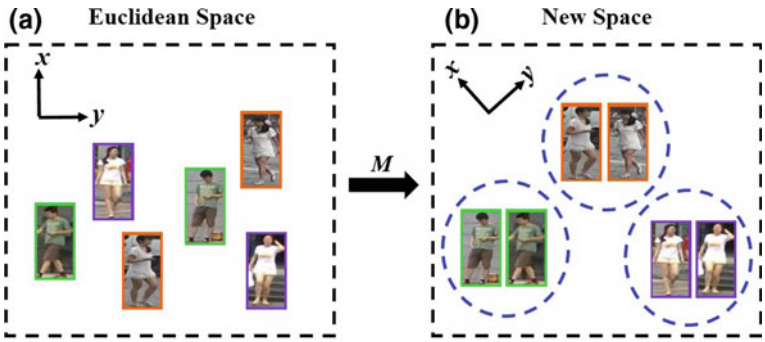



**Fig. 1.** Pedestrian images in the Euclidean space are mapped into a new space by the transformation matrix $M$.

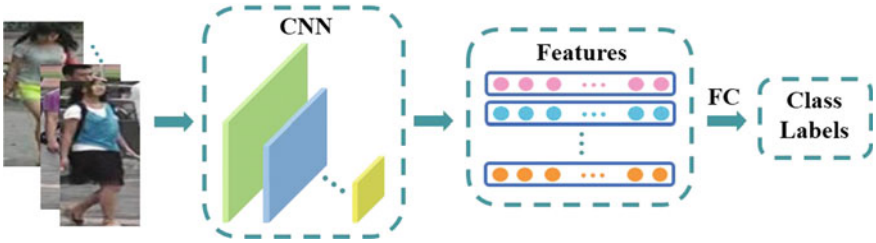



**Fig. 2.** End-to-end CNN architecture.

## 2 Distance Metrics for Pedestrian Retrieval

In this section, we introduce three types of distance metrics, i.e., the learning-free distance metric methods, metric learning methods, and CNN methods.

### 2.1 Learning-Free Distance Metric Methods

The learning-free distance metric methods utilize the ready-made distance functions, for example, Euclidean distance, Cosine distance, and Bhattacharyya distance, to compute the distance between two pedestrian images in the Euclidean space. We take the Euclidean distance and the Cosine distance as examples to describe the learning-free distance metric methods, and the smaller the distance is, the bigger the similarity score between two pedestrian images is. Given two feature vectors corresponding to two pedestrian images $P = \{p_1, p_2, \ldots, p_n\}$ and $Q = \{q_1, q_2, \ldots, q_n\}$, the Euclidean distance between two pedestrian images is expressed as

$$dist_E(P, Q) = (\sum_{i=1}^{n}(p_i - q_i)^2)^{1/2} \tag{1}$$

Different from the Euclidean distance that takes the absolute distance between two pedestrian images as the evaluation criterion of similarity score, the Cosine distance, also known as cosine similarity, exploits the cosine value to obtain the similarity score. If the two images are the same pedestrian, then the cosine value is close to 1, otherwise 0. It is formulated as:

$$dist_C(P, Q) = \frac{\sum_{i=1}^{n} p_i q_i}{(\sum_{i=1}^{n} {p_i}^2)^{1/2}(\sum_{i=1}^{n} {q_i}^2)^{1/2}} \tag{2}$$

The learning-free distance metric methods compute the similarity score between two pedestrian images in a fixed way, which neglects the properties of samples.

### 2.2 Metric Learning Methods

We first define two kinds of sample pairs. A pair of pedestrian images with the same class label is named positive sample pair, and a pair of pedestrian images with different class labels is called negative sample pair. Metric learning methods focus on learning a transformation matrix $M$ by positive and negative sample pairs. The transformation matrix $M$ maps the feature vectors of pedestrian images to a new space, which helps decrease the distance between positive sample pairs and enlarge the distance between negative sample pairs, simultaneously. The Mahalanobis distance proposed by Xing et al. [10] is one of the classical metric learning methods, and it is formulated as

$$dist_M(z_i, z_j) = [(z_i - z_j)^T M (z_i - z_j)]^{1/2} \tag{3}$$

where $z_i \in \mathbb{R}^n$ and $z_j \in \mathbb{R}^n$ represent the feature vectors of the $i$th and $j$th pedestrian images, respectively. $M = YY^T$ is the transformation matrix, and $Y \in \mathbb{R}^{n \times c}$ is the projection direction. $c$ is the dimension of feature vector in the new space.

Many metric learning methods based on the Mahalanobis distance are proposed to improve the performance of pedestrian retrieval. Weinberger et al. [11] describe a new kind of metric learning method termed large margin nearest neighbor (LMNN), and the goal of this method is that $k$-nearest neighbors usually possess the same identity, and meanwhile, pedestrian images with different identities are partitioned by a large margin. An information-theoretic approach [12] is presented to learn the Mahalanobis distance, which regards the optimal solution as searching the minimum of the LogDet divergence under linear constraints. Guillaumin et al. [13] propose the logistic discriminant metric learning (LDML), a logistic discriminant method, to force the positive sample pairs having smaller distance than that of negative sample pairs. Koestinger et al. [14] come up with the KISS metric learning method, which learns the distance function built on the statistical inference perspective. Liao et al. [15] present the metric learning method called cross-view quadratic discriminant analysis (XQDA) that maps the feature vectors of pedestrian images into a low-dimensional subspace. Specifically, they build two covariance matrices that are intra-class covariance matrix computed by positive sample pairs and extra-class covariance matrix calculated by the negative sample pairs. They then introduce the generalized Rayleigh quotient to solve the transformation matrix (Fig. 3).

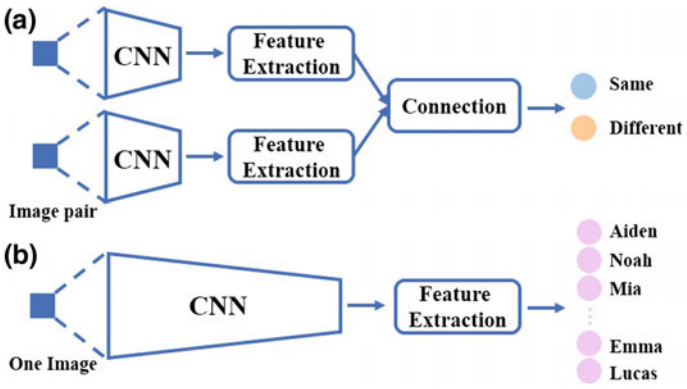



**Fig. 3.** Verification network and identification network.

### 2.3 CNN Methods

Two types of CNNs are illustrated in Fig. 2, i.e., verification network and identification network. Figure 2a shows the verification network. Its input is a pair of pedestrian images, and its output tells the two pedestrian images belonging to the same class or not. The verification network does not require explicit

image labels. Figure 2b shows the identification network, which considers one pedestrian image as input and output the class labels. Both the verification network and the identification network adopt the end-to-end CNN architecture to directly obtain the classification results. The CNN methods mainly rely on two kinds of CNNs and employ different loss functions. Here, we introduce the triplet loss and the cross-entropy loss corresponding to the verification network and the identification network, respectively.

The triplet loss [17] is an efficient CNN loss function which optimizes the embedding space so that samples with the same class are closer to each other than those with different classes. The triplet loss is defined as

$$\begin{aligned} Loss_{Tri} = & \sum_{\substack{x,b,e \\ L_x=L_b\neq L_e}} [s + dist_{x,b} - dist_{x,e}] \\ s.t.\ & [s + dist_{x,b} - dist_{x,e}] > 0 \end{aligned} \tag{4}$$

where $s$ is the value of margin, $x$ is the target image, $b$ denotes the pedestrian image having the same class label with the target image, $e$ represents the pedestrian image having the different class labels with the target image, and $L_x$, $L_b$, and $L_e$ are the class labels of three pedestrian images, respectively. Here, $dist_{x,b}$ and $dist_{x,e}$ express the distance between two pedestrian images in an embedding space, and the constraint term means that the distance between the same pedestrian is less than the distance between different pedestrians, by at least a margin $s$. The cross-entropy loss [16] is simple and easy to calculate and accelerates the speed of training network. The cross-entropy loss is represented as

$$Loss_{Cross-E} = \sum_{n=1}^{N} -x_n log\hat{x}_n \tag{5}$$

where $x_n$ is the true identity, and $\hat{x}_n$ indicates the prediction probability belonging to the $n$th pedestrian. $N$ is the number of identities.

## 3 Experimental Results

In this section, we evaluate the performances of three kinds of distance metrics on the Market-1501 database [9]. The Market-1501 database is composed of 32,668 pedestrian images of 1501 identities collected with six cameras, and it is split into the training set and test set. The training set contains 12,936 pedestrian images of 751 identities, and the test set includes 19,732 pedestrian images of 750 identities. There are still 3368 query images, and we adopt a single query setting. We find the same identity for each query in the test set. The Local Maximal Occurrence (LOMO) feature representation proposed by Liao et al. [15] is utilized for learning-free distance metric methods and metric learning methods, and the ResNet-50 network [18,19] is taken as the pre-trained CNN model for CNN methods. Table 1 lists the rank-1 accuracies of the different distance metrics.

From Table 1, the metric learning methods outperform the learning-free distance metric methods because the new space is more discriminative than the Euclidean space. Furthermore, the CNN methods achieve the best accuracy and surpass the first two kinds of methods by a large margin, especially, combining the verification network and the triplet loss. Experimental results show that the CNN methods have great potential to improve performance for pedestrian retrieval.

**Table 1.** Rank-1 accuracies (%) of different distance metrics on the Market-1501 database [9]. V represents the verification network, and I denotes the identification network.

| Methods | Rank-1 (%) |
|---|---|
| Euclidean distance | 7.83 |
| Cosine distance | 9.07 |
| LMNN | 17.88 |
| LDML | 22.43 |
| KISSME | 29.59 |
| XQDA | 45.31 |
| Cross-entropy loss + I | 73.69 |
| Triplet loss + V ($s = 0.1$) | **75.50** |

## 4 Conclusion

In this paper, three kinds of distance metrics are described which are learning-free distance metric methods, metric learning methods, and CNN methods. The learning-free distance metric methods ignore the properties of samples with the fix distance functions. The metric learning methods learn the similarity between two pedestrian images in a new space, and the CNN methods employ the end-to-end CNN architecture to obtain the classification results. Based on the results of the Market 1501 database, we conclude that the CNN methods are effective for pedestrian retrieval.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China under Grant No. 61711530240 and No. 61501327, Natural Science Foundation of Tianjin under Grant No. 17JCZDJC30600 and No. 15JCQNJC01700, the Fund of Tianjin Normal University under Grant No.135202RC1703, the Open Projects Program of National Laboratory of Pattern Recognition under Grant No. 201700001 and No. 201800002, the China Scholarship Council No. 201708120039 and No. 201708120040, the NSFC-Royal Society grant, and the Tianjin Higher Education Creative Team Funds Program.

## References


1. Zheng F, Shao L. Learning cross-view binary identities for fast person re-identification. In: International joint conference on artificial intelligence. New York: USA; 2016. p. 2399–406.
2. Chen J, Wang Y, Qin J, Liu L, Shao L. Fast person re-identification via cross-camera semantic binary transformation. In: IEEE conference on computer vision and pattern recognition. Honolulu, HI, USA; 2017. p. 5330–9.
3. Zhang Z, Wang C, Xiao B, Zhou W, Liu S, Shi C. Cross-view action recognition via A continuous virtual path. In: IEEE conference on computer vision and pattern recognition. Portland, OR, USA; 2013. p. 2690–7.
4. Zhang Z, Wang C, Xiao B, Zhou W, Liu S. Action recognition using context-constrained linear coding. IEEE Signal Proc Let. 2012;19(7):439–42.
5. Zhang Z, Wang C, Xiao B, Zhou W, Liu S. Attribute regularization based human action recognition. IEEE T Inf Foren Sec. 2013;8(10):1600–9.
6. Farenzena M, Bazzani L, Perina A, Murino V, Cristani M. Person re-identification by symmetry-driven accumulation of local features. In: IEEE conference on computer vision and pattern recognition. San Francisco, CA, USA; 2010. p. 2360–7.
7. Zhang D, Lu G. Evaluation of similarity measurement for image retrieval. In: International conference on neural networks and signal processing. Nanjing, China; 2003. p. 928–31.
8. Cheng D, Cristani M, Stoppa M, Bazzani L, Murino V. Custom pictorial structures for re-identification. In: British machine vision conference. Dundee, UK; 2011. p. 6.
9. Zheng L, Shen L, Tian L, Wang S, Wang J, Tian Q. Scalable person re-identification: a benchmark. In: IEEE international conference on computer vision. Santiago, Chile; 2015. p. 1116–24.
10. Xing E, Jordan M, Russell S, Ng A. Distance metric learning with application to clustering with side-information. In: Advances in neural information processing systems. Vancouver, British Columbia, Canada; 2003. p. 521–8.
11. Weinberger K, Blitzer J, Saul K. Distance metric Learning for large margin nearest neighbor classification. In: Advances in neural information processing systems. Vancouver, British Columbia, Canada; 2006. p. 1473–80.
12. Davis J, Kulis B, Jain P, Sra S, Dhillon I. Information-theoretic metric learning. In: International conference on machine learning. Cincinnati, Ohio, USA; 2007. p. 209–16.
13. Guillaumin M, Verbeek J, Schmid C. Is that you? Metric learning approaches for face identification. In: IEEE international conference on computer vision. Berthold K.P. Horn; 2009. p. 498–505.
14. Koestinger M, Hirzer M, Wohlhart P, Roth P, Bischof H. Large scale metric learning from equivalence constraints. In: IEEE conference on computer vision and pattern recognition. Providence, RI, USA; 2012. p. 2288–95.
15. Liao S, Hu Y, Zhu X, Li S. Person re-identification by local maximal occurrence representation and metric learning. In: IEEE Conference on computer vision and pattern recognition. Boston, Massachusetts; 2015. p. 2197–206.
16. Zheng Z, Zheng L, Yang Y. A Discriminatively learned CNN embedding for person re-identification. ACM T Multim Comput. 2017;14(1):13.

17. Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person re-identificationn. 2017. arXiv:1703.07737.
18. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition. Las Vegas, Nevada; 2016. p. 770–8.
19. Zhang Z, Huang M. Learning local embedding deep features for person re-identification in camera networks. Eurasip J Wirel Comm. 2018;1–9.

# Indoor Visible Light Positioning and Tracking Method Using Kalman Filter

Xudong Wang, Wenjie Dong(✉), and Nan Wu

Information Science Technology College, Dalian Maritime University, Dalian 116026, Liaoning, China
{wxd, dwenjie}@dlmu.edu.cn

**Abstract.** In order to improve the accuracy and tracking performance of the indoor positioning system based on visible light communication (VLC), an indoor positioning and tracking method is proposed in this paper. This method utilizes time difference of arrival (TDOA) solved by nonlinear least squares (NLLS) method to realize indoor positioning and uses Kalman filter to obtain the tracking capability. The performance of the proposed positioning method is evaluated in the room measuring 5 m × 5 m × 3 m. The simulation results show that the average location errors by adopting the NLLS method can reach to 2.99 cm and the accuracy of positioning can be promoted to 1.33 cm by using Kalman filter, the positioning accuracy increased by 55.52%.

**Keywords:** Indoor positioning · Visible light communication · Time difference of arrival · Kalman filter

## 1 Introduction

In recent years, visible light communication technology based on white light LED has gained more and more attention in academia and industry [1, 2]. White light LED with high-speed modulation and short response time as well as other characteristics makes the application of LED extended from illumination to communication. The dual function of illumination and communication can be achieved simultaneously. Visible light communication (VLC) as a new means of communication has obvious advantages in terms of electromagnetic radiation, application environment, and safety compared with the RF wireless communication [3]. For these reasons, visible light is an effective choice for indoor positioning. Recently, VLC-based positioning is of particular interest hotspot to be further investigated [4–6]. Visible light positioning technique including received signal strength (RSS), time of arrival (TOA), angle of arrival (AOA), and time difference of arrival (TDOA) can be used currently [7–9]. The technique of AOA can achieve high accuracy, but require deploying an array of image sensors which is expensive at the receiver side. The technique of RSS is relatively easier to realize, but it is hard to acquire a high accuracy. In indoor positioning, the traveling time of signal is so short due to the short distance between transmitters and receiver. This makes TOA method need additional hardware equipment because it requires a precision of the clock at the receiver side. But for TDOA technique, the synchronization between transmitters and the receiver is not necessary. As long as the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_141

information of time difference of arrival can be achieved, we can achieve high estimation accuracy.

In this paper, we use the TDOA method to realize indoor positioning. By considering the transmission time differences between each LED and positioning node are obtained to construct an objective function of positioning, and then the target location is computed through nonlinear least squares method. Furthermore, the Kalman filter is applied to modify the positioning results and promote the accuracy of mobile users tracking [10]. Simulation results show that a high positioning accuracy has been achieved when the Kalman filter is adopted.

The remainder of this paper is organized as follows. The principle of VLC positioning is presented in Sect. 2. In Sect. 3, a method of indoor positioning using NLLS and Kalman filter is proposed. Numerous simulation results and performance analysis are given in Sect. 4. Finally, the main conclusions of this paper are summarized in Sect. 5.

## 2 Principle of VLC Positioning

### 2.1 Positioning System Structure

VLC as the most promising technology of indoor wireless communication can provide various information transmission scheme. A VLC indoor location system structure diagram is shown in Fig. 1. Our system is modeled in the room measuring 5 m × 5 m × 3 m. Assuming the coordinate of LED, $Tx_i$ ($i$ = 1, 2, 3, 4) to be $Tx_i$ ($x_i$, $y_i$, $H$), where $H$ is the height of the room. The four LEDs located in (1 m, 1 m, $H$), (1 m, 4 m, $H$), (4 m, 1 m, $H$), and (4 m, 4 m, $H$), respectively, where $H$ = 3 m. When VLC is applied to indoor positioning, lighting LED can transmit reference signal, and the target position can be obtained by different positioning algorithm. This layout ensures that the mobile user in the radiation zone can receive information transmitted from different LEDs at the same time. The receiver can be placed at any point in the room.

### 2.2 VLC Channel Model

LED is used as the reference signal emission source in this positioning system based on VLC. Only direct signal is considered in this paper as direct signal strength is far greater than the reflection and diffraction signal strength. Lambertian radiation pattern is applied. In optical wireless channel, if the transmission distance $d$ is much larger than the detector size $A_r$, then the received irradiance is approximately constant over the surface of the detector and all of the signals arrive at the same time. Thus, the impulse response of VLC wireless channel can be expressed as

$$h(t) = \frac{A_r(m+1)}{2\pi d^2}\cos^m(\varphi)T_s(\phi)g(\phi)\delta\left(t-\frac{d}{c}\right) \quad 0 \le \phi \le \text{FOV} \tag{1}$$

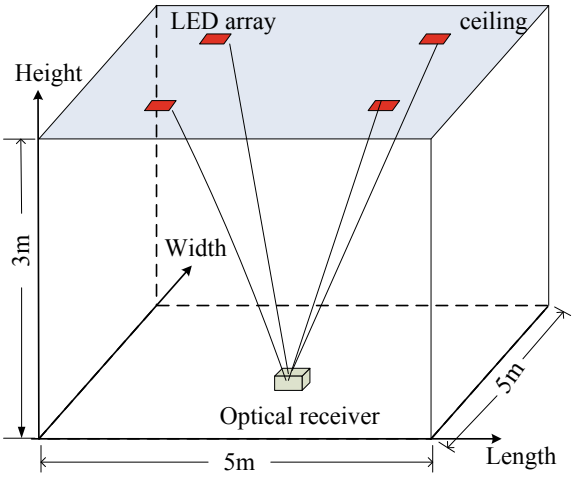



**Fig. 1.** System model of visible light positioning

where FOV is field of view of the receiver, $c$ is the speed of the light in free space (approximately $3.00 \times 10^8$ m/s), $\varphi$ is the light angle of incidence, $\phi$ is the angle of reception of the light at the receiver, $Ts(\phi)$ is the optical filter gain at the receiver, $g(\phi)$ is the gain of the optical concentrator, and m is the mode number of the radiation lobe given by $m = -\ln 2/\ln\left(\cos \varphi_{1/2}\right)$, where $\varphi_{1/2}$ is the LED semi-angle at half power. Thus, in VLC system, receiver signal $y(t)$ can be approximated as

$$y(t) = H(0) \cdot s(t) + n(t) \tag{2}$$

where $s(t)$ represents the incident optical signal, $H(0)$ is the channel DC gain given by $H(0) = \int_{-\infty}^{+\infty} h(t)\mathrm{d}t$, $n(t)$ denotes shot noise and thermal noise in the measurement and is modeled as AWGN with mean of zero and variance $\sigma^2$ given as $\sigma^2 = \sigma^2_{\text{shot}} + \sigma^2_{\text{thermal}}$. The variance for shot noise and thermal noise is $\sigma^2_{\text{shot}}$ and $\sigma^2_{\text{thermal}}$, respectively. The electrical *SNR* can be expressed in terms of the photodetector responsivity $R$, received optical power $P_r$, and noise variance as

$$\text{SNR} = \frac{(RP_r)^2}{\sigma^2} \tag{3}$$

## 3 Positioning Method

### 3.1 Time Delay Estimation

Each LED transmits a single pulse signal after every guard time period $\tau$. All signals have the same pulse width determined by the size of positioning room and width of

guard time. According to the simulation environment, the pulse width is chosen as 30 ns. The length of guard time period is set to be 100 ns so that the previous signal causes no or very little interference to the next signal. In the VLC channel model, the amplitude of the receiver signal is changed and arrives with different delay time. Assuming four transmitters is used in this model and the received four signals are $s_1(t)$, $s_2(t)$, $s_3(t)$, and $s_4(t)$, the waveforms with noise interference are shown in Fig. 2. The cross-correlation is used to detect each signal. Then the information of peak time of correlation signal is obtained. If correlation peak time is $\tau_1$, $\tau_2$, $\tau_3$, and $\tau_4$, respectively, the delay difference of arrival time between various signals can be given as $\Delta t_{21} = \tau_2 - \tau_1 - \tau$, $\Delta t_{31} = \tau_3 - \tau_1 - 2\tau$, $\Delta t_{41} = \tau_4 - \tau_1 - 3\tau$. Using this value, we can obtain the transmission distance difference $R_{21} = \Delta t_{21} \times c$, $R_{31} = \Delta t_{31} \times c$, $R_{41} = \Delta t_{41} \times c$. Figure 3 shows how the time delay difference is obtained using cross-correlation detection.

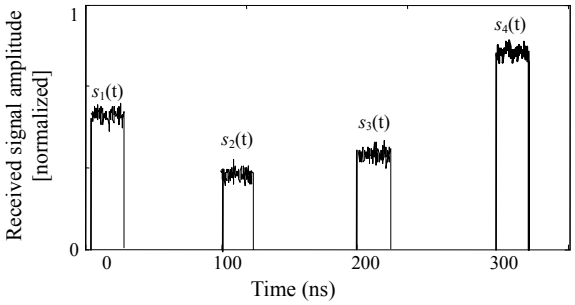



**Fig. 2.** Received signal waveform

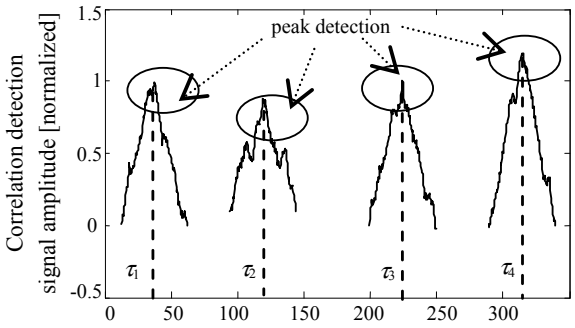



**Fig. 3.** Correlation detection signal

### 3.2 NLLS Method

As we all know, a number of methods can be applied to realize the TDOA positioning in the indoor scenario, such as the Chan method, the Fang method, and Taylor series

expansion method. Due to the short transmission distance in indoor environment, even very small detection error of time would lead to a large deviation of the eventual positioning results. Thus, they are not suitable for high precision indoor positioning problem in terms of VLC. In this paper, we convert position estimation into the problem of nonlinear optimization with constraints and then use nonlinear least squares (NLLS) algorithm to achieve high precision indoor positioning. We assume that the coordinates of mobile user and beacons used in location are set to be $(x, y)$ and $(x_i, y_i, z_i)$, $i = 1, 2, 3 \ldots M$, where $M$ is the number of beacons. We can get the time difference $\tau s1$ between the sth signal and the first signal (as reference) to the target receiver through correlation calculation. So, the distance difference can be expressed as $R_{s1} = R_s - R_1$ ($s = 2, 3 \ldots M$). Considering the case of $M = 4$, for TDOA positioning estimation, the objective function can be formulated as follows

$$\begin{aligned} &E(x,y) = (d_{21} - R_{21})^2 + (d_{31} - R_{31})^2 + (d_{41} - R_{41})^2 \\ &\text{s.t}\begin{cases} 0 \le x \le d_{\mathrm{L}} \\ 0 \le y \le d_{\mathrm{W}} \end{cases} \end{aligned} \tag{4}$$

where $d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + H^2}$, $d_{i1} = d_i - d_1$ ($i = 2, 3 \ldots M$), $d_{\mathrm{L}}$, and $d_{\mathrm{W}}$ are the length and width of the room, respectively. The coordinate $(x, y)$ of mobile user can be selected when $E(x, y)$ is minimized making use of nonlinear least squares algorithm to solve the above problem.

### 3.3 Kalman Filter

To realize indoor tracking for the mobile user, the Kalman filter (KF), which is a recursive state estimator of a linear dynamic system, is introduced into the positioning method mentioned above. The Kalman filter addresses the problem of estimating the state of a noisy system that can be described by a linear system and a linear measurement model. It generally describes the behavior of dynamical systems. The dynamic model can be described by

$$\boldsymbol{X}_k = \boldsymbol{A}\boldsymbol{X}_{k-1} + \boldsymbol{B}\boldsymbol{U}_{k-1} + \boldsymbol{W}_{k-1} \tag{5}$$

$$\boldsymbol{Z}_k = \boldsymbol{H}\boldsymbol{X}_k + \boldsymbol{V}_k \tag{6}$$

where $X_k$ is the vector of system state at time $k$, $Z_k$ is the vector of system measurements, $U_k$ is the control vector, $A$ is the state transition matrix, $B$ is the system control matrix, and $H$ is the measurement matrix. The state noise is $W_k \sim N(0, Q_k)$, zero mean, and Gaussian distributed noise with covariance $Q_k$, and $V_k \sim N(0, R_k)$ is the measurement noise with covariance $R_k$.

The Kalman filter can be divided into two steps as follow:

(1) The prediction process:

$$\boldsymbol{X}_{k,k-1} = \boldsymbol{A}\boldsymbol{X}_{k-1} + \boldsymbol{B}\boldsymbol{U}_{k-1} \tag{7}$$

$$\boldsymbol{P}_{k,k-1} = \boldsymbol{A}\boldsymbol{P}_{k-1}\boldsymbol{A}^T + \boldsymbol{Q} \tag{8}$$

where $X_k$,$k-1$ is the priori state estimate at time $k$ given knowledge of the process prior to time $k-1$, and $P_k$,$k-1$ is the covariance matrix of the state-prediction error.

(2) The updated process:

$$\begin{aligned} \boldsymbol{K}_k &= \boldsymbol{P}_{k,k-1}\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{P}_{k,k-1}\boldsymbol{H}^T + \boldsymbol{R})^{-1} \\ \boldsymbol{X}_k &= \boldsymbol{X}_{k,k-1} + \boldsymbol{K}_k(\boldsymbol{Z}_k - \boldsymbol{H}\boldsymbol{X}_{k,k-1}) \\ \boldsymbol{P}_k &= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H})\boldsymbol{P}_{k,k-1} \end{aligned} \tag{9}$$

where $X_k$ is the posteriori estimate at time $k$, $K$ denotes the Kalman gain, and $Z_k$ is the measurement matrix.

## 4 Simulation Results

Numerical simulations are performed to evaluate the performance of the considered TDOA-based positioning and Kalman filter-based tracking algorithm. We assume that the speed of mobile target is constant, and the average speed of moving on the $x$-axis and $y$-axis is 0.1 and 0.05 m/s, respectively. And the initial position of mobile user is set at (1, 1.5). The sampling interval is 1 s. In order to evaluate the performance of positioning, the position error is defined as

$$ER_i = \sqrt{(x_{ei} - x_{ci})^2 + (y_{ei} - y_{ci})^2} \tag{10}$$

where $ER_i$ is the position error of the node numbered $i$, and ($x_{ei}$, $y_{ei}$) and ($x_{ci}$, $y_{ci}$) is the estimation coordinate and real coordinate of the detection node, respectively.

The positioning error of proposed positioning method with and without KF is shown in Fig. 4. It can be seen that the average positioning error with and without KF is 1.33 and 2.99 cm, respectively. It is recorded that KF generates a 55.52% higher accuracy than without KF. The error of positioning is remarkably decreased throughout KF. Figure 5 illustrates the cumulative distribution of positioning error, which indicates that the error with and without KF is guaranteed to be within 2.78 and 6.03 cm 95% confidence interval. It proved that KF method has achieved higher performance.

It is generally acknowledged that SNR is lower around the room edge, and lower SNR will result in lower accuracy of positioning. We assume the initial position of mobile user is set at (0.5, 0.5) with SNR = −2 dB, and the user moves on along the

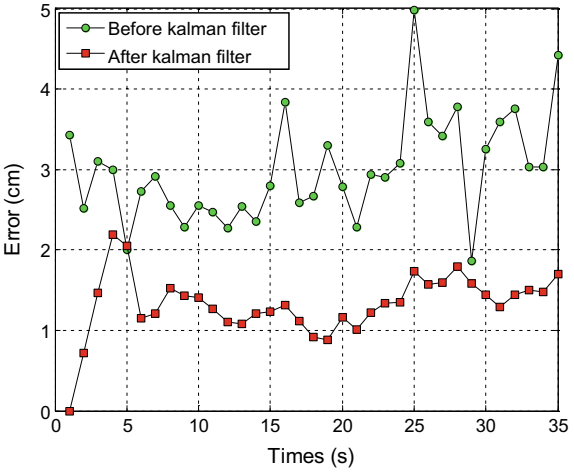



**Fig. 4.** Positioning error with and without KF

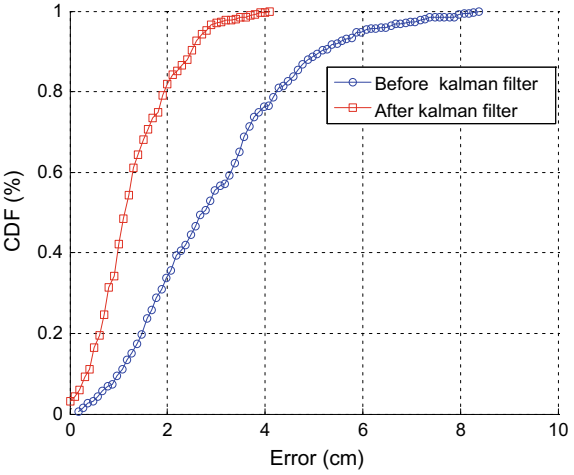



**Fig. 5.** CDF of positioning error

room edge. The positioning trajectory is shown in Fig. 6. The average positioning error with and without KF are respectively 1.79 and 3.62 cm. The positioning accuracy significantly increases by 50.55%. It proves that accuracy of positioning has been promoted when the KF is applied. Figure 7 shows the trajectory of mobile user move on the *x*-axis. As we can see, the filtering path is closer to actual path. This proves that the proposed algorithm is still valid when target moves to indoor corner areas.

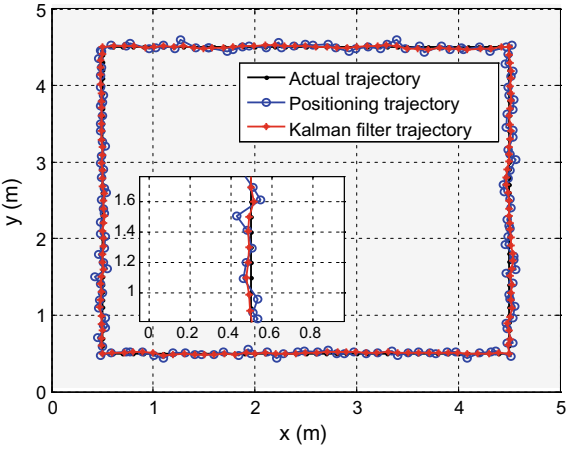



**Fig. 6.** Trajectory of users around the room edge

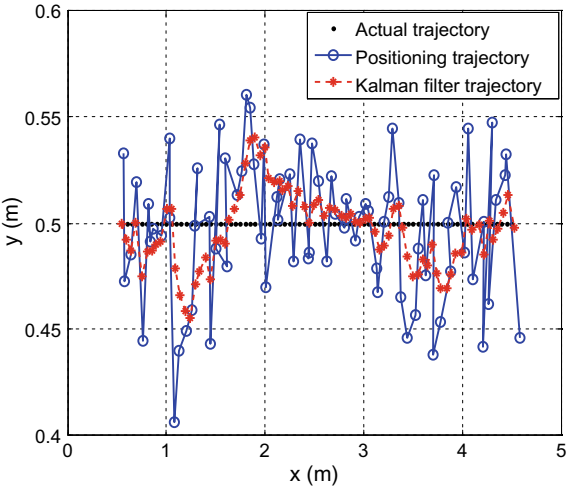



**Fig. 7.** Trajectory of users move on the *x*-axis

## 5 Conclusion

In this paper, we proposed an indoor positioning and tracking method for indoor positioning system based on VLC. In order to realize indoor positioning, a NLLS method was adopted. Main objective of this paper is then promoting the accuracy of proposed positioning method by using KF method. Evaluation by simulation proves that KF has achieved higher accuracy than the NLLS method, and the performance of positioning and tracking has been improved significantly.

## References

1. Pant K, Armstrong J. Indoor localization using white LEDs. Electron Lett. 2012;48(4):228–30.
2. Iturralde D, Azurdia-Meza C, Krommenacker N, et al. A new location system for an underground mining environment using visible light communications. In: International symposium on communication systems, networks and digital signal processing. IEEE, Manchester, UK; 2014. p. 1165–9.
3. Yan K, Zhou H, Xiao H, et al. Current status of indoor positioning system based on visible light. In: International conference on control, automation and systems. IEEE, Busan, South Korea; 2015. p. 565–9.
4. Dardari D, Closas P, Djuri PM. Indoor tracking: theory, methods, and technologies. IEEE Trans Veh Technol. 2015;64(4):1263–78.
5. Zhang W, Chowdhury M, Kavehrad M. Asynchronous indoor positioning system based on visible light communications. Opt Eng. 2014;53(4):045105.
6. Zhang W, Kavehrad M. A 2-D indoor localization system based on visible light. Seattle, USA: IEEE Photonics Society Summer Topical Meeting Series; 2012. p. 80–1.
7. Hassan NU, Naeem A, Pasha MA, et al. Indoor positioning using visible LED lights: a survey. ACM Comput Surv. 2015;48(2):20.
8. Do TH, Yoo M, An in-depth survey of visible light communication based positioning Systems. Sensors. 2016;16(5):678.
9. Ghassemlooy Z, Popoola W, Rajbhandari S. Optical wireless communications. Boca Raton: CRC Press, inc; 2012.
10. Eroglu YS, Guvenc I, Pala N, et al. AOA-based localization and tracking in multi-element VLC systems. In: Wireless and microwave technology conference. IEEE; 2015. p. 1–5.

# A Novel Method for 2D DOA Estimation Based on URA in Massive MIMO Systems

Bo Wang, Deliang Liu(✉), Dong Han, and Zhuanghe Zhang

Army Engineering University, Shijiazhuang, China
{liudeliang82,han58228}@sina.com

**Abstract.** Massive MIMO is one of the enabling technologies to cope with exponential data growth. It is very crucial for downlink precoding to accurately estimate direction-of-arrival. A lot of work has been done for 2D DOA estimation based on uniform rectangular array. However, in cases that snapshots are severely limited, such as extremely complex communication environment, conventional 2D DOA estimation method cannot work properly. Iterative adaptive approach (IAA) is one of the sparse algorithms which can handle heavy snapshot limitations. In this paper, we propose a novel 2D DOA estimation algorithm based on IAA for massive MIMO systems. Unlike conventional methods, an estimator in this algorithm is updated by previous iteration instead of the snapshots. The iteration ends until convergence. Simulation results demonstrate that the proposed algorithm is superior to conventional methods in low snapshot cases.

**Keywords:** 2D DOA estimation · Massive MIMO · Few snapshots

## 1 Introduction

It is a great challenge for future wireless systems to handle exponential traffic data growth. To meet such a challenge, the massive MIMO technology has received substantial attention for its higher data rates, enhanced link reliability, and potential power savings [1]. Accurate 2D DOA estimation is very important for massive MIMO systems to achieve reliable downlink communication.

In recent years, a lot of literatures have been addressed for 2D DOA estimation in massive MIMO. Classical multiple signal classification (MUSIC) is expanded into 2D version in [2] by using a unique steering matrix structure. This algorithm still requires spectral peak search and has high computational complexity. A low-complexity 2D DOA estimation algorithm is developed in [3] by utilizing the rotational invariance of two pairs of shifted URAs, in which a large scale of URA is needed. 2D DOA estimation based on ESPRIT without spectrum peak search is investigated in [4], at which DOA and the channel impulse responses are jointly estimated. 2D unitary ESPRIT is applied to massive MIMO systems in [5], which reduces computational complexity compared with 2D MUSIC. However, these methods are not suitable for coherent multipath signals.

The above-mentioned methods assume that snapshot numbers are large. However, in many practical application scenarios, such as underwater or indoor signal processing,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_142

signals are easily interfered or even interrupted, which result in snapshot numbers for 2D DOA estimation are very low. In this case, the performance of these algorithms deteriorates dramatically.

Sparsity-based algorithms can deal with snapshots limitation. In [6], Yardibi. T, etc. propose a user-parameter-free, weighted-least-square-(WLS)-based iterative adaptive approach for amplitude and phase estimation (IAA-APES), which has excellent estimation performance with few snapshots.

In this paper, we propose a novel 2D DOA estimation method based on IAA-APES for 2D DOA estimation in massive MIMO systems. In the first step, the optimal weighted least square estimator is obtained. In the second step, utilizing the estimator to update the energy of each scanning angle point. Step 1 and step 2 are iteratively calculated until convergence. Finally, the azimuth and elevation angles can be got from the spectrum peak without pairing process.

Notation: $(\cdot)^*, (\cdot)^T, (\cdot)^H$ denote the complex conjugate, transpose, Hermitian transpose, respectively. Symbol '$\otimes$' denotes Kronecker product, '$\odot$' stands for Khatri-Rao product.

## 2 Problem Formulation

In this paper, we consider a uniform rectangular array (URA) with $M \times N$ antenna elements is equipped with the base station. Assume that URA locates in the *xoz* plane. The positive direction of *y*-axis is determined by the right-hand rule from *x*-axis and *z*-axis. The azimuth angle with respect to *x*-axis, $\varphi_k \in (0°, 180°)$, and the elevation angle with respect to *z*-axis, $\theta_k \in (0°, 90°)$ (Fig. 1).

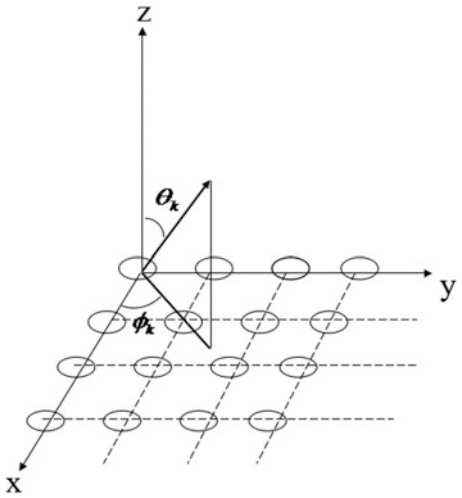



**Fig. 1.** URA for massive MIMO systems

Without loss of generality, the reference point is placed at the antenna element in origin $o$. The $(m,n)$th antenna element locates in the $xoz$ plane with coordinate, $(\frac{\lambda}{2}m, 0, \frac{\lambda}{2}n)$ where $\lambda$ is the wavelength of carrier frequency, antenna element spacing $d = \lambda/2$, and $m = 0, 1, \ldots, M-1$; $n = 0, 1, \ldots, N-1$. We assume that there are $K$ resolvable far-field narrowband signals, each of which has its own angle pair $\theta_k$ and $\varphi_k$, $k = 1, 2, \ldots, K$.

The received signal at $q$th snapshot on $(m,n)$th antenna element can be written as [7]

$$y_{m,n}(q) = \sum_{k=1}^{K} \gamma_k(q) \mathrm{e}^{j(m-1)\mu_k} \mathrm{e}^{j(n-1)\upsilon_k} + \omega_{m,n}(q) \tag{1}$$

where

$$\mu_k = \frac{2\pi d \cos\theta_k}{\lambda},\ \upsilon_k = \frac{2\pi d \cos\phi_k}{\lambda} \tag{2}$$

$\gamma_k(q)$ represents the unknown complex amplitude of the $k$th source at $q$th snapshot with power

$$E\{|\gamma_k(q)|^2\} = \sigma_k^2,\ k = 1, 2, \ldots, K \tag{3}$$

$\omega_{m,n}(q)$ is the Gaussian white noise with variance $\sigma^2$. The total number of snapshots is $Q$.

Let $\mathbf{a}(\mu_k)$ and $\mathbf{b}(\upsilon_k)$ be the $z$-axis and $x$-axis steering vector, respectively. Among them,

$$\mathbf{a}(\mu_k) = [1, \mathrm{e}^{j\pi\mu_k}, \mathrm{e}^{j2\pi\mu_k}, \ldots, \mathrm{e}^{j\pi(M-1)\mu_k}]^{\mathrm{T}} \tag{4}$$

$$\mathbf{b}(\upsilon_k) = [1, \mathrm{e}^{j\pi\upsilon_k}, \mathrm{e}^{j2\pi\upsilon_k}, \ldots, \mathrm{e}^{j\pi(M-1)\upsilon_k}]^{\mathrm{T}} \tag{5}$$

Assume that

$$\mathbf{y}(q) = [y_{1,1}(q), y_{2,1}(q), \ldots, y_{M,1}(q), \ldots, y_{M,N}(q)]^{\mathrm{T}} \tag{6}$$

be the data matrix at $q$th snapshot, then it has

$$\mathbf{y}(q) = [\mathbf{b}(\upsilon_1) \otimes \mathbf{a}(\mu_1), \ldots, \mathbf{b}(\upsilon_k) \otimes \mathbf{a}(\mu_k)]\boldsymbol{\gamma}(q) + \boldsymbol{\omega}(q) \tag{7}$$

where vector $\boldsymbol{\gamma}(q) = [\gamma_1(q), \gamma_2(q), \ldots, \gamma_K(q)]^{\mathrm{T}}$, and $(m,n)$ $\boldsymbol{\omega}(q)$ has the same data stacking form as $\mathbf{y}(q)$. Furthermore, defining $\mathbf{A} = [\mathbf{a}(\mu_1), \mathbf{a}(\mu_2), \ldots, \mathbf{a}(\mu_K)]^{\mathrm{T}}$, $\mathbf{B} = [\mathbf{b}(\upsilon_1), \mathbf{b}(\upsilon_2), \ldots, \mathbf{b}(\upsilon_k)]^{\mathrm{T}}$, and considering total $Q$ snapshots, (7) can be rewritten as a compact form

$$\mathbf{Y} = (\mathbf{B} \odot \mathbf{A})\mathbf{S} + \mathbf{Z} \tag{8}$$

where $\mathbf{Y} = [y(1), y(2), \ldots, y(Q)]$, $\mathbf{S} = [\gamma(1), \gamma(2), \ldots, \gamma(Q)]$ and $\mathbf{Z} = [\boldsymbol{\omega}(1), \boldsymbol{\omega}(2), \ldots, \boldsymbol{\omega}(Q)]$.

## 3 Proposed Algorithm

First of all, we refine the interested region into a grid with predefined increment. Each point of the grid is assumed as a potential signal with an azimuth and elevation angle. Actual signals are much smaller than potential ones so that sparse representation algorithm can be used in 2D DOA estimation.

We make an assumption that $\mathbf{P}$ is a $(m, n)K \times K$ diagonal matrix, whose diagonal is the power estimation of each scanning point. $\mathbf{P}$ is given by

$$P_k = \frac{1}{N} \sum_{n=1}^{N} |\gamma_k(q)|^2, \; k = 1, 2, \ldots, K \tag{9}$$

In addition, the noise covariance matrix is written as

$$\mathbf{Q}(\theta_k, \phi_k) = \mathbf{R} - P_k \mathbf{h}(\theta_k, \phi_k)\mathbf{h}(\theta_k, \phi_k)^H \tag{10}$$

in which

$$(m, n)\ \mathbf{h}(\theta_k, \phi_k) = \mathbf{b}(\upsilon_k) \otimes \mathbf{a}(\mu_k) \tag{11}$$

$\mathbf{R}$ is given by $\mathbf{R} \triangleq (\mathbf{B} \odot \mathbf{A})\mathbf{P}(\mathbf{B} \odot \mathbf{A})^H$, which is dependent on the unknown signal power.

Then, the weighted least square (WLS) cost function is expressed as

$$\sum_{t=1}^{Q} ||\mathbf{y}(t) - \gamma_k(t)\mathbf{h}(\theta_k, \phi_k)||^2_{\mathbf{Q}^{-1}} \tag{12}$$

where $\|\mathbf{x}\|^2_{\mathbf{Q}^{-1}} = \mathbf{x}^H \mathbf{Q} \mathbf{x}$. Minimizing (12) with respect to $\gamma_k(q)$, $t = 1, 2, \ldots, Q$, yields

$$\hat{\gamma}_k(t) = \frac{\mathbf{h}(\theta_k, \phi_k)^H \mathbf{Q}(\theta_k, \phi_k)^{-1} \mathbf{y}(t)}{\mathbf{h}(\theta_k, \phi_k)^H \mathbf{Q}(\theta_k, \phi_k)^{-1} \mathbf{h}(\theta_k, \phi_k)} \tag{13}$$

among them $\mathbf{Q}(\theta_k, \phi_k)$ and $\mathbf{h}(\theta_k, \phi_k)$ are known before. According to (10) and the matrix inversion lemma, (13) can be expressed as

$$\hat{\gamma}_k(t) = \frac{\mathbf{h}(\theta_k, \phi_k)^H \mathbf{R}^{-1} \mathbf{y}(t)}{\mathbf{h}(\theta_k, \phi_k)^H \mathbf{R}^{-1} \mathbf{h}(\theta_k, \phi_k)} \tag{14}$$

Unlike conventional DOA estimation methods, IAA-APES [6] gets $\mathbf{Q}_k$ from iteration.

Table 1 shows the calculation steps. The initialization of $\mathbf{P}_k$ is got from traditional DAS method

$$\hat{P}_k = \mathbf{h}(\theta_k, \phi_k)^H \mathbf{\Gamma}\, \mathbf{h}(\theta_k, \phi_k), k = 1, \ldots, K \tag{15}$$

$$\mathbf{\Gamma} = \frac{1}{Q}\sum_{t=1}^{Q} \mathbf{y}(q)\mathbf{y}^H(q) \tag{16}$$

The azimuth and elevation angles' estimation is obtained from spectrum peak after the above steps.

**Table 1.** Proposed algorithm

| |
|---|
| **Initialization** |
| calulate by $\hat{\mathbf{P}}_k^{(0)}$ (15) |
| **Iteration** |
| $\mathbf{R}^{(i)} = (\mathbf{B} \odot \mathbf{A})\hat{\mathbf{P}}^{(i-1)}(\mathbf{B} \odot \mathbf{A})^H$ |
| for $k$=1,2,…,$K$ |
| calculate $\hat{\gamma}_k^{(i)}(q)$ by (14), $q$=1,2,…,$Q$ |
| calculate $\hat{\mathbf{P}}_k^{(i)}$ by (9) |
| end for |
| **Termination** |
| $\lvert \hat{\gamma}_k^{(i)}(q) - \hat{\gamma}_k^{(i-1)}(q) \rvert^2$ is less than a specified tolerance |

## 4 Numerical Examples

In this section, we demonstrate the performance and advantages of proposed method in comparison with the 2D MUSIC algorithm. Simulation is based on a URA with $M$ = 11 and $N$ = 11, and the element space along the $X$-axis and $Y$-axis is $\lambda/2$. The number of signals is $K$ = 3, and the search step for 2D MUSIC is 0.2°.

Assume one mobile terminal with three coherent signals impinging on the antenna array from three directions $\theta \in \{66°, 62°, 61°\}$ and $\phi \in \{62°, 66°, 68°\}$. The SNR and the number of snapshots are set to 12 dB and 5 respectively. The spatial spectrum

obtained from the proposed algorithm and MUSIC-based algorithm is shown in Figs. 2 and 3. From the results, we can see that under the case that snapshots are only five and three DOAs are closely separated, the 2D MUSIC algorithm fails to estimate the three directions while the proposed method can accurately resolve the three directions with three sharp spectrum peaks. The estimated DOAs are $(66°, 62°)$, $(62°, 66°)$ and $(68°, 61°)$, respectively. Figures 2 and 3 show that the proposed method is much superior to the 2D MUSIC algorithm in conditions of low snapshot numbers.

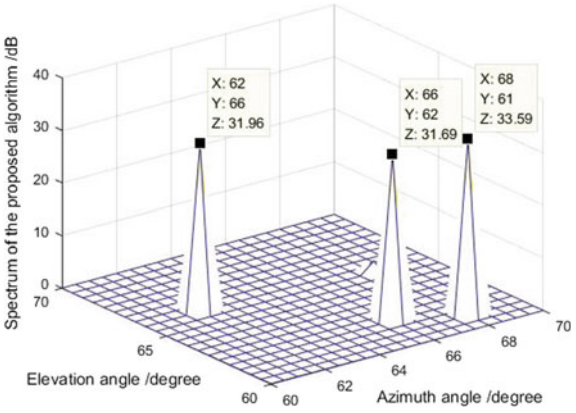



**Fig. 2.** Spectrum of the proposed method

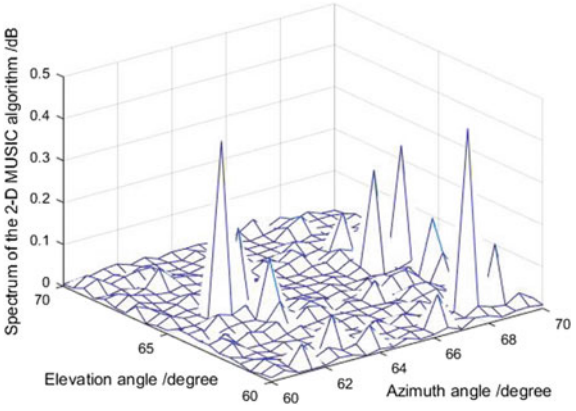



**Fig. 3.** Spectrum of a MUSIC-based algorithm

## 5 Conclusions

We present a new two-dimensional DOA estimation algorithm based on URA for massive MIMO systems. There are numerous practical applications that a large number of snapshots are unavailable, and this algorithm can effectively deal with this case. Moreover, it provides super-resolution. Simulation results verify the proposed algorithm has superior estimation performance with few snapshots.

**Acknowledgments.** This work is funded by the National Natural Science Foundation of China under Grant 61601494.

## References

1. Larsson E, Edfors O, Tufvesson F, Marzetta T. Massive MIMO for next generation wireless systems. IEEE Commun Mag. 2014;52(2):186–95.
2. Yang, K-Y, Wu J-Y, Li W-H. A low-complexity direction-of-arrival estimation algorithm for full-dimension massive MIMO systems. In: 2014 IEEE International Conference on Communication Systems (ICCS). IEEE; 2014. p. 472–6.
3. Meng H, Zheng Z, Yang Y, et al. A Low-complexity 2-D DOA estimation algorithm for massive MIMO systems. In: 2016 IEEE/CIC international conference on communications in China (ICCC). IEEE; 2016. p. 1–5.
4. Kuang J, Zhou Y, FEI Z. Joint DOA and channel estimation with data detection based on 2D unitary ESPRIT in massive MIMO systems. Front Inform Technol Electron Eng. 2017;18(6): 841–9.
5. Wang T, Ai B, He RS, Zhong ZD. Two-dimension direction-of-arrival estimation for massive MIMO system. IEEE Access. 2015;3:2122–8.
6. Yardibi T, Li J, Stoica P, et al. Source localization and sensing: a nonparametric iterative adaptive approach based on weighted least squares. IEEE Trans Aerosp Electron Syst. 2010;46(1):425–43.
7. Nion D, Sidiropoulos ND. Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar. IEEE Trans Signal Process. 2010;58(11):5693–705.

# Two-Dimensional DOA Estimation for 5G Networks

Zhuanghe Zhang(✉), Dong Han, Deliang Liu, and Bo Wang

Shijiazhuang Campus of Army Engineering University, Shijiazhuang, China
zhangzhuanghe1995@163.com

**Abstract.** Mobile communication is coming to the fifth generation (5G) networks. In the age of 5G, the three-dimensional (3D) beamforming is one of the highlighted technologies. 3D beamforming increases vertical dimension in terms of space domain, but also has brought a challenge to beamforming, especially the problem of direction-of-arrival (DOA) estimation. In this paper, a DOA estimation method based on multiple signal classification (MUSIC) of uniform circular array (UCA) with mutual coupling compensation is presented for the future 5G networks. 2D MUSIC does well in estimating DOA without an accurate number of signals. The MATLAB software is employed in order to conduct the modeling and 2D MUSIC algorithm simulations.

**Keywords:** 2D MUSIC · DOA · UCA · 5G networks

## 1 Introduction

In 5G networks, different from conventional beamforming techniques, the 3D beamforming shows us the possibility of combining the vertical dimension with horizontal dimension in order to enhance system performance [1]. So, the method for jointly estimating azimuth and elevation is important.

In [2], we can know the design of antenna array architecture, in which the antenna elements can be designed in the shapes of a hexagon, circle, or cross, in addition to the conventional rectangle. The simulation results indicate that while there always exists a non-trivial gain fluctuation in other common antenna arrays, the circular antenna array has a flat gain in the main lobe of the radiation pattern with varying angles. This makes the circular antenna array more robust to angle variations that frequently occur due to antenna vibration in a complex environment. So, in array signal processing, UCA with its unique array structure and widely spread, but in practical application, the mutual coupling effect is also very obvious, especially in the future 5G networks, which use the narrow millimeter wave beam technology and usually at a very high frequency of work.

DOA estimating techniques are of particular interest in communication systems, especially with antenna technology. There are many algorithms such as multiple signal classification (MUSIC), estimation of signal parameters via rotation invariance techniques (ESPRIT), but most of them have been proposed to estimating the one-dimensional direction of arrival with uniform linear array (ULA) due to their simplicity [3]. Recent years, more and more 2D DOA estimation algorithms are proposed,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_143

in [4], a 2D DOA based on iterative adaptive approach (IAA) algorithm has been introduced. As for [5], it presented a 2D DOA estimation with MUSIC algorithm using the modulus constraint, but most of the high-resolution DOA estimation algorithm for array mutual coupling effect caused by flow pattern error is very sensitive.

Aiming at the above problems, this article expounds how uniform circular array is established, and a DOA estimator is introduced to jointly estimate both azimuth and elevation based on MUSIC method under the condition of UCA mutual coupling effect, the computer simulation results by MATLAB verify the validity of it.

## 2 Data Model

In this section, consider an antenna with uniform circular array, in which N isotropic elements with radius $R$ as shown in Fig. 1. The center of the circle is at the origin of the frame, let K narrowband far-field sources at arbitrary locations whose angles are $\mathbf{a}(\theta, \varphi) = [(\theta_1, \varphi_1), (\theta_2, \varphi_2), \ldots, (\theta_K, \varphi_K)]$, respectively, impinge on this array, and wavelength of sources is denoted as $\lambda$ [6, 7].

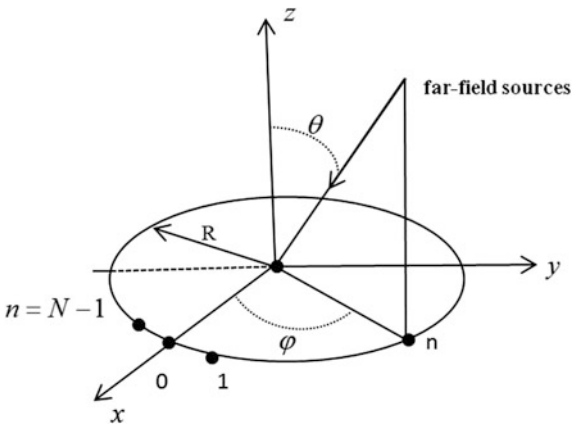



**Fig. 1.** Uniform circular array

In the ideal case, the received signals can be expressed as

$$X(t) = AS(t) + N(t) \tag{1}$$

And

$$A = [a(\theta_1, \varphi_1), a(\theta_2, \varphi_2), \ldots, a(\theta_K, \varphi_K)] \tag{2}$$

Matrix $A$ is a matrix of the $K$ steering vectors, which represents the possible value set of DOA.

In which

$$a(\theta_i, \varphi_i) = [e^{j2\pi R\sin\theta_i\cos\varphi_i/\lambda}, \ldots, e^{j2\pi R\sin\theta_i\cos(\varphi_i-\gamma_n)/\lambda}]^{\mathrm{T}} \tag{3}$$

where

$$\gamma_n = \frac{2\pi n}{N} \quad \begin{matrix} i = 1, 2, \ldots, K \\ n = 1, 2, \ldots, N \end{matrix} \tag{4}$$

And

$$S(t) = [s_1(t), s_2(t), \ldots, s_K(t)]^{\mathrm{T}} \tag{5}$$

$$N(t) = [n_1(t), n_2(t), \ldots, n_K(t)]^{\mathrm{T}} \tag{6}$$

where $S(t)$ is signal source vector of size $(K \times 1)$ and $N(t)$ is the received additive white Gaussian noise whose variance is $\sigma^2$. Here, it is assumed that there is no correlation between $S(t)$ and $N(t)$, based on (1), the correlation matrix of received vector can be computed as

$$R_x = E[XX^H] = AVA^H + \sigma^2 I = R_s + \sigma^2 I \tag{7}$$

$V$ is covariance matrix of signal vector $(S)$ which is a full rank matrix of order $(K \times K)$ given by

$$V = \begin{bmatrix} E\left[|S_1|^2\right] & \cdots & \cdots & 0 \\ 0 & E\left[|S_2|^2\right] & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & E\left[|S_K|^2\right] \end{bmatrix} \tag{8}$$

$R_s$ is a signal covariance matrix of order $N \times N$ with rank $K$ given by

$$R_s = \begin{bmatrix} E\left[|S_1|^2\right] & \cdots & \cdots & 0 & \cdots & 0 \\ 0 & E\left[|S_2|^2\right] & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & E\left[|S_K|^2\right] & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \tag{9}$$

## 3 UCA-MUSIC Algorithm

Given the decoupling effect, steering vector can be expressed as [8]

$$\tilde{a}(\theta, \varphi) = Z(\theta)a(\theta, \varphi) \tag{10}$$

where $Z(\theta)$ is mutual coupling matrix (MCM) of UCA, the elevation dependence of the mutual coupling effect is taken into account so that the MCM will vary with elevation angle for a fixed elevation angle, it is well known that a complex symmetric circular matrix provides a satisfactory model for the MCM of a UCA. If only consider about the mutual coupling between adjacent arrays of three elements, the cycle of cyclic matrix-vector can be defined as

$$z = [z(0), z(1), z(2), 0 \cdots, 0, z(2), z(1)] \tag{11}$$

According to (9), $R_S$ has $N-K$ eigenvectors corresponding to zero eigenvalues. Steering vector $\mathbf{a}(\theta, \varphi)$ in the signal subspace is orthogonal to noise subspace. Let $Q_n$ be such an eigenvector

$$R_s Q_n = AVA^H Q_n = 0 \tag{12}$$

Since $V$ is a positive definite matrix

$$a^H(\theta_i, \varphi_i)Q_n = 0 \tag{13}$$

This implies that signal steering vectors are orthogonal to eigenvectors corresponding to noise subspace. So, the MUSIC algorithm searches through all angles and plots the spatial spectrum

$$P_{\text{MUSIC}}(\theta, \varphi) = \frac{1}{(a^H(\theta, \varphi)Q_n Q_n^H a(\theta, \varphi))} \tag{14}$$

Therefore, when there is mutual coupling, the $m$th snapshot of the array can be expressed as

$$\tilde{X}(t) = \tilde{A}s(t) + n(t) = Z(\theta)\mathrm{A}s(t) + n(t) \tag{15}$$

Because of the existence of array mutual coupling error, the guiding vector of array becomes

$$\hat{A} = Z(\theta)\mathrm{A} \tag{16}$$

So, the actual direction vector is

$$\hat{a}(\theta, \varphi) = Z(\theta)a(\theta, \varphi) \tag{17}$$

Consider the effect of array mutual coupling error on array performance, the actual covariance matrix is

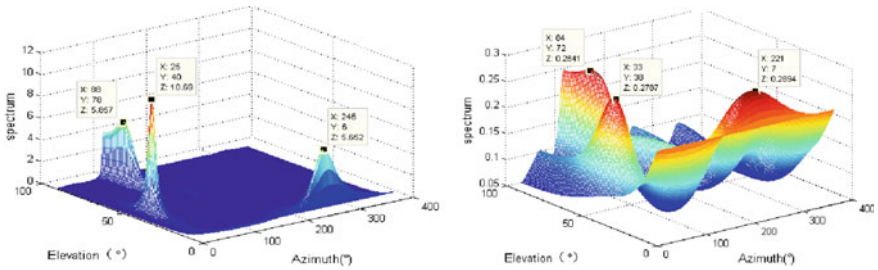
$$\hat{R} = ZR_xZ^H \tag{18}$$

Mutual coupling effect mainly affects the direction vector of array, thus to influence the array receiving signal matrix.

## 4 Simulation Results

In the first simulation, consider three sources which are uncorrelated with each other incident on a UCA consists of $N = 8$ elements, sources located at angles $(88°, 80°)$, $(25°, 40°)$, and $(245°, 5°)$. The number of snapshots is set to 500 and the signal-to-noise ratio (SNR) is 30 dB.

Figure 2 shows that three sources make three corresponding peaks in the plot when considering mutual coupling. However, it makes some pseudo peaks when decoupling error is not considered.

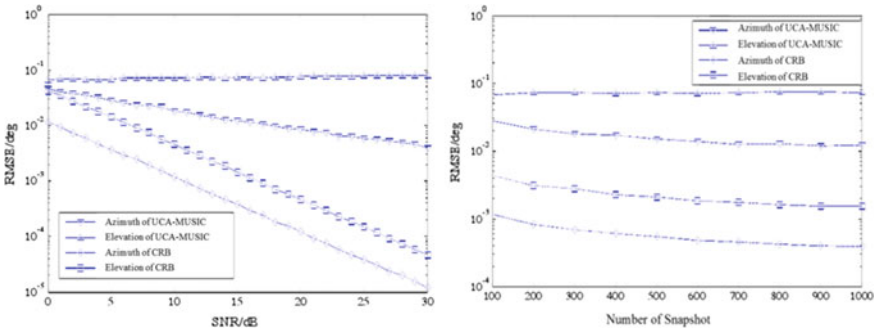



(a) consider coupling error (b) decoupling error is not considered

**Fig. 2.** Influence of the mutual coupling error

In the second experiment, we present 200 Monte Carlo simulations to illustrate the performance of proposed methods. Define root mean squared error (RMSE). The array structure is a uniform circular array (UCA) composed of $N = 8$ elements, and there are two noncoherent sources located at angles $(10°, 20°)$ and $(30°, 60°)$, the signal-to-noise ratio (SNR) is 15 dB in Fig. 3b. We demonstrate the performance of UCA-MUSIC algorithm with respect to SNR and the number of snapshots. Comparisons are performed with CRB.

From Fig. 3, RMES of the algorithm with mutual coupling compensation is in a downtrend with the increase of SNR, but elevation has no obvious changes. Also the performance of estimation gets better with the increase of snapshot.




(a) RMES performance of SNR (b) RMES performance of Snapshot

**Fig. 3.** Estimation performance of RMES

## 5 Conclusion

In this letter, An 2D DOA estimation algorithm based on MUSIC is proposed for UCA in the presence of mutual coupling. Simulations are presented to show the effectiveness of the proposed algorithm on SNR and snapshots. It can be concluded that the proposed algorithm improves the performance for both azimuth estimates and elevation estimates.

**Acknowledgments.** This work is supported by the National Natural Science Foundation of China under Grant 61601494.

## References

1. Kelif J-M, Coupechoux M, Mansanarez M. A 3D Beamforming analytical model for 5G wireless networks. In: 14th International Symposium on 2016. Modeling and optimization in mobile, Ad Hoc, and wireless networks. WiOPT 2016.
2. Zhang J, Ge X, Li Q, Guizani M, Zhang Y. 5G millimeter-wave antenna array: design and challenge. IEEE Wirel Commun. April 2017;106–12.
3. Yilmazer N, Koh J, Sarkar TK. Utilizaion of a unitary transform for efficient computation in the matrix pencil method find the direction of arrival. IEEE Trans Antenna Propag. 2006;54 (1):175–81.
4. Barcelo M, Vicario JL, Seco-Granados G. A reduced complexity approach to IAA beamforming for efficient DOA estimation of coherent sources. EURASIP J Adv Signal Process. 2011;1:1–1:16.
5. Cai JJ, Qin GD, Li P. Tow- dimensional DOA estimation with reduced- dimensional MUSIC algorithm using the modulus constraint. Syst Eng Electron. 2014;9:1681–6.
6. Mathews CP, Zoltowski MD. Performance analysis of the UCA-ESPRIT algorithm for circular ring arrays. IEEE Trans Signal Process. 1995;2535–9.
7. Yang L, Zhang H, Yang X, DOA estimation for wideband sources based on UCA. J. Electron. (China). 2006;128–31.
8. Hui HT. Improved compensation for the mutual coupling effect in a dipole array for direction finding, Antennas and Propagation. IEEE Trans. On. 2003;51:2498–503.

# A Grid-Map-Oriented UAV Flight Path Planning Algorithm Based on ACO Algorithm

Wei Tian and Zhihua Yang(✉)

Communications Engineering Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China
tianwei0323@foxmail.com, yangzhihua@hit.edu.cn

**Abstract.** With the extensive applications of unmanned aerial vehicle (UAV), typical algorithm for path planning is usually restricted for its low efficiency and easy failure, especially for the complex obstacle environments. Therefore, in this paper, a new UAV path planning algorithm is proposed based on ant colony optimization (ACO) for such complex obstacle environment. In particular, the proposed algorithm optimizes the distribution of pheromones and modifies the transfer probability by considering the regional security factors. As a result, it can increase search speed and avoid local optimum and deadlock. Simulation results verify the feasibility and effectiveness of the proposed method.

**Keywords:** UAV · Path planning · Grid map · ACO

## 1 Introduction

In these years, UAV takes more important roles in military missions and industrial applications, such as reconnaissance and tracking [1]. Generally, path planning enables an UAV with the capabilities of automatically deciding and executing a sequence of collision-free and safety motions in order to achieve certain tasks in a given environment [2]. Path planning is the basis and prerequisite for a series of autonomous control activities of UAV assignment planning system, such as formation control and multi-UAV coordination.

Currently, there are various types of path planning algorithms, such as A* algorithm, Dijkstra algorithm, artificial potential field algorithm, and genetic algorithm [3]. Among them, the ACO algorithm is widely used for its advantages on simple coding, strong robustness, and positive feedback. ACO algorithm was firstly used to solve traveling salesman problem (TSP) [4], by using a distributed feedback parallel computer system. Moreover, it is easy to merge with other algorithms, with strong robustness. Gradually, ACO algorithm was applied in the field of UAV path planning. However, this method also exists a few inherent flaws, including long search time, slow convergence, and easily trapped in a local optimum. In this regard, several researchers have done a lot of researches on the basis of the traditional ACO algorithm and have obtained good results [5]. From the early ACO algorithms, AS and ACS proposed by Dorigo M, the ant colony algorithm has been continuously improved and optimized. Stutzle et al. proposed MMAS [6] to improve the search in the early stages of the algorithm by limiting the concentration range of pheromones. In [7], the ant colony


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_144

algorithm is combined with the immune system to accelerate the convergence speed. Literature [8] combines the global search capability of genetic algorithm with the positive feedback of ACO algorithm to improve the efficiency of the algorithm.

Through the research on the ACO algorithm, it can be found that the bottleneck for slow search speed is the superfluous time consumed in initial search. At the beginning of the search, ants have few path information and the choices become relatively blind. With the increase of the cycle, the path is prosperous and the speed will be faster after convergence. Therefore, in this paper, aiming at the problem of too long searching time of the ant colony algorithm, the initial distribution of the pheromone is improved. In the initial stage, the ant is guided to reduce the wasted time in the initial search period and speed up the algorithm. In addition, considering that the algorithm is prone to fall into a local optimum, the improved algorithm optimizes the pheromone distribution strategy, by changing the pheromone within the effective range. As a result, the advantages of the excellent path are preserved, and the diversity of the solutions is increased by optimizing the probability transfer rule of global search.

## 2 System Model

### 2.1 Grid Map

This paper is based on two-dimensional environmental maps and uses grid method to establish the path planning environment model. In a raster map, each raster represents a region that has been rasterized to abstract the map with a uniform equal grid. Among them, each grid has a color attribute, i.e., white denotes a passable area, black denotes an impassable area and is regarded as an obstacle. The position information of the grid is described by the coordinate pair, in addition to the ID attribute that can be used to uniquely identify it. The map environment model established by the grid method is shown in Fig. 1a.

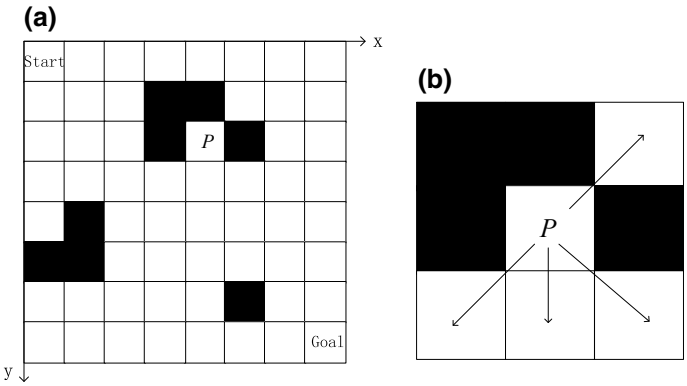



**Fig. 1.** **a** Grid map **b** Possible routes

As shown in Fig. 1a, there are several obstacles with uneven distribution and different shapes. The goal of path planning is to find an optimal path between the starting point and the end point. It is also required to keep away from obstacles while requiring the shortest distance, to strengthen the safety of the path. In Fig. 1a, the start and end points are: the grid with ID 0 in the upper left corner and the grid with ID (the grid number-1) in the lower right corner. In the path-finding process, each grid can be regarded as a node, and the UAV searched path is a set of line segments connected by the traversed nodes. In this paper, the default UAV flight range is one step at a time, that is, it passes through a grid node in any direction. The direction in which the UAV can fly is eight directions around the current position, allowing the selected node to be a node in the adjacent node that is not an obstacle and has not been traversed. Nodes that are not traversed here refer to nodes marked 0 in the tabu list. For example, in the position of *P* in Fig. 1a, the next optional flight path is shown in Fig. 1b.

### 2.2 Path Planning Problem

In the classic ACO algorithm, when the pheromone is initialized, the amount of pheromone in each path is unified to a certain value. Therefore, it takes a lot of time to clear the optimal path in the early stage of searching, which makes the convergence speed of the algorithm slow. In the grid map, there are eight directions for each step. In addition to the obstacles that are not accessible, there are still many types of next routes that can be selected. The same setting of the initial pheromone may cause the ant to choose the direction opposite to the endpoint. The regional nodes either take the route, which will lead to the increase of the search time, the length of the route, and for the physical UAV, the frequent conversion direction during the flight will increase the time consumption.

Besides, in a complex map environment, when an ant's current position in the next selectable node list is an obstacle or a traversed grid, it cannot select a node to transfer. At this time, the ant is caught in a deadlock. Figure 2 shows a schematic of a deadlock. The current path node sequence in the figure is 0–4–5–9–8–12. When the ant is in the 12th grid, there is no next node to reach. A rotor UAV is used in this paper. It is assumed that a certain track has a total of *N* nodes. As shown in Fig. 3, the distance from section $i$ is $l_i$, and the total path of the track can be expressed as $L = \sum_{i=1}^{N-1} l_i$.

## 3 Algorithm Descriptions

### 3.1 Algorithm Improvements

An ant releases a certain amount of pheromone after passing through a path. Later ants adjust the transition probability according to the number of pheromone on the path and make path selection. Therefore, the distribution of pheromone is very important for the advantages and disadvantages of ant's global path planning. This article makes the following improvements to the pheromone distribution strategy in the original algorithm.

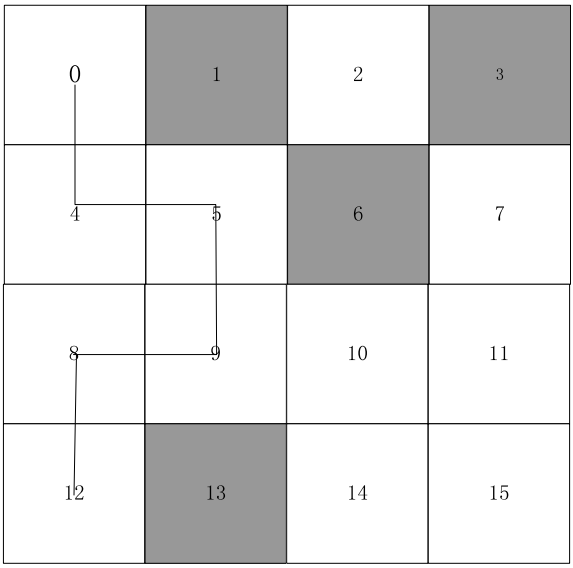



**Fig. 2.** Stuck in a deadlock

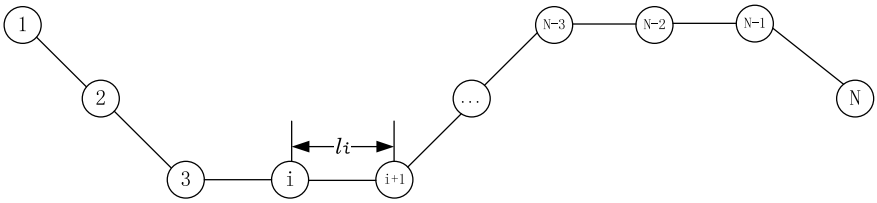



**Fig. 3.** Path diagram of UAV

**Pheromone Distribution Optimization**. For maps like the one shown in Fig. 1a, the starting point and the ending point have already been determined, so the global direction of the ant's advance can also be determined. This article uses the direction guidance to distribute the initial pheromone on each path. Take Fig. 1a as an example. The revised rules are as follows:

$$P_{mn} = \begin{cases} p & n = m - 1 \, or \, n = m - \text{ col or } n = m - \text{ col } - 1 \\ 2 * p & n = m - col + 1 \text{ or } n = m = \text{ col } - 1 \\ 3 * p & n = m + 1 \text{ or } n = m + \text{ col or } n = m + \text{ col } + 1 \\ 0 & \text{else} \end{cases} \tag{1}$$

where $p$ is the initial amount of pheromone, col is the number of rows in the raster map, and $mn$ is the path from raster $m$ to raster $n$. According to the location of the grid $n$ in the path relative to the grid $m$, different initial pheromones are set, and a global direction guidance is given. It can be seen here that $P_{mn} \neq P_{nm}$. The volatilization of

pheromone directly relates to the efficiency of the algorithm and the ability of global search. In the volatile of pheromones, we must maintain the advantages of the excellent path and concurrently do not affect the choice of follow-up ants too much. The pheromone volatility rule in the classical algorithm is: $\tau_{ij}(t+n) = \tau_{ij}(t) * (1-\rho)$, where $\rho$ is the volatility coefficient. This article proposes an adaptive volatility strategy based on path length:

$$\tau_{ij}^{*} = \begin{cases} 1 - \frac{L_{\text{best}}}{L_{\text{better}}} * \rho & (a) \\ 1 - \frac{L}{L_{\text{better}}} * \rho & (b) \end{cases}. \tag{2}$$

Among them, $L$ represents the length of the path searched by the ant, $L_{\text{best}}$ represents the shortest path of the current cycle, and $L_{\text{better}}$ represents the shortest path of history. If the path length obtained by the search is the shortest path length of the current period, use the formula (a) to volatilize the path segment on the path. Otherwise, use formula (b). Here, a lower limit of pheromone is given as the pheromone initial value $P$. The volatilization strategy based on path length gives the corresponding volatilization weight for different path lengths, which not only preserves the advantages of good path, but also guides the subsequent selection.

**Pheromone Updating Rule**. The pheromone increment calculation formula in the original algorithm is:

$$\Delta\tau_{ij}(t) = \sum_{k=1}^{\text{AntNum}} \Delta\tau_{ij}^{k}(t) \tag{3}$$

in which

$$\Delta\tau_{ij}^{k}(t) = \begin{cases} \frac{Q}{L} & \text{if ant k goes though } (i,j) \\ 0 & \text{else} \end{cases} \tag{4}$$

Among them, $Q$ is the total amount of pheromone, which is a fixed value. AntNum is the ant population. In order to manifest the advantages of the historical best path, preserve the excellent path information and speed up the convergence, the improved algorithm gives different update rules for each path segment in the local optimal solution and the current historical optimal solution. Apply the formula (5) to calculate the pheromone increment for the results of the path search in this cycle.

$$\Delta\tau_{ij}^{k} = \begin{cases} 2 * \frac{Q}{L} + \frac{Q}{L_{\text{better}}} & ij \in \text{optimal path} \\ 0 & \text{else} \end{cases} \tag{5}$$

and

$$\Delta\tau_{ij}(t) = \frac{\sqrt{\text{CityNum}}}{4} * \frac{Q}{L_{\text{better}}} \quad \text{Historical optimal path} \tag{6}$$

The increment of pheromone is calculated by applying formula (6) to each path segment in the optimal path of history. In the formula, CityNum is the number of grids in the map. Due to different map environments, the number of updates of the optimal path in the iterative process may be very different. Therefore, the weight setting is added to the consideration of scale to make it adapt to the environment.

**Improvement of Transfer Probability**. The transition probability is an important basis for the choice of the ant to the next arrival point. The transitional probability formula in the original method is as follows:

$$\mathrm{P}_{ij}^{k}(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^{\alpha}[\eta_{ij}(t)]^{\beta}}{\sum_{s \in T_{\text{allowed},k}} [\tau_{\text{is}}(t)]^{\alpha}[\eta_{\text{is}}(t)]^{\beta}}, & j \in T_{\text{allowed},k} \quad \text{(a)} \\ 0, & \text{else} \quad \text{(b)} \end{cases} \tag{7}$$

where $\tau_{ij}\,(t)$ denotes the amount of pheromone on the road between $i$-grid and $j$-grid at time $t$, $\eta_{ij}\,(t)$ denotes heuristic information, and the allowed table is the current set of selectable rasters. The value in original method is usually $1/d_{ij}$, which is the inverse of the distance between two grids. However, in this paper, the grids are adjacent to each other and the distance difference is small, so here we use the reciprocal of the Euclidean distance between the grid $j$ and the target grid as the heuristic information.

$$\eta_{ij}(t) = \frac{1}{d_{js}} \tag{8}$$

where $S$ is the target grid node. This allows the heuristic effect of distance length to be reflected.

Aiming at the problem of deadlock in the path planning mentioned in Sect. 2.2, we propose a probability selection strategy based on regional security information. In a grid map, each grid's neighboring grid has two possibilities, either passable or obstructed. If a grid has a majority of obstacles around it, the security of this grid area is relatively poor. Here, we use the variable black to represent the black state of the grid:

$$\text{black}_{\text{city}} = \frac{\text{AroundBlack[city]}}{\text{Around[city]}}. \tag{9}$$

In (9), AroundBlack[city] represents the number of obstacles in the grid adjacent to the current grid, and Around[city] represents the number of grids adjacent to the current grid. Therefore, black is actually the proportion of the barrier grid in the adjacent grid of the current grid. Consider the consideration of regional security factors, and the modified formula is as follows:

$$P_{ij}^{k}(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^{\alpha}[\eta_{ij}(t)]^{\beta}(1-\text{black}_j)^{\sigma}}{\sum_{s \in T_{\text{allowed},k}} [\tau_{is}(t)]^{\alpha}[\eta_{\text{is}}(t)]^{\beta}(1-\text{black}_s)^{\sigma}}, & j \in T_{\text{allowed},k} \\ 0, & \text{else} \end{cases} \tag{10}$$

where $\sigma$ is the weight of the security information. Considering that there is such a possibility that although the surrounding obstacles are in the majority, this grid is still a high possibility or even a necessary path. Therefore, the value of $\sigma$ should not be too high. In this paper, $\partial$: $\sigma = 2.5 : 1$.

### 3.2 Algorithm Description

In details, Algorithm 1 provides a pseudocode path planning for a complex environment with obstacles. In the algorithm, the obstacle grid map is generated randomly. According to the formula (1), initialize the path pheromone, update the allowed table of the ant $k$($k$ = 0, 1, 2 …, AntNum) at the current location. The algorithm calculates the probability of transfer of every feasible grid in the table through the improved probability transfer formula, and uses roulette to select the next grid. Finally, UAV moves a step of range along the planned path.

**Algorithm 1** Path Planning

**INPUT**: *map*

**OUTPUT**: *newState*，*path*

1: $rouletteWheel = 0$

2: $states =$ ant.getUnvisitedStates()

3:**for** $newState$ **in** *states* **do**

4: $rouletteWheel+=$

5: $\text{Math.pow}\big(\text{getPheromone}(state, newState), \text{getParam}('alpha')\big)$

6: $* \text{Math.pow}\big(\text{calcHeuristicValue}(state, newState), \text{getParam}('beta')\big)$

7: $* \text{Math.pow}(\text{getBlackValue}(state, newState), \text{getParam}('zeta'))$

8: **end for**

9:$randomValue =$ random()

10:$wheelPosition = 0$

11: **for** $newState$ **in** *states* **do**

12: $wheelPosition +=$

13: $\text{Math.pow}(\text{getPheromone}(state, newState), \text{getParam}('alpha'))$

14: $* \text{Math.pow}(\text{calcHeuristicValue}(state, newState), \text{getParam}('beta'))$

15: $* \text{Math.pow}(\text{getBlackValue}(state, newState), \text{getParam}('zeta'))$

16: $wheelPosition = wheelPosition \backslash rouletteWheel$

17: **if** $wheelPosition \geq randomValue$ **do**

18: **return** $newState$

19: **end for**

## 4 Simulation Results

In this section, we make experimental simulations for evaluations on the proposed algorithm. Two groups of simulation are carried out in MATLAB software. The first group is in a 20 × 20 grid map 1 and the second group is in a 40 × 40 grid map 2. For the convenience of analysis, Table 1 shows the parameters for simulations. The starting point is set to the upper left raster, and the end point is the lower right raster. In Figs. 4 and 5, the black dashed line represents the original ACO algorithm, and the red solid line represents the improved ACO algorithm.

**Table 1.** Parameters of simulation

| Parameter | Parameter specification | Value 1 | Value 2 |
|---|---|---|---|
| $m$ | Number of ants | 50 | 100 |
| $v$ | Volatile coefficient | 0.16 | 0.16 |
| *IterNum* | Number of iterations | 100 | 100 |
| $\beta$ | Weight of heuristic information. | 5.0 | 6.0 |
| $\partial$ | Weight of pheromones. | 1.5 | 1.5 |

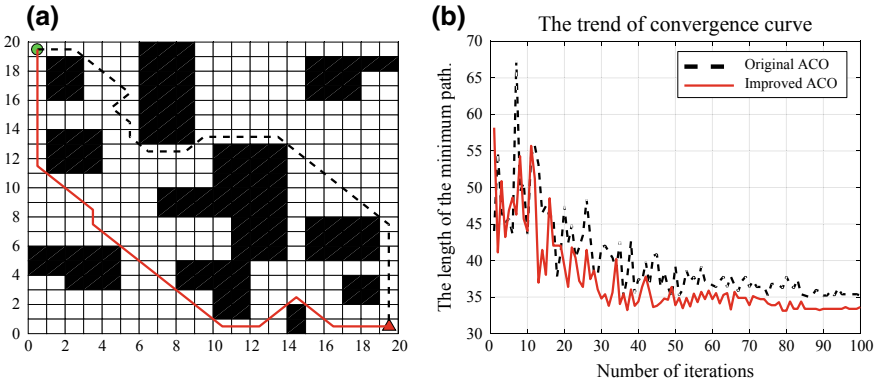



**Fig. 4.** **a** Comparison results in map 1. **b** Comparison with the original algorithm in map 1

Firstly, we chose value 1 in Table 1 for map 1. As shown in Fig. 4a, it is obvious that the path of the improved algorithm is shorter than that of the original algorithm. Moreover, the improved algorithm has fewer inflection points, so it is more suitable for actual flight scenarios. As shown in Fig. 4b, it can be observed that the improved ACO improves the convergence speed compared with the original algorithm, which shows the effectiveness of improving the convergence speed of the algorithm by optimizing the pheromone distribution strategy in the improved ACO. In Fig. 5, we chose value 2 to simulate randomly generated map with larger and more complex obstacles and obtained the same conclusion. Furthermore, we compare the results of the 50

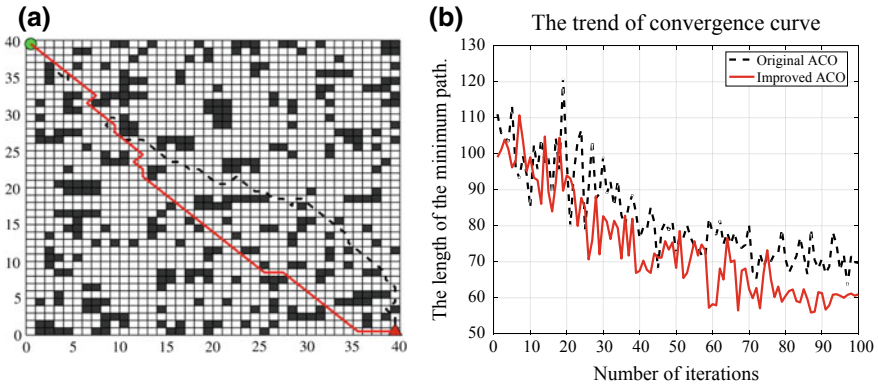



**Fig. 5.** **a** Comparisons in map 2. **b** Comparison with the original algorithm in map 2

**Table 2.** Comparison of experimental results

| Algorithm | Average time cost | Average path length | Shortest path length | Iteration times when converging | Standard deviation of path length |
|---|---|---|---|---|---|
| Original ACO | 16.54 | 35.3847 | 33.8944 | 95 | 3.024 |
| Improved ACO | 15.65 | 32.7990 | 30.9607 | 83 | 1.966 |

experiments in Map 1, and the specific results are shown in Table 2. From Table 2, it can be concluded that,

- The improvement of ACO greatly reduces the time required for path planning and verifies the effect of improving ACO on search speed.
- The path length obtained by the improved ACO is much shorter than that obtained by the original ACO. So, the effectiveness of directional guidance information optimization is verified.
- The improved ACO is more stable in many experiments.
- The improved ACO converges with fewer iterations.

In Fig. 6, the conditions are randomly distributed obstacles with density from 0.2 to 0.8. We carry out thousand experiments for calculating the success rate of path planning. It is seen that, with a small amount density of obstacles, the improved ACO can reach a probability of 98% by successfully avoiding obstacles. While, the original ACO is obviously lower than the improved method about 16 percent. With density of obstacles increasing, the improved APF method can still reach more than 80% probability by avoiding the obstacles.

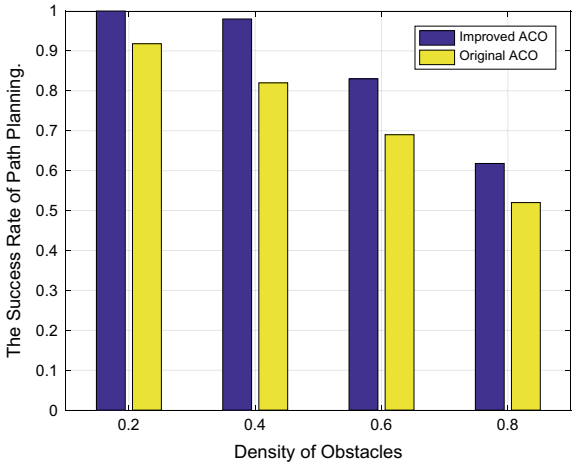



**Fig. 6.** Success rate of path planning

## 5 Conclusion

An UAV flight path planning algorithm based on ACO algorithm has been presented in this paper. By comparisons through the simulations, the proposed algorithm achieves a considerable reduction of traveling distance and improves the convergence speed by optimizing the pheromone distribution. Moreover, by adding the regional security factors to improving the probability transfer rules, it also greatly avoiding the possibility of deadlock. Simulation results verified that the method is effective for UAV path planning.

**Acknowledgments.** The authors would like to express their high appreciations to the supports from the Shenzhen Basic Research Project (JCYJ20150403161923521, JCYJ20170413110004682 and JCYJ20150403161923521).

## References

1. Khatib O. Real-time obstacle avoidance for manipulators and mobile robots. Int J Robot Res. 1986;5(1):90–8.
2. Weerakoon T, Ishii K, Nassiraei AAF. An artificial potential field based mobile robot navigation method to prevent From Deadlock. J Artif Intell Soft Comput Res. 2015;5(3):189–203.
3. Lazarowska A. Multi-criteria trajectory base path planning algorithm for a moving object in a dynamic environment. IEEE international conference on innovations in intelligent systems and applications. IEEE; 2017. p. 79–83.
4. Lin S. Computer solutions of the traveling salesman problem. Bell Labs Tech J. 2014;44 (10):2245–69.
5. Zhu QB, Zhang YL. An ant colony algorithm based on grid method for mobile robot path planning. Robot. 2005;27(2):132–6.

6. Dorigo M, Birattari M, Stutzle T. Ant colony optimization. IEEE Comput Intell Mag. 2007;1 (4):28–39.
7. Yuan M, Wang S, Li P. A model of ant colony and immune network and its application in path planning. In: IEEE conference on industrial electronics and applications. IEEE; 2008. p. 102–7.
8. Hu Y, Li D, Ding Y. A path planning algorithm based on genetic and ant colony dynamic integration. In: Intelligent control and automation. IEEE; 2015. p. 4881–6.

# Abnormal Event Detection and Localization in Visual Surveillance

Yonglin Mu and Bo Zhang(✉)

College of Information Science and Technology, Dalian Maritime University, Linghai Road 1, Dalian 116026, China
{yonglinmu, bzhang}@dlmu.edu.cn

**Abstract.** In this paper, we propose a framework for abnormal event detection and analysis in the field of visual surveillance based on the state-of-the-art deep learning techniques. We train a pair of conditional generative adversarial networks (cGANs) using the normal behavior samples, where one cGAN takes video frames as inputs and generates the corresponding optical flow features. While on the other hand, the other cGANs take optical flow features as inputs and generate the corresponding video frames. By analyzing the differences between the generated frames/optical flow features and the realistic samples, abnormal events can be detected and localized effectively. Moreover, for suspected regions, we adopt the faster RCNN to analyze the abnormal events. Experimental results demonstrate that the proposed framework can detect the abnormal events accurately and efficiently.

**Keywords:** Conditional GANs · Faster RCNN · Abnormal event detection · Visual surveillance

## 1 Introduction

With global urbanization and the rapid growth of population, detecting and analyzing abnormal events in public places are crucial in the field of visual surveillance. Despite a lot of research has been done in this area in the past 10 years [1–3], the problem is still open and far from being solved automatically. Two main challenges exist in this field: (1) Very few abnormal samples can be found in public benchmarks. This limitation constrains the applications of current machine learning methods, especially the deep learning-based models, due to the lack of training samples; (2) there is no clear definition on abnormality. For instance, it is very common to see vehicles moving fast on the highway, while cars moving fast on campus or on the squares can be regarded as abnormal. Typical normal/abnormal events in the UCSD and UMN datasets are shown in Fig. 1, where only pedestrians walking in the scenes are labeled as normal, while bicycles, vehicles, and running action appeared in the scenes are considered as abnormal events.

Suspected events always have multimodes and are difficult to model due to the limited number of training samples. The latest trend focuses on modeling merely the normal motion patterns of events using generative methods. Abnormal events can be


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_145

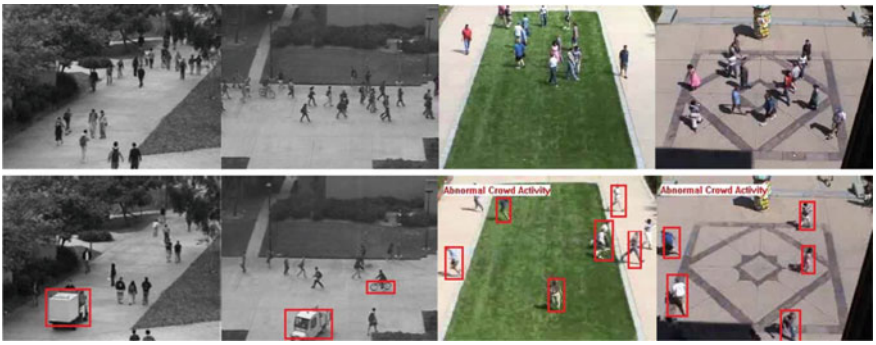



**Fig. 1.** First row: normal events; second row: abnormal events. These samples are taken from the UCSD and UMN datasets

detected as the outliers of normal event distribution. Since only normal samples are needed in the training phase, it is more feasible in realistic applications. Motivated by the fast development of generative adversarial networks (GANs), we propose a framework based on the conditional generative adversarial networks (cGANs) for abnormal events detection. Furthermore, we adopt the faster RCNN to analyze the suspicious events. The whole framework is shown in Fig. 2.

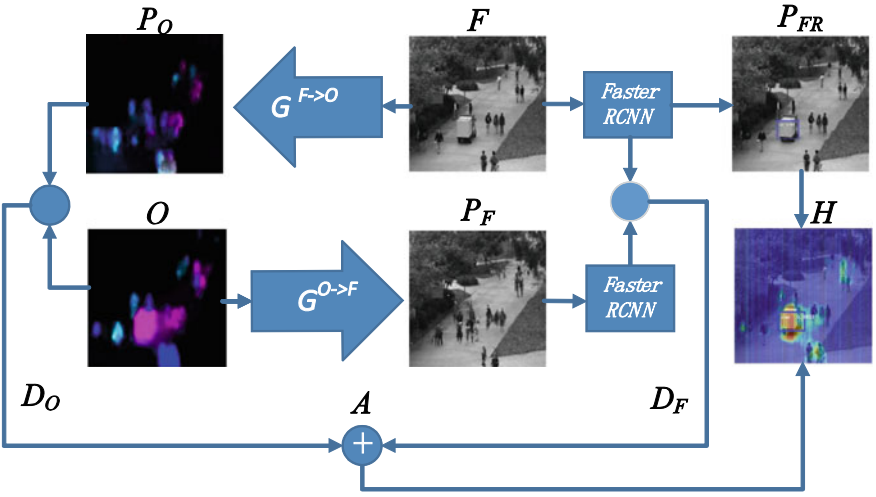



**Fig. 2.** Whole framework of abnormal events detection

GANs are originally used to generate images [4]. In this work, we propose to exploit cGANs [5] to learn the normal motion and appearance features in the training phase. Abnormal event detection is achieved by measuring the distance between the normal and abnormal distributions in terms of the motion and appearance features in

the test period. Once suspicious regions are obtained, we learn the prior knowledge of objects using the faster RCNN [6] based on the VOC2007 dataset, in order to analyze the semantic meaning of events in the scenes.

The main contributions of this paper are listed as follows:

- We propose a framework for abnormal event detection and analysis based on the state-of-the-art deep learning techniques, which does not rely on handcrafted features.
- The cGANs are exploited to learn the motion and appearance features only from normal events; thus, it is not necessary to incorporate negative samples in the training period.
- We adopt the faster RCNN to further analyze the semantic meaning of abnormal events in suspicious regions.

The rest of the paper is organized as follows: Sect. 2 reviews the related work briefly. We present the proposed method in Sect. 3 including the overall framework and the details for anomaly detection, localization, and analysis. The qualitative and quantitative results are provided in Sect. 4. We conclude our work in Sect. 5.

## 2 Related Works

In recent years, abnormal event detection has developed rapidly in the field of computer vision. Traditional approaches often rely on complex handcrafted features to represent visual information. Ionescu et al. [7] use the one-class SVM to detect abnormal events based on the spatiotemporal motion features extracted in 3-$D$ local cuboids. In [8], a space-time Markov random field model is proposed to localize abnormal events in videos. Based on the concept of proxemics [9], the so-called social force model is exploited to analyze abnormal pedestrian behaviors in crowded scenarios [1]. In [10], the histogram of oriented tracklets is used to detect abnormalities in the scenes.

With the rapid development of deep learning techniques, deep models gradually become effective tools for feature learning and representation. In [11], semantic information (learned by using the existing CNN models) and low-level optical flows are combined to measure the local abnormality in videos. In [12], a FCN-based framework is proposed to detect and localize abnormal events in crowded scenarios. Most recently, generative adversarial networks [4] and its variant [13] have shown great power in learning the latent distribution of data and generating samples. In [14], Ravanbakhsh et al. propose a generative deep learning model for abnormality detection in crowd analysis.

## 3 Proposed Method

### 3.1 Feature Learning

We use cGANs to learn the motion and appearance features from training samples. Specifically, let $F_t$ be the $t$th frame of a video and $O_t$ be the optical flow obtained using

$F_t$ and $F_{t+1}$, using the method in paper [15].$O_t$ consists of the horizontal, vertical, and the magnitude components.

We train a pair of networks: (1)$N^{O\to F}$, which generates video frames from the corresponding optical flow; and (2)$N^{F\to O}$, which generates optical flow from the corresponding frames. We adopt $G$ and $D$ to indicate the generator and discriminator, respectively. In the case of $N^{O\to F}$, $G$ takes an optical flow $O_t$ and a noise vector $z$ as the input and generates the corresponding frame $P_{Ft} = G(O_t, z)$, while $D$ evaluates the deviation of data generated by $G$ from the real distribution. Similarly, for network $N^{F\to O}$, $G$ takes a frame $F_t$ and a noise vector $z$ as the input and generates the corresponding optical flow $P_{Ot} = G(F_t, z)$. Taking $N^{F\to O}$, for example, training set $X$ consists of original video frames and the corresponding optical flow features, denoted as $X = \{F_t, O_t\}$. $G$ and $D$ are trained simultaneously using the loss $L_{cGANs}$ and loss $L_1$ as in Eqs. (1) and (2), respectively:

$$L_1(F_t, O_t) = ||O_t - G(F_t, z)||_1 \tag{1}$$

$$\begin{aligned} L_{\text{cGANs}}(G, D) &= E_{(x,y)\in X}[\log D(x, y)] \\ &+ E_{x\in\{F_t\},z\in Z}[\log(1 - D(x, (G(x, z))))] \end{aligned} \tag{2}$$

The training procedure of $N^{O\to F}$ is similar. Since only normal events are used in the training procedure, the model cannot reconstruct abnormal events in the test phase. As shown in Fig. 2, $G^{F\to O}$ takes an image $F$ as the input, which contains a vehicle moving on campus. However, in the corresponding map $P_O$, the optical flow in the area that the moving vehicle is passing by cannot reconstructed correctly. Similarly, $G^{O\to F}$ takes an optical flow $O$ as the input and generates the corresponding frame $P_F$. Comparing $F$ with $P_F$, the area where the vehicle is passing by cannot reconstructed correctly

### 3.2 Abnormality Detection and Analysis

At testing phase, we use the generators $G^{F\to O}$ and $G^{O\to F}$ to generate the reconstructed frame $P_F$ and optical flow $P_O$. Comparing the real optical flow $O$ and the generated optical flow $P_O$, we obtained $D_O = O - P_O$, where $D_O$ highlights the local differences between the real optical flow and its reconstruction. In abnormal areas, these differences are significantly higher. Similarly, we obtain $D_F = F - P_F$. However, the value of $D_F$ in abnormal regions is not that obvious; thus, we adopt the faster RCNN to obtain the differences. We use the first 13 layers of the faster RCNN to extract feature vector $F'$ and $P'_F$ from image $F$ and $P_F$ and compute $D_F = F' - P'_F$. Then, we upsample $D_F$ in order to obtain $D'_F$ with the same resolution as $D_O$. Next, for each video $V$, we compute the maximum value $M_O$ and $M_F$ of $D_O$ and $D'_F$ from all the frames.$D'_F$ and $D_O$ are normalized as in Eqs. (3) and (4):

$$N_O(i,j) = 1/M_O D_O(i,j) \tag{3}$$

$$N_F(i,j) = 1/M_F D'_F(i,j) \tag{4}$$

We create the so-called 'abnormal map' $A$ as in Eq. (5):

$$A = N_O + N_F \tag{5}$$

We adopt the faster RCNN to classify objects in the scene, which takes a frame $F$ as the input and outputs an image $P_{FR}$, where objects are labeled.$P_{FR}$ need to be upsampled in order to obtain $P'_{FR}$ with the same resolution as $A$. The labeled abnormal map $H$ is given by Eq. (6):

$$H = \alpha P'_{FR} + (1 - \alpha)A \tag{6}$$

In our experiments, $\alpha = 0.7$. Examples of typical $H$ are shown in Fig. 5, where we can not only observe the abnormal areas intuitively, but also recognize the semantic meanings of objects in the scenes so as to analyze what happened and whether the events are abnormal.

## 4 Experimental Results

### 4.1 Benchmarks and Experiment Setup

We evaluate the proposed approach on two standard benchmarks, namely the UCSD dataset and the UMN anomaly detection dataset. The UCSD dataset is divided into two parts: (1) **Ped1**, which contains 34 train and 16 test videos with the resolution of 158 by 238; and (2) **Ped2**, which contains 16 train and 12 test videos with the resolution of 240 by 360. This dataset is very challenging due to the low-resolution images. The crowd density in the walkways ranges from sparse to very dense. The UMN dataset consists of 11 videos that contain escape event in three different indoor and outdoor scenarios, with a total amount of 7700 frames. In our experiments, the video frames are resized to the resolution of 256 by 256. $N^{O\to F}$ and $N^{F\to O}$ are trained by stochastic gradient descent with momentum 0.5 and batch size 1. Each network is trained for ten epochs.

### 4.2 Results and Discussion

In our experiments, the criterion of detection is described as follows:

$$T(i,j) = \begin{cases} 1 & A(i,j) \geq \text{theta} \\ 0 & A(i,j) < \text{theta} \end{cases} \tag{7}$$

where $T(i,j) = 1$ indicates that the pixel $(i,j)$ is abnormal, while $T(i,j) = 0$ indicates that the pixel is normal, and theta is the threshold for anomaly detection. If the frame contains at least one abnormal pixel, it is labeled as abnormal. We compute the ROC curve according to different thetas in order to evaluate our approach in a quantitative way.

- **UCSD Dataset**

Quantitative results using the equal error rate (EER) and area under curve (AUC) measures are shown in Table 1, and the ROC curves are shown in Fig. 3.

**Table 1.** Comparison with other approaches on the UCSD dataset.

| Method | Ped1 | | Ped2 | |
|---|---|---|---|---|
| | EER (%) | AUC (%) | EER (%) | AUC (%) |
| MPPCA [8] | 40 | 59 | 30 | 69.3 |
| Social force (SF) [1] | 31 | 67.5 | 42 | 55.6 |
| SF + MPPCA [16] | 32 | 68.8 | 36 | 61.3 |
| SR [2] | 19 | – | – | – |
| MDT [16] | 25 | 81.8 | 25 | 82.9 |
| Detection at 150 fps [17] | 15 | 91.8 | – | – |
| AMDN (double fusion) [18] | 16 | 92.1 | 17 | 90.8 |
| *Proposed method* | ***13*** | ***92.5*** | ***17*** | ***89.1*** |

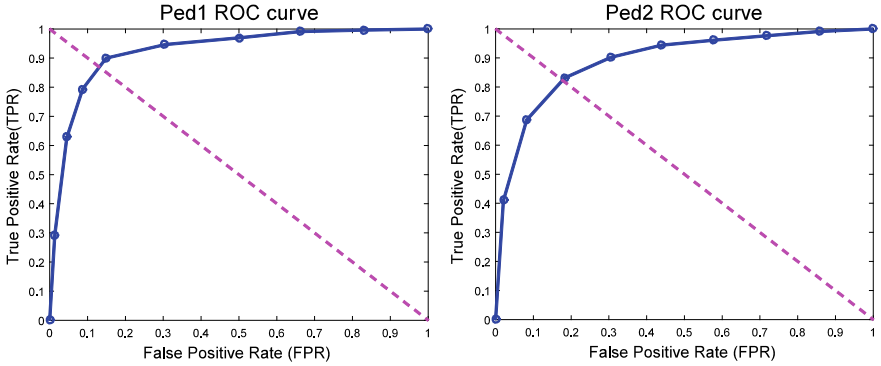



**Fig. 3.** ROC curves on the UCSD anomaly detection dataset.

- **UMN Dataset**

Quantitative results on the UMN dataset are shown in Table 2, and the ROC curve is shown in Fig. 4.

Typical examples of abnormality localization and analysis on the UCSD anomaly detection dataset are shown in Fig. 5. Our method can not only visualize abnormal regions, but also obtain the semantic labels of objects within the suspicious regions (blue boxes).

**Table 2.** Comparison with other approaches on the UMN dataset.

| Method | AUC |
|---|---|
| Optical flow [1] | 0.84 |
| SFM [1] | 0.96 |
| Sparse reconstruction [2] | 0.97 |
| Plug-and-Play CNN [11] | 0.97 |
| *Proposed method* | ***0.97*** |

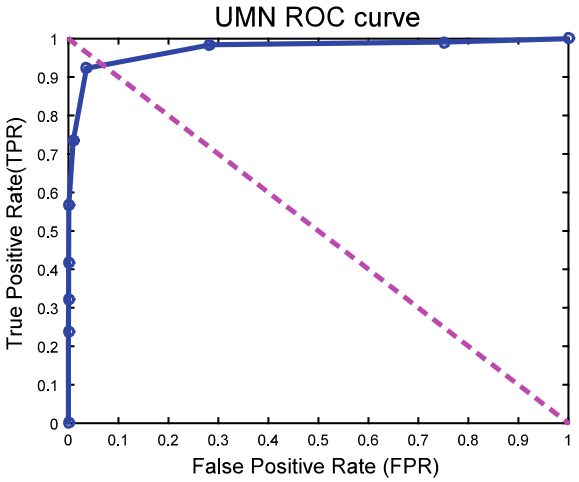



**Fig. 4.** ROC curve on the UMN anomaly detection dataset.

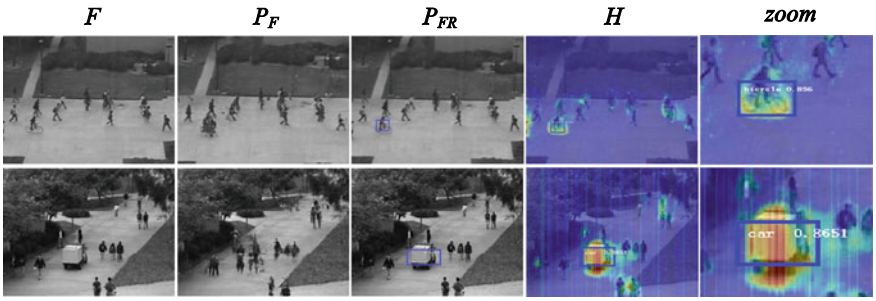



**Fig. 5.** Some examples of abnormality localization and analysis on the UCSD anomaly detection dataset.

## 5 Conclusions

In this paper, we propose a framework for abnormal event detection and analysis in crowd scenarios based on the state-of-the-art deep learning techniques. We train the cGANs using ordinary samples, which take video frames and optical flow features as inputs and generate the corresponding optical flow and frames. By analyzing the differences between the generated frames/optical flows and the realistic correspondences, the abnormal events can be detected and localized effectively. Moreover, for suspicious regions, we adopt the faster RCNN to recognize the semantic meanings of objects in the scenes so as to analyze what happened and determine whether the events are abnormal. Experimental results demonstrate the applicability of our proposed approach.

**Acknowledgments.** This work is partly supported by the National Natural Science Foundation of China (Grant No. 61702073) and the Fundamental Research Funds for the Central Universities (Grant No. 3132018190).

## References

1. Mehran R, Oyama A, Shah M. Abnormal crowd behavior detection using social force model. In: IEEE conference on computer vision and pattern recognition; 2009. p. 935–42.
2. Cong Y, Yuan J, Liu J. Sparse reconstruction cost for abnormal event detection. In: IEEE conference on computer vision and pattern recognition; 2011. p. 3449–3456.
3. Hinami R, Mei T, Shin. Joint detection and recounting of abnormal events by learning deep generic knowledge. In: IEEE international conference on computer vision; 2017. p. 3639–47.
4. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville AC, Bengio Y. Generative adversarial nets. In: International conference on neural information processing systems; 2014. p. 2672–80.
5. Isola P, Zhu J, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: IEEE conference on computer vision and pattern recognition; 2016. p. 5967–76.
6. Ren S, Girshick R, Girshick R, Sun J. Faster RCNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017;39(6):1137–49.
7. Ionescu RT, Smeureanu S, Popescu M, Alexe B. Detecting abnormal events in video using narrowed motion clusters; 2018. https://arxiv.org/abs/1801.05030.
8. Kim J, Grauman K. Observe locally, infer globally: A space-time MRF for detecting abnormal activities with incremental updates. In: IEEE Conference on Computer Vision and Pattern Recognition; 2015. 2921–7.
9. Choi W, Shahid K, Savarese S. Learning context for collective activity recognition. In: IEEE conference on computer vision and pattern recognition; 2011. p. 3273–80.
10. Mousavi H, Nabi M, Galoogahi HK, Perina A, Murino V. Abnormality detection with improved histogram of oriented tracklets. In: International conference on image analysis and processing; 2015. p. 722–32.
11. Ravanbakhsh M, Nabi M, Mousavi H, Sangineto E, Sebe N. Plug-and-play CNN for crowd motion analysis: an application in abnormal event detection. In: IEEE winter conference on applications of computer vision; 2018.

12. Sabokrou M, Fayyaz M, Fathy M, Klette R. Deep-anomaly: fully convolutional neural network for fast anomaly detection in crowded scenes; 2018. https://arxiv.org/abs/1609.00866.
13. Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. In: International Conference on Learning Representations; 2016.
14. Ravanbakhsh M, Nabi M, Sangineto E, Marcenaro L, Regazzoni C, Sebe N. Abnormal event detection in videos using generative adversarial nets. In: IEEE international conference on image processing; 2017.
15. Brox T, Malik J. Large displacement optical flow: descriptor matching in variational motion estimation. IEEE Trans Pattern Anal Mach Intell. 2011;33(3):500–13.
16. Mahadevan V, Li W, Vasconcelos N. Anomaly detection in crowded scenes. In: IEEE conference on computer vision and pattern recognition; 2010. p. 1975–81.
17. Lu C, Shi J, Jia J. Abnormal event detection at 150 FPS in MATLAB. In: IEEE international conference on computer vision; 2014. p. 2720–7.
18. Xu D, Yan Y, Ricci E, Sebe N. Detecting anomalous events in videos by learning deep representations of appearance and motion. Comput Vis Image Underst. 2016;156:117–27.

# Pseudorange Fusion Algorithm for GPS/BDS Software Receiver

Jiang Yi, Fan Yue(✉), Han Yan, and Shao Han

Information Science and Technology College, Dalian Maritime University, Dalian 116026, China
598150256@qq.com

**Abstract.** The multi-mode positioning of Global Navigation Satellite System (GNSS) could improve the positioning accuracy compared with the traditional single-mode positioning, as the number of observed satellites is increased and the geometry distribution of visual satellites is improved. In this paper, pseudorange fusion algorithm is proposed to combine the pseudorange observations of both Global Positioning System (GPS) and BeiDou Navigation System (BDS) to obtain the position equation for dual-mode positioning. Then the weighted least square method is used to solve this position equation. Besides that, the proposed pseudorange fusion algorithm is implemented in a GPS/BDS software receiver. According to positioning result comparison of single-mode and dual-mode positioning, it is concluded that the dilution of precision (DOP) of the dual-mode positioning is smaller and the positioning accuracy is more precision.

**Keywords:** Pseudorange fusion · Weighted least square method · Multi-mode positioning · Software receiver

## 1 Introduction

With the development of Global Navigation Satellite System (GNSS) technology and applications, many countries have been developing their own navigation systems. The US Global Positioning System (GPS) has been operated stably for more than 20 years and used widely in the world. The BeiDou Navigation System (BDS) which has been developing independently by China has already provided services for the Asia-Pacific region since 2012 and now is in the globalization phase. Meanwhile, dual-mode or multi-mode positioning technologies have become one of the developing trends and hot topics in GNSS receiver's algorithm research [1]. Besides that, many research groups studied GNSS software receivers based on a generic central processing unit (CPU) or an embedded platform, as software receivers could enhance its performance by replacing the internal software rather than changing the hardware structure [2].

The early research on the multi-mode GNSS positioning mainly focuses on a fusion of GPS and GLONASS or Galileo navigation system. With the development of BDS, more and more researchers began to study the multi-mode positioning with GPS and BDS. The dual-mode positioning of GPS/BDS is more reliable than that of single-mode positioning [3, 4]. In terms of software receivers, nowadays programming languages


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_146

are mainly MATLAB, C/C++, etc [2, 5, 6]. However, the MATLAB program execution is inefficient and inflexible. In contrast, the software receiver implemented in C++ is faster and more portable. Therefore, the pseudorange fusion algorithm is proposed for GPS/BDS dual-mode positioning in this paper, and a GPS/BDS combined software receiver based on real observation data is realized.

Firstly, the principle of GPS/BDS dual-mode positioning algorithm is investigated in this paper. Pseudorange fusion algorithm is used to combine the pseudorange observations of both GPS and BDS to obtain the position equation. The position of the receiver is estimated by the weighted least square method. Based on the above algorithm, a GPS/BDS software receiver based on C++ is designed and implemented. Position calculation both in single-mode and in GPS/BDS dual-mode positioning has been realized. Finally, the positioning results of GPS, BDS and GPS/BDS output from the software receiver are compared and analyzed. The results of this research support the idea that the dilution of precision (DOP) of GPS/BDS and the position accuracy of GPS/BDS are improved.

The rest of the paper is organized as follows. Section 2 investigates the pseudorange fusion algorithm in GPS/BDS dual-mode positioning technology. In Sect. 3, the proposed algorithm is implemented in the GPS/BDS dual-mode software receiver. In Sect. 4, the positioning results in different positioning mode output from the software are given and analyzed. Finally, conclusions are made in Sect. 5.

## 2 Pseudorange Fusion Algorithm

The principle of dual-mode GNSS positioning is similar to the single-mode positioning. The key technology of dual-mode positioning is the fusion algorithm. According to different fusion information, the dual-mode positioning can be divided into positioning result fusion algorithm and pseudorange fusion algorithm [7, 8]. In positioning result fusion algorithm, the position estimation from different navigation systems is obtained separately, and then the final positioning result is obtained by weighted averaging of these independent positioning results. In the pseudorange fusion algorithm, the pseudorange observation of different navigation systems is used to set up the position equation, thereby solving the final positioning result. Therefore, the former requires at least four visible satellites for each GNSS, while the latter only requires at least five visible satellites in total without restrictions for each GNSS. In terms of the number of visible satellite, pseudorange fusion algorithm is more applicable. For the GPS/BDS dual-mode receiver, signals from at least five satellites can be received generally under most circumstances.

Generally speaking, the time of arrival (TOA) technology is used in GNSS to determine a receiver's position [3]. Position equation used in single-mode positioning is shown in Eq. (1).

$$\sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} + c\delta t_u = \rho_{ci} - \varepsilon_{\rho i} \tag{1}$$

As the position of the satellite ($x_i$, $y_i$, $z_i$) can be computed according to the ephemerides, ($x_u$, $y_u$, $z_u$, $\delta t_u$) is the coordinates and clock offset of the receiver, $\rho_{ci}$ is

corrected pseudorange measurement, and $\varepsilon_{\rho i}$ is the measurement error of the pseudorange. As time reference of different GNSS systems is different, time unified problem must be considered in dual-mode GNSS positioning. Thus, the system time clock bias $\delta t_{GB}$ should be added in the GPS/BDS dual-mode position equations. The GPS/BDS dual-mode position equation contains five unknowns ($x_u$, $y_u$, $z_u$, $\delta t_u$, $\delta t_{GB}$).

Furthermore, in order to ensure that the receiver clock offset $\delta t_u$ in the position equation of each measurement is the same, all the pseudorange observations should be obtained at the same time in the receiver. Position equation of pseudorange fusion algorithm used in GPS/BDS dual-mode positioning is shown in Eq. (2).

$$\begin{cases} \sqrt{(x_{G_i}-x_u)^2+(y_{G_i}-y_u)^2+(z_{G_i}-z_u)^2}+c\delta t_u=\rho_{G_i}-\varepsilon_{\rho_{Gi}} \\ \sqrt{(x_{B_j}-x_u)^2+(y_{B_j}-y_u)^2+(z_{B_j}-z_u)^2}+c(\delta t_u+\delta t_{GB})=\rho_{B_j}-\varepsilon_{\rho_{B_j}} \end{cases} \tag{2}$$

The subscript G and B represent GPS and BDS, respectively. The term $i$ refers to the $i$th GPS satellite, and the term $j$ refers to the $j$th BDS satellite. As Eq. (2) is nonlinear, it needs to be linearized firstly before solving it. After expanding with the Taylor formula, Eq. (2) can be expressed as:

$$\boldsymbol{\delta\rho_G}=\mathbf{u_G}\cdot\mathbf{dx_0}+\boldsymbol{\varepsilon_{\rho_G}} \quad \boldsymbol{\delta\rho_B}=\mathbf{u_B}\cdot\mathbf{dx_0}+\boldsymbol{\varepsilon_{\rho_B}} \tag{3}$$

where defined the vector $\mathbf{u_G}$, $\mathbf{u_B}$ and $\mathbf{dx_0}$ as follows:

$$\begin{cases} \mathbf{u_G}\triangleq\left[\left.\frac{\partial\rho_{G_i}}{\partial x_u}\right|_{x_0},\left.\frac{\partial\rho_{G_i}}{\partial y_u}\right|_{y_0},\left.\frac{\partial\rho_{G_i}}{\partial z_u}\right|_{z_0},\left.\frac{\partial\rho_{G_i}}{\partial\delta t}\right|_{\delta t_0},\left.\frac{\partial\rho_{G_i}}{\partial t_{GB}}\right|t_{GB0}\right] \\ \mathbf{u_B}\triangleq\left[\left.\frac{\partial\rho_{B_j}}{\partial x_u}\right|_{x_0},\left.\frac{\partial\rho_{B_j}}{\partial y_u}\right|_{y_0},\left.\frac{\partial\rho_{B_j}}{\partial z_u}\right|_{z_0},\left.\frac{\partial\rho_{B_j}}{\partial\delta t}\right|_{\delta t_0},\left.\frac{\partial\rho_{B_j}}{\partial t_{GB}}\right|t_{GB0}\right] \\ \mathbf{dx_0}\triangleq[(x_u-x_0),(y_u-y_0),(z_u-z_0),(\delta t-\delta t_0),(t_{GB}-t_{GB0})]^T \end{cases} \tag{4}$$

where ($x_0$, $y_0$, $z_0$, $\delta t_0$, $t_{GB0}$) is the initial estimated value, the vector $\boldsymbol{\delta\rho_G}$, $\boldsymbol{\delta\rho_B}$ as follows:

$$\boldsymbol{\delta\rho_G}=[\,\delta\rho_{G1}\quad\ldots\quad\delta\rho_{Gm}\,]^T \quad \boldsymbol{\delta\rho_B}=[\,\delta\rho_{B1}\quad\ldots\quad\delta\rho_{Bn}\,]^T \tag{5}$$

and

$$\boldsymbol{\varepsilon_{\rho_G}}=\left[\varepsilon_{\rho_{G1}}\quad\ldots\quad\varepsilon_{\rho_{Gm}}\right]^T \qquad \boldsymbol{\varepsilon_{\rho_B}}=\left[\varepsilon_{\rho_{B1}}\quad\ldots\quad\varepsilon_{\rho_{Bn}}\right]^T \tag{6}$$

where $n$ and $m$ represent the number of BDS and GPS visual satellites, respectively. Equation (5) is generally called pseudorange residual, and the element of it can be expressed by

$$\delta\rho_{Gi}=\rho_{Gi}(x_u)-\rho_{Gi}(x_0) \quad \delta\rho_{Bj}=\rho_{Bj}(x_u)-\rho_{Bj}(x_0) \tag{7}$$

Position equation in the form of matrix of GPS/BDS dual-mode positioning is:

$$\boldsymbol{\delta\rho} = \mathbf{H}\mathbf{dx_0} + \boldsymbol{\varepsilon_\rho} \tag{8}$$

where

$$\boldsymbol{\delta}_\rho = \begin{bmatrix} \boldsymbol{\delta}_{\rho_G} & \boldsymbol{\delta}_{\rho_B} \end{bmatrix}^T \quad \boldsymbol{\varepsilon}_\rho = \begin{bmatrix} \boldsymbol{\varepsilon}_{\rho_G} & \boldsymbol{\varepsilon}_{\rho_B} \end{bmatrix}^T \quad \mathrm{H} = \begin{bmatrix} \mathbf{U}_{G1}^T & \cdots & \mathbf{U}_{Gm}^T & \mathbf{U}_{B1}^T & \cdots & \mathbf{U}_{Bn}^T \end{bmatrix}^T \tag{9}$$

The dimension of $H$ is $(n + m) \times 5$. We use the weighted least square method to solve the position equation:

$$\mathbf{dx_0} = \left(\mathbf{H^T W H}\right)^{-1} \mathbf{H^T W}\ \boldsymbol{\delta\rho} \tag{10}$$

$\mathbf{dx_0}$ is the initial estimation matrix of iteration. Matrix $\mathbf{W}$ is given by

$$\mathbf{W} = \mathrm{diag}(w_{G1} \cdots w_{Gm}\, w_{B1} \cdots w_{Bn}) \tag{11}$$

$w_{Gi}$ and $w_{Bj}$ are the weighting factor of the $i$th GPS satellite and the $j$th BDS satellite, respectively. They can be calculated based on the $\mathrm{CN_O}$ of satellite signals and the elevation angles of satellites [9]. The position of the receiver is estimated based on Eq. (10). For the $k$th iteration, the iteration process can be described as follows:

$$\begin{cases} \mathbf{dx_{k-1}} = (\mathbf{H}_{\mathbf{k-1}}^T \mathbf{W_{k-1}} \mathbf{H_{k-1}}) \mathbf{H}_{\mathbf{k-1}}^T \mathbf{W_{k-1}} \boldsymbol{\delta\rho_{k-1}} \\ \mathbf{x_k} = \mathbf{x_{k-1}} + \mathbf{dx_{k-1}} \end{cases} \tag{12}$$

## 3 Algorithm Implementation

The structure of the GPS/BDS dual-mode software receiver is shown in Fig. 1. GPS and BDS signals are received from the radio frequency (RF) front end and intermediate frequency (IF) data is processed by the baseband signal processing module. GNSS signals are acquired and tracked by the baseband processing module. After that, the navigation message data of GPS and BDS can both be decoded. The proposed pseudorange fusion algorithm is mainly implemented in the position estimation module. The function of the position mode selection submodule is selecting the positioning mode. Then the result display module could display positioning information, time information and satellite constellation.

The position mode selection submodule selects the positioning mode through the macro definition in C++. The single-mode positioning submodule is shared by GPS and BDS. It will distinguish between GPS and BDS through data interface. When GPS IF data is chosen to process, the single-mode positioning submodule is used to estimate the receiver position using GPS data. When BDS IF data is chosen to process, BDS data is used to estimate the receiver position in the single-mode positioning submodule.

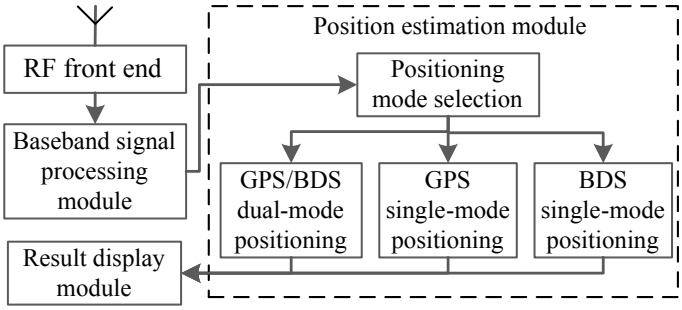



**Fig. 1.** Structure of GPS/BDS dual-mode software receiver

In the GPS/BDS dual-mode positioning submodule, GPS IF data and BDS IF data are processed simultaneously. The flow chart of the GPS/BDS dual-mode submodule is shown in Fig. 2.

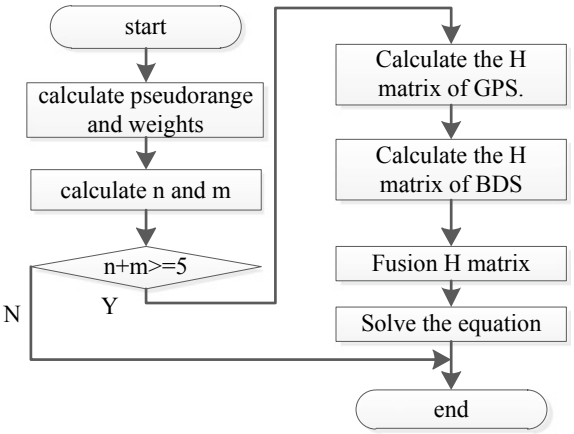



**Fig. 2.** Flow chart of GPS/BDS dual-mode positioning

Firstly, this submodule computes pseudoranges and weight factors for all tracked GPS and BDS satellites. Then the submodule determines whether the total number of tracked satellites meet the combination positioning requirement. If the condition is met, it will calculate the matrix **H** and matrix **W** given by Eqs. (9) and (11). Finally, the submodule computes the position of the receiver according to iteration process shown in Eq. (12).

## 4 Results and Analysis

The positioning experiments were carried out on the roof of Science Hall in Dalian Maritime University. With the Use of HG-SOFTGPS04 four-channel GNSS intermediate frequency signal collector, GPS and BDS signals are received simultaneously.

### 4.1 GPS Positioning Results

In the GPS single-mode positioning, the result shown by the result display module of the software receiver at a certain moment is illustrated in Fig. 3.

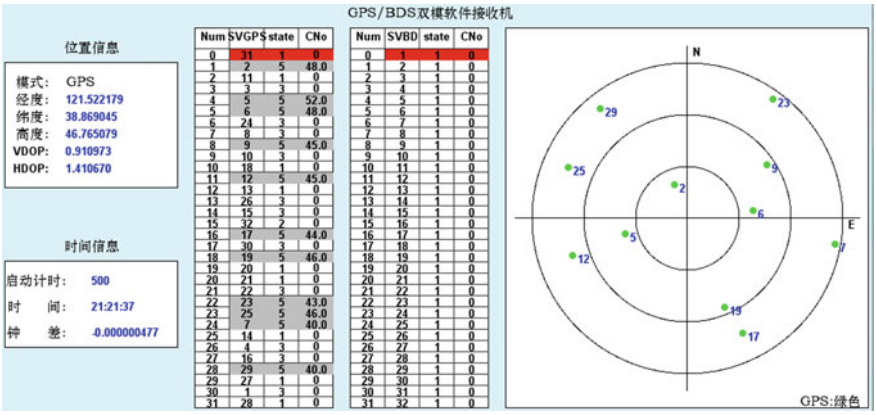



**Fig. 3.** GPS positioning results

From Fig. 3. It can be observed that a total of 10 satellites are tracked. The distribution of tracked satellites in the Earth-fixed coordinate system (ENU) is also shown. Positioning result is (121.522179°E, 38.869045°N).

### 4.2 BDS Positioning Results

In the BDS single-mode positioning, the output of the result display module of the software receiver at the same moment is shown in Fig. 4. It can be seen that the software receiver tracked a total of 12 satellites. The longitude of the receiver is 121.522141°E, and the latitude is 38.869015°N.

### 4.3 Dual-Mode Positioning Results

In GPS/BDS dual-mode positioning, the result given by the software receiver is illustrated in Fig. 5. As can be seen from Fig. 5, the software receiver tracked a total of 22 satellites. The receiver's positioning output from the software receiver is (121.522194°E, 38.869072°N).

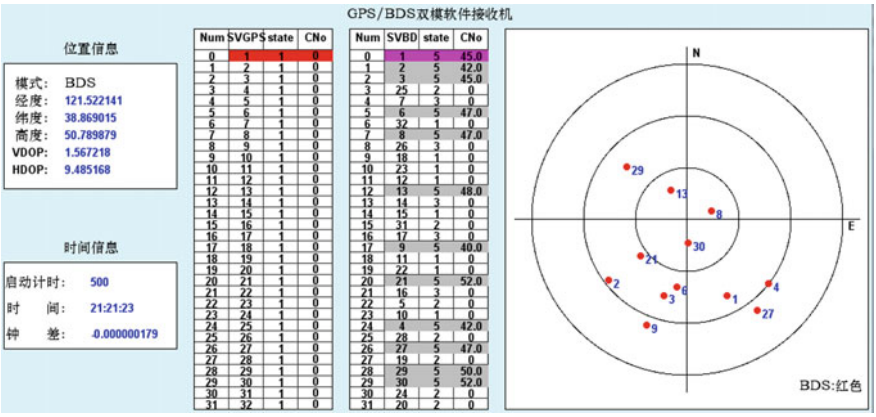



**Fig. 4.** BDS positioning results

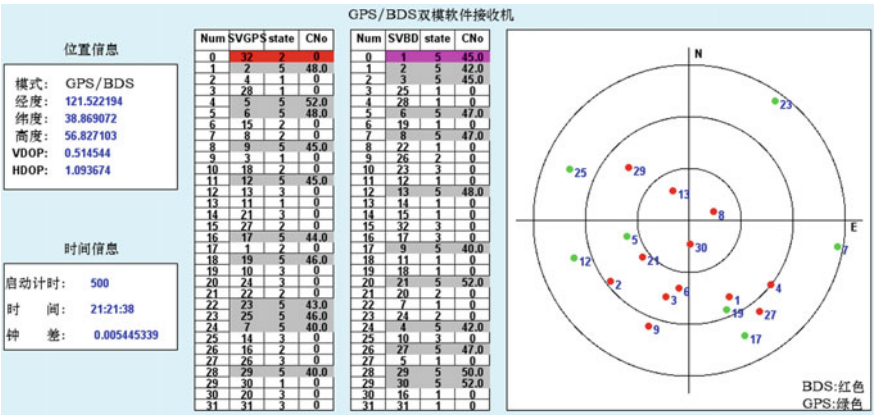



**Fig. 5.** GPS/BDS dual-mode positioning results

The DOP paramerters are defined as geometry factors that relate parameters of the user position and time bias errors to those of pseudrange errors. It only relates to the geometric distribution of visual satellites. In general, if there are more visual satellites, the corresponding DOP value will be smaller. DOP values are dimensionless. In general, they multiply the range errors corresponding to the position errors.

DOP values of GPS/BDS single-mode positioning and GPS/BDS dual-mode positioning are shown in Fig. 6. Figure 6a shows position dilution of precision (PDOP) of three kinds of positioning mode. Figure 6b shows horizontal dilution of precision (HDOP), and Figure 6c shows vertical dilution of precision (VDOP).

From Fig. 6, although the number of the visible satellites of BDS is larger than those of GPS, as the proportion of the GEO satellites in the tracked satellites is high, the satellite geometry distribution is poor. Therefore, DOP is large and the positioning

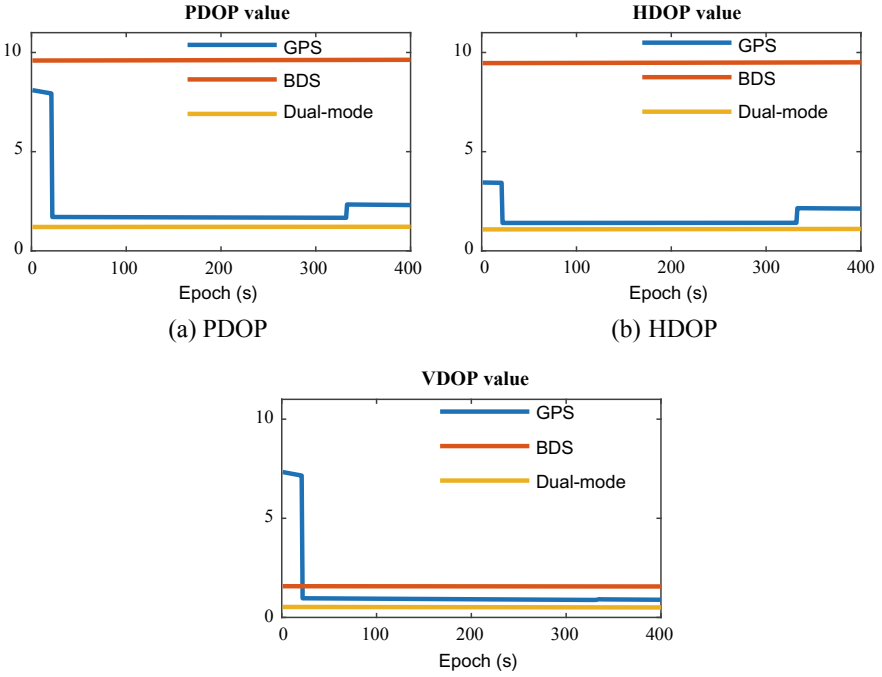



**Fig. 6.** DOP value of GPS/BDS dual-mode positioning system and single positioning system

accuracy is bad in BDS single-mode positioning. Meanwhile in the GPS single-mode positioning, the geometry distribution of the tracked satellites is better than BDS, and DOP value is relatively small. Additionally, the DOP values of GPS/BDS dual-mode positioning are apparently smaller compared to GPS/BDS single-mode positioning, which reveals that the satellite geometry distribution is further enhanced by the combination of GPS and BDS. As there are nearly 10 GPS satellites in this experiment, GPS positioning is well-geometry distribution. The DOP values of GPS/BDS dual-mode positioning are close to the DOP values of GPS positioning.

## 5 Conclusion

The pseudorange fusion algorithm is proposed to estimate the receiver's position in the GPS/BDS dual-mode positioning. In order to verify the proposed algorithm, the GPS/BDS software receiver based on C++ has been designed and implemented. This software receiver can achieve GPS and BDS single-mode positioning and BDS/GPS dual-mode positioning. Compared the position results of the dual-mode positioning output from the software receiver in the experiments, it can be concluded that the DOP of the GPS/BDS positioning mode is smaller and the position result of the GPS/BDS dual-mode positioning is more accurate.

**Acknowledgments.** This research is partially supported by the Chinese National Science Foundation (No. 61501079 and 61231006), Doctoral Scientific Research Starting Foundation of Liaoning Province (No. 2017011243 and No. 2017011149), Remote Sensing Youth Science and Technology Innovative Research, Dalian Technology Star Program, Foundation of Liaoning Educational Committee (No. L2015059) and the Fundamental Research Funds for the Central Universities (No. 3132018182 and 3132016317).

## References

1. Zhang P, Zeng Q, Zhu X, Pei L. Research progress of Global Navigation Satellite System software receiver. J Hebei Univ Sci Technol. 2016;3:220–9.
2. Juang J, Tsai C, Chen Y. Development of a PC-based software receiver for the reception of Beidou navigation satellite signals. J Navig. 2013;66:701–18.
3. Lu Y. Beidou/GPS dual-mode software receiver principle and implementation technology. Beijing: Publishing House of Electronics Industry; 2016.
4. Li X, Ge M, Dai X. Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo. J Geodesy. 2015;89(6):607–35.
5. Kim G, So H, Jeon S, Kee C, Cho Y, Choi W. The development of modularized post processing GPS software receiving platform. In: 2008 international conference on control, automation and systems, Seoul; 2008. p. 1094–8.
6. Zhang Z. Design and research of GPS software receiver and its key technologies based on VC ++. Tianjin: Nankai University; 2013.
7. Wang Q. The design and implementation positioning algorithm for GPS/BD2 dual-mode receiver. Southeast University; 2016.
8. Yu C, Qin H, Jin T. Design of weighted least squares positioning algorithm for BD2/GPS based on CNO. In: China Satellite Navigation Academic annual conference; 2015.
9. Afifi A, El-Rabbany A. Improved dual frequency PPP model using GPS and BeiDou observations. J Geodetic Sci. 2017;7(1):1–8.

# An Improved RSA Algorithm for Wireless Localization

Jiafei Fu[1(✉)], Jingyu Hua[1], Zhijiang Xu[1], Weidang Lu[1], and Jiamin Li[2]

[1] College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China
hanmaozi@hotmail.com, eehjy@163.com
[2] National Mobile Communications Research Lab, Southeast University, Nanjing 210096, China

**Abstract.** Wireless localization has become a hot issue in Internet of things, but the none-line-of-sight (NLOS) propagation will degrade the performance of traditional localization algorithms. Therefore, this paper proposed an improved range scaling algorithm (RSA) in the wireless sensor networks, where we use a two-step improvement to enhance the constrained optimization model. Simulations demonstrate that the proposed algorithm outperforms the compared algorithms, and effectively suppress the localization error caused by the none-line-of-sight propagation.

**Keywords:** Wireless location · None-line-of-sight propagation · Range scaling algorithm (RSA) · Quadratic programming

## 1 Introduction

Wireless localization had become a hot topic in recent years. Consequently, the positioning technology had become an important issue in the field of Internet of things (IoT) or communication networks [1].

In previous studies, the positioning schemes included methods based on time of arrival (TOA) [2], angle of arrival (AOA) [3], time difference of arrival (TDOA) [4], and received signal strength (RSS) [5]. However, the traditional location technology did not work effectively in practical wireless networks corrupted by NLOS error and measurement noise, in which the NLOS error acted as the major factor for the bad positioning performance [6].

There had been many studies on the elimination of NLOS effects, mainly including two kinds of methods. The first kind is to eliminate NLOS by appropriate weighting or optimization, such as approximate maximum likelihood algorithm [7], residual weighted algorithm [8], and quadratic programming algorithm [9]. The second kind is to identify NLOS BSs and then use only LOS BSs for localization; viz. the NLOS propagation is recognized [10].

This paper proposes an improved algorithm based on range scaling algorithm (RSA) [11]. This improved algorithm introduces a new constraint condition, and then,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_147

the quadratic programming is implemented in two steps. The first step is to use the original RSA algorithm to estimate a rough location and determine which area this rough position belongs to. The second step is to construct a tighter distance constraint based on this rough location region. Finally, we obtain the optimal solution according to this optimal model. Simulation results show that the proposed algorithm yields more accurate positioning performance than other compared algorithms.

## 2 Measurement Distance Model

We take into account the location algorithm based on TOA position estimation, and then, the distance between BS and MS is expressed as:

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}, \quad i = 1, \ldots, N \tag{1}$$

where $(x_i, y_i)$ and $(x, y)$ are the coordinate of the $i$th BS and the MS coordinate of MS position to be solved. If the corresponding measurement distance is $R_i$, the RSA method uses the following scaling relation:

$$r_i = \alpha_i R_i \tag{2}$$

where $\alpha_i$ represents the scaling factor, which satisfies $0 < \alpha_i < 1$ due to the influence of NLOS error.

Combining Eqs. (1) and (2), we obtain the following expression:

$$(x - x_i)^2 + (y - y_i)^2 = \alpha_i^2 R_i^2, i = 1, \ldots, N \tag{3}$$

We can define the weight vector as $\mathbf{v} = [v_1, \ldots, v_N]^{\mathrm{T}} = [\alpha_1^2, \ldots, \alpha_N^2]^{\mathrm{T}}$ to simplify the next derivations.

If the weight vector is accurately known, then the MS position estimation $(x, y)$ will be solved perfectly according to Eq. (3). Hence, how to obtain the accurate weight vector is the key issue as indicated in [11].

## 3 The Two-Step Optimization-Based RSA Algorithm

### 3.1 New Tighter Distance Constraint

From [12], we can derive the lower limit of $\mathbf{v}$, i.e.,

$$\mathbf{v}_{\min} = \{ \alpha_{1,\min}^2 \quad \cdots \quad \alpha_{N,\min}^2 \}^{\mathrm{T}} \tag{4}$$

where $\alpha_{i,\min}$ obeys,

$$\begin{cases} \alpha_{1,\min} = \max\left\{\dfrac{L_{1,2}-R_2}{R_1}, \dfrac{L_{1,3}-R_3}{R_1}, \cdots, \dfrac{L_{1,N}-R_N}{R_1}\right\} \\ \cdots \\ \alpha_{N,\min} = \max\left\{\dfrac{L_{1,N}-R_1}{R_N}\dfrac{L_{2,N}-R_2}{R_N}, \cdots, \dfrac{L_{N,N-1}-R_{N-1}}{R_N}\right\} \end{cases} \tag{5}$$

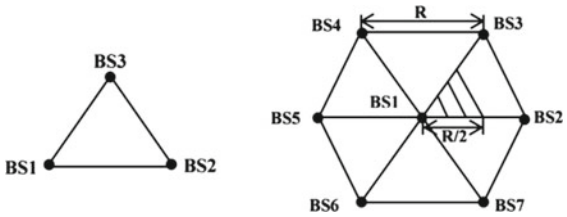
where $\max(\bullet)$ means the maximum operation, and $L_{i,j}$, $i \neq j$, refers to the distance between $i$th BS and $j$th BS.

In practice, the value of $v_i$ will not beyond one, viz. $\mathbf{v}_{\max} = [1, 1, \ldots, 1]^{\mathrm{T}}$. Hence, the original constraint in [11] is expressed as follows:

$$\mathbf{v}_{\min} \leq \mathbf{v} \leq \mathbf{v}_{\max} \tag{6}$$

In fact, the original constraint above is too loose to eliminate NLOS error enough. In order to suppress the NLOS error completely, we will refine the distance constraint next.

If MS is located in the region of {BS1, BS2, BS3} of Fig. 1a, i.e., MS is in the triangle surrounded by BS1, BS2 and BS3, and denoting that the distance between neighboring BSs is $R$, we have the inequality (7).




(a) MS in triangle region (b) MS in shade region

**Fig. 1.** Seven-BS topology

$$r_i \leq R, \quad i = 1, 2, 3 \tag{7}$$

Otherwise, the distance measurement of other BSs must obey the following constraint:

$$r_i \leq 2R, \quad i = 4, 5, 6, 7 \tag{8}$$

Moreover, we take into consideration the MS position which is within the regular hexagon as shown in Fig. 1b, and then, we obtain the tighter distance constraint

$$\begin{cases} 0 \le r_1 \le R \\ \dfrac{R}{2} \le r_i \le R, \quad i = 2, 3 \\ \dfrac{\sqrt{3}R}{2} \le r_i \le \dfrac{\sqrt{7}R}{2}, \quad i = 4, 7 \\ R \le r_i \le \dfrac{3R}{2}, \quad i = 5, 6 \end{cases} \tag{9}$$

In fact, Fig. 1b and formula (9) mean that MS belongs to a center BS (BS1 in this example). If we choose BS1 as the center base station, the inequalities above are determined. Finally, we can construct the optimal model with tighter constraints so long as the MS region is recognizable, which will be solved in Sect. 3.2.

### 3.2 The Two-Step RSA Algorithm

We utilize a two-step processing in our study. First, an original RSA is operated to find the coarse MS region, and then, the second step operated the RSA with refined distance constraint; finally, an improved localization performance can be obtained.

The optimization model of RSA at the first step is as follows:

$$\begin{aligned} &\underset{\mathbf{v}}{\text{Minimize}} \quad F(\mathbf{v}) = \sum_{k=1}^{d} \left| (x - x_k)^2 + (y - y_k)^2 \right| \\ &\text{s.t.} \quad \mathbf{v}_{\min} \le \mathbf{v} \le \mathbf{v}_{\max}, \ \text{MS} \in \text{FR} \end{aligned} \tag{10}$$

where FR denotes the feasible region and $(x_k, y_k)$ is the vertex of FR. We denote the vertex set as $\mathbf{X}_k = \{(x_k, y_k)\}_{k \in \{1,\ldots,d\}}$ with $d$ representing the number of vertexes.

Furthermore, we expand formula (3)

$$\begin{cases} v_1 R_1^2 - x_1^2 - y_1^2 = x^2 + y^2 - 2x_1 x - 2y_1 y \\ \qquad\qquad \cdots \\ v_N R_N^2 - x_N^2 - y_N^2 = x^2 + y^2 - 2x_N x - 2y_N y \end{cases} \tag{11}$$

Equation (11) can be written into a matrix form

$$\mathbf{Y} = \mathbf{A}\mathbf{x} \tag{12}$$

where $\mathbf{Y} = \begin{bmatrix} v_1 R_1^2 - x_1^2 - y_1^2 \\ v_2 R_2^2 - x_2^2 - y_2^2 \\ \cdots \\ v_N R_N^2 - x_N^2 - y_N^2 \end{bmatrix}$, $\mathrm{A} = \begin{bmatrix} -2x_1, -2y_1, x^2 + y^2 \\ -2x_2, -2y_2, x^2 + y^2 \\ \cdots \\ -2x_N, -2y_N, x^2 + y^2 \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$.

Then, we can solve the optimal solution of $\mathbf{x}$ on the basis of the least square method.

Taking $\hat{\mathbf{x}}$ as the position estimation, we rewrite the cost function of RSA as:

$$F(\mathbf{v}) = \sum_{k=1}^{d} \text{norm}(\hat{\mathbf{x}} - \mathbf{X}_k)^2 \tag{13}$$

Finally, from formula (9) and (13), we revise the new optimization model at the second step as:

$$\begin{aligned} &\underset{\mathbf{v}}{\text{Minimize}} \quad F(\mathbf{v}) \\ &\text{s.t.} \quad \begin{cases} \mathbf{v}_{\min} \le \mathbf{v} \le \mathbf{v}_{\max}, \ \text{MS} \in \text{FR} \\ \alpha_1 r_1 \le R \\ \dfrac{R}{2} \le \alpha_i r_i \le R, \ i = 2, 3 \\ \dfrac{\sqrt{3}R}{2} \le \alpha_i r_i \le \dfrac{\sqrt{7}R}{2}, \ i = 4, 7 \\ R \le \alpha_i r_i \le \dfrac{3R}{2}, \ i = 5, 6 \end{cases} \end{aligned} \tag{14}$$

## 4 Simulation and Analysis

Assume that MS is uniformly distributed in the shade of Fig. 1b with the classical seven-BS topology, i.e., the BS locates at $(0,0)$,$(R,0)$, $\left(\frac{R}{2}, \frac{\sqrt{3}R}{2}\right)$,$\left(-\frac{R}{2}, \frac{\sqrt{3}R}{2}\right)$, $(-R,0)$, $\left(-\frac{R}{2}, -\frac{\sqrt{3}R}{2}\right)$, $\left(\frac{R}{2}, -\frac{\sqrt{3}R}{2}\right)$. In our study, $R$ = 1000 m is employed. In addition, we deal with the measurement distance $R_i$ in terms of two main errors: the measurement error ($d_{\text{measure}}$) and the NLOS error ($d_{\text{NLOS}}$), i.e.,

$$R_i = r_i + d_{\text{NLOS}} + d_{\text{measure}} \tag{15}$$

where $d_{\text{NLOS}}$ is uniformly distributed in MIN and MAX, and $d_{\text{measure}}$ is a Gaussian variable with zero mean and standard deviation (SD) of 10 m. Moreover, the original RSA algorithm [11], the linear line of position (LLOP) algorithm [13], and the CLS algorithm [14] are used to compare with the proposed algorithm.

### 4.1 Influence of NLOS Error

We fixed the lower limit of NLOS at MIN = 100 m and study the effect of MAX on positioning performance.

Figure 2 shows that the performance of the proposed algorithm is improved about 12% compared to the original RSA algorithm. Meanwhile, the proposed algorithm outperforms other two comparison algorithms. Moreover, with the increase in MAX, the performance advantage is more obvious. However, the larger MAX leads to the worse accuracy.

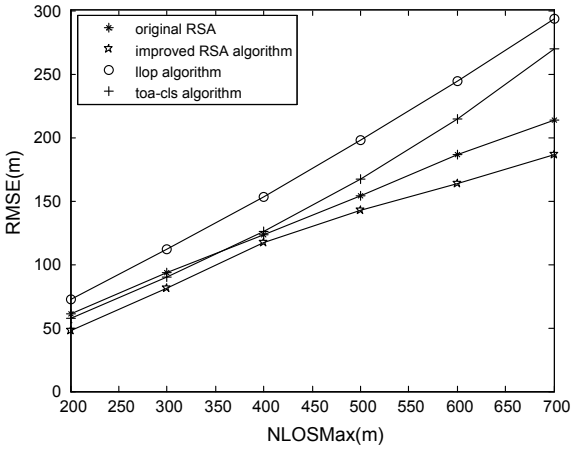
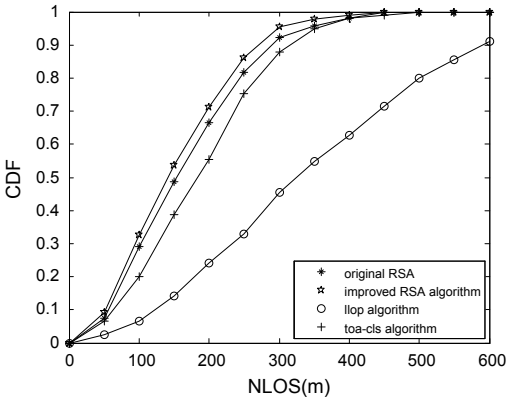



**Fig. 2.** Localization accuracy versus MAX: MIN = 200 m

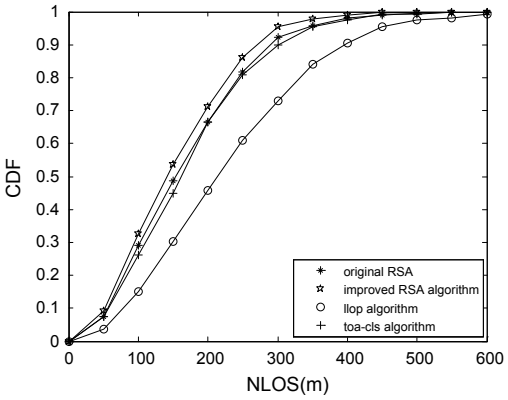
### 4.2 Effect of BS Number

Here, we observe the cumulative probability distribution function (CDF) in Fig. 3.

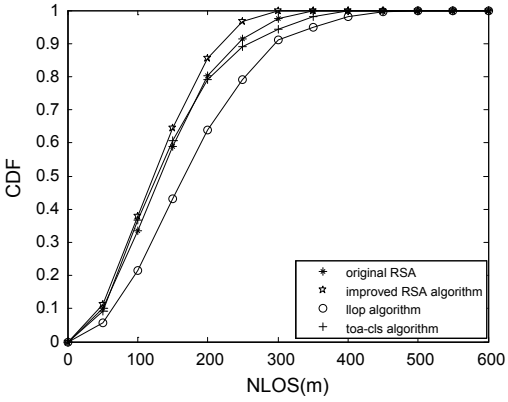
Explicitly from Fig. 3, we catch the sight of a significant improvement on positioning performance from 3BS topology to 7BS topology. In detail, the proposed algorithm outperforms other three conventional algorithms in all cases, whose CDF of 250 m error approximately equals 0.86 (5BS) and 0.96 (7BS).




(a) 3BS topology




(b) 5BS topology




(c) 7BS topology

**Fig. 3.** CDF variations versus different BS numbers

## 5 Conclusions

It is a key issue to eliminate NLOS error in wireless positioning accuracy. This paper constructs the new distance constraints on the basis of two-step RSA algorithm. Simulation results indicated that the proposed algorithm can effectively eliminate the NLOS error, and the algorithm performance is superior to the conventional algorithm.

**Acknowledgments.** This paper was sponsored by the National Natural Science Foundation of China under Grant No. 61471322.

## References

1. Spirito MA. On the accuracy of cellular mobile station location estimation. IEEE Trans Veh Technol. 2001;50(3):674–85.
2. Ke W, Wu L. Constrained least squares algorithm for TOA-based mobile location under NLOS environments. In: International conference on wireless communications, networking and mobile computing, Beijing; 2009. p. 1–4.
3. Kong F, Wang J, Zheng N, et al. A robust weighted intersection algorithm for target localization using AOA measurements. In: IEEE advanced information management, communicates, automation control conference, Xian; 2016. p. 23–8.
4. Compagnoni M, Pini A, Canclini A, et al. A geometrical–statistical approach to outlier removal for TDOA measurements. IEEE Trans Signal Process. 2017;65(15):3960–75.
5. Popleteev A. Indoor localization using ambient FM radio RSS fingerprinting: a 9-month study. In: 2017 IEEE international conference on computer and information technology (CIT), Helsinki; 2017. p. 128–34.
6. Caffery JJ, Stuber GL. Subscriber location in CDMA cellular networks. IEEE Trans Veh Technol. 1998;47(2):406–16.
7. Gazzah L, Najjar L, Besbes H. Improved hybrid AML algorithm without identification of LOS/NLOS nodes. In: IEEE Vehicular Technology Conference; 2015. p. 1–5.
8. Chen PC. A non-line-of-sight error mitigation algorithm in location estimation. In: IEEE wireless communications and networking conference; 1999. p. 316–20.
9. Zeyuan LI. Constrained weighted least squares location algorithm using received signal strength measurements. China Commun. 2016;13(4):81–8.
10. Muqaibel AH, Landolsi MA, Mahmood MN. Practical evaluation of NLOS/LOS parametric classification in UWB channels. In: IEEE international conference on communications; 2013. p. 1–6.
11. Zheng Z, Hua J, Jiang B, et al. A Novel NLOS mitigation and localization algorithm exploiting the optimization method. Chin J Sens Actuators. 2013;26(5):722–7.
12. Venkatraman S, Caffery JJ, You HR. A novel TOA location algorithm using LOS range estimation for NLoS environments. IEEE Trans Veh Technol. 2004;53(5):1515–24.
13. Zheng X, Hua J, Zheng Z, et al. LLOP localization algorithm with optimal scaling in NLOS wireless propagations. In: IEEE international conference on electronics information and emergency communication, Beijing; 2014. p. 45–8.
14. Wang X, Wang Z, O'Dea B. A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation. J China Inst Commun. 2001;52(1):112–6.

# A Quadratic Programming Localization Based on TDOA Measurement

Guangzhe Liu(✉), Jingyu Hua, Feng Li, Weidang Lu, and Zhijiang Xu

College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China
1599731854@qq.com

**Abstract.** With the popularity of smart devices, applications based on location services have been widely used, and wireless positioning technology can provide accurate positioning information. However, due to the effect of non-line-of-sight (NLOS) errors, the performance of the system can drop significantly. Accordingly, this paper introduces the theory of quadratic programming optimization based on the research of the time difference of arrival (TDOA) theory and proposes an optimization algorithm that can effectively suppress the influence of NLOS error. Simulation results show that compared with other common wireless location algorithms, the proposed algorithm has more reliable positioning accuracy under different environment models and has better system stability.

**Keywords:** Wireless localization · The time difference of arrival (TDOA) · Quadratic programming · Non-line-of-sight propagation

## 1 Introduction

In recent years, with the rapid development of wireless sensor networks, wireless positioning technology has become a research hot spot and is widely used in drone navigation, disaster rescue, and environmental monitoring [1]. The traditional localization algorithm has high accuracy when the radio wave is propagating in the ideal environment, but in the actual environment, the signal is affected by obstacles during the propagation process to generate multiple refractions and reflections. Adding non-line-of-sight (NLOS) propagation errors can severely impair the positioning accuracy of the algorithm [1]. Not only that, but the measurement error also produces negative gain.

At present, the use of wireless communication networks for location technology research has achieved a lot of results, and some advanced algorithms have been applied to production practices. The existing positioning algorithm mainly estimates the position of the mobile station through the time of arrival (TOA) [2], angle of arrival (AOA) [3], time difference of arrival (TDOA) [4], and received signal strength (RSS) [5].


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_148

In order to further reduce the impact of the NLOS error, the literature [6] uses the PDF of the adaptive kernel density estimation error and uses the Newton iteration scheme to improve the positioning accuracy. The strategy proposed in [7] is to first identify the NLOS error and mitigate it, then using the Kalman filter to estimate the location. Fascista [8, 9] proposed a new change-detection algorithm based on the generalized likelihood ratio test (GLRT) to improve the accuracy of position estimation by detecting the arrival angle of a specific signal.

Time difference of arrival (TDOA) positioning may be one of the widely used techniques for signal source positioning, and it has higher accuracy than other conventional algorithms [10, 11]. This article proposes a TDOA optimization strategy that solves the quadratic programming problem by combining the geometric relations between MN and SN and combining optimization theory. Simulations show that the proposed algorithm is obviously superior to the original TDOA algorithm, and it has advantages over other commonly used algorithms.

## 2 Original TDOA Positioning Algorithm

The TDOA algorithm is an improvement of the TOA algorithm. Assuming that there are $N$ base stations and one base station serves as a reference, the difference between the measurement distances of other base stations and the reference base station can be expressed as:

$$d_i = R_i - R_1, \quad i = 2, 3, \ldots, N \tag{1}$$

where $R_i$ is the measurement distance between the $i$th base station and the mobile station. Assuming the coordinate of the $i$th base station is $(x_i, y_i)$, the mobile station is $(x_{\text{ms}}, y_{\text{ms}})$, we have

$$R_i = \sqrt{(x_{\text{ms}} - x_i)^2 + (y_{\text{ms}} - y_i)^2} \tag{2}$$

So, according to Eqs. (1) (2) I can have

$$d_i = \sqrt{(x_{\text{ms}} - x_i)^2 + (y_{\text{ms}} - y_i)^2} - \sqrt{(x_{\text{ms}} - x_1)^2 + (y_{\text{ms}} - y_1)^2} \tag{3}$$

Writing (3) in matrix form gives

$$\mathbf{KX} = \mathbf{Y} \tag{4}$$

where $\mathbf{K} = 2\begin{bmatrix} x_2 - x_1 & y_2 - y_1 & d_2 \\ x_3 - x_1 & y_3 - y_1 & d_3 \\ \vdots & \vdots & \vdots \\ x_N - x_1 & y_N - y_1 & d_N \end{bmatrix}$, $\mathbf{Y} = \begin{bmatrix} \left[-d_2^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2\right] \\ \left[-d_3^2 + (x_3 - x_1)^2 + (y_3 - y_1)^2\right] \\ \vdots \\ \left[-d_N^2 + (x_N - x_1)^2 + (y_N - y_1)^2\right] \end{bmatrix}$

and $\mathbf{X} = [x_{\text{ms}} - x_1, y_{\text{ms}} - y_1, R_1]^T$.

Finally, according to the least squares theory, we can get the estimated value of $\mathbf{X}$,

$$\hat{\mathbf{X}} = \left(\mathbf{K}^{\mathrm{T}}\mathbf{K}\right)^{-1}\mathbf{K}^{\mathrm{T}}\mathbf{Y} \tag{5}$$

In the case of less environmental interference, the traditional TDOA algorithm has higher accuracy, but when the error is serious, the performance of the algorithm is significantly worse and is no longer applicable.

## 3 Quadratic Programming to Optimize TDOA Algorithm

### 3.1 Optimization Model

In order to alleviate the impact of errors in the actual situation, especially the NLOS error which has the most significant side effects [12, 13], we use a quadratic programming model to optimize the algorithm. Assuming there is a linear relationship between the measured distance and the actual distance, it can be expressed as:

$$R_i = \lambda_i \cdot r_i \tag{6}$$

where $r_i$ indicates the measurement distance between the $i$th base station and the mobile station. Obviously,

$$d_i = \lambda_i r_i - \lambda_1 r_1 \tag{7}$$

Combining Eqs. (4) and (8), we have

$$(x_i - x_1)(x_{\text{ms}} - x_1) + (y_i - y_1)(y_{\text{ms}} - y_1) = \frac{1}{2}\left(\lambda_1^2 r_1^2 - \lambda_i^2 r_i^2 + G\right) \tag{8}$$

where $G = (x_i - x_1)^2 + (y_i - y_1)^2$. And writing (9) in matrix form gives

$$\mathbf{AX} = \mathbf{Y} \tag{9}$$

where $\mathbf{A} = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \\ \vdots & \vdots \\ x_N - x_1 & y_N - y_1 \end{bmatrix}$, $\mathbf{Y} = \frac{1}{2}\begin{bmatrix} \lambda_1^2 r_1^2 - \lambda_2^2 r_2^2 + G \\ \lambda_1^2 r_1^2 - \lambda_3^2 r_3^2 + G \\ \\ \lambda_1^2 r_1^2 - \lambda_N^2 r_N^2 + G \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} x_{\text{ms}} - x_1 \\ y_{\text{ms}} - y_1 \end{bmatrix}$.

Therefore, we can get the LS solution of $\hat{\mathbf{X}}$, and the position estimate of the mobile station is obtained.

### 3.2 Constraint and Cost Function

In an actual cellular network, the measurement distance between the BS and MS must be greater than the actual distance due to the influence of NLOS error, i.e., $\lambda_i \subset (0, 1)$. For convenience, we define $\mathbf{v}$ as

$$\mathbf{v} = [v_1, v_2, \ldots, v_N] = \left[\lambda_1^2, \lambda_2^2, \ldots, \lambda_N^2\right] \tag{10}$$

where $\mathbf{v}_{\max} = [1, 1, \ldots, 1]$.

From [14], in the real cellular network, the lower limit of the $\lambda_i$ should satisfy the following conditions

$$\lambda_{i,\min} = \max\left\{\frac{l_{i,j} - r_j}{r_j}\middle| i, j \in [1, N], i \neq j\right\} \tag{11}$$

where max (•) represents the maximum operation. Therefore,

$$\mathrm{v}_{\min} = \left[\lambda_{1,\min}, \lambda_{2,\min}, \ldots, \lambda_{N,\min}\right] \tag{12}$$

Then, we can get the constraint of $\mathbf{v}$

$$\mathbf{v}_{\min} \leq \mathbf{v} \leq \mathbf{v}_{\max} \tag{13}$$

Second, we design the following optimization cost

$$\mathrm{F}(\mathrm{v}) = \sum_{i=2}^{N} \left| \left(\mathrm{norm}\left(\widehat{X} - \mathrm{BS}_i\right) - \mathrm{norm}\left(\widehat{X} - \mathrm{BS}_1\right)\right) - \left(\mathbf{v}_i r_i^2 - \mathbf{v}_1 r_1^2\right) \right| \tag{14}$$

where norm $(x)$ means the norm of vector $x$. When the value of $\mathbf{v}$ is closer to the real situation, it is obvious that the cost function is closer to zero.

Now the optimization model can be derived as

$$\begin{aligned} &\underset{\mathbf{v}}{\text{minimize}} \quad \mathrm{F}(\mathbf{v}) \\ &\text{subject to } \mathbf{v}_{\min} \leq \mathbf{v} \leq \mathbf{v}_{\max} \end{aligned} \tag{15}$$

Through the quadratic programming function, the final solution can be obtained by solving the above equations.

## 4 Simulation and Analysis

Assume that the reference base station coordinates are (0, 0), and other four base stations located at $\left(\frac{R}{2}, \frac{\sqrt{3}}{2}R\right)$, $\left(-\frac{R}{2}, \frac{\sqrt{3}}{2}R\right)$, $\left(-\frac{R}{2}, -\frac{\sqrt{3}}{2}R\right)$, $\left(\frac{R}{2}, -\frac{\sqrt{3}}{2}R\right)$. $R$ represents the side length of the cell, i.e., $R = 1000$ m in our study. In this paper, measurement distance errors are mainly caused by non-line-of-sight propagation ($\varepsilon_{\text{nlos}}$) and measurement error ($\delta_m$). Therefore, the distance $r_i$ can be expressed as

$$r_i = R_i + \varepsilon_{\text{nlos},i} + \delta_{m.i} \tag{16}$$

where $\varepsilon_{\text{nlos}}$ is uniformly distributed between 100 m and MAX, and $\delta_m$ is a zero-mean Gaussian with a standard deviation of 20 m.

In order to better reflect the superiority of the proposed algorithm, we compare it with the original TDOA algorithm [15], opt-LLOP [16] algorithm, and TDOA two-step maximum likelihood algorithm (TSML) [17]. We independently operate each simulation for 1000 times.

### 4.1 Influence of NLOS Error

In order to investigate the impact of NLOS error values on the performance of the algorithm, we set the value range from 100 to MAX. In this simulation, the measurement error variance is set to 20 m, and there are a total of 5 base stations. In order to show the performance of algorithm positioning intuitively, we use the root mean square error (RMSE) to represent.

Figure 1 shows that compared with the initial algorithm, the proposed algorithm has improved the performance greatly, and the optimization effect is obvious. Also, it can be seen that with the increase of NLOS error, the performance of the proposed algorithm has the best performance among these groups of algorithms. Therefore, the proposed algorithm can effectively resist the influence of NLOS errors while ensuring higher positioning accuracy.

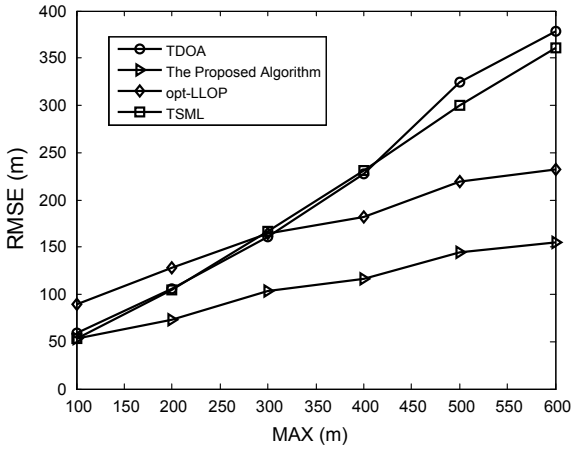
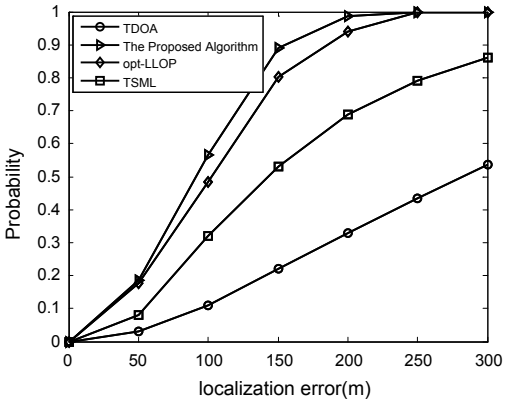
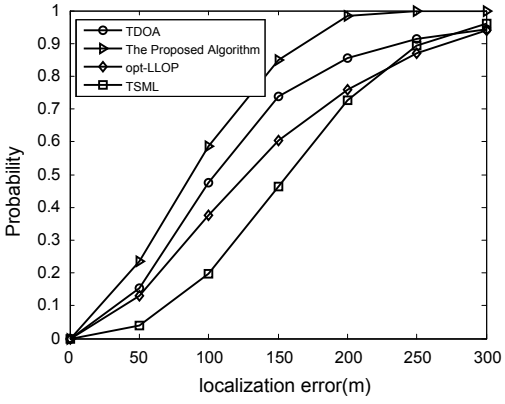


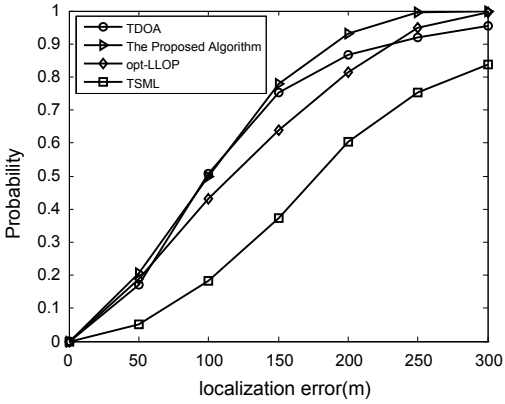

**Fig. 1.** RMSE variations versus different MAXs




(a) 4BS




(b) 5BS




(c) 6BS

**Fig. 2.** CDF of each algorithm with different BS number

### 4.2 Effect of BS Number

In order to explore the effect of the number of base stations on the performance of the algorithm, we control the NLOS error with a value of 300 m and a measurement error of 20 m, and the comparison is shown by cumulative distribution function (CDF)

Figure 2 shows the influence of BS number. It can be seen that the topology formed by the number of different base stations has different degrees of influence on the performance of these algorithms. Obviously, the proposed algorithm has the best positioning accuracy with different number of base stations.

## 5 Conclusion

In an actual wireless communication network, the most important factor affecting the positioning algorithm performance is the non-line-of-sight propagation error, so how to suppress its influence is the key to research. This paper proposes a new algorithm that combines TDOA and quadratic programming theory. Simulations demonstrate that compared with other common algorithms, the proposed algorithm has a higher ability to resist NLOS errors and better positioning accuracy.

**Acknowledgments.** This paper was sponsored by the National Natural Science Foundation of China under grant No. 61471322.

## References

1. Li H. Study of wireless sensor network applications in network optimization. Sens Transducers. 2013;157(10):180–9.
2. Vaghefi RM, Amuru SD, Buehrer RM. Improving mobile node tracking performance in NLOS environments using cooperation. In: IEEE international conference on communications; 2015. p. 6595–600.
3. Xu W, Quitin F, Leng M, et al. Distributed localization of a RF target in NLOS environments. IEEE J Sel Areas Commun. 2015;33(7):1317–30.
4. Kireev A, Fokin G, Al-odhari AHA. TOA measurement processing analysis for positioning in NLOS conditions. In: Systems of signals generating and processing in the field of on board communications; 2018. p. 1–4.
5. Wang Y, Ho KC. An asymptotically efficient estimator in closed-form for 3-D AOA localization using a sensor network. IEEE Trans Signal Process. 2015;14(12):6524–35.
6. Kim R, Ha T, Lim H, et al. TDOA localization for wireless networks with imperfect clock synchronization. In: International conference on information networking; 2014. p. 417–21.
7. Gholami MR, Vaghefi RM, Ström EG. RSS-based sensor localization in the presence of unknown channel parameters. IEEE Trans Signal Process. 2013;61(15):3752–9.
8. Feng Y, Fritsche C, Gustafsson F, et al. TOA-based robust wireless geolocation and Cramér-Rao lower bound analysis in harsh LOS/NLOS environments. IEEE Trans Signal Process. 2013;61(9):2243–55.
9. Long C, Wang Y, et al. A mobile localization strategy for wireless sensor network in NLOS conditions. China Commun. 2016;13(10):69–78.

10. Fascista A, Ciccarese G, Coluccia A, Ricci G. A change-detection approach to mobile node localization in bounded domains. In: Conference on information sciences and system; 2015. p. 1–6.
11. Martin RK, Yan C, Fan HH, et al. Algorithms and bounds for distributed TDOA-based positioning using OFDM signals. IEEE Trans Signal Process. 2011;59(3):1255–68.
12. Qi Y, Kobayashi H, Suda H. Analysis of wireless geolocation in a non-line-of-sight environment. IEEE Trans Wirel Commun. 2006;5(3):672–81.
13. Caffery JJ, Stuber GL. Subscriber location in CDMA cellular networks. IEEE Trans Veh Technol. 1998;47(2):406–16.
14. Venkatraman S, Caffery JJ, You HR. A novel TOA location algorithm using LOS range estimation for NLOS environments. IEEE Trans Veh Technol. 2004;53(5):1515–24.
15. Cheung KW, So HC, Ma WK, et al. A constrained least squares approach to mobile positioning: algorithms and optimality. EURASIP J Adv Signal Process. 2006. https://doi.org/10.1155/asp/2006/20858.
16. Zheng X, Hua J, Zheng Z, et al. LLOP localization algorithm with optimal scaling in NLOS wireless propagations. In: Proceedings of IEEE international conference on electronics information and emergency communication; 2014. p. 45–8.
17. Chan YT, Ho KC. A simple and efficient estimator for hyperbolic location. IEEE Trans Signal Process. 2002;42(8):1905–15.

# Study on Indoor Combined Positioning Method Based on TDOA and IMU

Chaochao Yang, Jianhui Chen(✉), Xiwei Guo, Deliang Liu, and Yunfei Shi

Missile Engineering Department, Army Engineering University of PLA, Shijiazhuang Campus, Shijiazhuang 050003, China
ddjqzd@126.com, 1415651643@qq.com

**Abstract.** This paper studies an indoor positioning method combining wireless sensor network (WSN) and inertial navigation system (INS). Because the positioning error of INS increases with time, the long-term positioning accuracy is poor, so the combination uses the wireless sensor network to measure the distance between unknown node and base station by TDOA method. The dead-reckoning data of inertial measurement unit (IMU) and the distance information of TDOA method are transmitted to the processing terminal, and then the particle filter algorithm is used to smooth the data to obtain the position estimation. The cumulative positioning error of INS is corrected, and the non-line-of-sight (NLOS) error in TDOA positioning method is reduced. The experimental results show that compared with the single TDOA localization method, the accuracy of the combined positioning method is higher.

**Keywords:** Indoor positioning · TDOA · IMU · Particle filter

## 1 Introduction

With the deepening of the construction of smart city, the accuracy of location service, especially indoor positioning system, is higher and higher. The importance of indoor positioning in hospitals, shopping malls, and supermarkets is self-evident. The outdoor positioning system is now very mature, such as the US GPS and China's BeiDou Navigation Satellite System (BDS), which can provide efficient and accurate outdoor location services. However, electromagnetic waves can be blocked or reflected by buildings and cannot enter the room, causing the moving objects to receive no satellite signals, so they cannot be located indoors.

At present, the commonly used indoor positioning technologies are ultrasound, Bluetooth, WIFI, radio frequency identification, ZigBEE and so on. Ultra-wideband (UWB) technology is actually a kind of bandwidth that can reach up to GHz, and the transmission rate reaches hundreds of Mbit/s, while the communication distance is about 100 m [1]. The signal bandwidth of UWB system is very wide, so the time resolution is very high, which can reach the nanosecond level, which can reach the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_149

positioning accuracy of centimeter-level, which is its biggest advantage. However, due to the influence of indoor complex environment, the UWB signal can be refracted and reflected, resulting in the appearance of the non-horizon error (NLOS error). The inertial navigation system (INS) mainly includes accelerometers and gyroscopes, which measure the acceleration $a$ and angular velocity, respectively. It belongs to calculate navigation method, that is, starting from the position $\omega$ of known points, through the continuous measurement of sports course angle and acceleration of body, the location of the below point is deduced, and body motion can be continuous measurement of the current position. However, due to the integration of the navigation information, the positioning error increases with time and the long-term accuracy is poor.

This article adopts the method of UWB and INS positioning, which uses the coordinates calculated value of UWB positioning as the main indoor localization method, uses INS correction under the condition of non-line-of-sight its positioning error, and improves the positioning accuracy of integrated positioning system.

## 2 Combination Positioning System Structure

The mobile node needs to obtain two kinds of information when realizing its positioning, the first is the measurement of inertial parameters, and the second is the estimate of the distance to the base station. For inertial information, using MEMSIC IMU380ZA—409, it was nine degrees of freedom inertial sensor board, can be measured, respectively, three axial linear acceleration and angular rate, there are three optional magnetometers, nine inertial parameters can be based [2].

Mobile node in a moment, on the basis of the optimal estimate position, mainly using UWB system orientation, under the condition of the stadia, precision can reach 5 and 16 cm [3], and under the condition of non-line-of-sight, due to reflection, diffraction and multipath effects, the positioning results will have large errors, then the IMU navigation a calculated as auxiliary positioning method. The self-navigation features of inertial navigation system can effectively compensate for the deficiency of UWB positioning system. Mobile node will be according to the features of UWB positioning signals to judge the moment is in the horizon or non-line-of-sight environments, if is stadia in UWB positioning, if non-line-of-sight, enable the INS system, because the INS positioning accuracy is very high in the short time. The structure diagram of the combined positioning system is shown in Fig. 1.

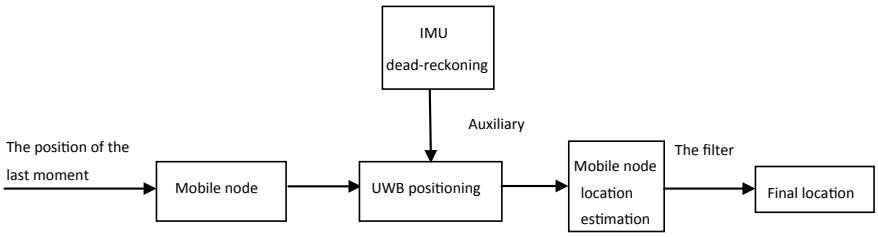



**Fig. 1.** Structure diagram of the combined positioning system

## 3 Combination Localization Algorithm

In indoor positioning, the common UWB-based location algorithm has time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA), received signal strength indicator (RSSI), etc. TOA technology through the measured signal transmission time to calculate the distance to the base station, this needs to be strictly accurate time synchronization between two communication nodes, but in the current hardware conditions it is difficult to achieve this goal; AOA ranging technology relies on installing antenna arrays on nodes to obtain angle information, requiring special hardware devices such as antenna arrays or antennas to support them. Signal strength and the space correlation are very strong, according to the intensity of waves in an indoor change model and the measured signal strength can determine the mobile end position, but most of these model parameters are estimated, and the accuracy will or changes in the environment and reduce with the increase of the scope, so the applicable scope is limited.

Different from TOA, TDOA don't need to two of the base station signal absolute time, instead of using time difference positioning, by measuring the time of arrival in the base station signal, can determine the distance of the signal source. Therefore, it only needs the clock synchronization between the base station, and the hardware requirements are low. By comparing the time difference of the signal to multiple base stations, the hyperbola with the focal length of the base station and the distance difference is a long axis, and the intersection point of the hyperbola is the position of the signal.

As shown in Fig. 2, BS1, BS2, BS3 for positioning the base station, the location is known and expressed as ($x1$, $y1$), respectively, ($x2$, $y2$), ($x3$, $y3$), to estimate the position of MS is unknown nodes ($x$, $y$), $d_i$ represents the distance between the $i$th positioning base station and the unknown node [4].

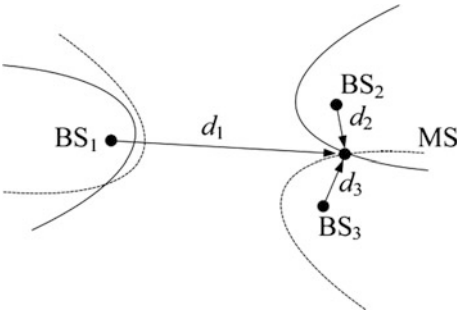



**Fig. 2.** Schematic diagram of hyperbolic rendezvous positioning

The mathematical relationship is as follows:

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} \tag{1}$$

$$d_{ij} = d_i - d_j = ct_{i,j} \tag{2}$$

$c$ represents the speed of the light, $t_{i,j}$ represents the measured value of TDOA.

$$d_{21} = d_2 - d_1 = \sqrt{(x_2 - x)^2 + (y_2 - y)^2} - \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \tag{3}$$

$$d_{31} = d_3 - d_1 = \sqrt{(x_3 - x)^2 + (y_3 - y)^2} - \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \tag{4}$$

Equation (3) can be translated into:

$$(d_{21} + d_1)^2 = d_2^2 \tag{5}$$

$$d_{21}^2 + 2d_1d_{21} = d_2^2 - d_1^2 = x_2^2 + y_2^2 - x_1^2 - y_1^2 - 2(x_2 - x_1)x - 2(y_2 - y_1)y \tag{6}$$

Making $z_i^2 = x_i^2 + y_i^2$, Eq. (6) can be translated into:

$$(x_2 - x_1)x + (y_2 - y_1)y = -d_1d_{21} + \frac{1}{2}\left(z_2^2 - z_1^2 - d_{21}^2\right) \tag{7}$$

Similarly, the Eq. (4) can be simplified to:

$$(x_3 - x_1)x + (y_3 - y_1)y = -d_1d_{31} + \frac{1}{2}\left(z_3^2 - z_1^2 - d_{31}^2\right) \tag{8}$$

Simultaneous (7) and (8) are expressed as:

$$Ax_m = Bd_1 + C \tag{9}$$

Among them

$$A = \begin{bmatrix} x_2 - x_1, y_2 - y_1 \\ x_3 - x_1, y_3 - y_1 \end{bmatrix} \qquad B = \begin{bmatrix} -d_{21} \\ -d_{31} \end{bmatrix}$$

$$C = \frac{1}{2}\begin{bmatrix} z_2^2 - z_1^2 - d_{21}^2 \\ z_3^2 - z_1^2 - d_{31}^2 \end{bmatrix} \qquad x_m = \begin{bmatrix} x \\ y \end{bmatrix}$$

The location of the test node MS is

$$x_m : x_m = A^{-1}Bd_1 + A^{-1}C \tag{10}$$

For the inertial navigation system, assuming that the moving node moves in a two-dimensional plane, the method is based on the dead-reckoning method:

$$\begin{cases} P_{dr}(k+1) = P_{dr}(k) + v_k\Delta t + \dfrac{1}{2}a_{k+1}\Delta t^2 \\ v_{k+1} = v_k + a_{k+1}\Delta t \\ \gamma_{k+1} = \gamma_k + \omega_{k+1}\Delta t \end{cases} \tag{11}$$

Among them $P_{dr}(k+1)$, $v_{k+1}$, $a_{k+1}$, $\gamma_{k+1}$, respectively, represents the position, velocity, line acceleration, and angular deviation of $k + 1$ moment.

Transform the local coordinate system to the reference frame and use the cosine formula:

$$\begin{bmatrix} a_{x_r} \\ a_{y_r} \end{bmatrix} = \begin{bmatrix} \cos\ \varphi, & -\sin\ \varphi \\ \sin\ \varphi, & \cos\ \varphi \end{bmatrix} \begin{bmatrix} a_{x_l} \\ a_{y_l} \end{bmatrix} \tag{12}$$

$\varphi = \int \omega_{z_r} \mathrm{d}t$ means the angle between the local coordinate system and the reference frame.

Particle filter is a suboptimal Bayes filter, which provides a non-gaussian environment, with appropriate weights using a set of random samples (often referred to as particles) to represent the state of the posterior distribution [5]. Basic idea of this method is an important choice probability density of random sampling, to get the corresponding weights of random sample, on the basis of state observation to adjust the weight and the size of the particle's position, and then use these sample approximation of the state of the posterior distribution, with these sample weighting and finally as a state estimate.

The state transfer model $p(s_k|s_{k-1})$ and measurement model $p(z_k|s_k)$ are introduced, where $s$ is the state variable of the system and $z$ is the measurement value of the system. For non-linear and non-Gaussian processes, the model can be expressed as:

$$\begin{aligned} s_k &= f(s_{k-1}, v_{k-1}) \\ z_k &= h(s_k, n_k) \end{aligned} \tag{13}$$

where $v_{k-1}$ and $n_k$ are independent and distributed process noise and measurement noise. Our task is to estimate unknown state variable $s_k$ according to the measured value $z_{1:k}$ of noise pollution, where $f(\ )$ and $h(\ )$ are known.

The posterior density of the system at the k moment can be approximately:

$$p(s_{0:k}|z_{1:k}) \approx \sum_{i=1}^{N_s} \omega_k^i \delta(s_{0:k} - s_{0:k}^i) \tag{14}$$

$s_k^i$ is the state value with weight $\omega_k^i$.

At this point, we get a discrete weighted estimation of the true posterior density function.

The option value is based on the importance sampling principle. In general, it's hard to sample from $p(x)$, let's say $p(x) \propto \pi(x)$. We might as well sample the density function $q(x)$ which is easy to sample, namely, $s_i \sim q(x),\ i = 1, \ldots, N_s$, $q(\ )$ is called importance density. The system obeys the first order Markov process, $q(s_k|s_{0:k-1}, z_{1:k}) = q(s_k|s_{k-1}, z_k)$ is established, and the weight is fixed as:

$$\omega_k^i \propto \omega_{k-1}^i \frac{p(z_k|s_k^i)p(s_k^i|s_{k-1}^i)}{q(s_k^i|s_{k-1}^i, z_k)} \tag{15}$$

The posterior filtering density can be approximated as:

$$p(s_k|z_{1:k}) \approx \sum_{i=1}^{N_s} \omega_k^i \delta(s_k - s_k^i) \tag{16}$$

Unfortunately, this process leads to degradation. After a series of iterations, the weight of most particles becomes very small, almost trivial, and the weight is concentrated only on a few particles. Effective sample size is introduced.

$$\text{Neff} = \frac{1}{\sum_{i=1}^{N_s} \left(\omega_k^i\right)^2} \tag{17}$$

Here $\omega_k^i$ is the normalized weight.

In order to solve the degradation problem by resampling, its basic idea is to eliminate the particles with smaller weights and to reproduce the particles with larger weights according to their weight. The specific practice is to resample $N_s$ times from an approximate discrete representation of $p(x_k|z_{1:k})$ and generate $p(x_k|z_{1:k})$ new set of particle $\{s_k^{i*}\}_{i=1}^{N_s}$, all weight of which are set to $\frac{1}{N_s}$.

Algorithm: Particle Filter

$\left[\{s_k^i, \omega_k^i\}_{i=1}^{N_s}\right] = PF\left[\{s_{k-1}^i, \omega_{k-1}^i\}_{i=1}^{N_s}, z_k\right]$

FOR $i = 1 : N_s$

- Draw $s_k^i \sim q(s_k|s_{k-1}^i, z_k)$
- Assign the particle a weight $\omega_k^i$
- END FOR
- Calculate total weight: $t = \text{SUM}[\{\omega_k^i\}_{i=1}^{N_s}]$

– FOR $i = 1 : N_s$
– Normalize: $\omega_k^i = t^{-1}\omega_k^i$
END FOR
Calculate Neff
IF Neff $< N_T$
RESAMPLE $\left[\left\{s_k^i, \omega_k^i\right\}_{i=1}^{N_s}\right]$
END IF

## 4 Algorithm Test Results

In order to evaluate and compare the algorithms we proposed, we set the experiment in a rectangular corridor 20 m long and 3 m wide, with corresponding obstacles in the corridor [6]. The six base stations are fixed around the corridor, essentially covering the path of the moving target. We then tested and recorded the estimated results of each algorithm in 48 different locations, using the cumulative distribution function (cdf) to observe the error patterns of all scenarios. Generally, as expected, the TDOA+IMU method performs better than the pure TDOA and inertial navigation methods. After the horizontal coordinate normalization processing, 0–0.5 is captured, as shown in Fig. 3.

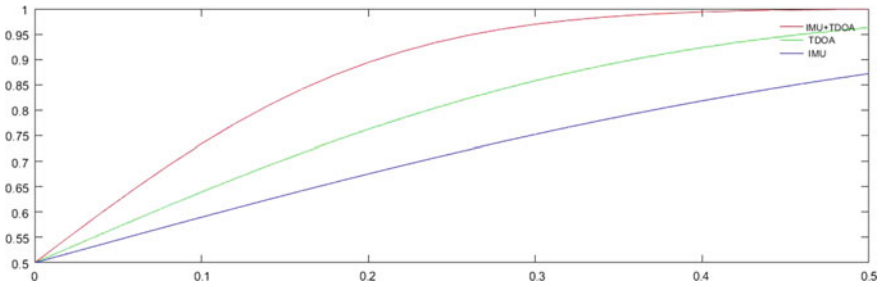



**Fig. 3.** Experimental result graph

## 5 Conclusion

In this paper, using a mobile sensor node to the environment of TDOA measurements in combination with the direction and distance from the inertial measurement unit adopts the minimum mean square estimation and maximum a posteriori estimation to get position estimation. TDOA positioning is suitable for passive positioning and does not require a signal time stamp. We make full use of the advantages of UWB positioning and inertial navigation system positioning to complement each other, reduce UWB non-line-of-sight error, and reduce the accelerometer and gyroscope drift and magnetometers, so as to improve the positioning accuracy.

According to the estimate of the unknown nodes with two stations distance through the simulation model of measurement noise, using the minimum mean square estimate and maximum a posteriori estimation algorithm, finally obtains the optimal estimation of unknown node position and suboptimal estimate position. Simulation and experiment show the effectiveness and robustness of this algorithm.

**Acknowledgments.** This work is supported by the National Natural Science Foundation of China under Grant 61601494.

## References

1. Cruz O, Ramos E, Ramírez M. 3D indoor location and navigation system based on Bluetooth. In: 2011 21st international conference on electrical communications and computers (CONIELECOMP). IEEE; 2011.
2. Duan Z, Cai Z. Adaptive particle filter for unknown fault detection of wheeled mobile robots. In: Proceedings of the 2006 IEEE/RSJ international conference on intelligent robots and systems; 2006. p. 1312–15.
3. Hol JD, Dijkstra F, Luinge H. Tightly coupled UWB/IMU pose estimation. In: 2009 IEEE international conference on ultra-wideband; 2009. p. 688–92.
4. Kok M, Hol JD, Schön TB. Indoor positioning using ultrawideband and inertial measurements. IEEE Trans Veh Technol. 2015;64:1293–303.
5. Hellmers H, Norrdine A, Blankenbach J. An IMU/magnetometer-based Indoor positioning system using Kalman filtering. In: Indoor Positioning and Indoor Navigation (IPIN); 2013. p. 1–9.
6. Megalingam RK, Rajendran AP, Dileepkumar D. LARN: implementation of automatic navigation in indoor navigation for physically challenged. In 2012 annual IEEE India conference (INDICON). IEEE; 2012.

# Research on the Fast Direction Estimation and Display Method of Far-Field Signal

Rong Liu[1], Jin Chen[1(✉)], Lei Yan[2], Ying Tong[1], Kai-kai Li[1], and Chuan-ya Wang[1]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
cjwoods@163.com
[2] Beijing Aerospace Measurement and Testing Technology Institute, Beijing, China

**Abstract.** In the array signal processing, the direction of arrival (DOA) is a very important parameter. It can be used to estimate the spatial parameter or source location of signals, so it has been deeply studied by scholars and has been widely applied in radar detection, underwater operation, and mobile communication. In this paper, an array signal DOA estimation and display system are proposed. It adopts the fast beamforming algorithm, which is an improved algorithm based on the MVDR algorithm. By comparing the output power of different positions, the angle of the arrival of the signal to the array can be obtained. The system can manually set the basic parameters, such as the number of sensors, the frequency of sampling, the parameter of the filter, the transmission speed of the sound wave, and so on. Then, the waveform characteristics of the time and frequency domain of the original acquisition signal and the filtered signal are displayed intuitively, and the real-time waveform of the signal is displayed at the same time. Through theoretical analysis and simulation experiments, we can get that the related algorithm in this system is not only suitable for microphone array, but also can meet the application requirements of radar monitoring, underwater operation, and other fields.

**Keywords:** Array signal processing · Direction of arrival (DOA) estimation · Beamforming algorithm

## 1 Introduction

The direction of arrival estimation (DOA) is a very important parameter in the spatial spectrum estimation of the array signal processing. It has been widely used in radar, communication, sonar, earthquake, astronomy, and other scientific and technological fields, which are widely used [1]. In this paper, an array signal DOA estimation and display system are proposed. First, the analog far-field signal is received, and then the DOA is estimated after filtering. In this system, the basic parameters can be set according to the actual conditions, such as the number of sensors, the spacing of the array element, the sampling frequency of the signal, the filter parameters, and so on. In


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_150

this way, many applications can be simulated, such as the acoustic signal receiving of the microphone, the radar, and the underwater operation. The system can directly display the time–frequency characteristics of the original acquisition signal, including waveform and phase. After the signal arrives, it will be filtered by the FIR filter to become a narrowband signal. The time and frequency characteristics of the filter can also be displayed in the system. The filtering effect can be easily seen through the comparison of the frequency-domain characteristics of the original signal and the filtered signal. After the basic parameters are set up, we can see the real-time waveform of the signal from a certain direction after running the system. Through theoretical analysis and simulation experiments, the feasibility of the system in the fields of microphone, radar, and underwater operation can be proved.

## 2 Signal Processing Model

In different application domains, the distance between the source and the receiving array is different. According to the size of the distance, the signal processing model can be divided into two parts: near-field model and far-field model. In the near-field model, the distance between the source and the receiving array is small, and the magnitude of the received signals is quite different. As shown in Fig. 1, the wavefront of the array at this time is the wavefront model of the sphere. The far-field model is the opposite. The distance between the source and the receiving array is far away, so the difference of the amplitude of the signal received by the array element can be ignored. At this time, the array wavefront can be regarded as the front model of the plane, as shown in Fig. 2. The signal received in this paper can be regarded as a far-field signal, and it can be considered that the amplitude of the signals received by each element is the same, which greatly reduces the complexity of the calculation.

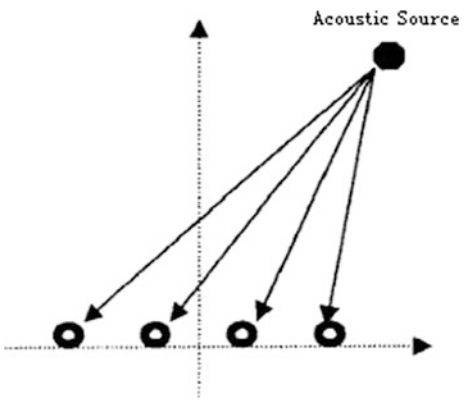



**Fig. 1.** Near-field model

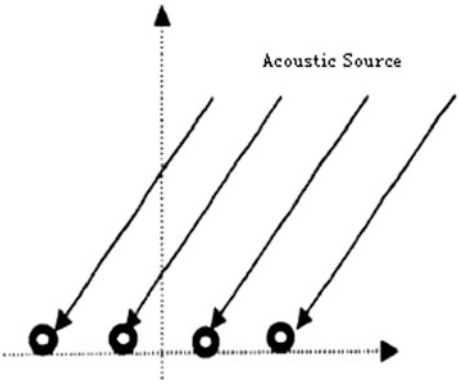



**Fig. 2.** Far-field model

Assuming that there are arrays of $M$ elements, $N$ far-field signals are received at the same time and then become $N$ far-field narrowband signals after filtering. The signal can be expressed as:

$$S_i(t) = u_i(t)\mathrm{e}^{j(w_0 t + \varphi(t))} \tag{1}$$

Then,

$$S_i(t-\tau) = u_i(t-\tau)\mathrm{e}^{j(w_0(t-\tau) + \varphi(t-\tau))} \tag{2}$$

Among them, $w_0$ is the frequency of the signal, $\varphi(t)$ is the phase of the signal, and $u_i(t)$ is the amplitude of the signal. Since the signal is far-field signal, it can be obtained as: $u_i(t-\tau) \approx u_i(t)$, and because the signal is a narrowband signal, it can be obtained as: $\varphi_i(t-\tau) \approx \varphi(t)$. Bring them into (2) and get the following:

$$S_i(t-\tau) \approx S_i(t)\mathrm{e}^{-jw_0\tau}, i = 1, 2, \ldots, N; \tag{3}$$

The reception signal of the $R$ element can be expressed as:

$$x_r(t) = \sum g_{ri} S_i(t-\tau_{ri}) + n_r(t), \quad r = 1, 2, \ldots, M; \tag{4}$$

$g_{ri}$ represents the gain of the $I$ signal received by the $R$ element; $\tau_{ri}$ represents the delay difference of the $I$ signal to the $R$ element; $n_i(t)$ $n$ represents the interference in the $R$ array element.

Therefore, we can list the signals received by $M$ elements at $t$ time, that is:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ . \\ . \\ . \\ x_M(t) \end{bmatrix} = \begin{bmatrix} e^{-jw_0\tau_{11}} & e^{-jw_0\tau_{12}} & \dots & e^{-jw_0\tau_{1N}} \\ e^{-jw_0\tau_{21}} & e^{-jw_0\tau_{22}} & \dots & e^{-jw_0\tau_{2N}} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ e^{-jw_0\tau_{M1}} & e^{-jw_0\tau_{M2}} & \dots & e^{-jw_0\tau_{MN}} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ . \\ . \\ . \\ s_N(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \\ . \\ . \\ . \\ n_M(t) \end{bmatrix}$$

If it is represented as a vector form, it can be obtained as:

$$X(t) = AS(t) + N(t) \tag{5}$$

Among them, $X(t)$ represents the data received by the matrix, $A$ represents the steering vector array, $S(t)$ represents the data emitted by the source, and $N(t)$ represents the noise parameter [2].

## 3 Far-Field Signal Arrival Estimation and Display System

The system includes the parameter setting module, the far-field signal receiving module, the filtered signal display module, the DOA estimation module, and the real-time waveform display module. The system block diagram is shown in Fig. 3. First, the basic parameters should be set up, including the receiving array parameters, filter parameters, and so on. After the operation, the system will receive the far-field signal; here, the signal can come from the analog signal source and can also come from the actual signal source. These signals will then be filtered to make them narrowband signals. Then, the direction of arrival is estimated and the real-time waveform of the filtered signal will be displayed. The direction of arrival is determined by the difference of color and luminance in different regions of the waveform.

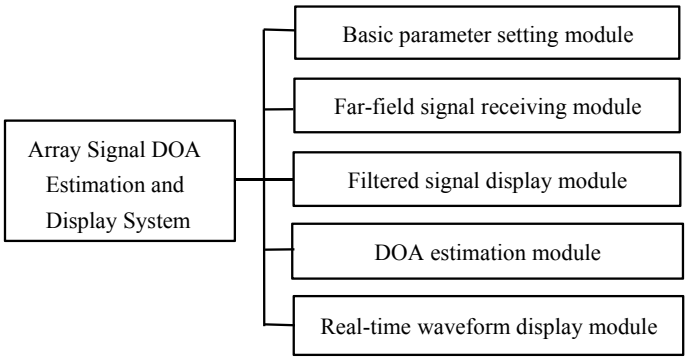



**Fig. 3.** System block diagram

### 3.1 Basic Parameter Setting Module

As shown in Fig. 4, the module is set for the basic parameters. Through this module, different parameters can be set according to various actual conditions, including setting parameters for receiving array: the number of sensors in the *X*-direction and the *Y*-direction, the spacing between the elements, the sampling frequency of the signal, and the intercepting frequency of the signal. The parameters of the FIR filter used for signal filtering, including its order, upper and lower cutoff frequencies, the gain fluctuation of the passband, and the attenuation multiple of the stopband, can also be set. In addition, it can also change the sound propagation speed $V$ and determine the wavelength of the array with the frequency $f$. Setting up different parameters can simulate a variety of working environment, so the system is theoretically applicable to the sound reception of the microphone, radar signal detection, underwater operation, and many other fields.

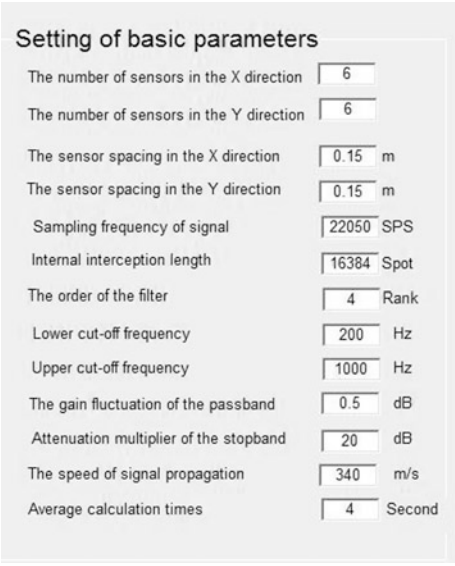



**Fig. 4.** Basic parameter setting module

### 3.2 Far-Field Signal Receiving Module

As shown in Fig. 5, it is the far-field signal receiving module, which can be analog or real signals. The parameters of the *X*-direction and the *Y*-direction are set according to the actual conditions. Through the "drawing" button, the time-domain waveform, frequency-domain phase, frequency-domain amplitude, and power density of the original far-field signal can be observed, respectively.

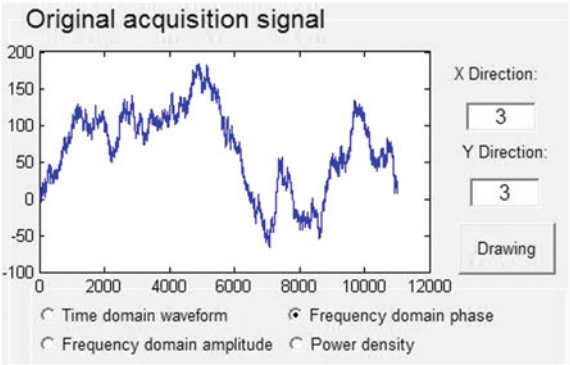



**Fig. 5.** Far-field signal receiving module

### 3.3 Filtered Signal Display Module

As shown in Fig. 6, it is the filtered signal display module. It shows the time–frequency characteristics of the narrowband signal formed by the original far-field signal filtered by the FIR filter. Similar to the far-field signal receiving module, this module can intuitively display the time-domain waveform, frequency-domain phase, frequency-domain amplitude, and power density of the filtered signal under different parameters. Compared with the original signal waveform in the frequency domain, it can clearly see its filtering effect. Compared with the original signal waveform in the frequency domain, the filtering effect can be displayed intuitively.

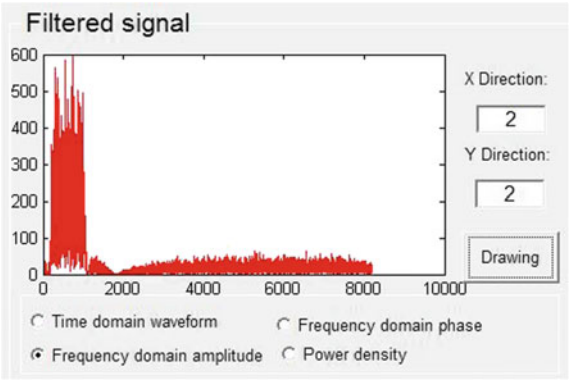



**Fig. 6.** Filtered signal display module

### 3.4 DOA Estimation Module

The module is the core module of the system. After processing the filtered data by correlation algorithm, the direction of arrival can be estimated. The beamforming algorithm is used here. The traditional beamforming algorithm is delay-and-sum

(DAS) beamforming. The distance between signal sources to various elements is different, so there will be delays, and each delay corresponds to a direction of arrival. The DAS algorithm is to apply the delay value set in advance to the output signal of each sensor and finally superimpose the output of all sensors. When the corresponding angle of the delay coincides with the direction of the signal source, the output power will reach the maximum value, thus obtaining the direction of the source [3]. The minimum variance beamforming used in this system is based on the DAS algorithm. Because in the DAS algorithm, the predetermined delay value is fixed and does not take into account the characteristics of the input signal. However, the delay in the minimum variance beamforming changes with the input signal, and finally, the output power is minimized without distortion. Finally, the power of the output signal will be represented by different brightness colors, so as to facilitate direct observation.

### 3.5 Real-Time Waveform Display Module

As shown in Fig. 7, it is the real-time waveform display module. After you have set the basic parameters, click "start," and the module will display the real-time waveform. The waveform is an upward view, where a rectangle window is intercepted. Since the signal is random, the graph is dynamic. The depth of the color and brightness represents the power of the signal in different directions. The deeper the color, the greater the signal power in the direction. By comparing the output power of various positions, the angle of signal arriving at the array can be obtained.

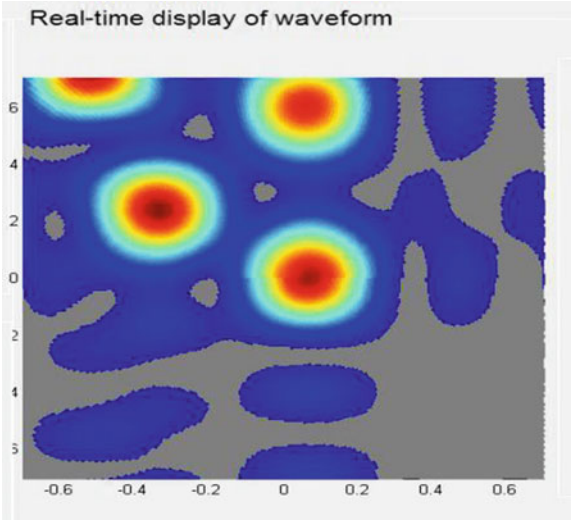



**Fig. 7.** Real-time waveform display module

### 3.6 Total Effect Diagram of the System

As shown in Fig. 8, the total effect diagram of the array signal DOA estimation and display system are presented.

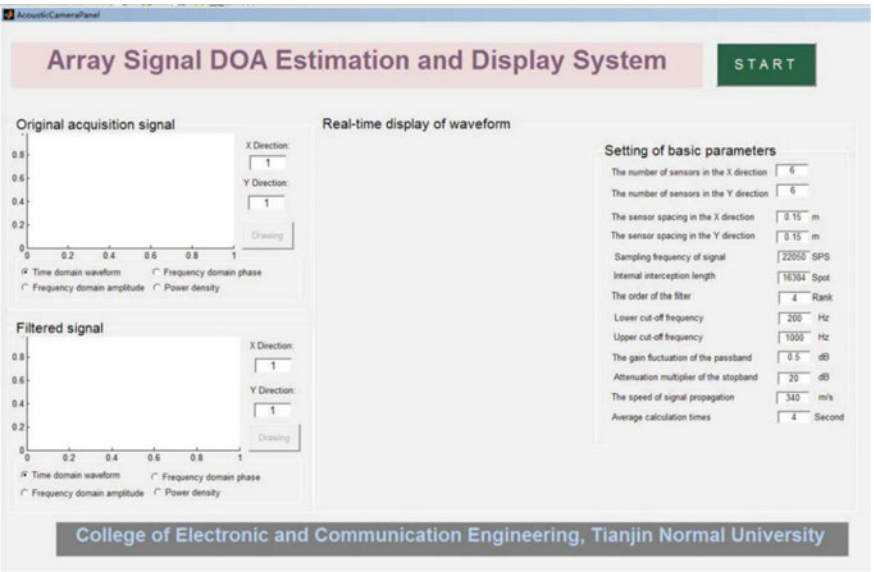



**Fig. 8.** Array signal DOA estimation and display system

## 4 Simulation

In order to verify the effectiveness of the system, the following simulation is carried out.

### 4.1 DOA Estimation Based on Acoustic Sensor

The velocity of the sound wave is $c$ = 340 m/s, where the central frequency of the wave signal is set to 1500 Hz, and the wavelength of the obtained wave is 0.23 m according to the $v = \lambda f$. So the spacing of the array element is set to 0.1 m. A two-dimensional (2D) uniform microphone array with a received array of 8 * 8 arrays is set; the sampling frequency is set as 22.05 k SPS; the filter order is set as 4; the upper and lower limits of the cutoff frequency are set as 20 and 500 Hz, respectively; the passband gain is set to 0.5 dB, and the stopband attenuation multiplier is set to 20 dB. The average number of times that the incoming wave is processed by smoothing filtering is set to 4. After the above parameters are set, the system will get the result shown in Fig. 9.

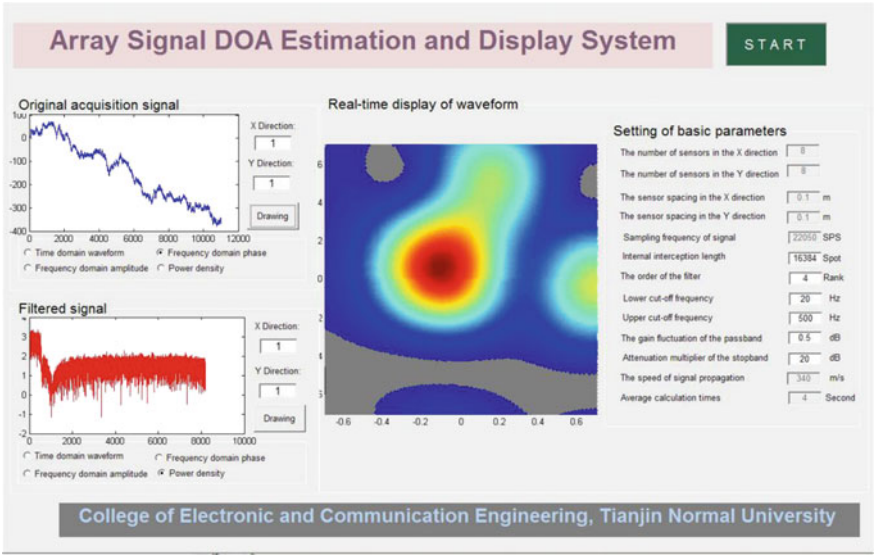



**Fig. 9.** DOA estimation based on acoustic sensor

## 4.2 DOA Estimation Based on Sonar Signal

Sonar consists mainly of active sonar and passive sonar. Active sonar mainly uses ultrasonic waves to work, and its working frequency is high, probably in the range of 3 K–97 KHz. The working frequency of passive sonar is relatively low, with a range of 3 Hz–97 KHz. The working environment of passive sonar is simulated. The center frequency of the wave is set as 1 kHz, and the transmission speed of the ultrasonic

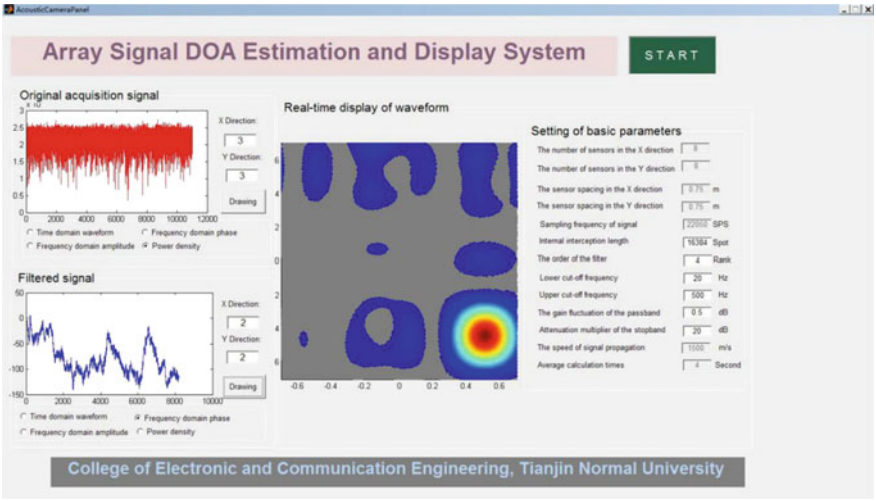



**Fig. 10.** DOA estimation based on sonar signal

wave is 1500 m/s. Thus, the wavelength of the wave is 1.5 m according to the formula $v = \lambda f$. So the spacing of the array element is 0.75 m, and the settings of other parameters are the same as those in 4.1. After setting the above parameters, the operation result is shown in Fig. 10.

## 5 Summary

This paper proposes a system which can estimate the DOA: the array signal DOA estimation and display system. The system can simulate various working environments by changing the values of basic parameters, such as acoustic signal reception of microphone array, radar, and sonar. After setting the parameters, the system can receive the analog far-field signal, and the filtered signal is estimated by DOA. Finally, the real-time waveform of the signal will be displayed. The output signal power of different positions can be judged by the difference of the color and luminance of various regions in the waveform graph, so that the direction of the incoming wave can be estimated.

Through theoretical analysis and simulation experiments, it can be proved that the system is feasible in the fields of microphone, radar, and underwater operation. However, there are still some problems that need to be further studied.

## References

1. Yan FG, Rong JJ, Liu S, Shen Y, Jin M. Joint cross-covariance matrix based fast direction of arrival estimation. Syst Eng Electron. 2017;39:1–2.
2. Raimondi FED, Farias RC, Michel OJ, Comon P. Wideband multiple diversity tensor array processing. IEEE Trans Signal Process. 2017;65:5334–45.
3. Li A, Masouros C, Sellathurai M. Analog-digital beamforming in the MU-MISO downlink by use of tunable antenna loads. IEEE Trans Veh Technol. 2018;67:3114–27.

# Compressive Sensing Approach for DOA Estimation Based on Sparse Arrays in the Presence of Mutual Coupling

Jian Zhang(✉), Zhenzhen Duan, Yang Zhang, and Jing Liang

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China
zzzjian@outlook.com

**Abstract.** In the process of direction-of-arrival (DOA) estimation, the difference co-array of sparse arrays can achieve high degrees of freedom, which can be utilized to detect more signal sources than physical sensors based on spatial smoothing (SS) algorithm. In this paper, we present a method for DOA estimation using sparse signal recovery through compressive sensing (CS) approach in the presence of mutual coupling. Compared with SS algorithm, CS approach achieves a lower estimation error. Additionally, simulation results show that the estimation error of CS approach increases with the increase of mutual coupling. Also, it increases with the increase of the grid interval of the entire DOA space.

**Keywords:** DOA estimation · Sparse arrays · Spatial smoothing · Compressing sensing · Mutual coupling

## 1 Introduction

Direction-of-arrival (DOA) estimation plays a significant role in radar and wireless communications [1,2]. It is well known that uniform linear arrays (ULAs) can only detect $N-1$ sources with $N$ sensors using traditional methods like MUSIC [3] and ESPRIT [4]. Recently, many researchers have proposed sparse arrays constructions, which can achieve higher degrees of freedom and detect more signal sources than physical sensors. In [5], minimum redundancy arrays (MRAs) are proposed, but the disadvantage of the MRA is that there is no closed-form expression to calculate the MRA. In addition, in [6,7], the authors proposed nested arrays and coprime arrays. Based on SS algorithm [6,8], coprime arrays can detect $MN+M-1$ sources with $N+2M-1$ sensors ($M$ and $N$ are coprime integers), and nested arrays can detect $N_2(N_1+1)-1$ sources with $N_1$ elements of the dense ULA part and $N_2$ elements of the sparse ULA part. However, coprime arrays have holes in the difference co-array and nested arrays have higher mutual coupling than coprime arrays. Considering the mutual coupling,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_151

Liu in [9] proposed super nested arrays, which can reduce mutual coupling by redistributing the elements of the dense ULA part of the nested array. Therefore, super nested arrays have higher estimation accuracy than nested arrays. As we know, the array aperture affects the estimation performance in the process of estimating DOA. Larger array aperture tends to have smaller estimation error [10,11]. SS algorithm [6,8] requires that the arrays are no-hole. Therefore, SS algorithm does not fully utilize the array aperture achieved in the difference co-array from the sparse arrays, and the utilizable degrees of freedom are approximately halved.

In this paper, we present sparse signal recovery through compressive sensing (CS) approach for DOA estimation in the presence of mutual coupling. The sensor arrays may not be continuous for CS approach, and CS approach fully utilizes the entire virtual array aperture and the degrees of freedom obtained in the difference co-array. Simulation results show that CS approach has a lower estimation error than that of SS algorithm.

The rest parts of this paper are organized as follows. In Sect. 2, we introduce the construction of coprime arrays, nested arrays, and super nested arrays. Then, in Sect. 3, CS approach are proposed. Section 4 shows simulation results of CS approach and SS algorithm. Conclusion and future work are provided in Sect. 5.

## 2 The Constructions of Three Categories of Sparse Arrays

Through the difference co-array of sparse arrays, we can obtain lots of virtual sensors, which can be utilized in DOA estimation. In this section, three categories of sparse arrays will be introduced.

### 2.1 Coprime Arrays

Coprime arrays can achieve lots of virtual sensors, which can be utilized in DOA estimation. Assume $M$ and $N$ are a coprime pair of positive integers with $M < N$; the locations of coprime arrays are

$$\mathbb{S} = \{0, M, 2M, \ldots, (N-1)M, N, 2N, \ldots, (2M-1)N\} \tag{1}$$

The number of physical sensors is $N + 2M - 1$, and the difference co-array is

$$\mathbb{D} = \pm\{Mn - Nm\}, 0 \le n \le N-1, 0 \le m \le 2M-1 \tag{2}$$

where $\mathbb{D}$ has $2MN + 2M - 1$ continuous elements between $-(MN + M - 1)$ and $(MN + M - 1)$. In Fig. 1a, it shows an example of coprime arrays with $M = 3$, $N = 5$, and $d = \lambda/2$. We can get the disadvantage of coprime arrays is that the difference co-array has holes.

### 2.2 Nested Arrays

Compared with coprime arrays, the difference co-array of nested arrays is no-hole. Assume there are $N_1$ elements in the dense ULA part and $N_2$ elements in the sparse ULA part. The locations of nested arrays are

$$\mathbb{S} = \{1, 2, \ldots, N_1, (N_1 + 1), 2(N_1 + 1), \ldots, N_2(N_1 + 1)\} \tag{3}$$

The number of physical sensors is $N_1 + N_2$, and the difference co-array is

$$\mathbb{D} = \{m(N_1 + 1) - n\} \tag{4}$$

where $1 \leq m \leq N_2$ and $1 \leq n \leq N_1$. We can obtain $2[(N_1 + 1)N_2 - 1] + 1$ continuous elements between $-(N_1 + 1)N_2 + 1$ and $(N_1 + 1)N_2 - 1$. In Fig. 1b, it shows an example of nested arrays with $N_1 = 5$, $N_2 = 5$, and $d = \lambda/2$. As the result of the inter-element spacing of the dense ULA part is small, the mutual coupling between sensors is very large.

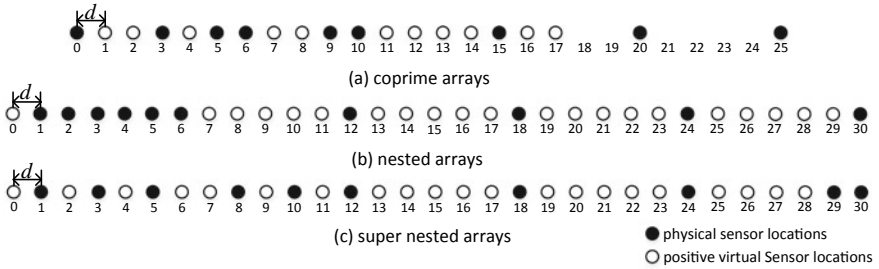



**Fig. 1.** Three categories of sparse arrays structures for physical sensors is 10

### 2.3 Super Nested Arrays

The dense ULA part of nested arrays is the main origin of the mutual coupling. In order to alleviate mutual coupling between the sensors. The authors in [9] proposed super nested arrays, which remove some sensors in the dense ULA part and relocate them, and the difference co-array of super nested arrays is no-hole. Assume $N_1$ and $N_2$ are integers satisfying $N_1 \geq 4$ and $N_2 \geq 3$. The locations of nested arrays are

$$\mathbb{S} = X_1 \cup X_2 \cup X_3 \cup X_4 \cup X_5 \cup X_6 \tag{5}$$

where

$$
\begin{aligned}
X_1 &= \{1+2\ell | 0 \le \ell \le A_1\} \\
X_2 &= \{(N_1+1)-(1+2\ell) | 0 \le \ell \le B_1\} \\
X_3 &= \{(N_1+1)+(2+2\ell) | 0 \le \ell \le A_2\} \\
X_4 &= \{2(N_1+1)-(2+2\ell) | 0 \le \ell \le B_2\} \\
X_5 &= \{\ell(N_1+1) | 2 \le \ell \le N_2\} \\
X_6 &= \{N_2(N_1+1)-1\}
\end{aligned}
$$

$$
(A_1,\ B_1,\ A_2,\ B_2) = \begin{cases} (r,\ r-1,\ r-1,\ r-2), & \text{if } N_1 = 4r, \\ (r,\ r-1,\ r-1,\ r-1), & \text{if } N_1 = 4r+1, \\ (r+1,\ r-1,\ r,\ r-2), & \text{if } N_1 = 4r+2, \\ (r,\ r,\ r,\ r-1), & \text{if } N_1 = 4r+3, \end{cases}
$$

Super nested arrays increase the spacing between physical sensors, and the difference co-array is equal to nested arrays. In Fig. 1c, it depicts an example of super nested arrays with $N_1 = 5$, $N_2 = 5$, and $d = \lambda/2$.

Based on the three categories of sparse arrays, we can utilize sparse signal recovery through CS approach to detect DOA. In the next section, CS approach is presented.

## 3 DOA Estimation with CS Approach in the Presence of Mutual Coupling

In this section, we present CS approach used in DOA estimation. Assume D narrowband sources with powers $\left[\sigma_1^2, \ldots, \sigma_D^2\right]$ impinge on the three categories of sparse arrays from direction $\theta_i, i = 1, 2, \ldots, D$. The locations of sensors are $nd$, where $n$ belongs to the set $\mathbb{S}$ and $d = \lambda/2$, where $\lambda$ is the wavelength of the signal. When mutual coupling is taken into account, the received signal of the arrays is

$$
\mathbf{x}_{\mathbb{S}}[k] = \mathbf{CAs}[k] + \mathbf{n}[k] \tag{6}
$$

where $\mathbf{C}$ is a mutual coupling matrix and can be approximated by a B-banded symmetric Toeplitz matrix in the ULA configuration [12,13] as follows:

$$
\langle \mathbf{C} \rangle_{n_1,n_2} = \begin{cases} c_{|n_1-n_2|}, & \text{if } |n_1-n_2| \le B, \\ 0, & \text{otherwise}, \end{cases} \tag{7}
$$

$n_1, n_2 \in \mathbb{S}$, and coupling coefficients $c_0, c_1, \ldots, c_B$ satisfy $1 = c_0 > |c_1| > \cdots > |c_B|$. $\mathbf{A} = [\mathbf{a}_{\mathbb{S}}(\theta_1), \mathbf{a}_{\mathbb{S}}(\theta_2), \ldots, \mathbf{a}_{\mathbb{S}}(\theta_D)]$ denotes the array manifold matrix, and $\mathbf{a}_{\mathbb{S}}(\theta_i) = e^{j(2\pi/\lambda)nd\sin\theta_i}$, $n \in \mathbb{S}$, $\mathbf{s}[k] = [s_1[k], s_2[k], \ldots, s_D[k]]$ is the source signal vector. The $\mathbf{n}[k]$ is assumed to be temporally and spatially white noise, which is also uncorrelated from the sources. The covariance matrix of $\mathbf{x}_{\mathbb{S}}[k]$ is

$$
\mathbf{R}_{\mathbb{S}} = \sum_{i=1}^{D} \sigma_i^2 \mathbf{a}_{\mathbb{S}}(\theta_i) \mathbf{a}_{\mathbb{S}}^H(\theta_i) + \sigma^2 \mathbf{I} \tag{8}
$$

where $\mathbf{a}_{\mathbb{S}}(\theta_i)\mathbf{a}_{\mathbb{S}}^H(\theta_i)$ is $e^{j\pi(n_1-n_2)\sin\theta_i}$, $n_1, n_2 \in \mathbb{S}$. Vectorizing $\mathbf{R}_{\mathbb{S}}$ yields

$$\mathbf{z} = \mathrm{vec}(\mathbf{R}_{\mathbb{S}}) = \mathbf{B}\mathbf{r} + \sigma_n^2\mathbf{I}_n \tag{9}$$

where $\mathbf{B} = [\mathbf{a}_{\mathbb{D}}(\theta_1), \ldots, \mathbf{a}_{\mathbb{D}}(\theta_D)]$, $\mathbf{a}_{\mathbb{D}}(\theta_i) = e^{j\pi n\sin\theta_i}$, $n \in \mathbb{D}$, $\mathbf{r} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_D^2]$, and $\mathbf{I}_n = \mathrm{vec}(\mathbf{I})$. $\mathbf{z}$ is received signal of the virtual sensors arrays. The virtual source signal becomes a single snapshot of $\mathbf{r}$, and the rank of $\mathbf{z}$ is 1. Therefore, the authors in [6,7] proposed SS algorithm. However, SS algorithm sacrifices the array aperture and halves the utilizable degrees of freedom. In this paper, we present CS approach to estimate DOA based on the three categories of sparse arrays in the presence of mutual coupling. The objective function of CS approach is defined as

$$\hat{\mathbf{r}} = \arg\min_{\tilde{\mathbf{r}}} \|\tilde{\mathbf{r}}\|_1 \quad \text{s.t.} \left\|\mathbf{z} - \tilde{\mathbf{B}}\tilde{\mathbf{r}}\right\|_2 < \beta \tag{10}$$

We define $\beta$ as follows:

$$\beta = \sqrt{\frac{\left(\mathbf{x}_L^2 - \|\mathbf{x}_{\mathbb{D}}\|_2\right)^2}{L}} \tag{11}$$

where $\mathbf{x}_L$ is the maximum in $\mathbf{x}_{\mathbb{D}}$ and $L$ is the number of virtual sensors. $\tilde{\mathbf{B}} = \left[\mathbf{a}_{\mathbb{D}}\left(\tilde{\theta}_1\right), \ldots, \mathbf{a}_{\mathbb{D}}\left(\tilde{\theta}_Q\right)\right]$, $\tilde{\theta} = \left[\tilde{\theta}_1, \ldots, \tilde{\theta}_Q\right]$ denotes the entire DOA space region, $Q \gg D$. $\tilde{\mathbf{r}} = [r_1, r_2, \ldots, r_Q]^T$ is a D-sparse vector and $\tilde{\mathbf{r}}$ is expressed as

$$\tilde{\mathbf{r}} = \begin{cases} \sigma_i^2, \tilde{\theta}_i \in [\theta_1, \theta_2, \ldots, \theta_D] \\ 0, \tilde{\theta}_i \notin [\theta_1, \theta_2, \ldots, \theta_D] \end{cases} \tag{12}$$

We can solve (10) by using the convex optimization toolbox CVX [14].

## 4 Simulations and Analysis

In this section, we will compare the estimation error of CS approach and SS algorithm based on the three categories of sparse arrays in the presence of mutual coupling. In the simulation, we assume there are $D = 7$ uncorrelated sources coming from $[-45°, -35°, -25°, 0°, 10°, 20°, 40°]$. The number of the physical sensors is 10, and the physical sensors' locations of the three types sparse arrays are shown in Fig. 1. The root-mean-squared error (RMSE) is defined as (13). Monte Carlo simulation is 5000, and the grid interval $\theta^{\mathrm{g}}$ is 1°. The mutual coupling is based on (7) with $c_1 = 0.3e^{j\pi/3}$, $B = 50$, and $c_\ell = c_1 e^{-j(\ell-1)\pi/4}/\ell$.

$$RMSE = \sqrt{\sum_{i=1}^{D}\left(\hat{\theta}_i - \theta_i\right)^2 \Big/ D} \tag{13}$$

Figure 2 shows RMSE of CS approach and SS algorithm for coprime arrays, nested arrays, and super nested arrays versus SNR. The red lines represent

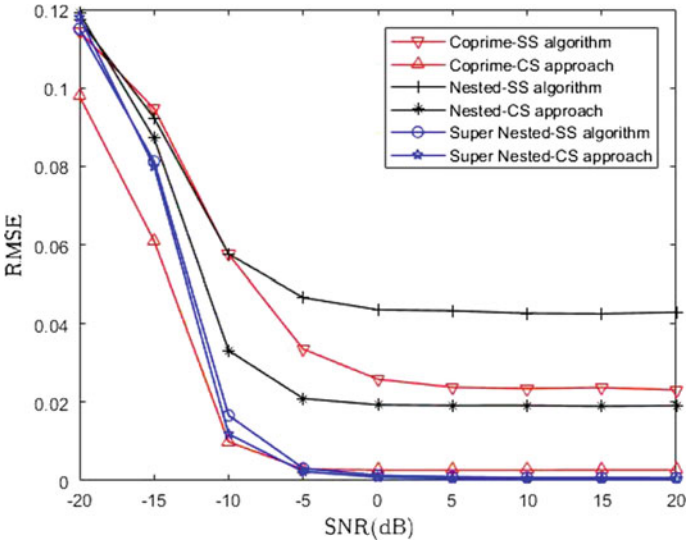



**Fig. 2.** RMSE of CS approach and SS algorithm in the presence of mutual coupling versus SNR; the number of snapshots is 300

coprime arrays, the black lines represent nested arrays, and the blue lines represent super nested arrays. It shows that the estimation error of CS approach is lower than that of SS algorithm for the three types sparse arrays, because CS approach utilizes all virtual aperture. With the increase of SNR, the estimation error of super nested arrays is the smallest in the presence of mutual coupling, and nested arrays is worst. Figure 3 shows RMSE of CS approach and SS algorithm versus the number of snapshots. We can draw the conclusion that RMSE decreases as the number of snapshots increases and the estimation error of CS approach is lower than that of SS algorithm.

RMSE of CS approach and SS algorithm versus the parameter $c(1)$ of the mutual coupling is plotted in Fig. 4. We choose coprime arrays as an example. It is quite obvious that the estimation performance heavily depends on the mutual coupling. RMSE is smallest if $c1$ is close to 0. The estimation error starts to increase dramatically for CS approach when $c1$ is larger than 0.24, and $c1$ is larger than 0.16 for SS algorithm. This phenomenon indicates that CS approach has a lower estimation error than that of SS algorithm in the presence of mutual coupling. Figure 5 shows RMSE of CS approach versus the grid interval $\theta^{\mathrm{g}}$. RMSE appears a rising trend with the increase of the grid interval $\theta^{\mathrm{g}}$. In order to get a lower estimation error, we trend to choose a smaller $\theta^{\mathrm{g}}$. However, the time needed to do one complete simulation will increase.

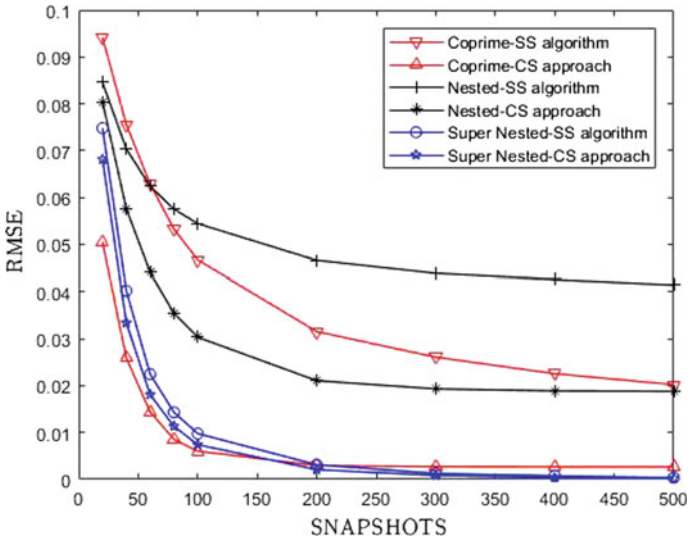



**Fig. 3.** RMSE of CS approach and SS algorithm in the presence of mutual coupling versus the number of snapshots, $SNR = 0$ dB

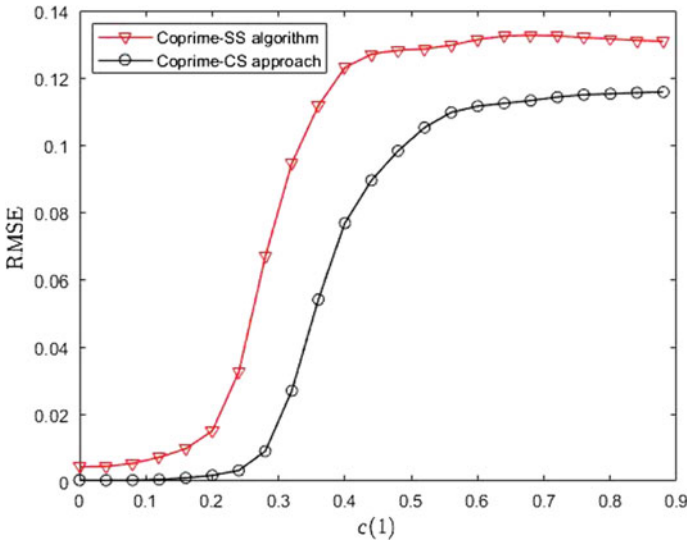



**Fig. 4.** RMSE of CS approach and SS algorithm for coprime arrays versus the parameter $c(1)$; the number of snapshots is 300 and $SNR = 0$ dB

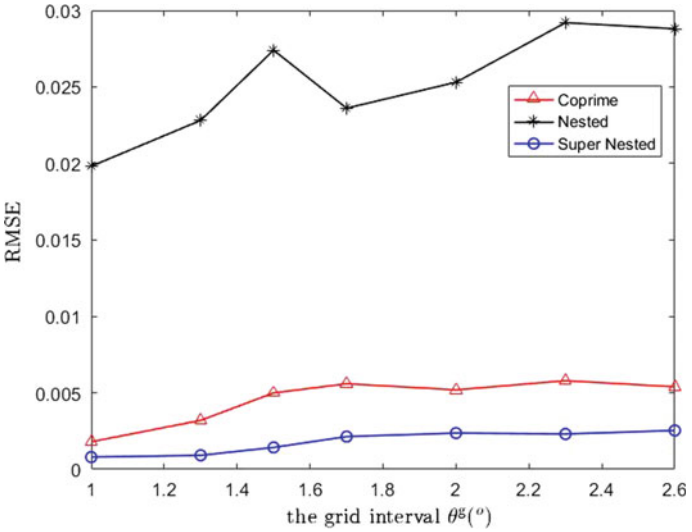



**Fig. 5.** RMSE of CS approach versus the grid interval $\theta^g$, the number of snapshots is 300 and $SNR = 0$ dB

## 5 Conclusion and Future Work

This paper utilizes three categories of sparse arrays and proposes a method for DOA estimation using sparse signal recovery through CS approach in the presence of mutual coupling. Compared with SS algorithm, CS approach fully uses all the virtual array aperture, which is extended based on the three categories of sparse arrays structures. The estimation error of CS approach is lower than that of SS algorithm. In addition, we discuss the effect of mutual coupling parameter $c(1)$ on the estimation error. The threshold of $c(1)$ for CS approach is lower than that of SS algorithm. Therefore, CS approach has the advantage over SS algorithm. Besides, the estimation error of CS approach presents a rising trend with the increase of the grid interval $\theta^g$. In the future, we may apply the sensor nodes moving algorithm to the three categories of sparse arrays structures.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61671138, 61731006) and was partly supported by the 111 Project No. B17008.

## References

1. Van Trees HL. Optimum array processing: part IV of detection, estimation and modulation theory. Publishing House of Elec; 2002.
2. Godara LC. Application of antenna arrays to mobile communications. ii. Beamforming and direction-of-arrival considerations. Proc IEEE. 2009;85(8):1195–245.

3. Schmidt R. Multiple emitter location and signal parameter estimation. IEEE Trans Antennas Propag. 1986;34(3):276–80.
4. Roy R, Kailath T. Esprit-estimation of signal parameters via rotational invariance techniques. IEEE Trans Acoust Speech Signal Process. 2002;37(7):984–95.
5. Moffet A. Minimum-redundancy linear arrays. IEEE Trans Antennas Propag. 2003;16(2):172–5.
6. Pal P, Vaidyanathan PP. Nested arrays: a novel approach to array processing with enhanced degrees of freedom. IEEE Trans Signal Process. 2010;58(8):4167–81.
7. Vaidyanathan PP, Pal P. Sparse sensing with co-prime samplers and arrays. IEEE Trans Signal Process. 2011;59(2):573–86.
8. Pal P, Vaidyanathan PP: Coprime sampling and the music algorithm. In: Digital signal processing workshop and IEEE signal processing education workshop; 2011. p. 289–94
9. Liu CL, Vaidyanathan PP. Super nested arrays: linear sparse arrays with reduced mutual coupling part I: fundamentals. IEEE Trans Signal Process. 2016;64(15):3997–4012.
10. Chambers C, Tozer TC, Sharman KC. Temporal and spatial sampling influence on the estimates of superimposed narrowband signals: when less can mean more. IEEE Trans Signal Process. 1996;44(12):3085–98.
11. Pillai SU, Bar-Ness Y, Haber F. A new approach to array geometry for improved spatial spectrum estimation. Proc IEEE. 1985;73(10):1522–4.
12. Friedlander B, Weiss AJ. Direction finding in the presence of mutual coupling. IEEE Trans Antennas Propag. 1991;39(3):273–84.
13. Svantesson T. Mutual coupling compensation using subspace fitting. In: Proceedings of the Sensor array and multichannel signal processing workshop, 2000; 2000. p. 494–8.
14. Grant M. CVX: Matlab software for disciplined convex programming, version 1.21; 2008. p. 155–210.

# DBSCAN-Based Mobile AP Detection for Indoor WLAN Localization

Wei Nie, Hui Yuan(✉), Mu Zhou, Liangbo Xie, and Zengshan Tian

Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
niewei@cqupt.edu.cn, yuanhui0128@foxmail.com, zhoumu@cqupt.edu.cn, xielb@cqupt.edu.cn, tianzs@cqupt.edu.cn

**Abstract.** The vast market of location-based services (LBSs) has brought opportunities for the rapid development of indoor positioning technology. In current indoor venues, by considering the fact that the wireless local area network (WLAN) infrastructure is widely deployed, the indoor WLAN localization method has become the focus of study. Nowadays, the WLAN module is used widely in a large number of advanced mobile devices, and meanwhile there are a variety of WLAN mobile access points (APs) in indoor environment. In this circumstance, due to the uncertainty of the state of mobile APs, the associated received signal strength (RSS) data are usually lowly dependent on the locations, which will consequently result in the decrease in localization accuracy. To solve this problem, a new method of mobile AP detection based on the density-based spatial clustering of applications with noise (DBSCAN) is proposed. This method aims to identify mobile APs in target area so as to eliminate the adverse impact of mobile APs on localization accuracy.

**Keywords:** Indoor localization · mobile AP detection · Location dependency · DBSCAN · WLAN

## 1 Introduction

Nowadays, people have been farther and farther away from the problem of getting lost, which has to be attributed to the rapid development of localization technology. As more and more electronic products such as mobile phones and tablet computer are able to access WLAN, indoor localization based on RSS in WLAN environment has gained a good opportunity for development. The reason why the traditional RSS-based indoor localization technology obtains high-accuracy lies in the high location dependency of RSS data. However, more and more electronic devices have mobile hotspot function, which leads to the existence of many mobile APs in the target area. The widespread existence of mobile APs significantly reduces the location dependency of the collected RSS data, causing large fluctuations in the collected RSS and serious localization errors. In this way, removing the impact of mobile APs on indoor


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_152

localization is particularly important. To this end, we propose a new approach of mobile AP detection based on density-based spatial clustering of applications with noise (DBSCAN) to identify mobile APs in WLAN. In this way, in the RSS-based localization process, after the construction of position fingerprint database in the offline phase, the RSS data from the mobile APs are removed, and the RSS data from the stationary APs are used as fingerprint database for matching location in the online phase. The structure of this paper is as follows. Localization-related technologies are introduced in Sect. 2. Mobile AP detection and localization system are described in Sect. 3. Then, Sect. 3 shows the experimental results. Finally, we conclude this paper in Sect. 4.

## 2 Related Work

In the outdoor environment, global positioning system (GPS) [1] can provide us with efficient localization and navigation services. However, GPS signals are not strong enough to provide high-accuracy localization service indoors due to the building obstruction. In this case, there are still a lot of achievements in indoor localization without GPS signals. Among the numerous approaches, the WLAN-based indoor localization has been widely welcomed by scholars because of its low infrastructure input costs and convenient RSS signal acquisition [2]. Indoor localization based on WLAN RSS is divided into offline phase and online phase [3]. In offline phase, a large number of reference points (RP) will be set in the target area and the RSS data collected from all APs at each RP constitute a piece of position fingerprint information. After that, the position fingerprint information at all RPs constitutes the position fingerprint database. In online phase, we will select several known location test points and newly collect RSS data at test points. Then, by comparing the online newly collected RSS data with the preconstructed fingerprint database, we will estimate the location of test points.

## 3 Approach Overview

### 3.1 RSS Sequence Generation

Previous research shows that indoor propagation model can well reflect the relationship between physical location and RSS data [4]. RSS sequences corresponding to the motion path are generated based on the indoor propagation model, as shown in Fig. 1.

For the convenience of the following discussion, the RPs is uniformly selected on every known location path of motion. Next, the RSS at each RP is calculated by the indoor propagation model. We assume that there are APs and Ł motion path in the target area. On each motion path, $N_l$ ($l = 1, \ldots, L$) RPs are been selected. Therefore, the set of RSS sequences, **RSS**, can be represented as

$$\begin{cases} \mathbf{RSS} = \{RSS_l,\ l = 1, \ldots, L\} \\ RSS_l = \begin{bmatrix} rss_{l1}^{1} & \cdots & rss_{lM}^{1} \\ \vdots & \ddots & \vdots \\ rss_{l1}^{N_l} & \cdots & rss_{lM}^{N_l} \end{bmatrix} \end{cases} \tag{1}$$

where $rss_{lm}^{n}$ $(n = 1, \ldots, N_l; m = 1, \ldots, M)$ is the RSS at the $n$th RP of the $l$th motion path from the $m$th AP, which is calculated by

$$rss_{lm}^{n} = p_0 - 10\eta \lg (d_{lm}^{n} d_0) - p_{\text{wall}} \tag{2}$$

where $p_0$ is the RSS value whose location is $d_0$ (usually set to 1m) away from the AP, $\eta$ is the path loss factor, $p_{\text{wall}}$ is the path loss caused by walls, and $d_{lm}^{n}$ is the distance between $n$th RP on the $l$th motion path and the $m$th AP. It can be seen that if the AP is mobile, the RSS generated by the indoor propagation model changes continuously with the movement of the AP. Similarly, it is not hard to imagine that the area where RPs can receive signals from mobile APs will be larger than a stationary one.

### 3.2 Mobile AP Detection

For each AP, we select the RPs whose RSS $r_m^{sel}$ is in the following range

$$\max \{r_{lm}^{n}\} - p_{\text{diff}}^{m} \leq r_m^{sel} \leq \max \{r_{lm}^{n}\} \tag{3}$$

where $\max \{r_{lm}^{n}\}$ is the maximum of RSS from the $m$th AP corresponding different RPs and $p_{\text{diff}}^{m}$ can be calculated by

$$p_{\text{diff}}^{m} = -10\eta \lg (1 + d_{\text{diff}} d_{\text{min}}^{m}) \tag{4}$$

where $d_{\text{diff}}$ is the signal reference distance (SRD) and $d_{\text{min}}^{m}$ is the minimum distances between the $m$th AP and different RPs. According to the electromagnetic field coherent diffraction theory [5], we usually set $d_{\text{diff}}$ to 10 m.

Then, we cluster the position coordinates of the RP selected for each AP through DBSCAN [6]. At this time, if the maximum distance, $d_{\max}$, between two RPs in a cluster is larger than $2d_{\text{diff}}$, the corresponding AP is identified as a mobile one, and otherwise it is identified as a stationary one. Here, the reason why the maximum distance between two RPs in the mobile AP's cluster is larger than the stationary one can be interpreted by the fact that if the AP is mobile, as the AP moves, more RPs can receive signals from the mobile AP and the RSS is much stronger. Therefore, for mobile APs, we will select more RPs for clustering, and the cluster will be larger. By comparing the sizes of $d_{\max}$ corresponding different APs and $2d_{\text{diff}}$, we can accurately determine whether the AP is mobile or stationary. To illustrate this process clearer, the pseudo-code of the proposed approach is shown in Algorithm 1.

## 4 Experimental Result

### 4.1 Environmental Layout

For this target area, we deploy five stationary APs and two mobile APs in five rooms and two corridors, as shown in Fig. 1.

**Algorithm 1** Pseudo-code of mobile AP detection

**Input:** $m$-th AP's RSS sequence
**Output:** $m$-th AP's mobility

$d_{\text{diff}} \leftarrow$ 10m; // Initialize SRD
Calculate $p^n_{\text{diff}}$;
Initialize $L$; //Number of motion paths
**for** $l$= 1:$L$ **do**
  **for** $n$ = 1:$N_l$ **do**
    **if** $\max\{r^n_{lm}\} - p^m_{\text{diff}} \leq r^{sel}_m \leq \max\{r^n_{lm}\}$ **then**
      $r^{sel}_m$ is selected;
    **end if**
  **end for**
**end for**
$C \leftarrow 0$; //Initialize clusters number
$\Gamma \leftarrow r^{sel}_m$; //Initialize unvisited RP set
**while** $\Gamma \neq \Phi$ **do**
  **for** each unvisited RP, $p \in \Gamma$ **do**
    $\Psi$ = getNeighbors(p, Eps); //Find p's neighbors with the radius Eps
    **if** sizeof($\Psi$) ¡ MinPts **then**
      Label $p$ as noise;
    **else**
      $C \leftarrow C + 1$
      Add $p$ into the cluster with the cluster ID $C$;
      **for** each point $p'$ in $\Psi$ **do**
        Label $p'$ as visited RP;
        $\Psi'$ = getNeighbor($p$, Eps);
        **if** sizeof($\Psi'$)≥ MinPts **then**
          $\Psi = \Psi \cup \Psi'$;
        **end if**
        **if** $p'$ is not included in any cluster **then**
          Add $p'$ into the cluster with the cluster ID C;
        **end if**
      **end for**
    **end if**
  **end for**
**end while**
Calculate $l_{\max}$;
**if** $l_{\max} > 2l_{\text{diff}}$ **then**
  The $n$-th AP is in mobile state;
**else**
  The $n$-th AP is in stationary state;
**end if**

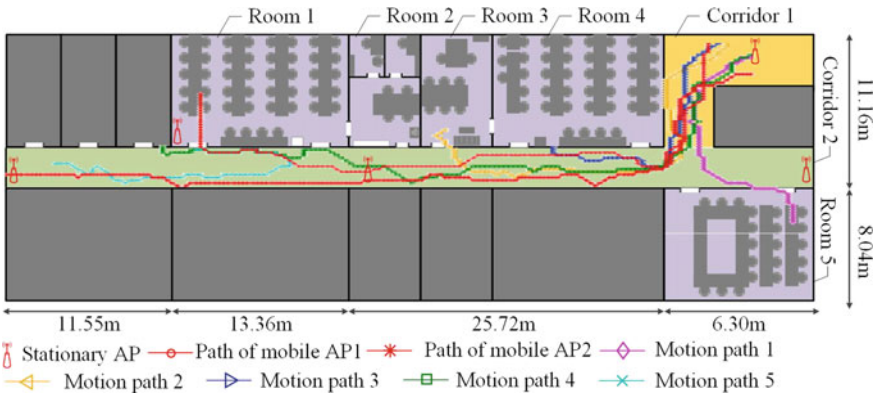



**Fig. 1.** Layout of experiment

### 4.2 Mobile AP Detection Results

For ease of testing, we randomly select five known motion paths in the target area as shown in Fig. 1. Then, we select one RP every second on this five paths and generate corresponding RSS sequences based on the indoor propagation model [4]. Next, we cluster the filtered RPs' position based on DBSCAN and obtain clustering results, as shown in Figs. 2 and 3.

After getting the clustering results, we calculate each AP's maximum distance between two RPs and the result is shown in Table 1. We preset the threshold 20 m to detect mobile AP. From Table 1, we can see that our approach can accurately detect mobile APs.

**Table 1.** Maximum of distance between different RPs in the cluster

| AP IDs | Maximum of distance (m) |
|---|---|
| AP1 | 8.49 |
| AP2 | 9.30 |
| AP3 | 13.15 |
| AP4 | 8.58 |
| AP5 | 11.68 |
| Mobile AP1 | 21.65 |
| Mobile AP2 | 24.36 |

### 4.3 Contrast of Localization Accuracy with and Without Mobile AP Detection

In order to verify the improvement of mobile AP detection for indoor localization accuracy, we use the traditional KNN algorithm to locate

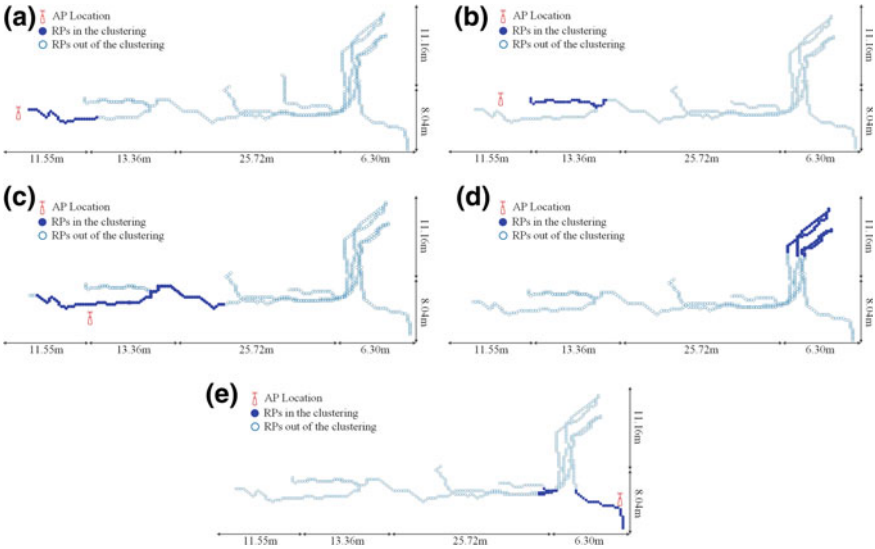



**Fig. 2.** Clustering result by the DBSCAN for stationary AP.

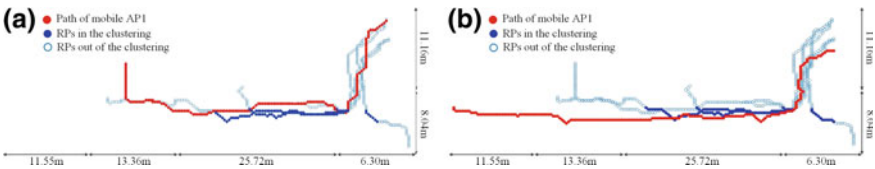



**Fig. 3.** Clustering result by the DBSCAN for mobile AP.

[7]. From Fig. 4, we can find that with the use of mobile AP detection, localization accuracy has been increased by about one time. Therefore, mobile AP detection can effectively improve the indoor localization accuracy.

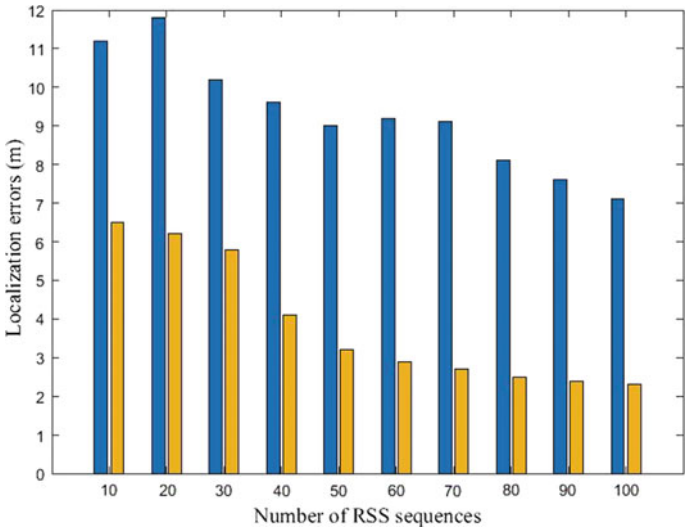



**Fig. 4.** Mean of localization errors with and without mobile AP detection.

## 5 Conclusion

We have proposed a new approach to detect mobile APs with the purpose of removing the RSS data from them and meanwhile claimed that the RSS data from stationary APs are featured with strong location dependency. At the same time, the KNN algorithm is applied to verify that the localization accuracy after detecting mobile APs can be significantly improved compared with the one without mobile AP detection.

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation of China (61771083, 61704015), Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), Special Fund of Chongqing Key Laboratory (CSTC), Fundamental Science and Frontier Technology Research Project of Chongqing (cstc2017jcyjAX0380, cstc2015jcyjBX0065), Scientific and Technological Research Foundation of Chongqing Municipal Education Commission (KJ1704083), and University Outstanding Achievement Transformation Project of Chongqing (KJZH17117).

## References

1. Saha HN, Basu S, Auddy S, et al. A low cost fully autonomous GPS (Global Positioning System) based quad copter for disaster management. In: IEEE annual computing and communication workshop and conference; 2014. p. 654–60.
2. Chan F, Chan YT, Inkol R. Path loss exponent estimation and RSS localization using the linearizing variable constraint. In: Military communications conference; 2016. p. 225–9.
3. Zhou M, Tang Y, Tian Z, et al. Semi-supervised learning for indoor hybrid fingerprint database calibration with low effort. IEEE Access. 2017;5(99):4388–400.
4. Chai P, Zhang L. Indoor radio propagation models and wireless network planning. In: IEEE international conference on computer science and automation engineering; 2012. p. 738–41.
5. Cheung KW, Sau JHM, Murch RD. A new empirical model for indoor propagation prediction. IEEE Trans Veh Technol. 1998;47(3):996–1001.
6. Wang J, Tan N, Luo J, et al. WOLoc: WiFi-only outdoor localization using crowdsensed hotspot labels. In: INFOCOM 2017—IEEE conference on computer communications; 2017. p. 1–9.
7. Markom MA, Adom AH, Shukor SAA, et al. Scan matching and KNN classification for mobile robot localisation algorithm. In: IEEE international symposium in robotics and manufacturing automation; 2017. p. 1–6.

# Error Bound Estimation for Wi-Fi Localization: A Comprehensive Survey

Mu Zhou, Yanmeng Wang(✉), Shasha Wang, Hui Yuan, and Liangbo Xie

Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
{zhoumu,xielb}@cqupt.edu.cn, hiwangym@gmail.com,
{w20ss08,yuanhui0128}@foxmail.com

**Abstract.** Applications on location-based services (LBSs) have driven the increasingly demand for indoor localization technology. Motivated by the widely deployed wireless local area network (WLAN) infrastructure and the corresponding easily accessible WLAN received signal strength (RSS) data, the Wi-Fi signal-based localization has become one of the superior positioning techniques in GPS-denied scenes. Meanwhile, the error bound estimation for the Wi-Fi localization has been attracting much attention due to its significant guidance meaning in practice. In this survey, the error bound estimation approaches for different categories of Wi-Fi localization approaches are overviewed and compared, including the error bound estimation with temporal and spatial signal features, and that with the RSS characteristics. Regarding the temporal and spatial signal feature-based Wi-Fi localization, we present how to utilize the time of arrival (TOA), the time difference of arrival (TDOA) as well as the arrival of angle (AOA) to analyze the error bound of localization systems. Regarding the received signal strength (RSS) characteristic-based Wi-Fi localization, we clarify the error bound estimation approaches for both the wireless signal propagation-based and location fingerprinting-based localization schemes. In addition, some future directions with respect to the error bound estimation for Wi-Fi localization are also discussed.

**Keywords:** Error bound estimation · Wi-Fi localization · Signal features

## 1 Introduction

For well over a decade, the rapid development of wireless communication technology has driven the increasing demand for the location-based services (LBSs). Because of the complicated indoor building structure and multipath effect, the performance of the outdoor positioning systems such as global positioning will dramatically deteriorate in indoor environment. Meanwhile, with the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_153

wide deployment of wireless Wi-Fi infrastructure, Wi-Fi has become one of the priorities for indoor positioning.

Wireless technology used for indoor positioning has been reviewed [1], while few of these works focus on the Wi-Fi localization error estimation, which has not yet been properly reviewed but has a significant guidance meaning in real practice. In this paper, we first introduce some typical localization techniques in Wi-Fi environment from following aspects.

(1) **Temporal and spatial feature-based Wi-Fi localization**. The temporal feature-based Wi-Fi localization includes the time of arrival (TOA) and time difference of arrival (TDOA)-based approach, while the spatial feature based localization contains the localization approach with the arrival of angle (AOA) information.
(2) **Received signal strength (RSS) characteristic-based Wi-Fi localization**. The RSS characteristic-based Wi-Fi localization utilizes the RSS from access points (APs) to estimate the collecting locations and can be divided into two categories including the wireless signal propagation based as well as the location fingerprinting.

Then, the error bound estimation approaches for different categories of Wi-Fi localization approaches are reviewed and compared from the following categories.

(1) **Error bound estimation with temporal and spatial signal features**. By utilizing the TOA, TDOA as well as the AOA information, the Wi-Fi localization based on temporal and spatial features is clearly clarified.
(2) **Error bound estimation with RSS features**. With the easily hearable RSS from APs, the error bound estimation for both the wireless signal propagation-based as well as the location fingerprinting-based Wi-Fi localization approaches is analyzed.

This survey is organized as follows. In Sect. 2, we give a comprehensive introduction about the existing Wi-Fi localization algorithms. In Sect. 3, the error bound estimation of the temporal and spatial signal feature-based Wi-Fi localization approaches is clearly discussed. In Sect. 4, the error bound estimation with respect to the error bound of the RSS-based Wi-Fi localization approaches is reviewed and compared. Finally, Sect. 5 concludes the paper and gives some future directions.

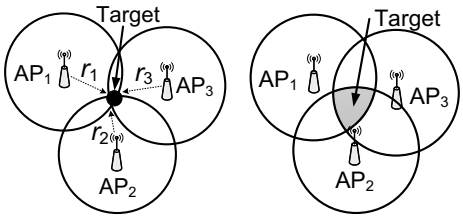
## 2 Review of Wi-Fi Localization Approaches

### 2.1 Temporal and Spatial Feature-Based Wi-Fi Localization

The temporal and spatial feature-based Wi-Fi localization utilizes the TOA, TDOA as well as AOA of the signals from APs to estimate the collecting locations and can be divided into two categories including the temporal feature-based as well as the spatial feature-based Wi-Fi localization approaches.

**Wi-Fi localization with temporal features.** Both TOA and TDOA are the spatial features in Wi-Fi localization. For the localization with TOA, according to the TOA from $AP_i$ to the target under noiseless environment, $t_i$, the distance between $AP_i$ and the target, $r_i$, can be calculated through multiplying $t_i$ by the signal propagation velocity $C$. Then, the location of target is at the intersection of circles as shown in Fig. 1a. However, the TOA-based Wi-Fi localization has extremely strict requirement for the time synchronization between APs and mobile terminal. In addition, the target is often in the non-line-of-sight (NLOS) range round APs, and this will cause large measurement deviation and make the circles in Fig. 1a cannot intersect at a point as shown in Fig. 1b.

To overcome the strict requirement for time synchronization, some researches utilize the TDOA between different APs to localize target. As shown in Fig. 2, the target is at the interaction of hyperbolas, where focal points are the locations of APs, and focal distance is the distance difference from different APs to target. This method just needs the time synchronization between APs, without demand for the time synchronization between APs and mobile terminals. However, the localization approach with TOA is also influenced by the NLOS factors.




(a) In ideal case. (b) In noisy case.

**Fig. 1.** Localization with TOA.

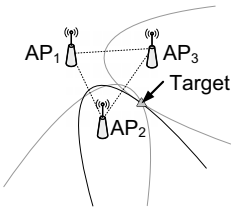



**Fig. 2.** Localization with TDOA.

**Wi-Fi localization with spatial features.** The AOA-based localization approaches utilize the directional antenna or array antennas to measure the direction angles of the signal arrived at APs from the target, which are used to estimate the targets location in Fig. 3. This kind of approach can achieve high even sub-meter level localization accuracy with good superiority in positioning principle and operability. However, the estimation of AOA requires special antenna array, which will increase the cost and the complexity of the localization system. Besides, the localization with AOA is also influenced by NLOS factors.

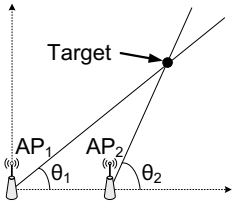



**Fig. 3.** Localization with AOA.

### 2.2 RSS Characteristic-Based Wi-Fi Localization

**Signal propagation model-based Wi-Fi localization.** In this approach, the wireless signal propagation model is established to describe the RSS value at each collecting location from APs, based on which the distance between collecting location and each hearable AP can be calculated, and the corresponding location of mobile terminal is estimated with triangulation method. However, due to the complicated building structure and multipath effect in indoor environment, the constructed signal propagation model is often unstable, which may dramatically decrease the localization accuracy.

**Location fingerprinting-based Wi-Fi localization.** Due to the easily accessible WLAN RSS, the location fingerprinting-based WLAN indoor localization systems have been widely researched. Generally, the location fingerprinting-based Wi-Fi localization consists of two phases, including the offline and online phases. Specifically, in the offline phase, the RSS value from APs collected at each reference point (RP) is regarded as the RPs' fingerprint, which is stored with the RPs' position together in the location fingerprint database. In the online phase, the newly collected RSS signal by mobile terminal is sent to the server and matched with the stored fingerprints to obtain the estimated location. However, due to the influential factors such as the distance interval of RPs, the signal fluctuation indoor environment as well as the reliability of location fingerprint database construction, the distance error between the real and estimated positions is inevitable in practical applications.

## 3 Error Bound Estimation with Temporal and Spatial Signal Features

### 3.1 Error Bound Estimation with Temporal Signal Features

As shown in Table 1, there are a batch of studies focusing on the error bounds of temporal signal feature-based Wi-Fi localization. With the TOA-based distance estimation, the authors in [2] develop an analytical framework for the relationships between multiple-access design parameters and localization error in Wi-Fi environment with a spread-spectrum physical layer, which is further utilized to

**Table 1.** Different methods utilizing temporal and spatial signal features

| Category | Scheme | Signal feature | Key algorithm | Limitations |
|---|---|---|---|---|
| Temporal signal features | [2] | APs' locations, TOA | CRLB | Targets should be located in the LOS range from APs |
| | [3] | APs' locations, TOA | CRLB | FIM may be singular and incalculable |
| | [4] | TOA | CRB | Cannot be applied in the NLOS case |
| | [5] | APs' locations, TOA | Generalized CRLB | FIM may be singular and incalculable |
| | [6] | APs' locations, TOA | CRB | Targets should be located in the LOS range from APs |
| Spatial signal features | [5] | APs' locations, AOA | CRB | Require the prior information of distance range distribution from AP to target |
| | [7] | APs' locations, AOA | CRB | Cannot be applied in the time-varying environment |
| | [8] | APs' locations, AOA | CRB | Cannot be applied in the NLOS environment |

extract the optimal network parameters for maximizing the localization accuracy. The authors in [3] compute the Cramer–Rao lower bound (CRLB) on positioning accuracy under different assumptions on the network synchronization to assess the ultimately achievable accuracy in practice, and the corresponding numerical examples show that the collaboration of multi-APs can eliminate the influence caused by AP deployment. The authors in [4] derive two kinds of error bound with respect to the TOA-based Wi-Fi localization, where the first kind utilizes the geographic information to calculate the upper and lower Cramer–Rao bound (CRB) of localization error with the prior information about at least three APs locations, and the second kind derives the approximate CRB for Wi-Fi localization without any prior knowledge of the APs.

However, these approaches all assume that the targets are localized in the line-of-sight (LOS) range round APs, which may be inconsistent with the real environment especially the complicated indoor environment. To overcome this, the authors in [5] propose a generalized lower CRB-based localization error estimation in both the LOS and NLOS environments, and the experimental results prove that the multipath information of Wi-Fi signal can effectively improve the localization performance. In addition, in the environment with dense AP deployment, some APs' locations cannot be accurately obtained, thus will cause with derived Fisher Information Matrix (FIM) in CRB singular and incalculable. To solve this, the authors in [6] construct an equivalent FIM to derive the localization error bound of the TOA-based system.

### 3.2 Error Bound Estimation with Spatial Signal Features

As one of the most representative studies, the authors in [4] construct the Fisher Information Matrix (FIM) as a function of the angles between the target and APs to derive the upper and lower CRB of localization error, respectively. Specifically, according to the AOA from the $i$th mobile terminal to the $k$th AP, $\alpha_{ik}$, the FIM for AOA-based localization can be constructed as $J=\frac{1}{\sigma^2}\left[\begin{array}{cc} \frac{W+\sum_{k=1}^{W}\cos(\alpha_{ik})}{2} & \frac{\sum_{k=1}^{W}\sin(\alpha_{ik})}{2} \\ \frac{W+\sum_{k=1}^{W}\cos(\alpha_{ik})}{2} & \frac{W-\sum_{k=1}^{W}\cos(\alpha_{ik})}{2} \end{array}\right]$, where $W$ is the AP number, and $\sigma^2$ is the variance of the distance range from each AP to the target. Then, with the unbiased estimator, the lower CRB with respect to the localization error of the target, $\tilde{d}$, can be obtained by $\tilde{d}^2 \geq \frac{4W\sigma^2}{W^2-\left(\sum_{k=1}^{W}\cos(\alpha_{ik})\right)^2-\left(\sum_{k=1}^{W}\sin(\alpha_{ik})\right)^2} \geq \frac{4\sigma^2}{W}$.

However, this method requires the prior assumption that the distance ranges from APs to the target are i.i.d. Gaussian with zero mean and common variance $\sigma^2$, which may be inconsistent with the real case and make the derived error bound cannot effectively describe the localization error in target environment. Different from this, the authors in [7] utilize AOA measurements to analyze the localization accuracy of moving target and extract the lower bound of distance error between real and estimated positions via CRLB. However, this approach requires to assume that the AOA of received signals at a location is stationary, which may be unsuitable for the real environment. Besides, the authors in [8] derive the lower error bound for AOA-based passive source localization.

## 4 Error Bound Estimation with RSS Features

The existing studies on the error bound estimation of RSS-based Wi-Fi localization are usually based on the wireless signal propagation model, kernel density estimation, and other approaches. Some typical works are shown in Table 2 and are described as follows.

Up to now, the overwhelming majority of studies working on the error bound of the RSS-based Wi-Fi localization are based on the CRB. For example, the

**Table 2.** Different methods utilizing RSS characteristics.

| Scheme | Signal feature | Key algorithm | Limitations |
|---|---|---|---|
| [9] | APs' location, SSD, AOA | Joint PDF of SSD, CRLB | Need to assume that the RSS follows Gaussian distribution |
| [10] | APs' location, RSS, AOA | Joint PDF of RSS, CRLB | Cannot be applied in the time-varying environment |
| [11] | APs' location, RSS, AOA | Joint PDF of RSS sum from APs, CRLB | Require accurate APs' locations |
| [12] | APs' location, RSS, AOA | Log-normal model of RSS, CRLB | Cannot be applied in the time-varying environment |
| [13] | APs' location, RSS, AOA | CRLB | RSS should follow Gaussian distribution |
| [14] | APs' location, RSS, AOA | Joint PDF of different distributed RSS, CRLB | Cannot be applied in the time-varying environment |
| [15] | RSS | Nonparametric kernel density estimation | Require accurate PDF of RSS |
| [16] | RSS | Hyper-parameters estimation | RSS should follow Gaussian distribution |

authors in [9] rely on the CRLB to analyze the error bound of Wi-Fi localization using the signal strength difference (SSD) data as the location fingerprint. In concrete terms, based on wireless signal propagation model, the SSDs expression can be obtained as $[p_{ik_1}/p_{ik_2}]_{dB} = -10\beta \log(d_{ik_1}/d_{ik_2}) + n_{ik_1} - n_{ik_2}$, where $p_{ik_1}$ and $p_{ik_2}$ denote the RSS value at $\mathrm{RP}_i$ from $\mathrm{AP}_{k_1}$ and $\mathrm{AP}_{k_2}$, respectively, in mW scale, $d_{ik_1}$ and $d_{ik_2}$ are the distances from $\mathrm{AP}_{k_1}$ and $\mathrm{AP}_{k_2}$, and $n_{ik_1}$ and $n_{ik_2}$ are the corresponding signal noise with the same variance $\sigma^2$. Utilizing SSD expression, the joint probability density function (PDF) of the independent SSD measurements with $W$ APs can be written as $f(p_i) = \prod_{k=1}^{W-1} \frac{1}{\sqrt{2\pi W\sigma^2}} \frac{10}{In10} \frac{p_{ik_1}}{p_{ik_2}} \times \exp\{-[10\log(p_{ik_1}/p_{ik_2}) + 10\beta\log(d_{ik_1}/d_{ik_2})]^2/2\sigma^2\}$. Based on this, the FIM of the localization is constructed, and the CRLB of target's location can be obtained.

Similarly to this, the authors in [10] adopt the joint PDF with respect to the RSS value from each AP to calculate the CRLB of Wi-Fi localization errors. The

authors in [11] rely on the sum of RSS from APs to establish a CRLB-based error bound estimation criterion for Wi-Fi localization. With the log-normal model for RSS measurement, the authors in [12] derive a closed-form solution to the CRLB of the variance with respect to the distance between real and estimated locations. The authors in [13] obtain the CRLB of the distance between real and estimated locations in dynamic environment. In addition, the authors in [14] construct the joint PDF of RSS by considering three basic RSS distributions (i.e., Gaussian, Rayleigh, and Rice distributions) in Wi-Fi environment, and derive the error bound of Wi-Fi localization under both the LOS and NLOS scenarios.

However, due to the difficulty of acquiring the accurate AP locations as well as complexity and dynamics of indoor signal propagation property, the CRLB-based approaches are most often ineffective and unstable for indoor WLAN localization error bound estimation [10]. Different from above estimation methods, the authors in [15] utilize the nonparametric kernel density function to estimate the PDF of errors and as well as the corresponding confidence regions for the generalized location fingerprint-based positioning, based on which the localization error bounds are derived. The authors in [16] derive the CRLB of location fingerprint-based Wi-Fi localization via analyzing the hyper-parameters of kernel function for RSS distribution. However, the kernel function-based error bound estimation schemes often require exact PDF expression of RSS as well as some prior information about the Wi-Fi signal distribution in the target environment.

## 5 Conclusion

In this paper, the error bound estimation approaches for different categories of Wi-Fi localization approaches are comprehensively reviewed and compared, including the error bound estimation with temporal and spatial signal features, and that with the RSS characteristics. For the temporal and spatial signal feature-based Wi-Fi localization, we present how to utilize the TOA, TDOA as well as the AOA to calculate the error bound of localization systems. For the RSS characteristic-based Wi-Fi localization, we clarify the error bound estimation approaches for both the wireless signal propagation-based and location fingerprinting-based localization schemes. However, the assumptions of the target environment in most existing error bound estimation approaches are ideal and simple and cannot be widely spread. Therefore, developing the localization error estimation approaches for more complicated and time-varying Wi-Fi environment forms an interesting work in future.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China (61771083, 61704015), Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), Special Fund of Chongqing Key Laboratory (CSTC), Fundamental and Frontier Research Project of Chongqing (cstc2017jcyjAX0380, cstc2015jcyjBX0065), University Outstanding Achievement Transformation Project of Chongqing (KJZH17117), and Postgraduate Scientific Research and Innovation Project of Chongqing (CYS17221).

## References


1. He S, Chan SHG. Wi-fi fingerprint-based indoor positioning: recent advances and comparisons. IEEE Commun Surv Tutorials. 2017;18(1):466–90.
2. Venkatesh S, Buehrer RM. Multiple-access insights from bounds on sensor localization. Pervasive Mob Comput. 2008;4(1):33–61.
3. Larssom E. Cramer-Rao bound analysis of distributed positioning in sensor netwroks. IEEE Signal Process Lett. 2004;11(3):334–7.
4. Chang C, Sahai A. Estimation bounds for localization. In: IEEE SECON; 2004, p. 415–24.
5. Qi Y, Kobayashi H, Suda H. On time-of-arrival positioning in a multipath environment. IEEE Trans Veh Technol. 2006;55:1516–26.
6. Shen Y, Wymeersch H, Win MZ. Fundamental limits of wideband cooperative localization via fisher information. In: IEEE WCNC; 2007. p. 3951–5.
7. Hejazi F, Norouzi Y, Nayebi MM. Lower bound of error in AOA based passive source localization using single moving platform. In: IEEE East-west design and test international symposium; 2013. p. 1–4.
8. Luo J, Zhang XP, Wang Z. A new passive source localization method using AOA-GROA-TDOA in wireless sensor array networks and its Cramer-Rao bound analysis. In: IEEE ICASSP; 2013. p. 4031–5.
9. Hossain AKMM, Soh WS. Cramer-Rao bound analysis of localization using signal strength difference as location fingerprint. In: IEEE INFOCOM; 2010. p. 1–9.
10. Stella M, Russo M, Begusic D. RF localization in indoor environment. Radioengineering. 2012;21(2):557–67.
11. Laitinen E, Lohan ES. Access Point topology evaluation and optimization based on Cramer-Rao lower bound for WLAN indoor positioning. In: International conference on localization and GNSS; 2016. p. 1–5.
12. Mazuelas S, Bahillo A, Lorenzo RM, et al. Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks. IEEE J Sel Top Signal Process. 2009;3(5):821–31.
13. Zhou M, Xu K, Tian Z, et al. Error bound analysis of indoor Wi-Fi location fingerprint based positioning for intelligent access point optimization via Fisher information. Comput Commun. 2016;86:57–74.
14. Kaemarungsi K. Indoor localization improvement via adaptive RSS fingerprinting database. In: International conference on information networking, vol. 19; 2013. p. 412–6.
15. Jin Y, Soh WS, Wong WC. Error analysis for fingerprint-based localization. IEEE Commun Lett. 2010;14(5):393–5.
16. Kumar S, Hegde RM, Trigoni N. Gaussian process regression for fingerprinting based localization. Ad Hoc Netw. 2016;51:1–10.

# Indoor WLAN Localization Based on Augmented Manifold Alignment

Liangbo Xie, Yaoping Li(✉), Mu Zhou, Wei Nie, and Zengshan Tian

Chongqing Key Lab of Mobile Communications Technology,
Chongqing University of Posts and Telecommunications,
400065 Chongqing, People's Republic of China
{xielb,zhoumu,niewei,tianzs}@cqupt.edu.cn
liyaopingna@foxmail.com

**Abstract.** With the dramatic development of location-based service (LBS), indoor localization techniques have been widely used in recent years. Among them, the indoor wireless local area network (WLAN) localization technique is recognized as one of the most favored solutions due to its low maintenance overhead and high localization accuracy. In this paper, we propose a new received signal strength (RSS)-based indoor localization approach using augmented manifold alignment. First of all, we construct the objective function in manifold space for indoor localization. Second, the optimal transform matrix is used to transform the coordinates of reference points (RPs) and the corresponding RSS vectors into manifold space. Finally, we locate the target at the RP with the transformed coordinates nearest to the transformation of the newly collected RSS vector in manifold space. The experimental results demonstrate that the proposed approach is able to achieve satisfactory localization accuracy with low overhead.

**Keywords:** Indoor localization · Augmented manifold alignment · Transform matrix · Lagrange multiplier · WLAN

## 1 Introduction

With the rapid development of wireless communication technology and the widespread popularity of mobile devices, the demand for location-based service (LBS) is continuously increasing, which has brought unprecedented development space to LBS. Therefore, the indoor localization systems have attracted great attention for domestic and foreign scholars. Among them, wireless local area network (WLAN) indoor localization system [1] is widely used due to the widespread deployment and low overhead of WLAN.

Currently, the measurements used in WLAN indoor localization systems are mainly composed of time of arrival (TOA) [2], time difference of arrival (TDOA) [3], angle of arrival (AOA) [4], andreceived signal strength (RSS) [5] according


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_154

to its basic principles and methods. For RSS-based localization technology, all it needs is to install the corresponding software on the mobile device to store and read the RSS values at different locations without adding any additional equipment, which results in high portability and feasibility. Therefore, RSS-based localization technology is preferred in most WLAN indoor localization systems.

In this paper, to further improve the indoor localization accuracy, a new WLAN RSS-based indoor localization approach using augmented manifold alignment is proposed. Through the objective function construction, we can get the optimized transform matrix that can transform RP coordinates and corresponding RSS vectors into manifold space and still maintains their neighborhood relations. Afterward, the target location can be estimated by finding the nearest RSS to the newly collected RSS vectors in manifold space.

The rest of this paper is organized as follows. In Sect. 2, we give some related work on manifold alignment. Section 3 describes the proposed approach in detail. The experimental results are presented in Sect. 4. Finally, Sect. 5 concludes the paper and provides an outlook of future work.

## 2 Related Work

Generally, semi-supervised learning approach is always used to improve the problem of time-consuming and costly site survey, which can achieve satisfactory localization accuracy through a small number of calibrated points. The authors in [6] propagate the labels to the unlabeled data by embedding labeled data and unlabeled data in common low-dimensional manifold to reduce calibration effort for localization. Without indoor space measurement, the authors in [7] automatically find the best correspondence between floor plan and RSS through graph matching-based manifold alignment process, which can achieve high localization and tracking accuracy.

Different from the works mentioned before, firstly, we augment the physical coordinates of RPs into the same dimension as the corresponding RSS vectors. Then the objective function of manifold alignment is constructed to obtain the optimal transform matrix. Finally, we convert the online RSS vector collected at the target location into manifold space and choose the nearest coordinate as the estimated target location.

## 3 System Overview

### 3.1 Algorithm Description

As shown in Fig. 1, the proposed algorithm is composed of two phases, namely offline phase and online phase, respectively. In offline phase, we first collet RSS vectors $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_m)^{\mathrm{T}}$ from $m$ RPs, where $\mathbf{r}_i = (r_{i1}, \ldots, r_{in})$ is the RSS measure collected at the $i$th RP and $r_{ij}$ $(j = 1, \ldots, n)$ denotes the RSS measure from the $j$th AP and $n$ is the number of APs. Then, we construct the objective

function and map RP coordinates and corresponding RSS vectors into manifold space, such that we can get the optimized transform matrix through solving the objective function.

In online phase, firstly the online RSS vector is mapped into manifold space and then we choose the nearest physical coordinate to it as the estimated location in manifold space.

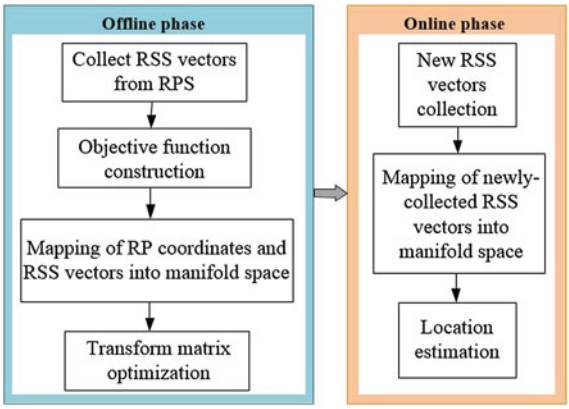



**Fig. 1.** Algorithm process

### 3.2 Objective Function

Let $\mathbf{c}_i^{'} = (x_i, y_i)^{\mathrm{T}}$ be the physical coordinate of the $i$th RP. According to the concept of manifold alignment that data should be connected in the same dimension [8], we firstly augment the two-dimensional coordinate into the same dimension as the corresponding RSS vector $\mathbf{r}_i = (r_{i1}, \ldots, r_{in})$. Thus, we can get

$$\mathbf{c}_i = \begin{cases} \left( \underbrace{x_i, y_i, \cdots, x_i, y_i}_{n2 \text{ paris of } (x_i, y_i)} \right)^{\mathrm{T}}, & \text{when } n \text{ is even} \\ \left( \underbrace{x_i, y_i, \cdots, x_i, y_i}_{(n-1)2 \text{ paris of } (x_i, y_i)}, x_i \right)^{\mathrm{T}}, & \text{when } n \text{ is odd} \end{cases} \tag{1}$$

It can be seen that the above formula still preserves the relative distance between different coordinates of RPs. Based on this, we construct the objective function

of the proposed algorithm for indoor localization as

$$\arg\min_{\mathbf{P_c},\mathbf{P_r}} \left\{ \sum_{i\in\{1,\cdots,m\}} \left\|\mathbf{P_c}^{\mathrm{T}}\mathbf{c}_i - \mathbf{P_r}^{\mathrm{T}}\mathbf{r}_i\right\|_2^2 + \sum_{i,i'\in\{1,\cdots,m\};i\neq i'} \left(\left\|\mathbf{P_c}^{\mathrm{T}}\mathbf{c}_i - \mathbf{P_c}^{\mathrm{T}}\mathbf{c}_{i'}\right\|_2^2 S_{ii'}^c\right) + \sum_{j,j'\in\{1,\cdots,m\};j\neq j'} \left(\left\|\mathbf{P_r}^{\mathrm{T}}\mathbf{r}_j - \mathbf{P_r}^{\mathrm{T}}\mathbf{r}_{j'}\right\|_2^2 S_{jj'}^r\right)\right\} \tag{2}$$

where the notation "$\|\cdot\|_2$" represents the two-norm operation, the transform matrices $\mathbf{P_c}$ and $\mathbf{P_r}$ are optimized to transform $\mathbf{c}_i$ and $\mathbf{r}_i$ into $\mathbf{a}_i = \mathbf{P_c}^{\mathrm{T}}\mathbf{c}_i$ and $\mathbf{b}_i = \mathbf{P_r}^{\mathrm{T}}\mathbf{r}_i$, respectively, in the $n$-dimensional manifold space, $S_{ii'}^c = \exp\left(-\|\mathbf{c}_i - \mathbf{c}_{i'}\|_2^2\right)$, $S_{jj'}^r = \exp\left(-\|\mathbf{r}_j - \mathbf{r}_{j'}\|_2^2\right)$. The first term stands for the proximity between the physical coordinates of RPs and corresponding RSS vectors in manifold space. The second and third term are to preserve proximity between coordinates and RSS vectors, respectively. Since

$$\sum_{i\in\{1,\cdots,m\}} \left\|\mathbf{P_c}^{\mathrm{T}}\mathbf{c}_i - \mathbf{P_r}^{\mathrm{T}}\mathbf{r}_i\right\|_2^2 = \mathrm{Tr}(\mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{C}^{\mathrm{T}}\mathbf{P_c} + \mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{R}^{\mathrm{T}}\mathbf{P_r} - \mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{R}^{\mathrm{T}}\mathbf{P_r} - \mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{C}^{\mathrm{T}}\mathbf{P_c}) \tag{3}$$

$$\sum_{i,i'\in\{1,\cdots,m\};i\neq i'} \left(\left\|\mathbf{P_c}^{\mathrm{T}}\mathbf{c}_i - \mathbf{P_c}^{\mathrm{T}}\mathbf{c}_{i'}\right\|_2^2 S_{ii'}^c\right) = 2\mathrm{Tr}\left(\mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{H_c}\mathbf{C}^{\mathrm{T}}\mathbf{P_c}\right) \tag{4}$$

$$\sum_{j,j'\in\{1,\cdots,m\};j\neq j'} \left(\left\|\mathbf{P_r}^{\mathrm{T}}\mathbf{r}_j - \mathbf{P_r}^{\mathrm{T}}\mathbf{r}_{j'}\right\|_2^2 S_{jj'}^r\right) = 2\mathrm{Tr}\left(\mathbf{P_r}^{\mathrm{T}}\mathbf{C}\mathbf{H_r}\mathbf{C}^{\mathrm{T}}\mathbf{P_r}\right) \tag{5}$$

where the notation "$\mathrm{Tr}(\cdot)$" is the trace of matrix, $\mathbf{C} = (\mathbf{c}_1, \cdots, \mathbf{c}_m)$, $\mathbf{R} = (\mathbf{r}_1, \cdots, \mathbf{r}_m)$, $\mathbf{H_c} = \mathbf{M_c} - \mathbf{N_c}$, $\mathbf{H_r} = \mathbf{M_r} - \mathbf{N_r}$, $\mathbf{N_c} = \left\{N_c^{ii'}\right\}$, $N_c^{ii'} = \exp\left(-\|\mathbf{c}_i - \mathbf{c}_{i'}\|_2^2\right)$, $\mathbf{N_r} = \left\{N_r^{jj'}\right\}$, $N_r^{jj'} = \exp\left(-\|\mathbf{r}_j - \mathbf{r}_{j'}\|_2^2\right)$, $\mathbf{M_c} = diag\left(\sum_{i'=1}^m N_c^{1i'}, \cdots, \sum_{i'=1}^m N_c^{mi'}\right)$, $\mathbf{M_r} = diag\left(\sum_{j'=1}^m N_r^{1j'}, \cdots, \sum_{j'=1}^m N_r^{mj'}\right)$, we have

$$\begin{aligned}&\arg\min_{\mathbf{P_c},\mathbf{P_r}} \left\{\mathrm{Tr}\left(\mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{C}^{\mathrm{T}}\mathbf{P_c} + \mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{R}^{\mathrm{T}}\mathbf{P_r} - \mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{R}^{\mathrm{T}}\mathbf{P_r} - \mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{C}^{\mathrm{T}}\mathbf{P_c}\right)\right.\\ &\left.+2\mathrm{Tr}\left(\mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{H_c}\mathbf{C}^{\mathrm{T}}\mathbf{P_c}\right) + 2\mathrm{Tr}\left(\mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{H_r}\mathbf{R}^{\mathrm{T}}\mathbf{P_r}\right)\right\}\end{aligned} \tag{6}$$

Considering the noise interference $\boldsymbol{\delta}$ on RSS, $\mathbf{R}$ is modified into $\mathbf{R} + \boldsymbol{\delta}$, thus the objective function is rewritten as

$$\begin{aligned}&\arg\min_{\mathbf{P_c},\mathbf{P_r}} \left\{\mathrm{Tr}\left(\mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{C}^{\mathrm{T}}\mathbf{P_c} + \mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{R}^{\mathrm{T}}\mathbf{P_r} - \mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{R}^{\mathrm{T}}\mathbf{P_r} - \mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{C}^{\mathrm{T}}\mathbf{P_c}\right)\right.\\ &+\mathrm{Tr}\left(\mathbf{P_r}^{\mathrm{T}}\mathbf{R}\boldsymbol{\delta}^{\mathrm{T}}\mathbf{P_r} + \mathbf{P_r}^{\mathrm{T}}\boldsymbol{\delta}^{\mathrm{T}}\mathbf{R}\mathbf{P_r} + \mathbf{P_r}^{\mathrm{T}}\boldsymbol{\delta}\boldsymbol{\delta}^{\mathrm{T}}\mathbf{P_r} - \mathbf{P_c}^{\mathrm{T}}\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}}\mathbf{P_r} - \mathbf{P_r}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{C}^{\mathrm{T}}\mathbf{P_c}\right)\\ &+2\mathrm{Tr}\left(\mathbf{P_c}^{\mathrm{T}}\mathbf{C}\mathbf{H_c}\mathbf{C}^{\mathrm{T}}\mathbf{P_c}\right) + 2\mathrm{Tr}\left(\mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{H_r}\mathbf{R}^{\mathrm{T}}\mathbf{P_r}\right) + 2\mathrm{Tr}\left(\mathbf{P_r}^{\mathrm{T}}\mathbf{R}\mathbf{H_r}\boldsymbol{\delta}^{\mathrm{T}}\mathbf{P_r}\right)\\ &\left.+ 2\mathrm{Tr}\left(\mathbf{P_r}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{H_r}\mathbf{R}^{\mathrm{T}}\mathbf{P_r}\right) + 2\mathrm{Tr}\left(\mathbf{P_r}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{H_r}\boldsymbol{\delta}^{\mathrm{T}}\mathbf{P_r}\right)\right\}\end{aligned} \tag{7}$$

Let $\mathbf{P} = \begin{pmatrix} \mathbf{P_c} \\ \mathbf{P_r} \end{pmatrix}$, $\mathbf{V} = \begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}$ and $\mathbf{G} = \begin{pmatrix} \mathbf{I} + 2\mathbf{H_c} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} + 2\mathbf{H_r} \end{pmatrix}$, where $\mathbf{I}$ is the $m \times m$ unit matrix, (7) can be converted into

$$\begin{aligned} &\arg\min_{\mathbf{P}} \{\mathrm{Tr}\left(\mathbf{P}^{\mathrm{T}}\mathbf{V}\mathbf{G}\mathbf{V}^{\mathrm{T}}\,\mathbf{P} + \mathbf{P}^{\mathrm{T}}\mathbf{V}\mathbf{G}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\delta} \end{pmatrix}^{\mathrm{T}}\mathbf{P} + \mathbf{P}^{\mathrm{T}}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\delta} \end{pmatrix}\mathbf{G}\mathbf{V}^{\mathrm{T}}\mathbf{P} + \right. \\ &\left. \mathbf{P}^{\mathrm{T}}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\delta} \end{pmatrix}\mathbf{G} \times \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\delta} \end{pmatrix}^{\mathrm{T}}\mathbf{P}\right)\Big\} \text{ s.t. } \mathbf{P}^{\mathrm{T}}\mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{P} = \mathbf{I},\ \mathbf{P}^{\mathrm{T}}\mathbf{V}\mathbf{e} = \mathbf{0} \end{aligned} \tag{8}$$

where $\mathbf{e}$ is the $2m \times n$ all-one matrix. By using the Lagrange multiplier approach and letting the partial derivatives of the above formula with respect to $\mathbf{P}$ equal to 0, we can obtain

$$\begin{aligned} &2\mathbf{V}\mathbf{G}\mathbf{V}^{\mathrm{T}}\mathbf{P} - 2\lambda\mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{P} - \mu\mathbf{V}\mathbf{e} \\ &+2\begin{pmatrix} -\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}} \\ -\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}}\mathbf{P_c} + \mathbf{R}\left(\mathbf{I} + 2\mathbf{H_r}\right)\boldsymbol{\delta}^{\mathrm{T}} + \boldsymbol{\delta}\left(\mathbf{I} + 2\mathbf{H_r}\right)\mathbf{R}^{\mathrm{T}} + \boldsymbol{\delta}\left(\mathbf{I} + 2\mathbf{H_r}\right)\boldsymbol{\delta}^{\mathrm{T}} \end{pmatrix}\mathbf{P} = 0 \end{aligned} \tag{9}$$

where $\lambda$ and $\mu$ are the Lagrange coefficients. By multiplying both sides of (9) with $\mathbf{P}'\mathbf{P}^T$, where $\mathbf{P}'\mathbf{P}^T = \mathbf{E}$ is the $2n \times 2n$ unit matrix and $\mathbf{P}'$ is the invertible matrix of $\mathbf{P}^T$, we have

$$\begin{aligned} &\left(\mathbf{V}\mathbf{G}\mathbf{V}^{\mathrm{T}} + \begin{pmatrix} \mathbf{0} & -\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}} \\ -\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}} & \mathbf{R}\left(\mathbf{I} + 2\mathbf{H_r}\right)\boldsymbol{\delta}^{\mathrm{T}} + \boldsymbol{\delta}\left(\mathbf{I} + 2\mathbf{H_r}\right)\mathbf{R}^{\mathrm{T}} + \boldsymbol{\delta}\left(\mathbf{I} + 2\mathbf{H_r}\right)\boldsymbol{\delta}^{\mathrm{T}} \end{pmatrix}\right)\mathbf{P} \\ &= \lambda\mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{P} \end{aligned} \tag{10}$$

By multiplying both sides of (10) with , we can obtain

$$\begin{aligned} &\mathbf{P}^{\mathrm{T}}\left(\mathbf{V}\mathbf{G}\mathbf{V}^{\mathrm{T}} + \begin{pmatrix} \mathbf{0} & -\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}} \\ -\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}} & \mathbf{R}\left(\mathbf{I} + 2\mathbf{H_r}\right)\boldsymbol{\delta}^{\mathrm{T}} + \boldsymbol{\delta}\left(\mathbf{I} + 2\mathbf{H_r}\right)\mathbf{R}^{\mathrm{T}} + \boldsymbol{\delta}\left(\mathbf{I} + 2\mathbf{H_r}\right)\boldsymbol{\delta}^{\mathrm{T}} \end{pmatrix}\right)\mathbf{P} \\ &= \lambda\mathbf{P}^{\mathrm{T}}\mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{P} = \lambda\mathbf{I} \end{aligned} \tag{11}$$

Let $\mathbf{J} = \mathbf{V}\mathbf{G}\mathbf{V}^{\mathrm{T}} + \begin{pmatrix} \mathbf{0} & -\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}} \\ -\mathbf{C}\boldsymbol{\delta}^{\mathrm{T}} & \mathbf{R}\left(\mathbf{I} + 2\mathbf{H_r}\right)\boldsymbol{\delta}^{\mathrm{T}} + \delta\left(\mathbf{I} + 2\mathbf{H_r}\right)\mathbf{R}^{\mathrm{T}} + \boldsymbol{\delta}\left(\mathbf{I} + 2\mathbf{H_r}\right)\boldsymbol{\delta}^{\mathrm{T}} \end{pmatrix}$ and $\mathbf{K} = \mathbf{V}\mathbf{V}^T$, we convert (11) into

$$\mathbf{J}\mathbf{P} = \lambda\mathbf{K}\mathbf{P} \tag{12}$$

Therefore, the optimal solution $\mathbf{P_{opt}} = \begin{pmatrix} \mathbf{P_{optc}} \\ \mathbf{P_{optr}} \end{pmatrix}$ is composed of $n$ non-zero minimum generalized eigenvalues corresponding to $n$ generalized eigenvectors, where $\mathbf{P_{optc}}$ and $\mathbf{P_{optr}}$ are the optimal solutions to $\mathbf{P_c}$ and $\mathbf{P_r}$ respectively.

During online phase, firstly, map the RSS vector $\mathbf{r_{tar}} = (r_1, \ldots, r_n)^{\mathrm{T}}$ collected at the target location into manifold space as $\mathbf{P_{optr}}^{\mathrm{T}}\mathbf{r_{tar}}$. And then, we select the coordinate of RP closet to $\mathbf{P_{optr}}^{\mathrm{T}}\mathbf{r_{tar}}$ in manifold space as the estimated location of the target, i.e.,

$$\arg\min_{\mathbf{c}_i}\left\{\left\|\mathbf{P}_{\mathbf{optc}}^{\mathrm{T}}\mathbf{c}_i - \mathbf{P}_{\mathbf{optr}}^{\mathrm{T}}\mathbf{r_{tar}}\right\|_2^2\right\} \tag{13}$$

# 4 Experimental Results

## 4.1 Environmental Layout

As shown in Fig. 2, the experimental environment is selected on the fifth floor in a building with the dimensions of 57 m × 25 m. There are five APs fixed in target environment. The 73 RPs are uniformly distributed and there are several RSS vectors collected at each of them.

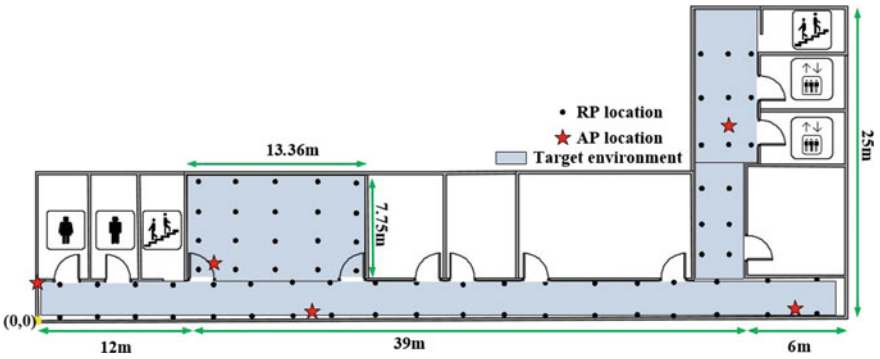



**Fig. 2.** Environmental layout

## 4.2 Localization Performance

We learn the noise interference on localization performance through experiment with measured RSS vectors. The experimental result in Fig. 3 shows that the increase in noise causes the deterioration of localization errors.

In Fig. 4, we compare the mean of localization errors of the proposed localization approach with CIMLoc [9] and WILL [10], respectively. As can be seen, our approach performs better than the other two existing approaches, although the mean of localization errors of CIMLoc is close to the proposed one under large number of RSS sequences. And there is more location information in large size of RSS vectors, which finally improves the performance of localization.

# 5 Conclusion

We propose an indoor WLAN localization approach based on augmented manifold alignment. By preserving the neighborhood relations of the transformed RP coordinates and the corresponding RSS vectors, we achieve satisfactory localization accuracy for the target. In addition, the impact of environmental noise and size of RSS vectors is also discussed through the experiments conducted in the actual indoor WLAN environment. In future, the application of the proposed approach in more complicated indoor environment forms an interesting direction.

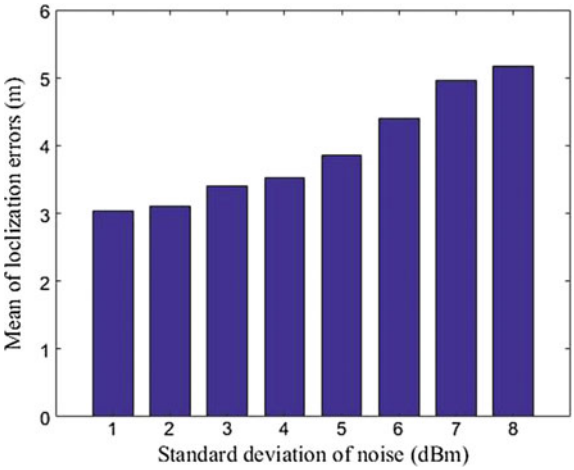



**Fig. 3.** Localization errors under different standard deviation of noise

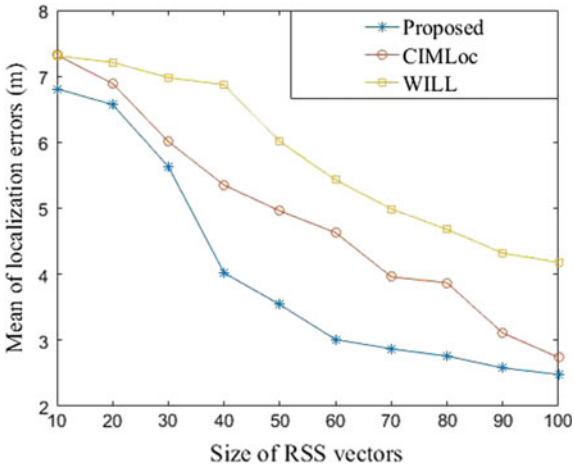



**Fig. 4.** Localization errors under different standard deviation of noise

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation of China (61771083, 61704015), Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), Special Fund of Chongqing Key Laboratory (CSTC), Fundamental Science and Frontier Technology Research Project of Chongqing (cstc2017jcyjAX0380, cstc2015jcyjBX0065), Scientific and Technological Research Foundation of Chongqing Municipal Education Commission (KJ1704083), and University Outstanding Achievement Transformation Project of Chongqing (KJZH17117).

## References

1. Nguyen GK, Nguyen TV, Shin H. Learning dictionary and compressive sensing for WLAN localization. In: IEEE WCNC; 2014. p. 2910–5.
2. Zheng Y, Wang H, Wan L, et al. A placement strategy for accurate TOA localization algorithm. In: Annual communication networks and services research conference; 2009. p. 166–70.
3. Zhang L, Yu X. A Kernel-based TDOA localization algorithm. In: International conference on computer application and system modeling; 2010. p. 412–5.
4. Dogancay K, Hmam H. Optimal angular sensor separation for AOA localization. Signal Process. 2008;88(5):1248–60.
5. Bahl P, Padmanabhan VN. RADAR: an in-building RF-based user location and tracking system. In: IEEE INFOCOM; 2000. p. 775–84.
6. Wang H, Zhang V, Zhao J, et al. Indoor localization in multi-floor environments with reduced effort. In: IEEE international conference on pervasive computing and communications; 2010. p. 244–52.
7. Jiang Z, Zhao J, Han J, et al. Wi-Fi fingerprint based indoor localization without indoor space measurement. In: IEEE international conference on mobile ad-hoc and sensor systems; 2013. p. 384–92.
8. Zhou M, Zhang Q, Tian Z. et al. Indoor WLAN localization using high-dimensional manifold alignment with limited calibration load. In: IEEE ICC; 2017. p. 1–6.
9. Zhang X, Jin Y, Tan HX, et al. CIMLoc: a crowdsourcing indoor digital map construction system for localization. In: IEEE International conference on intelligent sensors, sensor networks and information processing; 2014. p. 1–6.
10. Wu C, Yang Z, Liu Y, et al. WILL: wireless indoor localization without site survey. IEEE Trans Parallel Distrib Syst. 2013;24(4):839–48.

# Trajectory Reckoning Method Based on BDS Attitude Measuring and Point Positioning

Liangbo Xie, Shuai Lu(✉), Mu Zhou, Yi Chen, and Xiaoxiao Jin

Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, People's Republic of China
xielb@cqupt.edu.cn,lushuai.139@163.com,zhoumu@cqupt.edu.cn,
751796746@qq.com,jinxiaoxiaosx@163.com

**Abstract.** The traditional outdoor integrated positioning and navigation system is normally suffered by the disadvantages of accumulative error and high power consumption. To solve this problem, we propose a new trajectory reckoning method which use the BeiDou system (BDS) to conduct the attitude measuring and point positioning with respect to the target. In concrete terms, the target location is estimated by solving the pseudo-range observation equation, while the attitude angle is obtained from the dual-difference pseudo-range and carrier phase observation equations. Then, the trajectory of the target is constructed based on the estimated location and associated attitude angle. Finally, the extensive experimental results demonstrate the effectiveness of the proposed trajectory reckoning method with the BDS attitude measuring and point positioning.

**Keywords:** BeiDou System · Trajectory reckoning · Attitude measuring · Point positioning · Carrier phase

## 1 Introduction

The BDS has been used in all aspects of daily life, such as automatic driving, unmanned navigation, and air transportation. In recent years, with the rapid development of BDS, navigation satellites are used to obtain reliable and accurate attitude information, which has become a high-income, low-cost technology. Fan et al. [1] use a centralized extended Kalman filter to achieve attitude measuring based on GPS/gyroscope combination, conduct static and vehicle-borne dynamic experiments. Aboelmagd Noureldin achieves GPS/INS combined attitude measuring and positioning navigation by constructing a neural network model [2]. The traditional integrated navigation systems rely on satellites and


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_155

inertial sensors for integrated navigation to obtain track information, but inertial sensors have disadvantages such as high cost and accumulated error.

In order to solve the problems of the traditional integrated navigation system, we propose a trajectory reckoning method using BDS for positioning and attitude measuring. In attitude measuring phase, by solving the ephemeris data and observation data, the dual-difference pseudo-range and carrier phase observation equations are solved to obtain the attitude information. In point positioning phase, we establish the pseudo-range observation equations based on the satellite ephemeris data to solve the location of receiver.

The rest of the paper is organized as follows. In Sect. 2, we describe the proposed method in detail. In Sect. 3, we introduce the framework of the proposed trajectory reckoning system and also present the algorithm principle of the BDS attitude measuring and point positioning. Section 4 shows the experimental results and finally the conclusion of this paper is given in Sect. 5.

## 2 Related Work

In [3], the authors propose a GPS/BDS-combined baseline solution method and focus on the issue of cycle slips in relative positioning, and conduct a detailed study of how to detect cycle slips. Wu et al. [4] propose a multi-information fusion directional attitude measuring method, and use a micro-electro-mechanical System (MEMS) sensor and a GPS dual antenna to perform fusion attitude measuring. In [5], the authors use GPS and dead reckoning (DR) combined positioning method based on federated Kalman filter structure to carry out research on vehicle-mounted technology based on geographic information system, they use the unscented Kalman filter to achieve GPS/DR combined positioning to further improve the accuracy of the car navigation system.

Different from the works mentioned before, we firstly use the BDS pseudo-range observation equations to solve the location of moving target. Second, the carrier phase observation equations are solved by using the least-squares ambiguity decorrelation adjustment (LAMBDA) algorithm, and the algorithm is also used to solve heading angle. Finally, we use the path deduction method to output estimated trajectory of moving target.

## 3 System Description

As shown in Fig. 1, the proposed system contains heading angle and location calculation modules. In our system, we use BDS antenna to receive the original data. According to the BDS original data, we select BDS dual-antenna attitude measuring algorithm and point positioning algorithm to calculate the location and heading angle of carrier at each moment. Then, the location and heading angle information are used to calculate the trajectory of moving target.

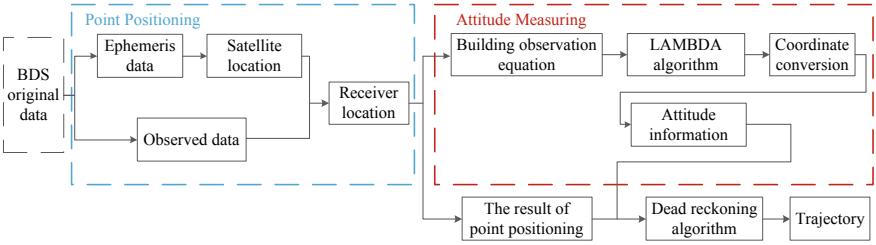



**Fig. 1.** System structure of the proposed BDS trajectory reckoning.

### 3.1 Attitude Measuring

In this paper, we mainly study single-baseline attitude measuring system, which is composed of dual antennas and the two BDS receiving antennas are fixed in the longitudinal direction of the carrier. As shown in Fig. 2, Antenna 1 is the main antenna and Antenna 2 is the secondary antenna.

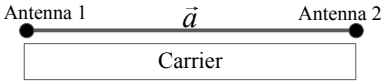



**Fig. 2.** Sketch map of single baseline attitude measuring.

Antenna 1 and Antenna 2 simultaneously perform pseudo-range and carrier phase measurements. As shown in Fig. 3, according to the attitude measuring algorithm, the location of baseline in the Earth-fixed Coordinate System (ENU) can be obtained, so as to calculate the attitude angles yaw and pitch.

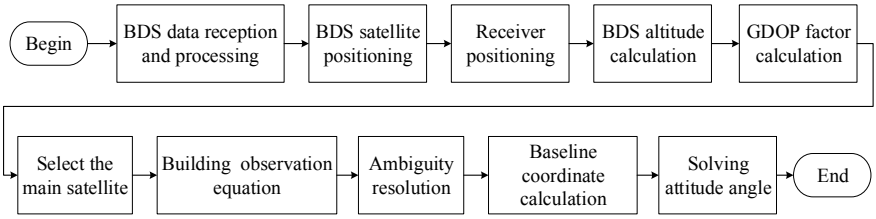



**Fig. 3.** Scheme of single baseline attitude measuring.

#### 3.1.1 Double-Difference Observation Equation

BDS broadcast satellite ephemeris. By analyzing the ephemeris, the location of satellite orbital can be calculated as

$$\begin{cases} X_{\mathrm{k}} = x_k \,.\cos \Omega_k - y_k \,.\cos i_k \,.\sin \Omega_k \\ Y_{\mathrm{k}} = x_k \,.\sin \Omega_k - y_k \,.\cos i_k \,.\cos \Omega_k \\ Z_{\mathrm{k}} = y_k \,.\sin i_k \end{cases} \tag{1}$$

where $[X_k, Y_k, Z_k]$ is the location of satellite and $[x_k, y_k, z_k]$ is the coordinates of satellite in orbital plane.

According to the geometric dilution Of precision (GDOP) factor selection results, the four minimum satellite combinations of GDOP factors are selected. Combined with the principle of double-difference observation model, six double-difference observation models are constructed to establish the following observation equation

$$Y = AX \tag{2}$$

where $A$ is the receiver-to-satellite unit vector matrix and $X$ is the required baseline vector and ambiguity solution. According to the least squares theory, $X$ [6,7] can be obtained as

$$X = \left(A^T \cdot P \cdot A\right)^{-1} \cdot A^T \cdot P \cdot Y \tag{3}$$

where $P$ is weight matrix.

Then, we set the double-difference integer ambiguity vector is $a$ and baseline correction vector is $b = [\delta X_j, \delta Y_j, \delta Z_j]^T$, the least squares result can be calculated by

$$\hat{X} = \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} \tag{4}$$

### 3.1.2 Gesture Solution

By searching for the integer ambiguity by the Lambda algorithm, a double-difference integer ambiguity vector can be obtained. The integer accuracy of the integer ambiguity can be used to further improve the baseline vector estimation accuracy [8]

$$\breve{b} = \hat{b} - Q_{\hat{b}\hat{a}} \cdot Q_{\hat{a}}^{-1} \cdot (\hat{a} - \breve{a}) \tag{5}$$

where the baseline $\breve{b}$ is based on the CGCS2000 (China Geodetic Coordinate System 2000), $\breve{b}_n$ is obtained by converting the coordinate transformation matrix to the ENU coordinate system.

$$\breve{b}_n = \begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} = R_L^n \cdot \breve{b}_l = \begin{pmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\cos\lambda\sin\phi & -\sin\lambda\sin\phi & \cos\phi \\ \cos\cos\varphi & \sin\lambda\cos\varphi & \sin\phi \end{pmatrix} \begin{bmatrix} \delta X \\ \delta Y \\ \delta Z \end{bmatrix} \tag{6}$$

where $\lambda$ is the longitude and $\phi$ is the latitude. The attitude rotation matrix $C_n^b$ is expressed as

$$C_n^b = R_x \cdot R_y \cdot R_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\gamma & 0 & -\sin\gamma \\ 0 & 1 & 0 \\ \sin\gamma & 0 & \cos\gamma \end{pmatrix} \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7}$$

$$\begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} = C_n^b \cdot \begin{bmatrix} x_b \\ y_b \\ z_b \end{bmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\cos\theta\sin\varphi & \cos\theta\cos\varphi & \sin\theta \\ \sin\theta\sin\varphi & -\sin\theta\cos\varphi & \cos\theta \end{pmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{8}$$

$$\begin{cases} x_n = -\cos\theta\sin\varphi \\ y_n = \cos\theta\cos\varphi \\ z_n = \sin\theta \end{cases} \tag{9}$$

The heading angle can be calculated by

$$\varphi = \arctan\frac{x_n}{y_n} \tag{10}$$

### 3.2 Point Positioning

According to the pseudo-range observation equation [9]

$$\rho^{(n)} = r^{(n)} + \delta t_u - \delta t^{(n)} + I^{(n)} + T^{(n)} + \varepsilon\rho^{(n)} \tag{11}$$

where $n = 1, 2, \ldots, N$ is temporary number of satellite measurements, $r^{(n)}$ is the geometry distance from satellite to receiver, $\delta t_u$ is receiver clock error, $\delta t^{(n)}$ is satellite clock error, $I^{(n)}$ is ionosphere delay, $T^{(n)}$ is tropospheric delay, and pseudo-range measurement noise is $\varepsilon\rho^{(n)}$. Error-corrected pseudo-range measurements $\rho_c^{(n)}$ is expressed as

$$\rho_c^{(n)} = \rho^{(n)} + \delta t^{(n)} - I^{(n)} - T^{(n)} \tag{12}$$

The corrected pseudo-range observation equation is described as

$$r^{(n)} + \delta t_u = \rho_c^{(n)} - \varepsilon\rho^{(n)} \tag{13}$$

where $r^{(n)}$ is geometric distance from the receiver to the satellite $n$ that can be expressed as

$$r^{(n)} = \left\| x^{(n)} - x \right\| = \sqrt{\left(x^{(n)} - x\right)^2 + \left(y^{(n)} - y\right)^2 + \left(z^{(n)} - z\right)^2} \tag{14}$$

where $x = [x, y, z]^T$ is an unknown receiver position coordinate vector, $\left[x^{(n)}, y^{(n)}, z^{(n)}\right]^T$ is the location coordinate vector of satellite $n$ [10]. Then, we subtract the pseudo-range error $\varepsilon\rho^{(n)}$ in Eq. (13) to obtain a quaternary nonlinear system of equations

$$\begin{cases} \sqrt{\left(x^{(1)} - x\right)^2 + \left(y^{(1)} - y\right)^2 + \left(z^{(1)} - z\right)^2} + \delta t_u = \rho_c^{(1)} \\ \sqrt{\left(x^{(2)} - x\right)^2 + \left(y^{(2)} - y\right)^2 + \left(z^{(2)} - z\right)^2} + \delta t_u = \rho_c^{(2)} \\ \cdots \\ \sqrt{\left(x^{(N)} - x\right)^2 + \left(y^{(N)} - y\right)^2 + \left(z^{(N)} - z\right)^2} + \delta t_u = \rho_c^{(N)} \end{cases} \tag{15}$$

By solving the four unknowns of the above equations, the location of receiver can be solved. Then, the iterative process is repeated through Newton iteration to obtain accurate positioning results. The iteration process is shown as below.

Step (1) Preparing data and setting initial solution.
Step (2) Linearizing nonlinear equations.
Step (3) Solving linear equations.
Step (4) Updating the root of the nonlinear equations.
Step (5) Determining the convergence of Newton's iteration.

## 4 Experimental Results

In this part, we verify the feasibility of the proposed method in a real-world environment. The platform consists of a BDS receiver, a personal computer (PC), and two BDS antennas, which is as shown in Fig. 4. During the test, the moving target carries the dual-antenna masts move in the direction of the predetermined trajectory, and saves the ephemeris data and observation data received by the BDS receiver on PC.

In the open environment (N 29.5389° and E 106.6039°), we set the heading angle in the north 0°, east 90°, west −90°, and south −180°. The real trajectories of moving target are L-shape and rectangle, respectively.

The L-shaped trajectory starts from direction of the east, and goes straight ahead by 10 m turn to the north. The positioning coordinates of moving target are shown in Fig. 5 and the heading angle is shown in Fig. 6. From Figs. 5 and 6, we can find that the actual positioning trajectory of L-shape test trajectory roughly consistent with the actual trajectory, the heading angle output clearly jumps from 90° to 0° at the corner.

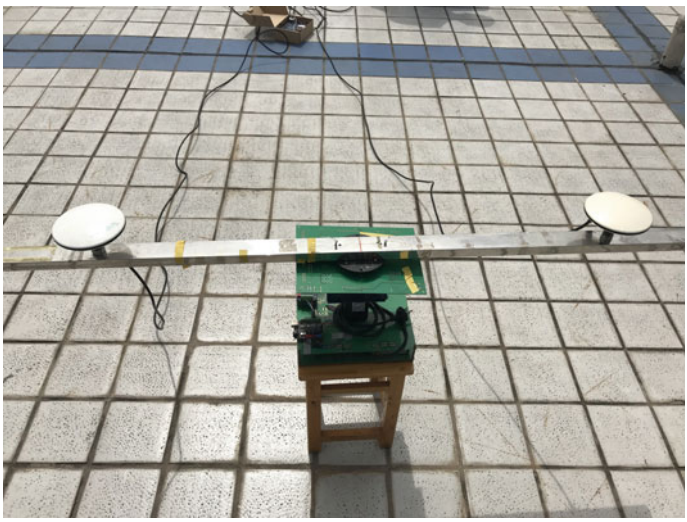


**Fig. 4.** Platform of BDS dual-antenna test.

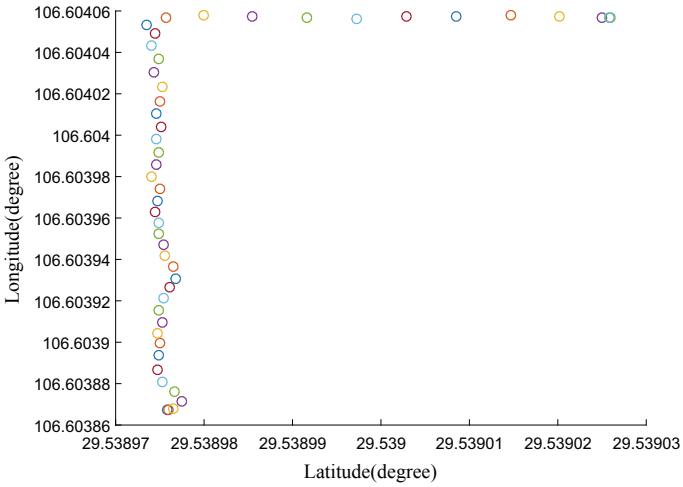



**Fig. 5.** Positioning result of L-shape trajectory.

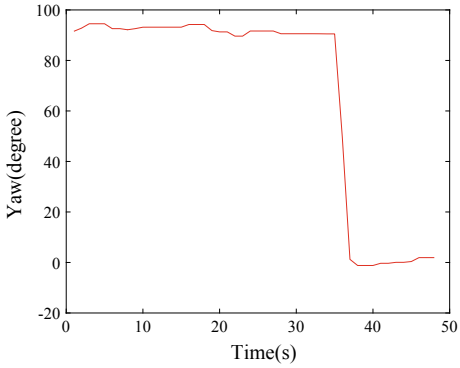



**Fig. 6.** Estimated heading angle of L-shape trajectory.

The rectangle test trajectory starts from direction of the east, and goes straight ahead 10 m turn to the north, then turns straight to the west and goes straight for 10 m, at last, turn left 90°, heading angle is −180°, which is a closed rectangle. The positioning coordinates are shown in Fig. 7 and the heading angle output is shown in Fig. 8. From the two figures, we can find that the estimated trajectory almost fits the true trajectory.

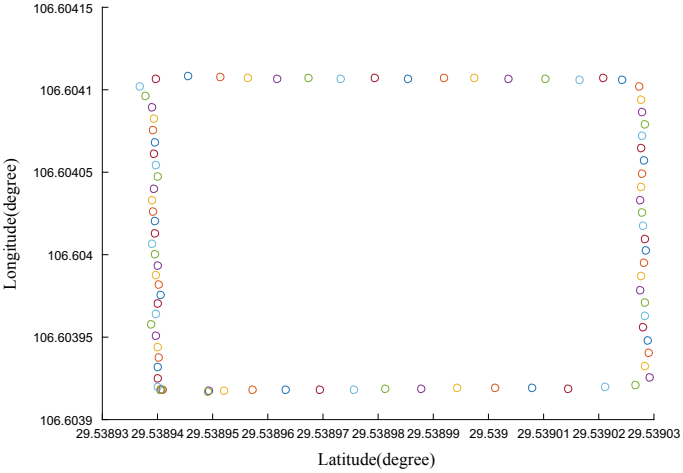



**Fig. 7.** Positioning result of rectangle trajectory.

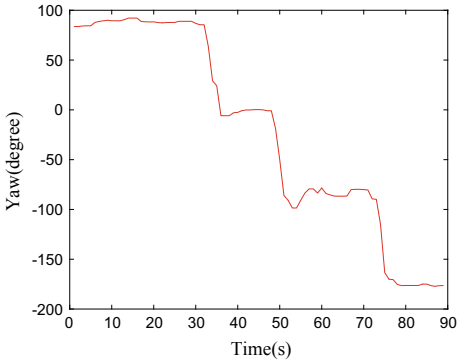



**Fig. 8.** Estimated heading angle of rectangle trajectory.

## 5 Conclusion

In this paper, we rely on the BDS single-baseline attitude measuring and point positioning to effectively improve the accuracy of target trajectory reckoning. By constructing the carrier phase and pseudo-range observation equations, we are capable of estimating the target location and the associated attitude angle. Compared with the traditional integrated positioning and navigation system, the proposed method is featured with high flexibility and low cost.

**Acknowledgments.** This work is supported in part by the National Natural Science Foundation of China (61771083, 61704015), Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), Special Fund of Chongqing Key Laboratory (CSTC), Fundamental Science and Frontier Technology Research Project

of Chongqing (cstc2017jcyjAX0380, cstc2015jcyjBX0065), Scientific and Technological Research Foundation of Chongqing Municipal Education Commission (KJ1704083), and University Outstanding Achievement Transformation Project of Chongqing (KJZH17117).

## References

1. Fan S, Zhang K, Wu F. Ambiguity resolution in GPS-based, low-cost attitude determination. Positioning. 2005;4(1):207–14.
2. Noureldin A, El-Shafie A, Bayoumi M. GPS/INS integration utilizing dynamic neural networks for vehicular navigation. Inf Fusion. 2001;12(1):48–57.
3. Chen J, Yue DJ, Zhao XW, Wang J. BDS/GPS combined single epoch baseline solution method. J Surv Mapp Sci Technol. 2017;34(3):232–5.
4. Wu JJ, Qian F. Research on directional posture method of multiple information fusion. Electron Meas Technol. 2012;35(2):41–5.
5. Yang C, Li RZ, Li Y. The application of GIS-based positioning correction in vehicle-borne inertial navigation system. Telemetry Remote Control. 2015;36(3):70–4.
6. Mosavi MR, Azarshahi S, Emamgholipour I, et al. Least squares techniques for BDS receivers positioning filter using pseudo-range and carrier phase measurements. Iran J Electr Electron Eng. 2014;10(1):18–26.
7. Park C, Teunissen PJG. Integer least squares with quadratic equality constraints and its application to GNSS attitude determination systems. Int J Control Autom Syst. 2009;7(4):566–76.
8. Wang JM, Ma TM, Zhu HZ. Improved LAMBDA algorithm to quickly resolve BDS dual-frequency integer ambiguities. Syst Eng Theory Pract. 2010;37(3):768–72.
9. Xie G. GPS principle and receiver design. Electronic Industry Press; 2011. p. 101–2.
10. He JL, Liu ZM. Beidou navigation satellite position calculation method. Glob Positioning Syst. 2013;38(5):5–10.

# An Adaptive Passive Radio Map Construction for Indoor WLAN Intrusion Detection

Yixin Lin(✉), Wei Nie, Mu Zhou, Yong Wang, and Zengshan Tian

Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
260150244@qq.com, linyixin_cqupt@foxmail.com,
{niewei,zhoumu,tianzs}@cqupt.edu.cn, dr_ywang@hotmail.com

**Abstract.** Indoor WLAN intrusion detection technique for the anonymous target has been widely applied in many fields such as the smart home management, security monitoring, counterterrorism, and disaster relief. However, the existing indoor WLAN intrusion detection systems usually require constructing a passive radio map involving a lot of manpower and time cost, which is a significant barrier of the deployment of WLAN intrusion detection systems. In this paper, we propose to use the adaptive-depth ray tree model to automatically construct an adaptive passive radio map for indoor WLAN intrusion detection. In concrete terms, the quasi-3D ray-tracing model is enhanced by using the genetic algorithm to predict the received signal strength (RSS) propagation feature under the indoor silence and intrusion scenarios, which improves the computational efficiency while preserving the accuracy of passive radio map. Then, the RSS mean, variance, maximum, minimum, range, and median are allied to increase the robustness of passive radio map. Finally, we conduct empirical evaluations on the real-world data to validate the high intrusion detection rate and low database construction cost of the proposed method.

**Keywords:** Indoor intrusion detection · Adaptive ray-tracing · Passive radio map · Genetic algorithm · WLAN

## 1 Introduction

With the wide deployment of wireless local area network (WLAN) and general support of WLAN protocol by various intelligent terminals, the intrusion detection with respect to the indoor target can be realized by using the existing WLAN infrastructure. Among the existing anonymous target intrusion detection techniques, the wireless local area network (WLAN) indoor target intrusion detection system [1–4] proposed by the University of Maryland performs outstandingly


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_156

because it can effectively protect the user's location privacy and work stably under non-line-of-sight and without special hardware at the same time. However, the main problem with this kind of algorithms is that the construction of the prior passive radio map takes a lot of manpower and time, which is a major barrier of WLAN intrusion detection systems deployment. On this basis, the WLAN indoor target intrusion detection algorithm proposed in this paper uses the adaptive-depth ray tree-based quasi-3D ray-tracing model to construct the passive radio map automatically, which requires less labor overhead compared with the traditional RSS feature database construction method. In addition, six signal characteristics of the passive radio map are constructed, which results in the better pattern recognition ability and learning convergence. The rest of this paper is structured as follows. In Sect. 2, we describe the proposed indoor WLAN intrusion detection method in detail, and the related experimental results are shown in Sect. 3. Finally, we conclude this paper in Sect. 4.

## 2 System Description

The overall flow of the system is shown in Fig. 1. First, a number of WLAN access points (APs) and monitor points (MPs) are arranged in the target area. Second, the GA algorithm is used to optimize the limited number of depth of the ray tree adaptively and the RSS characteristics under the indoor silence, and intrusion scenarios are constructed according to the optimized ray-tracing model. Then, the obtained RSS characteristics are used for probabilistic neural network (PNN) training. Finally, the trained PNN is used to classify the new observation RSS data by multiple classifications, so as to realize the intrusion detection and area localization.

### 2.1 Signal Prediction

Considering the limitations of the existing 2D and 3D ray-tracing models [5,6] on the accuracy of the signal prediction and the complexity of the algorithm respectively, the quasi-3D ray-tracing model used in this paper first carries out the ray-tracing in the 2D projection plane, and then transforms it into the propagation path in the 3D space, and this process significantly improves the computational efficiency while guaranteeing the accuracy of prediction. In this case, as shown in Fig. 2, a quasi-3D ray-tracing model based on the adaptive-depth ray tree is proposed in this paper, considering two factors: the model accuracy and calculation efficiency.

**The import of environmental information.** Figure 3 gives a 3D modeling of a simple environment and the corresponding 2D projection results. The gray and black parts of the diagram represent the boundary face of environment and the indoor facilities, respectively. In addition, in order to ensure the integrity of the imported environmental information, the 3D vertex coordinates, height information and relative permittivity, conductivity and permeability of the corresponding material will be recorded.

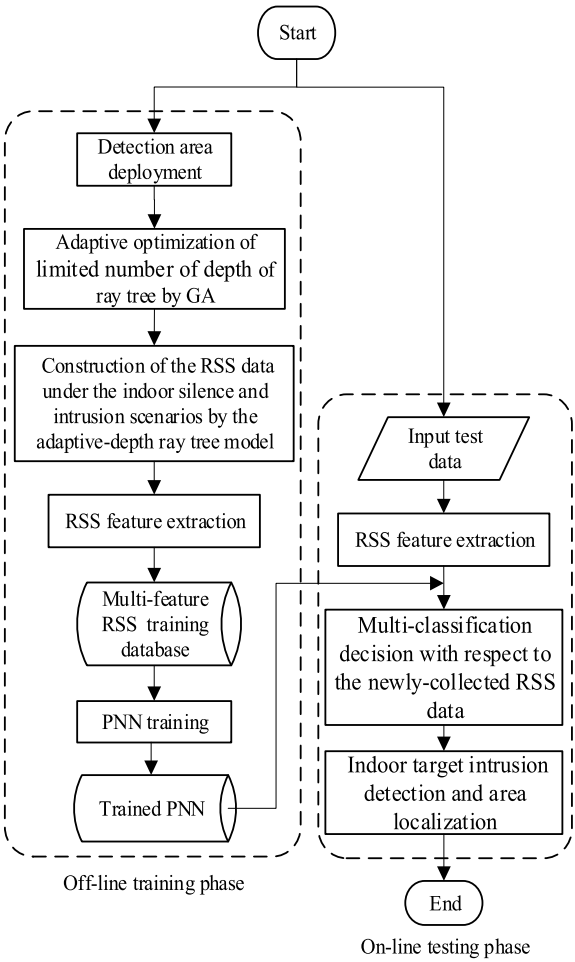



**Fig. 1.** Overall system flowchart

**Optimization of the limited number of depth.** In order to significantly improve the computing efficiency of the ray tree, the GA algorithm is used to optimize the limited number of depth of the ray tree in different environments. Specifically, first, the limited number of depth is initialized to 1; secondly, all the vertical planes and vertical lines of the 3D modeling of the environment are numbered; besides, the number of the functional parts of each ray is spliced into a chromosome in chronological order, and the field strength of each ray that reaches the MP is used as the fitness of its corresponding chromosome; then, the contribution rate of the ray to the field strength at MP under the condition of the current limited number of depth is calculated by Algorithm 1;

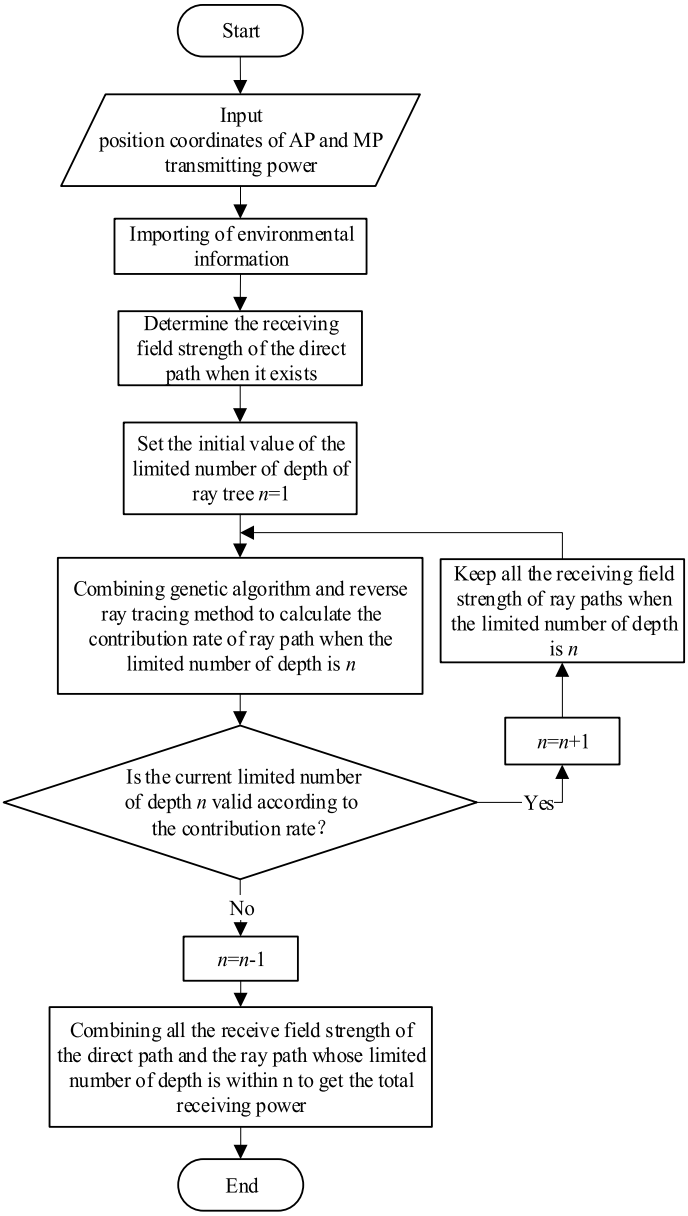



**Fig. 2.** Signal prediction flowchart

finally, determine whether the contribution rate of the ray under the current limited number of depth is greater than the preset threshold, and if so, add 1 to the limited number of depth and repeat the above steps, otherwise the current limited number of depth is the optimal limited number of depth (or the

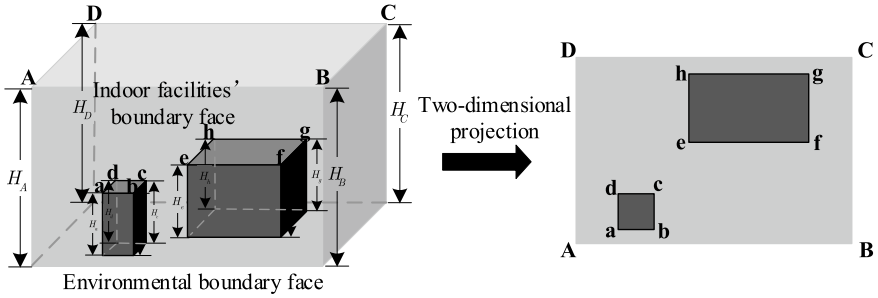



**Fig. 3.** 2D projection from 3D modeling of environment

**Algorithm 1** Calculation of contribution rate of ray to the field strength

**Input:** $n$ (the limited number of depth); $D_k(k{=}1, \ldots, N)$(the unique number of the $k$th functional part); $N$ (the total number of the functional parts); $P_{\mathrm{c}}$ (crossover probability); $P_{\mathrm{m}}$ (mutation probability); $M$ (population size); $\rho_{\mathrm{th}}$(calculation rate threshold); $e_{n-1}$(maximum field strength of the $n$-1 order ray)

**Output:** $C_n$(contribution rate of the ray to the $n$ order field strength)

The first generation population $T_1$ is randomly generated according to the limited number of depth $n$ (that is, $M$ random number sequences with a length of $n$), and set the current population $T = T_1$
**while** $\rho < \rho_{\mathrm{th}}$ **do**
    Use inverse ray-tracing method to calculate the fitness of each chromosome in current population $T$, namely $e_1$,...,$e_M$ (as described in Algorithm 2)
    **for** $i$= 1:$M$ **do**
        2 chromosomes were selected from $T$ by fitness ratio selection algorithm[7]
        **if** random(0,1)< $P_{\mathrm{c}}$ **then**
            implement crossover operation on the selected 2 chromosomes
        **end if**
        **if** random(0,1)< $P_{\mathrm{m}}$ **then**
            implement mutation operation on the selected 2 chromosomes
        **end if**
        Add this 2 new chromosomes to the updated population $T_{\mathrm{new}}$
    **end for**
    $T \Leftarrow T_{\mathrm{new}}$
    Compute $\rho = N_C/N^n$, and $N_C$ is the number of chromosomes types that have appeared from the initial population to the current population
**end while**
Compute $T_f = e_{n-1}/2$
Compute $m$ namely the number of chromosomes whose fitness> $T_f$ in the current population $T$
Compute $C_n = m/M$

optimal ray order). In Algorithm 1, the fitness of each chromosome in the current population is calculated by the reverse ray-tracing method, and its calculation process is described in Algorithm 2.

**Algorithm 2** Calculation of the $n$ order chromosome fitness

**Input:** $D_k$($k$=1, ..., $N$ (the unique number of the $k$th functional part); $N$ (the total number of the functional parts); $P_k$ and $H_k$ ($k = 1, \ldots, N_1$) (vertex coordinate and height information of the $k$th vertical lines); $N_1$ (the total number of the vertical lines); $c_k$ $\varepsilon_k$ and $\mu_k$ ($k = 1, \ldots, N_2$) (relative permittivity, conductivity and permeability of the $k$th vertical planes); $N_2$ (the total number of the vertical planes); $\lambda$ (working wavelength); $n$ (the limited number of depth); $P_{\mathrm{AP}}$ and $P_{\mathrm{MP}}$ (position coordinates of AP and MP); $P_{\mathrm{t}}$ (transmitting power of AP)

**Output:** $e_i$(chromosome fitness of the $i$th $n$ order ray)

Assign $B_k$ ($k = 1, \ldots, n+2$) according to the number sequence of functional part of the $i$th ($i = 1, \ldots, M$) $n$ order ray, in which $B_1 = 1$ and $B_{n+2} = 1$ is the initial position of the ray AP and the termination position MP respectively; $B_k$=0 and 1 ($k = 2, \ldots, n+1$)represents reflection and diffraction occurs in the $k$th functional part respectively.
The 2D projection coordinates of AP and MP are $L_1$ and $L_{n+2}$, respectively
**for** $k = 2$: $n + 1$ **do**
    **if** $B_k = 0$ **then**
        $L_k \Leftarrow$ 2D projection coordinates of the mirror points of $L_{k-1}$ with respect to the $k$th vertical planes
    **else if** $B_k = 1$ **then**
        $L_k \Leftarrow$ 2D projection coordinates of the $k$th vertical planes
    **end if**
**end for**
**while** there exists $k \in (1, ..., n)$ which makes $B_k = 1$ **do**
    **if** $B_k = 0$ and $B_{k+1} = 1$ **then**
        $L_k \Leftarrow$ 2D projection coordinates of the intersection point of the line connecting $L_k$ and $L_{k+1}$ and the $k$th vertical planes
        $B_k \Leftarrow 1$;
    **end if**
**end while**
**for** $k = 1 : n + 1$ **do**
    Set $T_{k,k+1}$ as the line connecting $L_k$ and $L_{k+1}$
    **if** there exists the intersection of $T_{k,k+1}$ and any functional part **then**
        $e_i \Leftarrow 0$
        break;
    **end if**
**end for**
**if** $e_i \neq 0$ **then**
    The 2D projection of the $i$th $n$ order ray extends to 3D space according to the Fermat principle, and then the fitness of the corresponding chromosome $e_i$ is calculated
**end if**

**Calculation of received signal power.** In order to calculate the received signal power of MP, direct and non-direct rays are considered respectively. All the direct and non-direct rays within the $n$ order are superimposed on the signal field strength, and the received signal power at MP can be obtained by the ray

power summation method [8] as

$$P_{total} = \sum_{i=1}^{l} \left( \frac{\lambda \left| \boldsymbol{E}_i \right|}{4\pi \left| \boldsymbol{E}_0 \right|} \right)^2, \tag{1}$$

in which $\boldsymbol{E}_0$ is the arrival signal field strength at 1 m from AP, and $\boldsymbol{E}_i$ is the arrival signal field strength of the $i$th ray, $l$ is the total number of rays.

### 2.2 Intrusion Detection

In this paper, the kernel density estimation method based on Bayesian decision theory is applied to train the PNN feature data under the indoor silence and intrusion scenarios.[1] In particular, the kernel density function is used to estimate the conditional probability of different states, and then the state of the maximum posterior probability is used as the PNN output [9] according to the Bias decision theory. In order to ensure the stability of RSS characteristic data between each pair of AP and MP, this paper uses a sliding window function to segment the original RSS data[2] and calculates the mean, variance, maximum, minimum, maximum, and middle value of each segment data. On the basis of these six signal characteristics, six PNN structures are trained respectively. Finally, according to the voting criterion, the indoor target detection and location are realized by the multiclassification decision of the newly acquired RSS data.

## 3 Experimental Result

### 3.1 Environmental Layout

Figure 4 shows an experimental environment, in which two APs (AP1 and AP2 with model D-Link DAP 2310) and three MPs (MP1, MP2, and MP3 with model SAMSUNG GT-S7568) are placed at 2 and 0.5 m high, respectively. At each MP, 5 min of RSS data from each AP are collected separately under the indoor silence and intrusion scenarios.

### 3.2 GA Optimization Result

Figure 5 shows the change of the overall fitness of each generation of population under the conditions of different values of $\rho_{th}$ when the GA was used to calculate the ray contribution rate. The overall fitness is defined as the ratio of $m$ to the population size $M$, and $m$ is the number of chromosomes whose fitness is greater than the threshold value $T_f$ in the population. It can be seen that with the increasing of population algebra, the overall fitness is on the rise and tends to be the same when the population algebra reaches 30. In addition, Table 1

[1] Considering the content of water in the human body more than 70%, the human body is modeled as a 3D water column [10] with a certain height.

[2] The length of each segment of the RSS data is the width of the sliding window.

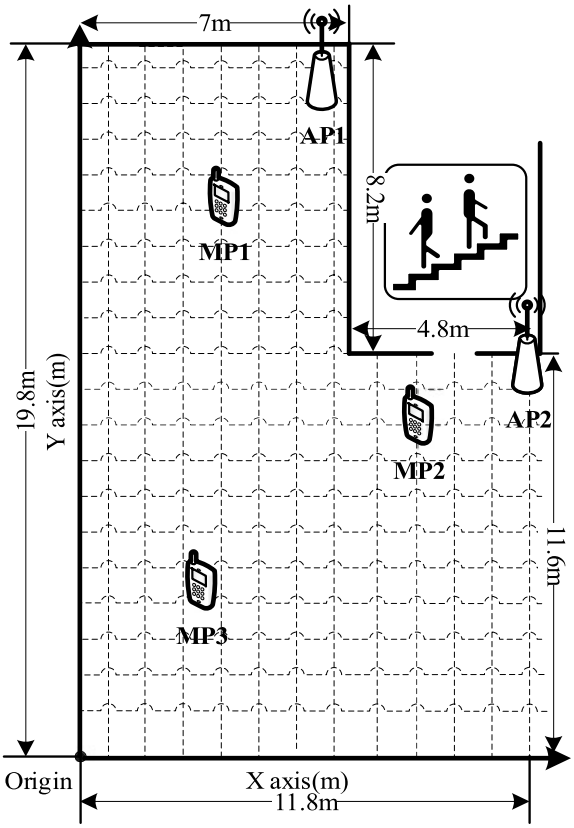



**Fig. 4.** Structure of experimental environment

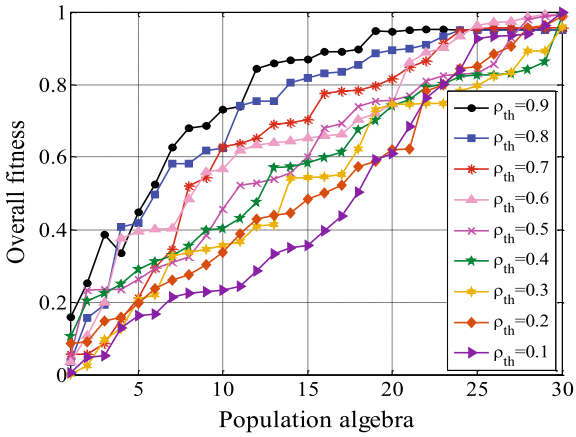



**Fig. 5.** Change of overall fitness

compares the average time overhead required by the 3D ray-tracing model [5], the traditional 2D ray-tracing model [6], and the proposed method for the ray modeling between each pair of AP and MP under the condition of the limited number of depth of 3. It can be seen from the table that this method performs obviously better than the methods used in the literature [5,6] in terms of time overhead.

**Table 1.** Average time cost for ray modeling between each pair of AP and MP

| Performance index | Paper [6] | Paper [5] | The proposed |
|---|---|---|---|
| Time overhead (s) | 6.03 | 7.25 | 3.41 |

### 3.3 Signal Prediction Result

Figures 6 and 7 compare the cumulative density function (CDF) of RSS prediction errors by the proposed method and the ones in [5,6] under the limited number of depth of 3, from which we can find that the proposed method performs better than the others.

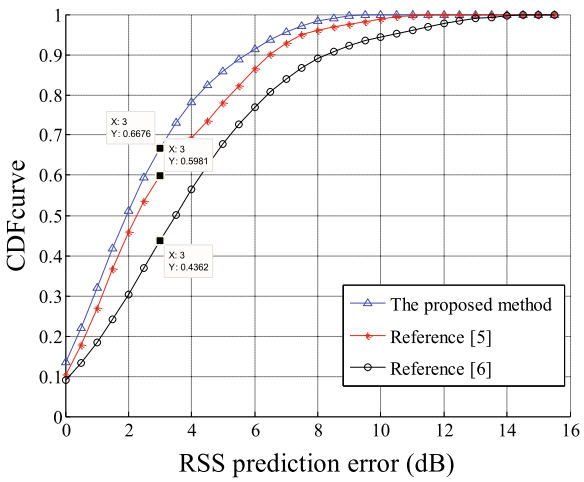



**Fig. 6.** CDF of errors for AP1

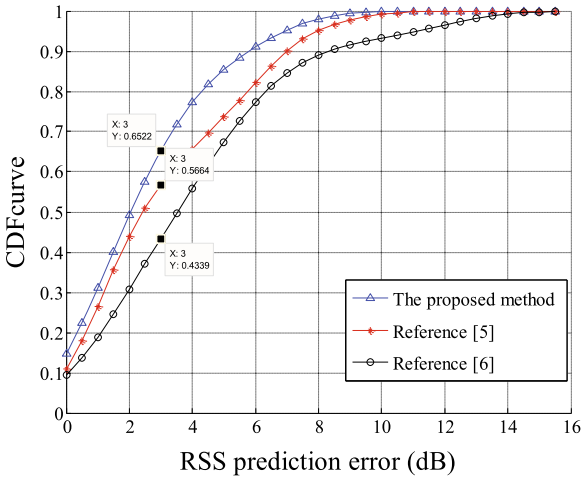



**Fig. 7.** CDF of errors for AP2

## 4 Conclusion

In this paper, we propose the adaptive-depth ray tree model, which can be used to adaptively construct a passive radio map for indoor WLAN intrusion detection. For one thing, we use the genetic algorithm to enhance the traditional quasi-3D ray-tracing model to depict the RSS variation under the indoor silence and intrusion scenarios with low labor and time cost. For another, six common signal features are allied to ensure the stability of RSS data and robustness of passive radio map. In future, we will continue to investigate a more effective passive radio map construction method to accurately locate multiple targets in the anonymous indoor WLAN environment.

**Acknowledgments.** This work is supported in part by the Fundamental Science and Frontier Technology Research Project of Chongqing (cstc2017jcyjAX0380).

## References

1. Youssef M, Mah M, Agrawala A. Challenges: device-free passive localization for wireless environments. In: ACM international conference on mobile computing and networking; 2007. p. 222–9.
2. Jin S, Choi S. A seamless handoff with multiple radios in IEEE 802.11 WLAN. IEEE Trans Veh Technol. 2014;63(3):1408–18.
3. Wang Q, Yigitler H, Jantti R, et al. Localizing multiple objects using radio tomographic imaging technology. IEEE Trans Veh Technol. 2016;65(5):3641–56.
4. Deak G, Curran K, Condell J, et al. Detection of multi-occupancy using device-free passive localization. IET Wirel Sens Syst. 2014;4(3):130–7.
5. Liu Z, Guo L, Tao W. Full automatic preprocessing of digital map for 2.5D ray tracing propagation model in urban microcellular environment. Waves Random Complex Media. 2013;23(3):267–78.

6. Jong YLCD, Herben MAHJ. Prediction of local mean power using 2-D ray-tracing-based propagation models. IEEE Trans Veh Technol. 2001;50(1):325–31.
7. Sabar NR, Ayob M, Kendall G, et al. A dynamic multiarmed bandit-gene expression programming hyper-heuristic for combinatorial optimization problems. IEEE Trans Cybern. 2015;45(2):217–28.
8. Erceg V, Rustako AJ, Roman R. Diffraction around corners and its effects on the microcell coverage area in urban and suburban environments at 900 MHz, 2 GHz, and 6 GHz. IEEE Trans Veh Technol. 1994;43(3):762–6.
9. Dutt V, Chaudhry V, Khan I. Different approaches in pattern recognition. Comput Sci Eng. 2011;1(2):32–5.
10. Queiroz A, Trintinalia LC. An analysis of human body shadowing models for ray-tracing radio channel characterization; 2016. p. 1–5.

# An Iris Location Algorithm Based on Gray Projection and Hough Transform

Baoju Zhang(✉) and Jingqi Fei

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
wdxyzbj@163.com

**Abstract.** In order to improve the performance of the existing iris location algorithm, a transform algorithm based on gray projection and Hough is proposed. The algorithm uses the grayscale transformation of the binary image to obtain a graph of the gray projection. At the same time, according to the value of the peak or trough in the graph, the maximum radius of the circle is obtained. The result of experiment shows that: The algorithm can get the parameters needed in Hough transform, which greatly improves the speed and accuracy of iris positioning.

**Keywords:** Iris recognition · Canny edge detection · Binarization · Grayscale projection · Hough transform

## 1 Introduction

With the development of society, the importance of identity recognition is increasingly evident. In recent years, in the fields of maintaining national security, aviation safety, financial security, social security, and network security, there is a need for more accurate, reliable, and more practical authentication methods for identifying and authenticating identity. However, relying on identity documents, user names, and identity authentication in the form of passwords is far from meeting the requirements of the information age for the validity of authentication and the accuracy of identification. As an important identification feature, iris has the advantages of uniqueness, stability, collectability, and non-invasion. One of the keys to the iris recognition algorithm is to accurately locate the iris region from the acquired iris image. It mainly includes the edge position of the pupil and the iris, the iris and the sclera, hereinafter referred to as the inner edge and the outer edge.

There are many articles on iris recognition at domestic and foreign, including Daugman [1], who used the characteristics of the inner and outer edge of the iris as an approximate circle, a circular difference operator is proposed to extract the edge of the iris. Wildes [2] uses a two-step method combining edge detection and Hough transform to locate the iris region. Both of these iris location methods have high accuracy, but both searching in three-dimensional space, positioning speed is slow and is unable to meet the requirements of real-time systems. Tisse proposes to extract iris features using a time-phase technique. Weiqi Yuan [3] uses the traditional gray projection method to roughly locate the pupil center and position and then uses Hough transform to


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_157

accurately position. When the iris image contains thick eyelashes, the gray projection method can easily misplace the eyelash position as the pupil position, so that the accuracy of this method is lower.

In order to overcome the above defects and further improve the iris recognition system, according to the characteristics of the iris itself, an iris location algorithm based on gray projection and Hough transform is proposed. This method can not only suppress noise interference, but also improve the accuracy of iris location.

## 2 Iris Localization Algorithm

### 2.1 Iris Image Smoothing and Edge Extraction

The collected iris images have different levels of interference. Filtering before the iris boundary location can help to eliminate the effect of interference on the boundary location. The image should be smoothed. On the one hand, it can highlight a large area, a low-frequency component and a trunk part, on the other hand, smoothed can suppress image noise and interfere with high-frequency components. At the same time, it can reduce a sudden gradient, and improve image quality [4]. In addition, taking into account the edge extraction of the grayscale image of the iris, the inner and outer boundaries of the iris belong to the detailed information. In order to not only retain the iris edge information but also effectively eliminate high-frequency interference, Gaussian filter templates are used for smoothing. The formula is as follows:

$$H_{ij} = \frac{1}{2\pi\delta^2} \exp\left( -\frac{[i-(k+1)]^2 + [j-(k+1)]^2}{2\delta^2} \right) \tag{1}$$

When Gaussian template filtering is used, the point farther from the center point will have less effect on the smoothing effect. According to the distance, selecting a weighting coefficient can construct Gaussian templates of different sizes whose filter parameters can be determined by the variance of the Gaussian function. The Gaussian template can well protect the contour information of the iris region while filtering the noise, so as not to cause the edges to be too fuzzy. Figure 1a is the original. The result is shown in Fig. 1b.

For effective edge extraction, Canny edge detection is used in this paper. In the first step, the gradient intensity and direction of each pixel in the image are calculated for the smoothed image. Edges in the image can point in all directions, so the Canny algorithm uses four operators to detect horizontal, vertical, and diagonal edges in the image. The operator of the edge detection (such as Roberts, Prewitt, and Sobel) returns the first derivative value of the horizontal $G_x$ and vertical $G_y$ directions, thereby determining the gradient $G$ and the direction theta of the pixel [5]. The formula is as follows:

$$G = \sqrt{G_x^2 + G_y^2} \tag{2}$$

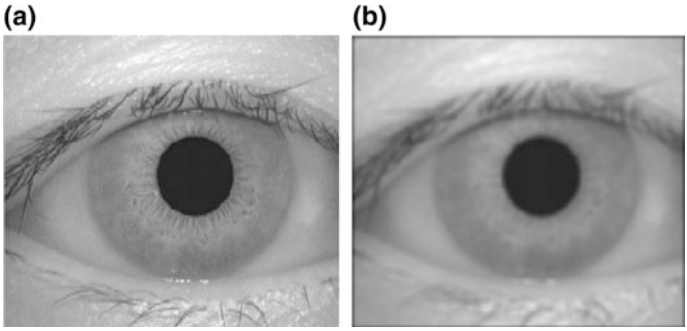



**Fig. 1.** (a) Original image (b) Gaussian image

$$\theta = \arctan\left(G_y / G_x\right) \tag{3}$$

In the second step, non-maximum suppression is performed to compare the gradient intensity of the current pixel with two pixels in the positive and negative gradient directions. If the gradient intensity of the current pixel is the largest compared to the other two pixels, the pixel remains as an edge point, otherwise the pixel will be suppressed. The third step, it is dual threshold detection. After non-maximal suppression, the remaining pixels can more accurately represent the actual edges in the image. However, there are still some edge pixels caused by noise and color changes. In order to solve this type of edge pixels, weak gradient values are used to filter edge pixels, and edge pixels with high gradient values are retained, which is achieved by selecting high and low thresholds [6]. If the edge pixel's gradient value is higher than the high threshold, it is marked as a strong edge pixel; if the edge pixel's gradient value is less than the high threshold and greater than the low threshold, it is marked as a weak edge pixel; if the edge pixel's gradient value is less than low thresholds, it will be suppressed. The choice of threshold depends on the content of a given input image. The result is shown in Fig. 2.

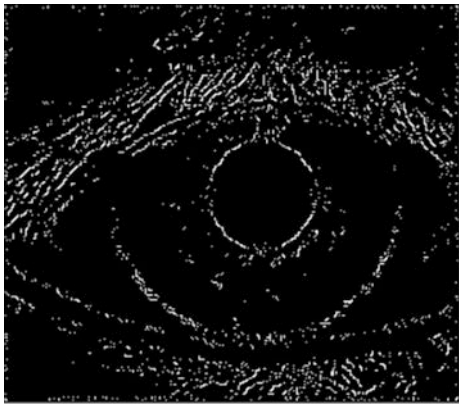


**Fig. 2.** Edge contour extraction map of the entire eye

### 2.2 Iris Boundary Localization

To locate the inner circle, we need to extract the inner circle. By setting the threshold, the image is binarized and the pupil is extracted. The formula is as follows:

$$\begin{cases} B(m,n) = 1, & \text{if} \quad I(m,n) > \alpha \\ B(m,n) = 0, & \text{if} \quad I(m,n) \leq \alpha \end{cases} \tag{4}$$

According to the formula, in this paper, $\alpha$ has a value of 47, the binarized picture obtained, at the same time, we should extract the outline, as shown in Fig. 3a, b.

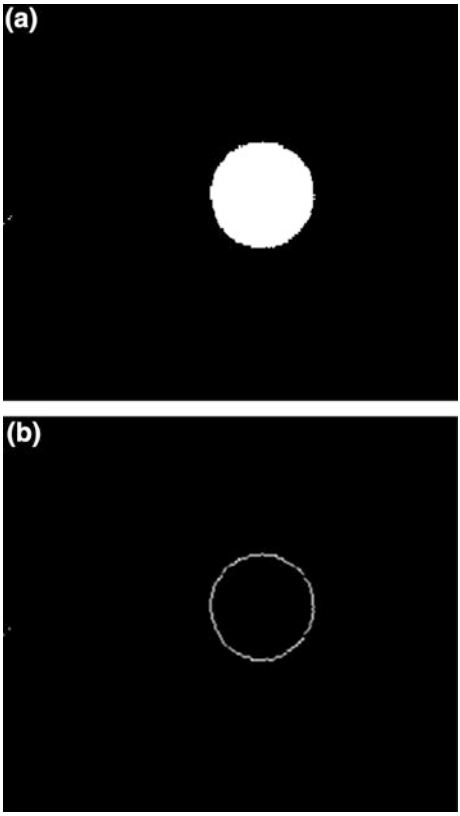



**Fig. 3.** (a) and (b) show further extraction of contours by binarization

Thus leaving only the outline of the inner circle. Then we can use the information of the gray change to determine the parameters of the pupil. The algorithm is as follows:

Step 1. Find the horizontal grayscale projection curve. The image is represented by a matrix of $M * N$, then $M$ can be regarded as a matrix, $i$ and $j$ are the matrix rows and columns, and the horizontal grayscale projection curve of the image, then the

abscissa is the row number of each row of the matrix, and the ordinate is a matrix [7]. The sum of the row elements, that is, the sum of the pixel values for each row in the picture. Actually, the curve contains the change in the sum of the pixels in each row of the image.

Step 2. Based on the calculation of the gray curve, the peaks and troughs of the curve are obtained, as shown in Fig. 4a. Since the black gray value is 0 and the white gray value is 255, the region coordinates information with a smaller gray value is included in the trough.

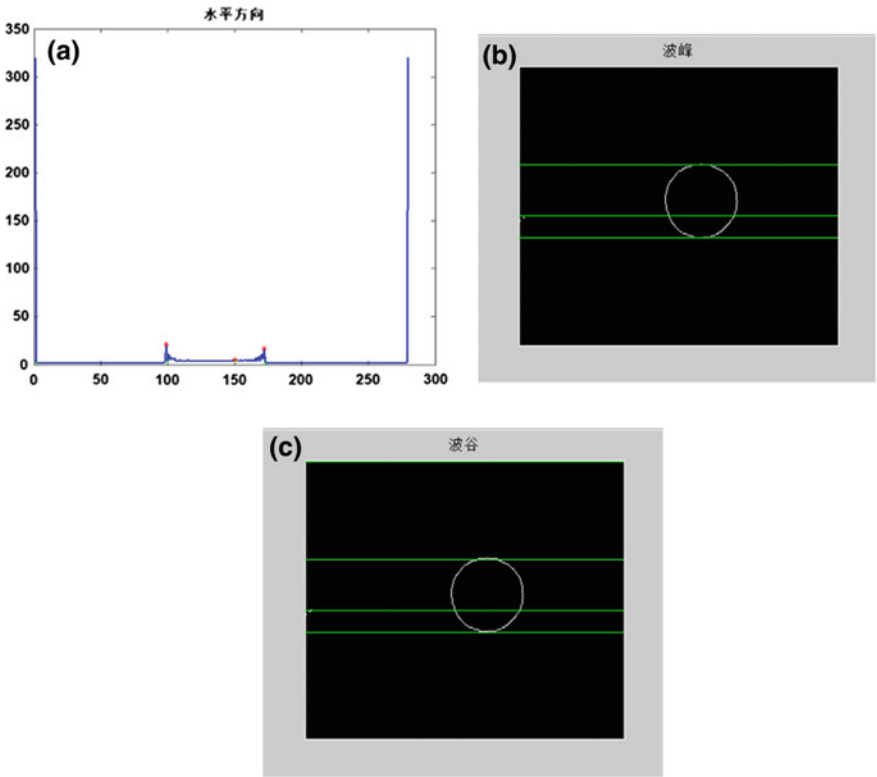



**Fig. 4.** (a) Shows a chart of pixels in the horizontal direction (b) and (c) show the peaks and valleys of the outline according to the chart

Step 3. Drawing the line information of the image represented by the abscissa of the wave trough, as shown in Fig. 4b, c, and the maximum value of the radius can take the maximum difference *r*_max in the crest or trough. According to the prior knowledge, the minimum radius can be obtained by subtracting 13 from the maximum difference *r*_max.

The Hough transform is used to locate the circle parameter. A circle is a typical and regular geometric shape with fewer parameters, namely the center coordinates and

radius [8]. Therefore, detection of a circle becomes a process of voting on a parameter group. Define a three-dimensional array as

$$H(x_m, y_n, r) = \sum_{j=1}^{n} h(x_m, y_m, x_n, y_n, r) \tag{5}$$

Here, $H(x_m, y_n, r)$ is an accumulator corresponding to the parameter group $(x_m, y_j, r)$ formed by the center point coordinate and the radius size, and the obtained value is used to accumulate the votes of the group of parameters. The number of votes is represented by the number of boundary points passed by the circle drawn by the parameter. If the edge point falls on the circle corresponding to the parameter group, it is equivalent to the edge point casting a vote for the parameter group, and the corresponding array element value is increased by 1, otherwise the corresponding array element value is not changed [9]. The formula is as follow:

$$h(x_m, y_m, x_n, y_n, r) = \begin{cases} 1 & g(x_m, y_m, x_n, y_n, r) = 0 \\ 0 & g(x_m, y_m, x_n, y_n, r) = 1 \end{cases} \tag{6}$$

In the formula, $g(x_m, y_m, x_n, y_n, r) = (x_m - x_n)^2 + (y_m - y_n)^2 - r^2$ is a decision function that satisfies the circular equation of the parameter $(x_n, y_n, r)$. When $g(x_m, y_m, x_n, y_n, r) = 0$, the center of the circle is $(x_n, y_n)$, and the boundary circle with the radius $r$ passes through the edge point $(x_m, y_m)$, indicating that the point will vote for the parameter $(x_n, y_n, r)$. After passing through all votes of columnar-(5) statistic, the circle with the most votes will be used to determine the circle of the boundary equation [10]. In the end, the results obtained are shown in Fig. 5.

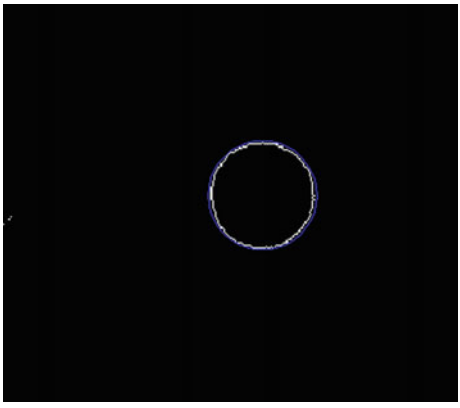


**Fig. 5.** Inner edge contour extraction image

## 3 Experiments Results

The computer CPU used in the experiment was clocked at 2.01 GHz and the memory was 1 G. The programming tool was MATLAB 2014a. The iris images used were all from the CASIA 1.0 database of the Chinese Academy of Sciences. According to the method mentioned above, the iris is positioned inside and outside the circle. The final result is shown in Fig. 6.

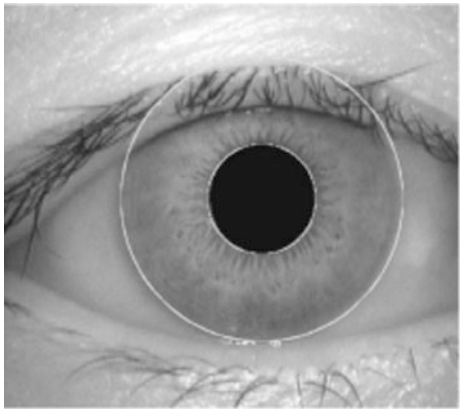


**Fig. 6.** The final result

## 4 Conclusions

Based on previous study of iris recognition, this paper proposes an algorithm which can locate iris by grayscale changes. Firstly, image which should be detected is binarized and the outline is extracted. Then, grayscale graph is drawn by taking advantage of gray change of the image. Finally, according to the peaks and valleys in the graph, the parameters required by the Hough transform are obtained. Experiments show that the method used in this paper reduces Hough transform traversal, which can make iris localization improve efficiently.

**Acknowledgments.** This paper is supported by Natural Youth Science Foundation of China (61501326, 61401310), the National Natural Science Foundation of China (61731006) and Natural Science Foundation of China (61271411). It also supported by Tianjin Research Program of Application Foundation and Advanced Technology (15JCZDJC31500), and Tianjin Science Foundation (16JCYBJC16500). This work was also supported by the Tianjin Higher Education Creative Team Funds Program.

## References

1. Daugman J. How iris recognition work. IEEE Trans Circ Syst Video Technol. 2004;14 (1):21–30.
2. Wildes RP. Iris recognition an emerging biometric technology. Proc IEEE. 1997;85 (9):1348–63; 2000;15(10):939–57.
3. Yuan W, Lin Z, Xu L. A rapid iris location method based on the structure of human eyes. In: 27th annual international conference; 2006. p. 17–8.
4. Daugman J. New methods in Iris recognition. IEEE Trans Syst Man Cybern Part B Cybern. 2007;37(5):1167–75.
5. Kawaguchi T, Rizon M. Iris detection using intensity and edge information. Pattern Recogn. 2003;36:549–62.
6. Wang J-G, Sung E. Study on eye gaze estimation. IEEE Trans Syst Man Cybern Part B. 2002;32(3):332–50.
7. Daugman J. High confidence visual recognition of person by a test of statistical independence. IEEE Trans Pattern Anal Mach Intell. 1993;15(11):1148–61.
8. Canny J. A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell. 1986;8(6):679–714.
9. Williams GO. Iris recognition technology. IEEE Aerosp Electron Syst Mag. 1997;12(4): 23–9.
10. Park KR, Kim J. A real-time focusing algorithm for Iris recognition camera. IEEE Trans Syst Man Cybern Part C. 2005;35(3):441–4.

# Robust Tracking via Dual Constrained Filters

Bo Yuan(✉), Tingfa Xu, Bo Liu, Yu Bai, Ruoling Yang, Xueyuan Sun, and Yiwen Chen

School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China
{bityuanbo,ciom_xtf1}@bit.edu.cn

**Abstract.** In this paper, we propose a novel correlation filter framework constrained by dual filters. The Minimum Output Sum of Squared Error (MOSSE) filter is the unbiased estimate of the filter which easily to cause overfitting. The trained filter by linear ridge regression is the biased estimate of the filter which can deal with the overfitting. We combine the advantages of the two filters to constrain the trained filter which optimizes our model. To deal with background clutter, clipping background patches around the target position up, down, left, and right, we add the cropped background patches to the learning filter. To overcome the challenge of occlusion, we introduce a novel criterion, Average Peak-to-Correlation Energy (APCE). Extensive experiments on the CVPR 2013 Benchmark well demonstrate that our tracker can effectively solve the background clutter and occlusion. Both quantitative analysis and qualitative analysis show that our tracker outperforms some state-of-the-art trackers.

**Keywords:** Dual filters · Background clutter · Occlusion · CVPR 2013 Benchmark

## 1 Introduction

In recent years, visual tracking has attracted the attention of researchers as an important part of machine vision field. A large number of tracking algorithms have been proposed [1–6]. In general, visual tracking is divided into single-target tracking and multi-target tracking. The paper focuses on the theoretical exploration of single-target tracking. The problem of object tracking is to estimate the trajectory of an object in a video sequence when given object's initial state (position and size) in the first frame. There are still many challenges that will affect the target tracking effect although some trackers have demonstrated superior performance in both precision and success rate, such as illumination variation,


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_158

scale variation, occlusion, deformation, background clutter, and fast motion. So how to design a robust target tracking algorithm is still a difficult problem in the current target tracking field. In this paper, we are mainly focused on designing a novel visual tracking approach to correlation filtering and solving the problems of background clutter and occlusion.

Correlation Filter (CF) tracking algorithms have attracted the attention of researchers in recent years, due to their excellent performance. Bolme et al. [1], proposed the MOSSE tracker, it introduces correlation filtering into the target tracking field for the first time. The Fourier transform is used to directly convert the operation to the frequency domain, which greatly reduces the amount of computation. Henriques et al. [2], proposed Circulant Structure with Kernel (CSK) has superior performance. Furthermore, many CF-based trackers have been proposed to solve a kind of tracking challenges such as DSST proposed in [3] to deal with scale variation and KCF [4] proposed a CF-based tracker combining multi-channel features which improves tracking performance. Although these methods above have greatly improved in both precision and success rate, they still can not solve the tracking drift, due to scale variation, occlusion, and deformation.

Current correlation filter trackers still have some shortcomings; the first one is that tracking box is fixed, and the tracker cannot adapt to the scale change of the target. In other words, the model will be polluted due to the introduction of a lot of background information. And model training is incorrect due to the introduction of few of background information. The second shortcoming is that model update learning rate is fixed. That is, the filter parameters would be learned from occluded objects when occlusion occurs, which may lead to wrongly update of the model. As time passes, it will be tracking drift.

To solve the challenging problem of occlusion, Bolme et al. [1] proposed to adaptively update the model based on Peak-to-Sidelobe Ratio (PSR), which measures the strength of a correlation peak and can be used to detect occlusions or tracking failure. The confidence degree (APCE) is proposed in [6]; in the method, both the response peaks and the APCE are a certain proportion and greater than their respective historical averages. Then update the tracking model.

In this paper, we mainly focus on designing a robust CF-based tracker to solve the above-mentioned problems which are the occlusion and background clutter. We aim to build a comprehensive correlation filter that can handle the background clutter and occlusion. We propose a novel correlation filter framework constrained by dual filters. The Minimum Output Sum of Squared Error filter is the unbiased estimate of the filter which easily to cause overfitting. The trained filter by linear ridge regression is the biased estimate of the filter which can deal with the overfitting. We combine the advantages of the two filters to constrain the trained filter which optimize our model. To deal with background clutter, clipping background patches around the target position up, down, left, and right, we add the cropped background patches to the learning filter. To overcome the challenge of occlusion, we introduce a novel criterion, APCE. Our

method shows superior tracking performance than the existing CF, especially when dealing with the tracking problems of occlusion and background clutter.

## 2 Proposed Method

### 2.1 Dual Constrained Filters

Bolme et al. [1], proposed the MOSSE tracker, which trained filter by the Minimum Output Sum of Squared Error. Henriques et al. [2], proposed the CSK tracker, which trained filter by ridge regression. We propose a novel correlation filter framework constrained by dual filters, which combined the advantages of the above filters. Then we will introduce our model in detail. The goal of the CF trackers is to learn a discriminative correlation filter which can be applied to the region of interest in consecutive frames to infer the location of the target. It can be seen to find a linear regression function $f(x) = \mathbf{w}^T\mathbf{x}$ to minimize the training sample error. Where $x$ is a training sample, and the learned correlation filter is represented by the vector $\mathbf{w}$. In order to achieve this with dual constrained filters, the model can be optimized as follows:

$$\min_{\mathbf{w}_1} \sum_i (f(x_i) - y_i)^2 + \lambda_1 \parallel \mathbf{w}_1\|^2 + \lambda_2||\mathbf{w}_1 - \mathbf{w}_2||^2 \tag{1}$$

where $L(y_i, f(x_i)) = (y_i - f(x_i))^2$ express as a Loss function, and $y_i$ is the corresponding regression label to the sample $x_i$. And $\lambda_1$ is a regularization coefficient to prevent the model from overfitting. $\lambda_2$ represents penalty coefficients for double-constrained filters. $\mathbf{w}_2$ is the trained filter by the Minimum Output Sum of Squared Error. The optimization model can be seen as a ridge regression problem, which has a closed-loop solution. Equation 1 can be solved as follows:

$$\mathbf{w_1} = (\mathbf{X}^T\mathbf{X} + \lambda_1\mathbf{I} + \lambda_2\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{y} + \lambda_2\mathbf{w_2}) \tag{2}$$

where $\mathbf{X}$, $\mathbf{y}$ represent the sample matrix and the label matrix, respectively. $\mathbf{I}$ represents a unit matrix. However, the inversion process in Eq. 2 is a problem to solve. We can solve Eq. 2 easily by introducing the property (Eq. 3) that the circulant matrix is diagonalizable in the Fourier domain.

$$\mathbf{x} = \mathbf{F}diag\left(\widehat{x}\right)\mathbf{F}^H \tag{3}$$

The solution to Eq. 2 is as follows:

$$\widehat{\mathbf{w}}_1 = \frac{\widehat{\mathbf{x}}^* \odot \widehat{\mathbf{y}} + \lambda_2\widehat{\mathbf{w}}_2}{\widehat{\mathbf{x}}^* \odot \widehat{\mathbf{x}} + \lambda_1 + \lambda_2} \tag{4}$$

In the case of no-linear regression, kernel trick $f(\mathbf{z}) = \mathbf{w}^T\mathbf{z} = \sum_{i=1}^n \alpha_i k(\mathbf{z}, \mathbf{x}_i)$ is applied to allow more powerful classifier. For the most commonly used kernel functions, the circulant matrix trick can also be used. Therefore, the dual domain resolves as follows:

$$\widehat{\alpha} = \frac{\widehat{\mathbf{y}} + \lambda_2\widehat{\beta}}{\widehat{\mathbf{k}} + \lambda_1 + \lambda_2} \tag{5}$$

where $\widehat{\beta}$ is the solution of the dual domain of $\widehat{\mathbf{w}}_2$. And $\widehat{\beta} = \frac{\widehat{\mathbf{y}}}{\widehat{\mathbf{k}}}$. Therefore, Eq. 5 can be rewritten as follows:

$$\widehat{\alpha} = \frac{\widehat{\mathbf{k}}^* \odot \widehat{\mathbf{y}} + \lambda_2 \widehat{\mathbf{y}}}{\widehat{\mathbf{k}}^* \odot \widehat{\mathbf{k}} + \lambda_1 \widehat{\mathbf{k}} + \lambda_2 \widehat{\mathbf{k}}} \tag{6}$$

In order to predict the position of target in the next frame and the location corresponding to the peak response is the position of the target, the target detection formula can be expressed as follows:

$$\widehat{f} = \widehat{\mathbf{k}} \odot \widehat{\alpha} \tag{7}$$

The above is a correlation filter general framework proposed by us, suitable for other related filter trackers. To solve the background clutter, we introduce a context-aware correlation filter tracker [8] in our model. Clipping background patches around the target position up, down, left, and right, we add the cropped background patches to the learning filter. Equation 1 can be rewritten as:

$$\min_{\mathbf{w}_1} = ||\mathbf{A}_0\mathbf{w}_1 - \mathbf{y}||^2 + \lambda_1||\mathbf{w}_1||^2 + \lambda_2||\mathbf{w}_1 - \mathbf{w}_2||^2 + \lambda_3 \sum_i^k ||\mathbf{A}_i\mathbf{w}_1||^2 \tag{8}$$

where $\mathbf{A}_0$ represents the cyclic matrix generated by the cyclic shift of the base sample. $\mathbf{A}_i$ is the cyclic matrix generated by the cyclic shift of the background patch. The number of background patches is represented as $k$. Similar as Eqs. 1, 8 can be solved as:

$$\mathbf{w_1} = (\mathbf{A}_0^T\mathbf{A}_0 + \lambda_1\mathbf{I} + \lambda_2\mathbf{I} + \lambda_3 \sum_i^k \mathbf{A}_i^T\mathbf{A}_i)^{-1}(\mathbf{A}_0^T\mathbf{y} + \lambda_2\mathbf{w}_2) \tag{9}$$

which can also be rewritten as:

$$\widehat{\mathbf{w}}_1 = \frac{\mathbf{a}_0^* \odot \mathbf{y} + \lambda_2 \widehat{\mathbf{w}}_2}{\mathbf{a}_0^* \odot \mathbf{a}_0 + \lambda_1 + \lambda_2 + \lambda_3 \sum_i^k \mathbf{a}_i^* \odot \mathbf{a}_i} \tag{10}$$

### 2.2 Model Updating

The traditional CF-based trackers update the target model with a fixed learning rate. However, the model will be degenerated if a fixed high learning rate to update the model when the target is occluded. To overcome this problem, we introduce a novel criterion, APCE, which indicates the degree of oscillation of the response map. And the confidence degree (APCE) is defined as follows:

$$APCE = \frac{|F_{\max} - F_{\min}|^2}{mean\left(\sum_{w,h}(F_{\omega,h} - F_{\min})^2\right)} \tag{11}$$

where $F_{\max}$ represents the peak of the response map. $F_{\min}$ is the minimum value of the response map. The response value on the coordinate $(w, h)$ is expressed as $F_{\omega,h}$. We update the model only when the response peaks and APCE of the current frame are greater than their respective historical averages by a certain percentage $\beta_1$, $\beta_2$.

$$\begin{cases} response_{\max} \geq \beta_1 \cdot mean\_response \\ APCE \geq \beta_2 \cdot mean\_APCE \end{cases} \tag{12}$$

where $mean\_response$ and $mean\_APCE$ represent the historical average of the response peak and the historical average of APCE, respectively. Then we update the model with a learning rate parameter $\eta$ as

$$\begin{cases} \widehat{\alpha}_t = (1-\eta) \cdot \widehat{\alpha}_{t-1} + \eta \cdot \widehat{\alpha}_t \\ \widehat{\mathbf{x}}_t = (1-\eta) \cdot \widehat{\mathbf{x}}_{t-1} + \eta \cdot \widehat{\mathbf{x}}_{tt} \end{cases} \tag{13}$$

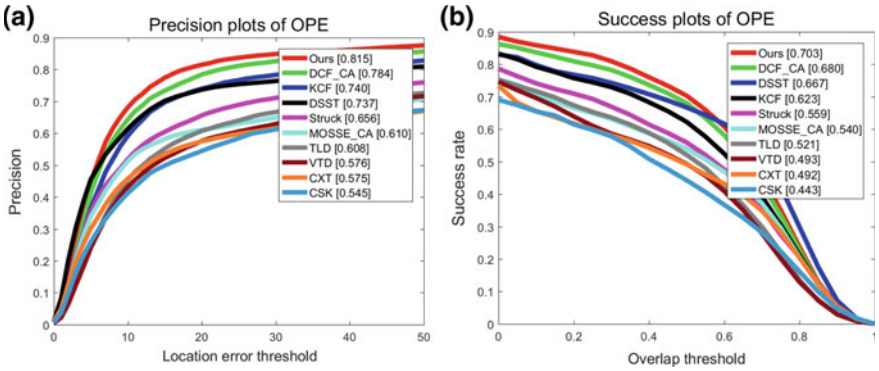



**Fig. 1.** Overall quantitative evaluation on CVPR 2013 Benchmark: precision plots (**a**) and success plots (**b**).

## 3 Experimental

### 3.1 Experimental Setup

We consider doing some work on feature extraction; the powerful features we use is HOG features. The proposed method in this paper is implemented in MATLAB 2016a. We perform the experiments on a PC with Intel i5-6500 CPU (3.2 GHz) and 8 GB RAM memory. We test the performance of the proposed tracker with the total 51 video sequences using in the CVPR 2013 benchmark [7] and compare with the top nine state-of-the-art trackers which include DSST [3], DCF_CA [8], KCF [4], CSK [2], MOSSE_CA [8], Struck [9], TLD [10], VTD

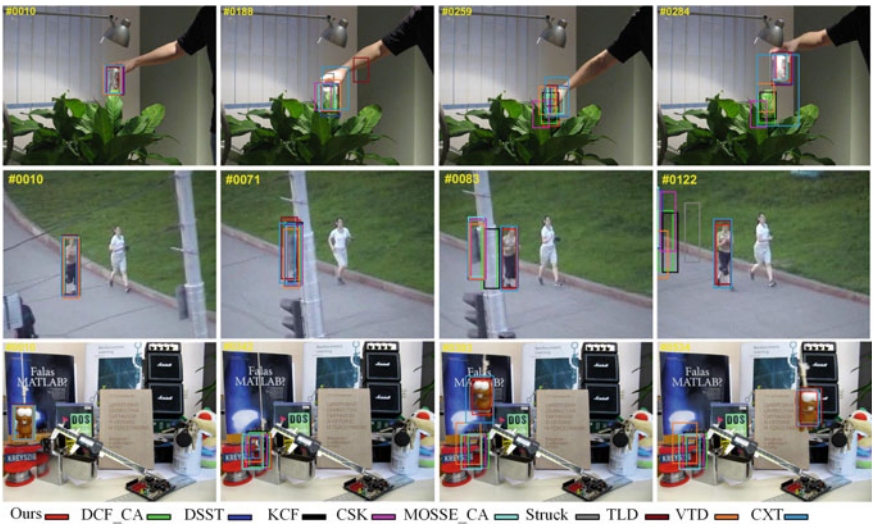



**Fig. 2.** Qualitative results of ten trackers on the video sequence with Coke, Jogging-1, and Lemming, which in the challenges of occlusion.

[11], and CXT [12], where Struck, VTD, and CXT are the three best-performed ones demonstrated in the CVPR 2013 Benchmark.

The parameters are set as follows. The regularization parameters $\lambda_1$, $\lambda_2$, $\lambda_3$ are set to be $10^{-4}$, 0.0045 and 25. The learning rate parameter $\eta$ and search box (padding) are set to 0.015 and 2, respectively. $\beta_1$, $\beta_2$ are set to be 0.5 and 0.4.

### 3.2 Experimental Results

Figure 1 contains the precision plots and success plots that overall quantitative evaluation with One-Pass Evaluation (OPE). Both precision plots and success plots show that our tracker is more robust than some state-of-the-art trackers in terms of the total 51 video sequences in CVPR 2013 Benchmark. Our trackers have achieved 0.703 and 0.815 in success rate and precision, respectively, which both rank the first. The baseline tacker KCF has achieved 0.623 and 0.740 in success rate and precision, respectively. Our tracker is better than the baseline tracker.

Seen in Fig. 2 the qualitative results of 10 trackers on the video sequence with Coke, Jogging-1, and Lemming in the attribute of occlusion.

As seen in Fig. 3, we give the results of our tracker in CVPR 2013 Benchmark on the attributes. Our tracker have achieved 0.684 and 0.804 in success rate and precision in the attribute of occlusion, respectively, which both rank the first.

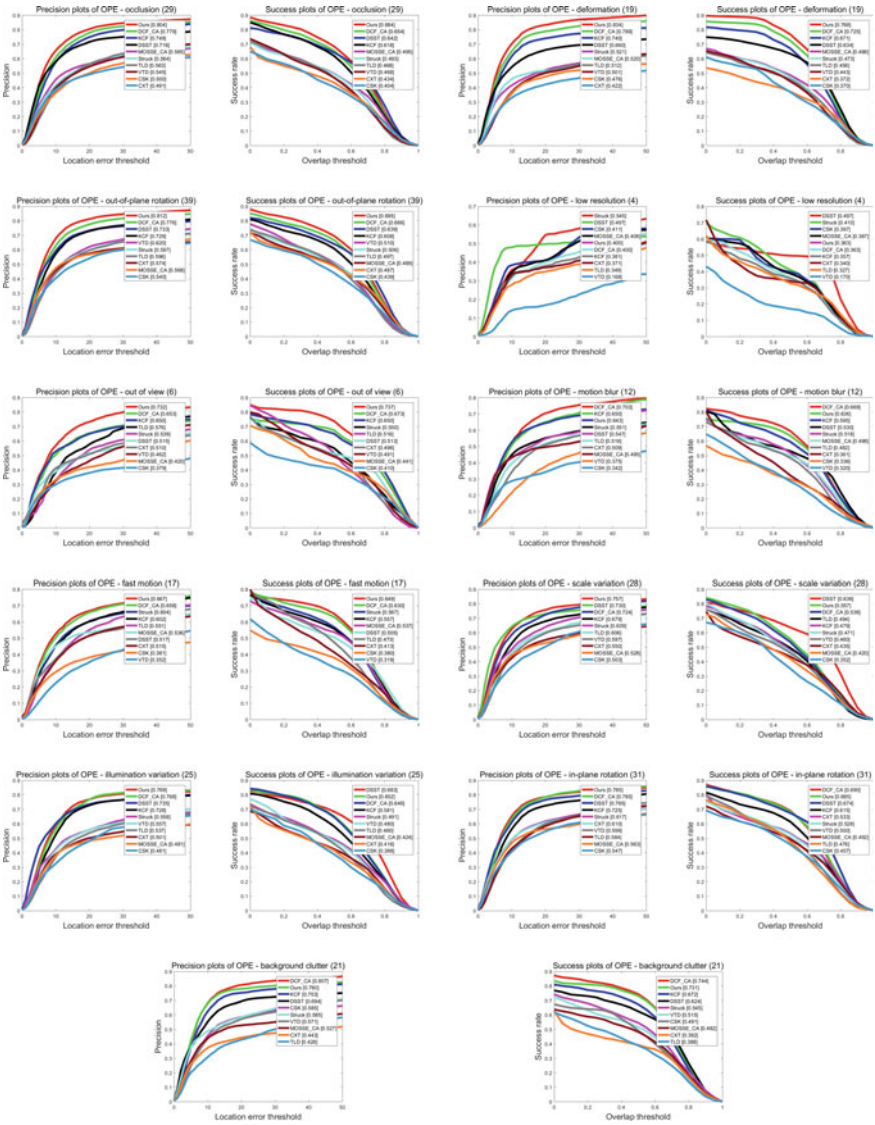



**Fig. 3.** Attribute-based comparison of our tracker with some state-of-the-art trackers in precision rate and success rate.

## 4 Conclusion

This paper has shown that the correlation filter framework constrained by dual filters, which combine the advantages of the MOSSE filter and the linear ridge regression filter, has better robustness with excellent performance. With adding the cropped background patches to the learning filter and introducing a

novel criterion, APCE, our designed tracker can effectively solve the background clutter and occlusion. Extensive experiments on the CVPR 2013 Benchmark well demonstrate that our tracker outperforms some state-of-the-art trackers in both quantitative and qualitative analysis. The correlation filter framework constrained by dual filters can also be extended to other similar CF-based trackers to improve the robustness and performance.

**Acknowledgments.** This work was supported by the Major Science Instrument Program of the National Natural Science Foundation of China under grant 61527802, and the General Program of National Nature Science Foundation of China under grants 61371132 and 61471043.

## References

1. Bolme DS, Beveridge JR. Visual object tracking using adaptive correlation filters. In: IEEE computer vision and pattern recognition; 2010. p. 2544–50.
2. Henriques JF, Rui C, Martins P. Exploiting the circulant structure of tracking-by-detection with kernels. In: Computer vision—ECCV. Berlin: Springer; 2012. p. 702–15.
3. Danelljan M, Gustav H. Discriminative scale space tracking. IEEE Trans Pattern Anal Mach Intell. 2017;39(8):1561–75.
4. Henriques JF, Rui C, Martins P. High-speed tracking with kernelized correlation filters. IEEE Trans Pattern Anal Mach Intell. 2015;37(3):583–96.
5. Li Y, Zhu J. Adaptive kernel correlation filter tracker with feature integration, vol. 8926. Springer; 2014. p. 254–65.
6. Wang MM, Liu Y, Huang ZY. Large margin object tracking with circulant feature maps. In: IEEE computer vision and pattern recognition; 2017. p. 4800–8.
7. Wu Y, Lmi J, Yang MH. Online object tracking: a benchmark. In: IEEE computer vision and pattern recognition; 2013. p. 2411–8.
8. Mueller M, Smith N, Ghanem B. Context-aware correlation filter tracking. In: IEEE computer vision and pattern recognition; 2017. p. 1387–95.
9. Sunando S. Struck: structured output tracking with kernels. IEEE Trans Pattern Anal Mach Intell. 2015;38(10):2096–109.
10. Kalal Z, Mikolajczyj K, Matas J. Tracking-learning-detection. IEEE Trans Pattern Anal Mach Intell. 2012;34(7):1409–22.
11. Kwon J, Lee KM. Visual tracking decomposition. In: IEEE computer vision and pattern recognition; 2010. p. 1269–76.
12. Medioni G. Exploring supporters and distracters in unconstrained environments. In: IEEE computer vision and pattern recognition; 2011. p. 1177–84.

# Grid-Based Monte Carlo Localization for Mobile Wireless Sensor Networks

Qin Tang(✉) and Jing Liang

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China
tangqin0228@163.com

**Abstract.** Localization is an important requirement for wireless sensor networks (WSNs), but the inclusion of GPS receivers in sensor network nodes is often too expensive. Therefore, many solutions focus on static networks and do not consider mobility. In this paper, we analyze the Monte Carlo location (MCL) algorithm and propose an improved method—grid-based MCL. It applies the mobility of nodes to reduce the sampling area and to build an internal grid to predict the behavior of nodes. We investigate the properties of our technology and analyze its performance. The simulation and analysis show that the proposed grid-based MCL not only reduces localization error, but also improves the sampling efficiency.

**Keywords:** WSNs · Grid-based MCL · Mobility

## 1 Introduction

Location awareness is important for wireless sensor networks since many applications such as environment monitoring, vehicle tracking, and mapping depend on knowing the locations of sensor nodes [1,2]. Localization denotes the process of identifying the own position in space, but placing GPS receivers in every node or manually configuring locations is not cost-effective. Therefore, localization schemes for sensor networks typically let a small number of seed nodes whose location and protocols are already known(e.g., they are all equipped with GPS receivers) [3] to broadcast their location messages around and then estimate the location of other nodes according to the messages they received.

Solutions for localization in WSNs can be classified into two groups of algorithms: range-based and range-free [4–6]. Range-based localization means that distances between sensor nodes are estimated by using some physical properties of communication signals [7,8], while range-free algorithms are often based only on connectivity. A prominent example of range-free is the Monte Carlo localization (MCL). MCL is designed for applications in which all nodes are able to move freely in the deployment area. In certain applications, one can assume that nodes


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_159

are mainly moving on a group of paths [9]. The authors in [10] proposed an path-oriented approach (PO_MCL). It directly applies the original MCL algorithm. PO_MCL builds grids for static paths but requires additional hardware (magnetometer) on the nodes, and it may cost many times of sampling to find a valid sample. MCB [11,12] was proposed to reduce the sampling area. In this paper, we propose grid-based mobility behavior and present an adapted solution based on the MCL algorithm. We not only use the idea of MCB to improve sampling efficiency, but also use a forecast grid to divide the application area into cells to prognosticate the node's direction of movement. The grid is updated dynamically based on new observations from seed nodes. We show that by scheming anchor boxes and the grid, a better sample prediction of node localization can be achieved.

The rest of the paper is organized as follows. In Sect. 2, we present the main idea of MCL and proposes grid-based MCL. In Sect. 3, we present our experimental results. Finally, we conclude our work in Sect. 4.

## 2 Grid-Based MCL

### 2.1 MCL Algorithm

There are two main steps of MCL localization algorithms.

• Prediction step: In this step, a new sample is drawn from a circular sampling area with radius $r_{sarea} = v_{\max} \times t_{check}$ around its current position given by a transition equation $p(l_t|l_{t-1})$. The probability of the current location based on the previous location estimation is given by a uniform distribution (1).

$$p\left(l_t|l_{t-1}\right) = \begin{cases} \frac{1}{\pi \times {v_{\max}}^2}, & d\left(l_t, l_{t-1}\right) \in [0, v_{\max}) \\ 0, & d\left(l_t, l_{t-1}\right) \notin [0, v_{\max}) \end{cases} \tag{1}$$

• Filtering step: In this step, the set generated by the prediction step is put into the filtering step which uses the observations to filter out impossible node locations from the sample set. Each node keeps track of its rst-hop neighbor seeds $S$ and of its second-hop neighbor seeds $T$. The filtering condition for a sample $l$ is given in Eq. (2).

$$filter\left(l\right) = \forall\ s \in S,\ d\left(l, s\right) \leq r \wedge \forall s \in T,\ r < d\left(l, s\right) \leq 2r \tag{2}$$

However, the MCL algorithm has a large sampling area and the sampling efficiency is not high. In order to obtain a sufficient number of valid samples, continuous sampling is required, which reduces positioning efficiency. For making better use of new observations from seed nodes, we propose a grid-based Monte Carlo positioning algorithm in the next subsection. In Sect. 3, we will show the results of the MCL algorithm simulation and compare it with the algorithm (centroid and APIT) in the static network.

### 2.2 The Key Idea of the Grid-Based MCL

The main process of this algorithm is similar to the MCL algorithm, with the main difference being the selection of the sampling area and the establishment of the prediction grid. The choice of sampling area lies in the construction of anchor boxes and sample boxes and then the use of node mobility to reduce the sampling area. The sampling area will be divided into cells by using the prediction grid. Based on the observation from the seed node, the grid is updated. As long as seed node information is available, the original MCL algorithm is performed except that the sample is assigned the weight of its corresponding grid cell. The grid-based MCL algorithm is shown in Algorithm 1.

**Algorithm 1** Grid-based MCL algorithm

**Require:** Initial $L_t = \{\ \}$
**Require:** Initial $|O_t| > 0$
  **if** $|O_t| > 0$ **then**
    **MCL**();
    **UpdateGrid**();
  **end if**
  **EstOnGrid**();

### 2.3 Theoretical Analysis of the Grid-Based MCL

• Selection of the sampling area: We use the idea of the MCB algorithm [6] to select the sampling area and then rely on the mobility of the node to narrow down the sampling area (Fig. 1).

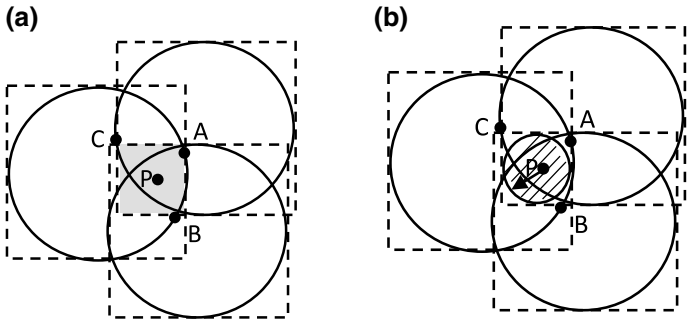



**Fig. 1.** Selection of the sampling area: **a** build sample boxes based on seed nodes; **b** reducing the sampling box based on node mobility

As shown in Fig. 1a, the superposition region formed by the seed nodes $A$, $B$, and $C$ contains the node $P$ to be measured, and then, the four endpoint

coordinate expressions of the superposition region are calculated according to Eq. 3. The maximum moving speed of the node $P$ to be measured is $v_{\max}$, and the other nodes to be tested are the same. Therefore, each node can construct a circular sampling area with its own position as the center and radius $v_{\max}$, as shown in Fig. 1b. These nodes have Markov characteristics, and the current position of each node is only related to the previous moment.

$$\begin{cases} X_{\max} = \max_i^n (x_i - r) \\ Y_{\max} = \max_i^n (y_i - r) \end{cases} \qquad \begin{cases} X_{\min} = \min_i^n (x_i - r) \\ Y_{\min} = \min_i^n (y_i - r) \end{cases} \tag{3}$$

• The construction of prediction grid: A node can only be located in one cell, and the entire grid cell can be divided into eight adjacent cells exactly. The neighboring cells are marked with the corresponding basic directions. We assign values for all grid cells, which indicate the probability of moving to this cell next, as shown in Fig. 2.

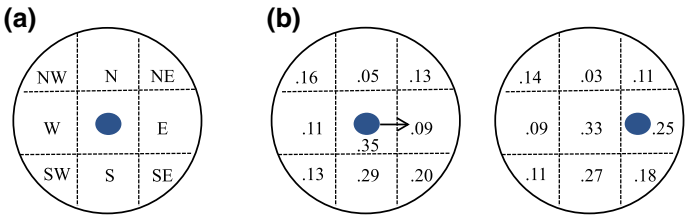
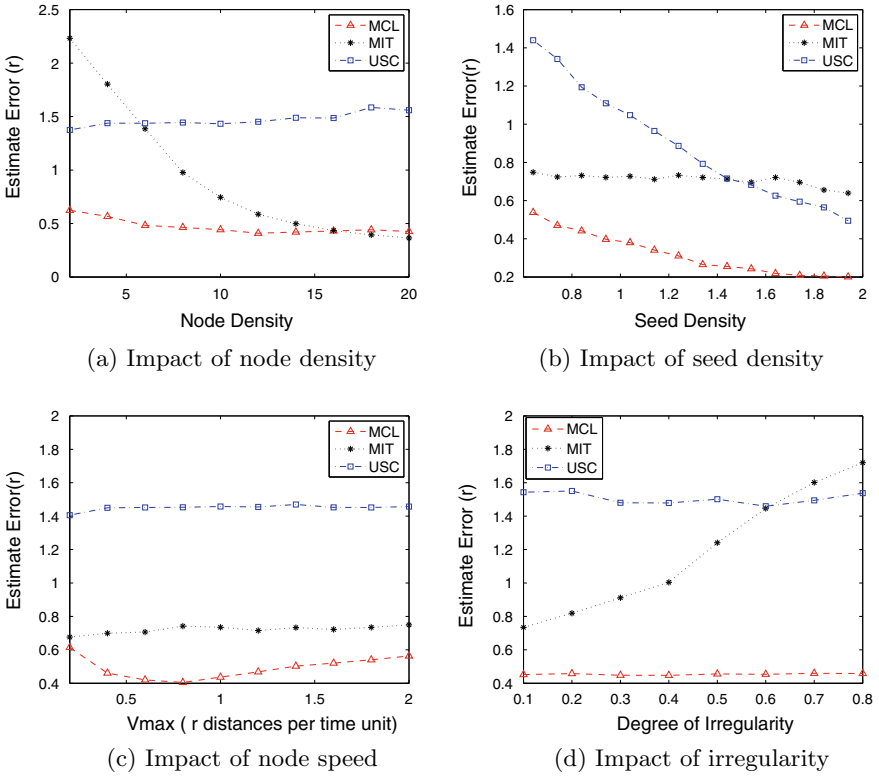



**Fig. 2.** The construction of prediction grid: **a** grid directions; **b** grid update process.

The size of the grid cells is an important parameter because they mainly determine the memory overhead of the grid-based MCL. We are adjusting the size of the grid cell based on $v_{\max}$ and $t_{check}$. Since the maximum distance that a node can travel between two location estimates is $d = v_{\max} \times t_{check}$, we also define the size of the grid unit as $2d/3$, as shown in Fig. 2a. Obviously, as the value of $d$ is smaller, the resolution of the grid is higher, and the positioning of the nodes can be more accurately mapped to the grid.

At the beginning, each of the grid cell is given an initial value of 0.1, because the information about the route has not yet been collected and any block of the nine cells of the grid totals 1. Based on the observation from the seed node, the grid is updated so that the probability of the cell to which the node has moved increases and the values of all other neighbor cells decrease. The value of the probability increase $\Delta_{inc}$ and the decrease $\Delta_{dec}$ is determined according to the values of cell $c_t$ in Eq. 4, where $\alpha$ is a tunable parameter for adjusting the increase level. A smaller value of $\alpha$ leads to a slower rate of increase in the probability that the grid convergence rate will slow down.

$$\Delta_{inc} = \frac{\alpha}{value}, \qquad \Delta_{dec} = \frac{\Delta_{inc}}{8} \tag{4}$$




(a) Impact of node density (b) Impact of seed density

(c) Impact of node speed (d) Impact of irregularity

**Fig. 3.** Comparison of the estimated error of different localization techniques

## 3 Simulation Results

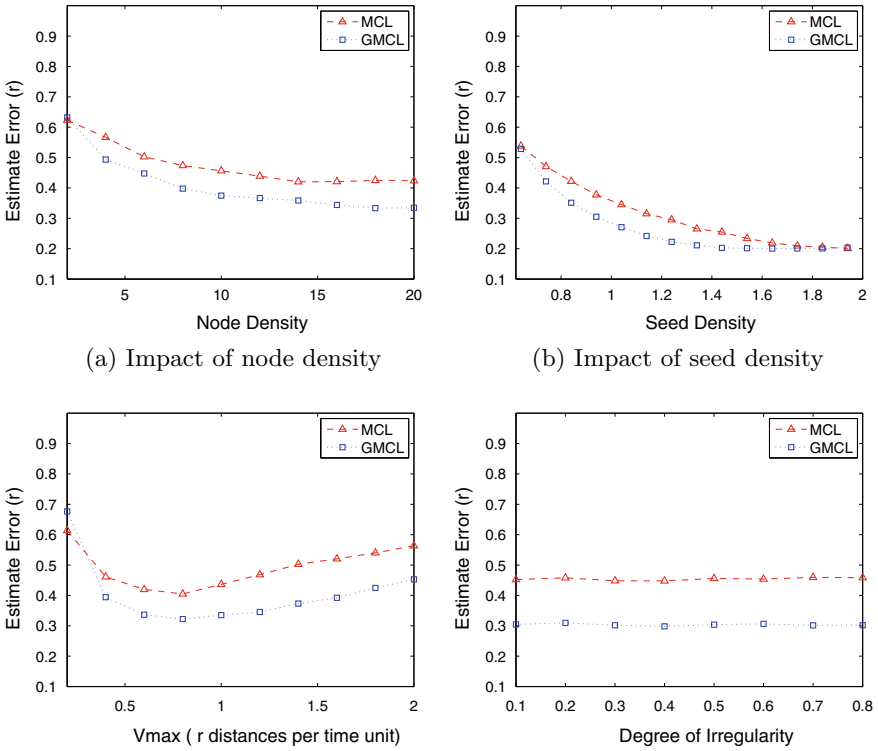
### 3.1 Analysis of MCL

We compare the MCL algorithm with the positioning algorithms (centroid algorithm and APIT algorithm) in the static network under the main performance indicators, such as node speed, seed density, node density, and degree of irregularity, as shown in Fig. 3.

Figure 3 shows that mobility can improve the accuracy in WSNs. Our simulation experiments reveal that the MCL technique can provide accurate localization even when the seed density is low, and network transmissions are highly irregular.

### 3.2 Simulation Results of Grid-Based MCL

We compare the grid-based MCL with the MCL in terms of node speed, seed density, node density, and degree of irregularity. As shown in Fig. 4, the trend of the simulation curve of the grid-based MCL is similar to the MCL, because the




(a) Impact of node density (b) Impact of seed density

(c) Impact of node speed (d) Impact of irregularity

**Fig. 4.** Comparison of the estimated error of grid-based MCL and MCL algorithm

grid-based MCL still follows the main idea of the MCL, although the grid-based MCL has higher positioning accuracy and faster convergence speed.

Figure 4a illustrates the impact of node density. It performs poorly when network density is below 6, but performs best when network density is larger than 10. This is due to the fact that network density has a great impact on the accuracy of hop count. Figure 4b shows that the accuracy of both grid-based MCL and MCL improves as seed density increases since nodes will receive more location announcements. The localization error of MCL and grid-based MCL will almost converge to a single curve, although grid-based MCL still benefits little from its improved particle weighting. In Fig. 4c, we investigate different node velocities. Both algorithms benefit from an increasing node velocity in the beginning, as periods without seed information are getting shorter as nodes are moving faster. However, since the node velocities are larger than .8r, the localization error is growing as nodes lose contact to seed nodes more often. We use the degree of irregularity ($doi$) to denote the maximum radio range variation in the direction of radio propagation. For example, if $doi = 0.1$, then the actual

radio range in each direction is randomly chosen from [0.9r, 1.1r]. Figure 5d shows the grid-based MCL and MCL techniques are not substantially affected by *doi*.

## 4 Conclusion and Future Work

In this paper, we analyzed the MCL algorithm and proposed grid-based MCL, a localization solution for mobility applications in WSNs. The grid-based MCL uses node's mobility to reduce the sampling area and builds grids to predict the node's movement behavior. In our approach, either the estimated error or the sampling area can be reduced. The proposed grid-based MCL approach can be applied to, but not limited to applications where nodes are moving on predefined paths as in wildlife monitoring, car traffic, and so on. Future work may include how different types of movement will affect localization and how our technologies will be extended to provide security.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61671138, 61731006) and was partly supported by the 111 Project No. B17008.

## References

1. Chien CH, Hsu CC, Wang WY, Kao WC, Chien CJ: Global localization of Monte Carlo localization based on multi-objective particle swarm optimization. In: 2016 IEEE 6th international conference on consumer electronics-Berlin (ICCE-Berlin). IEEE; 2016. p. 96–7.
2. Saarinen J, Andreasson H, Stoyanov T, Lilienthal AJ. Normal distributions transform Monte-Carlo localization (NDT-MCL). In: 2013 IEEE/RSJ international conference on intelligent robots and systems; Nov 2013. p. 382–9.
3. Nuss D, Yuan T, Krehl G, Stuebler M, Reuter S, Dietmayer K. Fusion of laser and radar sensor data with a sequential Monte Carlo Bayesian occupancy filter. In: 2015 IEEE intelligent vehicles symposium (IV); June 2015. p. 1074–81.
4. Doherty L, Pister KSJ, Ghaoui LE. Convex position estimation in wireless sensor networks. In: Proceedings IEEE INFOCOM 2001. Conference on computer communications. Twentieth annual joint conference of the IEEE computer and communications society (Cat. No. 01CH37213), vol. 3; 2001. p. 1655–63.
5. Li CY, Li IH, Chien YH, Wang WY, Hsu CC. Improved Monte Carlo localization with robust orientation estimation based on cloud computing. In: 2016 IEEE congress on evolutionary computation (CEC); July 2016. p. 4522–7.
6. Adewumi OG, Djouani K, Kurien AM. RSSI based indoor and outdoor distance estimation for localization in WSN. In: 2013 IEEE international conference on industrial technology (ICIT); Feb 2013. p. 1534–9.
7. Bellili F, Amor SB, Affes S, Samet A. A new importance-sampling ml estimator of time delays and angles of arrival in multipath environments. In: 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP); May 2014. p. 4219–23.

8. Al-Jazzar SO, Strangeways HJ, McLernon DC. 2-d angle of arrival estimation using a one-dimensional antenna array. In: 2014 22nd European signal processing conference (EUSIPCO); Sept 2014. p. 1905–9.
9. Militzer B, Driver KP. Development of path integral Monte Carlo simulations with localized nodal surfaces for second-row elements. Phys Rev Lett. 2015;115:176403. https://doi.org/10.1103/PhysRevLett.115.176403.
10. Hartung S, Kellner A, Rieck K, Hogrefe D. Monte Carlo localization for path-based mobility in mobile wireless sensor networks. In: 2016 IEEE wireless communications and networking conference; Apr 2016. p. 1–7.
11. Baggio A, Langendoen K. Monte Carlo localization for mobile wireless sensor networks. Ad Hoc Netw. 2008;6(5):718–33. http://www.sciencedirect.com/science/article/pii/S1570870507001242
12. Zhu H, Mao J, Wang L, Fu L, Guo N. The study on point average energy consumption by Monte Carlo in large-scale wireless sensor networks. In: 2015 IEEE international conference on information and automation; Aug 2015. p. 1700–3.

# WalkSLAM: A Walking Pattern-Based Mobile SLAM Solution

Lin Ma[1(✉)], Tianyang Fang[1], and Danyang Qin[2]

[1] School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, China
malin@hit.edu.cn

[2] Electrical Engineering College, Heilongjiang University, Harbin, China

**Abstract.** In indoor localization scenarios, a sheer coordinate with respect to a basis is insufficient to indicate the users' situation due to a lack of information about landmarks distributed in the environments. To extract landmarks' information manually, however, is inefficient and thus vulnerable to changes of the environments. Simultaneous localization and mapping can solve the localization and landmarks' information extracting problems. This paper presents WalkSLAM, a SLAM solution that estimates both the path taken by the user and the locations of Wi-fi devices in the indoor space, using a smartphone. This solution extends the previous work by introducing human walking patterns into the specific SLAM problem. Experiments demonstrate that the improvement consists of increased efficiency of the particle filter, and hence, of the overall algorithm, and a better estimation of the user's location and path.

**Keywords:** Indoor localization · Simultaneous localization and mapping · Walking pattern · Walk ratio

## 1 Introduction

In the last decades, both military and civil have seen increasing need for efficient and accurate location-based service. Location-based service, as its name indicates, controls features and provides functions regarding location data of the users. Thus, the basis of location-based service is to provide a location estimate of the user according to the measurements of devices distributed in the environment taken by the user's device, which requires mapping measurements into exact locations.

In order to provide massive location-based service in large-scale indoor scenario both efficiently and economically, the idea of using a smartphone to run SLAM algorithm while the user carries it and walk around the indoor [1–3] space comes out and draws researchers' immediate attention. But the inaccurate nature of smartphone sensors will lead to an undesirable uncertainty of both the location estimates and landmark estimates.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_160

### 1.1 Related Work

Previous work has addressed this accuracy issue of smartphone SLAM in different levels. SmartSLAM [3] used path-smoothing model and building orientation model to supervise the heading orientation of the user in order to smooth the noisy path estimate; simultaneously localization and configuring [1] used fixed step length model and EKF to overcome the unreliability of motion measurements. However, fixed step length is still not optimal. Some researches also focused on elimination of Gaussian noise in a RSSI-based localization scenario [4].

This paper presents WalkSLAM, a SLAM solution runs on smartphones. WalkSLAM contributes previous work by adding human walking pattern elements into the SLAM algorithm, both on prior stage and posterior stage. The evaluation on live data shows a noticeable improvement on the path estimation and landmark estimation.

The rest of the paper is organized as follows: In Sect. 2, the problem statement and the model are discussed. In Sect. 3, we introduce a new SLAM algorithm named WalkSLAM. The implantation of WalkSLAM and some refining maneuvers are introduced shortly after. In evaluation section, live data is analyzed and the average errors of both path estimation and landmark estimation are discussed separately with a comparison between WalkSLAM solution and common FastSLAM solution.

## 2 Problem Definition and Modeling

The original aim of the research is to map an indoor environment using a smartphone without any prior information from the environment. The smartphone is capable of motion sensing and RSSI observing, and in a common SLAM system, the motion sensing ability guarantees the control info of a robot, where location estimation can be extracted, and the RSSI observing achieved by Wi-fi module inside the smartphone can obtain observation of the accessing points working as the landmarks. After this thought, a hidden Markov model [5–7] can be applied to our scenario.

Figure 1 represents a hidden Markov model. The poses make the core Markov chain, and in our situation, they are the locations on the path during the navigation; the state transition means a step of the user is taken. Because we do not know the actual coordinate of the locations, they become the hidden states.

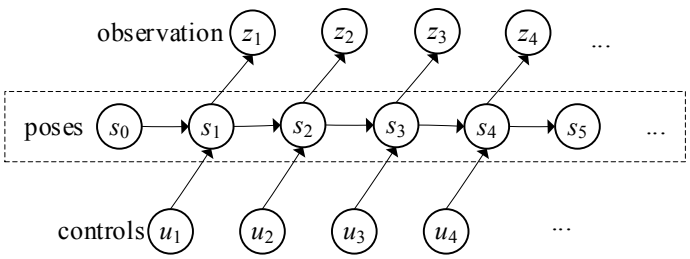



**Fig. 1.** Hidden Markov model in SLAM problems

Now, we can effectively describe our problem as a SLAM problem. The poses are the user's location, and the current pose will be denoted $s_t$. Poses evolve according to the motion model [7]:

$$s_t \sim p(s_t|u_t, s_{t-1}) \tag{1}$$

Among which $s_t$ is the current pose, $s_{t-1}$ the previous, and $u_t$ the current control. The control consists of motion measurements from the device, which in our case is the smartphone. After initializing each pose, to map the surroundings, the device senses landmarks using sensors. Sensor observations follow the measurement model:

$$z_t \sim p(z_t|s_t, \theta) \tag{2}$$

where $\theta$ is a set consisting of all the landmarks. Now, the SLAM problem can be formulated. SLAM problem is to make a refined estimation of all landmarks $\theta$ and all poses $s^t = \{s_0, s_1, s_2, \ldots\}$ along the user's path from the controls and observations, which can be described as $p(s^t, \theta|z^t, u^t)$. The controls, as stated above, are the motion measurements from motion sensors on the phone, which indicates a step's heading and orientation. After an initialization of the starting pose, given previous and current control, an estimate of the current state can be calculated. However, the motion sensors are alarmingly inaccurate; thus, in further discussion, we decide to add some prior information into the controls for better performance.

## 3 WalkSLAM

### 3.1 Particle Path Sampling

With the control and the observation of our SLAM system defined, we can present our SLAM problem solution in the following steps with implantation details. As shown in Fig. 2, the algorithm begins with the path particle sampling after the input of motion measurements $u_t$ from motion sensors. Due to the nature of human walking pattern, we do not sample the pose $s_t$ directly but rather estimate the transition from $s_{t-1}$ to $s_t$ which is denoted $\overrightarrow{(s_t - s_{t-1})}$.

The transition can be described as a combination of the fixed step length and the heading. For the filtering of the heading direction, instead of the path-smoothing method, we present a new filtering model:

$$\omega_t = \begin{cases} \omega_{t-1}, & |\omega_{t-1} - \omega_t| > \varepsilon_\omega \\ \omega_t, & |\omega_{t-1} - \omega_t| \leq \varepsilon_\omega \end{cases} \tag{3}$$

The rationale behind this model is that first the motion sensor is pretty noisy and unreliable according to previous reports [1, 8], and because of that, in some degree we can safely ignore small direction changes without worrying too much about losing the details of heading since Monte Carlo method involved in the sampling phase can guarantee the coverage of most minor direction changes.

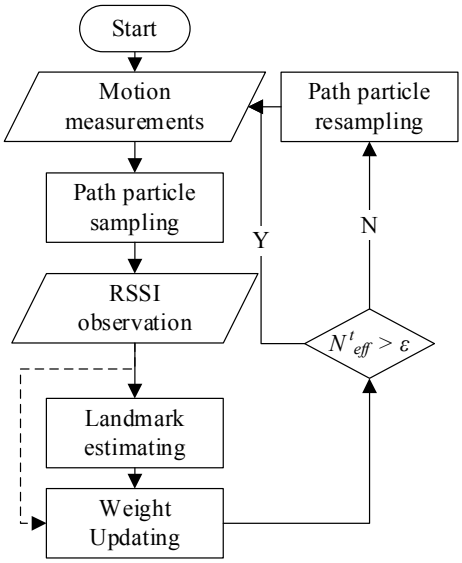



**Fig. 2.** SLAM algorithm loop. Notice that the dotted line connecting RSSI observation with weight updating step means the observation takes part in the weight's calculation

After the path sampling, we get a set of path particles, and according to our design, the next step should be landmark estimation, where we initialize and update the landmarks' locations using the observations. However, since our observation is the 1D RSSI measurements, a single observation is not enough to initialize a landmark's location. Thankfully this problem can be solved using the range-only localization model presented in [9], and the major requirement of this method is several consecutive poses and relative observations.

The updating of landmarks is guided by the observations, so we need a filter to refine the landmark location, in this situation a 2D coordinate, with 1D RSSI as the observations. An extended Kalman filter can be applied to this situation, using the method mentioned in [1].

### 3.2 Importance Weight Updating

In robotic SLAM systems, the only indicator of the reliability of a particle is how the observation matches up the landmark estimations, and it is called importance weight factor:

$$w_t^{[m]} = \frac{\text{target distribution}}{\text{proposal distribution}} = \frac{p\left(s^{t,[m]}|z^t, u^t\right)}{p\left(s^{t,[m]}|z^{t-1}, u^t\right)} \tag{4}$$

However, in our WalkSLAM scenario, another factor should be taken into consideration to evaluate the degree of the particle path fitting human walking pattern [10], specifically the walk ratio indicator [11]. For convenience, we call the original

importance weight factor observation factor (OF), and the walk ratio indicator walk factor (WF). Then, the new weight becomes:

$$w_t^{[m]} = \mathrm{OF} \cdot \mathrm{WF} \tag{5}$$

As the matter of implantation, for the calculation of OF, we can use the conclusion from [7]:

$$\mathrm{OF} = \frac{p\left(s^{t,[m]}|z^t, u^t\right)}{p(s^{t,[m]}|z^{t-1}, u^t)} \overset{\mathrm{EKF}}{\approx} \sum_k p\left(z_t|\theta_k^{[m]}, s_t^{[m]}\right) p\left(\theta_k^{[m]}\right) \tag{6}$$

Since we are always sure about the identity of each RSSI scan result, the equation can be simplified to:

$$\mathrm{OF}^{[m]} = \sum_k p\left(z_t|\theta_k^{[m]}, s_t^{[m]}\right) \tag{7}$$

Then, we can calculate OF using the Gaussian function:

$$\mathrm{OF}^{[m]} = \sum_k f\left(z_t|\left\|\theta_k^{[m]} - s_t^{[m]}\right\|, \delta_\omega\right) \tag{8}$$

For the calculation of WF, since the proposal distribution is made to be a Gaussian distribution, according to the Bayesian theory we can simplify the equation as:

$$\mathrm{WF}^{[m]} \propto p\left(d^{t,[m]}|\Delta^{t,[m]}, \delta_{\mathrm{WR}}\right) \tag{9}$$

For implementation purpose, we first calculate the average WR of the path particle,

$$\overline{\mathrm{WR}} = \frac{1}{n}\sum_{k=1}^{n} d_k \cdot \Delta_k \tag{10}$$

Then, through a Gaussian function, we can get the WF of the particle as:

$$\mathrm{WF}^{[m]} = f\left(\overline{\mathrm{WR}}|R, \delta_{\mathrm{WR}}\right) \tag{11}$$

where $d_k$ is the step length and $\Delta_k$ the time that step costs. $R$ is the average walk ratio of a healthy man, and $\delta_{\mathrm{WR}}$ the variance of walk ratio.

To sum up, the new weight updating model becomes:

$$w_t^{[m]} = f\left(\frac{1}{n}\sum_{k=1}^{n} d_k \cdot \Delta_k|R, \delta_{\mathrm{WR}}\right) \sum_k p\left(z_t|\theta_k^{[m]}, s_t^{[m]}\right) \tag{12}$$

Then, for further use, we normalize the weights:

$$w_t^{[m]} = w_t^{[m]} \left( \sum_{k=1}^{n} w_t^k \right)^{-1} \tag{13}$$

### 3.3 Resampling Particles

After getting the weights of the particle set, we now resample the particle set due to the weight of each particle. To evaluate the effectiveness of a particle set using the weight factor, the effectiveness index is introduced:

$$N_{\text{eff}}^t = \left( \sum_{k=1}^{n} w_t^k \right)^{-1} \tag{14}$$

As shown in Fig. 2, every time the value of $N_{\text{eff}}^t$ gets low enough, the resampling step is triggered; this way we can resample only when we need to, to effectively use all the particles and lower the time consumption.

## 4 Implementation and Evaluation

In the test, we choose a typical indoor scenario which is a floor of office with a long corridor. The longest part of the corridor is 47.3 m at length, and the overall corridor is 3 m at width. Thirteen access points were accessible along the corridor. A live test in comparison with the ground truth is given in Figs. 3 and 4. Since our goal is to estimate both the path and the landmarks, the localization biases of them are evaluated.

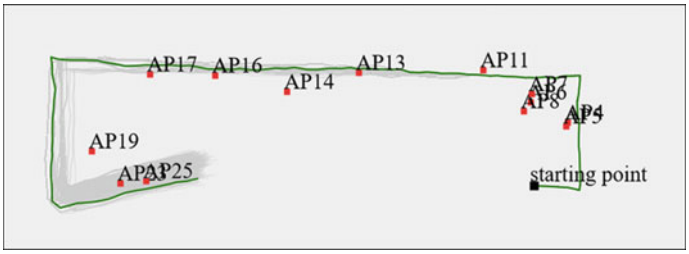



**Fig. 3.** Replay of a live test

Considering the additive nature of the pedometer bias, tests were run on different lengths to indicate the difference made by navigating different lengths and the results are given in Table 1, although due to the variance of step length the tests can only be grouped by a range of distance they cover.

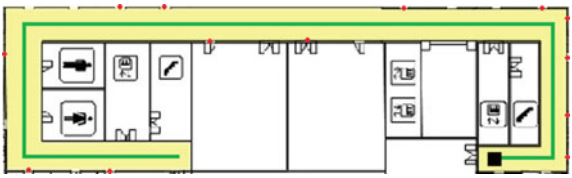


**Fig. 4.** Ground truth of the same test with a floorplan

**Table 1.** Result of three long tests and three short tests

| Path length | Test | Localization error (m) | Standard deviation |
|---|---|---|---|
| 90–100 m | 1 | 2.97 | 0.37 |
| | 2 | 2.34 | 0.29 |
| | 3 | 2.78 | 0.30 |
| 35–40 m | 4 | 1.17 | 0.11 |
| | 5 | 1.41 | 0.17 |
| | 6 | 0.94 | 0.10 |

We can still conclude from the results that for a short run of the WalkSLAM, the average localization error falls below 3.0 m, which is good enough for a room-level localization scenario. Another essential factor of the evaluation is the localization error of the landmarks. In the three long runs, 12 access points are all observable; thus, we use them to calculate the error.

To better demonstrate the overall performance when localizing the landmarks, we take each landmark estimation in each run as an individual landmark estimation and use the cumulative distribution function to bring out the result. When compared to a common FastSLAM run with an S1 model applied and one without, as shown in Fig. 5, the improvement in performance of WalkSLAM is quite noticeable.

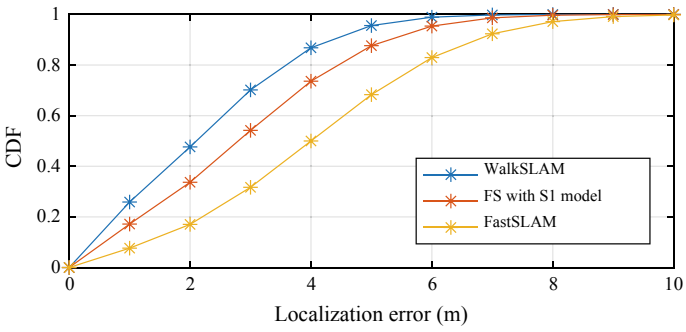



**Fig. 5.** Comparison of landmark estimating performance between WalkSLAM and traditional FastSLAM

## 5 Conclusion

In this research, a simultaneous localization and mapping solution that focuses on running on smartphones was developed. Compared to traditional SLAM solutions, in this new solution, the human walking pattern was considered into the algorithm, with a new sampling model and an advanced weight updating model presented, and the implantation methods given in detail. Six live tests were run in a typical indoor environment. The result indicated that the accuracy could fit a room-level localization scenario and the landmark estimation performance of WalkSLAM is noticeably better than that of a traditional solution.

**Acknowledgments.** This paper is supported by National Natural Science Foundation of China (61571162), Ministry of Education—China Mobile Research Foundation (MCM20170106).

## References

1. Bulten W, Van Rossum AC, Haselager WFG. Human SLAM, indoor localisation of devices and users. In: 2016 IEEE first international conference on internet-of-things design and implementation (IoTDI). IEEE; 2016. p. 211–22.
2. Ferris BD, Fox D, Lawrence N. WiFi-SLAM using Gaussian process latent variable models. Science. 2007;7:2480–5.
3. Shin H, Chon Y, Cha H. Unsupervised construction of an indoor floor plan using a smartphone. IEEE Trans Syst Man Cybern Part C Appl Rev. 2012;42(6):889–98.
4. Patri A, Rath SP. Elimination of Gaussian noise using entropy function for a RSSI based localization. In 2013 IEEE second international conference on image information processing (ICIIP-2013). IEEE; 2013. p. 690–4.
5. Arulampalam MS, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans Signal Process. 2002;50(2):174–88.
6. Dellaert F, Fox D, Burgard W, Thrun S. Monte Carlo localization for mobile robots. In: Proceedings 1999 IEEE international conference on robotics and automation (Cat. No.99CH36288C), vol. 2; May, 1999. p. 1322–8.
7. Montemerlo M, Thrun S, Koller D, Wegbreit B. FastSLAM: a factored solution to the simultaneous localization and mapping problem. In: Proceedings of 8th national conference on artificial intelligence/14th conference on innovative applications of artificial intelligence, vol. 68, issue 2; 2002. p. 593–8.
8. Zhao N. Full-featured pedometer design realized with 3-axis digital accelerometer. Analog Dialogue. 2010;44:1–5.
9. Olson E, Leonard JJ, Teller S. Robust range-only beacon localization. IEEE J Oceanic Eng. 2006;31(4):949–58.
10. Sekiya N, Nagasaki H. Reproducibility of the walking patterns of normal young adults: test-retest reliability of the walk ratio(step-length/step-rate). Gait Posture. 1998;7(3):225–7.
11. Rota V, Perucca L, Simone A, Tesio L. Walk ratio (step length/cadence) as a summary index of neuromotor control of gait: application to multiple sclerosis. Int J Rehabil Res. 2011;34(3): 265–9.

# Time-Frequency Spatial Smoothing MUSIC Algorithm for DOA Estimation Based on Co-prime Array

Aijun Liu(✉), Zhichao Guo, and Mingfeng Wang

Harbin Institute of Technology, Weihai, China
mylaj@hitwh.edu.cn

**Abstract.** In this paper, the time-frequency spatial smoothing MUSIC algorithm (TF-SSMUSIC) for DOA estimation based on co-prime array is proposed. The spatial smoothing MUSIC (SSMUSIC) is a typical DOA estimation algorithm based on co-prime array. TF-SSMUSIC replaces SSMUSIC's data covariance matrix with a time-frequency distribution matrix, which leads to a better DOA estimation performance. By selecting points in the time-frequency domain, not only the signal-to-noise ratio (SNR) can be improved effectively, but the signal interference in different time-frequency domains can be isolated. The improvement of SNR makes TF-SSMUSIC have a more accurate DOA estimation than SSMUSIC in the case of low SNR. Especially, if source signals are separable in the time-frequency domain, TF-SSMUSIC can process them solely. In this way, the angle resolution and the number of predictable source signals can be improved greatly.

**Keywords:** DOA estimation · Time-frequency · Co-prime array · Spatial smoothing MUSIC

## 1 Introduction

In recent years, co-prime array has attracted extensive attentions because of its own characteristics: simple construction, large degree of freedom, etc. However, the DOA estimation of co-prime array is faced with difficulties. Because the distance between array elements includes multiple scales, the traditional algorithm is no longer applicable. The spatial smoothing MUSIC (SSMUSIC) algorithm has solved this problem, which is introduced in detail in [4–7]. Because of the increased array aperture, the angle resolution is improved. By making effective use of the degree of freedom, the number of predictable source signals can be improved obviously.

In order to further improve the performance of the SSMUSIC algorithm, the time-frequency analysis method was introduced. The spatial time-frequency distribution matrix is applied to the subspace-based algorithm and used to replace the covariance matrix in the SSMUSIC algorithm for DOA estimation, which is called the time-frequency spatial smoothing MUSIC algorithm. PWVD enhances the signal-to-noise ratio and sort source signals in different time-frequency domains, which greatly


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_161

improve the performance of DOA estimation [1–3, 8]. In the case of low SNR, the SNR of the useful signal can be effectively improved by selecting points on the time-frequency ridge of the signal, so the DOA estimation is more accurate. For source signals which are separable in the time-frequency domain, TF-SSMUSIC can process them solely. In this way, the angle resolution and the number of predictable source signals in the time-frequency domain can be improved greatly.

The manuscript is organized as follows. The system model and the SSMUSIC algorithm are briefly introduced in Sect. 2. In Sect. 3, we shall introduce the TF-SSMUSIC algorithm in detail. In order to verify the performance of the TF-SSMUSIC algorithm, the simulation results are given in Sect. 4. Lastly, the paper is concluded in Sect. 5.

## 2 System Model

Figure 1 shows a co-prime array, consisting of a co-prime pair of uniform linear sub-arrays. The first uniform linear array includes $N$ array elements with an array element spacing of $Md$. The second uniform array includes $2M - 1$ arrays, and the element spacing is $Nd$, where $M$ and $N$ are co-primes. The basic array element spacing is $d = \lambda/2$, where $\lambda$ is the wavelength of signals.

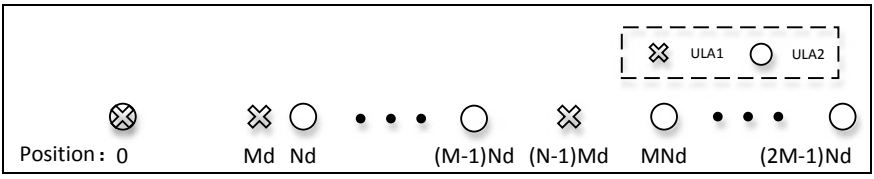



**Fig. 1.** System model

Since $M$ and $N$ are prime numbers, other array elements will not overlap except the first array element. So, there are $N + 2M - 1$ array elements in total, which will result in a continuation position from $-MNd$ to $MNd$ sets at least. In this way, we can get $2MN + 1$ degrees of freedom from $N + 2M - 1$ physical elements. These elements are located at:

$$S = \{Mnd, 0 \le n \le N-1\} \cup \{Nmd, 0 \le m \le 2M-1\} \tag{2.1}$$

Assuming there are $P$ narrowband signals in the far field, the data received by the co-prime array can be expressed as:

$$\mathbf{X}[k] = \mathbf{A}\mathbf{S}[k] + \mathbf{N}[k] \tag{2.2}$$

Among them, $\mathbf{A} = [a(\theta_1)\ a(\theta_2)\ \cdots\ a(\theta_p)]$ denotes the array manifold matrix, $a(\theta)$ denotes the $N + 2M - 1$ dimension steering vector and $\mathbf{S}[k] = [s_1(k)\ s_2(k)\ \cdots\ s_p(k)]^{\mathrm{T}}$

denotes the $k$th snapshot of the source vector. The noise $\mathbf{N}[k]$ is assumed to be white Gaussian noise.

Therefore, the array's covariance matrix of the received data $\mathbf{X}(k)$ can be expressed as:

$$\mathbf{R}_{xx} = E\left[\mathbf{X}(k)\mathbf{X}^H(k)\right] \tag{2.3}$$

Vectorize $\mathbf{R}_{xx}$:

$$\mathbf{z} = \mathrm{vec}(\mathbf{R}_{xx}) = \mathrm{vec}\left[\sum_{i=1}^{P} \sigma_i^2 a(\theta_i) a^H(\theta_i)\right] + \rho^2 l_n = \mathbf{B}(\theta_1, \ldots, \theta_P)\mathbf{p} + \rho^2 \mathbf{L} \tag{2.4}$$

where

$$\mathbf{B} = [b(\theta_1)\, b(\theta_2) \cdots b(\theta_P)] \tag{2.5}$$

$$b(\theta_i) = a^*(\theta_i) \otimes a(\theta_i) \tag{2.6}$$

$$\mathbf{p} = \left[\sigma_1^2, \ldots, \sigma_P^2\right]^T \tag{2.7}$$

$$\mathbf{L} = \left[\mathrm{e}_1^T, \ldots, \mathrm{e}_L^T\right] \tag{2.8}$$

Among them, $\mathrm{e}_i$ ($i$ = 1, …, $L$) is a column vector where the $i$th position is 1 and the other positions are 0. $z$ can be seen as a received signal model whose array manifold is $B$, where $p$ denotes the source vector and $\rho$ $2L$ can be seen as a noise term.

At this time, each column of $B$ is a longer array manifold, and its specific position is determined by the position differences, including the position differences between the sub-arrays:

$$\{\pm(Mn - Nm)d, 0 \le n \le N - 1, 0 \le m \le 2M - 1\} \tag{2.9}$$

and the position differences of the two sub-arrays:

$$\{(Mn_1 - Nn_2)d, 0 \le n_1, n_2 \le 2M - 1\} \tag{2.10}$$

$$\{(Nm_1 - Nm_2)d, 0 \le m_1, m_2 \le 2M - 1\} \tag{2.11}$$

Sort these differences and reject the excesses. From these differences, continuous position differences can be obtained from $-MN$ to $MN$. This results in a new array manifold $B1$, which is a subset of $B$, and the dimension is $(2MN + 1)K$. $B1$ can be seen as an array manifold corresponding to an array containing $2MN + 1$ array elements whose position difference is from $-MNd$ to $MNd$. Then, it is decomposed into $MN + 1$ overlapping sub-arrays, where the position of the $i$th sub-array is: $\{(-i + 1 + n)d \mid n = 0, 1, \ldots, MN\}$ and the degree of freedom of each sub-array is $MN + 1$. Assuming that Ai represents the array manifold of the $i$th sub-array, the corresponding sub-array

receiving model is: $z_{li} = \mathbf{A}_i\mathbf{p} + \sigma_n^2 \mathrm{e}_i$. The $i$th position of $\mathrm{e}_i$ is 1, and the rest is 0. The covariance matrix of each sub-array is: $\mathbf{R}_i = z_{li} z_{li}^H$. Take the average of them to get the final expression:

$$\mathbf{R} = \frac{1}{MN+1} \sum_{i=1}^{MN+1} \mathbf{R}_i \tag{2.12}$$

This matrix is called a spatial smoothing matrix, which is full rank. Therefore, the subsequent steps can be performed with the MUSIC algorithm. In this way, *MN* source signals can be estimated.

## 3 Time-Frequency Spatial Smoothing MUSIC

Before introducing the time-frequency spatial smoothing MUSIC algorithm, we first introduce the spatial time-frequency distribution matrix. In discrete Cohen class bilinear time-frequency transforms, express $\mathbf{x}(t)$ as $\mathbf{x}_p(t)$ and $\mathbf{x}_q(t)$ ($p$, $q$ = 1, 2, …, $M$), and then we can obtain the mutual time-frequency distribution among elements. We can construct the spatial time-frequency distribution matrix by using these elements.

$$\mathbf{C}_{\mathbf{xx}}(t,f) = \sum_{l=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \phi(k,l)\ \mathbf{x}(t+k+l)\mathbf{x}^H(t+k-l)\mathrm{e}^{-j4\pi fl} \tag{3.1}$$

where $t$ denotes time, $f$ denotes frequency and $\phi(k, l)$ is a kernel function, which is a function of time and delay. Substituting the received array signal model (2.2) into (3.1) and taking the statistical average:

$$\boldsymbol{E}\{\mathbf{C}_{\mathbf{xx}}(t,f)\} = \mathbf{A}\mathbf{C}_{\mathbf{ss}}(t,f)\mathbf{A}^H + \sigma^2\mathbf{I}_{\mathbf{M}} \tag{3.2}$$

The formula (3.2) has a similar structure to the covariance matrix of the data received from arrays, so it is called an array reception model in the time-frequency domain.

When the kernel function $\phi(k, l) = \delta(k - m)$, the time-frequency distribution is Pseudo-Wigner–Ville distribution (PWVD), which has a good time-frequency flocculability. Spatial PWVD is defined as:

$$\mathbf{C}_{\mathbf{xx}}(t,f) = \sum_{\tau=-(L-1)/2}^{(L-1)/2} \mathbf{x}(t+\tau)\ \mathbf{x}^H(t-\tau)\mathrm{e}^{-j4\pi f\tau} \tag{3.3}$$

where $L$ denotes the length of the window function.

The PWVD of the signal is expressed as:

$$\mathbf{C_{ss}}(t,f) = \sum_{\tau=-(L-1)/2}^{(L-1)/2} \mathbf{s}(t+\tau)\ \mathbf{s}^H(t-\tau)\mathrm{e}^{-j4\pi f\tau} \tag{3.4}$$

The spatial time-frequency distribution matrix can still be expressed by formula (3.2).

The model shown in Eq. (3.2) is suitable for all the time-frequency points, and we can use the time-frequency average method to choose points. For chirp signals which contain noise, the energy of the signal in the time-frequency plane concentrates on the vicinity of instantaneous frequency. The noise is evenly distributed in the entire time-frequency plane. Therefore, selecting the time-frequency points along the signal time-frequency ridge can increase signal-to-noise ratio effectively.

The *i*th diagonal element of the signals' spatial time-frequency distribution matrix $\mathbf{C}_{ss}$ $(t, f)$ is:

$$\mathbf{C_{d_i d_i}}(t,f) = \sum_{\tau=-(L-1)/2}^{(L-1)/2} A_i^2 \mathrm{e}^{j[\varphi_i(t+\tau)-\varphi_i(t-\tau)]}\mathrm{e}^{-j4\pi f\tau} \tag{3.5}$$

where $\varphi_{\mathrm{i}}(t)$ denotes the phase of the *i*th signal and $A_i$ denotes the amplitude of the *i*th signal.

For each signal, take $N - L + 1$ time-frequency points along the time-frequency ridge, and then average them to obtain the average spatial time-frequency matrix, as shown in the formula:

$$\hat{\mathbf{C}} = \frac{1}{n_0(N-L+1)}\sum_{q=1}^{n_0}\sum_{i=1}^{N-L+1}\mathbf{C_{xx}}\left(t, f_{q,i}(t)\right) \tag{3.6}$$

$f_{q,i}$ $(t)$ represents the instantaneous frequency of the *q*th signal at the *i*th sampling point, taking a mathematical expectation on $\hat{\mathbf{C}}$:

$$\mathbf{C} = E\left(\hat{\mathbf{C}}\right) = \frac{L}{n_0}\mathbf{A^0}\mathbf{R_{ss}^0}\left(\mathbf{A^0}\right)^H + \sigma^2\mathbf{I} \tag{3.7}$$

where

$$\mathbf{R_{ss}^0} = \boldsymbol{diag}\left[A_i^2,\ i = 1, 2, \ldots, n_0\right] \tag{3.8}$$

$$\mathbf{A^0} = \left[\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_{n_0}}\right] \tag{3.9}$$

only, $n_0$ signals are considered here.

It has been proved that the spatial time-frequency distribution matrix has a similar structure as a traditional array covariance matrix. Therefore, the STFD matrix can be applied to the subspace-based algorithm to replace the covariance matrix in the spatial

smoothing MUSIC algorithm. The spatial spectrum estimation of the signal can be obtained by processing the STFD matrix. We call this the time-frequency smoothing MUSIC algorithm.

## 4 Simulation Results

In this section, we present several numerical examples to evaluate the proposed method. We will compare the performance of SSMUSIC and TF-SSMUSIC under different conditions by simulation. Construct a co-prime array. In this array, $M = 2$, $N = 3$, basic array element spacing $d = \lambda/2$.

First, we consider a single source coming from 21°. Using the Monte Carlo method, 500 simulations, take 512 snapshots, and the SNR is varied from −20 to 10 dB in increments of 2 dB. The comparison of SSMUSIC and TF-SSMUSIC is shown in Fig. 2. In DOA estimation success rate, the estimated angle is rounded to an integer.

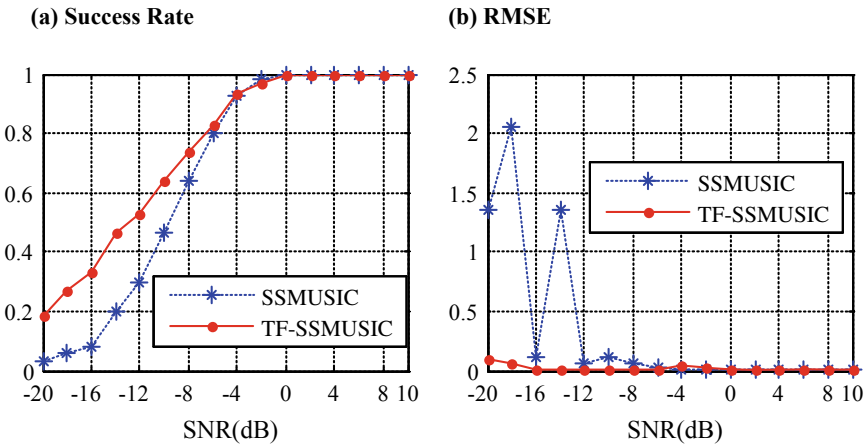



**Fig. 2.** DOA estimation in the case of single source signal

As shown in the simulation results, the TF-SSMUSIC algorithm is more accurate and stable, especially in the case of low SNR. The energy of the source signal has a good aggregation on the time-frequency ridge, but the noise is evenly distributed on the time-frequency domain. By selecting points on the time-frequency ridge, the actual SNR can be effectively improved. In this way, TF-SSMUSIC can have a better DOA estimation performance.

In order to analyze the angle resolution, set two chirp sources and make the direction of them approach gradually. Keep the SNR constant (SNR = 0 dB), and the start and end frequencies of the two sources are: 0.2–0.3 and 0.3–0.4.

From Fig. 3, we can see that when the direction of arrivals is 0° and 4°, respectively, the SSMUSIC algorithm cannot distinguish them, but the TF-SSMUSIC algorithm can get the DOA clearly. The reason is that the source signals are separable in the time-frequency domain, and the interference of others can be avoided by selecting points in the time-frequency domain. So, as long as the two source signals are separable in the time-frequency domain, they can be distinguished using the TF-SSMUSIC algorithm no matter how close they are.

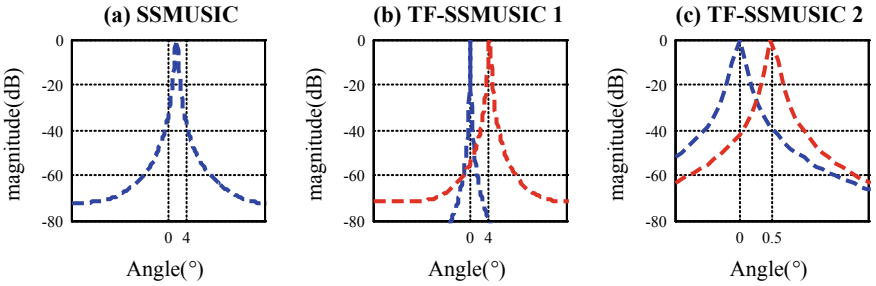



**Fig. 3.** Angle resolution of SSMUSIC and TF-SSMUSIC

For co-prime array, using the SSMUSIC algorithm, continuous position differences can be obtained from $-MN$ to $MN$ at least. In this part, we consider $MN = 6$ uncorrelated chirp sources with same SNR. The direction and the begin–end frequency of them are shown in Table 1.

**Table 1.** Source signal settings

| Signals | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Frequency | 0.1–0.15 | 0.15–0.2 | 0.2–0.25 | 0.25–0.3 | 0.3–0.35 | 0.35–0.4 |
| DOA | −55 | −35 | −15 | 0 | 20 | 40 |

The result is shown in Fig. 4, and it can be seen from the simulation that both algorithms can estimate the DOA of six source signals well. But, for the TF-SSMUSIC algorithm, as long as the signal is separable in the time-frequency domain, it can estimate more source signals regardless of $M$ and $N$. There are eight sources that estimate the DOA of them using TF-SSMUSIC. It can be seen from Fig. 4c TF-SSMUSIC 2 that the DOA of them can be well estimated.

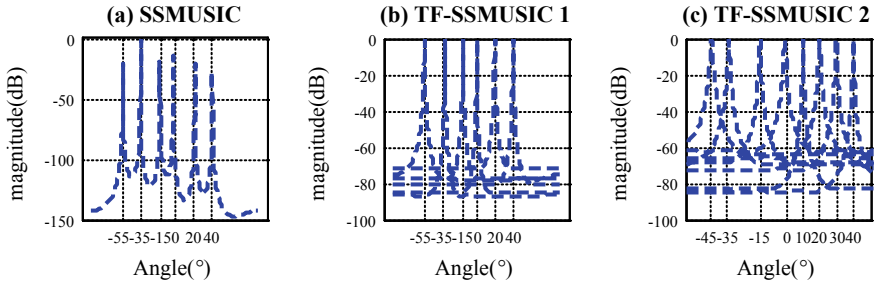



**Fig. 4.** Maximum number of predictable source signals

Consider six sources as shown in Table 1, **and** change the spacing between the elements. The influence of the reduction in element spacing on TF-SSMUSIC is shown in Fig. 5. For the SSMUSIC algorithm, the interference between source signals will be more serious when the spacing between array elements is reduced, which makes it difficult to reduce the element spacing. By using the frequency division feature of the TF-SSMUSIC algorithm, high angle resolution can be maintained on the premise of narrower array element spacing. However, if different source signals are inseparable in the time-frequency domain, the DOA estimation performance of the TF-SSMUSIC algorithm will be drastically reduced.

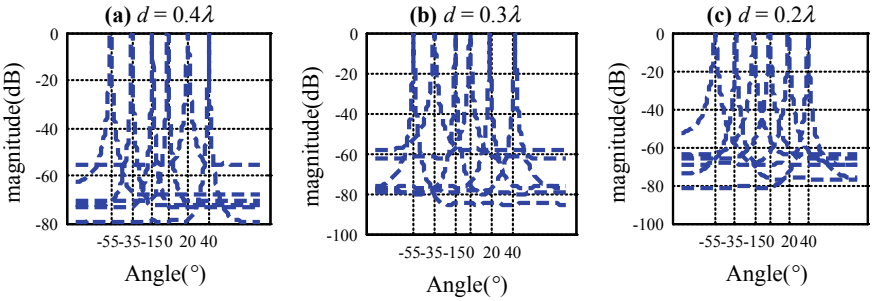



**Fig. 5.** TF-SSMUSIC

It can be seen from the simulation that the reduction in element spacing leads to a decrease in the DOA estimation performance of SSMUSIC, and it has been unable to estimate the direction of arrival correctly. However, TF-SSMUSIC can still estimate the DOA correctly. When the spacing between the elements is further reduced, the results of the TF-SSMUSIC algorithm are shown in Fig. 5.

## 5 Conclusion

In this paper, the TF-SSMUSIC algorithm is proposed by replacing SSMUSIC's data covariance matrix with a time-frequency distribution matrix. TF-SSMUSIC algorithm has the function of SNR enhancement and signal screening, which has been verified through theoretical analysis and simulation. Therefore, TF-SSMUSIC algorithm has a better estimation performance than SSMUSIC algorithm. However, TF-SSMUSIC algorithm has its own limitation too. The signals processed by TF-SSMUSIC should be separable in the time-frequency domain.

**Acknowledgments.** This work was supported in part by the National Key R&D Program of China under Grant 2017YFC1405202, in part by the National Natural Science Foundation of China under Grant 61571159 and Grant 61571157, and in part by the Public Science and Technology Research Funds Projects of Ocean under Grant 201505002.

## References

1. Belouchrani A, Amin MG. Time-frequency MUSIC. IEEE Signal Process Lett. 2002;6(5): 109–10.
2. Chen S, Li X, Shao Z. Study on the performance of DOA estimation algorithms. In: IEEE international conference on communication problem-solving. IEEE; 2016. p. 475–7.
3. Li F. Time-frequency DOA estimation and application comprehensive polarization information. Harbin: Harbin Institute of Technology; 2016.
4. Liu CL, Vaidyanathan PP. Remarks on the spatial smoothing step in coarray MUSIC. IEEE Signal Process Lett. 2015;22(9):1438–42.
5. Pal P, Vaidyanathan PP. Co-prime sampling and the music algorithm. In: Digital signal processing workshop and IEEE signal processing education workshop. IEEE; 2011. p. 289–94.
6. Tan Z, Eldar YC, Nehorai A. Direction of arrival estimation using co-prime arrays: a super resolution viewpoint. IEEE Trans Signal Process. 2014;62(21):5565–76.
7. Xiong J, Chen S. Direction of arrived estimation based on the co-prime arrays. In: Fourth international conference on information science and industrial applications. IEEE; 2016. p. 21–4.
8. Zhang Y, Mu W, Amin M. Subspace analysis of spatial time-frequency distribution matrices. IEEE Trans Signal Process. 2001;49(4):747–59.

# Non-uniform Sampling Scheme Based on Low-Rank Matrix Approximate for Sparse Photoacoustic Microscope System

Ting Liu(✉) and Yongsheng Zhao

Department of Control Science and Engineering, Dalian Maritime University, Dalian 116026, China
{liuting0910, yszhao}@dlmu.edu.cn

**Abstract.** Optical-resolution photoacoustic microscopy (OR-PAM) has rapidly emerging as tool for label-free morphology and function imaging of the microvasculature in vivo with a high resolution. However, it is difficult to achieve real-time imaging due to the limitation of data acquisition time. Therefore, a sparse PAM (SPAM) has been proposed to obtain a high-resolution PAM image with relatively low sampling density. In order to successfully set up a SPAM system, the two key problems that we need to keep focus on are designation of the compressive sampling scheme and the corresponding image recovery algorithm. Typically, a random uniform sampling scheme is adopted. In this paper, a non-uniform sampling scheme based on low-rank matrix approximate is proposed to replace the conventional point-by-point scanning scheme to implement fast data acquisition. The effectiveness of the proposed non-uniform scanning scheme is validated using both numerical analysis and PAM experiments. As compliments for SPAM system, the total sampling points are dramatically decreased for a relatively high-resolution PAM vascular image and to implement accelerated data acquisition. Thus, OR-PAM is of great potential to find board biomedical applications in the pathophysiology studies of tumor and treatments for anti-angiogenesis.

**Keywords:** Tumor angiogenesis · Optical-resolution photoacoustic microscopy · Low-rank matrix approximate

## 1 Introduction

Angiogenesis is a hallmark of tumor growth, invasion, metastasis, etc. all [1–3]. Take tumor growth as an example, regulated by a number grown factors secreted by tumor cells, the remodeling of existing vasculature and forming of new microvascular results in the morphology alteration of microvascular network [3, 4].

Medical imaging modalities such as magnetic resonance imaging and positron emission tomography have been employed to study angiogenesis, but the micrometer-level resolution limits their capacity to visualize the fine feature changes [3, 5, 6]. To clearly visualize the finest micro-vessels in tumor angiogenesis, a spatial resolution of $\sim$5 μm is required, which can be provided by optical-resolution photoacoustic


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_162

microscopy (OR-PAM)—a technique that uses a tightly focused laser beam for photoacoustic excitation [7, 8].

OR-PAM images are typically acquired with a point-by-point sampling scheme in the field of view [9–11]. For OR-PAM, more researches are focused on the resolution enhancement. Unfortunately, most of the methods sacrifice the imaging speed, such as employing more measurements to get high-resolution images [12]. For example, the spatially Fourier-encoded photoacoustic microscopy [13] proposed by Liang et al. In their system, the measurements required are twice as many as the conventional OR-PAM. Besides, the more the number of measurements the great challenge for the system memory and the more is the radiation to tissues. Thus, a fast OR-PAM with a relatively high resolution is preferred.

Like many other medical images [14–16], the sparsity of the PAM has been proven [16–18] and thoughtfully explored by incorporated the technique of compressed sensing. The final image is reconstructed from far less measurements [19–21]. Based on that analysis, we have proposed a SPAM system to achieve fast imaging [11]. The effectiveness of a SPAM system depends on many critical techniques, the most important one being the compressive sampling scheme.

Typically, a random uniform scanning has been adopted. In our research [11], the random uniform sampling mask is generated by a random uniform expander graph, for it can be perfectly recovered by low-rank matrix completion (LRMC) [22, 23]. However, the significant part of the PAM image is ignored for uniform scanning. Thus, a non-uniform sampling scheme is preferred, which could be acquired based on the edge expander graphs. Therefore, different sampling density could be utilized for the region of interest and other background areas to put more measurements on the significant region of the image. This operation can further reduce the measurements.

However, the region of interest and other background areas could not be separated. In this situation, more measurements may be focused on the principal components [24]. Based on that analysis, low-rank matrix approximate (LRMA) based non-uniform sampling scheme is proposed for vascular imaging by OR-PAM.

In this paper, a LRMA-based non-uniform sampling scheme is investigated for vascular PAM imaging. The paper is organized as follows. The proposed method is presented in Sect. 2, including the theory of LRMA and the LRMA based non-uniform sampling scheme. In Sect. 3, the results of both numerical and real PAM system have been analyzed and discussed to validate the non-uniform sampling scheme. Conclusions are drawn in the final section.

## 2 Methods

### 2.1 Low-Rank Analysis for PAM Vascular Images

The sparsity of the photoacoustic image has been verified in many researches. More recently, the low-rank property has been incorporated into the field of photoacoustic imaging.

To our knowledge, photoacoustic imaging is typically used in tumor and vasculature imaging. As for the low-rank and non-uniform properties for PAM vascular images,

compressive random sampling scheme has been proposed to cope with the redundancy (sparsity) of the observed photoacoustic data to realize rapid data acquisition.

Therefore, the image reconstruction process is formulated into a LRMC problem. Based on that analysis, the compressive sampling model for our system is defined as follow:

$$D = Kx, \tag{1}$$

where $D$ is the incomplete compressive measurement data, $x$ is the final desired PAM image, and $K$ is the random sampling mask. The goal is to recover $x$ from $D$, which is a typically LRMC problem. The problem is defined as:

$$\begin{array}{ll} \min\limits_{x} & \operatorname{rank}(x) \\ \text{s.t.} & D = Kx \end{array}, \tag{2}$$

Unfortunately, the rank minimization problem in (2) is NP-hard, and so we solve the following convex relaxation problem instead:

$$\begin{array}{ll} \min\limits_{x} & \|x\|_* \\ \text{s.t.} & D = Kx \end{array}, \tag{3}$$

where $\|x\|_*$ is the nuclear norm of $x$.

According to the compressive sampling model, our interest is focused on two key problems: the sampling scheme and the corresponding image recovery method. The major contribution of our paper is the former problem. For the designation of $K$, the edge expander has been incorporated to fully explore the low-rank and non-uniform properties. However, the non-uniform property is not obvious for vascular image, which means the region of interest cannot be point out from the background. But the edge expander can only be used as uniform random sampling scheme in this situation. Therefore, a LRMA-based sampling scheme has been proposed for PAM vascular images in this paper.

### 2.2 Low-Rank Matrix Approximation

LRMA method is a fast way to solve the LRMC problem. The method is utilized to obtain the non-uniform sampling mask to focus more measurements on the principal component of vascular.

Given a matrix $A \in R^{m \times n}$ and an integer $r$, $r \leq \min(m, n)$, finding its LRMA matrix $B$ is a critical task in many technical fields. The problem has been described as:

$$\begin{array}{ll} \min\limits_{x} & \|A - B\|_F^2 \\ \text{s.t.} & \operatorname{rank}(B) < r \end{array}, \tag{4}$$

Thus, the constraint here is the same for low-rank matrix completion problem. Typically, the problem of LRMA can be solved by the truncated singular value

decomposition (TSVD). In order to obtain $B$ with rank $(B) = r$ of $A$, the SVD of $A$ should be calculated first:

$$A = \sum_{i=1}^{n} u_i \sigma_i v_i^T \tag{5}$$

For TSVD, the first-$r$ singular values approximation is chosen as these of LRMA results for $A$, which is

$$B = A_r = \sum_{i=1}^{r} u_i \sigma_i v_i^T \tag{6}$$

Take a numerical vascular image, for example, the dimension for the full rank image, which is shown in Fig. 1a, is 256 × 256. The LRMA image obtained by TSVD with $r = 50$ is shown in Fig. 1b. From Fig. 1, all the principal components of the vascular were exacted by TSVD-based LRMA.

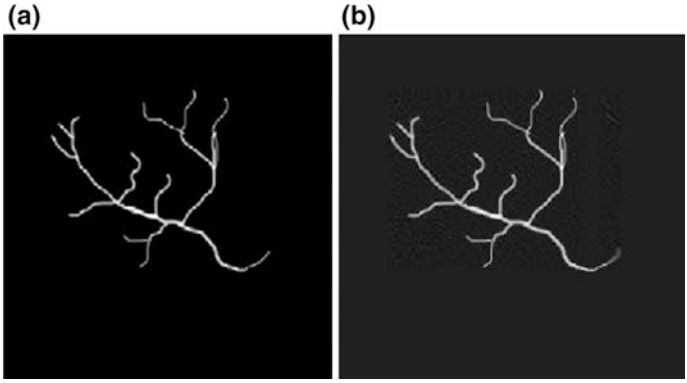



**Fig. 1.** Low-rank approximation image of photoacoustic image. **a** Original image. **b** Low-rank approximation image, rank is 50.

With the help of low-rank matrix approximate $B$, the non-uniform sampling mask $K$ could be obtained by the comparison between $A$ and $B$,

$$K_{ij} = \begin{cases} 1 & \text{if } |A_{ij} - B_{ij}| \geq T \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $T$ is the threshold, which is determined by experience.

### 2.3 LRMA-Based Fast Non-uniform Sampling Scheme

In general, TSVD method is used in LRMA. However, it is time-consuming that all the singular values should be calculated for TSVD. In this section, the bilateral random

projections (BRP) of matrix $A \in R^{m\times n}$ are used to replace the SVD computation to significantly reduce the time cost.

Assume that $\tilde{A}$ is a fast rank $r$ BRP approximation of $A$, the bilateral random projection is

$$\begin{aligned} Y_1 &= AA_1 \\ Y_2 &= A^T A_2 \end{aligned} \tag{8}$$

where both $A_1 \in R^{n\times r}$ and $A_2 \in R^{m\times r}$ are random matrixes. Therefore, the fast rank $r$ BRP approximation of $A$ is

$$\tilde{A} = Y_1 \left(A_2^T Y_1\right)^{-1} Y_2^T \tag{9}$$

The computation for $\tilde{A}$ included a matrix inversion operation and three matrix multiplication. Therefore, the computation cost for $\tilde{A}$ is $r^2(2n+r)+mnr$ which is dramatically less than TSVD.

However, the decreasing rate of the singular values for vascular images is not fast enough. In this case, the effect of Eq. (9) will be worsened. Therefore, a modified version of BRP is utilized.

For the modified BRP, $A' = (AA^T)^q A$ is used to instead of $A$. Because these two matrixes have the same singular values, while they have equal eigenvalues $\lambda_i(A') = \lambda_i(A)^{2q+1}$. However, the decreasing rate of the singular values for $A'$ is much faster than $A$. Based on that analysis, the BRP for $A'$ is defined as

$$\begin{aligned} Y_1 &= A'A_1 \\ Y_2 &= A'A_2^T \end{aligned} \tag{10}$$

And the fast rank $r$ matrix approximate for $A'$ is

$$\tilde{L} = Y_1 \left(A_2^T Y_1\right)^{-1} Y_2^T \tag{11}$$

Here, the $QR$ decomposition of $Y_1$ and $Y_2$ are computed to obtain $\tilde{L}$. Thus, the LRMA is

$$\begin{aligned} L = (\tilde{L})^{2q+1} &= Q_1 \left[R_1 \left(A_2^T Y_1\right)^{-1} R_2^T\right]^{\frac{1}{2q+1}} Q_2^T \\ Y_1 &= Q_1 R_1 \\ Y_2 &= Q_2 R_2 \end{aligned} \tag{12}$$

In order to show a comparison study, Fig. 1a is still used to test the performance of BRP-based LRMA method. The LRMA image obtained by BRP with $r$ = 50 is shown in Fig. 2b. From the result, all the principal components of the vascular were exacted

by BRP-based LRMA. Even though there are many artifacts in BRP-based LRMA image, there is no influence on the non-uniform sampling scheme with the help of threshold $T$.

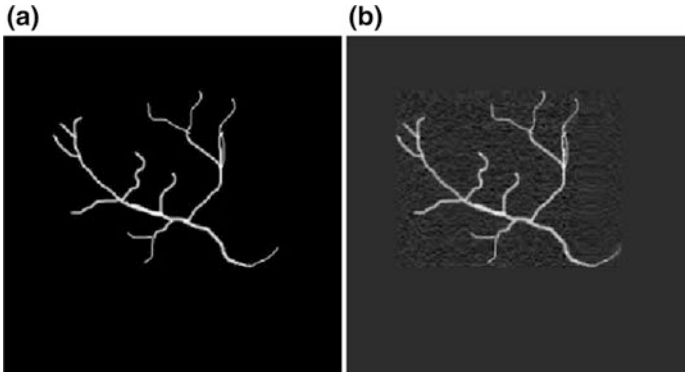



**Fig. 2.** Low-rank approximation image of photoacoustic image. **a** Original image. **b** BRP-based Low-rank approximation image, rank is 50.

## 3 Experimental Results

The performance of the proposed non-uniform sampling mask based on LRMA was investigated both by simulated and real PAM images involving quantitative and qualitative analysis.

### 3.1 Results from Simulated Study

The simulated study is conducted to provide both qualitative and quantitative analysis. Besides, the reference image has been carefully selected to simulate the property of vascular, as shown in Fig. 3a. The dimension of the image is 500 × 500. The same with our previous results, the compressive sampling image is acquired by downsampling on the reference image according to sampling mask $K$ [11].

In our paper, a non-uniform sampling mask based on LRMA has been proposed to effectively improve the imaging speed by dramatically reduce the sampling points with relative high resolution. As for vascular, the region of interest is not obvious, the edge expander can only be used as uniform sampling mask. Our method focused on the principal components is more appropriate.

Figure 3 shows the non-uniform sampling mask and the comparison results between LRMA-based non-uniform scheme and edge expander-based uniform schemes. Figure 3b shows the non-uniform sampling mask base on LRMA with the sampling rate fixed at 0.5. From the sampling mask, the outline of the vascular could be

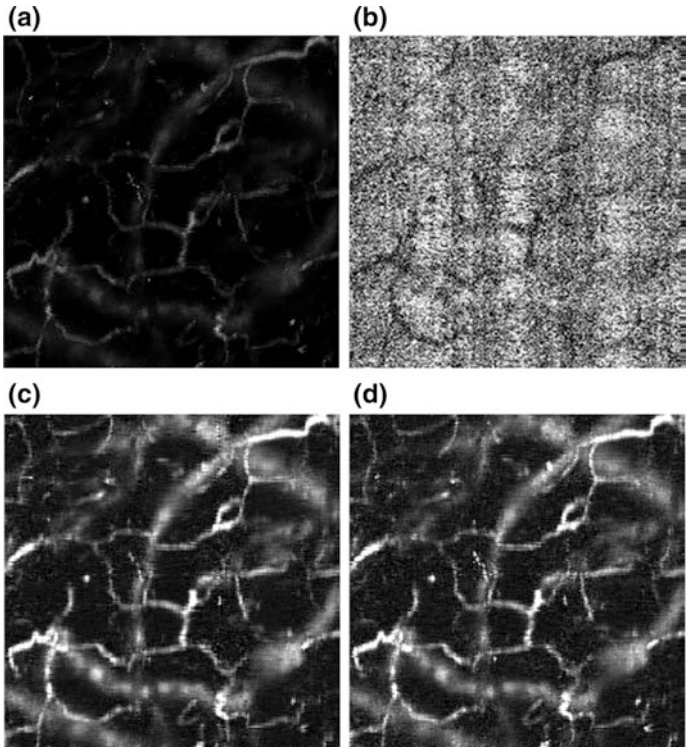



**Fig. 3.** Result of non-uniform sampling based on low-rank matrix approximation with BRP. **a** Original image. **b** Non-uniform sampling mask. **c** Recovered image with non-uniform sampling scheme. **d** Recovered image with uniform sampling scheme.

found for more sampling points are focused on the principal components. The recovered image under two sampling schemes is shown in Fig. 3c, d. The results show that the two sampling schemes could acquire considerable results.

In order to better compare the two schemes, three quantitative parameters are evaluated to illustrate the performance, mean square error (MSE), peak signal to noise ratio (PSNR), and structural similarity (SSIM), which are defined as

$$\mathrm{MSE} = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\|X - Y\|^2/mn\right) \tag{13}$$

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right) \tag{14}$$

$$\begin{aligned} \mathrm{SSIM} &= l(X,Y)\cdot C(X,Y)\cdot S(X,Y) \\ l(X,Y) &= \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \\ C(X,Y) &= \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \\ S(X,Y) &= \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \end{aligned} \tag{14}$$

where $X$ is the reference image and $Y$ is the recovered image. $\mu_x$ and $\mu_y$ are the mean value of $X$ and $Y$, while $\sigma_x$ and $\sigma_y$ are the variance. Besides, $\sigma_{xy}$ is covariance between $X$ and $Y$. The results are shown in Table 1. From all the three parameters, the proposed sampling scheme shows a better result.

**Table 1.** Performance indices for comparison between uniform and non-uniform sampling mask

| | Uniform | Non-uniform |
|---|---|---|
| SSIM | 0.9949 | 0.9957 |
| MSE | 3.5351e−06 | 2.8853e−06 |
| PSNR | 54.5160 dB | 55.3980 dB |

### 3.2 Results from Real PAM System

In this section, the feasibility of our non-uniform sampling scheme in practice has been verified with our system. Figure 4 shows a schematic of the experimental OR-PAM system, adapted from a conventional OR-PAM system. In order to achieve non-uniform sampling, a compressive sampling scheme is utilized to realize rapid data acquisition, and the corresponding image reconstruction problem is formulated as a low-rank matrix completion problem.

As shown in Fig. 4, our system is modified from an OR-PAM system and can be transferred to many other existing OR-PAM systems. The main modification is the compressive sampling mode achieved by the optical scanner. The compressive sampling employs a sampling mask $K$, which is a $0 \times 1$ matrix generated by the PC with a user-defined sampling density $k$. Based on the sampling mask $K$, the $X$–$Y$ linear stage is controlled by a controller to achieve sparse optical scanning.

The mouse ears have complete vascular structure. Moreover, imaging tumor angiogenesis on the mouse ear could avoid the respiratory and cardiac motion artifacts from living animals [3]. So, the mouse ears are chosen as our imaging sample to image the vascular as other previous studies.

The full sampling image (Fig. 5a) has been acquired to be a reference image to better compare the non-uniform and uniform scheme in real system. The resolution of conventional OR-PAM system could reach microvascular level. For comparison between non-uniform and uniform sampling, both schemes are utilized to acquire the incomplete observed data, as shown in Fig. 5b, c. The sampling rate is both fixed at 0.48. From the sampling mask, more sampling pints are focused on vascular for non-

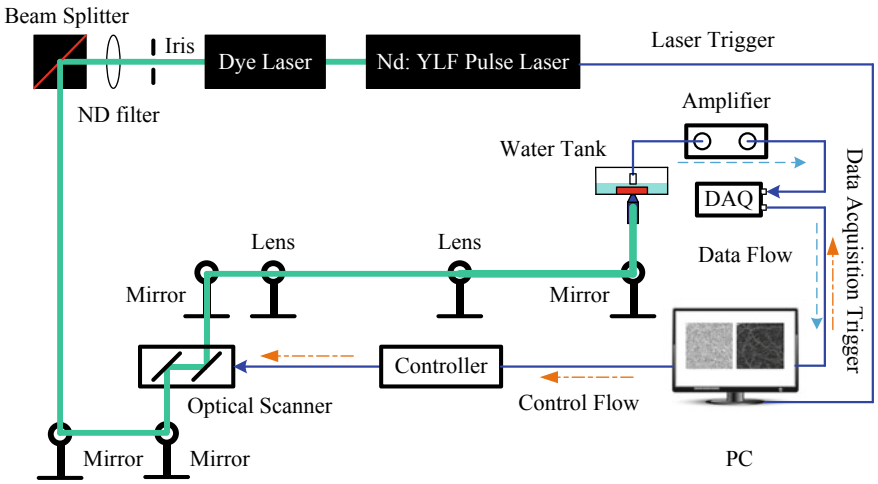



**Fig. 4.** Schematic of the experimental sparse PAM system.

uniform sampling mask. The corresponding recovery image is shown in Fig. 5e, f. The recovery method we used is OPTSPACE. All the microvascular could be recovered by both schemes. But there are more artifacts shown in Fig. 5e, which is non-uniform sampling scheme.

The same quantitative analysis with numerical study, SSIM and PSNR are also studied. The SSIM and PSNR for uniform sampling are 0.9172 and 39.5907 dB, respectively. Further, the SSIM and PSNR for non-uniform sampling are 0.9898 and 47.0978 dB, respectively. Both parameters have shown significantly improvement.

In summary, non-uniform sampling scheme proposed for photoacoustic microscopy imaging on vascular has obtained more complete sparse sampling data than the uniform random sampling in real PAM experiments. Besides, in terms of the final image recovery performance, non-uniform scheme still has achieved a better result.

## 4 Conclusion

In this paper, a LRMA-based non-uniform sampling scheme has been proposed. The main idea of the paper is to obtain higher resolution image than uniform sampling scheme by giving more observation to the principal component and recovery the images from limited number of acquisitions. For the designation of LRMA-based non-uniform sampling, BRP has been incorporated to accelerate the speed of sampling mask generation by avoid the computation of SVD. Finally, the numerical and PAM experiment results have validated that the proposed method can achieve fast data better resolution imaging than uniform sampling in the compressive OR-PAM system.

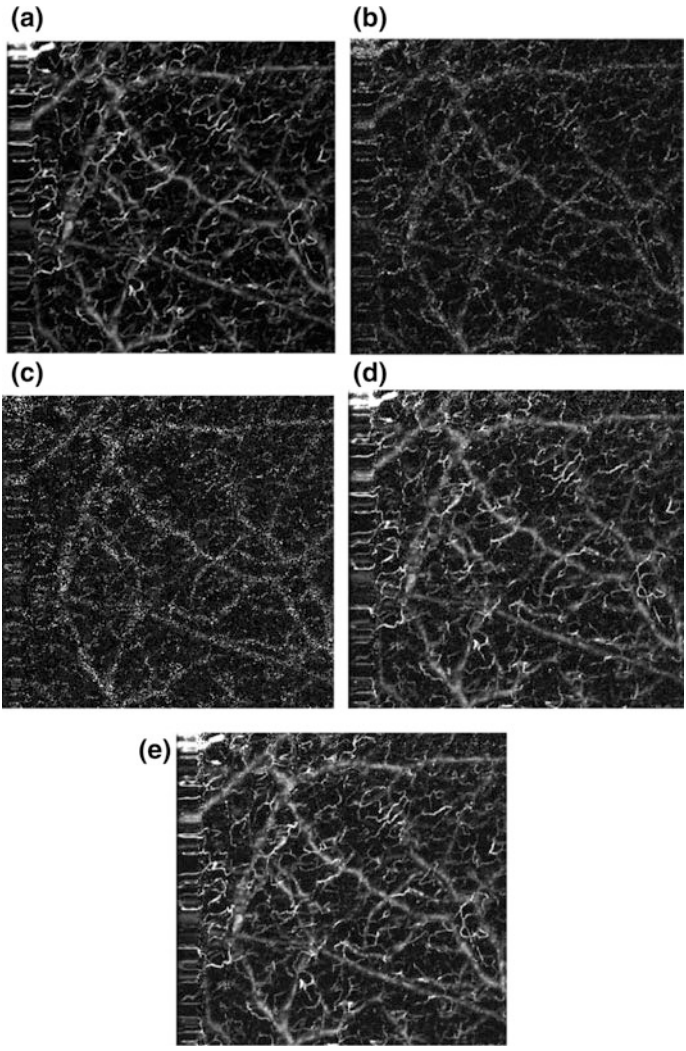



**Fig. 5.** Results of non-uniform sampling based on low-rank matrix approximation. **a** The full sampling PAM image. **b** The uniform sampling PAM data. **c** Non-uniform sampling PAM data with low-rank matrix approximation method. **d** Recovered image by uniform sparse sampling. **e** Recovered image by non-uniform sparse sampling.

**Acknowledgments.** This work is supported by Fundamental Research Funds for the Central Universities (Grant No. 3132017127).

## References


1. Weis SM, Cheresh DA. Tumor angiogenesis: molecular pathways and therapeutic targets. Nat Med. 2011;17:1359–70.
2. Carmeliet P, Jain RK. Molecular mechanisms and clinical applications of angiogenesis. Nature. 2011;473:298–307.
3. Lin R, Chen J, Wang H, et al. Longitudinal label-free optical-resolution photoacoustic microscopy of tumor angiogenesis in vivo. Quant Imaging Med Surg. 2015;5(1):23–9.
4. Albini A, Tosetti F, Li VW, Noonan DM, Li WW. Cancer prevention by targeting angiogenesis. Nat Rev Clin Oncol. 2012;9:498–509.
5. Emblem KE, Mouridsen K, Bjornerud A, Farrar CT, Jennings D, Borra RJ, Wen PY, Ivy P, Batchelor TT, Rosen BR, Jain RK, Sorensen AG. Vessel architectural imaging identifies cancer patient responders to anti-angiogenic therapy. Nat Med. 2013;19:1178–83.
6. Haubner R, Beer AJ, Wang H, Chen X. Positron emission tomography tracers for imaging angiogenesis. Eur J Nucl Med Mol Imaging. 2010;37:S86–103.
7. Maslov K, Zhang HF, Hu S, Wang LV. Optical-resolution photoacoustic microscopy for in vivo imaging of single capillaries. Opt Lett. 2008;33:929–31.
8. Rao B, Li L, Maslov K, Wang L. Hybrid-scanning optical-resolution photoacoustic microscopy for in vivo vasculature imaging. Opt Lett. 2010;35:1521–3.
9. Lihong VW, Liang G. Photoacoustic microscopy and computed tomography: from bench to bedside. Annu Rev Biomed Eng. 2014;16:155–85.
10. Zhenghua W, Mingjian S, Qiang W, Ting L, Naizhang F, Jie L, Yi S. Photoacoustic microscopy image resolution enhancement via directional total variation regularization. Chin Opt Lett. 2014;12(12):121701.
11. Liu T, Sun M, Meng J, et al. Compressive sampling photoacoustic microscope system based on low rank matrix completion. Biomed Signal Process Control. 2016;26:58–63.
12. Meng J, Wang LV, Ying L, Liang D, Song L. Compressed-sensing photoacoustic computed tomography in vivo with partially known support. Opt Express. 2012;20:16510.
13. Liang J, Gao L, Li C, Wang LV. Spatially Fourier-encoded photoacoustic microscopy using a digital micro mirror device. Opt Lett. 2014;39:430–4.
14. Provost J, Lesage F. The application of compressed sensing for photoacoustic tomography. IEEE Trans Med Imaging. 2009;28(4):585–94.
15. Liang D, Zhang HF, Ying L. Compressed sensing photoacoustic imaging based on random optical illumination. Int J Funct Inform Pers Med. 2009;2:394.
16. Liu X, Dong P, Wei G, Xibo M, Xin Y, Jie T. Compressed sensing photoacoustic imaging based on fast alternating direction algorithm. Int J Biomed Imaging. 2012;2012:1–8.
17. Guo Z, Li C, Song L, et al. Compressed sensing in photoacoustic tomography in vivo. J Biomed Opt. 2010;15(2):021311–021311-6.
18. Sun M, Feng N, Shen Y, et al. Photoacoustic imaging method based on arc-direction compressed sensing and multi-angle observation. Opt Express. 2011;19(16):14801–6.
19. Meng J, Wang LV, Liang D, Song L. In vivo optical-resolution photoacoustic computed tomography with compressed sensing. Opt Lett. 2012;37:4573.
20. Meng J, Liu CB, Zheng JX, Lin RQ, Song L. Compressed sensing based virtual-detector photoacoustic microscopy in vivo. J Biomed Opt. 2014;19:036003.
21. Zhenghua W, Mingjian S, Qiang W, Naizhang F, Yi S. Compressive sampling photoacoustic tomography based on edge expander codes and TV regularization. Chin Opt Lett. 2014;12 (10):101102.
22. Wu Z, Wang Q, Liu J, Sun M, Shen Y. Compressive sensing theory based on edge expander graphs. Acta Autom Sin. 2014;12:2824–35.

23. Li W, Lei Z, Zhijie L, Duanqing X, Dongming L. Non-local image inpainting using low-rank matrix completion: non-local image inpainting using low-rank matrix completion. Comput Graph Forum. 2014;00:1–12.
24. Liu T, Sun M, Feng N, et al. Non-uniform sampling photoacoustic microscope system based on low rank matrix completion. In: Instrumentation & measurement technology conference. IEEE; 2016.

# An Improved Monte Carlo Localization Algorithm in WSN Based on Newton Interpolation

Lanjun Li(✉) and Jing Liang

School of Information and Communication Engineering,
University of Electronic Science and Technology of China, Chengdu, China
frankl29@163.com

**Abstract.** In recent years, with the development of sensor technology and wireless communication technology, wireless sensor network (WSN) as the technology for information acquisition and processing is widely applied in many fields. It is important for nodes to know their localizations for further applications. In this article, a range-free localization algorithm in WSN that builds upon the Monte Carlo Localization (MCL) algorithm is proposed. It concentrates on improving the sampling efficiency by changing the weights of samples. More specifically, mobility is used to improve the sampling efficiency to make sure MCL can perform well even when the sample number is low.

**Keywords:** Wireless sensor network · Monte Carlo Localization algorithm · Newton interpolation method · Sample weight

## 1 Introduction

WSN is a new technology filed based upon wireless communication technology, sensor technology, and microelectromechanical system. WSN system includes sensor node, sink node, and management node. Plenty of nodes can be set randomly inside or near the sensor filed by throwing them down from a plane or a rocket. Nodes in the sensor filed can form wireless networks through self-organizing [1].

Nowadays, there are many applications of WSN. For instance, since WSN can be set easily and the cost is not high, people are able to set sensors in those places that are hard for humans to reach, sensors in WSN can collect the information human needs to explore the area. However, receiving monitoring information of nodes without knowing the position is meaningless, so it is crucial for nodes in WSN to know their locations.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_163

In order to determine the locations of sensor nodes, the global positioning system (GPS) is used, it provides with 3D localization based on direct line-of-sight with at least four satellites. Theoretically, people can put GPS on every single sensor nodes to get the location information. Those nodes with GPS which knows their locations are called beacon nodes, and those nodes which have no idea of their locations are called unknown nodes. However, if every node is equipped with GPS, the price is often too high. As a result, we need to use as few beacon nodes as possible to determine more unknown nodes locations.

The rest of the paper is organized as follows. Section 2 presents several localiztion methods adapted in WSN. Section 3 presents the main idea of MCL and the improvement for MCL using Newton interpolation. Section 4 presents our simulation results. Finally, we conclude our work in Sect. 5.

## 2 Localization in WSNs

In this section, we mainly introduce several localization protocols designed for wireless sensor networks. Those protocols are traditionally classified in two types, range-free and range-based.

### 2.1 Range-Based Localization Methods

Range-based localization methods mainly depend on special hardware to determine the distance or angles between different nodes, then use mathematical methods to determine the location of the unknown node.

In common, there are several mathematical methods to determine the position of unknown nodes using angle or distance information between nodes, for example, trilateration method and triangulation method. As a result, it is crucial to obtain the distance and angle information in WSN using other techniques.

Mostly, distance and angles can be determined using four techniques (AOA [2], TOA [3], TDOA [4], and RSSI [5]). Although range-based localization methods always have high accuracy, most range-based methods depend on special hardware making the cost pretty high. Also, in real environment, nodes in WSN can be affected easily, making those methods unsuitable when nodes are moving. In order to be independent of hardware and counter range inaccuracies, researchers developed range-free methods that depend uniquely on the information a node receive from its neighbors.

### 2.2 Range-Free Localization Methods

Compared to range-based localization methods, range-free localization methods do not rely on distance information or angle information. They can finish localization with data as network connectivity and hop numbers. Range-free methods can achieve high localization accuracy and do not require special hardware or expensive facilities, making them one of the most potential localization techniques in the future. For example, centroid method [6], DV-HOP method [7], and APIT method [8], are commonly used in localization.

# 3 Improved Monte Carlo Localization Algorithm Based on Newton Interpolation

## 3.1 Monte Carlo Localization Algorithm

In 2004, Hu and Evans firstly come up with the idea that using Monte Carlo method in WSN localization [9]. Normally, Monte Carlo method is used in determining location of robots. It is a range-free method so that it is low cost and does not have high requirement for hardware. More importantly, Monte Carlo localization (MCL) performs much better than other localization methods when beacon nodes and unknown nodes are moving, which suits the real world environment.

In MCL, the time is divided into discrete intervals. A sensor node realizes in each time interval (the location can be described as $L_t$). The whole algorithm has three steps, which are prediction step, filtering step, and sampling step.

In prediction step, a node starts its moving from its location in $t-1$, which is $L_{t-1}$. Assume the node has no idea of its moving speed and direction, but is aware of its maximum velocity, $v_{max}$. If a node position in $t-1$ is $L_{t-1}$, then $L_t$ should be contained in the circle with a radius $v_{max}$. Once the location of a node in previous step is known, we are able to obtain the possible area the node may be in the next time interval.

However, only knowing the possible area of nodes is not enough, so in filtering step, we use the relationship between unknown nodes and beacon nodes to obtain filtering conditions, hoping to restrain the possible area. In MCL, we assume that all messages nodes send are received instantly. Hence, at time $t$, every node within the radio range $R$ of beacon nodes, is able to hear announcement from the beacon nodes. If a beacon node can communicate directly with unknown nodes, which means it is within the radio range $R$, it is called one-hop node. If a beacon node cannot communicate with unknown nodes directly by itself, but it is able to communicate with nodes via one-hop neighbors, then it is called two-hop node. Formula (1) is used to present the filtering condition. In a time interval, we can obtain one-hop list, $S$, and two-hop list, $T$ from a node. According to $S$ and $T$, we are able to get filtered area from (1).

$$filter\,(l) = \forall\ s \in S,\ d\,(l,s) \leq R \wedge \forall s \in T,\ R < d\,(l,s) \leq 2R \tag{1}$$

$d(l,s)$ stands for the Euclidean distance between unknown node and beacon node. With the filtering condition, the circle in prediction step can be reduced.

In sampling step, we use importance sampling, if a sample meets the filtering condition, the weight is set as 1, otherwise, the weight is 0. After collecting enough samples MCL needs, the average coordinate of samples are regarded as the coordinate of the estimating position of unknown node in a time interval. However, the degeneracy of the importance sampling is unavoidable since the unconditional variance if the importance weight will increase. It is necessary to do re-sampling if the number of samples is not enough. In MCL, a sample threshold $N_{eff}$ is set as a standard for sampling.

In the simulation, they adopt the random waypoint mobility model and vary several parameters for both sensor nodes and WSN. In MCL, when localization error reaches a balance phase, error in MCL is less than centroid algorithm and amorphous algorithm. Maintaining more samples for MCL algorithm can improve accuracy, but needs additional memory. From this simulation, we find that MCL performs well in WSN when nodes are moving, and mobility can even improve accuracy and reduce the costs of localization.

However, since MCL has high requirement for sample numbers, when the area after filtering is small, it is difficult to obtain enough effective samples to meet the threshold $N_{eff}$, which means the algorithm will do re-sampling over and over again. It increases the cost of the algorithm and cause time waste. As a result, we suggest an improved MCL algorithm based on Newton interpolation, fixing the problem by using the mobility of nodes.

### 3.2 Improved Monte Carlo Localization Algorithm Based on Newton Interpolation

In MCL, the weight of samples is either 1 or 0. In our algorithm, we change sample weight into numbers between 0 and 1 to increase localization accuracy. Prediction step and filtering step stay the same. Because of inertia, the moving track of nodes is continuous. Theoretically, a present position of node $L_t$ can be predicted by its previous location $L_{t-1}$, $L_{t-2}$, *etc.* In numerical analysis, interpolation is often used to predict data from the information of its previous value. So we adopt Newton interpolation in our prediction. Assume the position of a node in $L_{t-1}$, $L_{t-2}$, $L_{t-3}$ is known, we use second phase Newton interpolation (2) to determine its previous position $L_t$.

$$f(x) = f(x_0) + (x - x_0)f[x_1, x_0] + (x - x_0)(x - x_1)f[x_2, x_1, x_0] \tag{2}$$

$f[x_0, x_1]$ stands for difference quotient of $x_0$ and $x_1$, the same as $f[x_0, x_1, x_2]$. After computation, the predicted ndoe's coordinates are shown as (3) and (4).

$$x_{pre} = 3x_{t-1} - 3x_{t-2} + x_{t-3} \tag{3}$$

$$y_{pre} = 3y_{t-1} - 3y_{t-2} + y_{t-3} \tag{4}$$

Since the predicted position follows the mobility track, a node has higher probability of moving to the position near the predicted position. By using this feature, we can change samples' weights. The nearer a node is to the predicted position, the more contribution it has for localization, which means the importance weight is also higher. In order to make the computation easy, we consider the reciprocal of the distance of samples and predicted location as the sample weight as shown in (5). After normalization shown in (6), the sample weights are changed to numbers between 0 and 1.

$$d_i = \sqrt{(x_i - x_{pre})^2 + (y_i - y_{pre})^2} \tag{5}$$

$$w_i = \frac{\frac{1}{d_i}}{\sum_{i=1}^{N} \frac{1}{d_i}} \tag{6}$$

The final coordinate the algorithm provides is shown as (7) and (8)

$$x_t = \sum_{i=1}^{N} x_i w_i \tag{7}$$

$$y_t = \sum_{i=1}^{N} y_i w_i \tag{8}$$

If the predicted position does not follow the filtering condition, we still use traditional MCL algorithm to finish localization in this time interval. And we also have to set a sample number threshold in case of degeneration. However, the threshold can be smaller than traditional MCL, for after fixing the importance weights, a sample has more contribution to localization. This method improves sampling efficiency by shortening the time MCL wastes on doing re-sampling. When initializing the algorithm, the location of nodes, $L_1$, $L_2$, and $L_3$ are determined by traditional MCL.

## 4 Evaluation

In WSN, doing simulation for nodes localization is an essential way to see the advantages and disadvantages of an algorithm. In this paper, we use MATLAB to run the test to show the localization result for improved MCL based on Newton interpolation.

### 4.1 Simulation Parameters

In our experiment, the simulation parameters are shown as follows.

- Sensor nodes are randomly distributed in a 500 m × 500 m rectangular region.
- The number of unknown nodes is 320, while number of beacon nodes is 96.
- Both beacon nodes and unknown nodes can move in a velocity between 0 and vmax. In our experiment, a node cannot move faster than 50 m in a time interval, which means vmax equals to 50 m.
- The radio range of nodes is 100 m, and all nodes are able to communicate with other nodes if are within the radio range.
- The mobility model of nodes is random waypoint mobility model.

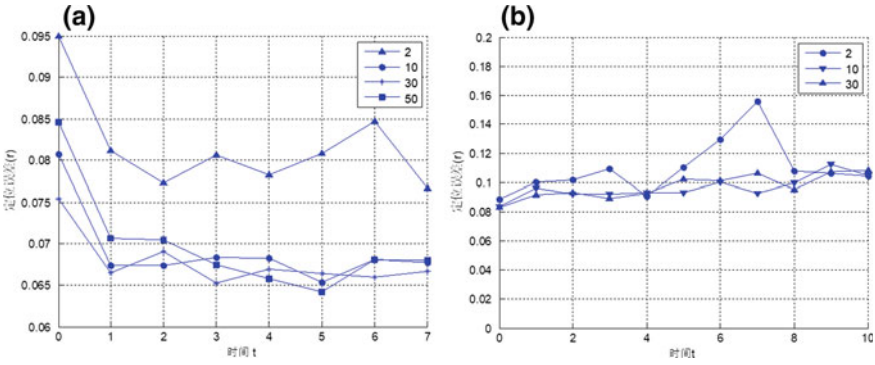



**Fig. 1.** Impact of sample number: **a** for traditional MCL; **b** for improved MCL

### 4.2 Number of Samples

As we mentioned in the last section, improved MCL increases sampling efficiency. We are going to prove that by comparing the influence the sample number has for localization accuracy. For traditional MCL, samples numbers threshold are set as four groups, 2 samples, 10 samples, 30 samples, and 50 samples. When the sample number is lower than threshold, do re-sampling until sample number equals to (or is higher than) threshold. The simulation result is shown as Fig. 1a.

The figure shows that maintaining more samples for MCL can definitely increase localization accuracy. However, when sample number is around 30, the effect of sample number is not obvious, especially when localization reaches its balance phase. And if there is too many samples in sampling phase, the accuracy is certainly high, but the computation cost can be high. Also, we cannot sacrifice localization accuracy only to make the computation simple. For improved MCL, sample numbers are set as three groups, 2 samples, 10 samples, and 30 samples. In order to make the effect of Newton interpolation more obvious, we assume that nodes move in certain track. The simulation result is shown as Fig. 1b.

The figures show that localization with high accuracy can also be achieved even when sample number is low. However, when sample number is too small, localization result may not be stable. This problem can be fixed when sample number is high, for example, over 10 samples are enough for a stable localization. This method fixes the problem that too much re-sampling process decreases the sample efficiency.

### 4.3 Localization Accuracy Comparison

In last section, we can see that improved MCL algorithm increases sampling efficiency. However, they are simulated in two different environments. In this section, we will simulate two algorithms in the same environment, which means nodes in both two algorithms follow the same moving track, to see whether the accuracy has been sacrificed to decrease the sampling efficiency. The simulation result is shown as Fig. 2.

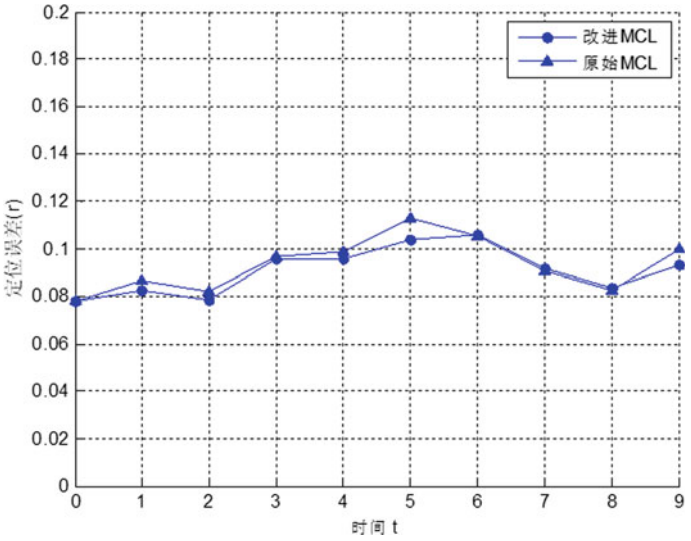



**Fig. 2.** Comparison of both algorithms

As is shown in Fig. 2, we can see that improved MCL does not affect the localization accuracy, even increases it, although the improvement in accuracy is not obvious.

## 5 Conclusion

Monte Carlo Localization algorithm is the first work to study range-free localization in the presence of mobility, and it shows that mobility improves localization accuracy, which is essential for localization in moving WSN.

However, MCL suffers from low sampling efficiency, mostly because of resampling. Our work, Monte Carlo Localization based on Newton interpolation increases sampling efficiency by adjusting sample weights due to their contribution to localization. While increasing sampling efficiency, the improved MCL algorithm still has high localization accuracy, even higher than traditional MCL.

However, there are still some disadvantages in the improved MCL algorithm. For example, computing the distance between a sample and the predicted position can be complex, which will increase the cost of the network. And when predicted position does not follow the filtering condition, the algorithm is useless. As a result, many issues remain to be explored in future work to improve those disadvantages mentioned above.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China (61671138, 61731006), and was partly supported by the 111 Project No. B17008.

## References


1. Akyildiz IF, Su W, Sankarasubramaniam Y. Wireless sensor networks: a survey. Comput Netw. 2002;38:393–441.
2. Niculescu D, Nath B. Ad hoc positioning system (APS) using AoA. In: IEEE INFOCOM, 2003, San Francisco, vol. 3; 2003. p. 1734–43.
3. Girod L, Estrin D. Robust range estimation using acoustic and multimodal sensing. In: IEEE/RSJ international conference on intelligent robots and systems (IROS 01), Maui, vol. 3; 2001. p. 1312–20.
4. Priyantha NB, Miu AKL, Balakrishnan H, et al. The ricket compass for context-aware mobile applications. In: 7th annual international conference on mobile computing and networking, Rome; 2001. p. 1–14.
5. Girod L, Bychovskiy V, Elson J, et al. Locating tiny sensors in time and space: a case study. In: 2002 IEEE international conference on computer design: VLSI in computers and processors, Freiburg; 2002. p. 214–9.
6. Bulusu N, Heidemann J, Estrin D. GPS-less low cost outdoor localization for very small devices. IEEE Pers Commun Mag. 2000;7(5):28–34.
7. Nicolescu D, Nath B. Ad-Hoc positioning systems (APS). In: Proceedings of the 2001 IEEE global telecommunications conference, vol. 5. San Antonio: IEEE Communications Society; 2001, p. 2926–31.
8. He T, Huang C, Blum BM et al. Range-free localization schemes for large scale sensor networks. In: Proceedings of the 9th annual international conference on mobile computing and networking. New York, NY, USA: ACM Press; 2003, p. 81–95.
9. Hu L, Evans D. Localization for mobile sensor networks. In: Proceedings of the 10th annual international conference on mobile computing and networking. ACM Press; 2004. p. 45–7.

# UAV Autonomous Path Optimization Simulation Based on Multiple Moving Target Tracking Prediction

Bo Wang(✉), Jianwei Bao, and Li Zhang

Nanjing University of Aeronautics and Astronautics,
Nanjing 21001, Jiangsu, China
wangbo_nuaa@nuaa.edu.cn

**Abstract.** In the UAV path planning study, due to the relative movement of multiple targets and the drone, long-term and large-scale UAV autonomous tracking has not been achieved. Therefore, aiming at this problem, this paper uses multiple moving target tracking algorithm to provide a real-time feedback on target position, estimates the later motion state of the target according to its position, and then performs the dynamic path planning by combining the feedback data and the state estimation result. Finally, The UAV path is optimized in real time. Experiments show that the proposed scheme can better plan the UAV path when multiple targets are in motion, thus improving the intelligence of the drone and the capability of long-time tracking.

**Keywords:** UAV · Target tracking · Motion estimation · Path planning

## 1 Introduction

Relying on the flexibility of UAV with the fast and low-cost remote sensing of the high-definition cameras on board, UAVs have achieved a wide application in the field of military and civil aviation remote sensing, communication relay, and so on. In the aspect of moving target tracking by UAV, the existing autonomous tracking of the single moving target has been extensively studied and applied [1]. However, as to the problem of multiple moving target tracking, the UAV operation still relies on operator's observing to control the UAV flight, and the autonomous flight path design has not been realized. Therefore, this paper combines the tracking, prediction, and path planning of single moving target by UAV, so as to realize UAV autonomous path planning for multiple targets. In order to achieve long-term and large-scale UAV autonomous tracking of multiple moving targets, this paper presents a solution. There are N moving targets to be tracked, and a drone starts from a point to reach the positions of these targets. After each target is visited once, the drone will return to the starting point. The UAV path problem [2] means how to find the shortest path. To realize the UAV autonomous path planning, this paper divides the scheme into three steps.

The first step is to obtain the specific motion equation of the selected moving targets. Because of the continuous relative movement of the drone and the targets, the


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_164

target's position will constantly change. To obtain the motion equation of the targets, this paper adopts the method of target tracking. Correlation filter method is widely used in the field of target tracking, due to its speed and accuracy, such as the KCF algorithm proposed by Henriques [3]. KCF constructs a training sample of a circulant matrix structure and transforms the solution of the problem into a discrete Fourier domain, avoiding the process of matrix inversion, reducing the complexity of the algorithm, and improving the tracking real-time performance. Therefore, the target position can be effectively updated by the KCF algorithm.

The second step is to predict the positions of moving targets. Because the position of a target is constantly changing, tracking its present position will lead to longer tracking time. Therefore, this paper uses the method of estimating the target position to optimize the UAV path. Due to the state change and uncertainty of moving targets, Kalman filtering method [4] is used in this paper. Kalman filtering method is a time-domain method which solves the problem based on state space. It has high applicability to the estimation of moving targets that frequently change motion states, thus becoming a universal machine learning method.

The last step is to use predicted positions to plan the UAV path, which means to quickly and accurately design a non-repeating shortest path after obtaining the positions of the targets. This problem is called the traveling salesman problem (TSP) in the unified research [5]. Due to the complexity of TSP, the time complexity of the exact solution method is long. Therefore, ant colony algorithm is adopted instead of the exact algorithm. The basic principle of ant colony algorithm uses pheromone to control the direction of ant's movement and autonomously and effectively approach the optimal path ultimately [6]. Then, Stutzle [7] proposed the maximum and minimum ant colony algorithm. Only the pheromone of the ant with the optimal algorithm is updated; thus, the convergence speed of the algorithm is improved. Li [8] uses prior knowledge to limit the increment of pheromone and increase the road weight factor to effectively avoid the stagnation of the algorithm. Therefore, this paper combines the improved methods of scholars to optimize the ant colony algorithm and thus solves the problem of long solution time and falling into the local optimum in the route plan using ant colony algorithm.

In general, to solve the problem of UAV autonomous flight path planning, this paper presents a solution. First, the KCF algorithm is used to track multiple moving targets to obtain the position and motion information of multiple moving targets. Then, the Kalman filter method is used to predict the positions of the moving targets. Finally, the improved ant colony algorithm is used to process the acquired information to design a shortest UAV flight path to track multiple moving targets.

The scheme can realize the near-optimal drone flight path by means of computer vision under the scene of UAV multiple moving targets tracking, thereby reducing manual operations and improving the UAV autonomous flying capability.

## 2 Algorithm Implementation

### 2.1 Overall Algorithm Flow

This paper presents an algorithmic process for UAV autonomous planning. Firstly, the targets to be tracked are selected on the video sequence returned by the drone. Then, the tracking algorithm is used to obtain the positions and state changes of the moving targets, and the positions of the targets after a certain period of time are estimated according to the position changes. After obtaining the estimated motion states, the present path is optimized based on the present UAV path setting and the motion estimation of the moving targets using the ant colony algorithm. Finally, when the drone performs the calculated path flight, it updates the positions of the targets and the path in real time to achieve the UAV path planning requirement for multiple moving targets. The specific flowchart is shown in Fig. 1.

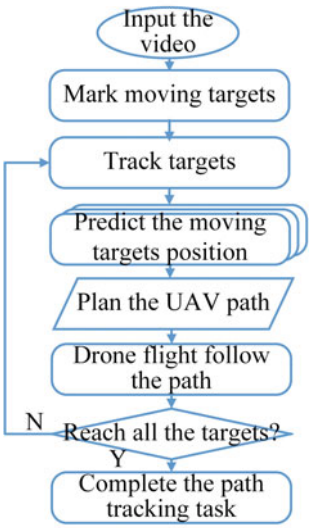



**Fig. 1.** Flowchart of the algorithmic process

### 2.2 Target Tracking Algorithm

The KCF tracking algorithm completes target tracking by designing a ridge regression classifier. The purpose of training the ridge regression classifier is to find a function $f(x) = w^T x$ to obtain the minimum loss function:

$$\min_{w} \sum_{i} (f(x_i - y_i)^2 + \lambda \|w\|^2) \tag{1}$$

where $x = (x_1, x_2, \ldots, x_n)$ represents a sample of the tracking model, $\lambda$ represents a regular term which prevents over-fitting, and $w$ represents a solution parameter. Solve the above equation and obtain:

$$w = (X^T X + \lambda I)^{-1} X^T y \tag{2}$$

where $X$ is the cyclic matrix, and $y$ is a set of tag values for each sample. For samples that cannot be classified in the original space, the kernel function is needed to map the linearly inseparable patterns in the low-dimensional space to the high-dimensional space. The format of the kernel function is as follows:

$$k(x, z) = \varphi(x)\varphi(z) \tag{3}$$

where $k(x, z)$ is the kernel function, and $\varphi(x)$ and $\varphi(z)$ are mapping functions from low-dimensional space to high-dimensional space. When the kernel function is used to map the sample $x$ to $\varphi(x)$, in $f(x) = w^T x$, the coefficient $w$ is converted into $a$ in the dual space. Combine with the above formula and obtain:

$$a = (K + \lambda I)^{-1} y \tag{4}$$

where $K$ represents the mapped kernel matrix, $K_{ij} = k(x_i, x_j)$. In the tracking phase, dense sampling is performed near the targets to obtain different tracking samples, and then the ridge regression classifier is used to determine the final position information of the targets. After that, the position information is further trained by the ridge regression classifier again. The process is performed repeatedly to complete the tracking of targets. The specific process is shown in Fig. 2.

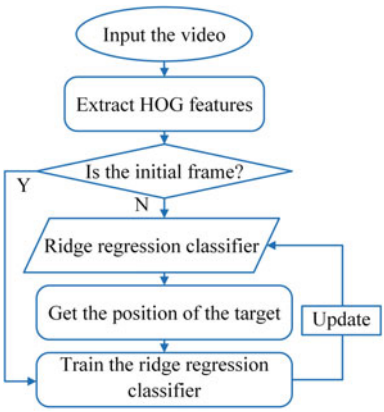



**Fig. 2.** Flowchart of the KCF tracking algorithm

### 2.3 Target Position Prediction

Kalman filter is a recursive estimator that implements the estimation of motion states through the principle of feedback control. From the dynamic target tracking in the previous step, the motion vector of the targets can be obtained. According to the historical motion states of the targets, the current states of the targets can be predicted. The basic state model of the targets includes the observation equation and the state equation as follows:

$$\begin{cases} Z_k = H_k X_k + V_k \\ X_{k,k-1} = A_{k,k-1} X_{k-1} + B_{k-1} U_{k-1} \end{cases} \tag{5}$$

where $Z_k$ represents the observed value of the target state, $H_k$ represents the observation matrix, $X_k$ represents the real state at time $k$, $X_{k,k-1}$ represents the target state estimated from the previous state, $A_{k,k-1}$ represents the state transition matrix from $k-1$ to $k$, $B_{k-1}$ represents the system control matrix at time $k-1$, $U_{k-1}$ represents the noise that influences the current motion law of the targets, such as acceleration and deceleration of the moving targets, $Q$ is the self-covariance of $U_{k-1}$, and $V_k$ represents the observation noise, which in this experiment represents the deviation during the tracking process. Therefore, the prediction equation of the next coordinate $X_{k+1,k}$ and its covariance matrix $P_{k+1,k}$ can be obtained as:

$$\begin{cases} X_{k+1,k} = A_{k+1,k} X_k + B_k U_k \\ P_{k+1,k} = A_{k+1,k} P_k A_{k+1,k}^T + B_k Q_k B_k^T \end{cases} \tag{6}$$

### 2.4 UAV Path Planning

This paper focuses on the optimization of ant colony algorithm which is slow in convergence and easy to fall into the local optimal solution. The probability of stagnation of the algorithm is reduced by limiting the increment of pheromone update and the dynamic change of volatile factors, also by using blending inheritance algorithm; thus, the UAV dynamic path planning can be realized. The specific steps are as follows:

Step 1: Information initialization:

$$\tau_{ij}(0) = m / d_{\min} \tag{7}$$

where $\tau_{ij}(0)$ represents the pheromone from target $i$ to target $j$, $m$ represents the number of ants, and $d_{\min}$ represents the distance between the closest two targets.

Step 2: Path construction by all ant state transition.

Calculate the probability of transition of each ant $k$ from target $i$ to target $j$ by the following formula at time $t$:

$$p_{ij}^k(t) = \begin{cases} \dfrac{\tau_{ij}^{\alpha}(ij)\eta_{ij}^{\beta}}{\sum_{s \in \text{allowed}_k} \tau_{ij}^{\alpha}(ij)\eta_{ij}^{\beta}}, & j \in \text{allowed}_k \\ 0, & \text{other} \end{cases} \tag{8}$$

where $\eta_{ij}(t) = 1/d_{ij}$ represents the degree of expectation of ants from target $i$ to target $j$, $\alpha$ represents the information heuristic factor, which indicates the influence of pheromone on the transition probability, $\beta$ represents the expected heuristic factor, which indicates the influence of visibility on the transition probability of ants, allowed $= \{1, 2, \ldots, n\} - \text{tabu}_k$ represents the targets which ant $k$ has not visited, and $\text{tabu}_k$ represents the targets that the ant $k$ has passed.

Step 3: Pheromone update according to the steps which ants construct.

Ants will release pheromone on the path and update the pheromone locally according to the following formula. In order to increase the difference between the best and the worst paths, this paper adopts the maximum and minimum ant system to strengthen the optimal solution and weaken the worst solution:

$$\tau_{ij}(t+1) = \begin{cases} (1-\rho)\tau_{ij}(t) + \varepsilon(L_{\text{worst}} - L_{\text{best}}), & \text{if } ij \text{ is the best path.} \\ (1-\rho)\tau_{ij}(t) - \varepsilon(L_{\text{worst}} - L_{\text{best}}), & \text{if } ij \text{ is the worst path.} \\ (1-\rho)\tau_{ij}(t), & \text{other} \end{cases} \tag{9}$$

where $\rho$ represents the pheromone volatility coefficient, $\varepsilon$ represents the pheromone enhancement coefficient, and $L_{\text{best}}, L_{\text{worst}}$ represent the length of the optimal path and the worst path, respectively, for all ants in this time.

Step 4: Update the number of iterations.

If the number of iterations is less than the set value $N_{\max}$, repeat step 2.

Finally, the ant colony will gradually converge on the best path through the constant updating of pheromone. The specific process is shown in Fig. 3.

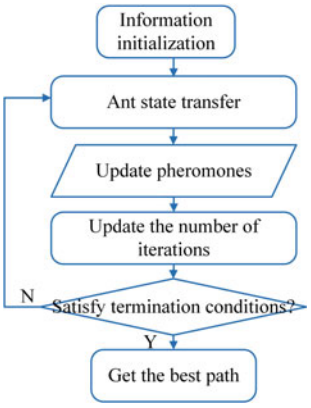



**Fig. 3.** Flowchart of ant colony algorithm

# 3 Simulation Experiment

## 3.1 Single-Target Tracking Experiment

In order to demonstrate the effectiveness and feasibility of the algorithm, there are two simulation experiments for the single-target tracking and multi-target tracking of UAV.

The first experiment is the simulation of UAV's approaching a single moving target. At first, a moving target is marked for tracking. Then, the marked target is tracked by the KCF tracking algorithm to obtain the target motion and position information, and the Kalman filter method is used to predict the moving target position. There are two UAVs for comparison in the experiment, one of which is an experimental UAV and the other is a comparison UAV. The experimental UAV is the UAV that tracks pre-targeted positions, and the comparison UAV is the UAV that tracks the current position of the target. In the experiment, the UAVs' movement speed was set to 2 pixels per frame, and the target movement speed was about 1 pixel per frame. The experimental results are shown in Fig. 4.

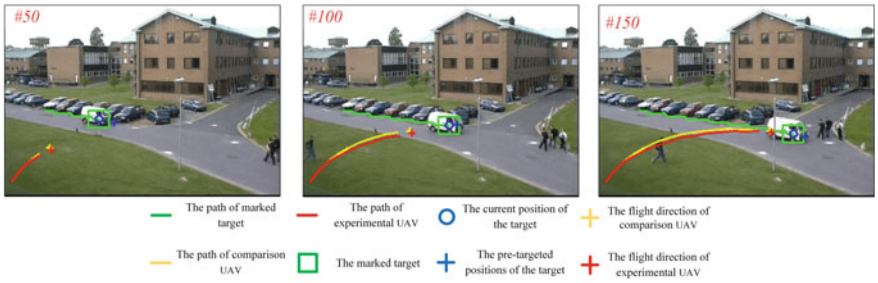



**Fig. 4.** Experimental results of the simulation of a single moving target

The result shows that the flight path of the experimental UAV is shifted in advance to the moving direction, and the required path is relatively shorter. It means the algorithm can improve tracking efficiency and optimize the UAV path.

## 3.2 Multi-target Tracking Experiment

The second experiment is the simulation of the path planning of multiple moving targets for UAV. Based on the first experiment, this experiment expands the single target to multi-targets, then by using the improved ant colony algorithm to plan the UAV's paths. The experimental results are as shown in Fig. 5.

The experimental result shows that the algorithm can effectively adapt to track the targets' movement and get the tracking sequence of the moving target to plan the UAV's paths. It means that the UAV path planning algorithm is feasibility. And at the 50th frame, two UAVs have the same sequence of tracking the targets. But at 70 frames, the experimental UAV changed the tracking experimental, but the tracking sequence of the comparison UAV did not change. At 90 frames, the tracking sequence

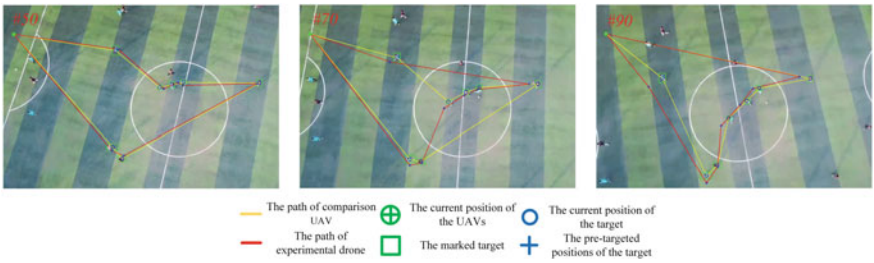



**Fig. 5.** Experimental results of the simulation of the path planning of multiple targets

of the comparison UAV changes to the same as that of the experimental UAV. It means that the UAV path planning algorithm can effectively predict the moving target positions, and the algorithm can update the tracking sequence of the moving targets in advance to optimize the UAV tracking path.

## 4 Conclusions

In this paper, aiming at the problem of UAV path planning for multiple moving targets, firstly, the KCF algorithm is used to track the moving targets to obtain the positions of the targets. Then, Kalman filter is used to predict the positions of the moving targets, and finally, the improved ant colony algorithm is used to complete the dynamic path planning to achieve the UAV path planning for multiple moving targets. The experimental results show that the proposed scheme can effectively complete the dynamic optimization of the UAV path, thus improving the intelligence of the drone in the scene of multiple moving target tracking.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China under Project No. 41701531. It was also supported in part by the Natural Science Foundation of Jiangsu Province under Project No. BK20170782. And this work was supported by the Open Research Fund of State Key Laboratory of Tianjin Key Laboratory of Intelligent Information Processing in Remote Sensing under grant No. 2016-ZW-KFJJ-01.

## References

1. Wang X. Vision-based detection and tracking of a mobile ground target using a fixed-wing UAV. Int J Adv Robot Syst. 2014;11(156):1–11. https://doi.org/10.5772/58989.
2. Han P, Chen M, Chen SD, et al. Path planning for UAVs based on improved ant colony algorithm. J Jilin Univ. 2013;31(1):66–72. https://doi.org/10.3969/j.issn.1671-5896.2013.01.011.
3. Henriques JF, Rui C, Martins P, et al. High-speed tracking with kernelized correlation filters. IEEE Trans Pattern Anal Mach Intell. 2015;37(3):583–96. https://doi.org/10.1109/tpami.2014.2345390.

4. Kalman RE. A new approach to linear filtering and prediction problems. Trans ASME J Basic Eng. 1960;82(1):35–45. https://doi.org/10.1115/1.3662552.
5. Dantzig G, Johnson S. Solution of a large-scale traveling-salesman problem. Oper Res. 2010;2(4):393–410. https://doi.org/10.2307/166695.
6. Dorigo M, Maniezzo V, Colorni A. Ant system: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern B. 1996;26(1):29. https://doi.org/10.1109/3477.484436.
7. Stutzle T, Hoos H. MAX-MIN ant system and local search for the traveling salesman problem. In: IEEE international conferences on evolutionary computation; 2002. p. 309–14. https://doi.org/10.1109/icec.1997.592327.
8. Li S, Zhang Y, Gong Y. The research on the optimal path of intelligent transportation based on ant colony algorithm. J Changchun Univ Sci Technol. 2015;4:122–6. https://doi.org/10.3969/j.issn.1672-9870.2015.04.027.

# A Least Square Dynamic Localization Algorithm Based on Statistical Filtering Optimal Strategy

Xiaozhen Yan[1,2(✉)], Zhihao Han[1], Yipeng Yang[1], Qinghua Luo[1,2], and Cong Hu[2]

[1] School of Information Science and Engineering, Harbin Institute of Technology at Weihai, Weihai, China
{yxz_heu, luoqinghua081519}@163.com
[2] Guangxi Key Laboratory of Automatic Detecting Technology and Instruments, Guilin University of Electronic Technology, Guilin, China
yiqi@guet.edu.cn

**Abstract.** In wireless sensor network localization, many anchor nodes and target node exchange information at specified time intervals to obtain the distance information between each anchor node and the target node. With this information, the coordinates of the target node can be achieved through the calculation of the positioning algorithm. However, as there are numerous negative factors like non-line-of-sight measurement, complex multipath fading, which leads to high-level localization error. To improve localization accuracy, an improved least square localization algorithm is proposed, which combines the least square localization method with the statistical filtering optimization strategy. The simulation results show that this algorithm can effectively reduce localization error and achieve more accurate localization.

**Keywords:** Wireless sensor networks · Localization · Least square method · Extended Kalman filter · Particle filter

## 1 Introduction

With the development of science and technology, people's demand for localization is increasing day by day, and the demand for localization accuracy is also greatly increased. As a mature positioning scheme, GPS has good user experience and positioning accuracy in an open and unobstructed environment. However, after encountering obstacles, the satellite signal strength decreases greatly, which leads to a significant decline in GPS positioning accuracy. And it is difficult to meet the requirements of high-precision positioning results. Wireless sensor network localization could solve the problem.

The use of wireless sensor network in a small range of high-precision positioning has become the current trend. Wireless sensor network is a self-organizing network composed of a large number of randomly distributed small nodes, which integrated with sensors, data processing units, and communication modules. There are two import steps in range-based localization method. Firstly, the distances between anchor nodes


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_165

and unknown node (to be localized) should be estimated via a specific method. And then localization result can be gained based on distances estimation results through specific localization computation method. When we choose a suitable algorithm, we can get high-precision positioning results.

The remainder of this paper is organized as follows: In Sect. 2, we review the related research works. Section 3 introduces the least square method and its implementation principle. Section 4 describes the proposed improved localization algorithm based on least square and statistical filtering strategy. Section 5 simulates the proposed algorithm and compares it with other related works. Finally, we conclude this paper.

## 2 Related Works

There are numerous negative factors during distance estimation and localization computation, which lead to different degrees of error in distance estimation and localization error, or even more serious consequence. Therefore, in recent years, scholars of various countries have proposed many optimization algorithms.

Literature [1] proposed a localization algorithm based on mobile anchor nodes. The author used mobile anchor node localization to avoid the accumulation of localization errors caused by multi-hop and long-distance transmission in wireless sensor networks. Literature [2] proposed a moving anchor node localization estimation algorithm based on weighted least square method. The author moves the anchor nodes along the linear model [3, 4]; at the same time, using the weighted least square method reduces the distance estimation error. This algorithm has good performance on localization estimation [5]. In the literature [6], the author designed an indoor wireless localization platform based on NanoLoc and used an improved particle filter algorithm. Compared with the ordinary least square method, it had a certain improvement in accuracy. In the literature [7], the problem of limited precision of the traditional least square localization algorithm based on the received signal strength indicator (RSSI) is addressed. Starting from the lower signal-to-noise influence, the author proposed a modified least square-BFGS localization algorithm based on dynamic *t*-test [8]. In reference [9], the author put forward a Monte Carlo mobile node localization algorithm based on least square method. Based on the continuity of the movement, this algorithm used the method of least square curve fitting and calculated where the unknown nodes may be in the next moment.

## 3 The Least Square Localization Algorithm

The principle of least square localization algorithm is shown in Fig. 1. Take $n$ reference nodes as $P_1$, $P_2$, $P_3$, …, $P_n$. We suppose their coordinates are $(x_1, y_2)$, $(x_2, y_2)$, $(x_3, y_3), \ldots, (x_n, y_n)$, respectively. The distance between them and the unknown node $D$ is $d_1$, $d_2$, $d_3, \ldots, d_n$. And suppose the coordinate of point $D$ is $(x, y)$, which to be determined. According to geometric constraint relation among anchor nodes and unknown node, we can get the equations as shown in (1).

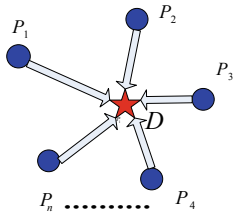



**Fig. 1.** Principle of least square localization algorithm

$$\begin{cases} (x_1 - x)^2 + (y_1 - y)^2 = d_1^2 \\ \vdots \\ (x_n - x)^2 + (y_n - y)^2 = d_n^2 \end{cases} \tag{1}$$

According to the idea of least square method, the target node coordinate $(x, y)$ should get the smallest error sum of squares between all measured distances $d_i (i = 1, 2, \ldots, n)$ and its corresponding actual distances $\sqrt{(x_i - x)^2 + (y_i - y)^2}$. The equation can be formed as followed.

$$\min_{x,y} \sum_{i=1}^{n} (\sqrt{(x_i - x)^2 + (y_i - y)^2} - x)^2 \underset{m}{\Delta} e(x, y) \tag{2}$$

Taking the partial derivative to $e(x, y)$ and let it be zero, we can get the equation as followed.

$$\frac{\delta e}{\delta x} = \sum_{i=1}^{n} 2(\sqrt{(x_i - x)^2 + (y_i - y)^2} - d_i) \frac{2(x_i - x) \cdot (-1)}{2\sqrt{(x_i - x)^2 + (y_i - y)^2}} = 0 \tag{3}$$

$$\frac{\delta e}{\delta y} = \sum_{i=1}^{n} 2(\sqrt{(x_i - x)^2 + (y_i - y)^2} - d_i) \frac{2(y_i - y) \cdot (-1)}{2\sqrt{(x_i - x)^2 + (y_i - y)^2}} = 0 \tag{4}$$

Then, we can simplify them and get the equations as followed.

$$\begin{cases} \sum_{i=1}^{n} \frac{\mu_i - d_i}{\mu_i} (x_i - x) = 0 \\ \sum_{i=1}^{n} \frac{\mu_i - d_i}{\mu_i} (x_i - x) = 0 \end{cases} \tag{5}$$

The coordinates of target nodes can be obtained by solving the equations.

## 4 The Improved Least Square Localization Algorithm

This section mainly introduces the optimized least square localization algorithm, which uses extended Kalman filter or particle filter to optimize localization results to improve localization accuracy. The following is mainly discussed for the nonlinear model.

### 4.1 The Least Square Localization Algorithm Based on Extended Kalman Filtering

For nonlinear filtering system, transforming it into an approximate linear filtering problem by linearization technique is a commonly used treatment method. And one of the most widely used is extended Kalman filter method. Extended Kalman filter based on linear Kalman filter, the core idea is that below. For a nonlinear system, we can expand the nonlinear function $f(*)$ and $h(*)$ into Taylor series around the filter value $X(k)$, and omit the second order and above terms. Then, we will get an approximate linear model, and apply linear Kalman filter to estimate filter result and other processes.

The flowchart of extended Kalman filter localization algorithm is shown in Fig. 2.

First of all, an approximate linear model is obtained by local linearization of the nonlinear motion model. For a nonlinear system, its dynamic equations can be expressed as shown in (6) and (7).

$$X(k+1) = f[k, X(k)] + G(k)W(k) \tag{6}$$

$$Z(k) = h[k, X(k)] + V(k) \tag{7}$$

Based on the local linearization characteristics of nonlinear functions, we can expand the nonlinear function (6) and (7) into Taylor series in first order around the filter value $\hat{X}(k)$. The equations can be formed as followed.

The state equation is presented as following:

$$X(k+1) = \phi(k+1|k)X(k) + G(k)W(k) + \phi(k) \tag{8}$$

The observational equation is noted as follows:

$$Z(k) = H(k)X(k) + \mathrm{y}(k) + V(k) \tag{9}$$

where the matrix is defined as follows, respectively:

$$\frac{\partial f}{\partial \hat{X}(k)} = \frac{\partial f\left[\hat{X}(k), k\right]}{\partial \hat{X}(k)}\bigg|_{\hat{X}(k)=X(k)} = \phi(k+1|k) \tag{10}$$

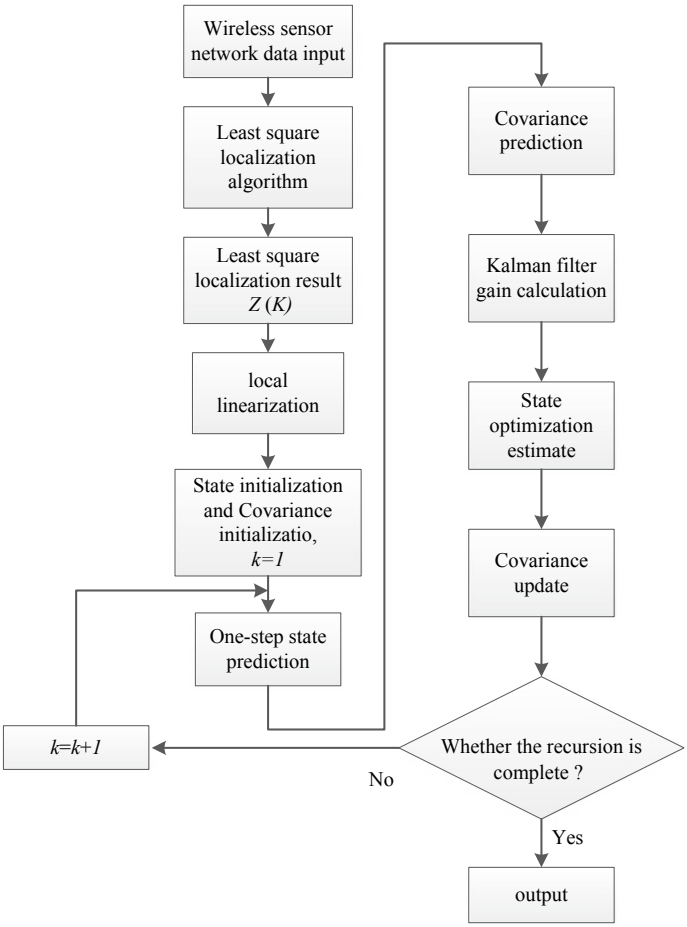



**Fig. 2.** Framework of extended Kalman filter localization algorithm based on least square method

$$f\left[\hat{X}(k),k\right]-\left.\frac{\partial f}{\partial X(k)}\right|_{\hat{X}(k)=X(k)}\hat{X}(k)=\phi(k) \tag{11}$$

$$\left.\frac{\partial h}{\partial \hat{X}(k)}\right|_{X(k)=\hat{X}(k)}=H(k) \tag{12}$$

$$y(k)=h\left[\hat{X}(k|k-1),k\right]-\left.\frac{\partial h}{\partial \hat{X}(k)}\right|_{X(k)=\hat{X}(k)}\hat{X}(k|k-1) \tag{13}$$

After local linearization, we can use the basic equations of linear Kalman filter to get the recurrence equations of extended Kalman filter.

One-step state prediction equation and covariance prediction equation are formed as followed.

$$\hat{X}(k|k+1) = f(\hat{X}(k|k)) \tag{14}$$

$$P(k+1|k) = \phi(k+1|k)P(k|k)\phi^T(k+1|k) + Q(k+1) \tag{15}$$

Then, we can use the formula (16) to get the Kalman filter gain.

$$K(k+1) = P(k+1|k)H^T(k+1)\left[H(k+1)P(k+1|k)H^T(k+1) + R(k+1|k)\right]^{-1} \tag{16}$$

Next, we can have the following equation for state optimization estimate:

$$\hat{X}(k+1|k+1) = \hat{X}(k+1|k) + K(k+1)\left[Z(k+1) - h\left(\hat{X}(k+1|k)\right)\right] \tag{17}$$

where $Z(k+1)$ is the localization result obtained by the least square localization algorithm.

Finally, update the covariance with the following equation so that it can be used for the next recursive calculation.

$$P(k+1) = [I - K(k+1)H(k+1)]P(k+1|k) \tag{18}$$

We can recursively compute the above equations to obtain the filtering estimate of the least square localization result.

### 4.2 The Least Square Localization Algorithm Based on Particle Filtering

Particle filter is an approximate Bayesian filtering algorithm based on Monte Carlo simulation. Its core idea is to use some discrete random sampling points to approximate the probability density function of system random variables. Based on that, we can use the sample mean instead of the integral operation to obtain the minimum variance estimate of the state.

The flowchart of particle filter localization algorithm is presented in Fig. 3.

For $N$ particles, we need to make predictions based on the state transfer function during sampling. The relationship is as follows.

$$X_k^{(i)} \sim q\left(X_k | X_{0:k-1}^{(i)}, Z_{1:k}\right) \tag{19}$$

When calculating the weight, substitute $\Delta(t) = Z(t) - Zpre(t)$ as the error value into the weight calculation equation, where $Zpre(t)$ is the observed value of the sampling point and $Z(t)$ is the localization result obtained by the least square method. The choice of $Z(t)$ affects the accuracy of the filter estimation.

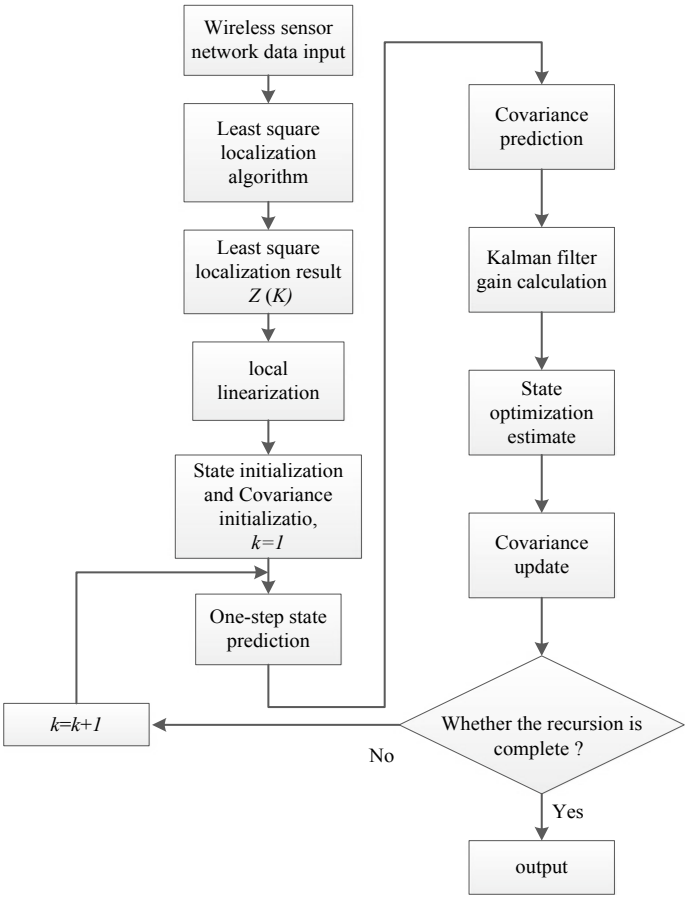



**Fig. 3.** Framework of particle filter localization algorithm based on least square method

Based on the position of the particle, we can get the corresponding weight. This estimation method will determine the effect of each particle, and by re-sampling, it will remove some particles that have little effect on the estimation of the filter. Then, we can estimate a result with the highest probability, which is our optimized localization result, by the weighted average method.

## 5 Evaluation

In this section, we will validate and evaluate the performance of our optimized least square localization algorithm. The performance of the localization is evaluated in terms of localization accuracy and computation time. We adopt absolute localization error ($e_{ER}$) to indicate localization accuracy, which is calculated according to (20), where $e_{ER}$ is the absolute localization error, $(\hat{x}_n, \hat{y}_n)$ is the localization result of the unknown node,

$(x_n, y_n)$ is the exact coordinate of unknown node, n is the serial number of the localization point, and $N$ is the total number of localization experiments. The smaller the $e_{ER}$ is, the more accurate the localization result becomes. Computational time can reflect the cost of the algorithm.

$$e_{ER} = \sqrt{(x_n - \hat{x}_n)^2 + (y_n - \hat{y}_n)^2} \tag{20}$$

During the simulation, five anchor nodes were randomly distributed in the 100 m * 100 m motion region. The simulation results are shown in Figs. 4, 5, and 6.

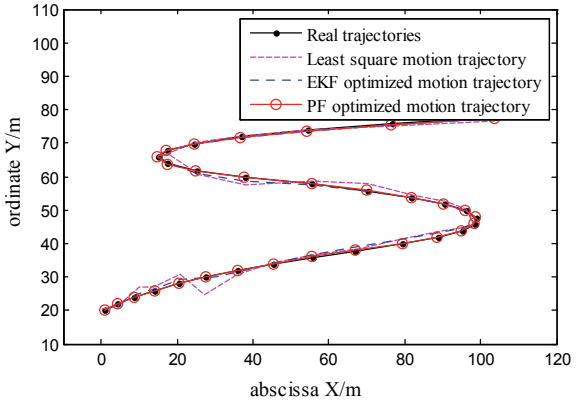



**Fig. 4.** Trajectory of different dynamic localization methods

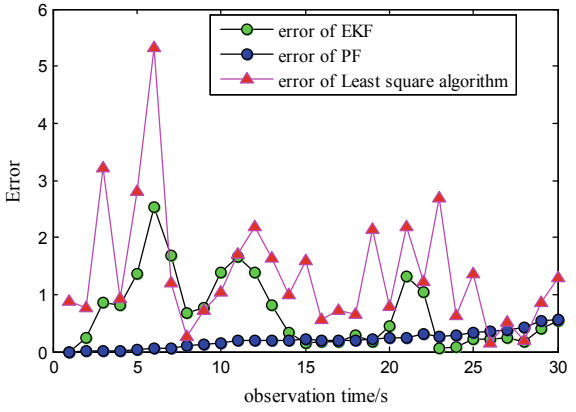



**Fig. 5.** Localization error of different methods

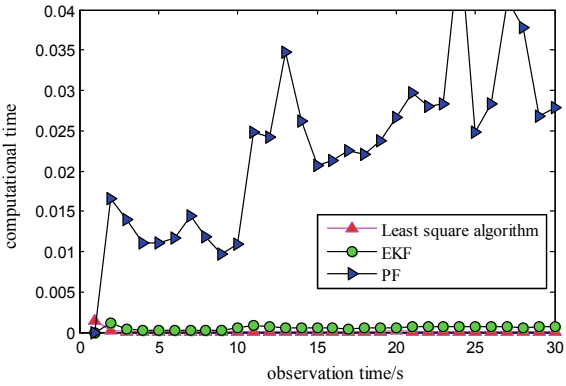



**Fig. 6.** Computational time of different methods

Under the nonlinear model, the localization trajectory obtained by different localization methods is illustrated in Fig. 4. It can be seen from Fig. 4 that all three localization methods can basically complete the task of localization.

Figure 5 is a comparison of localization errors of the three localization methods. Based on the distance between nodes measured by the wireless sensor, the least square localization algorithm can get the location result by solving the equation. However, due to noise and error, this algorithm has a relatively large localization error. As for extended Kalman filter algorithm based on the least square method, the localization error is reduced compared with the result before optimization. At the same time, localization errors fluctuate less. Least square localization algorithm based on particle filter has the best localization accuracy most of the time. However, because of accumulative error, the localization error will gradually increase.

Comparison of calculation time of three localization methods is shown in Fig. 6. The calculation time can reflect the overhead problem of the algorithm. In areas where real-time requirements are relatively high, this indicator is extremely concerned. As can be seen from Fig. 6, the least square method and the extended Kalman filter localization algorithm based on the least square method have shorter calculation time and better real-time performance. As for the particle filter localization algorithm based on the least square method, it has longer calculation time and more overhead than the previous two algorithms. Therefore, it is suitable for applications where the demand for localization accuracy is high, but the demand for real-time performance is not superior.

## 6 Conclusion

To improve the localization accuracy with a nonlinear dynamic model, we proposed an optimized least square localization algorithm. It can also be divided into two types of algorithms, which are based on the extended Kalman filter algorithm and the particle filter algorithm, respectively. In these two methods, we take advantage of filtering estimates to minimize the impact of negative factor during the localization procedure.

Relative to least square localization method, EKF algorithm based on the least square method has higher accuracy with approximate efficiency. As for PF algorithm based on the least square method, it has higher accuracy but less computational efficiency relative to the EKF.

**Acknowledgments.** The research presented in this paper is supported by the National Natural Science Foundation of China (61671174, 61601142), the Natural Science Foundation of Shandong Province of China (ZR2015FM027), WeiHai Research program of Science and Technology (16), the Laboratory of Satellite Navigation System and Equipment Technology (EX166840037, EX166840044), the Guangxi Key Laboratory of Automatic Detecting Technology and Instruments (YQ18206,YQ15203), the Natural Scientific Research Innovation Foundation of the Harbin Institute of Technology (HIT.NSRIF.2015122), the State Key Laboratory of Geo-information Engineering (SKLGIE2014-M-2-4), and Discipline Construction Guiding Foundation in Harbin Institute of Technology (Weihai) (WH20150211).

## References

1. Peng Y, Wang D. A review: wireless sensor network localization. J Electron Meas Instrum. 2011;25(5):389–99.
2. Kim E, Kim K. Distance estimation with weighted least squares for mobile beacon-based localization in wireless sensor networks. IEEE Signal Process Lett. 2010;17(6):559–62.
3. Cheng J, Yang L, Li Y, et al. Seamless outdoor/indoor navigation with WIFI/GPS aided low cost Inertial Navigation System. Phys Commun. 2014;13(PA):31–43.
4. Ades M, Van Leeuwen PJ. An exploration of the equivalent weights particle filter. Q J R Meteorol Soc. 2013;139(672):820–40.
5. Huang CH, Lee LH, Ho CC, et al. Real-time RFID indoor positioning system based on Kalman-filter drift removal and Heron-Bilateration location estimation. IEEE Trans Instrum Meas. 2015;64(3):728–39.
6. Jiaojiao W. Research on wireless localization technology of indoor moving target. TianJing University master thesis; 2013. p. 19–31.
7. Juan M, Li H, Yanan L, et al. Modified least squares-BFGS positioning algorithm based on dynamic T-test. Comput Appl Softw. 2016;22(6):126–9.
8. Mirzaei HR, Akbari A, Gockenbach E, et al. A novel method for ultra-high-frequency partial discharge localization in power transformers using the particle swarm optimization algorithm. IEEE Electr Insul Mag. 2013;29(2):26–39.
9. Yao Y. Polaronic quantum diffusion in dynamic localization regime. New J Phys. 2017;19(4): 043015.

# Design and Implementation of an UWB-Based Anti-lose System

Yue Wang(✉) and Yunxin Yuan

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
ywang_tjnu@163.com

**Abstract.** Accurate and reliable distance information is essential to a variety of wireless applications, and ultra-wideband (UWB) signal can theoretically achieve centimeter-level ranging accuracy. In this paper, we design and implement an UWB-based anti-lose system, using commercial ScenSor DWM1000 module. The centimeter-level ranging accuracy of the system is shown under both line-of-sight (LoS) and non-line-of-sight (NLoS) experimental conditions.

## 1 Introduction

Accurate and reliable distance information is essential to a variety of wireless applications, such as anti-lose devices, indoor/outdoor positioning and navigation, and wireless sensor networks [1]. High-accuracy ranging can be achieved by using high-resolution time measurement, which is reversal to the bandwidth of the ranging signal [2]. Ultra-wideband (UWB) signal has huge bandwidth, i.e., more than 500 MHz, thus it can theoretically achieve centimeter-level ranging accuracy [3].

In this paper, we design and implement an UWB-based anti-lose system using ScenSor DWM1000 module, which is the first commercial module conformed to the IEEE 802.15.4a standard with low cost [4]. In the experiments, ranging performance is measured under two perpendicular antenna radiation angles and two signal prorogation conditions, i.e., line-of-sight (LoS) and non-line-of-sight (NLoS).

The rest of the paper is organized as follows. Section 2 introduces the design and implementation of the anti-loss system. Section 3 gives the experimental ranging results. A conclusion in Sect. 4 wraps up this paper.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_166

## 2 System Design and Implementation

The anti-loss system contains a pair of nodes; the distance between them is measured using two-way time-of-arrival (TW-ToA) ranging protocol [5], as shown in Fig. 1. The distance between node A and node B is calculated by

$$\hat{d}_{AB} = c \cdot \hat{t}_p = c \cdot \left( \frac{t^A_{RTT} - t^B_{TAT}}{2} \right), \tag{1}$$

where $c$ is the propagation speed of UWB signal, $\hat{t}_p$ is the calculated propagation time, $t^A_{RTT}$ is the round-trip-time (RTT) measured by node A, and $t^B_{TAT}$ is the turn-around-time (TAT) measured by node B.

Each node in Fig. 1 contains four subsystems, as shown in Fig. 2; they are power subsystem, UWB ranging subsystem, micro-controller subsystem, and alarm subsystem.

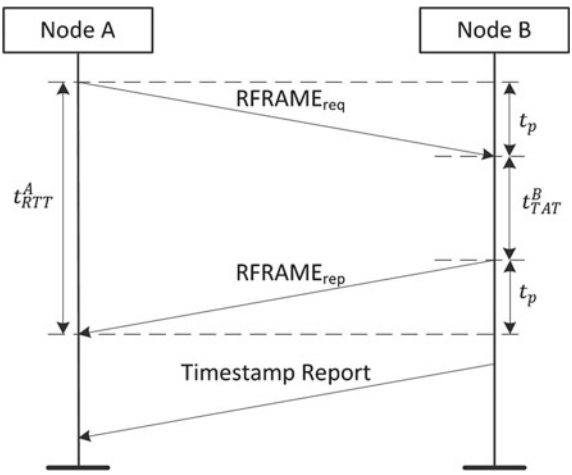



**Fig. 1.** TW-ToA ranging protocol.

The key element of the UWB ranging subsystem is ScenSor DWM1000 module and its peripheral circuit, which integrates DW1000 chip, onboard antenna, power management, clock management, and serial peripheral interface (SPI). By using DWM1000 module, the size of ranging node can be reduced and the RF circuit design is omitted. The micro-controller subsystem can configure the registers of DWM1000 and control its send/receive status via SPI.

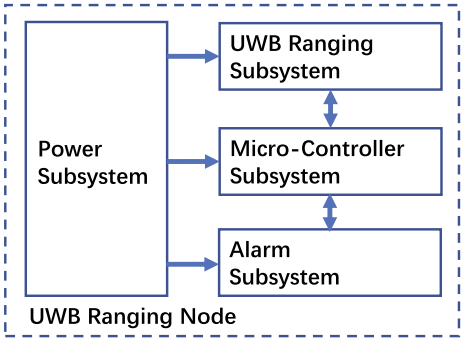



**Fig. 2.** Hardware structure of the ranging node.

The micro-controller subsystem adopts low-power and ARM-embedded STM32F411 chip. By designing the embedded program, the micro-controller can read the ranging information from the UWB ranging subsystem and transmit this information to the computer display via SPI to conduct ranging experiments.

The power subsystem uses single lithium battery TP054 to provide 5V power source, which can be charged by USB port.

The alarm subsystem contains a buzzer and a 2-bit dial switch. By using the 2-bit dial switch, there are four kinds of alarm distance settings; they are 0.5, 1, 5, and 10 m. If the ranging result is larger than the setting distance, the buzzer will produce the alarm sound.

The implemented ranging node is shown in Fig. 3.

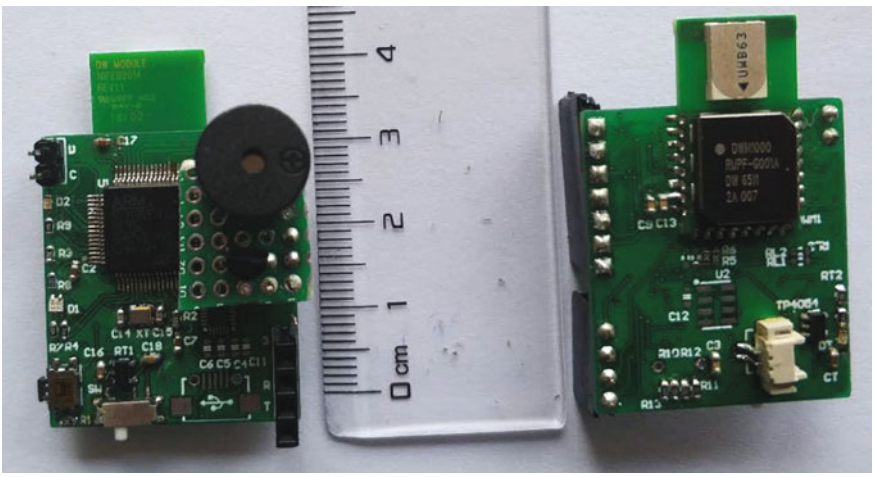


**Fig. 3.** UWB ranging node prototype.

## 3 Experimental Ranging Results

In our ranging experiments, one ranging node is fixed on a tripod, the other ranging node is mounted on a moving cart. The distance between two nodes is preset, and the increasing distance step is 1 m. For each distance, 50 ranging measurements are carried out and recorded. We define the ranging error between node A and node B as

$$e_{AB} = \left| \hat{d}_{AB} - d_{AB} \right|, \tag{2}$$

where $\hat{d}_{AB}$ is the measured distance between node A and node B, and $d_{AB}$ is the real distance between them.

In the first experiment, we measure the ranging accuracy for two perpendicular directions, i.e., 0° and 90°, between the antenna on each ranging node in the indoor environments, since the antenna gain is different for different directions. Two antennas on two nodes are parallel for 0°, and two antennas on two nodes are perpendicular for 90°.

The cumulative distribution function (CDF) for two directions is shown in Figs. 4 and 5. Under both 0° and 90° directions, the ranging error is less than 0.07 m with about 50% probability, and less than 0.15 m with about 90% probability.

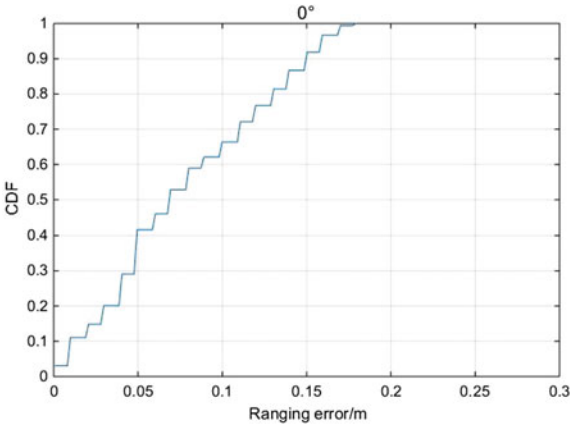



**Fig. 4.** CDF of 0° direction.

The box plot for two directions is shown in Figs. 6 and 7. Although the median for both directions is small, the outliers for 90° are larger than that for 0°.

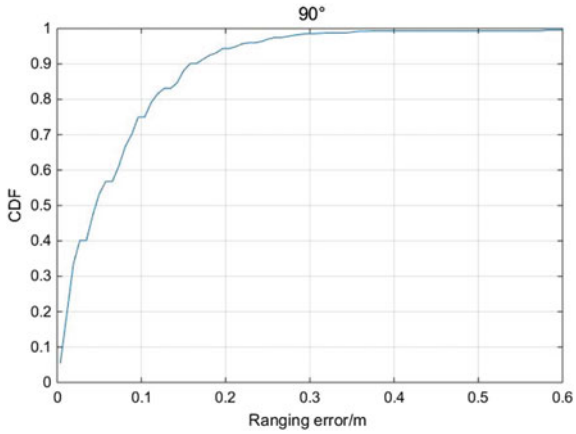



**Fig. 5.** CDF of 90° direction.

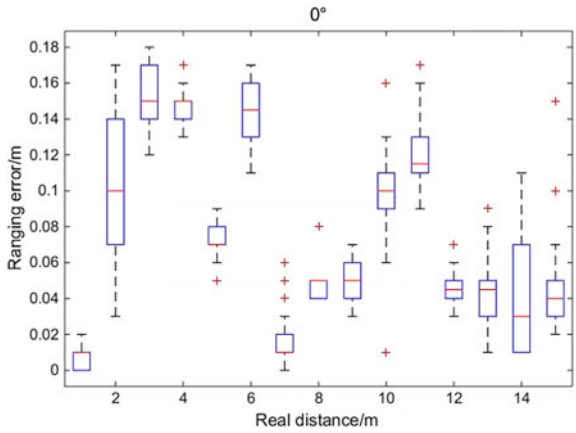



**Fig. 6.** Box plot of 0° direction.

In the second experiment, we compare the ranging accuracy for LoS and NLoS conditions. The experimental setting is shown in Fig. 8, where the NLoS condition is caused by a wooden door with 10 cm thickness.

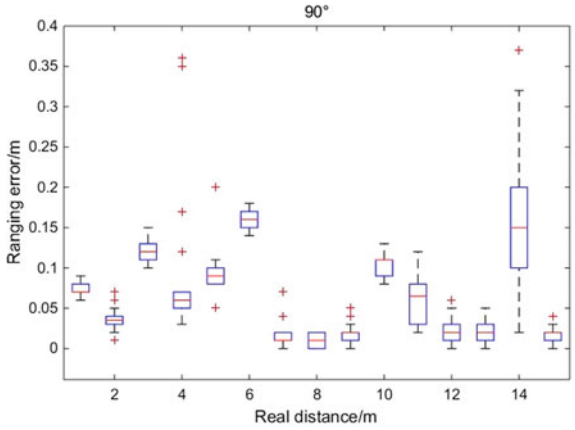



**Fig. 7.** Box plot of 90° direction.

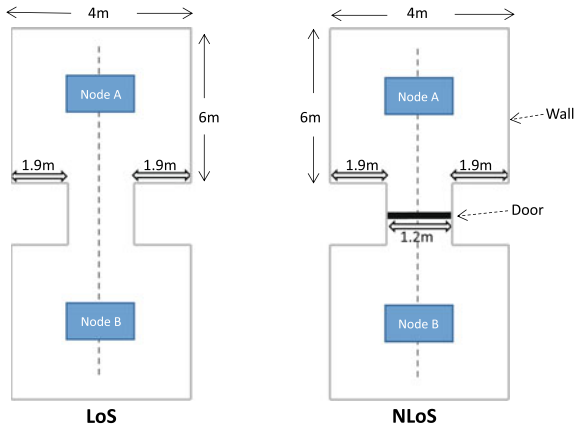



**Fig. 8.** LoS and NLoS ranging experimental settings.

The CDF for ranging error in LoS and NLoS conditions is shown in Figs. 9 and 10, respectively. Under LoS condition, the ranging error is less than 0.1 m with about 50% probability, and less than 0.13 m with about 90% probability. Under NLoS condition, the ranging error is less than 0.17 m with about 50% probability, and less than 0.25 m with about 90% probability.

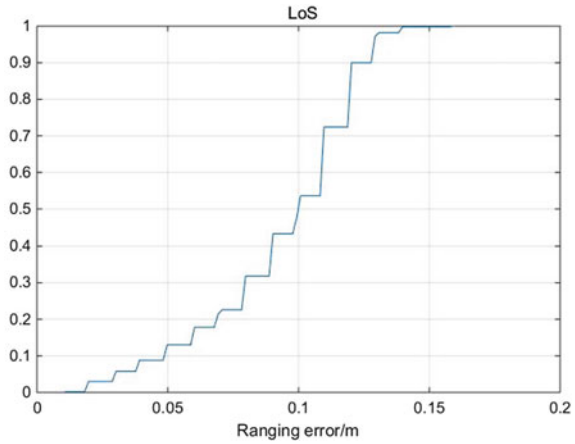



**Fig. 9.** CDF of LoS condition.

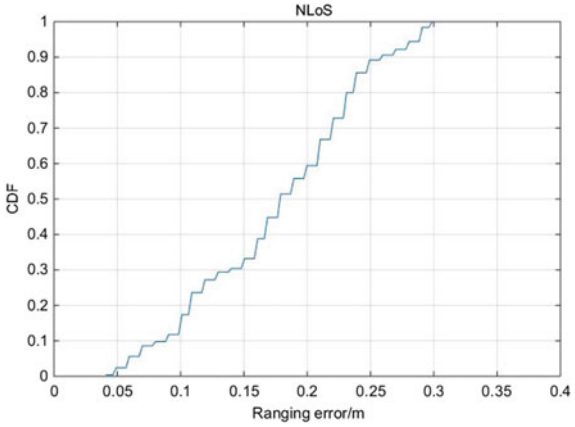



**Fig. 10.** CDF of NLoS condition.

The box plot for ranging error in LoS and NLoS conditions is shown in Figs. 11 and 12, respectively. The ranging error becomes larger in NLoS condition due to the signal block by the door. Even in the NLoS condition, the ranging error is relatively low thanks to the good penetrating property of UWB signal.

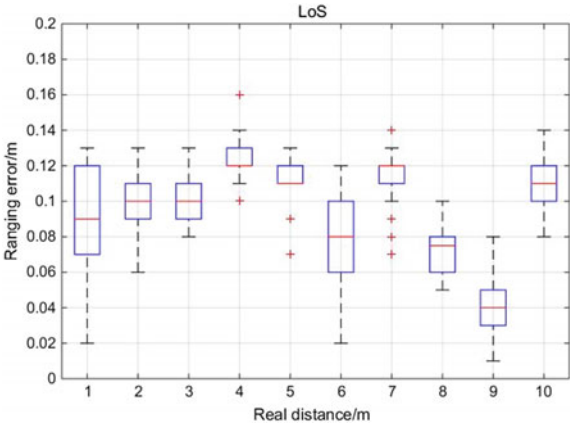



**Fig. 11.** Box plot of LoS condition.

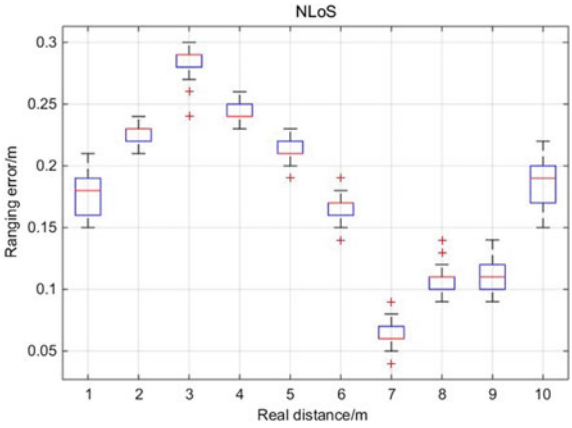



**Fig. 12.** Box plot of NLoS condition.

## 4 Conclusions

In this paper, an UWB-based anti-loss system is design and implemented. The centimeter-level accuracy of the system is shown under both LoS and NLoS experimental environments.

**Acknowledgments.** This work was supported by the Scientific Research Plan Project of the Committee of Education in Tianjin (No. JW1708).

## References

1. Dardari D, Luise M, Falletti E. Satellite and terrestrial radio positioning techniques: a signal processing perspective. Oxford: Elsevier; 2012.
2. Gezici S, Poor HV. Position estimation via ultra-wide-band signals. Proc IEEE. 2009;97(2):386–403.
3. Dardari D, Conti A, Ferner U, Giorgetti A, Win MZ. Ranging with ultrawide bandwidth signals in multipath environments. Proc IEEE. 2009;97(2):404–26.
4. www.decawave.com
5. Wang C, Wang Y, et al. Ultra-wideband ranging system prototype design and implementation. In: Proceedings of the international conference in communications, signal processing and systems (CSPS), Harbin; July 14–16; 2017. p. 1848–55.

# Indoor and Outdoor Seamless Localization Method Based on GNSS and WLAN

Yongliang Sun, Jing Shang(✉), and Yang Yang

School of Computer Science and Technology,
Nanjing Tech University, Nanjing 211816, Jiangsu, China
1254938049@qq.com

**Abstract.** Localization technology has been widely applied in various fields such as military investigation, natural disaster prevention, address search, and travel route planning. In order to guarantee the coverage range and localization performance of localization technology in both indoor and outdoor environments, research on indoor and outdoor seamless localization using global navigation satellite system (GNSS) and wireless local area network (WLAN) has attracted lots of attention. In this paper, a seamless localization method based on GNSS and WLAN is proposed. The method is able to switch smoothly from GNSS to WLAN localization in indoor and outdoor environments and outperforms either the GNSS localization or WLAN trilateration localization.

**Keywords:** Global Navigation Satellite System · Wireless Local Area Network · Seamless Localization · Trilateration

## 1 Introduction

In recent years, with the emergence of smart terminal devices and the rapid development of Internet of Things, people have more demands on location-based services (LBS) [1–3]. In outdoor environments, people can get satisfactory LBS with global navigation satellite system (GNSS) that consists of Beidou, GLONASS, global positioning system (GPS), and so on [4,5]. However, in indoor environments, satellite signals that are blocked by buildings are usually very weak [6]. The terminal devices can occasionally receive the satellite signals indoors. So indoor localization has been one of the research hot spots. Currently, wireless local area network (WLAN) is almost ubiquitous in people's daily life. WLAN-based indoor localization systems are able to meet the people's requirements for high-accuracy and low-cost LBS [7,8]. Therefore, in order to achieve seamless localization in indoor and outdoor environments, we propose an indoor and outdoor seamless localization method based on GNSS and WLAN.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_167

The proposed localization method has strong localization reliability and is able to switch smoothly from GNSS to WLAN localization in indoor and outdoor environments.

## 2 Related Localization Methods

### 2.1 Satellite-Based Localization

Satellite-based localization technology usually locates a terminal device with four satellites as shown in Fig. 1. The ground terminal device measures the satellite signals and computes the location coordinates and clock bias. Let the device location coordinates be $(x, y, z)$, $i$th satellite location and satellite time be $(x_i, y_i, z_i, s_i)$, $i = 1, 2, 3, 4$, the signal reception time denoted by the terminal device clock be $t'_i$, and the true signal reception time from $i$th satellite be $t_i = t'_i - b$, where $b$ is the clock bias of the terminal device. For $i$th satellite, we can have the distance between the terminal device and $i$th satellite as follows:

$$d_i = (t'_i - b - s_i) \times c \tag{1}$$

where $c$ is the speed of light and $d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}$. When four satellites are available, we have four equations to determine the four unknown parameters $(x, y, z, b)$.

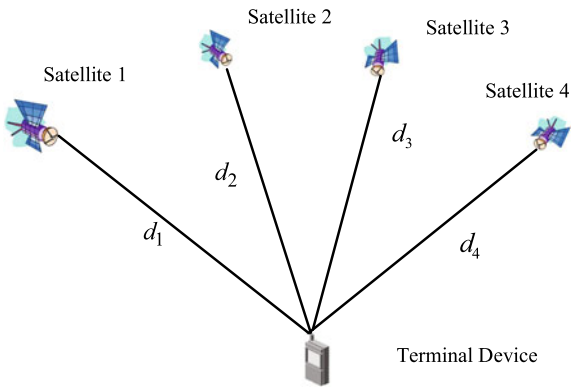



**Fig. 1.** Satellite-Based Localization Method

### 2.2 WLAN Trilateration Localization

When WLAN signals are transmitted, there will be power loss that is closely related to the propagation distance. WLAN trilateration localization is actually based on this relationship to obtain the propagation distances between a terminal

device and access points (APs). In our experiment, we use the propagation model denoted by (2) to estimate the distances.

$$r = -(10N \log_{10} d + A) \tag{2}$$

where, $r$ is the RSS value measured by a WLAN terminal device, $d$ is the distance between the terminal device and an AP, $N$ and $A$ are propagation model parameters. With the propagation model, three distances between the terminal device and APs are estimated, then WLAN trilateration localization can be performed [5].

## 3 Experimental Setup and Results of GNSS and WLAN Localization

### 3.1 Experimental Setup

The experimental area is a part of the top floor of the School of Computer Science and Technology in Nanjing Tech University. As shown in Fig. 2, there is an outdoor terrace with dimensions of 20.4 m × 9.2 m, which makes the floor a proper experimental area for indoor and outdoor seamless localization. We totally select 46 testing points (TPs) with 0.6 m gaps in the outdoor terrace and indoor office area denoted by the green dots. We collect satellite signals and RSS samples using an OPPO smart phone. The smart phone is placed on a tripod with a height of 1.2 m.

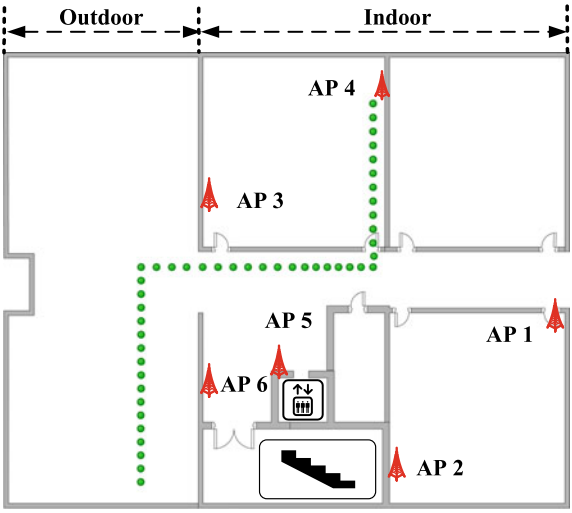



**Fig. 2.** Experimental Area Plan.

### 3.2 GNSS Localization Experiment

As shown in Fig. 2, 20 TPs are in the outdoor terrace and the other 26 TPs are in the indoor environment. We find that when we enter into the building from the terrace about 4 meters, satellite signals are difficult to obtain. Therefore, besides the 20 outdoor TPs, we collect satellite signals at extra six indoor TPs and calculate the localization results of the 26 TPs using GNSS. Then, we transform the localization results into a coordinate system we establish based on the experimental area plan. Within the signal collection duration, some latitude and longitude outliers appear, which may be caused by the ionosphere or troposphere effects, satellite clock errors, or human interference factors. In order to improve the localization accuracy, we filter the outliers and then sample the data. The means of the latitude and longitude are used to compute the localization coordinates. As shown in Fig. 3, the localization errors of the 26 TPs are calculated and the mean error is 5.13 m.

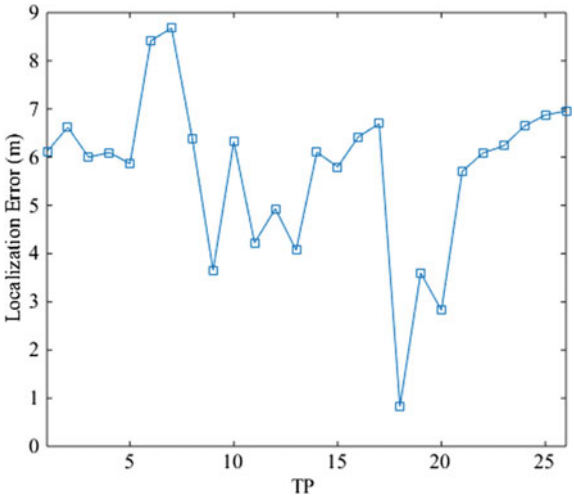



**Fig. 3.** GNSS Localization Errors.

### 3.3 WLAN Localization Experiment

The WLAN localization experiment is also carried out under the experimental area shown in Fig. 2. A total of six APs of WLAN are mounted in the experimental area with a height of 2.2 m and 60 RSS samples are collected at each TP. After filtering the RSS outliers from the six APs, the means of RSS samples collected at each TP are shown in Fig. 4.

According to empirical data, we set the propagation model parameters $N$ and $A$ be 1.73 and 21.31, respectively. Using the RSS data measured at each

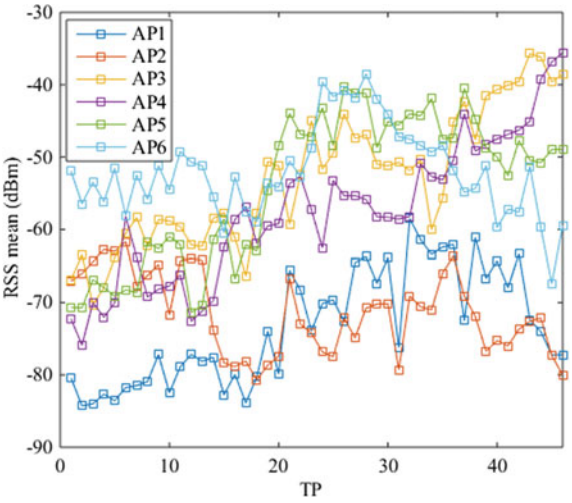



**Fig. 4.** Means of RSS Samples Collected at TPs.

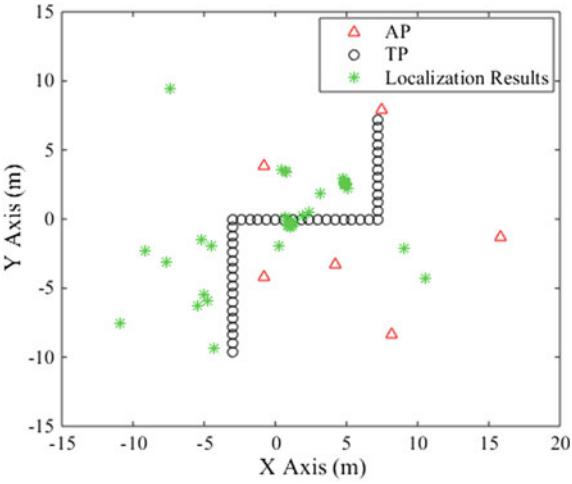



**Fig. 5.** Trilateration Localization Results.

TP, the distances between each TP and three corresponding APs are computed, so trilateration localization can be performed. The localization results as well as actual locations of TPs and APs are shown in Fig. 5. The mean error of the WLAN trilateration localization is 6.91 m.

## 4 Experiment of Indoor and Outdoor Seamless Localization

The experimental area can be divided into three subareas: GNSS localization subarea, handover subarea, and WLAN localization subarea. Satellite signals are used in the GNSS localization subarea and RSS data are used in the WLAN localization subarea. In the handover subarea, RSS may be strong and satellite signals are also available, so localization method can be selected according to a handover mechanism. We select the nearest three APs to the handover subarea that are AP3, AP5, and AP6, and set an RSS threshold $T_{rss}$ be $-65$ dBm for localization method selection. When the measured RSS values from the nearest three APs are smaller than $-65$ dBm, then GNSS localization is selected in the handover subarea. If these RSS values are equal or greater than the threshold $T_{rss}$, then the counting parameter $\alpha$ increases by 1, otherwise the counting parameter $\alpha$ will multiply with a reduction rate $\beta$, which can be denoted by:

$$\begin{cases} \alpha(i) = \alpha(i-1) + 1, rss(i) \geq T_{rss} \\ \alpha(i) = \alpha(i-1) \times \beta, rss(i) < T_{rss} \end{cases} \tag{3}$$

After the counting parameter $\alpha$ is greater than a counting threshold, which means the RSS values are equal or greater than the threshold $T_{rss}$ for a while, then WLAN-based trilateration localization is applied. The general flowchart of the handover mechanism is shown in Fig. 6.

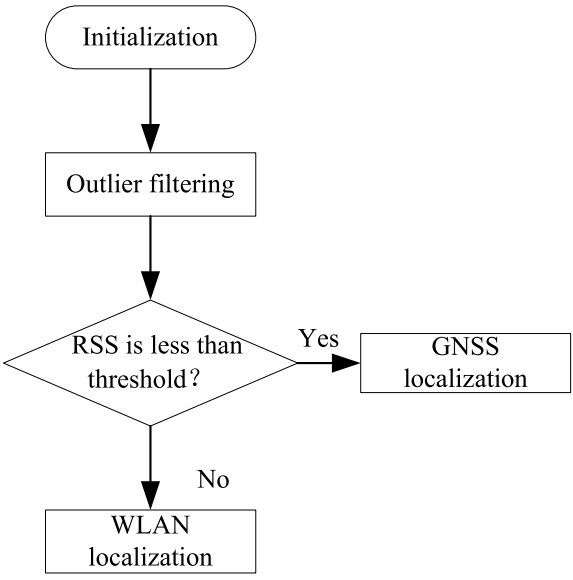



**Fig. 6.** Means of RSS Samples Collected at TPs.

From the experimental results, we find that the localization coordinates of outdoor TP1, TP2, ..., TP18 are computed using GNSS localization and the localization coordinates of the other TPs are computed using WLAN trilateration localization. The localization errors of all the TPs are shown in Fig. 7. After performing the indoor and outdoor seamless localization, the mean error of the localization results is reduced to 4.60 m. Compared with the mean errors of 5.13 m using GNSS localization only and 6.91 m using WLAN trilateration localization only, the proposed indoor and outdoor seamless localization method outperforms any of the two localization methods.

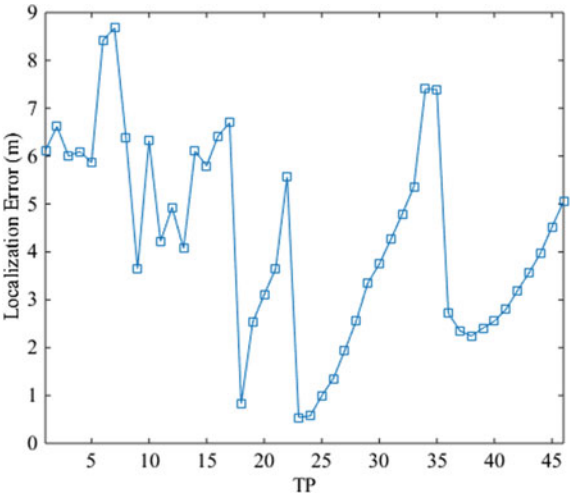



**Fig. 7.** Seamless Localization Errors.

## 5 Conclusions

This paper proposes an indoor and outdoor seamless localization method based on GNSS and WLAN. The method utilizes a handover mechanism to achieve seamless localization using GNSS and WLAN. The handover mechanism is used to select the GNSS localization or WLAN trilateration localization in the handover subarea. The experiment is performed in a real environment with an outdoor terrace and indoor office area. The experimental results demonstrate that the proposed indoor and outdoor seamless localization method using GNSS and WLAN outperforms either the GNSS localization or WLAN trilateration localization.

**Acknowledgments.** This work was supported by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant No. 16KJB510014, the Natural Science Foundation of Jiangsu Province under Grant No. BK20171023, and the National Natural Science Foundation of China under Grant No. 61701223.

## References


1. Gu YY, Lo A, Niemegeers I. A survey of indoor positioning systems for wireless personal networks. IEEE Commun Surv Tutorials. 2009;11(1):13–32.
2. Han GJ, Jiang JF, Zhang CY, et al. A survey on mobile anchor node assisted localization in wireless sensor networks. IEEE Commun Surv Tutorials. 2016;18(3):2220–43.
3. Zhou, M., Tang, Y.X., Tian, Z.S., Xie, L.B, Nie, W. Robust neighborhood graphing for semi-supervised indoor localization with light-loaded location fingerprinting. IEEE Internet Things J. 2017;(99): 1–1.
4. Kim J, Cheng J, Guivant J, Nieto J. Compressed fusion of GNSS and inertial navigation with simultaneous localization and mapping. IEEE Aerosp Electron Syst Mag. 2017;32(8):22–36.
5. Zou DY, Meng WX, Han S, He K, Zhang ZZ. Toward ubiquitous lbs: multi-radio localization and seamless positioning. IEEE Wirel Commun. 2016;23(6):107–13.
6. Sun YL, Xu YB. Error estimation method for matrix correlation-based Wi-Fi indoor localization. KSII Trans Internet and Inf Syst. 2013;7(11):2657–75.
7. Zhou M, Tang YX, Tian ZS, Xie LB, Geng XL. Semi-supervised learning for indoor hybrid fingerprint database calibration with low effort. IEEE Access. 2017;5(1):4388–400.
8. Sun YL, Meng WX, Li C, Zhao N, Zhao KL, Zhang NT. Human localization using multi-source heterogeneous data in indoor environments. IEEE Access. 2017;5:812–22.

# Land Subsidence Monitoring System Based on BeiDou High-Precision Positioning

Yuan Chen[1], Xiaorong Li[2], Yue Yue[1], and Zhijian Zhang[1(✉)]

[1] Nanjing University, Nanjing, China
{281508958,765727692}@qq.com,njuzzj@nju.edu.cn
[2] Traffic Business Department of CIECC, Beijing, China
13911537198@163.com

**Abstract.** Land subsidence is a geological disaster caused by natural or human activities. The rate of change in early settlements is often extremely small and presents a challenge to monitoring. This experiment includes BeiDou positioning, multiple antenna, and high-precision baseline solution. It developed the BeiDou deformation monitoring system and used static relative positioning for high-precision land subsidence monitoring. We have adopted integrated hardware design, equipped with a variety of communication modules, satellite receivers, and embedded module in one. In addition, we have developed the corresponding communication protocol for data transmission. Finally, a corresponding monitoring interface software was designed on the client to intuitively reflect the settlement process in a graphical manner.

**Keywords:** Land subsidence · Baseline solution · High-precision positioning

## 1 Introduction

Since entering the twenty-first century, China has made remarkable achievements in the process of modernization. However, in the process of rapid development, many potential hidden dangers are often ignored by people. Land subsidence is one of them, which is a kind of local subsidence movement caused by the loosening of underground structures and stratum crushing due to human engineering activities, resulting in the reduction of the elevation of the upper crust.

The secondary consequences of land subsidence have caused heavy casualties and loss of social wealth. In order to avoid or reduce the disasters caused by land subsidence, it is necessary to carry out effective monitoring in the early stage of land subsidence.

Current monitoring methods are mainly divided into traditional manual monitoring and automatic monitoring systems. In the last century, land subsidence monitoring has been dominated by traditional methods which is often difficult


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_168

to meet the monitoring needs due to the poor accuracy. With the development of information, communication, and automation technology, automatic monitoring system has been rapidly developed in recent years. Compared to traditional methods, the automatic monitoring system has advantages such as real time, accuracy, and stability.

## 2 Measurement Technology

There are many ways in traditional geodesy, such as leveling, GNSS, InSAR, and layerwise mark [1]. Among them, GNSS measurement technology has been developed rapidly in the past decades. It may provide high-precision three-dimensional coordinate which can be widely used in land subsidence monitoring. The relative baseline precision of GNSS can achieve $110^{-7}$.

Current GNSS measurement technology mainly includes precise point positioning (PPP), differential positioning, and network real-time kinematic (RTK).

### 2.1 Precise Point Positioning

Precise point positioning, also called as absolute positioning, uses single receiver for positioning. GNSS performances are optimal in an open sky when many satellites are in view and the signals are uncorrupted [2]. After years of development, PPP has been widely used in high-precision measurement, satellite orbit determination, aeronautical measurements, and surface deformation monitoring. The system architecture of PPP is relatively simple, while precise satellite orbit and clock products required by it always suffer a latency [3].

Traditional precise point positioning using undifferenced and ionosphere-free pseudo-range and phase combination measurements can obtain positioning results of centimeter level [4].

### 2.2 Differential Positioning

Differential positioning, also known as relative positioning, places some GPS receivers as reference station for observation. Based on the known precise coordinates of the reference station, the distance correction to satellite is calculated, and the reference station sends data continuously. After years of improvement, real-time differential positioning technology has been well developed which can significantly improve the positioning accuracy and reliability [5].

### 2.3 Network RTK

Network RTK is a new technology based on continuous operation reference station (CORS) network and conventional RTK technology which have been widely used in real-time high-precision navigation, survey, and mapping [6].

## 3 Land Subsidence Monitoring System

For a real-time precise positioning service, at least three components including precise orbit determination (POD), precise clock estimation (PCE), and precise point positioning (PPP) are necessary.

Multi-global navigation satellite system (GNSS) combined positioning [7] has become an inevitable trend in GNSS-based navigation. In addition to BeiDou as the main reference, the system also introduced GPS, GLONASS, and GALILEO auxiliary navigation systems.

### 3.1 System Schematic

The automatic land subsidence monitoring system mainly includes data receiving system, communication network, and monitoring center as Fig. 1 shows.

The receiving system consists of several GNSS stations, and satellite raw data will be transmitted to communication modules and then be forwarded to communication network to the cloud server.

Data of all the station will be transmitted to the solution terminal. When completed, results will be uploaded to the cloud server and broadcast to the monitoring center.

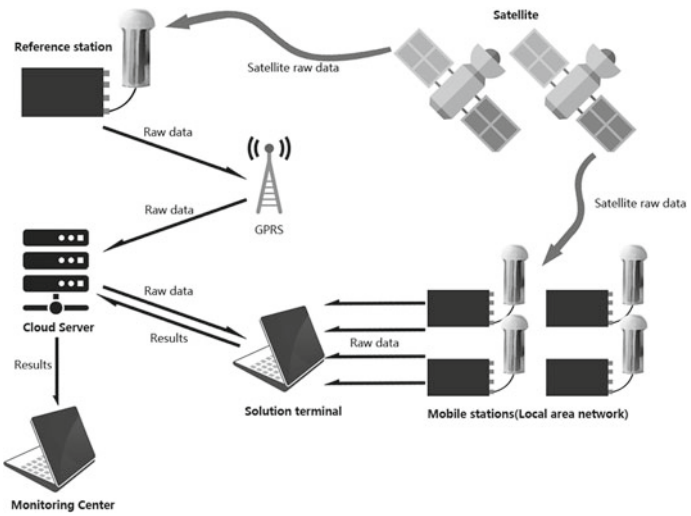



**Fig. 1.** Land subsidence monitoring system

### 3.2 Receiving System

In order to obtain data in a specific format and verify it, commands need inputting to the serial port of GNSS receiver. As for stability and safety, and rel-

atively low real-time requirements, ARM embedded modules are recommended to be applied in the subsystem.

The GNSS receiver and ARM embedded module together form a receiving system. Orbit and ephemeris parameters will be transmitted to the GNSS receiver through high-frequency (HF) antenna. ARM embedded module is responsible for messaging with the receiver and as the controlling core.

### 3.3 Communication Network and Monitoring Center

In actual situations, the reference point of geological stability often exceeds the distance of the monitoring point beyond the scope of the general local area network, so that the cloud server will play the role of forwarding data.

## 4 Data Processing

The processing of GPS measurements adopts GAMIT/GLOBK, which is a comprehensive GPS analysis package developed by MIT, Scripps Institution of Oceanography, and Harvard University [8]. IRTF2000 is adopted as reference frame, and the model is set to the baseline solution (detailed in Fig. 2) with an interval of 24 h.

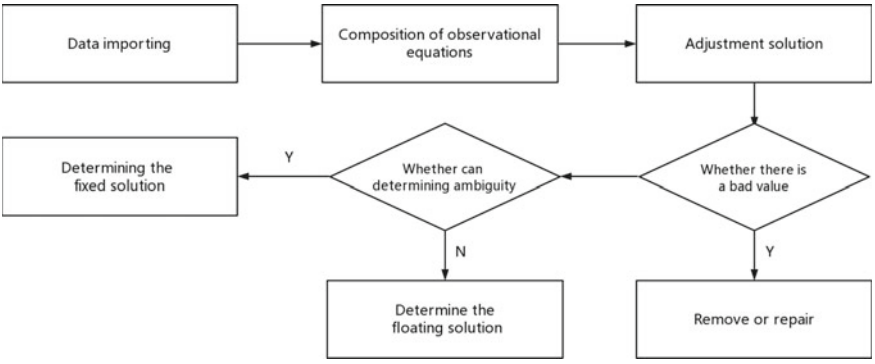



**Fig. 2.** Process of baseline solution

### 4.1 Basline Solution

The principle of the baseline solution can be summarized in the following steps:

(1) Adjustment initialization: Uncertain parameters can be solved based on the double-differenced observables, and the error equation is as follows.

Undetermined parameters:

$$\widehat{X} = \begin{bmatrix} \widehat{X}\ C \\ \widehat{X}\ N \end{bmatrix} \tag{1}$$

Cofactor matrix of the undetermined parameters:

$$Q = \begin{bmatrix} Q_{\widehat{X_C}\widehat{X_C}} & Q_{\widehat{X_C}\widehat{X_N}} \\ Q_{\widehat{X_N}\widehat{X_C}} & Q_{\widehat{X_N}\widehat{X_N}} \end{bmatrix} \tag{2}$$

After initialization, the variable of integer ambiguity is obtained.

(2) Determination of ambiguity: There are many ways to confirm the ambiguity. Now, a more reliable and reliable way is based on the search method which takes each ambiguity as the origin and then uses the error as a radius to confirm the integer solution of all the ambiguities.

(3) Determination of baseline fix solution: When confirming the integer solution of ambiguity, baseline integer solution [9] that the integer ambiguity located can be obtained. The integer solution can be used as a reference, which is helpful to the subsequent baseline solution, and to evaluate the quality of baseline.

### 4.2 Kalman Filtering

The essence of GLOBK is a Kalman filter, which makes the best estimate of the system state by analyzing the input and output observation data [10].

The Kalman filter estimates the process through a form of feedback control: The filter estimates the process state at a certain time and then obtains the feedback as a measurement. Therefore, the Kalman filter equations are divided into two groups: time-updated and measured equations, which can be expressed as follows:

$$\widehat{x_k}^- = A\widehat{x_{k-1}} + B\widehat{u_{k-1}} \tag{3}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{4}$$

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{5}$$

$$\widehat{x_k} = \widehat{x_k}^- + K(z_k - H\widehat{x_k}^-) \tag{6}$$

$$P_k = (I - K_k H)P_k^- \tag{7}$$

The schematic diagram of Kalman filter is detailed in Fig. 3.

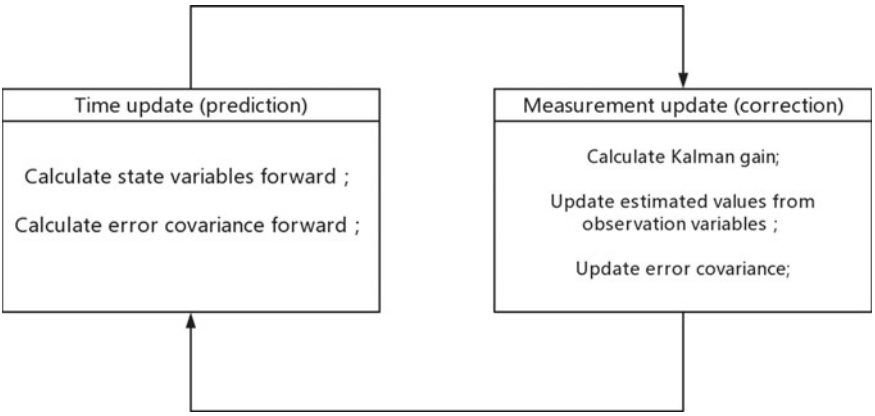



**Fig. 3.** Process of Kalman filtering

## 5 Results and Analysis

To simulate the actual settlement, observation antenna was placed on a precision mobile platform with an precision of 0.1 mm. The platform was moved in the vertical direction of 1.0 cm on March 24.

### 5.1 Results

Table 1 and Fig. 4 show the daily average observations from 23 to 28, which reflect the average change trend of the elevation. Compared with the actual adjustment, the absolute measurement accuracy has reached mm level.

**Table 1.** Daily average settlement data

| Station | Date | Altitude (m) | Settlement (mm) |
|---|---|---|---|
| NJ-01 | March 23 | 154.0375 | 0 |
| NJ-01 | March 24 | 154.0327 | 4.8 |
| NJ-01 | March 25 | 154.0275 | 10 |
| NJ-01 | March 26 | 154.0269 | 10.6 |
| NJ-01 | March 27 | 154.0273 | 10.2 |
| NJ-01 | March 28 | 154.0273 | 10.2 |

### 5.2 Error Analysis

The data of March 23 is taken as a benchmark to analyze the accuracy of the system.

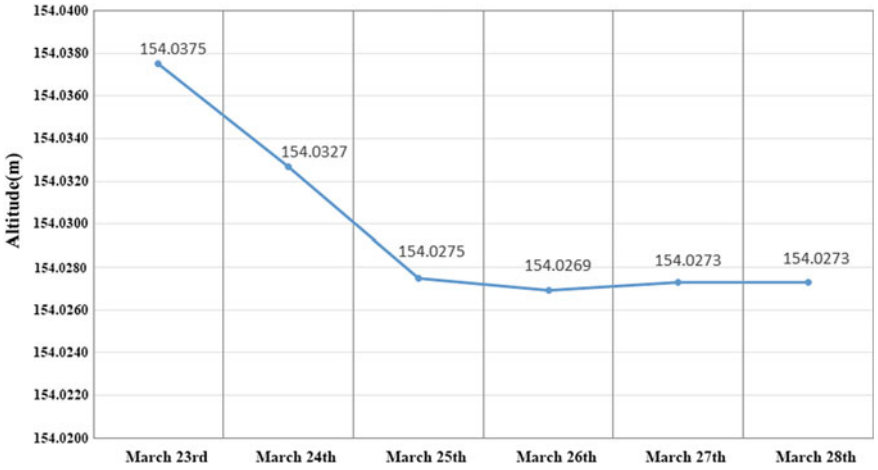



**Fig. 4.** Settlement curve

In order to quantitatively reflect the accuracy, it is necessary to introduce absolute difference and standard deviation which can be expressed as follows:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (S_i - S_{avg})^2}{N}} \tag{8}$$

According to the data in Table 1, substituting them into formula (8) respectively, the average absolute error and standard deviation can be calculated as 1.2 and 2.2 mm, having reached the mm level.

## 6 Conclusions

The chapter first introduced the current situation, background, and hazards of land subsidence in China. In order to take necessary measures in time for the hidden safety hazards of land subsidence, high-precision areas need to be monitored in real time. However, traditional monitoring methods have been difficult to meet the monitoring needs. Therefore, automated subsidence monitoring systems using modern communications, computer science, and satellite navigation technologies have emerged. With the completion of the BeiDou navigation system, more and more GNSS receivers are now beginning to be compatible with BeiDou Navigation Satellite System (BDS) information. Compared with other navigation systems, BeiDou provides unique short message communication means on the basis of ensuring safety, reliability, and high performance and can be used as an emergency communication method when a disaster occurs. Therefore, this experiment uses BDS as the main body and uses GPS, GLONASS, etc., as auxiliary references to establish a multi-mode fusion ground subsidence monitoring system to ensure the positioning accuracy in the case of real time and stability.

**Acknowledgments.** This work was supported by State Key Laboratory of Smart Grid Protection and Control of NARI Group Corporation.

## References

1. Wang A, Sun Z. Multi-geodesy techniques data fusing and analyzing for land subsidence monitoring. In: International workshop on earth observation and remote sensing applications, 2014. p. 345–8.
2. Angrisano A, Gaglione S, Gioia C. Performance assessment of GPS/GLONASS single point positioning in an urban environment. Acta Geod Geophys. 2013;48(2):149–61.
3. Shi J, Yuan X, Cai Y, Wang G. GPS real-time precise point positioning for aerial triangulation. GPS Solutions, 2017. p. 1–10.
4. Qu L, Zhao Q, Guo J, Wang G, Guo X, Zhang Q, Jiang K, Luo L. BDS/GNSS real-time kinematic precise point positioning with un-differenced ambiguity resolution. Lect Notes Electr Eng. 2015;342:13–29.
5. Wang L, Li Z, Yuan H, Zhao J, Zhou K, Yuan C. Influence of the time-delay of correction for BDS and GPS combined real-time differential positioning. Electron Lett. 2016;52(12):1063–5.
6. Yang C, Wu D, Lu Y, Yu Y. Research on network RTK positioning algorithm aided by quantum ranging. Sci China Inf Sci. 2010;53:248–57.
7. Li X, Ge M, Dai X, Ren X, Fritsche M, Wickert J, Schuh H. Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo. J Geodesy. 2015;89(6):607–35.
8. Li X, Xu L, Fang Y, Zhang Y, Ding J, Liu H, Deng X. Estimation of the precipitable water vapor from ground-based GPS with GAMIT/GLOBK. IEEE. 2010;1:210–4.
9. Schwarz KP, Lachapelle G. Kinematic systems in geodesy, surveying, and remote sensing. New York: Springer; 1991.
10. King RW. Documentation for the GAMIT GPS analysis software. Massachusetts Institute of Technology, 1995.

# Multi-layer Location Verification System in MANETs

Jingyi Dong(✉)

School of Electrical Engineering and Telecommunications,
The University of New South Wales, Sydney, Australia
jennied_jiyue@163.com

**Abstract.** The mobility and feasibility of mobile ad hoc networks (MANETs) have to deeply rely on the accurate location information to support multiple applications. A wrong announced location of a node may cause some serious consequences. Thus, the localization and the location security should be considered as important parts of the whole design of the MANETs. In the traditional way, the location verification schemes need complex calculation and multi-step communication with base stations and other vehicles, which are strengthen the burden of the MANETs routings and increased the complexity of protocol. In this paper, an improved multi-layer location verification system (MLVS) based on the optimal common neighbor's knowledge between claimer and verifier in MANETs is been proposed and discussed. In this system, each node in MANETs could have a trust value, and a mutually shared token scheme is provided to make the decision of the MLVS. Furthermore, the MLVS shows a reliable performance on the ability of attacker defense and accuracy in high node density networks.

**Keywords:** Mobile ad hoc networks · Location verification system · MLVS · Location security

## 1 Introduction

As several innovative technologies of multi-hop ad networks, the mobile ad hoc networks (MANETs), it has been one of the most popular and interesting topics for researchers. The major characteristic of MANETs is that many mobile terminals and low-cost sensors are consisted, and the packets are delivered between wireless interfaces following the geographic routing. To achieve a better performance and support possible applications, some localization schemes should be deployed into sensors and terminals, such as Global Positioning Systems (GPS). However, the openness and public ability of the wireless communication makes that the localization of MANETs could be easily attacked by malicious nodes [1]. Hence, the accuracy of the location information plays a significant role to protect the sensors' privacy, and the way to authenticate location information has been concerned [2–6]. Figure 1 illustrates the basic concept of location verification.


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_169

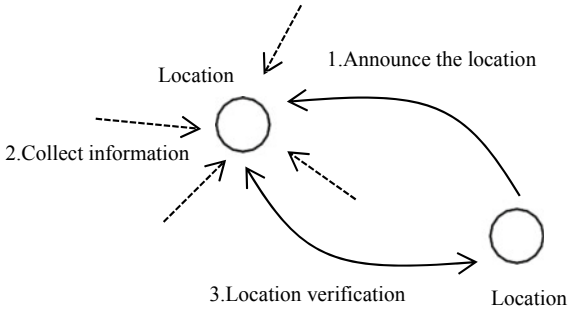



**Fig. 1.** Basic concept of location verification

A basic binary rule is proposed for location verification system between claimer and verifier in [7]. However, some measurements are difficult to acquire without line-of-sight (LOS). Therefore, the cooperative nodes' location communication scheme is proposed with the protocol and routing techniques' help [8, 9]. Compared to the other routing techniques, geographic routing has got foremost attention for information transmission in vehicular communication [10–12], and the information of neighbors can be obtained [13].

In this paper, a two-layer location verification system with neighbors' knowledge in mobile ad hoc networks is proposed; it follows the basic principle of mutually shared region-based location verification (MSRLV) in WSNs [14]. In layer one, LVS would roughly eliminate some malicious nodes and create a trust value table for nodes. Then, in layer two, the trust weight of common neighbors would be signed as the preparation to create a mutually shared model; the selected common neighbors in shared region between verifier and claimant provide the common knowledge to both verifier and claimant. With the information collected from neighbors, the system shows a good performance with higher node density. However, the location error would be increased in sparse MANETs.

## 2 System Model

In this section, some assumptions are considered for the network as follows: The mobile terminals and sensors have the same wireless transmission range $r$; assume that the network contains $N$ nodes. Some of them are malicious nodes $n_m$, which may provide a fake location to others. The nodes' location is assumed to be obtained by GPS which is considered as the true location neglecting the localization error; we assume that the node (source) who sends the data first is not a malicious node, and following the protocol, the hops between source and destination must do the location verification before the data transmission.

The transmitted data of nodes is $n_i\{\text{data}_i, L_i, \text{RV}_i\}$, where $\text{data}_i$ is the data transmitted by the node $n_i$, $L_i$ is the announced location, and $\text{RV}_i$ is a random value which would be discussed in next section.

### 2.1 Layer 1: Trust Value Table (TV)

The trust value table (TV) is constructed by several factors which are RSSI, moving direction and relative speed.

The RSSI is capable of roughly eliminating some malicious nodes who provide fake locations. The node $n$ could obtain the received signal power $P^r_{n_j}$ from its one-hop neighbors $n_j$ to measure the distance $d_{n_j}$ using RSSI. At the same time, the $P^r_{n_j^*}$ could be calculated with announced location from $n_j$ with $d_{n_j^*}$. In ideal state, the malicious node could pass the estimation only if $d_{n_j} = d_{n_j^*}$. However, considering the mobility of the MANETs, some error may be caused by complex wireless environment. In this term, a filter could be proposed to eliminate some malicious points with the threshold $\alpha$; the neighbors who pass the distance filter can consist into a set $S_{\text{neighbor}}$ with their different level.

$$\beta_i = 1 - \left| P^r_{n_j} - P^r_{n_j^*} \right| / P^r_{n_j^*} \tag{1}$$

where $\beta_i$ is the $i$-th neighbor's distance different level of $n_i$. The $\beta_i$ would be stored in $S_{\text{neighbor}} \left\{ n, \beta_n; \left| P^r_{n_j} - P^r_{n_j^*} \right| < \alpha \right\}$. However, the RSSI could only eliminate some malicious nodes which fake their location with a large distance difference; in some cases, the system may be suffered from the similar distance-based malicious node (SDM attack). Hence, a mutually shared region token scheme (MSR) is considered to deploy as layer two of the LVS [14]. To achieve a better performance, a trust value is regarded as the standard about common neighbors' selection in layer two. The probability that the neighbors passed the distance filtering, the relative speed and the relative distance between verifier and claimant are the factors contributing to the trust value. The model of the network is shown in Fig. 2, the node $n$ could be seen as the verifier $n_v$, and the next-hop node could be seen as the claimant $n_c$.

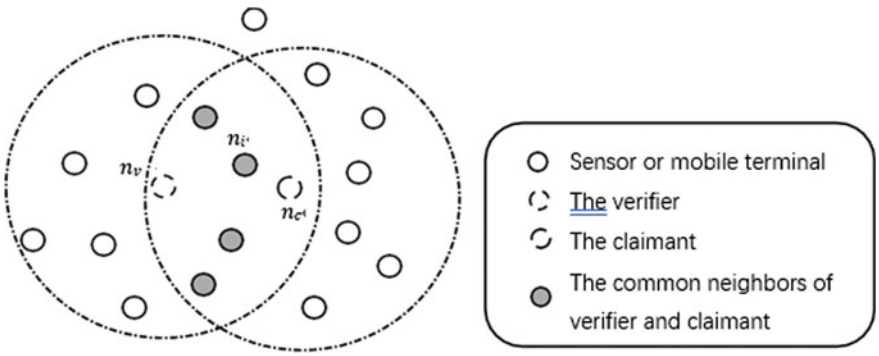



**Fig. 2.** Sensors and mobile terminals in network

We assume the speed of $n_v$,$n_c$, and $n_i$ are $s_v$, $s_v$ and $s_i$, respectively, which the unit is m/s. The speed difference level $\gamma$ is calculated by the following equation:

$$\gamma_i = 1 - \frac{|s_v - s_i| + |s_v - s_i|}{2s_i} \tag{2}$$

The distance difference between the verifier and claimant level μ is measured with:

$$\mu_i = \frac{r}{d_{v-i} + d_{v-c}} \tag{3}$$

where $d_{v-i}$ is the distance between the verifier and the $i$-th neighbor, and $d_{v-c}$ is the distance between the claimant and $i$-th neighbor; the location information is acquired from the GPS. $r$ is the communication range. Hence, the trust value $\mathrm{TV}_{\mathrm{n_i}}$ of the $i$-th neighbor is represented by Eq. (4):

$$\mathrm{TV}_{\mathrm{n_i}} = \gamma_i + \mu_i + \beta_i \tag{4}$$

### 2.2 Layer 2: Mutually Shared Token Scheme

As Fig. 2 shows, when the nodes finished the distance filtering, it could collect $\{\mathrm{RV}_i\}$ into a packet $\mathrm{data}_{\mathrm{mst}}^{n_c}$ from the common neighbors, the mutually shared token follows the bellowing calculation:

$$\mathrm{MST}_{n_c} = \{n_i \in \mathrm{SCN}: \quad \mathrm{TV}_{n_i} > \varphi, \quad \mathrm{data}_{\mathrm{mst}}^{n_c} = \sum \mathrm{RV}_i\} \tag{5}$$

where $\varphi$ is a threshold of trust value depending on the wireless environment, and SCN is the set of common neighbors. And then $\mathrm{data}_{\mathrm{mst}}^{n_c}$ would be sent to $n_v$. As the packet received by the verifier $n_v$, the $n_v$ would do mutually shared token in the same way:

$$\mathrm{MST}_{n_v} = \{n_j \in \mathrm{SCN}^*: \quad \mathrm{TV}_{n_j} > \varphi, \quad \mathrm{data}_{\mathrm{mst}}^{n_c} = \sum \mathrm{RV}_j\} \tag{6}$$

and if the $\mathrm{MST}_{n_c}$ matched with $\mathrm{MST}_{n_v}$ $(\mathrm{MST}_{n_c} = \mathrm{MST}_{n_v})$, the location information of $n_c$ could be accepted and regarded the $n_c$ as the next-hop node.

## 3 Performance Analysis

We assume that the malicious nodes $n_{m_k}$ fake its location and broadcast a wrong location to the transmission environment $n_{m_k}^* \left\{ L_{m_k}^* \right\}$, $L_{m_k}^* \neq L_{m_k}^{\mathrm{true}}$. With the suitable selection of common neighbors, the verifier and the claimant can receive the same MST. However, if a malicious node pretends its position at a wrong place, such as the situation shows in Fig. 3, even it is an SDM attack, it is hard to collect all necessary packets from signed nodes. Hence, the multi-layer location verification system can defense several attacks discussed above.

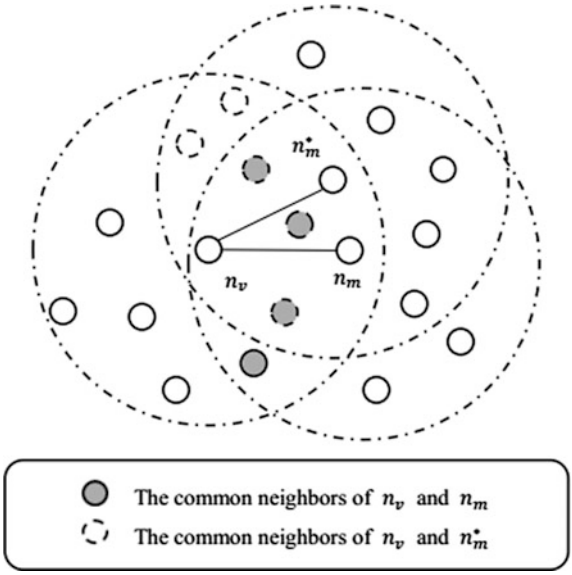



**Fig. 3.** Mutually shared region model with malicious nodes

The performance is evaluated in terms of the ability of the malicious nodes rejection and the location error in different situation. The simulations are built on the MATLAB. The parameter of the simulation is shown in Table 1 which is similar to the [15].

**Table 1.** Parameters of the simulation

| Parameters | Values | Parameters | Values |
|---|---|---|---|
| Simulation area | 1000 × 1000 $m^2$ | Propagation | Shadowing |
| Nodes' speed | 0–10 m/s | MAC data rate | 10 Mps |
| Simulation time | 20 | MAC protocol | IEEE 802.11 |
| Number of nodes | 0–500 nodes | Frequency | 5.9 GHz |
| Sources and destinations | 20 nodes | Packet type | UDP |
| Transmission range | 150 m | Packet size | 512 bytes |

The ability of the malicious nodes' defense means that the probability of location error has been discovered under the MLVS. It is assumed that the total number of the malicious nodes is $N_m$, and the times that packets sent to a malicious node in *i-th* simulation can be regarded as $T_m^i$. Then the location error $e_L$ happened in the MANETs with multi-layer location verification is following the equation:

$$e_L(\%) = \left\{ \frac{\sum_{i=1}^{20} \frac{T_m^i}{N_m}}{20} \right\} \times 100 \tag{7}$$

In addition, we assume that a single-fake location error in this simulation. Single-fake location error is considered that only one malicious node fakes its location around honest nodes. And Fig. 4a illustrates the performance of MLVS in MANETs with different maximum nodes which contains 10%, 20%, and 30% malicious nodes respectively. The MLVS is a location verification scheme which verifies nodes relying on the neighbors' information; thus, the density of the MANETs deeply impacts the performance of the system. As Fig. 4a shows, the system has a lower location error as the maximum number of nodes is more than 400.

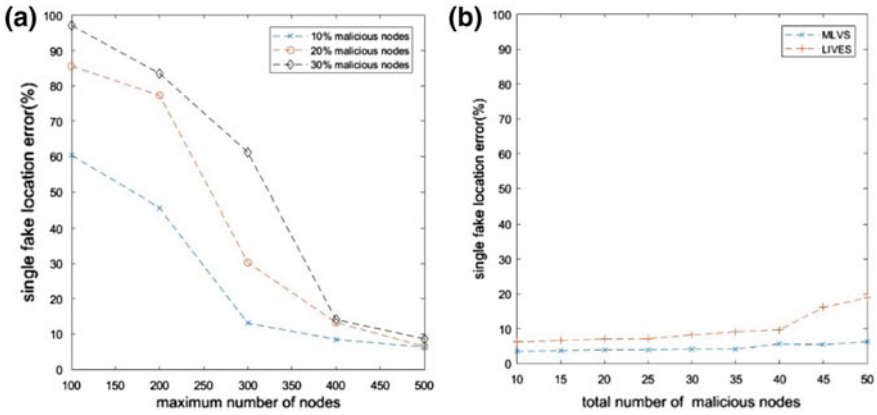



**Fig. 4.** **a** Performance of MLVS with difference number of malicious nodes and **b** the performance of MLVS compares with LIVES

In Fig. 4b, it shows the performance of MLVS with different number of malicious node and the maximum number of nodes is 500 compared with LIVES illustrated in [15]. In LIVES, the number of malicious nodes impacts the probability of fake position rejection. However, the MLVS has a stable performance with higher node density.

## 4 Conclusion

In this paper, a multi-layer location verification technique is discussed based on geographic routing in MANETs. The first layer performs a distance filtering for verifiers and claimants and then calculates the trust value for their neighbors as the preparation of common neighbor selection. The second layer evaluates the information collected from common based on the principle of mutually shared token. The trust value calculated in layer one considers the mobility of the network to ensure the verifier, and the claimant can collect information from the same group of neighbors. The MLVS shows a good performance with high-density network; however, it cannot be deployed in

sparse MANETs or higher mobility ad hoc networks, such as VANETs. In future, an improved location verification system for VANETs will be considered with the concept of cooperative position.

## References


1. Tippenhauer NO, Rasmussen KB, Pöpper C, Capkun S. iPhone and iPod location spoofing: attacks on public WLAN-based positioning systems. SysSec technical Report. Swiss Federal Institute of Technology; 2012 Apr. 2008.
2. Malaney RA. A location enabled wireless security system. In Global telecommunications conference, 2004. GLOBECOM'04. IEEE; 2004. pp. 2196–200.
3. Faria DB, Cheriton DR. Detecting identity-based attacks in wireless networks using signalprints. In Proceedings of the 5th ACM workshop on wireless security; 2006. pp. 43–52.
4. Papadimitratos P, Gligor V, Hubaux JP. Securing vehicular communications-assumptions, requirements and principles. Proc ESCAR. 2006:5–14.
5. Papadimitratos P, Buttyan L, Holczer T, Schoch E, Freudiger J, Raya M, Hubaux JP. Secure vehicular communication systems: design and architecture. IEEE Commun Mag. 2009;46 (11):100–9.
6. Bauer K, McCoy D, Anderson E, Breitenbach M, Grudic G, Grunwald D, Sicker D. The directional attack on wireless localization: how to spoof your location with a tin can. In: Proceedings of the 28th IEEE conference on global telecommunications; 2009. pp. 4125–30.
7. Yan S, Malaney R. Location verification systems in emerging wireless networks. arXiv preprint; 2013 arXiv:1307.3348.
8. Abumansoor O, Boukerche A. A secure cooperative approach for nonline-of-sight location verification in VANET. IEEE Trans Veh Technol. 2012;61(1):275–85.
9. Vora A, Nesterenko M. Secure location verification using radio broadcast. IEEE Trans Dependable Secure Comput. 2006;3(4):2006.
10. Kaiwartya O, Kumar S. Cache agent-based geocasting in VANETs. Int J Inf Commun Technol. 2015;7(6):562–84.
11. Suthaputchakun C, Sun Z. Routing protocol in intervehicle communication systems: a survey. IEEE Commun Mag. 2011;49(12):150–156.
12. Ansari K, Feng Y, Singh J. Study of a geo-multicast framework for efficient message dissemination at unmanned level crossings. IET Intell Transp Syst. 2013;8(4):425–34.
13. Cao Y, Sun Z, Wang N, Riaz M, Cruickshank H, Liu X. Geographic-based spray-and-relay (GSaR): an efficient routing scheme for DTNs. IEEE Trans Veh Technol. 2015;64(4): 1548–64.
14. Kim IH, Kim BS, Song J. An efficient location verification scheme for static wireless sensor networks. Sensors. 2017;17(2):225.
15. Kargl F, Klenk A, Schlott S. Weber M. Advanced detection of selfish or malicious nodes in ad hoc networks. In: European workshop on security in Ad-hoc and sensor networks; 2014. pp. 152–165.

# Design of Multi-antenna BeiDou High-Precision Positioning System

Kunzhao Xie[1(✉)], Zhicong Chen[1], Rongwu Tang[1], Xisheng An[1], and Xiaorong Li[2]

[1] Laibin Power Supply Bureau of Guangxi, Power Grid Co., Ltd., Laibin, China
275856140@qq.com,
{chen_zc.lbg,tang_rw.lbg,an_xs.lbg}@gx.csg.cn
[2] Traffic Business Department of CIECC, Beijing, China

**Abstract.** With the progress of society and economic development and the expanding deployment scale of substations, the automatic monitoring of the foundation settlement in construction station becomes an important issue of the operation and maintenance in the power grid. BeiDou satellite navigation system is a global navigation and positioning system, which is independently developed to provide navigation, positioning, and timing services in China and its surrounding areas. This article describes the principles of BeiDou high-precision positioning. According to the land subsidence monitoring and warning requirements for substations, multi-antenna technology is introduced in order to solve the problem of the cost of large-scale land subsidence monitoring system. The monitoring interface software is designed to visually reflect the settlement situation. The monitoring system can effectively improve the monitoring efficiency and early warning capability of the ground subsidence in the power company's substation. The system will also provide a basis for decision-making management and improve the reliability, safety, and stability of power grid operation based on high-precision field accuracy and trend.

**Keywords:** Land subsidence · Multi-antenna · BeiDou satellite positioning

## 1 Introduction

With the development of social economy, land resources are decreasing more and more land and low-lying areas are being exploited and utilized. Grid transmission and transformation should be synchronized with the development of social economy, when choosing the location of the substation, and some of them have to be established in some areas that are not suitable for them. Meanwhile, substation foundation sedimentation problem is increasingly prominent. Hidden dangers that may cause substation equipment accidents happen sometimes, which


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_170

bring great influence to the safety of substations [1,2]. Substation foundation sedimentation problems are divided into two parts: uniform sedimentation and non-uniform sedimentation. On the occasion of uniform sedimentation, there is little internal stress on electrical equipment on the ground, doing no harm to the substation. However, on the other occasion, the internal stress is much larger. It is easy to cause the electric equipment to break, tilt, or even collapse, resulting in accidents. In order to avoid above-mentioned problems and keep substations work under safe conditions, monitoring the non-uniform sedimentation of the substation foundation is of great significance. Nowadays, the automatic level of land subsidence and deformation monitoring systems is limited [3,4].

According to the ground sedimentation monitoring and warning requirements for substations, this paper studies and develops a BeiDou high-precision network deformation monitoring system, focusing on the development of multi-antenna technology to achieve time-sharing reception of multi-channel antenna signals by satellite receiver boards. That is, a receiver board can receive antenna signals at multiple points, which greatly reduce the cost of large-scale deployment of satellite positioning monitoring networks.

The system provides the power company with an advanced, practical, and reliable online monitoring and analysis system for ground sedimentation in substation areas. It is able to remotely monitor and analyze the geological deformation of the site in the substation area and provide timely warning of various dangerous situations that may occur. Based on BeiDou's high-precision field accuracy and trend, it will also provide a basis for decision management and improve the reliability, safety, and stability of power grid operation [5].

## 2 Problems

With the improvement of satellite receiver hardware performance and software processing technology, satellite positioning technology has been applied and popularized in many fields such as slope deformation monitoring, geodetic survey, crustal deformation monitoring, and precision engineering measurement. The application of satellites to slope deformation monitoring is not only highly accurate, unaffected by climatic conditions, but also enables unattended automated real-time monitoring modes.

BeiDou satellite navigation system is a global satellite navigation system independently developed by China. At present, the system is actively promoting and building in accordance with the principles of openness, autonomy, compatibility, and gradual progress. Now, it has the ability to cover the positioning, navigation, and communication services in the Asia-Pacific region. Around 2020, the system will include 5 GEO satellites, 3 IGSO satellites, and 27 MEO satellites, eventually achieving global coverage [6]. In Asia-Pacific areas, it is even better than GPS, which is verified by researchers from different countries. The basic method of ground sedimentation monitoring is to obtain ground elevation changes through high-precision positioning technology and to analyze vertical displacement components based on complex algorithms to achieve monitoring

of ground subsidence. The BeiDou second-generation satellites' high-precision positioning technology plays an active and important role in it. The application of satellite technology to a wide range of geological monitoring is mainly limited by the following factors:

- The precision measurement satellite receiving equipment are expensive. In a conventional way, one satellite antenna is equipped with one receiver. If it is utilized in a large-scale deformation monitoring, the cost could be high.
- Generally, the landslide that needs to be monitored has a large monitoring range and many monitoring points. The amount of original satellite monitoring data collected at the monitoring site is large, and how to use effective means to ensure the massive and reliable transmission of massive data to the control center is worth focusing.

## 3 Receiver with Multi-antenna Design

One receiver is connected to multiple antennas. At each monitoring point, satellite antennas are installed with no receivers. Multiple monitoring points share one satellite receiver. Compared to the conventional idea that one satellite antenna uses one satellite receiver, the cost of such a monitoring system can be significantly reduced by the reduction in the number of satellite receivers [7,8]. The satellite multi-antenna controller is shown in Fig. 1.

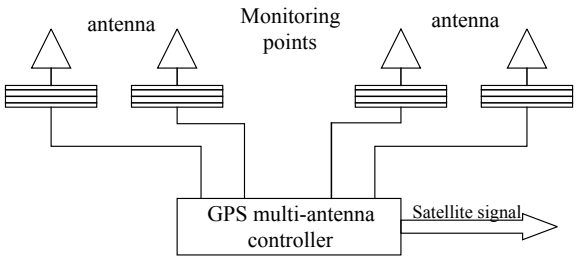



**Fig. 1.** Multi-antenna technology

The satellite multi-antenna controller is a key device in the landslide monitoring system. It is an organic integrated body of multiple satellite antennas and microwave switches with multiple channels. Each satellite antenna is connected to a corresponding channel of the microwave switch, and the on/off state of the plurality of signal channels of the microwave switch is controlled by the switch control circuit in real time through computer programming. The rotation of the monitoring point data acquisition is realized by controlling the on and off states of the channel.

Figure 2 provides a multichotomous antenna signal controller, including a power supply module, a control module, and a switch module. The power supply

module provides the voltage for the system. The control system consists of a microprocessor unit and a relay drive unit. The switch system includes a relay group, receiving antenna bases, and a sending antenna base. Through the selector in the module, signals are sent to the sending antenna base and transmitted to the receiver board.

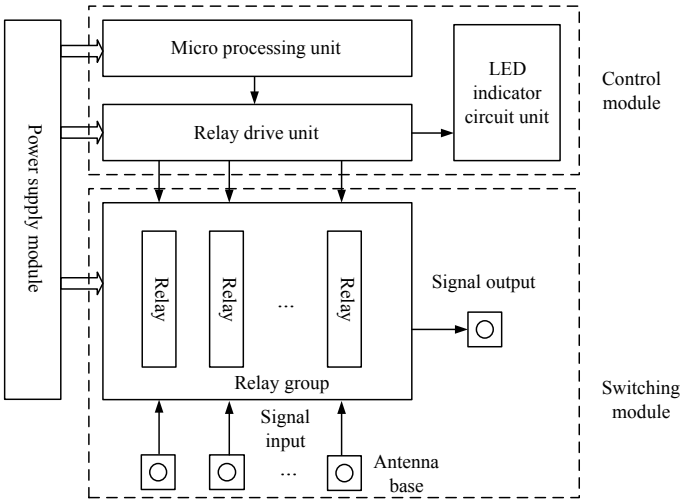



**Fig. 2.** Multi-antenna controller

## 4 Multi-antenna BeiDou High-Precision Positioning System

According to the requirements of data acquisition and power system security considerations, the ground sedimentation monitoring system consists of monitoring terminals, a reference terminal, a substation local server, a cloud server, and a client, which is illustrated in Fig. 3. To monitor one substation, there are always more than one monitoring terminals that are distributed around the substation. The monitoring terminal is deployed at where is geological hard and about 3 kms away from the substation.

### 4.1 System Structure

The receiving terminal is mainly responsible for receiving the radio-frequency signal transmitted by the satellite through the antenna and performing preliminary processing of the satellite signal to obtain the original data and transmitting it to the local server terminal.

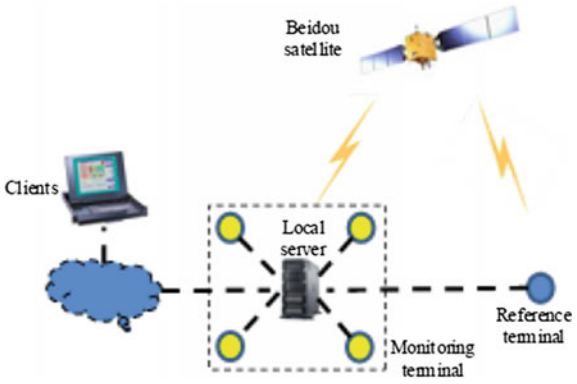



**Fig. 3.** Overall design of land subsidence monitoring system

### 4.2 Reference Terminal

The reference station terminal consists of a power supply module, a GNSS board, a control module, and a communication module. The power supply module mainly contains DC-DC voltage conversion circuits. The GNSS board is responsible for receiving satellite radio-frequency signals through the satellite antenna and transmitting the coordinate data to the control module. The control module is made up of a microprocessor unit, a storage unit, an LED indicator unit, and an associated external circuit. It is mainly responsible for transmitting the satellite positioning coordinate data of the desired format to the receiver board module. Meanwhile, it transmits the coordinate data of the desired format to the communication module. The coordinate data is stored locally. The communication module includes a wired network unit and a 4G mobile network unit.

### 4.3 Monitoring Terminal with Multi-antenna

This part includes power supply modules, a module that has one receiver with multiple antennas, a GNSS board, a control module, and a communication module. Different from the reference station terminal, the monitoring station terminal has a module that has one receiver with multiple antennas. Other parts are consistent with the reference station terminal. The multichotomous switch module is composed of a relay driving unit, an RF relay group, and a receiving antenna base. The microprocessor unit outputs a control signal at a preset time interval, and the control signal can cause the corresponding pin of the relay driving unit to output a current which is high enough to drive the corresponding relay to realize the function of switching the control circuit. This module enables the GNSS board to receive signals from multiple antennas in a time-division manner.

### 4.4 Substation Local Server

The substation local server deploys the data solving subsystem and receives the terminal data. The main functions implemented include: (1) Receive raw data sent by the satellite positioning information receiving subsystem, and store it locally in real time. (2) Set the solution interval, sampling interval, filter settings according to the instructions sent by the center, and start solving. (3) Keep a long connection with the central station, send the solution result (site coordinates) to the center through the cloud center, and receive instructions from the center.

### 4.5 Client

In order to better realize the management and maintenance of the settlement monitoring system, the central station software is designed. The main functions realized are: (1) Download the solution data forwarded from the cloud center, and store it to a local database in real time. (2) Display the solution data by graphics. (3) Control the receiving terminal subsystem partially.

## 5 Experiments and Results

### 5.1 Experiments

The ground subsidence monitoring system designed in this paper was tested at a substation in Laibin, Guangxi province. The offset of XYZ was adjusted by the coordinate adjustment table, compared with the system data measurement. Figure 4 is the measurement data in the X-direction; Fig. 5 is measurement data in the Y-direction; Fig. 6 is measurement data in the Z (elevation)-direction.

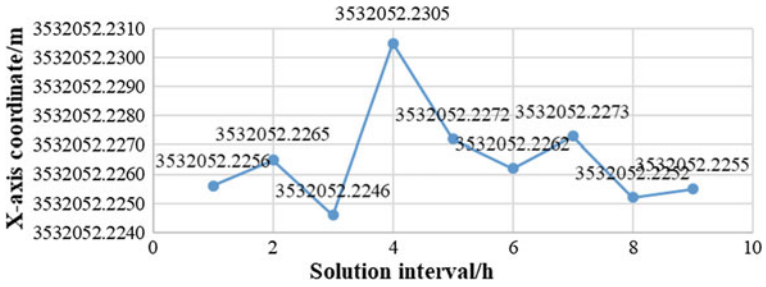



**Fig. 4.** Measurement data of X-direction

Figure 6 shows the monitoring data for the experiment of simulation settling. For measurement data of the elevation test, we reduced the height of the monitoring station by 16 mm during the sixth data solution calculation. We find that the offset between 1–5 sets of data and 7–9 sets of data is about 15 mm.

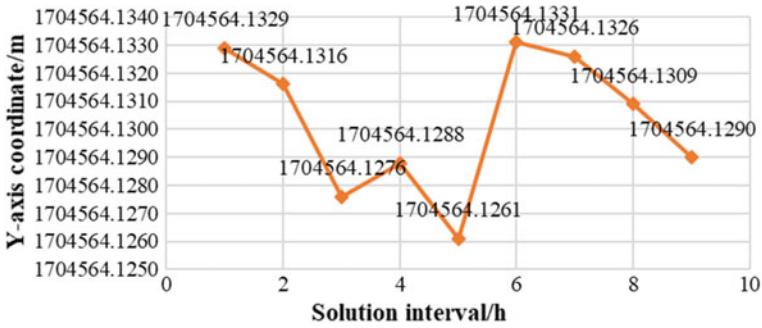



**Fig. 5.** Measurement data of Y-direction

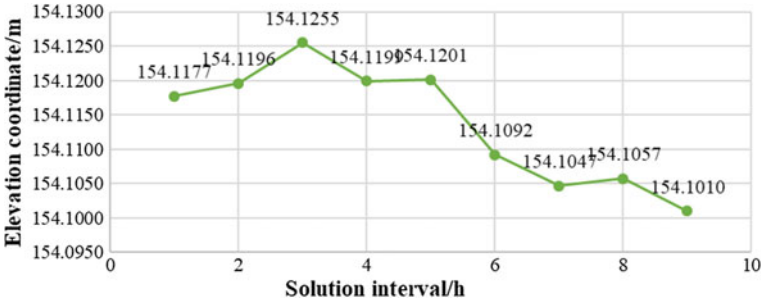



**Fig. 6.** Measurement data of Z-direction

## 5.2 Results

We use the range and standard deviation as the evaluation criteria to analyze the above test data. Table 1 shows the test data. The standard deviation is within 3 mm, and the error meets the design accuracy requirements.

**Table 1.** The data analysis

| | Range(R) (mm) | Standard deviation($\sigma$) (mm) |
|---|---|---|
| X | 6 | 1.7 |
| Y | 7 | 2.5 |
| Z(1–5) | 7.8 | 2.9 |
| Z(7–9) | 4.6 | 2.5 |

The range is the difference between the maximum and minimum values of a set of measurements, which is also known as range error or full distance, expressed as R:

$$Range : R = X_{max} - X_{min}, \tag{1}$$

$$Standarddeviation : \sigma = \sqrt{\frac{1\sum_{i=1}^{N}(x_i - \mu)^2}{N}}. \tag{2}$$

## 6 Conclusion

According to the ground subsidence monitoring and early warning requirements of the substation, this paper proposed a BeiDou high-precision network deformation monitoring system, especially developed a multi-antenna technology and module. The system reduced the system cost and effectively realized the automatic collection, monitoring, early warning, and management of ground settlement monitoring data, which greatly improved the automation level of grid operation and maintenance. The results of experiment show that the system function is stable, and the positioning accuracy meets the design requirements.

## References

1. Gang Liu. On the safety management of substation operation and accident prevention. Value Eng. 2011;1656(1):113–23.
2. Mi Chen, Roberto Toms, Zhenhong Li, Mahdi Motagh, Tao Li, Leyin Hu, Huili Gong, Xiaojuan Li, Jun Yun, Xulong Gong. Imaging land subsidence induced by groundwater extraction in beijing (china) using satellite radar interferometry. Remote Sens. 2016;8(6).
3. Nasipuri A, Cox R, Conrad J, Van Der Zel L. Design considerations for a large-scale wireless sensor network for substation monitoring. In: Local computer networks; 2010. p. 866–73.
4. Guoqing Yao, Jingqin Mu. D-insar technique for land subsidence monitoring. Earth Sci Front. 2008;15(4):239–43.
5. Yue Yang, Qi Kang. Operation situation analysis of substation equipment online monitoring system in inner mongolia power grid. Inner Mongolia Electric Power. 2014.
6. Xuefeng Lu, Yongfeng Liao, Bo Li, Lan Deng. Beidou integrated disaster reduction application platform. China Telecom. 2015;12(8):169–82.
7. Juang JC, Tsai CT, Chen YH. Development of a multi-antenna gps/beidou receiver for troposphere/ionosphere monitoring. In: Proceedings of international technical meeting of the satellite division of the institute of navigation; 2012. p. 947–952.
8. Yongqi Chen, Xiaoli Ding, Dingfa Huang, Jianjun Zhu. A multi-antenna gps system for local area deformation monitoring. Earth Planets Space. 2000;52(10):873–6.

# Research on Sound Source Localization Algorithm of Spatial Distributed Microphone Array Based on PHAT Model

Yong Liu, Jia qi Zhen(✉), Yan chao Li, and Zhi qiang Hu

School of Electronics Engineering, Heilongjiang University, Harbin 150080, China
zhenjiaqi2011@163.com

**Abstract.** With the development of artificial intelligence voice technology, array signal processing technology has been widely used in intelligent human–computer interaction. In this field, the sound source localization technology of the microphone array system is also one of the key technologies in the composition of intelligent systems, and it is also a research hotspot technology, which is of great significance for improving human–computer interaction ability. In this paper, the microphone array technology is used to demonstrate and improve the sound source localization of indoor speech signals. A spatial distributed microphone array localization algorithm is proposed to improve the accuracy of the indoor voice source signal localization in the presence of indoor noise and reverberation.

**Keywords:** Spatial distributed microphone array · Sound source localization · Room impulse response · Phase transformation weighted generalized cross-correlation algorithm

## 1 Introduction

The microphone array technology uses a focused method to generate a beam and align it with the estimated position of the sound source, thereby reducing the interference signals of other unwanted signals, environmental noise, indoor reverberation, etc., to a certain extent, so that the collected elements are acquired [1, 2].

The time delay of arrival (TDOA) sound source localization method is the most commonly used in all positioning fields [3]. It first estimates the time delay of the sound source reaching each element of the receiver and obtains the distance difference. The mathematical method of space geometry is used to solve the position of the target sound source [4].

However, the sound source localization technology of steerable beamforming is difficult to apply in practice [5]. Even though some iterative methods can be used to reduce the computation time and enhance the real-time performance, this may result in


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_171

the inability to search for effective global peaks and make the positioning effect less than ideal [6]. In addition, the algorithm relies on the spectral characteristics of the sound source. In practical applications, a series of interference signals in the sound source environment make the spectral characteristics of the sound source difficult to obtain, which also limits the application range of the algorithm [7].

## 2 Sound Source Localization Algorithm Based on PHAT for Spatial Distributed Microphone Array

Establishing a spatial distributed microphone array model, the indoor space is a cube, as shown in Fig. 1. The origin of the coordinate is also the coordinate position of the microphone element $M_1(0,0,0)$, and the positions of the remaining seven elements are $M_2(L,0,0)$, $M_3(L,L,0)$, $M_4(0,L,0)$, $M_5(0,L,L)$, $M_6(0,0,L)$, $M_7(L,0,L)$, $M_8(L,L,L)$, and L represents the length of the room model. The coordinates of the speaker are $S(x,y,z)$. The azimuth is $\varphi(0^\circ \le \varphi \le 360^\circ)$. The pitch angle is $\theta(0^\circ \le \theta \le 90^\circ)$. The distance from the first microphone receiver element $M_1$ is $r$. $d_i(2 \le i \le 8)$ is the distance difference between the sound source to the position 2 of the first microphone element and the position of the sound source to the third element position 3. $d_i(2 \le i \le 8)$ is the distance difference between the position $M_1$ of the sound source to the first microphone element and the position $M_i$ of the source to the $i$th element.

According to the geometric relationship between each receiver element and the position of the sound source:

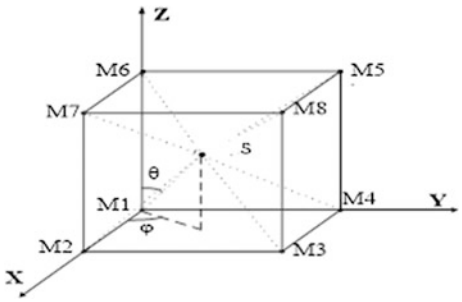



**Fig. 1.** Spatial distributed microphone array

$$\begin{cases} S(O)M_1 : x^2 + y^2 + z^2 = r^2 \\ SM_2 : (L-x)^2 + y^2 + z^2 = (r+d_2)^2 \\ SM_3 : (L-x)^2 + (L-y)^2 + z^2 = (r+d_3)^2 \\ SM_4 : x^2 + (L-y)^2 + z^2 = (r+d_4)^2 \\ SM_5 : x^2 + (L-y)^2 + (L-z)^2 = (r+d_5)^2 \\ SM_6 : x^2 + y^2 + (L-z)^2 = (r+d_6)^2 \\ SM_7 : (L-x)^2 + y^2 + (L-z)^2 = (r+d_7)^2 \\ SM_8 : (L-x)^2 + (L-y)^2 + (L-z)^2 = (r+d_8)^2 \end{cases} \tag{2.1}$$

From formula (2.1):

$$r = \frac{c \cdot \left[\sum_{i=2}^{8} \hat{\tau}_i^2 - 4(\hat{\tau}_2 + \hat{\tau}_4 + \hat{\tau}_6)\right]}{8(\hat{\tau}_2 + \hat{\tau}_4 + \hat{\tau}_6) - 2\sum_{i=2}^{8} \hat{\tau}_i} \tag{2.2}$$

From formula (2.1) can obtain the elevation angle $\theta$ and the azimuth angle $\varphi$:

$$\begin{cases} \tan\ \theta = \dfrac{L^2 - 2rd_4 - d_4^2}{L^2 - 2rd_2 - d_2^2} \\ \tan\ \varphi = \dfrac{\sqrt{(L^2 - 2rd_2 - d_2^2)^2 + (L^2 - 2rd_4 - d_4^2)^2}}{L^2 - 2rd_6 - d_6^2} \end{cases} \tag{2.3}$$

From formula (2.3), $d_i = c\,\hat{\tau}_i$, c is the speed of sound. $\hat{\tau}_i = (2 \le i \le 6)$ represents an estimated value of the signal delay received by the first microphone element $M_1$ and the $i$th microphone element $M_i$.

Using the above method, the azimuth of the sound source and the tangent of the pitch angle are obtained. At the same time, since the arctangent function belongs to the transcendental function, it is calculated by the coordinate rotation numerical calculation method Coordinate Rotation Digital Computer (CORDIC). This method is a simple displacement and addition and subtraction operation, which will improve the accuracy of the iterative calculation.

The azimuth angle $r$ obtained by formula (2) and the azimuth angle $\varphi$ obtained by formula (3), and the elevation angle $\theta$, the spherical coordinates can be obtained as $(r, \varphi, \theta)$, which can be expressed as (4):

$$\begin{cases} x = r\ \sin\ \varphi\ \cos\ \theta \\ y = r\ \sin\ \varphi\ \sin\ \theta \\ z = r\ \cos\ \varphi \end{cases} \tag{2.4}$$

Formula (2.4) can determine the position coordinates of the sound source.

## 3 Conclusion

It can be seen from above that in the case of indoor background noise and reverberation in the room, there is a certain deviation between the actual result and the simulation result. Compared with the traditional quaternary cross-indoor positioning effect, the indoor positioning effect of the spatial distributed microphone array has the characteristics of improved positioning accuracy.

**Acknowledgments.** This paper is supported by the Natural Science Foundation of China 61501176 and Natural Science Foundation of Heilongjiang province F2018025.

## References

1. Dmochowski JP, Goubran RA. Decoupled beamforming and noise cancellation. IEEE Trans. Instrumentation Meas. 2007;56(1):80–8.
2. Talagala Dumidu S, Zhang W, Abhayapala TD. Broadband DOA estimation using sensor arrays on complex-shaped rigid bodies. IEEE Trans Audio, Speech Lang Processing. 2013;21 (8):1573–85.
3. Minotto VP, Jung CR, Lee B. Simultaneous-speaker activity detection and localization using mid-fusion of SVM and HMMs. IEEE Trans Multimed. 2014;16(4):1032–44.
4. Yong-guang C, Xiu-he L. Indoor positioning technology based on signal strength. Electron. J. 2010;7(9):1457–8.
5. Zuo-liang Y. Experimental research on MUSIC sound source location based on microphone array. Ha er bin: Harbin Institute of Technology, 2011.
6. Huan Y, Meng-rao Z, Xiao-qiang Z. Microphone array consistency analysis based on time delay estimation. Fudan J (Nat Sci Ed.), 2017;56(02):175–81.
7. Yi-bo Z, Wei J, Li-fu W. Speech signal endpoint detection method based on microphone array adaptive nonlinear filtering. Sci Bull. 2017;33(04):199–203.

# A Research on the Improvement of Resource Allocation Algorithm for D2D Users

Yan-Jun Liang[1] and Hui Li[1,2(✉)]

[1] College of Information Science and Technology, Hainan University, Haikou, China
1307431141@qq.com, hitlihuilll2@163.com
[2] Marine Communication and Network Engineering Technology Research Center of Hainan Province, Hainan University, Haikou, China

**Abstract.** In the 5G communication network, D2D technology is considered to be one of the important components of the future [1]. D2D users bring multiplexing gain to the channel resources of multiplexed cellular users, improve system communication capacity, and reduce communication delay and terminal performance [2]. D2D also brings considerable interference to the reuse of cellular user channel resources [3]. The research on the resource allocation algorithm in D2D communication system mainly aims to reduce the interference between D2D users and cellular users. The improved bilateral rejection algorithm based on the best algorithm improves the stability and system performance of the whole communication system to a large extent [4].

**Keywords:** Wireless communication · Resource allocation · D2D · Interference suppression · Bilateral rejection algorithm

## 1 Introduction

In the past research on network communication systems, D2D technology has become an important research direction in the 5G network framework due to its low latency, higher transmission rate, and higher spectrum utilization [5]. The D2D user brings multiplexing gain to the channel resources of the multiplexed cellular users, improves the system communication capacity, and reduces the communication delay and the power consumption of the terminal [6]. However, while D2D users access traditional cellular communication networks with the reuse of communication resources, the problem of inter-user interference is caused [7]. An effective resource management strategy will solve the interference problem, and it's an important way to make D2D users and traditional cellular users gain higher spectrum utilization and increase the capacity of communacation system [8]. This paper proposes a new scheme to satisfy the communication between D2D users and cellular devices sharing spectrum resources, and studies the optimization of D2D communication resource allocation algorithm under this model.

The main work of this paper includes: (1) Propose and build a new D2D into the existing cellular network communication system model. (2) Try an existing resource allocation algorithm for the model to satisfy the frequency reuse of D2D users and cellular users. (3) Improve the existing resource allocation algorithm, propose a new


Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1_172

bilateral elimination algorithm, and compare the simulation. (4) Simulation shows that the improved bilateral culling algorithm can effectively improve the stability of the system communication link and increase the system communication capacity (Fig 1).

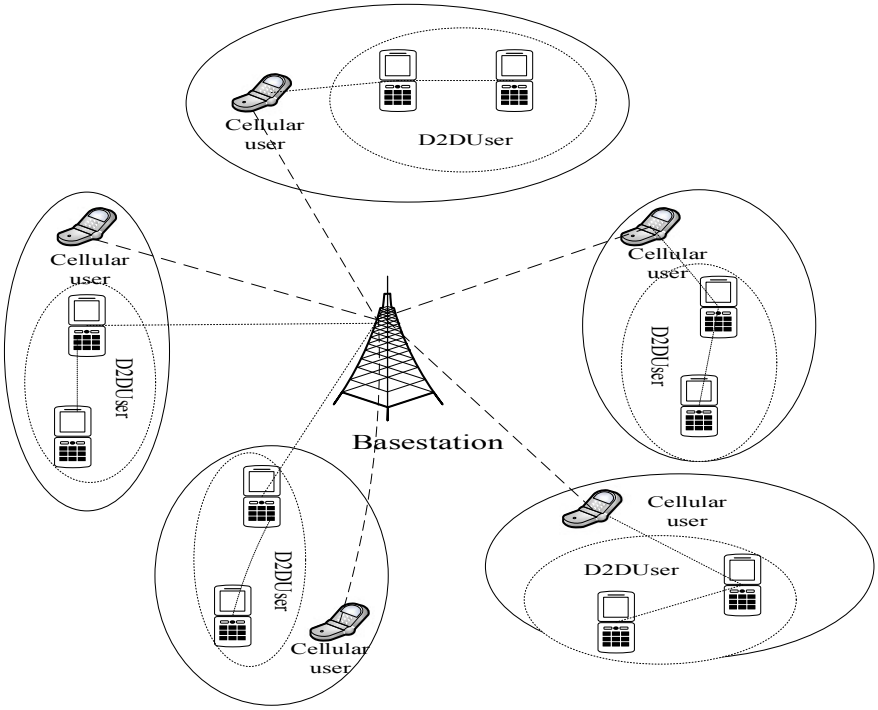



**Fig. 1.** Cell model

## 2 System Model

There are $M$ cellular users and $N$ D2D users in one cell, and each pair of D2D users includes one transmitting user $D_n$ and receiving user $D_{n'}$, and when the nth pair of D2D users multiplex the channel resources of the $m$th cellular user, the D2D user interference occurs for the $m$th cellular user, and the base station in the system also interferes with the D2D receiving user [9]. The number of channel resources in the cell is $K$, because some D2D users in the cell multiplex cellular user channel resources, and some D2D users occupy dedicated channel resources, so the total channel resource quantity of the system is greater than the number of cellular users, and is smaller than the sum of the number of D2D users and the number of cellular users [10], that is, $N < M < K < N + M$.

The SINR of receiving part in the D2D pairing-user could be calculated by formula $\gamma_{n,m} = \frac{\alpha_{n,m} P_{n,m} G_{n,n'}}{I_{BS,n'} + N_0}$, $\alpha_{n,m}$ is the nth pair of D2D users multiplex the channel resources of the $m$th cellular user, $P_{n,m}$ represents the transmit power of the D2D pair of users transmitting the user $D_n$ of the $m$th cellular user communication resource, $I_{BS,n'}$ is the

transmission interference from the base station received by the D2D to the receiving user in the user, and $N_0$ is the received noise power of $D_n$, $G_{n,n'}$ is the link gain between the $n$th D2D user pair [11].

The signal-to-noise ratio (SINR) of cellular users is:

$$\gamma_m = \frac{P_{BS,m} G_{BS,m}}{I_{n,m} + N_0} \tag{1}$$

where $P_{BS,m}$ is the base station transmit power received by the $m$th cellular user, $G_{BS,m}$ is the link gain between the $m$th cellular users of the base station, and $I_{n,m}$ is the interference of the sender $D_n$ from the nth pair D2D users to the cellular user. $I_{n,m} = P_{n,m}, G_{n,m}, G_{n,m}$ represents the link gain for the sender for the cellular user and the D2D user.

According to the Shannon formula, the communication capacity of the system is as follows:

$$C = \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{k=1}^{K} \log_2(1 + \mathrm{SINR}_{n,m}^{k}) + \sum_{m=1}^{M} \sum_{k=1}^{K} \log_2(1 + \mathrm{SINR}_{m}^{k}) \tag{2}$$

## 3 Algorithm Analysis

Using the appropriate resource allocation algorithm to improve the communication capacity of the communication system is the focus of this paper. Based on the above-mentioned system model, the related algorithms are proposed to realize the rational allocation of resources.

### 3.1 Game Playing Algorithm

Establish a game model and define the communication network power control model for cellular users and D2D users to coexist with users as follows:

$$G = (L, P, C) \tag{3}$$

L Link set between communication users.

P The set of user transmit powers participating in the game, values $[0, P_D^{\max}]$, $P_D^{\max}$ is the maximum transmit power of the D2D user.

C Cost function set.

$$C_n(P_n, P_{-n}) = \left( \frac{G_{n,m} P_n}{I_{m,n} + N_0} - \mathrm{SINR} \right)^2 + \beta_n P_n \tag{4}$$

$P_n$ The transmit power policy value of the $n$th D2D user pairs.

$P_{-n}$ Other users (except the nth pair D2D user pairs) interfere with the link's power policy value.

Under this model, each D2D user can choose its own optimal cost function to determine its own transmit power, and establish a minimum cost function for the communication link:

$$\min_{p_n \in [0, P_D^{\max}]} C_n(P_n, P_{-n}) \tag{5}$$

Then, the distributed iterative algorithm is used to find the Nash equilibrium point power that satisfies the objective function, and the partial derivative of the cost function is obtained.

$$\frac{\partial C_n}{\partial P_n} = \frac{2G_{n,m}}{I_{m,n} + N_0} \left( \frac{G_{n,m} P_n}{I_{m,n} + N_0} - \text{SINR} \right) + \beta_n \tag{6}$$

Let the formula be equal to 0 and we can get the follows:

$$P_n = \frac{I_{m,n} + N_0}{G_{n,m}} \left( \text{SINR}_{n,m} - \frac{\beta_n (I_{m,n} + N_0)}{2G_{n,m}} \right) \tag{7}$$

The best response function of the link $l_n$ can be obtained by using the above formula.

$$P^* = f(P_n) \tag{8}$$

According to the given optimal power response function, the optimal power can be achieved by the distributed power allocation algorithm. We get the simplification of the above formula:

$$I_{m,n} + N_0 = I_n \tag{9}$$

$$P_n^{q+1} = \frac{I_{-n}^q}{G_{n,m}} \left( \text{SINR}_{n,m} - \frac{\beta_n I_{-n}^q}{2} \right) \tag{10}$$

In one cycle, all transmitting users should adjust the transmission power of the user by accepting the interference value and the best response function of the user feedback until the transmission power satisfies $\left| P_n^{q+1} - P_n^q \right| < \Delta_P$, and the task link converges to the Nash equilibrium state.

### 3.2 Optimal Algorithm

The process of multiplexing the *m*th cellular user for the *n*th D2D user is to be analyzed. During the D2D link establishment process, it is interfered by the cellular user and the signal-to-noise ratio that the D2D communication establishment condition needs to satisfy is:

$$\gamma_{n,m} = \frac{P_{n,m}H_n}{I_{m,n} + N_0} \geq \gamma_{th}^{D} \tag{11}$$

$P_{n,m}$ is the transmit power of the D2D pair user transmitting user $D_n$ multiplexed with the mth cellular user communication resource, $H_n$ is the link gain between $D_n$ and $D_n$, $I_{m,n}$ is the D2D for the user receiving user $D_n$ The transmission interference from the cellular user, $N_0$ is the noise received by $D_n$. As we have set the threshold of SINR for the communication system,only when the receiving SINR of the receiving part is greater than the threshold that the quality of the D2D user can be guaranteed.

Then, the minimum transmission power of the sender of the D2D user pair can be calculated as:

$$P_{n,m} = \frac{\gamma_{th}^{D}(I_{m,n} + N_0)}{H_{n,n'}} \tag{12}$$

In this communication system, cellular users have higher priority than D2D users. So base stations can determine whether D2D users can access cellular communication systems. To facilitate the algorithm calculation, a control factor $\theta$ is set to indicate whether the D2D user is allowed to multiplex the channel resources of the cellular user, and the value is 0/1. $I_0$ is set as the interference threshold of the base station. If $I_{n,m} \leq I_0$, the system allows the nth D2D users to multiplex the channel resources of the $m$th cellular user, $\theta = 1$, and vice versa,$\theta = 0$.

If $N_1$ D2D user pairs are unable to multiplex cellular user channel resources, when $N_I \leq K + M$, these multiplexed D2D users can still access the communication system through dedicated communication channel resources. When $N_1 > K + M$, insufficient dedicated communication channel resources result in $N_1 - K + M$ D2D user pairs being completely excluded from the communication system.

The threshold for setting the interference of the D2D user to the base station is $C_1$, and the objective function for establishing the D2D communication link is:

$$\min_{\rho_{n,m}} \sum_{n=1}^{N_2} \sum_{m=1}^{K} p_{n,m} I_{n,m} \tag{13}$$

If $\min_{p_{n,m}} \sum_{n=1}^{N_2} \sum_{m=1}^{K} p_{n,m} I_{n,m} > C_1$, then the final channel allocation matrix is calculated.

### 3.3 Bilateral Rejection Algorithm

Set the link gain between cellular users and D2D users as a fixed value. The user acceptance of noise power is also a fixed value. The transmission power of the user was determined by the interference value. Then the objective function of the maximum system capacity can be converted to the maximum interference value that the cellular user can withstand.

$$\min_{\beta_{n,m}^k} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{k=1}^{K} \beta_{n,m}^k I_{m,n}^k + \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{k=1}^{K} I_{n,m}^k \tag{14}$$

In the communication system, some D2D users multiplex cellular user communication resources, and another part of D2D users use dedicated communication resources. D2D users do not interfere with other cellular users when using dedicated communication resources, and are not interfered by other cellular users. The above formula can be turned into:

$$\min_{\beta_{n,m}^k} \sum_{n=1}^{N'} \sum_{m=1}^{M} \sum_{k=1}^{K} \beta_{n,m}^k I_{m,n}^k + \sum_{n=1}^{N'} \sum_{m=1}^{M} \sum_{k=1}^{K} I_{n,m}^k \tag{15}$$

$$\begin{cases} \beta_{n,m}^k = \{0,1\}, \forall n,m,k; \\ \sum\limits_{n=1}^{N'} \sum\limits_{m=1}^{M} \beta_{n,m}^k = 1, \forall k; \\ \sum\limits_{n=1}^{N'} \sum\limits_{m=1}^{M} \beta_{n,m}^k = 1, \forall m; \\ \sum\limits_{m=1}^{M} \sum\limits_{k=1}^{K} \beta_{n,m}^k = 1, \forall n; \end{cases} \tag{16}$$

Algorithm steps:

1. Setting parameters(number of communication resource blocks $K$, number of cellular users $M$, number of D2D users $N$, interference threshold of D2D users $I_{th}^D$, and interference threshold of cellular users $I_{th}^c$ )
2. Calculating the interference matrix of the D2D user multiplexed cellular user communication resources $I_{N\times M}$ and the interference matrix of the D2D users of the multiplexed communication resources by the cellular users $I_{M\times N}$.
3. Calculating $I_{N\times M}$, we can get the minimum value $I_{n,m}^{\min}$, and if $I_{n,m}^{\min} C_{th}$, we find $I_{m,n}$ in $I_{M\times N}$, when $I_{m,n} D_{th}$, the BS will allocate resource block $k$ to the cellular user $m$ and the D2D user $n$. And then Excluding resource block $k$ from $K$ to indicate that the D2D user $n$ and the cellular user $m$ are no longer involved in the resource allocation progress. If it is not satisfied, the same operation is performed to the second smallest $I_{n,m}$ in the sorting, and so on to the end.
4. If there is a cellular user failing to allocate communication resources, then the resource blocks in $K$ will be preferentially assigned to these cellular users, and after the cellular users are fully satisfied, the allocation of dedicated communication resources will be performed.

## 4 Simulation Analysis

We simulate and analyze the above optimal algorithm, game algorithm, and bilateral culling algorithm. To reduce the complexity of the algorithm, a base station is set in a cell, so that the D2D user pair and the cellular user are evenly distributed in the cell, and each cellular user is preset to occupy only one channel resource, and only one D2D user pair can be restored. Using a cellular user channel, the average is taken after multiple simulations (Fig. 2).

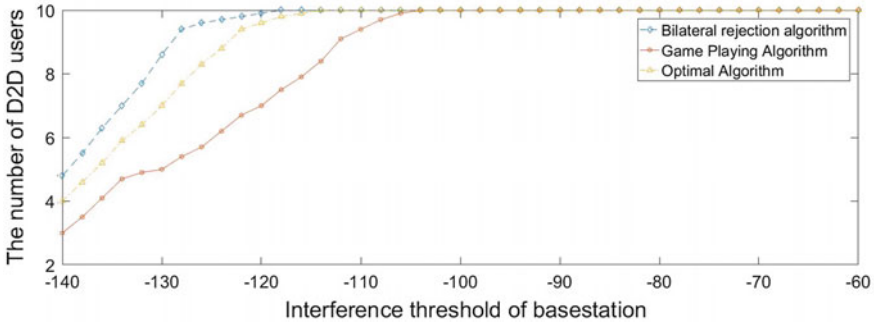



**Fig. 2.** Relationship between the interference threshold of a base station and the number of D2D users allowed to access

The figure above shows the relationship between the interference threshold of the base station and the number of D2D users allowed to communicate. It can be seen that as the interference threshold of the base station becomes larger, the number of D2D users that can be allowed to communicate is gradually increasing, but eventually an equilibrium value will be reached, i.e., the system will allow up to 10 pairs of D2D users to communicate. The simulation results can clearly see the comparison of the three algorithms. We name the improved algorithm proposed in this paper as the bilateral culling algorithm.It could reach the maximum allowable communication D2D user pair when the interference threshold is −120 dBm. The signal is reached at −115 dBm, and the game algorithm reaches the D2D user-to-communication upper limit at −108 dBm. From −140 dBm, the bilateral culling algorithm can allow five pairs of D2D users to communicate with each other, and the best algorithm and game. The algorithm can only do four pairs and three pairs.

As shown in Fig. 3, it can be seen that as the distance between users increases, the number of D2D users that can access the system will decrease. This is because the D2D user's interference value to the system will also vary with the D2D users. The maximum distance increases and increases, but the stability of the resource allocation strategy based on the double-quenching algorithm has obvious advantages. When the distance keep under 160m, the maximum number of accessers could be maintained at the number 10, and is exceeded. In the case of more than 160 m, the system still maintains considerable stability. In the case of 160 m, it still maintains considerable stability. In contrast, the communication link based on the best algorithm and the game algorithm shows considerable instability.

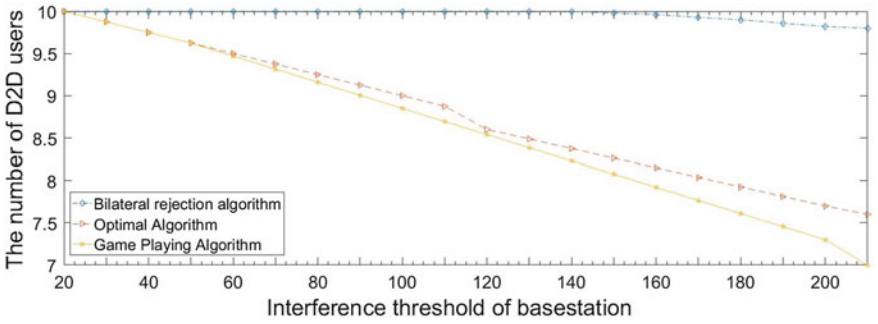



**Fig. 3.** Relationship between the maximum distance between D2D users and the number of D2D users allowed to access

## 5 Conclusions

In this paper, the D2D communication system based on cellular users is established firstly, and the optimal algorithm and the resource allocation method of the game algorithm in D2D communication are analyzed in detail. Finally, an improved method is proposed based on the optimal algorithm. The simulation results of the bilateral culling algorithm shows that the proposed bilateral culling algorithm could effectively reduce the complexity of algorithm, and it keeps the considerable advantage of the system communication link. The bilateral culling algorithmgreatly greatly increases the accessibility of D2D users to the cellular communication systems.

**Acknowledgments.** This work was supported by Hainan Provincial Key R. & D. Projects of China (ZDYF2016010 and ZDYF2018012) and the National Natural Science Foundation of China (No. 61661018).

## References

1. Fodor G. Design aspects of network assisted device-to-device communications. IEEE Commun Mag. 2012;50(3):170–7.
2. Yu CH, Tirkkonen O, Doppler K, Ribeiro C. On the performance of device-to-device underlay communication with simple power control. In: Proceeding of IEEE Vehicular Technology Conference; 2009. p. 1–5.
3. Moltchanov D. Distance distributions in random networks. Ad Hoc Netw. 2012;10(6):1146–66.
4. Lin X, Andrews J. Optimal spectrum partition and mode selection in device-to-device overlaid cellular networks. In: Proceedings of IEEEGLOBECOM; 2013. p. 1837–42.
5. Lei L, Zhong Z, Lin C, Shen X. Operator controlled device to-device communications in LTE-advanced networks. IEEE Wireless Commun. 2012;19(3):96–104.
6. Xiao X, Tao X, Lu J. A QoS-aware power optimization scheme in OFDMA systems with integrated device-to-device (D2D) communications. In: Proceedings of IEEE Vehicular Technology Conference; 2011. p. 1–5.

7. Jindal N, Weber S, Andrews JG. Fractional power control for decentralized wireless networks. IEEE Trans Wireless Commun. 2008;7(12):5482–92.
8. Zhang X, Haenggi M. Random power control in Poisson networks. IEEE Trans Commun. 2012;60(9):2602–11.
9. Gu J, Bae SJ, Choi BG, Chung MY. Dynamic power control mechanism for interference coordination of device-to-device communication in cellular networks. In: Proceeding of 3rd International Conference on Ubiquitous and Future Networks; 2011. p. 71–5.
10. Corson M. Toward proximity-aware internetworking. IEEE. Wireless Commun. 2010;17 (6):26–33.
11. Janis P. Device-to-device communication underlaying cellular communications systems. Int J Commun Netw Syst Sci. 2009;2(3):169–78.

# Author Index




Q. Liang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 516,
https://doi.org/10.1007/978-981-13-6504-1

**M**