Lecture Notes in Electrical Engineering 873

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

Communications, Signal Processing, and Systems

Proceedings of the 11th International Conference on Communications, Signal Processing, and Systems, Vol. 2

Springer

Lecture Notes in Electrical Engineering 873

Series Editors

Leopoldo Angrisani, *Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Napoli, Italy*
Marco Arteaga, *Departamento de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico*
Samarjit Chakraborty, *Fakultät für Elektrotechnik und Informationstechnik, TU München, München, Germany*
Jiming Chen, *Zhejiang University, Hangzhou, Zhejiang, China*
Shanben Chen, *School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China*
Tan Kay Chen, *Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore*
Rüdiger Dillmann, *Gebäude 07.21, University of Karlsruhe (TH) IAIM, Karlsruhe, Baden-Württemberg, Germany*
Haibin Duan, *Beijing University of Aeronautics and Astronautics, Beijing, China*
Gianluigi Ferrari, *Dipartimento di Ingegneria dell'Informazione, Sede Scientifica Università degli Studi di Parma, Parma, Italy*
Manuel Ferre, *Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain*
Faryar Jabbari, *Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA*
Limin Jia, *State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China*
Janusz Kacprzyk, *Intelligent Systems Laboratory, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*
Alaa Khamis, *Department of Mechatronics Engineering, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt*
Torsten Kroeger, *Intrinsic Innovation, Mountain View, CA, USA*
Yong Li, *College of Electrical and Information Engineering, Hunan University, Changsha, Hunan, China*
Qilian Liang, *Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA*
Ferran Martín, *Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain*
Tan Cher Ming, *College of Engineering, Nanyang Technological University, Singapore, Singapore*
Wolfgang Minker, *Institute of Information Technology, University of Ulm, Ulm, Germany*
Pradeep Misra, *Department of Electrical Engineering, Wright State University, Dayton, OH, USA*
Subhas Mukhopadhyay, *School of Engineering, Macquarie University, NSW, Australia*
Cun-Zheng Ning, *Department of Electrical Engineering, Arizona State University, Tempe, AZ, USA*
Toyoaki Nishida, *Department of Intelligence Science and Technology, Kyoto University, Kyoto, Japan*
Luca Oneto, *Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Genova, Genova, Italy*
Bijaya Ketan Panigrahi, *Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India*
Federica Pascucci, *Department di Ingegneria, Università degli Studi Roma Tre, Roma, Italy*
Yong Qin, *State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China*
Gan Woon Seng, *School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore*
Joachim Speidel, *Institute of Telecommunications, University of Stuttgart, Stuttgart, Germany*
Germano Veiga, *FEUP Campus, INESC Porto, Porto, Portugal*
Haitao Wu, *Academy of Opto-electronics, Chinese Academy of Sciences, Haidian District Beijing, China*
Walter Zamboni, *Department of Computer Engineering , Electrical Engineering and Applied Mathematics, DIEM—Università degli studi di Salerno, Fisciano, Salerno, Italy*
Junjie James Zhang, *Charlotte, NC, USA*

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, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

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

Communications, Signal Processing, and Systems

Proceedings of the 11th International Conference on Communications, Signal Processing, and Systems, Vol. 2

Springer

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

Xin Liu
Dalian University of Technology
Dalian, China

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

Wei Wang
Tianjin Normal University
Tianjin, China

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

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-99-1259-9 ISBN 978-981-99-1260-5 (eBook)
https://doi.org/10.1007/978-981-99-1260-5

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Synthetic Aperture Radar Image Ship Detection Based on YOLO-SARshipNet

Bing Chen[1], Yuting Zhu[2], Chen Wang[1], and Xiaoqing Wang[2(✉)]

[1] Transport Planning and Research Institute Ministry of Transport, Beijing, China
[2] Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou, China
xqwang@mail.sysu.edu.cn

Abstract. Automatic ship detection from synthetic aperture radar (SAR) imagery plays a significant role in many urban applications. Recently, owing to the impressive performance of deep learning, various SAR ship detection methods based on the convolution neural network (CNN) have been proposed for optical SAR images. However, existing CNN-based methods and spatial-domain-based methods exhibit certain limitations. Some algorithms do not consider the detection speed and model scale when improving the detection accuracy, which limits the real-time application and deployment of SAR. To solve this problem, a lightweight, high-speed and high-accurate SAR ship detection method based on yolov3 has been proposed. First, the backbone part of the model is improved, the pure integer quantization network is applied as the core to reduce a small amount of accuracy while reducing the model scale by more than half; Second, modify the feature pyramid network to improve the detection performance of small-scale ships by enhancing feature receptive fields; third, introduce the IoU loss branch to further improve the detection and positioning accuracy; finally, the feature distillation is applied to handle the problem of accuracy decrease caused by model integer quantization. The experimental results on the two public SAR ship datasets show that this algorithm has certain practical significance in the real-time SAR application, and its lightweight parameters are helpful for future FPGA or DSP hardware transplantation.

Keywords: Ship detection · Synthetic aperture radar · Pure integer quantization · Feature pyramid network · Feature distillation

1 Introduction

Synthetic aperture radar (SAR) has important applications in various fields, such as military, agriculture, and oceanography, owing to its high resolution and all-weather features among other advantage. [1] Ship target detection is an

Bing Chen and Yuting Zhu contributed to the work equally and should be regarded as co-first authors.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 1–11, 2023.
https://doi.org/10.1007/978-981-99-1260-5_1

important link in the implementation of marine surveillance, and has important applications in military reconnaissance, marine transportation management, and maritime crime combat. In this way, SAR images become an effective remote sensing data source for ship targets, and SAR ship target detection has been widely studied. [2]

The traditional ship detection methods are mainly based on the statistical distribution of sea clutter, such as the constant false alarm rate (CFAR) method, and the method based on machine learning to extract features [3]. However, these traditional methods are highly dependent on the pre-defined distribution of ship geometric features or semantic features, which degrades the ship detection performance in SAR images. Therefore, it is difficult for these methods to perform ship detection accurately and robustly. In addition, some super-pixel-based ship detection methods have been proposed, which are also difficult to accurately detect ships in nearshore and offshore scenes [4]. Following the advancement of deep learning technology, many methods have been designed to use deep learning to solve the problem of SAR ship detection. The target detection algorithm based on deep learning is not limited by the scene, does not need sea-land-separation, and can learn the characteristics of ships spontaneously, which improves the shortcomings of traditional methods to a certain extent [5].

Supervised deep learning methods often require many training samples as support. In the field of SAR image ship detection, the Naval Engineering University released the first public SAR Ship Detection Dataset (SSDD) in China [6]. Since then, there are mainly high-resolution SAR ship target datasets (SAR-Ship-Dataset) constructed by the team of the Institute of Aeronautical and Astronautical Information, Chinese Academy of Sciences [7]. On this basis, SAR ship detection methods based on deep learning have developed rapidly. On the SSDD dataset, [8] uses the optimized Faster R-CNN model, with feature fusion and transfer training to achieve a detection accuracy of 78.8%. [9] uses the improved YOLOv3 [10] algorithm, combined with the feature pyramid structure, to improve the detection performance of small-sized ships, the accuracy is improved but the detection speed is reduced to a certain extent. [11] proposed a detection method SARShipNet-20 based on the depth separation convolutional neural network, and adopted the attention mechanism, which achieved good results in detection accuracy and detection speed. [7] introduces the SAR-Ship-Dataset dataset in detail and uses different object detection algorithms to conduct experiments on this dataset. [12] uses Retinal-Net as the basis, with the pyramid network and class imbalance loss function to improve the accuracy, reaching 96% accuracy. The above methods have achieved good results in SAR ship detection, but they all have the characteristics of large number of parameters, which improves the accuracy but reduces the detection speed. Therefore, there are certain defects in the hardware transplantation of the model and the application of SAR ship detection with high real-time requirements.

To solve the above shortcomings, this paper proposes a lightweight high-speed high-precision SAR ship detection method YOLO-SARshipNet based on the improved YOLOv3 framework. First, a pure integer quantization network is applied as the backbone network. The network structure is compact and efficient,

including depth separation convolution, and the size of the model parameters is 1/6 of the original YOLOv3 framework; second, the spatial pyramid pooling module is introduced into the feature pyramid, which improves the detection ability of the model for small-sized ships by enhancing feature receptive field; third, introduce the IoU loss branch to further improve the accuracy of detection and positioning; finally, feature distillation is applied to solve the problem of accuracy decrease caused by model integer quantization. The experimental results demonstrate that our method can achieve an excellent SAR image ship detection effect. The contributions of this study are as follows:

- We propose a YOLO-SARshipNet detection model that successfully reduces the model size by one-sixth without sacrificing accuracy. The detection model can effectively detect ships in the SAR images;
- We introduce feature distillation to successfully solve the problem of model accuracy degradation;
- The experimental results on SSDD and SAR-Ship-Dataset show that YOLO-SARshipNet has achieved relatively good results in terms of accuracy, speed, and model size, and its lightweight parameters are helpful for FPGA or DSP hardware porting.

The rest of this paper is organized as follows. In Sect. 2, the principles and implementation of the proposed algorithm is introduced. Section 3 presents experimental results and discussion. Finally, a summary is provided in Sect. 4.

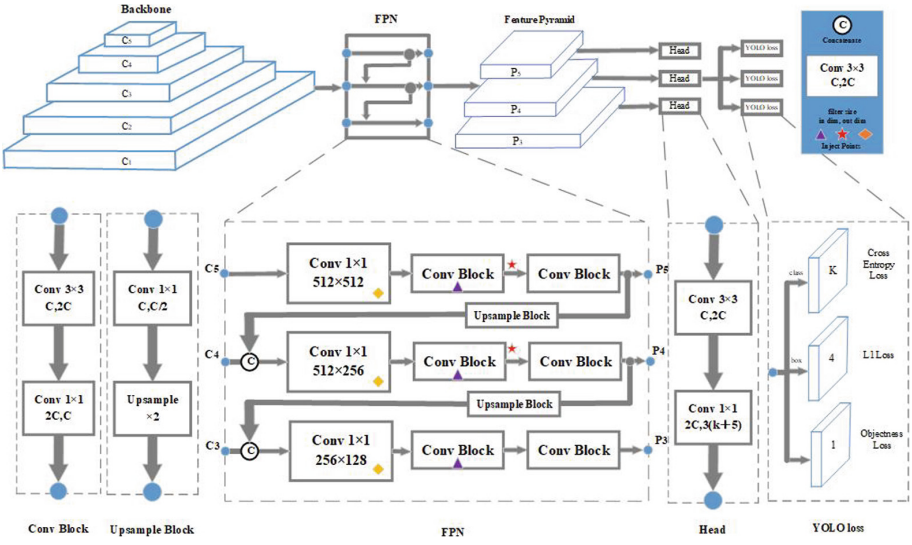


Fig. 1. Schematic of YOLO-SARshipNet algorithm flow.

2 Model Structure and Improvement Methods

Synthetic aperture radar (SAR) has important applications in various fields, such as military, agriculture, and oceanography, owing to its high resolution and all-weather features among other advantage. Ship target detection is an important link in the implementation of marine surveillance, and has important applications in military reconnaissance, marine transportation management, and maritime crime combat. In this way, SAR images become an effective remote sensing data source for ship targets, and SAR ship target detection has been widely studied.

Target detection models based on deep learning have achieved great success in natural image target detection tasks, but most of those models are not suitable for SAR image ship detection. On the one hand, there is a big difference between SAR images and ordinary optical images; on the other hand, most detectors are designed for multi-target detection. If they are directly applied to SAR detection, there must be a certain degree of redundancy. Therefore, different from the existing optical image target detection models, this paper designs a YOLO-SARshipNet network that is more suitable for SAR ship detection and deployment based on the YOLOv3 network [10].

The overall network structure of the designed YOLO-SARshipNet is shown in Fig. 1. First, a deep convolutional backbone network is used to extract features of ship targets in SAR images; second, a feature pyramid network is constructed using the extracted feature layers to fuse features of different scales; finally, on the fused feature map, the position information of the target, the confidence information of the category and the intersection ratio information of the position are applied for regression prediction.

2.1 Backbone

The traditional YOLOv3 [10] uses Darknet-53 as the feature extraction skeleton network, which consists of a series of 1 and 3 convolutional layers (each convolutional layer is followed by a normalization layer and a pooling layer), the network contains a total of 53 convolutional layers, and a residual structure is introduced [13], so it is called Darknet-53. However, the YOLOv3 is aimed at optical images containing multi-scale and multi-target, while the SAR image ship target is small, the size is relatively single, and the gradient between the target area and the background area is not obvious. In addition, the size of the YOLOv3 model using Darknet-53 as the backbone network that reaches 139.25 MB, which is not suitable for hardware transplantation of FPGA or DSP. To handle these problems, this paper constructs a pure integer quantization network as a backbone network.

The pure integer quantization network is based on the Efficient-Net network proposed in [14]. Applying EfficientNet-B4 as the skeleton network of the model can reduce the number of parameters of the model by half without reducing the detection accuracy. However, there are still two problems in the deployment of

this model on embedded devices. First, many ship detection devices have limited floating-point operation support; second, each hardware supports different operation acceleration operations, and several operations in the network cannot be supported by hardware. In response to the above problems, the tools provided by Google [15] are used to quantize the model, that is, the floating-point operation of the model is converted into an integer operation. This operation not only helps the device-side operation, but also successfully reduces the size of the model. Second, the attention mechanism in Efficient-Net is removed, which facilitates device deployment and reduces the ship detection time of the model. The constructed pure integer quantization network is based on EfficientNet-B4, which contains 19 network layers and 5 down-sampling layers.

2.2 Feature Pyramid Networks

The structure of Feature Pyramid Networks (FPN) [16] is shown in Fig. 1. When the image passes through the backbone network, feature blocks with larger receptive fields and richer semantic information are finally obtained through continuous convolution and down-sampling operations. However, larger down-sampling will lose part of the semantic information, which cause missed detection of small-sized objects. In order to solve this problem, this paper uses FPN to perform convolution and up-sampling operations on the feature blocks obtained by the backbone network and adopts a feature fusion strategy to fuse the shallow semantic information and deep semantic information, thereby improving the detection ability of small-sized ships.

In addition, this paper introduces the Spatial Pyramid Pooling Layer (SPP) [18] in FPN. The SPP module applied in this paper concatenates the output of max pooling with kernel size$k = 1, 5, 9, 13$, the stride is 1, and the insertion position is the "red star" position in the Fig. 1. Introducing the SPP module to the feature blocks containing rich semantic information has the following advantages:

- The SPP module is multi-scale, which improves the receptive field of the feature blocks;
- The SPP module does not lead additional parameters, effectively ensuring the model;
- The detection accuracy of the model is effectively improved after the introduction of the SPP module.

2.3 Loss Function

After passing through the FPN, three feature layers are selected for two convolutions to obtain the following feature parameters of the SAR image: (1) The center coordinates (x, y) of the detection anchor; (2) The width and height coordinates (w, h) of the detection anchor; (3) Ship confidence C of detection anchor. We utilize these feature parameters to construct a new well-designed loss function:

$$\text{loss} = \text{loss}_{box} + \text{loss}_{obj} + \text{loss}_{IoU} \tag{1}$$

where loss_{box} is the coordinate loss of the detection anchor, defined as:

$$\begin{aligned} \text{loss}_{\text{box}} &= \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} \left(2 - w_i \times h_i\right) \\ &\left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (w_i - \widehat{w}_i)^2 + \left(h_i - \hat{h}_i\right)^2 \right] \end{aligned} \tag{2}$$

where x_i^j, y_i^j, w_i and h_i are the real coordinates of the j-th frame of the i-th grid, and $\hat{x}_i^j, \hat{y}_i^j, \widehat{w}_i$ and $\hat{h}_i$ are the corresponding predicted coordinates; $I_{ij}^{obj} = 1$ when the grid contains a ship or part of a ship, otherwise $I_{ij}^{obj} = 0$. And λ_{coord} is the weight coefficient of coordinate loss, B is the number of detection frames, and S is the number of divided grids. loss obj is the confidence loss function, which is defined as:

$$\text{loss}_{obj} = \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{noobj} \left(c_i - \hat{c}_i\right)^2 + \lambda_{obj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} \left(c_i - \hat{c}_i\right)^2 \tag{3}$$

where λ_{noobj} and λ_{obj} are the weight coefficients of the j-th box of the i-th grid with and without the target, respectively, and the value is 0.5.

loss_{IoU} is the intersection ratio loss function, defined as:

$$\text{loss}_{IoU} = -\ln IoU \tag{4}$$

where IoU is the intersection ratio of the predicted anchor and the label anchor.

$$IoU = \frac{\text{area}\left(B_G \cap B_P\right)}{\text{area}\left(B_G \cup B_P\right)} \tag{5}$$

where B_P is the predicted box, B_G is the label box, that is, IoU is the area ratio of the intersection and union of the predicted frame and the real frame area.

2.4 Feature Distillation

In this paper, the integer quantization technology is applied to reduce the number of parameters and calculation of the model, but it has a bad impact on the detection accuracy of the model. To solve this issue, we use a feature distillation-based approach.

The main idea of feature distillation is to train a small network model to learn a large pre-train network model, the large network is called the "teacher network", and the small network is the "student network". Feature distillation expects the student network to achieve accuracy like or even better than the teacher network with fewer parameters and a smaller scale. The FSP matrix is used to characterize the data association between different layers of the "teacher

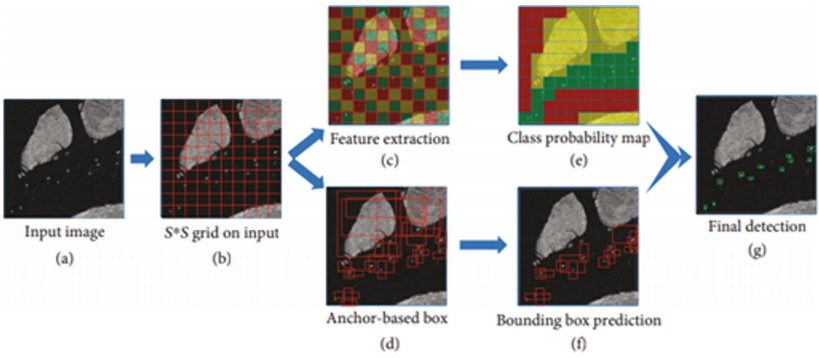


Fig. 2. The proposed YOLO-SARshipNet with feature distillation flow.

network" and the "student network", and then the $L2$ loss is used to fit the FSP matrix of the corresponding layer of the small model and the FSP matrix of the corresponding layer of the large model, as shown in Fig. 2. The complete Efficient-Net is applied as the teacher model and YOLO-SARshipNet as the student model to construct the FSP matrix of the corresponding layers. The advantage of this algorithm is that it allows the small model to learn the intermediate process and obtain more prior knowledge. We verified the effectiveness of this method through experiments, and the specific analysis is shown in Sect. 3.2.

3 Experimental Results and Discussion

In this section, we verify the effect of YOLO-SARshipNet on SSDD and SAR-Ship-Datase. All experiments were done using a single graphics card (NVIDIA GeForce GTX TITAN XP).

3.1 Dataset and Experiment Setup

In the experiment, we tested the effectiveness of the proposed algorithm using SSDD and SAR-Ship-Dataset. The SSDD dataset contains 1160 SAR images with a total of 2358 ships, and the smallest ships marked are 7×7 pixels and the largest are 211×298 pixels. The SAR-Ship-Dataset contains $43,819$SAR image slices of size 256×256, whose data sources are the Gaofen-3 satellite and the European Sentinel-1 satellite. The SAR images in the two datasets have multiple polarization modes, multiple scenes, and multiple resolutions, which can effectively verify the effect of the ship detection algorithm.

The two datasets are randomly divided into training set, validation set, and test set according to the ratio of 7 : 2 : 1. The SGD optimizer is used to iteratively update the network parameters. The initial learning rate is set to 0.0013 and the decay coefficient is 0.1. A total of 10000 iterations are performed, and the decay is once at the 8000st and 9000st times. The momentum parameter of SGD is set to 0.9. Batch-size for each training is set to 8.

The evaluation index is an important method for effectively measuring the detection effect of the algorithm. Generally, the detection quality of an image is comprehensively evaluated based on the following five aspects: precision, recall, inference time, model size, and mean Average Precision (mAP). We used the following three objective indicators to evaluate the effectiveness of the proposed algorithm:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{7}$$

$$OA = \frac{TP + FP}{TP + FP + TF + TN} \tag{8}$$

where TP, FP, TN,and FN represent the number of true positives, false positives, true negatives and false negatives, respectively.

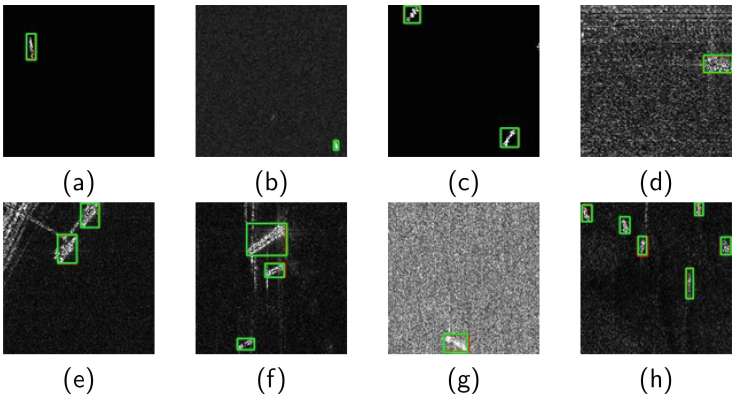


Fig. 3. Visualization of detection results on SSDD dataset.

3.2 Experiment Results

Table 1 shows the ship detection performance of different deep learning-based target detection methods on the SSDD dataset, and compares each method from Recall, Precision, and mAP. In order to ensure fairness, the compared methods all use the uniformly divided data in the experiments in this paper, and the

Table 1. Comparison of detection performance of different methods on SSDD dataset.

Method	Recall (%)	Precision (%)	mAP (%)
Faster-RCNN [10]	75.16	81.15	78.86
RetinaNet [12]	86.70	83.11	85.68
YOLOv3 [10]	93.62	88.54	90.33
YOLOv4 [17]	93.51	87.55	92.38
YOLO-SARshipNet	93.23	90.51	92.44

parameter settings, training methods and training environments are the same as those of YOLO-SARshipNet. It can be found that YOLO-SARshipNet achieves good results compared to other advanced models.

3.3 Ablation Experiment

YOLO-SARshipNet based on YOLOv3. The experiments in this section are to verify the effectiveness of the proposed improved module. This experiment used the SAR-Ship-Dataset dataset, the results are shown in Table 2, and "✔" means that this module is used.

As can be seen from Table 2, YOLO-SARshipNet is about 1/6 of the YOLOv3 model, but the inference time and accuracy are better than YOLOv3. Although the accuracy of the model is reduced by integer quantization, the prediction time is shortened again, and such a lightweight integer network is more convenient for FPGA or DSP porting. And the model overcome the above problems after feature distillation.

Figures 3 and 4 shows the ship detection results of some samples of YOLO-SARshipNet on the SAR-Ship-Dataset dataset, where the red box is the real ship, and the green box is the detected ship. It shows that YOLO-SARshipNet has a good detection effect whether it is a ship target in a blurred background, a near-shore ship, or a small target ship.

Table 2. Ablation experiments on the SAR-Ship-Dataset dataset.

Method	Backbone	FCN	IoU-Loss	Integer quantization	Feature distillation	mAP	Scale (MB)	Reference (ms)
YOLOv3						93.32	139.25	28.32
	✔					95.54	62.44	16.13
	✔	✔				96.23	62.44	17.78
	✔	✔	✔			97.19	62.44	17.78
	✔	✔	✔	✔		96.13	21.38	13.23
YOLO-SARshipNet	✔	✔	✔	✔	✔	97.34	21.38	13.23

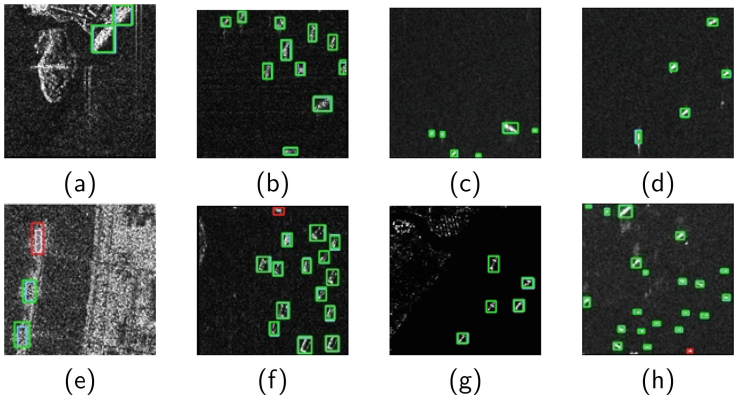


Fig. 4. Visualization of detection results on SAR-Ship-Dataset dataset.

4 Conclusion

This paper proposes a high-speed and high-precision SAR ship detection algorithm based on the YOLOv3 algorithm, namely YOLO-SARshipNet. Through the improvement of the backbone network, the accuracy is improved while reducing the size of the model; the proposed feature pyramid network effectively improves the small target detection ability of the model, and the loss function is modified for the detection of a single target SAR ship target, it also brings a certain improvement to the detection accuracy; finally, the integer quantization operation of the model allows the model to have a smaller number of parameters and calculations, and the feature distillation method has been used to solve the problem of integer quantization. The experimental results on SSDD and SAR-Ship-Dataset show the correctness and effectiveness of the method proposed in this paper, and it has certain practical significance in the real-time SAR application field. It can be transplanted to the hardware of FPGA or DSP later.

References

1. Hui, M., Xiaoqing, W., Jinsong, C., Xiangfei, W., Weiya, K.: Doppler spectrum-based NRCS estimation method for low-scattering areas in ocean SAR images. *Remote Sensing* (2017)
2. Chen, G., Li, G., Liu, Y., Zhang, X.P., Zhang, L.: Sar image despeckling based on combination of fractional-order total variation and nonlocal low rank regularization. IEEE Trans. Geosci. Remote Sens. **58**(3), 2056–2070 (2021)
3. Gui, G.: A parzen-window-kernel-based CFAR algorithm for ship detection in SAR images. 8(3), 557–561 (2011)
4. Jiang, S., Chao, W., Bo, Z., Hong, Z.: Ship detection based on feature confidence for high resolution SAR images. In: Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International (2012)

5. Wang, C., Bi, F., Liang, C., Jing, C.: A novel threshold template algorithm for ship detection in high-resolution SAR images. In: IGARSS 2016 - 2016 IEEE International Geoscience and Remote Sensing Symposium (2016)
6. Li, J., Qu, C., Shao, J.: Ship detection in SAR images based on an improved faster R-CNN. In: Sar in Big Data Era: Models, Methods & Applications (2017)
7. Wang, Y., Wang, C., Zhang, H., Dong, Y., Wei, S.: A SAR dataset of ship detection for deep learning under complex backgrounds. Remote Sensing, 11(7) (2020)
8. Li Jianwei, Q., Changwen, P.S., Bing, D.: Ship target detection in SAR images based on convolutional neural network. Syst. Eng. Electron. Technol. **40**(9), 7 (2018)
9. Changhua, H., Chen, C., Chuan, H., Hong, P., Jianxun, Z.: Small target detection of ships. In: SAR images based on deep convolutional neural networks (3) (2022)
10. Redmon, J., Farhadi, A.: YOLOv3: an incremental improvement. arXiv e-prints (2018)
11. Xiaoling, Z., Tianwen, Z., Jun, S., Shunjun, W.: High-speed and high-precision SAR ship detection based on deep separation convolutional neural network. J. Radar **8**(6), 11 (2019)
12. Wang, Y., Wang, C., Zhang, H., Dong, Y., Wei, S.: Automatic ship detection based on retinanet using multi-resolution gaofen-3 imagery. Remote Sens., 11(5) (2019)
13. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. IEEE (2016)
14. Tan, M., Le, Q.V.: Efficientnet: rethinking model scaling for convolutional neural networks (2019)
15. Abadi, M., Barham, P., Chen, J., Chen, Z., Zhang, X.: A system for large-scale machine learning. USENIX Association, Tensorflow (2016)
16. Lin, T.Y., Dollar, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. IEEE Computer Society (2017)
17. Bochkovskiy, A., Wang, C.Y., Liao, H.: YOLOv4: optimal speed and accuracy of object detection (2020)

Target Recognition and Tracking Based on Point Cloud Fusion of Automotive Millimeter-Wave Radar and Camera

Chi Zhang(✉), Siwei Li, Xu Si, and Jing Liang

University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave, Chengdu 611731, China
13708039695@163.com

Abstract. In this paper, we propose a point-cloud-based millimeter-wave radar and camera fusion system, which aims to improve the recognition rate and tracking accuracy of the fusion system. For point cloud clustering, our method adds the velocity difference parameter to the density based spatial clustering of applications with noise (DBSCAN) algorithm and selects different parameter values according to three different kinds of targets. Aiming at the problem of inaccurate estimation of monocular ranging when the angle between the target and the sensor is formed, a distance-pixel-based algorithm is proposed to significantly reduce the above error. Experimental results show that the fusion system can improve the detection rate and tracking efficiency compared with the single camera scheme under the requirement of real-time.

Keywords: Target Recognition · Target Tacking · Multimodal Sensor Fusion · Density Based Spatial Clustering · Monocular Depth Estimation

1 Introduction

With the improvement of the national economy and technology, cars have entered thousands of households and become an important means of transportation in people's daily life. However, the increase of vehicle ownership has caused many social problems. To improve the traffic efficiency, autonomous driving technology has come into being. The scene involved in autonomous driving is complicated. How to obtain a higher target recognition rate and a better tracking accuracy under the requirement of real-time is the problem that autonomous driving needs to solve. This depends on using many different sensors and fusing them.

Sensor fusion technology is widely used in intelligent medical treatment [1], image processing [2], etc. Common sensors for autonomous driving include cameras, millimeter-wave radar, and lidar. For the fusion of camera and lidar, Cheng et al. proposed a fusion algorithm based on two-dimensional projection [3], which presents lidar data on two-dimensional images and uses end-to-end convolutional neural network (CNN) to predict target confidence degree and bounding box.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 12–19, 2023.
https://doi.org/10.1007/978-981-99-1260-5_2

However, this method needs to project 3D dense point cloud data, resulting in the loss of information in dimension. PointNet uses point-based fusion to solve the problem of information loss [4]. The network uses multilayer perceptron (MLP) to learn high-dimensional local features and then connects the global features with the local features of each point for following label prediction. However, the number of parameters in the fusion of lidar and camera is large, which leads to a long processing time. For the fusion of millimeter-wave radar and camera, Ren et al. proposed a fusion system [5,6]. The system uses the mapping relationship between radar and image to generate the region of interest (ROI), and then calculates the intersection ratio with the detection boxes of the image to locate and recognize the target. The system meets the real-time requirements, but the recognition rate and positioning accuracy are reduced.

In this paper, we design an automotive millimeter-wave radar and camera fusion system, which can effectively meet the needs of road targets' (including pedestrians, bicycles, and cars) recognition and tracking.

Our major contributions are as follows:

1. Aiming at the problem of large redundant computation of point cloud data in millimeter wave radar, we proposed an improved DBSCAN point cloud clustering algorithm with speed difference (V) as a new parameter, and the parameters of three types of targets are designed respectively to improve the accuracy of point cloud clustering.
2. Aiming at the problem of poor image depth calculation by traditional methods, distance-pixel-based monocular depth estimation is proposed to reduce the range estimation error when there is an angle between the target and the camera.

The rest of this article is arranged as follows. Section 2 introduces the point cloud clustering algorithm, monocular depth estimation algorithm, target recognition, and tracking algorithm used in this paper. In Sect. 3, we discuss the accuracy of monocular depth estimation for different types of targets in real measurements and compare the recognition accuracy before and after fusion. Section 4 is the conclusion.

2 Methodology

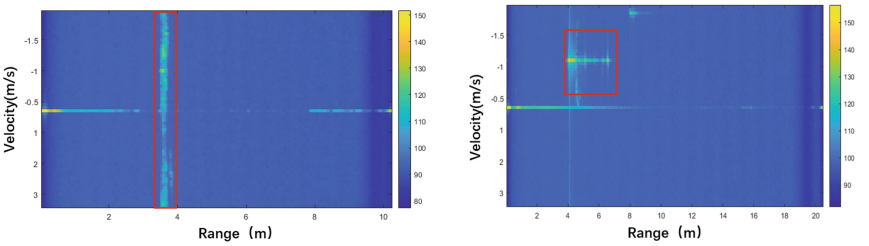
2.1 Point Cloud Clustering Algorithm Based on Automotive Millimeter Wave Radar

After channel correction, range and doppler fast Fourier transform (FFT) transform, digital beamforming, constant false-alarm rate detection (CFAR), and other processing, radar echoes are transformed into point cloud data with Angle, range, and velocity information. Usually, because the range resolution of millimeter-wave radar is smaller than the size of common targets, multiple point clouds can be resolved from the same target. If these point cloud data are not clustered, it will lead to data redundancy and greatly increase the amount

of computation. For the irregular shape of data and the unknown number of targets, DBSCAN [7]can cluster radar point clouds well.

DBSCAN algorithm has two parameters, one is that the radius (Eps) represents the minimum distance between two classes, and the other is the minimum number of points (MinPts) represents the least number of points in a class. Setting an appropriate MinPts can not only reduce the number of anomalistic points but also avoid the error of merging small clusters with large clusters.

We collect data for pedestrians, bicycles, and cars. The range-Doppler feature maps of the pedestrians and cars are shown in the Fig. 1.



(a) (b)

Fig. 1. Range doppler feature maps of various targets.(a) Pedestrian (b) Car

It can be seen that point clouds generated by pedestrians are narrow in distance dimension and have wide expansion in velocity dimension. On the contrary, for large targets such as cars, point clouds are wider in range dimension but narrower in velocity dimension. Based on this property, we add a variable named speed difference V as the third parameter of clustering based on the original DBSCAN algorithm, that is, the speed difference between any two data points in the same class must be less than V. The specific steps of the algorithm are in the Algorithm 1.

After multiple experiments, the DBSCAN parameter selection table for different targets is obtained, as shown in Table 1. We believe that the car category has a larger size and a larger minimum number of points. The speed difference between these points is more stringent and therefore smaller, while pedestrians are the opposite.

Table 1. DBSCAN parameter selection for different categories of targets

Category	E_{ps} (m)	Min_{pts}	V (m/s)
Pedestrian	0.6	1	2
Bicycle	1.6	2	1.5
Car	4	3	1

Algorithm 1. Improved DBSCAN Algorithm for point cloud processing of automotive millimeter wave radar

Input

Point cloud set DB, E_{ps}, Min_{pts}, V

for each point P in database DB **do**
  if P is marked **then**
    goto Step 1.
  else
    if the number of points in $E_{ps}(P)$ is less than Min_{pts} **then**
      Mark P as boundary point or noise point
    else
      Mark P as the core point, establish a new cluster C, and add points in the neighborhood of P that satisfy the intra-class velocity difference less than V to C
    end if
    for each unmarked point Q within the neighborhood of P **do**
      if the number of points in $E_{ps}(Q)$ is less than Min_{pts} **then**
        goto Step 9.
      else
        Add unmarked points in the $E_{ps}(Q)$ neighborhood that satisfy the intra-class velocity difference less than V to C
      end if
    end for
  end if
end for

2.2 Monocular Depth Estimation Algorithm Based on Distance-Pixel

Unlike radar, which calculates the range of the target by echo, monocular depth estimation uses the image to estimate the range of the target. The common method of monocular depth estimation is to use target detection bounding boxes combined with the keyhole imaging model. [8] So the depth can be calculated by the following formula.

$$Z = \frac{H}{y} \times f \tag{1}$$

H represents the actual height of the target to be measured, y represents the height of the detection bounding box, f represents the focal length of the camera, and Z represents the distance between the target and the camera. In practical scenes, there is an angle between the target and the optical axis of the camera (the transverse distance is not 0), which leads to a large result compared with the true value when directly using the above formula. Our improved calculation method is shown in the following formula.

$$\begin{cases} d_x = \frac{(x-x_0)}{d_y} \times \beta \\ d_y = \frac{F \times H_i}{h} \end{cases} \tag{2}$$

The coordinates of the center point of the target detection bounding box are [x, y, w, h], H_i is the true height of the target of class i (1.7 m for pedestrians, 1.5 m for vehicles, and 0.9 m for bicycles), and the coefficient β represents the product value of the vertical distance and the number of pixels corresponding to each meter in the horizontal direction which will change with the target class and need to obtain through experiments. x_0 represents the x-coordinate of the camera's optical center, while F represents the distance between the camera's optical center and the imaging plane.

2.3 Target Recognition and Tracking Algorithm

We designed a target recognition and tracking architecture based on the fusion of automotive millimeter-wave radar and camera. For the recognition part, the radar obtains the target type by calculating the radar cross section (RCS) value, while the camera uses the YOLOV3 network. For the tracking part, the camera adopts the distance-pixel-based monocular depth estimation algorithm, and the radar adopts the traditional CFAR detection and the improved DBSCAN algorithm. The position measurement results of the two methods are fused by the extended Kalman filtering method. The overall tracking algorithm flow is shown in the Fig. 2. The entire processing and fusion framework is shown in Fig. 3.

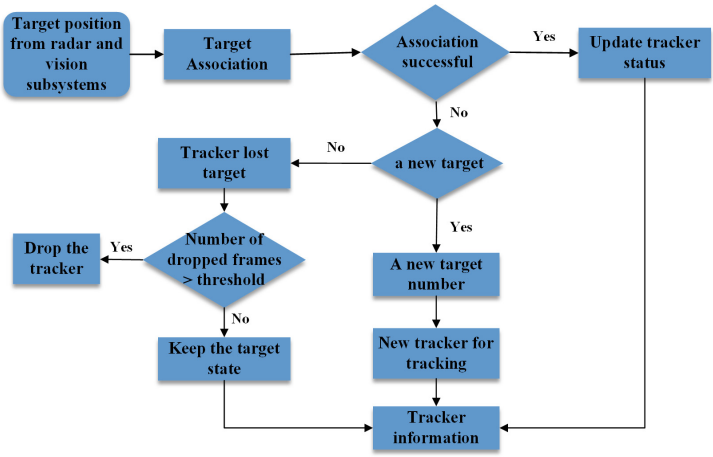


Fig. 2. Target Tracking Flowchart

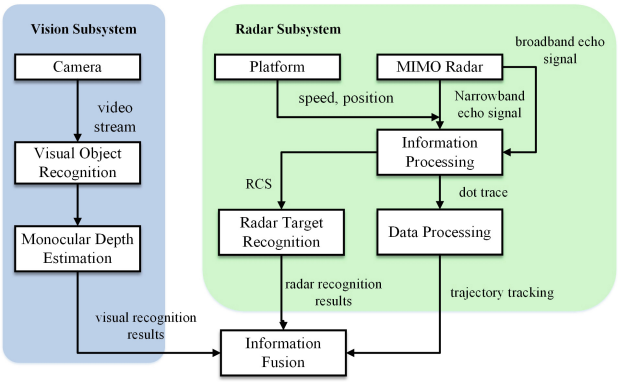


Fig. 3. Target recognition and tracking framework for millimeter-wave radar and camera fusion

3 Experiment and Analysis

3.1 Experiment on the Range-Pixel-Based Monocular Depth Estimation Algorithm

In order to obtain the value of coefficient β in the range-pixel-based monocular depth estimation algorithm, we conducted extensive experiments on targets with different longitudinal and lateral distances. Analyzing these data, it can be seen that the number of pixels corresponding to 1m laterally is inversely proportional to the longitudinal distance. That is, the longer the longitudinal distance, the larger the lateral offset corresponding to one-pixel width. In this algorithm, the coefficient β is 1800 for pedestrian and bicycle targets and 1500 for car targets. In addition, the experiment also compares the accuracy of monocular depth estimation for different categories, and the results are shown in the following table (Tables 2, 3 and 4).

Table 2. Rangig accuracy of monocular depth estimation for different classes of targets

Category	Longitudinal ranging (m)	Longitudinal ranging error (%)	Lateral ranging (m)	Lateral ranging error (%)
Pedestrian	5–20	3.74	0–3	8.95
Pedestrian	25–35	4.82	0–3	11.26
Bicycle	5–20	4.21	0–3	10.69
Bicycle	20–40	5.06	0–3	12.40
Car	5–40	4.83	0–3	14.73
Car	40–60	6.18	0–3	16.87

As shown in the table above, long-range measuring has a slightly larger error than short-range measuring due to the location prediction error of the target detection network. The calculation of the lateral distance depends on the acquisition of the longitudinal measuring, so it will also produce a certain amount of error accumulation, resulting in the lateral ranging error being larger than that of the longitudinal ranging.

3.2 Experiment on the Target Recognition and Tracking Algorithm

The fusion system designed by us will calculate the RCS of each target through millimeter-wave radar echoes, and then combine it with the classification results of image branches during fusion. Therefore, the recognition accuracy of the system is 4% higher than that of the image-based recognition. This is shown in the following table. Especially for small targets at a distance, the improvement is more obvious.

Table 3. Comparison of recognition accuracy before and after fusion

Category	Image-based accuracy (%)	Fusion-based accuracy (%)
Pedestrian	67.27	70.83
Bicycle	78.88	80.38
Car	62.53	66.63

The following table shows the performance of our fusion system and that of the existing fusion algorithm, although our system performance is a bit weak on the average accuracy. However, it takes less time to process a single frame, which is better than most fusion algorithms based on lidar and camera. It can better meet the real-time requirements of autonomous driving and achieve a balance between accuracy and speed.

Table 4. Comparison of performance indicators of different fusion algorithms

Fusion algorithm	Average accuracy (%)	Single frame processing time (ms)
Our system	76.36	70
3D-FCN	79.30	1000
MV3D	81.10	357
LaserNet	74.77	30

We also experimentally tested the average number of tracking frames for pedestrians, bicycles, and cars when walking laterally. Pedestrians, bicycles, and cars can track 74%, 68%, and 55% of the total frames, respectively. Our tracking algorithm can also track when the targets are briefly occluded from each other.

4 Conclusion

In this paper, we design a millimeter-wave radar and camera fusion system. In the visual part of the system, an improved distance-pixel-based monocular depth estimation algorithm is proposed to solve the problem that the target and the camera form a certain angle, which leads to the decrease of the accuracy of monocular ranging. The millimeter wave radar part uses the point cloud RD features of pedestrians, bicycles, and cars to propose an improved DBSCAN point cloud clustering algorithm, which reduces the amount of calculation and improves the clustering accuracy. Finally, the two results are fused by the extended-Kalman filter to realize target recognition and tracking. We tested it in a real environment, compared with the target recognition method only using vision, the recognition accuracy is improved by 4%, and both real-time and accuracy are taken into account, which better meets the needs of the autonomous driving scene. In the future, we will conduct further research on how to retain more sensor information for pre-fusion and how to add more heterogeneous sensors to fuse together.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (61731006), and was partly supported by the 111 Project No. B17008.

References

1. Sun, D., Shi, H., Ma, H.: Application of multimedia technology in health service management system under the background of big data fusion. In: 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 1134-1137 (2022)
2. Ma, J., Chen, C., Li, C., Huang, J.: Infrared and visible image fusion via gradient transfer and total variation minimization In: Information Fusion, vol. 31, pp. 100-109 (2016)
3. Cheng, H., Zheng, N., Zhang, X., Qin, J., van de Wetering, H.: Interactive road situation analysis for driver assistance and safety warning systems: framework and algorithms. In: IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 1, pp. 157-167 (2007)
4. Charles, R.Q., Su, H., Kaichun, M., Guibas, L.J.: PointNet: deep learning on point sets for 3D classification and segmentation. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 77-85 (2017)
5. Ren, J., et al: Information fusion of digital camera and radar. In: 2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC). IEEE (2019)
6. Aziz, Kheireddine, et al.: Radar-camera fusion for road target classification. In: 2020 IEEE Radar Conference (RadarConf20). IEEE (2020)
7. Ester, M., Kriegel, H.P., Sander, J., et al.: A Density-based algorithm for discovering clusters in large spatial databases with noise. AAAI Press (1996)
8. Stein, G.P., Mano, O., Shashua, A.: Vision-based ACC with a single camera: bounds on range and range rate accuracy. Intelligent Vehicles Symposium. Proceedings. IEEE (2003)

Check for updates

Micro-doppler Feature Extraction of UAVs Based on Synchrosqueezed Transform and Ridge Path Regrouping

Siwei Li(✉), Xu Si, Chi Zhang, and Jing Liang

University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave, Chengdu 611731, China
1376099505@qq.com

Abstract. With the rapid growth of the number of UAV, the research of UAV detection and target recognition is a hot spot at present. When detecting UAVs, in addition to the movement of the UAV's body, the micro-Doppler features of the rotor blades of the UAV are the unique characteristics of detecting UAV. Through the extracted micro-Doppler features, the rotor speed and blade length of the UAV can be calculated, which can be used to recognize the UAV. In this work, the echo models of multiple UAVs are established. In order to extract the micro-Doppler features, the synchrosqueezed transform is used to do time-frequency analysis. There are many components in the micro-Doppler features of multi-UAV. In order to recognize the parameter of each UAV, a ridge path extraction method based on synchrosqueezed transformation is used to separate the components in time-frequency domain. Experimental results show that the proposed method achieves high accuracy in extracting component micro-Doppler features and recognizing UAV information.

Keywords: Micro-doppler extraction · UAV · Synchrosqueezed transform · Signal separation

1 Introduction

With the widespread use of unmanned aerial vehicles (UAV), UAV's number is increasing rapidly, which not only brings convenience to people life, but also brings great security to citizens' privacy and public security. Therefore, the detection, recognition and classification of UAV targets have been focused on by researchers [1–3].

For UAV, its micro-Doppler feature is its typical characteristic. Specifically, UAV 's rotor rotation have sinusoidal modulation on the frequency of radar echo, and V.C.Chen called this modulation phenomenon micro-Doppler effect [4]. From the micro-Doppler feature of echo signal, UAV rotor blade size, rotation speed and other information can be obtained, which plays an important role in UAV target automatic recognition. The common method to extract micro-Doppler features is time-frequency analysis, such as short-time Fourier transform [5],

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 20–27, 2023.
https://doi.org/10.1007/978-981-99-1260-5_3

wavelet transform [6] and so on. Since UAV rotor have at least two blades, the microwaving characteristics of UAV rotor also contain multiple components. In order to analyze each component of the microwaving characteristic, it is necessary to separate the echo signal of UAV in the time-frequency domain. The commonly-use signal separation methods include empirical mode decomposition [7] and intrinsic chirp component decomposition [8]. Moreover, since the multiple components of the UAV's microwaving characteristics are often overlapped in the time-frequency domain, S. Chen et al. proposed the ridge path regrouping [9] to separate the overlapped non-stationary signals. We applied this signal separation method to extract the micro-Doppler features of multiple UAVs and extract the parameter information of each UAV. In order to improve the accuracy, we integrate the synchrosqueezed transform into the ridge path regrouping.

In this work, echo model of multiple UAVs is established, and the micro-Doppler features of UAVs are extracted based on RPRG to recognize their blade information. When extracting instantaneous frequency, we use the synchrosqueezed transformation(SST) to improve the accuracy. Through simulation experiments, the micro-Doppler features is extracted and the blade information is recognized. The rest of the paper is arranged as follows. In Sect. 2, UAV rotor blade echo model is established, and in Sect. 3, the steps of RPRG algorithm based on SST are given. In Sect. 4, the simulation experiment is carried out and the experiment results are analyzed. Finally, the conclusion and future work are given.

2 UAV Blade Echo Signal Model

In this section, the geometric diagram of UAV and radar is given, then the model employs the physical optics model to calculate the RCS of the blade. Finally the mathematical representation of the echo signal is obtained.

2.1 Geometric Model

Considering two UAVs A and B, one of whose rotor has two blades, the geometry of the radar and rotor blades is shown in Fig. 1. (X_0^A, Y_0^A, Z_0^A), (X_0^B, Y_0^B, Z_0^B), (X_1, Y_1, Z_1) represents the location of the rotor center of A, the rotor center of B and the location of the radar respectively. L_A, W_A respectively represents the length of A's blades, the width of A's blades. L_B, W_B respectively represents the length of B's blades, the width of B's blades. The rotation rates of two UAVs are Ω_A and Ω_B respectively, the azimuth angle is ϕ, and the aspect angle is θ.

2.2 Echo Signal Model

The blade is simplified as a rigid, homogeneous, linear rectangular flat plate. Adopt the physical optics method to predict the RCS of the rotor blades, the RCS of the rectangular flat plate is

$$\sigma = \frac{\pi L^2 W^2}{\lambda^2}(cos\theta \frac{sinx_k}{x_k} \frac{siny_k}{y_k})^2 \tag{1}$$

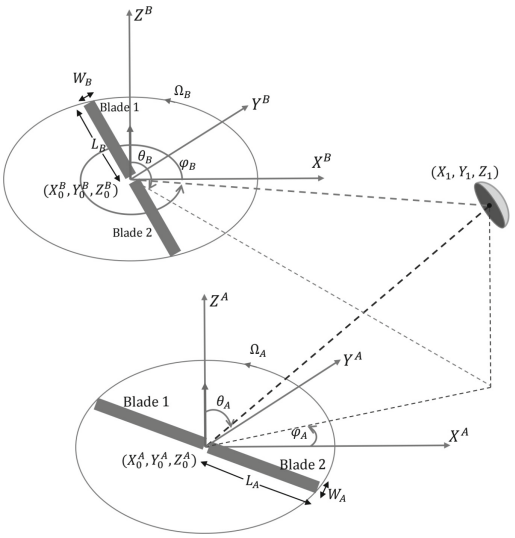


Fig. 1. The geometry of the radar and rotor blades

where (x_k, y_k) is the location of the blades in time t_k. The echo signal in the radar receiver is

$$s(t) = \sum_{k=1}^{n_p} \sum_{n=1}^{N_B} \sqrt{\sigma(t)} rect\{t - k\Delta T - \frac{2R_{n,m}(t)}{c}\} exp\{-j2\pi f_c \frac{2R_{n,m}(t)}{c}\} \quad (2)$$

where N_B is the total numbers of the blades,n_p is the pulse number during the radar observation time interval, $R_{n,m}(t)$ is the distance between radar and the nth blade, c is the velocity of light, f_c is the radar transmitted frequency.

3 Micro-doppler Feature Extraction Method

The time frequency analysis of UAV rotor blade echo is carried out and its micro-Doppler feature is extracted from the time frequency representation(TFR). In order to improve the resolution of time-frequency representation, we use synchrosqueezed transform to extract micro-Doppler features. Because of UAV blade echo signal contains at least two blades echo information, the time-frequency representation can have multiple overlapping signal component. In order to better analyze each component, we separate each component and then extract them.

3.1 Synchrosqueezed Transform

Synchrosqueezed transform (SST) is based on wavelet transform, using synchronous squeezing operator to improve the resolution of the time-frequency

ridge in the spectrum. Sharpen the time-frequency representation by reallocating values on the plane to different time-frequency points. By concentrating the energy of the spectrum, the instantaneous frequency is extracted and reconstructed [10].

$\Phi(b)$ denote the wavelet basis function, the continuous wavelet transform of signal s(t) is

$$W(a,b) = \int s(t)a^{-\frac{1}{2}}\Phi^*(\frac{t-b}{a})dt \tag{3}$$

where a is a scale factor, usually chosen as 32. b is a shift factor. In the wavelet domain, the instantaneous frequency information of position is

$$\omega_s(a,b) = \frac{-j}{W(a,b)}\frac{\partial}{\partial b}W(a,b) \tag{4}$$

For $\omega_s(a,b)$, it is integrated along the scale a direction and classified to the position of $\omega = \omega_s(a,b)$ in the frequency domain. So the synchrosqueezed transform is

$$s_{st}(\omega,b) = \int_{A(b)} W(a,b)a^{-\frac{3}{2}}\delta[W(a,b)-\omega]da \tag{5}$$

$$A(b) = [a, W(a,b) \neq 0] \tag{6}$$

3.2 Extraction Algorithm

The micro-Doppler feature of UAV rotor blades contain multiple overlapping components, especially in the case of multiple UAVs. In order to analyze each component, it is necessary to separate each component from their time-frequency representation. Rigde path regrouping based on intrinsic chirp component decomposition is a method proposed by Shiqian Chen to separate overlapping components in the time-frequency domain. On this basis, we consider to integrate the synchrosqueezed transform into the algorithm. The complete algorithm is as follows.

Step 1. The echo signals $s(t)$ of multiple UAVs are obtained according to the echo model in Sect. 2.

Step 2. Detect the instantaneous frequency(IF) ridges from the time frequency representation.

i. The TFR of $s(t)$ is obtained by SST, and then find the time t_{max} and frequency f_{max} corresponding to the maximum value in the TFR. Traverse the entire timeline from t_max, along left and right respectively, find the frequency unit corresponding to the maximum value at each time. The resulting frequency sequence $[f_1\ f_2\ \ldots]$ is the ridge IF_i of TFR.

ii. Smooth the extracted ridges IF_i by Fourier series fitting. After that, remove it from the original TFR. Extract the corresponding signal of the frequency sequence from $s(t)$ by Fourier model fitting.

iii. Remove the extracted signal $s_e xt(t)$ from the origin signal $s(t)$, return i and Repeat the preceding step I several times and all the ridges have been extracted.

Step 3. Break the extracted ridges and regroup them according to the slope rate of the intersections.

i. Take any two curves, find their intersection. Merge the intersection intervals whose distance is less than threshold thf.
ii. Break the ridges in the intersection intervals and then reconnect them according to their slopes. Finally, can obtain all the regrouped frequency sequences.

Step 4. Obtain each component corresponding to the regrouped ridge by Fourier model fitting.

4 Experiment and Analysis on UAV Blade Micro-doppler

4.1 Parameter Setting for the Simulation Experiment

The echo model of two UAVs A and B is simulated, and the micro-Doppler features are extracted by the algorithm in Sect. 3. A and B are hovering. The parameters of A and B are as follows: the rotation speed of A and B is $\Omega_A = 30r/s$, $\Omega_B = 40r/s$ respectively. The length and width of rotor blade of A are $L_A = 0.2$ m, $W_A = 0.02$ m. The length and width of each rotor blade of B are $L_B = 0.13$ m, $W_B = 0.02$ m. Their echo signal in time domain is shown in figure.

4.2 Experiment Result and Analysis

Using the echo model in Sect. 2, the echo signal of the UAV A and B shows in Fig. 2.

From the Fig. 2, it can be clearly seen that the frequency of echo is sinusoidal modulated from the micro-Doppler feature, which is composed of multiple components. In order to further analyze each component and obtain all the information carried by each component, use the extraction algorithm in Sect. 3 to separate each component and extract them.

Figure 3 shows the intermediate results using the extraction algorithm. (a) is the extracted ridge. It can be seen that this step can only extract the ridge from the time-frequency representation, which is confusing. (b) is the result of regroup the ridges in (a) by using RPRG algorithm. This step can get the ridges of the components, but it's rough. (c) is the result of separating each component from the original signal by Fourier model fitting based on the regrouped ridge in (b). (c) shows that all the components are completely separated out.

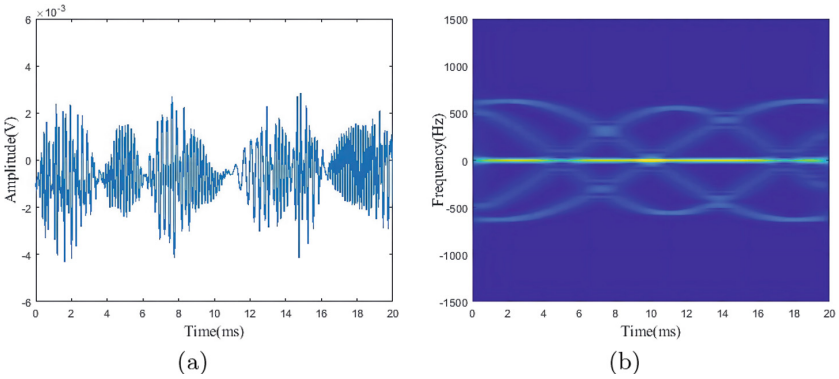


Fig. 2. Echo signal of UAV A and B : (a)Signal in time domain (b)The time-frequency representation

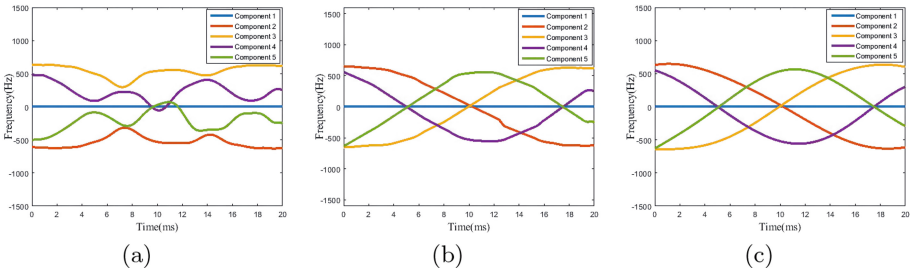
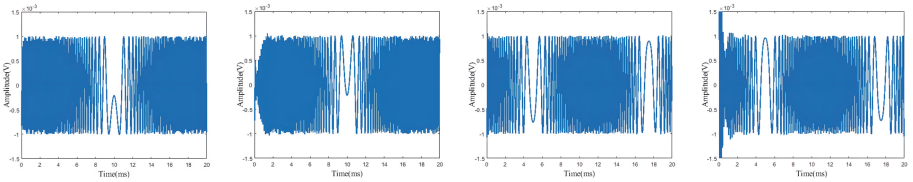


Fig. 3. Intermediate results using the extraction algorithm: (a)The extracted ridge (b)The regrouped ridge (c)The final separation result

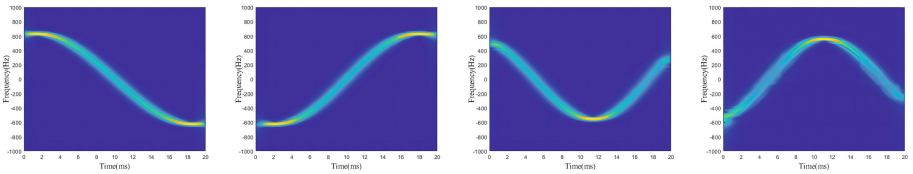
The extracted signals of each component in time domain is shown in Figs. 4 and 5 represents the time frequency representation of the each component, which is clear. From the extracted components, the blade speed and blade size of UAV can be calculated.

From the extracted components, it can be obtained that the rotation period of A's rotor is $T_A = 0.0248$ s, the rotational speed is $\Omega_A^* = 1/T_A = 40.32$ r/s, the maximum Doppler frequency shift caused by the blade tip is $\Delta f_A = 565.15$ Hz, the velocity of the blade tip is $v_A = \Delta f_A \times \lambda = 33.91$ m/s, and the blade length is $L_A^* = v_A/(2\pi\Omega_A^*) = 0.1318$ m. Similarly, it can be obtained that the rotation period of the B rotor is $T_B = 0.035$ s, the rotational speed is $\Omega_B^* = 28.58$ r/s, the maximum Doppler frequency shift caused by the blade tip is $v_B = 38.86$ m/s, the velocity of the blade tip is, and the blade length is $L_A^* = 0.2165$ m. It can be calculated that the error of A speed is 0.8%, and the error of blade length is 1.38%. The error of B speed is 4.77%, and the error of blade length is 8.25%.



(a) (b) (c) (d)

Fig. 4. The obtained individual components in time domain: (a)-(d) represent four blades TFR of two UAVs



(a) (b) (c) (d)

Fig. 5. The obtained individual components in time-frequency domian : (a)-(d) represent four blades TFR of two UAVs

5 Conclusion

In this work, we realized the extraction of UAV micro-Doppler features and rotor blade information recognition. Firstly, the rotor blade echo of two UAVs is modeled. Then a ridge path regrouping method based on synchrosqueezed transform is used to extract the micro-Doppler features of UAV. Finally, the micro-Doppler feature extraction experiment of two UAVs is simulated, which proved that the method has high accuracy in micro-Doppler features extraction and separation of multiple UAVs, so as to accurately recognize the parameters of multiple UAVs. The recognition error is as follows: the error of two UAVs' rotation speed are 0.8% and 4.77%, respectively. And the error of two UAVs' blade length are 1.38% and 8.25%, respectively. In future work, we will improve our method so as to extract the micro-Doppler features of each component in the more complex scenarios.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (61731006), and was partly supported by the 111 Project No. B17008.

References

1. Dudush, A.S., Tyutyunnik, V.A., Reznichenko, A.A., Hohoniants, S.U.: State of the art and problems of defeat of low, slow and small unmanned aerial vehicles. Mod. Inf. Technol. Sphere Secur. Defence **1**(31), 121–131 (2018)
2. Karimov, A.K.: The goals and tasks solved by unmanned aviation complexes of a new generation. In: Proceedings of the MAI (in Russian), no. 47 (2011)

3. United States Army: Counter-Unmanned Aircraft System (C-UAS) Strategy Extract (2016)
4. Chen, V.C., Li, F., Ho, S.-S., Wechsler, H.: Micro-Doppler effect in radar: phenomenon, model, and simulation study. IEEE Trans. Aerosp. Electron. Syst. **42**(1), 2–21 (2006)
5. Li, L., Cai, H., Han, H., Jiang, Q., Ji, H.: Adaptive short-time Fourier transform and synchrosqueezing transform for non-stationary signal separation. Sig. Process. **166**, 107231 (2020). ISSN 0165-1684. https://doi.org/10.1016/j.sigpro.2019.07.024
6. Gilles, J.: Empirical wavelet transform. IEEE Trans. Signal Process. **61**(16), 3999–4010 (2013)
7. Huang, N.E., et al.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. Roy. Soc. London Ser. A, Math., Phys. Eng. Sci. **454**(1971), 903–995 (1998)
8. Chen, S., Peng, Z., Yang, Y., Dong, X., Zhang, W.: Intrinsic chirp component decomposition by using Fourier Series representation. Signal Process. **137**, 319–327 (2017). ISSN 0165–1684. https://doi.org/10.1016/j.sigpro.2017.01.027
9. Chen, S., Dong, X., Xing, G., Peng, Z., Zhang, W., Meng, G.: Separation of overlapped non-stationary signals by ridge path regrouping and intrinsic chirp component decomposition. IEEE Sensors J. **17**(18), 5994–6005 (2017). https://doi.org/10.1109/JSEN.2017.2737467
10. Daubechies, I., Lu, J., Wu, H.-T.: Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. Appl. Comput. Harmonic Anal. **30**(2), 243–261 (2011). ISSN 1063–5203. https://doi.org/10.1016/j.acha.2010.08.002

Radial Basis Function Neural Network Based Fast Directional Modulation Design

Mingjie Zhou[1], Bo Zhang[1(✉)], Baoju Zhang[1(✉)], Wei Wang[2], Bo Wang[2], Lin Han[3], Xiaonan Zhao[1], and Cheng Wang[1]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
b.zhangintj@tjnu.edu.cn, wdxyzbj@163.com
[2] Sitonholy (Tianjin) Technology Co., Ltd., Tianjin, China
[3] Tianjin Vocational Institute, Tianjin, China

Abstract. For directional modulation (DM) based on antenna arrays, the associated weight coefficients are normally calculated by some traditional optimisation methods. In this paper, a radial basis function (RBF) based neural network is proposed for fast directional modulation design, which is trained by data corresponding to sets of transmission direction and interference direction angles with the corresponding weight coefficients. The proposed method takes advantage of the strong nonlinear function approximation capability of the RBF neural network, and simulation results are provided to show the effectiveness of the proposed design.

Keywords: Directional modulation · Uniform linear antenna array · Radial basis function (RBF) · Neural networks

1 Introduction

Directional modulation (DM) as a physical layer security technology can transmit the digitally modulated signal to desired directions and distort the constellation diagram of the signal in other directions with the aid of optimised weight coefficients of the carrier signal for each array element [1,2]. In [3], a directional modulation technique based on phased arrays was proposed, where the desired amplitude and phase of each symbol can be generated in a given direction in digital modulation by varying the phase shift value of each array element. In [4,5], a practical secure communication system for DM was built and the bit error rate of the received signal was analysed. A two array element based directional modulation method was mentioned in [6] by adding a two-bit phase shifter to each antenna, so that different beams were sent according to the corresponding phase shift values. In [7], the directional modulation system was redesigned based on the sparse array theory using the compressive sensing method to select fewer antenna array elements from the reserved antenna grid compared to the uniform antenna array.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 28–35, 2023.
https://doi.org/10.1007/978-981-99-1260-5_4

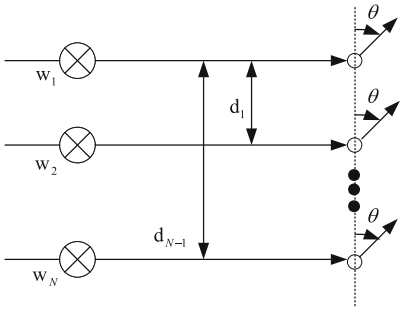


Fig. 1. A narrowband transmit beamforming structure.

However, the above designs for weigh coefficients optimisation are normally based on some traditional optimisation methods and it takes a long time to obtain a satisfactory design result. To solve the problem, in this paper, we introduce radial basis function (RBF) based neural network [8,9] to approximate the value of weight coefficients. RBF neural network is a single hidden layer feed forward network with good performance, featuring fast convergence, simple structure and small operation.

The remaining part of this paper is structured as follows. A review of DM technique based on a traditional optimisation method is given in Sect. 2. RBF neural network based DM design is presented in Sect. 3. Design examples are provided in Sect. 4, followed by conclusions in Sect. 5.

2 Review of Directional Modulation Based on a Traditional Optimisation Method

2.1 Principle of Narrowband Beamforming

A narrowband linear array for transmit beamforming is shown in Fig. 1, consisting of N omni-directional antennas with the array element spacing d_n for $n = 1, 2, \ldots, N-1$. Then the corresponding weight vector is given by

$$\mathbf{w} = [w_1, w_2, \cdots, w_N]^T. \tag{1}$$

The steering vector of the array is a function of angular frequency ω and transmission angle θ, given by

$$\mathbf{s}(\omega, \theta) = [1, e^{j\omega d_1 \cos\theta/c}, \ldots, e^{j\omega d_{N-1} \cos\theta/c}]^T, \tag{2}$$

where $\{\cdot\}^T$ is the transpose operation, c is the speed of propagation and the transmission angle $\theta \in [0°, 180°]$. Based on the given weight vector and steering vector, the beam response of the array can be represented by

$$\mathbf{p}(\theta) = \mathbf{w}^H \mathbf{s}(\omega, \theta), \tag{3}$$

where $\{\cdot\}^H$ represents the Hermitian transpose.

2.2 DM Design Based on a Traditional Optimisation Method for a Given Array Geometry

For M-ary signaling, such as multiple phase shift keying (MPSK), there are M sets of desired array responses $p_m(\theta)$, with a corresponding weight vector $\mathbf{w}_m = [w_{m,1}, w_{m,2}, \cdots, w_{m,N}]^T$, $m = 1, 2, \cdots, M$. According to the value of θ, each desired response $p_m(\theta)$ can be classified into either mainlobe responses $\mathbf{p}_{m,ML}$ or sidelobe responses $\mathbf{p}_{m,SL}$. Here, we sample r points in the mainlobe directions and $R - r$ points in the sidelobe region. Then, we have

$$\begin{aligned} \mathbf{p}_{m,ML} &= [p_m(\theta_1), p_m(\theta_2), \cdots, p_m(\theta_r)], \\ \mathbf{p}_{m,SL} &= [p_m(\theta_{r+1}), p_m(\theta_{r+2}), \cdots, p_m(\theta_R)]. \end{aligned} \tag{4}$$

Similarly, steering vectors in the mainlobe and sidelobe directions can be represented by $\mathbf{S}_{ML}$ with a size of $N \times r$ and $\mathbf{S}_{SL}$ with a size of $N \times (R - r)$, respectively. Then, for the m-th constellation point, its corresponding weight coefficients can be found by

$$\begin{aligned} \min \quad & \left\| \mathbf{p}_{m,SL} - \mathbf{w}_m^H \mathbf{S}_{SL} \right\|_2 \\ \text{subject to} \quad & \mathbf{w}_m^H \mathbf{S}_{ML} = \mathbf{p}_{m,ML}, \end{aligned} \tag{5}$$

where $\|\cdot\|_2$ denotes the l_2 norm. The objective function and constraints in (5) ensure the minimum difference between the desired and designed responses in the sidelobe directions, and the same value for desired and designed in the mainlobe directions.

3 Directional Modulation Design Based on RBF Neural Network

3.1 The RBF Neural Network Model

As shown in formulation (5), the optimisation of weight vector for DM is a nonlinear function with respect to the desired transmission directions. Therefore, we propose a RBF neural network based DM design. The RBF neural network structure is shown in Fig. 2, which is a three-layer feedforward network consisting of an input layer, a hidden layer and an output layer [10]. The activation function of the hidden layer uses a Gaussian function, given by

$$\varphi_i(x) = e^{-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{2\sigma_i^2}}, \quad 1 \le i \le I, \tag{6}$$

where $\mathbf{x} = [x_1, x_2, \cdots, x_K]$ is the input vector of the network, $\mathbf{c}_i = [c_{i1}, c_{i2}, \cdots, c_{iK}]$ is the center of the Gaussian function of the i-th node of the hidden layer ($i = 1, 2, \ldots, I$), and σ_i represents the variance of the Gaussian function of the i-th neuron node [11], given by

$$\sigma_i = \frac{d_{max}}{\sqrt{2i}}. \tag{7}$$

Here d_{max} is the maximum Euclidean distance between the selected central vectors and i is the number of hidden layer neurons.

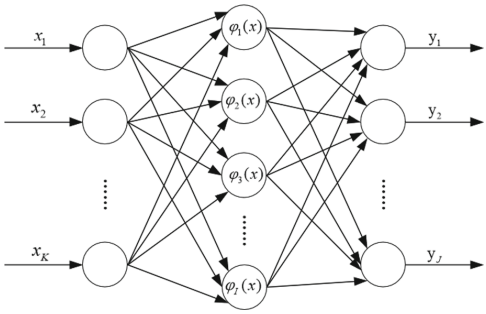


Fig. 2. RBF learning process in DM case.

3.2 Orthogonal Least Squares Training Algorithm for RBF Networks

The orthogonal least squares method has already been shown in applications to be stable, computationally low, and able to save a lot of storage space [12]. In order to eliminate the influence of magnitudes between indexes, data standardization is needed to deal with the comparability issue between data indexes. The sample normalization formula is given by

$$\hat{x}_k = \frac{x_k - min(\mathbf{x}_k)}{max(\mathbf{x}_k) - min(\mathbf{x}_k)} \tag{8}$$

where $\hat{x}_k$ is the normalization result, $max(\mathbf{x}_k)$ and $min(\mathbf{x}_k)$ are the maximum and minimum values of the parameters in the k-th index. Cluster the input samples and then select the centroid $\mathbf{c}_i$ for the i-th neuron of the hidden layer by k-means clustering algorithm [13].

According to Eq. (7), the Gaussian function variance σ_i can be calculated. The hidden layer output φ_i can be calculated by Eq. (6). We use orthogonal linear least squares to determine the weights of the linear layer neurons. The implementation of RBF network mapping with a regression model is given by

$$z_j = \sum_{i=1}^{I} \varphi_i(x)\tilde{w}_{ji} + e_j \tag{9}$$

where $\tilde{w}_{ji}$ $(i = 1, 2, \cdots, I, j = 1, 2, \cdots, J)$ is the weight from i-th node of the hidden layer to the j-th node of the output layer, z_j and e_j are the expected output and the error between the expected output and the model output. Then, Eq. (9) is transformed to the following matrix form

$$\mathbf{z} = \boldsymbol{\Phi}\tilde{\mathbf{w}} + \mathbf{e} \tag{10}$$

where $\mathbf{z} = [z_1, z_2, \cdots, z_J]$, $\Phi = [\varphi_1(x), \varphi_2(x), \cdots, \varphi_I(x)]$, $\tilde{\mathbf{w}} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{J1} \\ w_{12} & w_{22} & \cdots & w_{J2} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1I} & w_{2I} & \cdots & w_{JI} \end{bmatrix}$ and $\mathbf{e} = [e_1, e_2, \cdots, e_J]$. The problem of least squares can be defined as finding the parameter $\tilde{\mathbf{w}}$ that minimizes $\|\mathbf{z} - \Phi\tilde{\mathbf{w}}\|$.

The matrix Φ is decomposed by Gram-Schmidt orthogonalization method as

$$\Phi = \mathbf{QR} = \left[\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_I\right] \begin{bmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1I} \\ 0 & 1 & a_{23} & \cdots & a_{2I} \\ \vdots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}, \tag{11}$$

where $\mathbf{Q}$ is an orthogonal matrix with mutually orthogonal columns, $\mathbf{R}$ is the upper triangular matrix with the main diagonal being 1. Here, $\mathbf{Q}$ satisfies

$$\mathbf{Q}^T\mathbf{Q} = \mathbf{H}, \tag{12}$$

where $\mathbf{H}$ is diagonal matrix. Substitute Eq. (11) into Eq. (10) and let $\mathbf{R}\tilde{\mathbf{w}} = \mathbf{g}$, and then we have

$$\mathbf{z} = \mathbf{Qg} + \mathbf{e}, \tag{13}$$

where the output vector $\mathbf{z}$ is expected to be expressed as a linear combination of mutually orthogonal columns of a matrix $\mathbf{Q}$. Then, the least squares problem can be solved by

$$\min_{\mathbf{g}} \psi(\mathbf{g}) = \|\mathbf{z} - \mathbf{Qg}\|_2^2 \tag{14}$$

we have

$$\begin{aligned} \psi(\mathbf{g}) &= \mathbf{e}^T\mathbf{e} \\ &= (\mathbf{z} - \mathbf{Qg})^T(\mathbf{z} - \mathbf{Qg}) \\ &= \mathbf{z}^T\mathbf{z} - 2\mathbf{g}^T\mathbf{Q}^T\mathbf{z} + \mathbf{g}^T\mathbf{Q}^T\mathbf{Qg} \end{aligned} \tag{15}$$

The gradient of Eq. (15) is

Table 1. Algorithm time comparison

Experiment	Algorithm	
	DM-cvx	DM-rbf
Experiment 1	27.17 mins	0.03 mins
Experiment 2	29.21 mins	0.05 mins
Experiment 3	32.32 mins	0.02 mins

$$\frac{\partial \psi}{\partial \mathbf{g}} = 2\mathbf{Q}^T\mathbf{Qg} - 2\mathbf{Q}^T\mathbf{z} \tag{16}$$

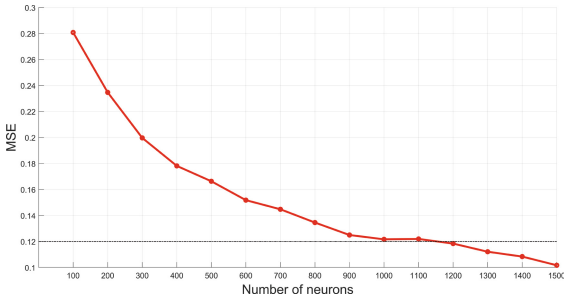


Fig. 3. Relationship between the number of hidden layer neurons and MSE.

With $\frac{\partial \psi}{\partial \mathbf{g}} = 0$, the least squares solution of the coordinate vector is

$$\hat{\mathbf{g}} = (\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{Q}^T\mathbf{z} = \mathbf{H}^{-1}\mathbf{Q}^T\mathbf{z} \tag{17}$$

Through the above process, we know that the orthogonal least squares can reduce the nonlinear model solving problem to an expression finding the maximum value problem. The orthogonalization process is the characteristic of the method, which can reduce the computational effort, enhance the stability and make the whole model more generalizable.

4 Simulation Results

In the design example, we use the orthogonal least squares (OLS) based radial basis network to find the optimised weights of the output layer. Here, we assume the number of antenna array elements is 8 and there is a desired direction θ_s and an interference direction θ_I ranging from 0° to 180° as the input of the RBF neural network at the same time. The array element spacing is 0.5λ (λ is the signal wavelength).

Since the performance of the OLS-based RBF neural network and the running time are affected by the number of neurons in the hidden layer, we did the following experiments to solve this tradeoff problem. As shown in Fig. 3, we can see the RMSE value under a given number of neurons in the validation set. Here, we seek the appropriate number of neurons by setting the RMSE threshold to 0.12. By comparison, we select 1200 neurons in the design, where the RMSE drops below 0.12. Moreover, the sum of the running times of the four phases on the two different methods are compared, as shown in Table 1. We compare the time required by the two algorithms through three experiments, where the running time for 1656 samples in validation set using the traditional method are all more than 27 mins, while the running time for the RBF-based DM design is around 0.03 mins, much lower than the the running time for the counterpart, which demonstrates a better real-time performance.

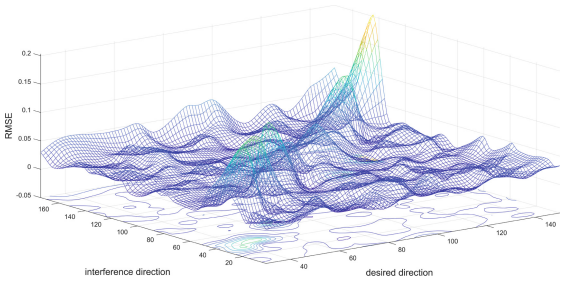
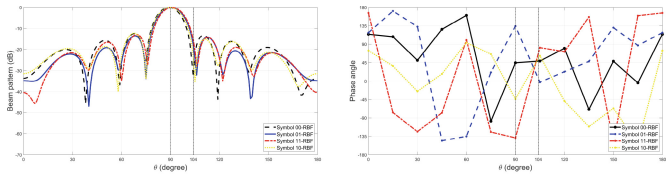


Fig. 4. The relationship between expectation angle, interference angle and error.



(a) (b)

Fig. 5. Beam patterns and phase patterns based on RBF neural network in the broadside design:(a)beam patterns,(b)phase patterns

A three-dimensional error plot for symbol 00 is shown in Fig. 4, where the x, y, z axes are the desired transmission direction, interference direction and the root mean square error (RMSE) between the actual weights and the predicted weights, respectively. Here, we can see that the prediction results show a small error and a good fitting effect for the RBF neural network. The error plot for symbols 01, 11 and 10 are not given as the corresponding plots are similar to that of symbol 00.

Moreover, with the optimised RBF network, the magnitude and phase responses based on one sample are drawn. Here, we assume the desired transmission direction is 90° and the interference direction is 104°. As shown in Figs. 5(a) and 5(b), we can see that the mainbeams for all symbols point to the desired direction 90° with a very low magnitude at 104°. The phase at the desired direction follows the standard QPSK modulation, while phases are scrambled in other directions.

5 Conclusions

In this paper, a radial basis function (RBF) neural network based directional modulation design is proposed, which realizes the mapping of any angle of the transmission direction to the weight vector of the given antenna array, according to the feature of nonlinear mapping from the input space to the output space of

the RBF network. Simulation results show that the method can quickly approximate the weight vector required for any transmission direction, and has the advantages of good approximation performance and fast assignment compared with the traditional method.

Acknowledgments. The work was partially supported by the National Natural Science Foundation of China (62101383) and Science and technology project of Hedong District (TJHD-JBGS-2022-08).

References

1. Babakhani, A., Rutledge, D.B., Hajimiri, A.: Transmitter architectures based on near-field direct antenna modulation. IEEE J. Solid-State Circuits **43**(12), 2674–2692 (2008)
2. Babakhani, A., Rutledge, D.B., Hajimiri, A.: Near-field direct antenna modulation. IEEE Microwave Mag. **10**(1), 36–46 (2009)
3. Daly, M.P., Bernhard, J.T.: Directional modulation technique for phased arrays. IEEE Trans. Antennas Propag. **57**(9), 2633–2640 (2009)
4. Daly, M.P., Daly, E.L., Bernhard, J.T.: Demonstration of directional modulation using a phased array. IEEE Trans. Antennas Propag. **58**(5), 1545–1550 (2010)
5. Daly, M.P., Bernhard, J.T.: Directional modulation and coding in arrays. In: 2011 IEEE international symposium on antennas and propagation (APSURSI), pp. 1984-1987. IEEE (2011)
6. HongZhe, S., Alan, T.: Direction dependent antenna modulation using a two element array. In: Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP), pp. 812-815. IEEE (2011)
7. Zhang, B., Liu, W., Gou, X.: Sparse antenna array design for directional modulation. In: 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), pp. 1-5. IEEE (2016)
8. Roy, A., Govil, S., Miranda, R.: A neural-network learning theory and a polynomial time RBF algorithm. IEEE Trans. Neural Networks **8**(6), 1301–1313 (1997)
9. Vt, S.E., Shin, Y.C.: Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems. IEEE Trans. Neural Networks **5**(4), 594–603 (1994)
10. Nabney, I.T.: Efficient training of RBF networks for classification. Int. J. Neural Syst. **14**(03), 201–208 (2004)
11. Suykens, J.A.K., Vandewalle, J.: Least squares support vector machine classifiers. Neural Process. Lett. **9**(3), 293–300 (1999)
12. Soussen, C., Gribonval, R., Idier, J., et al.: Joint k-step analysis of orthogonal matching pursuit and orthogonal least squares. IEEE Trans. Inf. Theory **59**(5), 3158–3174 (2013)
13. Likas, A., Vlassis, N., Verbeek, J.J.: The global k-means clustering algorithm. Pattern Recogn. **36**(2), 451–461 (2003)

A Precoding Algorithm Based on Weighted MMSE for MmWave MU-MIMO System with Discrete Lens Array

Dehao Wang[1,2(✉)], Le Yang[2], Xupeng Li[2], Qian Wang[2], Feng Zhang[2], and Dingyuan Shi[2]

[1] Communication Research Center, Harbin Institute of Technology, Harbin, China
2726747345@qq.com

[2] Research and Development Department China Academy of Launch Vehicle Technology, Beijing, China

Abstract. Existing researches for Beamspace massive MIMO system mainly concentrate on simple precoders, which severely limits the system performance. Inspired by the relationship between received minimum mean square error (MMSE) and spectural efficiency (SE) for single-antenna user, we formulate the problem to maximize system SE as a minimization problem about sum of received mean square error, whose objective function is concave with respect to each variable when all other variables are fixed. Based on this property, we propose a precoding algorithm based on weighted MMSE, which can find a satisfactory solution by blocked coordinate descent method with acceptable complexity. Simulation results demonstrate the performance of proposed precoding algorithm.

Keywords: Precoding · Beamspace MIMO · MMSE

1 Introduction

Hybrid analog/digital beamforming technology for mmWave massive Multiple Input Multiple Output (MIMO) allows trade-off between the number of RF chains and system performance in 5G network [1]. To simplify the structure of analog beamformer, the concept of Beamspace MIMO based on discrete lens array (DLA) is proposed. Typical DLA includes an electromagnetic lens and a matching antenna array, whose elements are located in the focal region of the lens [2]. Beamspace MIMO can be regarded as a special hybrid architecture composed of baseband precoding and analog beam selection. Beam selection is achieved by switching network, such as Maximization of the SINR selection (MSS), maximization of the capacity selection (MCS) [3], etc. However, current researches mainly concentrate on simple precoder, such as zero force (ZF), minimum mean square error (MMSE), wiener filter (WF), matched filter (MF) and other mechanism [4–6], which brings severe limitation to the perfomance of spectural efficiency.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 36–43, 2023.
https://doi.org/10.1007/978-981-99-1260-5_5

The rest of this paper is organized as follows: Sect. 2 represent the system model and problem formulation. The proposed precoding alogrithm based on weigthed MMSE is described in Sect. 3. Sections 4 and 5 show the simulation results and conculsion of this paper.

2 System Model and Problem Formulation

2.1 Beamspace MU-MIMO System Model

As shown in Fig. 1, we consider a mmWave Beamspace MU-MIMO system where the base station (BS) employs N-dimension DLA and N_b beam to serve K single-antenna users.

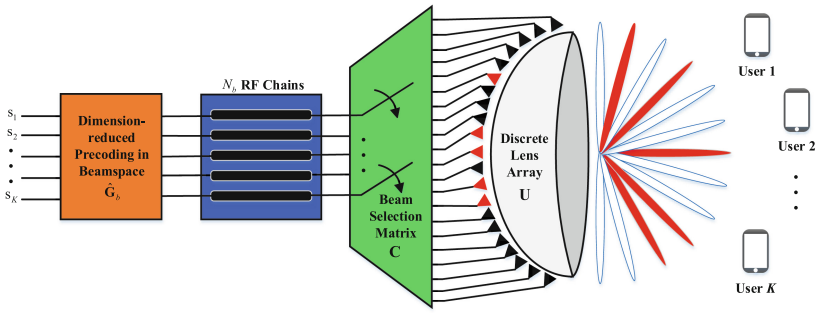


Fig. 1. Beamspace MU-MIMO system with DLA

Then, the received signal vector $\mathbf{y}$ in the downlink is

$$\mathbf{y} = \mathbf{H}_b^H \mathbf{C}\, \hat{\mathbf{G}}_b \mathbf{s} + \mathbf{n} \tag{1}$$

where $\mathbf{s} \in \mathbb{C}^{K\times 1} = [s_1, s_2, \ldots, s_K]$, $\mathbf{y} \in \mathbb{C}^{K\times 1} = \left[y_1, y_2, \ldots, y_K\right]$ are transmitted and receivied signal vectors respectively, and $\mathbf{s}$ satisfies $\mathbb{E}(\mathbf{s}\mathbf{s}^H) = \mathbf{I}_K$. $\mathbf{n} \in \mathbb{C}^{K\times 1} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_K)$ is the additive White Gaussian noise vector, where σ^2 is the noise power. $\mathbf{H}_b^H \in \mathbb{C}^{K\times N}$ is the channel matrix, $\mathbf{C} \in \mathbb{R}^{N\times N_b}$ is the beam selection matrix, $\hat{\mathbf{G}}_b \in \mathbb{C}^{N_b\times K} = \left[\hat{\mathbf{g}}_{b,1}, \hat{\mathbf{g}}_{b,2}, \ldots, \hat{\mathbf{g}}_{b,K}\right]$ is the precoding matrix in beamspace, $\hat{\mathbf{g}}_{b,k} \in \mathbb{C}^{N_{RF}\times 1}$ is precoding vector for the k-th user. By employing DLA, the channel matrix $\mathbf{H}_b \in \mathbb{C}^{N\times K}$ in beamspace can be obtained by of physical domian channel.

$$\mathbf{H}_b = \left[\mathbf{h}_{b,1}, \mathbf{h}_{b,2}, \ldots, \mathbf{h}_{b,K}\right] = \mathbf{U}^H[\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K] \tag{2}$$

where $\mathbf{U}$ denotes the analog beamforming matrix achieved by DLA, $\mathbf{U}$ is a $N \times N$ unitary Discrete Fourier transformation (DFT) matrix for carefully designed DLA [2]. To grip channel characteristic in mmWave propagation environments, we adopt the Saleh-Valenzuela (SV) channel model [3]-[6], channel vector $\mathbf{h}_k$ in physical domain is the same as that in [6].

2.2 Problem Formulation

In Beamspace MU-MIMO system described in Fig. 1, the k-th user's received siganl is $y_k = \mathbf{h}_{b,k}^H \mathbf{C}\,\hat{\mathbf{G}}_b I_k s_k + \sum_{i\neq k}^{K} \mathbf{h}_{b,k}^H \mathbf{C}\,\hat{\mathbf{G}}_b I_i s_i$, where n_k is the gaussian noise, I_k denotes the k-th column vector of unit matirx $\mathbf{I}_K$. . The spectrual efficiency (SE) of MU-MIMO system can be defined as $R = \sum_{k=1}^{K} R_k$, the SE of the k-th user is

$$R_k = \log(1 + SINR_k) = \log\left(1 + \frac{\left|\mathbf{h}_{b,k}^H \mathbf{C}\hat{\mathbf{G}}_b I_k\right|^2}{\sum\limits_{i\neq k}^{K}\left|\mathbf{h}_{b,k}^H \mathbf{C}\,\hat{\mathbf{G}}_b I_i\right|^2 + \sigma^2}\right) \tag{3}$$

In this paper, we assume that the beam selection matrix $\mathbf{C}$ is obtained in advance. When the transmitt power of BS is P, the SE maximization problem can be mathematically formulated as

$$\begin{aligned} &\mathrm{P1}: \max_{\hat{\mathbf{G}}_b} \sum_{k=1}^{K} R_k \\ &s.t.\ tr\left(\hat{\mathbf{G}}_b \hat{\mathbf{G}}_b^H\right) = P \end{aligned} \tag{4}$$

3 Proposed Precoding Algorithm Based on Weighted MMSE

3.1 The Relationship Between Received MMSE and SE

The estimated signal of k-th user is $\hat{s}_k = u_k y_k$, where $u_k \in \mathbb{C}$ is the receivers of K users. The mean square error (MSE) between transmitted signal s_k and $\hat{s}_k$ is

$$e_k = \mathbb{E}\left(\left|\hat{s}_k - s_k\right|^2\right) = |u_k|^2 Q_k - \left(I_k^T \hat{\mathbf{G}}_b^H \mathbf{C}^T \mathbf{h}_{b,k} u_k^* + u_k \mathbf{h}_{b,k}^H \mathbf{C}\,\hat{\mathbf{G}}_b I_k\right) + 1 \tag{5}$$

where $Q_k = \sum_{i=1}^{K}\left|\mathbf{h}_{b,k}^H \mathbf{C}\,\hat{\mathbf{G}}_b I_i\right|^2 + \sigma^2$. When receiver u_k satisfies

$$u_k^{opt} = Q_k^{-1} I_k^T \hat{\mathbf{G}}_b^H \mathbf{C}^T \mathbf{h}_{b,k} \tag{6}$$

The k-th user's minimum mean square error (MMSE) e_k^{MMSE} is

$$e_k^{MMSE} = 1 - Q_k^{-1}\left|\mathbf{h}_{b,k}^H \mathbf{C}\,\hat{\mathbf{G}}_b I_k\right|^2 \tag{7}$$

Based on (3) and (7), we get the relationship between R_k and e_k^{MMSE} as follows

$$\begin{aligned} \log\left(\left(e_k^{MMSE}\right)^{-1}\right) &= \log\left(\left(1 - Q_k^{-1}\left|\mathbf{h}_{b,k}^H \mathbf{C}\,\hat{\mathbf{G}}_b I_k\right|^2\right)^{-1}\right) \\ &= \log(1 + SINR_k) = R_k \end{aligned} \tag{8}$$

3.2 Problem Reformulation

Noting that problem P1 is unconvex, we introduce two groups of auxiliary variables, and obtain problem P2 of weighted MSE.

$$\begin{aligned} \text{P2}: \min_{\hat{\mathbf{G}}_b,\{u_k\},\{w_k\}} & \sum_{k=1}^{K} w_k e_k - \log(w_k) \\ s.t.\ & tr\left(\hat{\mathbf{G}}_b \hat{\mathbf{G}}_b^H\right) = P \end{aligned} \tag{9}$$

where $\{w_k \in \mathbb{R} | w_k > 0, k = 1, 2, \ldots, K\}$ is the group of weighted factors for each user's MSE e_k. Then, we can get the folloing theorem.

Theorem 1: Problem P2 of weighted MSE is equivalent to problem P1, which indicates that the global optimal solution of $\hat{\mathbf{G}}_b$ for the two problems are identical.

Proof: Firstly optimze variables $\{u_k\}$ and $\{w_k\}$ with fixed variable $\hat{\mathbf{G}}_b$, problem P2 can be simplified as

$$\text{P3}: \min_{\{u_k\},\{w_k\}} \sum_{k=1}^{K} w_k e_k - \log(w_k) \tag{10}$$

We can see that problem P3 is unconstrained, and $\{u_k\}$ and $\{w_k\}$ are decoupled. As descirbed before, when fixing $\hat{\mathbf{G}}_b$ and $\{w_k\}$, the optimal solution of $\{u_k\}$ is $\left\{u_k^{opt}\right\}$ in (6), which is the well-kown MMSE receiver, the MMSE of k-th user is e_k^{MMSE} in (7). Then fixing $\hat{\mathbf{G}}_b$ and $\left\{u_k^{opt}\right\}$ in (10), whose objective function is concave with respect to $\{w_k\}$, the optimal solution of w_k is

$$w_k^{opt} = (\ln 2)^{-1} \left(e_k^{MMSE}\right)^{-1} \tag{11}$$

Substituting $\left\{u_k^{opt}\right\}$ and $\left\{w_k^{opt}\right\}$ into objective function of problem P2, , based on relationship in (8), problem P2 can be tanrsformed into

$$\begin{aligned} \text{P4}: \min_{\hat{\mathbf{G}}_b} & \sum_{k=1}^{K} -R_k \\ s.t.\ & tr\left(\hat{\mathbf{G}}_b \hat{\mathbf{G}}_b^H\right) = P \end{aligned} \tag{12}$$

We find that the global optimal solutions of precoding matrix $\hat{\mathbf{G}}_b$ for problem P3 and P2 are identical, which proves the correctness of Theorem 1.

3.3 Proposed Precoding Alogrithm

When fixing the variables $\{u_k\}$ and $\{w_k\}$, problem P2 would be simplifed as a quadratic programming problem in (13), which is concave with respect to $\hat{\mathbf{G}}_b$.

$$\begin{aligned} \text{P5}: \min_{\hat{\mathbf{G}}_b} & \sum_{k=1}^{K} w_k e_k \\ s.t.\ & tr\left(\hat{\mathbf{G}}_b \hat{\mathbf{G}}_b^H\right) = P \end{aligned} \tag{13}$$

Typical Lagrange Multiplier Method can be employed to settle problem P5. By introducing a Lagrange Multiplier $\lambda > 0$ and calculating the derivative of Lagrange Multiplier function with respect to $\hat{\mathbf{G}}_b$, the closed-form solution is

$$\hat{\mathbf{G}}_b = \left(\mathbf{A} + \lambda \mathbf{I}_{N_b}\right)^{-1} \times \mathbf{B} \tag{14}$$

where $\mathbf{A} = \sum_{k=1}^{K} w_k |u_k|^2 \mathbf{C}^T \mathbf{h}_{b,k} \mathbf{h}_{b,k}^H \mathbf{C} \in \mathbb{C}^{N_b \times N_b}$, $\mathbf{B} = \sum_{k=1}^{K} w_k u_k^* \mathbf{C}^T \mathbf{h}_{b,k} \mathbf{h}_{b,k}^T \in \mathbb{C}^{N_b \times K}$. To satisfy the transmitt power constrain, precoding matrix $\hat{\mathbf{G}}_b$ satisfies

$$\begin{aligned} tr\left(\hat{\mathbf{G}}_b \hat{\mathbf{G}}_b^H\right) - P &= tr\left(\left(\mathbf{\Lambda} + \lambda \mathbf{I}_{N_b}\right)^{-2} \mathbf{D}^H \mathbf{B} \mathbf{B}^H \mathbf{D}\right) - P \\ &= \sum_{m=1}^{N_b} \frac{\mathbf{\Gamma}(m,m)}{\left(\mathbf{\Lambda}(m,m) + \lambda\right)^2} - P = 0 \end{aligned} \tag{15}$$

where $\mathbf{D}\Lambda\mathbf{D}^H$ is the eigende composition form of $\mathbf{A}$, $\mathbf{\Lambda} \in \mathbb{R}^{N_b \times N_b}$ is diagonal matrix of eigenvalues for $\mathbf{A}$. We can find that $tr\left(\hat{\mathbf{G}}_b \hat{\mathbf{G}}_b^H\right)$ is monotonically increasing in terms of λ, the optimal solution λ^{opt} can be obtained by bisection method.

Based on the analysis above, we find that the objective function of problem P3 is concave with respect to each variable respectively when fixing all other variables, and the optimal solutions in cosed-form for $\{u_k\}$, $\{w_k\}$ and $\hat{\mathbf{G}}_b$ are represented in (6), (11) and (14). Inspired by this property, blocked coordinate descent (BCD) method is used to solve problem P2. In the t-th iteration of the proposed precoding algorithm, firstly fix $\hat{\mathbf{G}}_{b,t}$, $\{w_{k,t}\}$, and update $\{u_{k,t}\}$ according to (6). Next step, fix $\hat{\mathbf{G}}_{b,t}$, $\{u_{k,t}\}$, and update $\{w_{k,t}\}$, $R_t = \sum_{k=1}^{K} R_{k,t}$ on the basis of (11), (8). Then, fix variables $\{u_{k,t}\}$, $\{w_{k,t}\}$, calculate λ_{t+1} by bisection method, and update $\hat{\mathbf{G}}_{b,t+1}$ by (14). Repeat the iteration until the maximum number of iterations is reached or $R_t - R_{t-1} \leq \eta$, η is a threshold assigned to avoid unneccessary iteration.

In summary, the proposed precoding algorithm based on weighted MMSE is described in Table 1. we design typical zero-force (ZF) precoder as the initial configuration $\hat{\mathbf{G}}_{b,1} = \hat{\alpha}\mathbf{C}^T\mathbf{H}_b\left(\mathbf{H}_b^H \mathbf{C}\mathbf{C}^T\mathbf{H}_b\right)^{-1}$, where $\hat{\alpha}$ is the power normalization factor, $\mathbf{C}$ is the beam selection matrix obtained in advance.

Table 1. Precedure of proposed precoding algorithm based on weighted MMSE

Algorithm 1 Precoding algorithm based on weighted MMSE

Input: Maximum number of iterations T ; Threshold value η .

Initialization: Obtain initial $\hat{\mathbf{G}}_{b,1}$;

For $1 \le t \le T$ **do**

1. Fix $\hat{\mathbf{G}}_{b,t}$ and $\{w_{k,t}\}$, upadate $\{u_{k,t}\}$ by (6);
2. Fix $\hat{\mathbf{G}}_{b,t}$ and $\{u_{k,t}\}$, update $\{w_{k,t}\}$ and R_t on basis of (11) and (8);
3. Fix $\{u_{k,t}\}$ and $\{w_{k,t}\}$, calculate λ_{t+1} and update $\hat{\mathbf{G}}_{b,t+1}$ by (14);
4. **If** $t = T$ or $R_t - R_{t-1} \le \eta$, **break**; **else** $t = t+1$;

End for

Output: Precoding matrix $\hat{\mathbf{G}}_b^{opt}$ 。

4 Simulation Results and Analysis

In this section, we present the simuation results to demonstrate the performance of proposed algorithm. It is assumed that BS employs ideal DLA with demension $N = 64$ and serves $K = 8$ or $K = 40$ users equipped with single antenna. The channel vector $\mathbf{h}_k$ is set the same as that in [2], and all the following simulation results are averaged over 100 channel realizations.

Figure 2 illustrates the system SE of different precoding algorithms against N_b, when SNR = 5dB. In Fig. 2, "WMMSE-MCS", "MMSE-MCS" and "ZF-MCS" represent the hybrid beamformers consist of MCS algorithm and proposed algorithm, MMSE, ZF precoder, respectively. "WMMSE-Full", "MMSE-Full" and "ZF-Full" respectively indicate the corresponding fully digital. It can be seen that, fully digital proposed algorithm always performs better than typical MMSE and ZF precoders. Moreover, when number of beams $N_b = 20$, three kinds of hybrid beamformers achieve almost the same performance as fully digital precoders, which demonstrates that beam selection can significantly reduce hardware complexity without obvious system perfomance loss. When $N_b \ge 11$ and $N_b \ge 12$, WMMSE-MCS respectively achieves higher system SE than ZF-Full and MMSE-Full, which further verifies the good performance of proposed algorithm.

Figure 3 shows the system SE of different precoding algorithms agaist SNR, when $K = 40$, $N_b = 42$. It can be observed that, the SE of WMMSE-MCS is almost the same as that of WMMSE-Full. Moreover, the number of beams for proposed precoding algorithm $N_b = 42$ is the half of the latter two fully digital precoder with $N_b = N = 64$, which proves that WMMSE-MCS achieves much better performance than ZF-Full and MMSE-Full. That is because, proposed precoding algorithm is designed to find the global optimal solution of the minimization problem of weighted MMSE. Through iterations, it can keep approaching the global optimal solution of the best precoder matrix which can achieve the highest system SE, which guareetes the good perfomance of proposed algorithm.

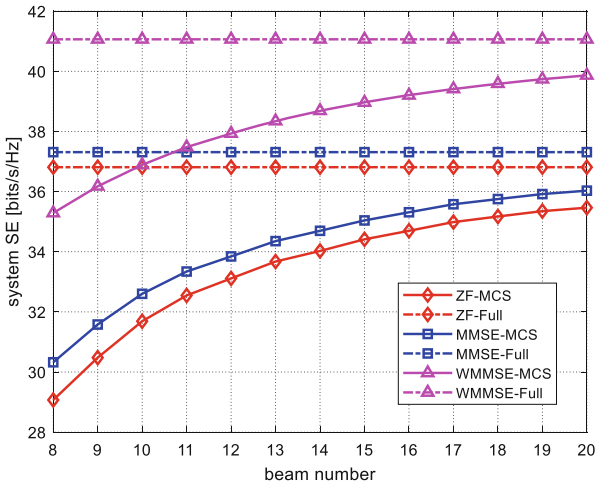


Fig. 2. System SE of different precoding algorithms against N_b

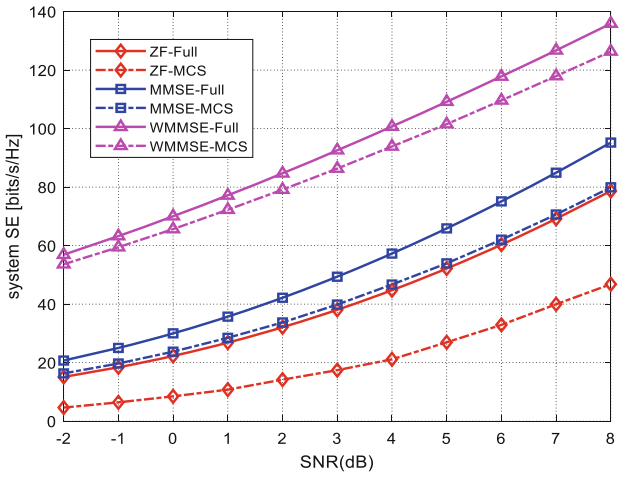


Fig. 3. System SE of different precoding algorithms against SNR

5 Conclusion

In this paper, inspired by the relationship between received MMSE and spectral efficiency for single-antenna user, we reformulate the problem of maximization for system SE as the problem of minimization for weighted MMSE. The proposed precoding algorithm based on weighted MMSE is verified to have a good property of convergency, and it can achieve monotonically increasing system SE in terms of iteration time. Simulation results prove that proposed precoding algorithm performs much better than typical ZF or MMSE precoder. Moreover, we are going to study efficient joint optimization of precoding and beam selelction matrixs in wideband Beamspace MIMO system in future.

References

1. Lu, L., Li, G.Y., Swindlehurst, A.L., Ashikhmin, A., Zhang, R.: An overview of massive MIMO: benefits and challenges. IEEE J. Selected Top. Signal Process. **8**(5), 742–758 (2014). https://doi.org/10.1109/JSTSP.2014.2317671
2. Brady, J., Behdad, N., Sayeed, A.M.: Beamspace MIMO for millimeter-wave communications: system architecture, modeling, analysis, and measurements. IEEE Trans. Antennas Propag. **61**(7), 3814–3827 (2013)
3. Amadori, P.V., Masouros, C.: Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection. IEEE Trans. Commun. **63**(6), 2212–2223 (2015)
4. Gao, X., Dai, L., Chen, Z., Wang, Z., Zhang, Z.: Near-optimal beam selection for beamspace MmWave massive MIMO systems. IEEE Commun. Lett. **20**(5), 1054–1057 (2016)
5. Qiu, S., Luo, K., Jiang, T.: Beam selection for mmWave massive MIMO systems under hybrid transceiver architecture. IEEE Commun. Lett. **22**(7), 1498–1501 (2018)
6. Wang, D., Zhang, W., Zhu, Q.: Heuristic search inspired beam selection algorithms for mmWave MU-MIMO system with discrete lens array. IEEE Access **9**, 61324–61333 (2021)

Directional Modulation Design Based on Hybrid Beamforming Structure for Massive MIMO

Maolin Li[1], Bo Zhang[1(✉)], Baoju Zhang[1(✉)], Wei Liu[2], Taekon Kim[3], and Cheng Wang[1]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
limaolin0302@163.com, {b.zhangintj,cwang}@tjnu.edu.cn, wdxyzbj@163.com
[2] Communications Research Group, Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield S1 4ET, UK
w.liu@sheffield.ac.uk
[3] Department of Electronics and Information Engineering, Korea University, Seoul, Korea
taekonkim@korea.ac.kr

Abstract. As a physical layer security technique, directional modulation (DM) is normally realized using a fully digital structure due to its easy and flexible hardware design. However, for a large antenna array, such as the case of massive MIMO, to reduce the cost and power consumption, a hybrid digital and analog beamforming structure is widely adopted. In this paper, the DM design problem is studied based on such a hybrid implementation for the first time. As a representative example, the interleaved subarrays are used for the proposed design. Design examples are provided to show the effectiveness of the proposed design.

Keywords: Directional modulation · Hybrid beamforming · Interleaved subarrays

1 Introduction

Traditional wireless communication transmitters implement digital modulation at baseband, where the constellation pattern of signals received in different azimuth angles are only different in their overall amplitudes. As a physical layer security technique, direction modulation (DM) can keep the constellation pattern in the desired direction but scramble it in other directions, further improving the security of a communication system [1–4]. In [5], a phased array was proposed that scrambles the amplitude and phase of the signal in undesired directions and maintains the modulation pattern in the desired direction by phase-shifting each element. A four-symbol modulation was proposed in [6], and the phase shift was chosen to maximize the bit error rate in the eavesdropping direction.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 44–50, 2023.
https://doi.org/10.1007/978-981-99-1260-5_6

In [7], a four-dimensional (4-D) modulation scheme was proposed on two polarization diversity channels using the polarization information of the signal, and the number of antennas was reduced based on the compressive sensing technique. A dual-beam directional modulation technique was proposed in [8], which utilizes in-phase and quadrature (IQ) baseband signals to interfere with phase in undesired directions.

The above DM designs are normally applied on either a fully analogue or digital structure. The advantage of analogue implementation is low cost, but its performance is sub-optimal compared with the digital one due to practical constraints on the analogue coefficients. The advantage of digital structure is its easy and flexible hardware design, where the amplitude and phase of the signal can be adjusted with very high precision. However, in a fully digital implementation each antenna needs to be equipped with an independent digital to analogue converter (DAC), which incurs high energy and hardware costs when a large number of antennas are needed, in particular in the case of massive MIMO for future communication systems [9]. To solve the problem, various hybrid digital and analog beamforming structures and designs have been studied extensively for massive MIMO, which greatly reduce the number of required DACs [10–12]. One representative structure for hybrid beamforming is the sub-aperture or subarray based, such as the side by side or localised implementation and the interleaved one [13–15], while the interleaved structure can provide a much narrower beam but with a higher sidelobe level.

In this work, the directional modulation design problem is studied for the first time employing the sub-aperture based hybrid beamforming structure. Although the proposed design can be applied to both the localised and the interleaved structures, without loss of generality, we adopt the interleaved subarray structure as a representative example.

The remaining part of this paper is structured as follows. The hybrid DM design problem based on interleaved subarrays is formulated in Sect. 2. Design examples are provided in Sect. 3, followed by conclusions in Sect. 4.

2 DM Design Based on Interleaved Subarrays

As shown in Fig. 1, the uniform linear array (ULA) T_x is divided into N interleaved subarrays, each subarray consists of M array elements, and the distance between the zeroth and the m-th antenna is d_m ($m = 1, \ldots, MN-1$). Then, the spacing between adjacent antennas in each subarray is $d_M = Nd_1$. Each subarray is connected to a DAC for digital precoding to form an interleaved subarray structure. The steering vector of the n-th subarray can be represented by

$$\begin{aligned} &\mathbf{s}_n(\omega, \theta) = [e^{j\omega d_n \cos\theta/c}, \ldots, e^{j\omega d_{n+(M-2)N} \cos\theta/c}, e^{j\omega d_{n+(M-1)N} \cos\theta/c}]^T, \\ &\text{for} \quad n = 0, \ldots, N-1; \quad \text{with} \quad d_0 = 0 \end{aligned} \tag{1}$$

where $\{\cdot\}^T$ is the transpose operation, ω is the angular frequency, $\theta \in [0°, 180°]$ is the transmission angle, and c is the speed of propagation. The digital beamformer

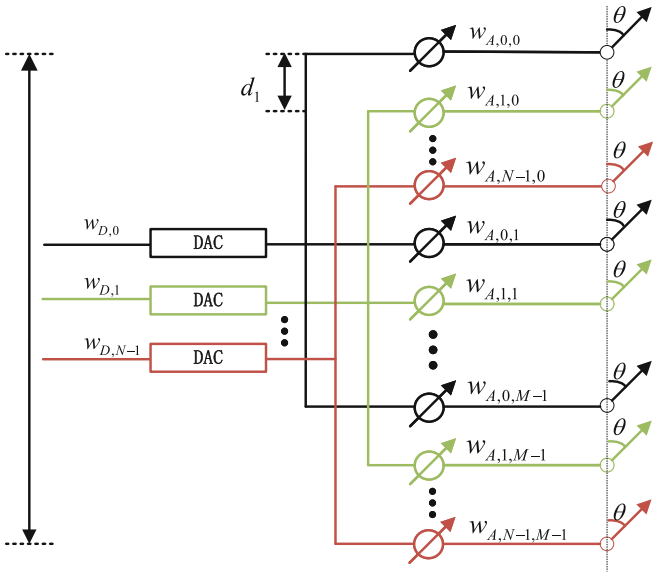


Fig. 1. Interleaved subarrays based hybrid DM structure.

coefficient vector and the analog weighting factors can be represented by

$$\begin{aligned} \mathbf{w}_D &= [w_{D,0}, w_{D,1}, \ldots, w_{D,N-1}], \\ \mathbf{w}_{A,n} &= [w_{A,n,0}, w_{A,n,1}, \ldots, w_{A,n,M-1}], \end{aligned} \tag{2}$$

where $w_{D,n}$ $(n = 0, \ldots, N-1)$ corresponds to the weight corresponding to the n-th subarray, which is a complex number. Suppose r points are sampled in the mainlobe region and $R - r$ points are sampled in the sidelobe region. Then, we have the corresponding steering matrices where $\mathbf{S}_{h,n}$ and $\mathbf{S}_{l,n}$ corresponding to the n-th subarray are represented by

$$\begin{aligned} \mathbf{S}_{h,n} &= [\mathbf{s}_n(\omega, \theta_0), \mathbf{s}_n(\omega, \theta_1), \ldots, \mathbf{s}_n(\omega, \theta_{r-1})], \\ \mathbf{S}_{l,n} &= [\mathbf{s}_n(\omega, \theta_r), \mathbf{s}_n(\omega, \theta_{r+1}), \ldots, \mathbf{s}_n(\omega, \theta_{R-1})]. \end{aligned} \tag{3}$$

In DM design, for B-ary signaling, the beam responses of the two regions are $\mathbf{p}_{b,H}$ and $\mathbf{p}_{b,L}$, respectively, where $b = 0, \ldots, B-1$. Then, the digital beamformer coefficient vector $\mathbf{w}_{b,D}$ and the analog weighting factors $\mathbf{w}_{b,A,n}$ for the b-th symbol can be written as

$$\begin{aligned} \mathbf{w}_{b,D} &= [w_{b,D,0}, w_{b,D,1}, \ldots, w_{b,D,N-1}], \\ \mathbf{W}_{b,A} &= \begin{bmatrix} \mathbf{w}_{b,A,0} \\ \mathbf{w}_{b,A,1} \\ \cdots \\ \mathbf{w}_{b,A,N-1} \end{bmatrix} \\ \text{with} \quad \mathbf{w}_{b,A,n} &= [w_{b,A,n,0}, w_{b,A,n,1}, \ldots, w_{b,A,n,M-1}]. \end{aligned} \tag{4}$$

Accordingly, the designed responses for each symbol at the R sampling points can be expressed as

$$\begin{aligned} \mathbf{p}_{b,h} &= [p_b(\omega, \theta_0), \ldots, p_b(\omega, \theta_{r-1})], \\ \mathbf{p}_{b,l} &= [p_b(\omega, \theta_r), \ldots, p_b(\omega, \theta_{R-1})]. \end{aligned} \tag{5}$$

Then, to maintain the desired response in the mainlobe region, and disturb the phase of the received signal with a low magnitude level in the sidelobe region, optimisation of the set of digital coefficients and analog coefficients for DM design for the b-th symbol can be formulated as

$$\begin{aligned} &\min_{\mathbf{w}_{b,D}, \mathbf{W}_{b,A}} ||\mathbf{p}_{b,l} - \sum_{n=0}^{N-1} \mathbf{w}_{b,D}(n)\mathbf{W}_{b,A}(n,:)\mathbf{S}_{l,n}||_2 \\ &\text{subject to} \quad \sum_{n=0}^{N-1} \mathbf{w}_{b,D}(n)\mathbf{W}_{b,A}(n,:)\mathbf{S}_{h,n} = \mathbf{p}_{b,h}, \end{aligned} \tag{6}$$

where $b = 0, 1, \ldots, B-1$ and $||\cdot||_2$ represents the l_2 norm. The analog matrix $\mathbf{W}_{b,A}$ is implemented by phase shifters and satisfies the constant modulus constraint, i.e., $|\mathbf{W}_{b,A}| = \mathbf{I}_{N\times M}$, where $\mathbf{I}_{N\times M}$ represents an all-ones matrix of size $N \times M$. To solve the non-convex constant modulus constraint problem, we assume a new constraint

$$||vec(|\mathbf{W}_{b,A}|)||_\infty \leq 1, \tag{7}$$

where $vec(\cdot)$ and $||\cdot||_\infty$ represent matrix vectorization and l_∞ norm, respectively. Adding constraint (7) to (6), we have

$$\begin{aligned} &\min_{\mathbf{w}_{b,D}, \mathbf{W}_{b,A}} ||\mathbf{p}_{b,l} - \sum_{n=0}^{N-1} \mathbf{w}_{b,D}(n)\mathbf{W}_{b,A}(n,:)\mathbf{S}_{l,n}||_2 \\ &\text{subject to} \\ &\sum_{n=0}^{N-1} w_{b,D}(n)\mathbf{W}_{b,A}(n,:)\mathbf{S}_{h,n} = \mathbf{p}_{b,h} \\ &||vec(|\mathbf{W}_{b,A}|)||_\infty \leq 1. \end{aligned} \tag{8}$$

Then, in order to guarantee the constraint that the constant modulus is one, $\mathbf{w}_{b,D}$ and $\mathbf{W}_{b,A}$ are optimized alternately by the following method:

1. Initialize $\mathbf{W}_{b,A} = e^{j\pi}\mathbf{I}_{N\times M}$.
2. Optimize $\mathbf{w}_{b,D}$ by (8) based on $\mathbf{W}_{b,A}$.
3. Taking $\mathbf{w}_{b,D}$ as a given value, $\mathbf{W}_{b,A}$ is optimized using (8).
4. Let $\mathbf{W}_{b,A} = e^{j\angle \mathbf{W}_{b,A}}\mathbf{I}_{N\times M}$, i.e., the amplitudes of analog coefficient factors are all corrected to one; then use (8) to re-optimize $\mathbf{w}_{b,D}$.
5. Go back to step 3 until the cost function no longer changes and the iteration ends.

The above optimization problem can be solved with the cvx toolbox in MATLAB and a minimum cost function can be found [16].

3 Design Examples

In this section, design examples are provided for the proposed method. The number of antennas for the ULA transmitter T_x is 32 with $d_1 = \lambda/3$, and each

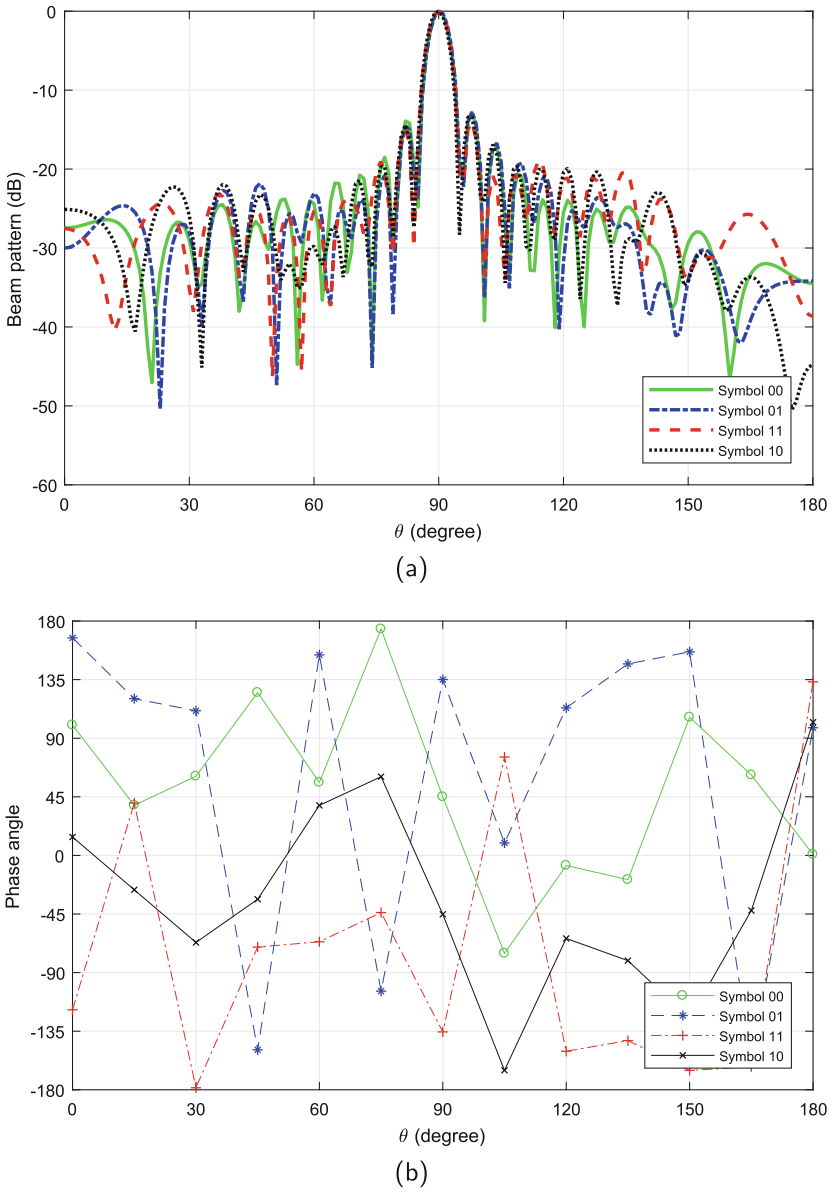


Fig. 2. (a) Resultant beam responses and (b) phase responses for DM design.

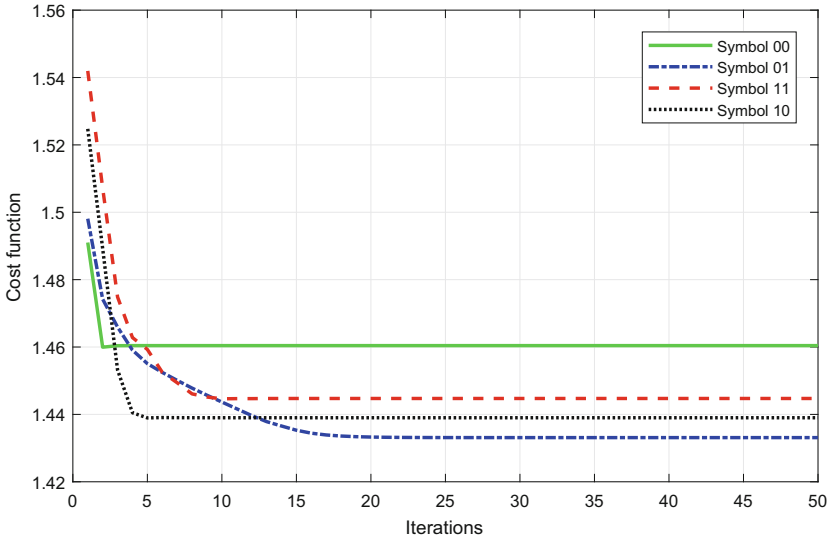


Fig. 3. Cost function vs Iterations in (8).

subarray consists of eight antennas, i.e., $M = 8$. The mainlobe direction is $\theta_h = 90°$ and the sidelobe region is $\theta_l \in [0°, 85°] \cup [95°, 180°]$, sampled every 1°. The desired response received in the direction of the mainlobe is a value of one in magnitude (the gain is 0 dB) with 90° phase shift, i.e. symbols '00', '01', '11', '10' correspond to 45°, 135°, −135° and −45°, respectively, and a value of 0.1 (magnitude) with randomly generated phase shifts in the sidelobe region.

The resultant magnitude and phase responses for all sampling angles are shown in Figs. 2a and 2b. The beam response level for all sampling angles $\theta_l \in [0°, 85°] \cup [95°, 180°]$ in the sidelobe region is lower than 0dB which is the magnitude response for the mainlobe direction, and the phases of signal at these sampling angles θ_l are random. The change of the cost function value with the number of iterations is shown in Fig. 3, which shows that the cost function values corresponding to all symbols converge.

4 Conclusions

In this paper, the interleaved subarrays based hybrid DM design is investigated and a new method is proposed. As shown in the resultant magnitude and phase responses, signals with a given modulation pattern can only be received in the desired direction, but scrambled in other directions. The proposed optimization process is effective, as demonstrated by the convergence of the cost function values for different symbols.

Acknowledgement. The work was supported by the National Natural Science Foundation of China (62101383).

References

1. Babakhani, A., Rutledge, D.B., Hajimiri, A.: Near-field direct antenna modulation. IEEE Microwave Mag. **10**(1), 36–46 (2009)
2. Xie, T., Zhu, J., Li, Y.: Artificial-noise-aided zero-forcing synthesis approach for secure multi-beam directional modulation. IEEE Commun. Lett. **22**(2), 276–279 (2018)
3. Shu, F., et al.: Directional modulation: a physical-layer security solution to B5G and future wireless networks. IEEE Netw. **34**(2), 210–216 (2020)
4. Zhang, B., et al.: Directional modulation design under a given symbol-independent magnitude constraint for secure IoT networks. IEEE Internet Things J., 8(20), 15:140–15:147 (2021)
5. Daly, M., Bernhard, J.: Directional modulation technique for phased arrays. IEEE Trans. Antennas Propag. **57**(9), 2633–2640 (2009)
6. Daly, M.P., Daly, E.L., Bernhard, J.T.: Demonstration of directional modulation using a phased array. IEEE Trans. Antennas Propag. **58**(5), 1545–1550 (2010)
7. Zhang, B., Liu, W., Lan, X.: Orthogonally polarized dual-channel directional modulation based on crossed-dipole arrays. IEEE Access **7**, 34:198–34:206 (2019)
8. Hong, T., Song, M.-Z., Liu, Y.: Dual-beam directional modulation technique for physical-layer secure communication. IEEE Antennas Wireless Propag. Lett. **10**, 1417–1420 (2011)
9. Larsson, E.G., Edfors, O., Tufvesson, F., Marzetta, T.L.: Massive MIMO for next generation wireless systems. IEEE Commun. Mag. **52**(2), 186–195 (2014)
10. Sohrabi, F., Yu, W.: Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays. IEEE J. Selected Areas Commun. **35**(7), 1432–1443 (2017)
11. Satyanarayana, K., El-Hajjar, M., Kuo, P.-H., Mourad, A., Hanzo, L.: Hybrid beamforming design for full-duplex millimeter wave communication. IEEE Trans. Veh. Technol. **68**(2), 1394–1404 (2019)
12. Molisch, A.F., et al.: Hybrid beamforming for massive MIMO: a survey. IEEE Commun. Mag. **55**(9), 134–141 (2017)
13. Zhang, J.A., Huang, X., Dyadyuk, V., Guo, Y.J.: Massive hybrid antenna array for millimeter-wave cellular communications. IEEE Wireless Commun. **22**(1), 79–87 (2015)
14. Shimizu, M.: Millimeter-wave beam multiplexing method using subarray type hybrid beamforming of interleaved configuration with inter-subarray coding. Int. J. Wireless Inf. Netw. **24**(3), 217–224 (2017)
15. Zhang, J., Liu, W., Gu, C., Gao, S.S., Luo, Q.: Multi-beam multiplexing design for arbitrary directions based on the interleaved subarray architecture. IEEE Trans. Veh. Technol. 69(10), 11:220–11:232 (2020)
16. Grant, M., Boyd, S., Ye, Y.: CVX: matlab software for disciplined convex programming, version 2.0 beta (2013)

Research on Ranging Range Based on Binocular Stereo Vision Plane-Space Algorithm

Qing Shi[1], Xiaohua Zhang[2], Huiyu Zhu[1], Xuefei Li[1], and Shubin Wang[1(✉)]

[1] College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China
wangshubin@imu.edu.cn

[2] Department of Foreign Languages, Guizhou University of Commerce, Guiyang, China

Abstract. In the Internet of Vehicles, the binocular stereo vision ranging method has the advantages of high efficiency, simple system structure and low cost, and its ranging function is an indispensable part of the intelligent network vehicle terminal, but how to use it to determine the ranging range is rarely studied. To solve the problem that the binocular stereo vision plane-space algorithm cannot evaluate the distance range on the limited surface, this paper proposes a surface constraint-range sweep algorithm, which is applied to evaluate the distance range between pedestrians and vehicles in Internet of Vehicles environment. First, according to the image sensor information of the ranging model, the model parameters of the plane-space algorithm are calculated. Next, whether the projection of the target point is on the image sensor is calculated according to the principle of geometric optical imaging. For the point projected on the image sensor, the error margin of the algorithm point can be set according to the working environment. If the error margin is met, the measurement point is considered to be within the limited surface value range. Finally, according to the actual object, the point cloud of the target surface is obtained by using the geometric optics theory, and the range of the plane-space algorithm on the limited surface is calculated, so as to construct the ranging evaluation model of the plane-space algorithm. The experimental results show that the efficiency of the algorithm can reach 99. 8% by adjusting the parameters of the surface constraint-range sweep algorithm. The surface constraint-range sweep algorithm can more accurately evaluate the range of the plane-space algorithm on the limited surface.

Keywords: Internet of Vehicles · Plane-space algorithm · Surface constraint-range sweep algorithm · Target surface point cloud

1 Introduction

Based on the Intranet, Internet, and Mobile Internet of the vehicle, the Internet of Vehicles wirelessly connects the vehicle centric system in accordance with the agreed communication protocol and data exchange standard [1], and the vehicle terminal interconnection is one of the main ways to realize the function of the Internet of Vehicles. Traditional intelligent network connected vehicle terminals include T-Box, Tracker, OBD, ETC OBU,

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 51–58, 2023.
https://doi.org/10.1007/978-981-99-1260-5_7

etc. The tracker can realize positioning management and obtain real-time vehicle location information. Binocular stereo vision is the process of obtaining depth information from a pair of left and right camera images [2]. It does not require the priori conditions of the target object, and the distance measurement can be completed through the parallax of the left and right image sensors, which is widely used in vehicle positioning.

As an important branch of computer vision, binocular stereo vision is mainly based on the parallax principle, which uses two image acquisition devices located in different positions to obtain two images of the measured object, and then calculates the three-dimensional information of the object by calculating the coordinate difference of a point in the reference image and the corresponding point in the target image. There are many examples for the realization of binocular stereo vision ranging and positioning functions [3–8], but there is little research on using it to determine the ranging range. In the 1980s, Marr [9] proposed a visual computing theory and applied it to binocular matching. By calculating the disparity of two plane images, a stereo image with depth can be further generated, thus laying the theoretical foundation for the development of binocular stereo vision. Reference [10] focuses on the image matching method based on Canny edge detection and Harris corner detection, and calculates the distance according to the principle of parallax. Its experiments show that binocular stereo vision ranging is related to the focal length and resolution of the camera, and the ranging error can be maintained at about 6%. Reference [3] proposes an algorithm that combines MATLAB calibration and OpenCV matching to achieve binocular stereo vision ranging. Simulation experiments show that when the ranging range is less than 2 m, the measurement error of the algorithm does not exceed 5%, and the measurement accuracy is affected by the texture of the measured object. The description of the ranging range of the algorithm is not exhaustive. In [11], Ye Qingzhou et al. proposed a flame localization and ranging algorithm based on binocular stereo vision. The accuracy of the algorithm's flame location is higher than 95% within the measurement range of 0.5 m to 5 m, and the ranging error is less than 5%. The experiment mainly measures the distance of target objects within a fixed range, and does not explain how to evaluate the range of the algorithm. Zhang Jiaxu [12] et al. proposed a target distance measurement algorithm based on deep learning and binocular stereo vision, which realized obstacle target recognition and localization. Through the ranging experiment at a fixed distance, the error analysis between the measured value and the actual value is used to evaluate the accuracy of the algorithm, and when the distance is 4.5 m, the algorithm has the highest accuracy, and the corresponding error is 0.9%. Reference [13] mainly studies the distance measurement of vehicles, and proposes an automatic driving distance measurement system based on image processing. The system selects camera bases with different heights to perform ranging experiments at a fixed distance. When the selected base is 0.6 m, the error is the smallest. Reference [14] proposes a distance estimation system based on pedestrian detection results to obtain the distance between pedestrians and cameras. The experiment uses NVIDIA's graphics processor to accelerate pedestrian detection and distance estimation. When the ranging range is less than 5 m, the relative error of the ranging system is less than 8%. Reference [15], Zhang Enshuo proposed a new binocular stereo vision ranging algorithm, which uses the plane coordinates measured by the image sensor to calculate the three-dimensional coordinates of the target point, and reflects the space information of the target object by

selecting the appropriate target boundary points. Experimental data show that the error rate of the algorithm is less than 2% within 1 m.

The algorithms or systems proposed in different ranging studies mainly carry out ranging experiments at a fixed distance, and the determination of ranging range is not explained in some studies. Based on the plane space algorithm proposed by Zhang Enshuo, this paper studies the ranging range of the plane-space algorithm on a limited surface according to geometric optics theory, and proposes a surface constraint-range sweep algorithm, which effectively solves the problem of obtaining the distance range of vehicles or pedestrians in the environment of Internet of vehicles.

2 Plane-Space Algorithm

Firstly, the plane-space algorithm proposed by Zhang Enshuo is introduced. According to literature [15], the plane - space algorithm model is shown in Fig. 1:

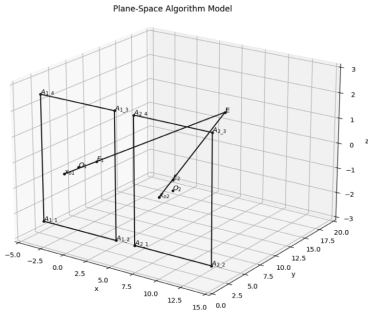


Fig.1. Mathematical model of plane-space algorithm

The rectangle $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$ and $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$ are the positions and sizes of the left and right image sensors of the binocular ranging system. The midpoint of $A_{1\_1}A_{1\_4}$ is on the origin of the standard coordinate system, its normal vector is parallel to the y axis of the standard coordinate system, and the x and y axes of image sensors $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$ and $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$ are parallel to the x and z axes of the standard coordinate system respectively. The point E is the measured target, and points E_1 and E_2 are the projections of point E on image sensors $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$ and $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$. x_{o1} and x_{o2} are the rear focus corresponding to image sensors $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$ and $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$, and x_{o1} and x_{o2} are on the x-axis. O_1 and O_2 are the projection coordinates of the optical axis corresponding to image sensors $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$ and $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$ respectively. The distance between the left and right image sensors and x_{o1}, x_{o2} is y_1. Let $E_1(x_1, y_1, z_1), E_2(x_2, y_1, z_2), x_{o1}(x_{o1}, 0, 0)$ and $x_{o2}(x_{o2}, 0, 0)$ be known. If the target is calculated according to points $E_1(x_1, y_1, z_1)$

and $E_2(x_2, y_1, z_2)$, the coordinates of the target point are $E(x, y, z)$:

$$\begin{cases} x = \frac{x_1 x_{o2} - x_2 x_{o1}}{(x_{o2} - x_{o1}) + (x_1 - x_2)} \\ y = \frac{(x_{o2} - x_{o1}) y_1}{(x_{o2} - x_{o1}) + (x_1 - x_2)} \\ z = \frac{(x_{o2} - x_{o1}) z_2}{(x_{o2} - x_{o1}) + (x_1 - x_2)} \end{cases} \tag{2.1}$$

3 Surface Constraints - Range Sweep Algorithm

In actual working environment, it is usually necessary to have a detailed knowledge of the target, such as the outline of a pedestrian, the shape of a vehicle, and so on. Therefore, the performance of the algorithm can be estimated from one or more sets of normative values for that type of contour or shape. The surface constraint-value range scanning algorithm is a method related to the plane-space algorithm for calculating the value range on limited surface. Therefore, the surface constraint-range sweep algorithm is derived using the relevant parameters of the plane-space algorithm and the principle of optical imaging. The parameters related to the model in the plane-space algorithm are x_{o1}, x_{o2}, y_1, where x_{o1}, x_{o2} determines the distance of the left and right sensors, and y_1 determines the focal length of the left and right image sensors. Let the length and width of the left and right image sensors are w and h, the distance between the two sensors is d, and the focal length is f, then

$$\begin{cases} x_{o1} = \frac{w}{2} \\ x_{o2} = \frac{3}{2}w + d \\ y_1 = f \end{cases} \tag{3.1}$$

The input parameters of the model are x_1, x_2, z_2. If the target point can be detected at this time, it must meet the requirements of the existence of the projection of these three points on the corresponding sensor, that is

$$\begin{cases} x_1 \in (0, w) \\ x_2 \in (w + d, 2w + d) \\ z_2 \in (-\frac{1}{2}h, \frac{1}{2}h) \end{cases} \tag{3.2}$$

Bring the model input parameters x_1, x_2, z_2 into formula (2.1), let the coordinates of the target point be $E_{text}(x_{text}, y_{text}, z_{text})$, the standard coordinates of the target point are $E_{std}(x_{std}, y_{std}, z_{std})$, and set the error sequence $(error_x, error_y, error_z)$ according to the actual situation. If met

$$\begin{cases} |x_{text} - x_{std}| \le error_x \\ |y_{text} - y_{std}| \le error_y \\ |z_{text} - z_{std}| \le error_z \end{cases} \tag{3.3}$$

If any point on the target point cloud $E_{std}(x_{std}, y_{std}, z_{std})$ corresponds to $E_{text}(x_{text}, y_{text}, z_{text})$) and x_1, x_2, z_2 meet both formula (3.2) and formula (3.3), the

point is considered to be within the algorithm value range. Let the number of valid points of the estimated model be Cnt_e, and the total number of points is Cnt_s, then the effective rate is

$$r = \frac{Cnt_e}{Cnt_s} \tag{3.4}$$

4 Experimental Analysis

4.1 Experimental Principle and Steps

According to the principle of geometrical optics, after calculating the projection of all points on the surface with known average shape on the image sensor, the corresponding algorithm value is calculated according to the projection value. If the difference between the algorithm value and the standard value is within the set range, the point is considered to be within the algorithm value domain. The experimental steps are as follows:

1. Obtain the point cloud of the measured target surface according to the actual situation;
2. For any point, calculate the projection of the point on the image sensor;
3. Calculate the algorithm value of the target point according to the projection coordinates obtained in step 2;
4. Compare the algorithm value with the real value of the target point. If the difference between the coordinates of the algorithm value and the real value is less than the error redundancy, and the target point is projected within the imaging range of the image sensor, that is, if formula (3.2) and (3.3) are satisfied at the same time, the point is considered to be within the algorithm value range;
5. Draw the simulation results by marking the marked target point cloud as blue, the points in the value range as red, the points that do not conform to the algorithm value as black, and the invisible points as green. Then calculate the efficiency of the algorithm for the target according to formula (3.4).

4.2 Experimental Data and Analysis

w, h, d, f, p and $disVec$ are respectively the width and height of image sensors, distance between the left and right image sensors, focal length, pixel size and target displacement of image sensors. The unit of width and height of the image sensor is pixel, the unit of distance between left and right image sensors is millimeter, the unit of focal length is millimeter, the unit of pixel size is micrometer, and the unit of target displacement is meter. This project mainly conducts simulation experiments on the outline of people and the shape of vehicles.

4.2.1 Human Silhouette Simulation

The standard posture of a person is taken as an example for simulation, in which the width of the image sensor is 1920p, the height of the image sensor is 1080p, the resolution

of the input image at this time is 1920 × 1080, the distance between the left and right image sensors is 500 mm, the focal length of the image sensors is 5 mm, the pixel size is 6 μm, the error margin sequence $(error_x, error_y, error_z)$ is uniformly set as $1e^{-5}$, and the target displacement is [−120.5 m, 105 m, 0]. The simulation results are shown in Fig. 2, and the effective rate is $r = 89.83\%$.

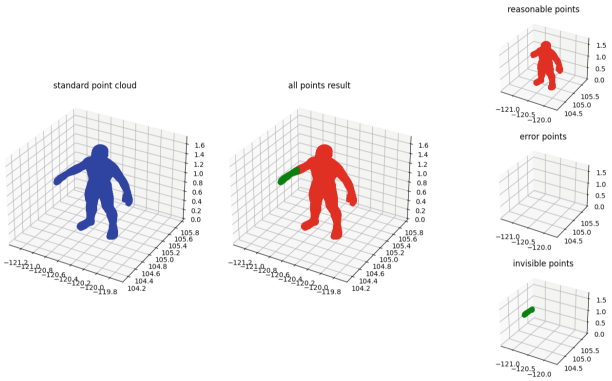


Fig.2. Range evaluation image of human standard contour

By adjusting the simulation parameters, the target displacement is adjusted to [−120.7 m, 105.5 m, 0], the simulation result is shown in Fig. 3, and the effective rate is $r = 98.97\%$.

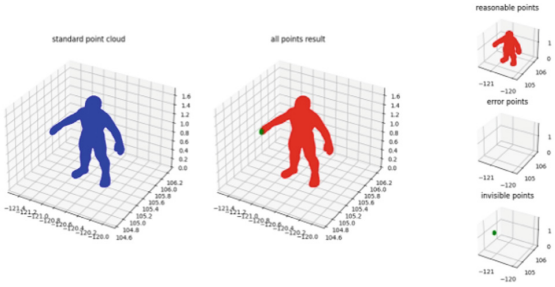


Fig.3. Value range evaluation image of standard posture after adjusting parameters

4.2.2 Vehicle Shape Simulation

Taking the shape of the vehicle as an example for simulation, in which the width of the image sensor is 1920*p*, the height of the image sensor be 1080*p*, the resolution of the input image is 1920 × 1080, the distance between the left and right image sensors is 500 mm, the focal length of the image sensors is 5mm, the pixel size is 6μm, the error margin sequence $(error_x, error_y, error_z)$ is uniformly set as $1e^{-5}$, and the target

displacement is [−117.7 m, 107.4 m, 0]. The simulation results are shown in Fig. 4, and the effective rate is $r = 99.80\%$.

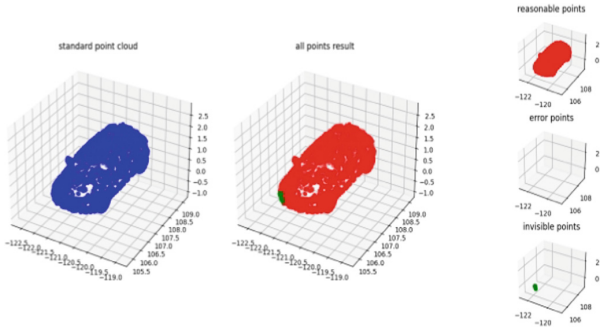


Fig.4. Value range evaluation image of vehicle shape

5 Conclusion

In the context of the Internet of Vehicles, this paper studies how to evaluate the plane-space algorithm to determine the ranging range on a limited surface, and proposes a surface constraint-range sweep algorithm. The experimental results show that the ranging range of the surface constraint-range sweep algorithm can be adjusted by changing the target displacement and other parameters, and in the simulation of the vehicle shape, the efficiency can reach 99.8% when the target displacement is $[-117.7m, 107.4m, 0]$. When similar model parameters are brought into the plane-space algorithm, the results of the above simulations do not appear. Therefore, the surface constraint-range sweep algorithm performs well in evaluating the distance range of the target object, and effectively obtains the distance range of the target object in the environment of the network of vehicle. However, due to the complexity of the model data, the algorithm has certain requirements for the computing power of the device, and the real-time performance is not good enough. This problem can be solved by adding the number of cores of the device processor or optimizing the calculation amount of the algorithm.

Acknowledgment. Shubin Wang (wangshubin@imu.edu.cn) is the correspondent author and this work was supported by the National Natural Science Foundation of China (61761034), the Natural Science Foundation of Inner Mongolia (2020MS06022).

References

1. Wang, Y., Zhao, Y.: Research on service monitoring system of intelligent connected vehicle based on cloud platform. In: 2021 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), pp. 963–966. IEEE (2021)

2. Geiger, D., Ladendorf, B., Yuille, A.: Occlusions and binocular stereo. Int. J. Comput. Vision **14**(3), 211–226 (1995)
3. Sun, X., Jiang, Y., Ji, Y., et al.: Distance measurement system based on binocular stereo vision. IOP Conf. Ser.: Earth Environ. Sci. IOP Publish. **252**(5), 052051 (2019)
4. Wang, C., Zou, X., Tang, Y., et al.: Localisation of litchi in an unstructured environment using binocular stereo vision. Biosys. Eng. **145**, 39–51 (2016)
5. Shuai, G., Wenlun, M., Jingjing, F., et al.: Target recognition and range-measuring method based on binocular stereo vision. In: 2020 4th CAA International Conference on Vehicular Control and Intelligence (CVCI), pp. 623–626. IEEE (2020)
6. Guo, S., Chen, S., Liu, F., et al.: Binocular vision-based underwater ranging methods. In: 2017 IEEE International Conference on Mechatronics and Automation (ICMA), pp. 1058–1063. IEEE (2017)
7. Zhang, Q., Ma, C.: Automobile braking performance detection system based on binocular stereo vision. J. Phys.: Conf. Ser. IOP Publishing **2037**(1), 012082 (2021)
8. Mu, Q., Wei, J., Yuan, Z., et al.: Research on target ranging method based on binocular stereo vision. In: 2021 International Conference on Intelligent Computing, Automation and Applications (ICAA), pp. 81–85. IEEE (2021)
9. Marr, D., Poggio, T.: A computational theory of human stereo vision. Proc. R. Soc. Lond. **204**(1156), 301–328 (1979)
10. Liu, Z., Chen, T.: Distance measurement system based on binocular stereo vision. International Joint Conference on Artificial Intelligence. IEEE (2009)
11. Ye, Q., Cheng, Y., Zhang, M., et al.: Research on flame location and distance measurement method based on binocular stereo vision. In: 2020 Chinese Automation Congress (CAC), pp. 4089–4094. IEEE (2020)
12. Zhang, J., Hu, S., Shi, H.: Deep learning based object distance measurement method for binocular stereo vision blind area. Methods, 9(9) (2018)
13. Zaarane, A., Slimani, I., Al Okaishi, W., et al.: Distance measurement system for autonomous vehicles using stereo camera. Array **5**, 100016 (2020)
14. Li, J., Wu, J., You, Y., et al.: Parallel binocular stereo-vision-based GPU accelerated pedestrian detection and distance computation. J. Real-Time Image Proc. **17**(3), 447–457 (2020)
15. Zhang, E., Wang, S., Sun, Y.: A new binocular stereovision measurement by using plane-space algorithm. In: Liang, Q., Jiasong, Mu., Jia, M., Wang, W., Feng, X., Zhang, B. (eds.) Communications, Signal Processing, and Systems: Proceedings of the 2017 International Conference on Communications, Signal Processing, and Systems, pp. 622–628. Springer Singapore, Singapore (2019). https://doi.org/10.1007/978-981-10-6571-2_76

Research on Plane-Space Algorithm of Binocular Stereo Vision Under Zoom Condition

Xuefei Li[1], Xiaohua Zhang[2], Qing Shi[1], Huiyu Zhu[1], and Shubin Wang[1(✉)]

[1] College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China
wangshubin@imu.edu.cn

[2] Department of Foreign Languages, Guizhou University of Commerce, Guiyang, China

Abstract. In the paper, a zoom-plane-space algorithm based on binocular stereo vision is presented to expand the range of the plane-space algorithm. On the basis of the plane-space algorithm, a mathematical model of the zoom-plane-space algorithm is established by introducing the zoom function. According to the optical imaging principle, the imaging model of the zoom image sensor is established and the input point coordinates of the left image sensor without zoom are calculated with this model. Finally, the position coordinates of the target point can be obtained by bringing the converted input point coordinates into the plane-space algorithm. The algorithm is simulated by using the Matplotlib library based on Python3.6. The simulation data show that the error of the algorithm is about 0.04 m when the horizontal coordinates of the target point are moved in steps of 0.1 for 100 steps.

Keywords: Binocular stereo vision · Optical zoom · Zoom-plane-space algorithm

1 Introduction

In the future development of automobile and transportation technology, Internet of vehicles (IoV) technology is regarded as the main technical means to solve the existing traffic problems. In the IoV application scenarios, a multi-sensor fusion network is used to detect information about the surrounding environment, in which the visual sensors are mainly responsible for obstacle detection and distance measurement. Binocular stereo vision (BSV) technology can effectively obtain the distance information in three-dimensional spatial scene. The advantages like high efficiency, suitable accuracy and simple system structure make it very appliable in such fields as unmanned driving and target tracking [1–4].

The main task of BSV algorithm is to obtain the coordinate information of the target object by calculating the position deviation between the corresponding points of images, according to which target matching is an important step in BSV ranging algorithms, therefore many scholars have improved the performance of BSV algorithms by enhancing the accuracy of target matching in binocular ranging algorithms [5–7]. The

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 59–67, 2023.
https://doi.org/10.1007/978-981-99-1260-5_8

BSV ranging algorithm determines the mapping relationship between the target point in the world coordinate system and the pixel coordinate system through the camera stereo calibration. Hence, some researchers also improve the solution accuracy of 3D world coordinate by improving the accuracy of binocular vision calibration [8–10]. Similarly, based on the parallax principle and the structural characteristics of the elements of image sensor, Enshuo Zhang et al. of our research group established the similarity transformation model, and proposed the plane-space algorithm, which can calculate the three-dimensional coordinates of the target point according to the binocular imaging position, and this calculation is high in precision and simple in algorithm structure [11]. However, most current stereo vision systems, including the plane-space algorithm, are mainly fixed focus stereo vision systems. Considering the scene range, in the IoV application environment, the field of view and image resolution of the fixed focus camera are very limited. Compared with the traditional fixed focus stereo vision system, the zoom stereo vision system has obvious advantages, i.e., it can obtain more and farther traffic scene information. The detection, recognition and tracking of the target objects with different depths in the traffic scene can be achieved in different scenes by adjusting the focal length of the camera.

In the environment of IoV, the maximum range of object detection of driverless vehicles limits the maximum speed of the vehicle. Therefore, based on the plane- space algorithm proposed by our project team, the paper aims to solve the problem that the traditional fixed focus stereo vision system cannot warn the collision of high-speed vehicles due to the short measurement distance, and proposes a zoom-plane-space algorithm based on BSV, which can effectively expand the range of the algorithm and play a positive role in vehicle collision warning and improving road traffic safety, etc.

2 Zoom-Plane-Space Algorithm

2.1 Algorithm Model

Figure 1 is the zoom-plane-space algorithm model, where rectangles $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$, $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$ are the positions and sizes of the photosensitive element when the focal lengths of the left and right image sensors of the binocular ranging system are the same, and rectangle $A_{3\_1}A_{3\_2}A_{3\_3}A_{3\_4}$ are the positions and sizes of the photosensitive element of the left image sensor after zooming. The plane-space algorithm model is composed of $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$ and $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$, while in the zoom-plane-space algorithm model, an additional $A_{3\_1}A_{3\_2}A_{3\_3}A_{3\_4}$ overlaps with $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$, therefore the two rectangles are drawn as shown in Fig. 1 for convenience of observation. Point E is the measured target, and points E_1, E_2 and E_3 are the projections of point E on image sensors $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$, $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$ and $A_{3\_1}A_{3\_2}A_{3\_3}A_{3\_4}$. X_{o1}, X_{o2} and X_{o3} are the rear focus corresponding to image sensors $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$, $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$ and $A_{3\_1}A_{3\_2}A_{3\_3}A_{3\_4}$. X_{o1} and X_{o2} are on the x-axis, and X_{o3} is always on the optical axis of the left image sensor. O_1, O_2 and O_3 are the projection coordinates of the optical axis corresponding to image sensors $A_{1\_1}A_{1\_2}A_{1\_3}A_{1\_4}$, $A_{2\_1}A_{2\_2}A_{2\_3}A_{2\_4}$ and $A_{3\_1}A_{3\_2}A_{3\_3}A_{3\_4}$.

The distance between the left (right) image sensor and the point X_{o1} (X_{o2}) is y_1, let the coordinates of points $E_1(x_1, y_1, z_1)$, $E_2(x_2, y_1, z_2)$, $E_3(x_3, y_1, z_3)$, $O_1(x_{o1}, y_1, 0)$,

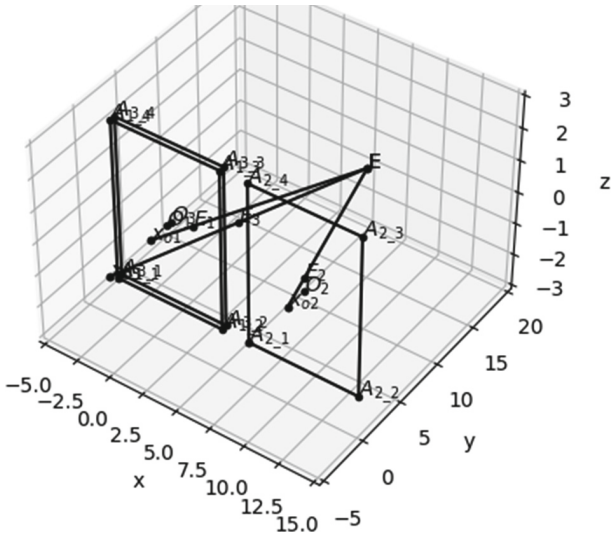


Fig. 1. Zoom-plane-space algorithm model

$O_2(x_{o2}, y_1, 0)$, $O_3(x_{o3}, y_1, 0)$, $x_{o1}(x_{o1}, 0, 0)$, $x_{o2}(x_{o2}, 0, 0)$, $x_{o3}(x_{o1}, -df, 0)$ be known. If the target is calculated according to points $E_1(x_1, y_1, z_1)$ and $E_2(x_2, y_1, z_2)$, the coordinates E(x, y, z) of the target point can be obtained according to the plane space algorithm:

$$\left\{\begin{array}{l} x = \dfrac{x_1 x_{o2} - x_2 x_{o1}}{(x_{o2} - x_{o1}) + (x_1 - x_2)} \\ y = \dfrac{(x_{o2} - x_{o1}) y_1}{(x_{o2} - x_{o1}) + (x_1 - x_2)} \\ z = \dfrac{(x_{o2} - x_{o1}) z_2}{(x_{o2} - x_{o1}) + (x_1 - x_2)} \end{array}\right\} \tag{2.1}$$

When the system zooms, the coordinates E(x, y, z) of the target point must be calculated by points $E_2(x_2, y_1, z_2)$ and $E_3(x_3, y_1, z_3)$. Bring:

$$\left\{\begin{array}{l} x_1 = \dfrac{x_3 - x_{o1}}{k} + x_{o1} \\ z_1 = \dfrac{z_3}{k} \\ k = \dfrac{y_1 + df}{y_1} \end{array}\right. \tag{2.2}$$

into formula (2.1), and the coordinates of the target point E(x, y, z) is obtained.

2.2 Algorithmic Proof

When the system is not zoomed, the algorithm model is the same as the plane-space algorithm in literature [11], so the formula (2.1) can be verified.

For formula (2.2), the model (top view) of the zoom image sensor (left image sensor) is shown in Fig. 2, where $x_{o1}O_1$ and $x_{o3}O_3$ are the optical axes of the left image sensor.

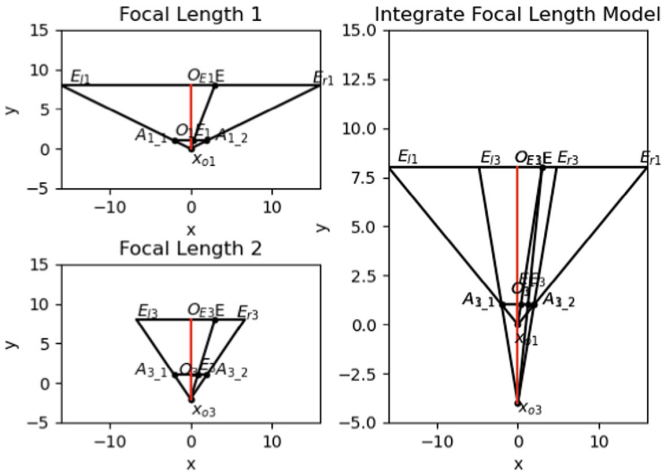


Fig. 2. Imaging model of zoom image sensor (top view)

When the camera zooms to see a targets farther away, it can be derived from the principle of optical imaging:

$$\begin{cases} \dfrac{O_1E_1}{O_{E1}E} = \dfrac{x_{o1}E_1}{x_{o1}E} = \dfrac{E_1A_{1\_2}}{EE_{r1}} \\ O_{E1}E = O_{E3}E \\ \dfrac{O_3E_3}{O_{E3}E} = \dfrac{x_{o3}E_3}{x_{o3}E} = \dfrac{E_3A_{3\_2}}{EE_{r3}} \end{cases} \tag{2.3}$$

According to the formula (2.3), it can be deduced that:

$$\frac{O_1E_1/O_3E_3}{O_{E1}E/O_{E3}E} = \frac{E_1A_{1\_2}/E_3A_{3\_2}}{EE_{r1}/EE_{r3}} \tag{2.4}$$

$$\frac{O_1E_1/O_3E_3}{O_{E1}E/O_{E3}E} = \frac{x_{o1}E_1/x_{o3}E_3}{x_{o1}E/x_{o3}E} \tag{2.5}$$

And there are:

$$\left\{ \begin{array}{l} \dfrac{x_{o1}E_1}{x_{o1}E} = \dfrac{x_{o1}O_1}{x_{o1}O_{E1}} \\ \dfrac{x_{o3}E_3}{x_{o3}E} = \dfrac{x_{o3}O_3}{x_{o3}O_{E3}} \end{array} \right\} \tag{2.6}$$

According to the formula (2.6), the following can be obtained:

$$\frac{x_{o1}E_1/x_{o3}E_3}{x_{o1}E/x_{o3}E} = \frac{x_{o1}O_1/x_{o3}O_3}{x_{o1}O_{E1}/x_{o3}O_{E3}} \tag{2.7}$$

According to the formulas (2.5), (2.7), the following can be deduced:

$$\frac{O_1E_1/O_3E_3}{O_{E1}E/O_{E3}E} = \frac{x_{o1}E_1/x_{o3}E_3}{x_{o1}E/x_{o3}E} = \frac{x_{o1}O_1/x_{o3}O_3}{x_{o1}O_{E1}/x_{o3}O_{E3}} \tag{2.8}$$

That is:

$$\frac{O_1E_1/O_3E_3}{O_{E1}E/O_{E3}E} = \frac{x_{o1}O_1/x_{o3}O_3}{x_{o1}O_{E1}/x_{o3}O_{E3}} \tag{2.9}$$

From Fig. 2:

$$\frac{O_{E1}E}{O_{E3}E} = 1 \tag{2.10}$$

$$x_{o1}O_{E1} - x_{o3}O_{E3} = x_{o1}O_1 - x_{o3}O_3 \tag{2.11}$$

From the formula (2.11), it is deduced that:

$$\frac{x_{o1}O_{E1}}{x_{o3}O_{E3}} = \frac{x_{o3}O_{E3} + x_{o1}O_1 - x_{o3}O_3}{x_{o3}O_{E3}} \tag{2.12}$$

Usually the focal length of the image sensor is in millimeters, while the distance of the target measured is in meters, so it can be obtained:

$$\frac{x_{o1}O_1 - x_{o3}O_3}{x_{o3}O_{E3}} \approx 0 \tag{2.13}$$

According to the formulas (2.12), (2.13), the following can be calculated:

$$\frac{x_{o1}O_{E1}}{x_{o3}O_{E3}} \approx 1 \tag{2.14}$$

By substituting formulas (2.10), (2.14) into formulas (2.9), it can be deduced that:

$$\frac{O_1E_1}{O_3E_3} = \frac{x_{o1}O_1}{x_{o3}O_3} \tag{2.15}$$

Therefore, the following results can be derived:

$$\frac{y_1 + df}{y_1} = \frac{x_3 - x_{o1}}{x_1 - x_{o1}} \tag{2.16}$$

The left view of the imaging model of the zoom image sensor, as shown in Fig. 3. From Fig. 3, a formula (2.15) can also be obtained, and the following results can be derived:

$$\frac{y_1 + df}{y_1} = \frac{z_3}{z_1} \tag{2.17}$$

Therefore, according to the formulas (2.16), (2.17), the formula (2.2) can be proved. To sum up, the algorithm is validated.

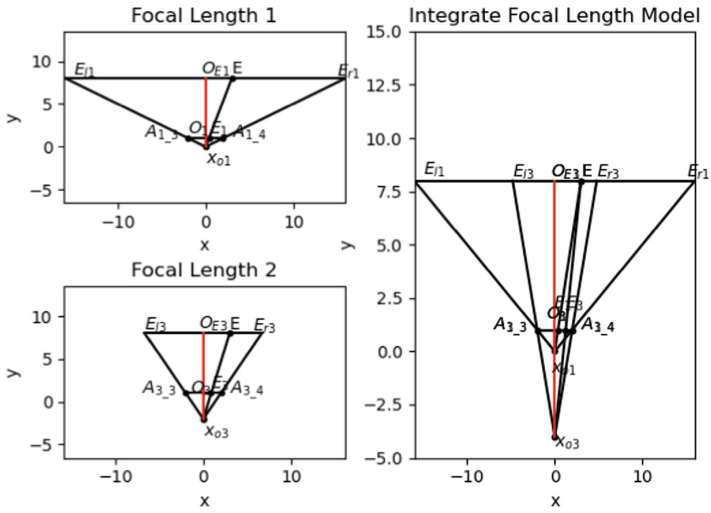


Fig. 3. Imaging model of zoom image sensor (left view)

3 Simulation and Analysis

Based on Matplotlib library of Python3.6, this paper completes all simulation analysis. First, the target point is set and the projected coordinates of the target on the image sensor are calculated according to the principle of optical imaging. Then the input values of the algorithm are determined according to the projected coordinates, and the algorithm values of the target coordinates are calculated by substituting the input values into the algorithm. Finally, the algorithm accuracy is evaluated by comparing the target algorithm values with the set standard values.

3.1 Parameters and Methods of Simulation

The parameters needed for simulation include the pixel width of the image sensor: w, the pixel height of the image sensor: h, the focal length of the left image sensor before or after zooming: f_1 *or* f_2, the spacing between left and right image sensors: d, the pixel size: sizePixel, the location of the target point: ObjPont(x, y, z) and the number of simulation experiments: Cnt.

The method of calculating model parameters by simulation parameters is as follows:

$$\begin{cases} x_{o1} = \dfrac{w}{2} \\ x_{o2} = \dfrac{3}{2}w + d \\ y_1 = 1000 \times \dfrac{f_1}{sizePixel} \end{cases} \tag{3.1}$$

First, the parameters x_{o1}, x_{o2} and y_1 of the simulation model are calculated according to the formula (3.1). Then a target ObjPont(x, y, z) is set and named the standard value,

and the input values of the plane-space algorithm and the zoom-plane-space algorithm are calculated according to the principle of optical imaging. Next, the target values of the plane-space algorithm and the zoom-plane-space algorithm are calculated according to the parameters obtained, and the values are named as the algorithm values. Finally, the coordinate performance curve is drawn with the algorithm value as the transverse coordinate and the standard value as the vertical coordinate, and the performance of the algorithm is evaluated with the coordinate performance curve. The ideal coordinate performance curve is a straight line with a slope of 1 and an intercept of 0.

3.2 Result Analysis

The initial values of simulation parameters are: w = 1920p,h = 1080p,f_1 = 5 mm, f_2 = 10 mm, d = 300 mm, sizePixel = 6 μm, ObjPont(50 m, 250 m, 8 m),Cnt = 100.

When the abscissa of the target point moves laterally by 100 steps in steps of 0.1, the simulation results are shown in Fig. 4. In the figure, the coordinate performance curve of the plane-space algorithm is marked with “singleFocal”, while the coordinate performance curve of the zoom-plane-space algorithm is marked with “multiFocal”.

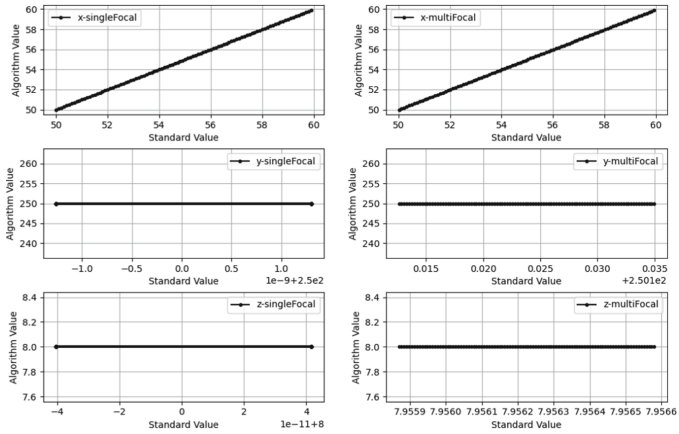


Fig. 4 Simulation image of the target moving along the X-axis

According to Fig. 4, when the measured point moves along the x-axis, the slope of the coordinate performance curve of the x-coordinate is 1, which indicates that the algorithm performs well. When measuring the x-coordinate value of the target point, the zoom-plane-space algorithm and the plane-space algorithm have the same performance. But for y or z values, the error of the algorithm value measured with the plane-space algorithm is between 0–1 m or 0–4 m, while the error of the algorithm value measured with the zoom-plane-space algorithm is about 0.01m or 0.04m.

Therefore, the performance of the zoom-plane-space algorithm is better than that of the plane-space algorithm. In addition, comparing with the plane-space algorithm, when the zoom-plane-space algorithm is used, the measurement range of equipment is relatively larger by adding zoom function, so the comprehensive performance of the

device carrying this algorithm is better than that of the device carrying the plane-space algorithm.

4 Conclusion

Based on the plane-space algorithm, this paper presents a zoom-plane-space algorithm. Using the principle of optical imaging, the algorithm converts the coordinates of the zoomed target point and uses the converted coordinates as input values to the plane-space algorithm. The calculated results are the measured values of the algorithm. The system with the zoom-plane-space algorithm has a larger measurement range than the system with the plane-space algorithm, so the range of the zoom-plane-space algorithm is farther than that of the plane-space algorithm. The simulation data show that the performance of the zoom-plane-space algorithm is better than that of the plane-space algorithm, and the range is larger than that of the plane-space algorithm.

Acknowledgment. Shubin Wang (wangshubin@imu.edu.cn) is the correspondent author and this work was supported by the National Natural Science Foundation of China (61761034), the Natural Science Foundation of Inner Mongolia (2020MS06022).

References

1. Feng, M., Liu, Y., Jiang, P., Wang, J.: Object detection and localization based on binocular vision for autonomous vehicles. J. Phys: Conf. Ser. **1544**, 012134 (2020). https://doi.org/10.1088/1742-6596/1544/1/012134
2. Fang, L., Guan, Z., Li, J.: Automatic roadblock identification algorithm for unmanned vehicles based on binocular vision. Wireless Communications and Mobile Computing, vol. 2021, Article ID 3333754, p. 7, (2021)
3. Xu, Y., Dong, Y., Li, J., Wu, X.: Research on target tracking algorithm based on parallel binocular camera. In: 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), pp. 1483–1486 (2019)
4. She, H., Yang, X., Shi, Y., Fen, B., Ye, H., Liu, W.: Design and implementation of a target tracking and ranging system based on binocular vision. IEEE Int. Conf. Recent Adv. Syst. Sci. Eng. (RASSE) **2021**, 1–5 (2021). https://doi.org/10.1109/RASSE53195.2021.9686859
5. Zhang, L., et al.: The SGM Algorithm based on Census Transform for Binocular Stereo Vision. Int. Conf. Mach. Learn. Knowl. Eng. (MLKE) **2022**, 50–54 (2022). https://doi.org/10.1109/MLKE55170.2022.00015
6. Gai, Q.: Optimization of stereo matching in 3D reconstruction based on binocular vision. J. Phys. Conf. 960(1), 012029 (2018)
7. Y. Wei and Y. Xi, "Optimization of 3-D Pose Measurement Method Based on Binocular Vision," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1–12, 2022, Art no. 8501312, doi: https://doi.org/10.1109/TIM.2022.3149334
8. Gao, Z., et al.: Stereo camera calibration for large field of view digital image correlation using zoom lens. Measurement 185, 109999 (2021). ISSN 0263–2241
9. Li, Y.: A calibration method of computer vision system based on dual attention mechanism. Image Vis. Comput. 103, 104039 (2020). ISSN 0262–8856

10. Huang, S., Gu, F., Cheng, Z., Song, Z.: A joint calibration method for the 3D sensing system composed with ToF and stereo camera. IEEE Int. Conf. Inf. Autom. (ICIA) **2018**, 905–910 (2018)
11. Zhang, E., Wang, S., Sun, Y.: A new binocular stereovision measurement by using plane-space algorithm. In: Liang, Q., Jiasong, Mu., Jia, M., Wang, W., Feng, X., Zhang, B. (eds.) Communications, Signal Processing, and Systems: Proceedings of the 2017 International Conference on Communications, Signal Processing, and Systems, pp. 622–628. Springer Singapore, Singapore (2019). https://doi.org/10.1007/978-981-10-6571-2_76

Research on Tablet Crack Detection Algorithm Based on Improved YOLOv5

Huiyu Zhu[1], Xiaohua Zhang[2], Xuefei Li[1], Qing Shi[1], and Shubin Wang[1(✉)]

[1] College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China
wangshubin@imu.edu.cn

[2] Department of Foreign Languages, Guizhou University of Commerce, Guiyang, China

Abstract. Tablet crack detection is an important part to ensure drug quality in the process of drug production. In view of the problems of target detection algorithm in the actual tablet crack detection, such as few training samples and the embedded device based on ARM architecture is difficult to accommodate the algorithm when the system is implemented, this paper proposes a tablet crack detection algorithm based on improved YOLOv5. The collected data set is processed by data enhancement, and YOLOv5 algorithm is improved by using fusion involution and varifocal loss to detect tablet cracks. The improved tablet crack detection algorithm improves the accuracy and speed of model detection. In the detection system designed with Raspberry Pi as the core, the crack detection mAP_{50} is 99.5%, and the reasoning speed is reduced from 797 ms to 299 ms. After the model is converted and simplified, the reasoning speed is reduced to 86 ms, which meets the design requirements of the crack detection system based on ARM architecture.

Keywords: Tablet crack detection · YOLOv5 · involution · varifocal loss · Raspberry Pi

1 Introduction

In the production process, the quality of drug products is always affected by the appearance of tablet cracks. Because of the dynamic changes and the high complexity in the production process, as well as the lack of effective testing technology and equipment, the situation of cracks can't be detected directly by machines. In manual detection of cracks, there are such problems as high labor intensity, high cost, low accuracy and low efficiency. In recent years, with the continuous development of machine vision and in-depth learning, traditional manual detection and mechanical detection can no longer meet the requirement of modern industrial production, while visual detection technology gradually realizes the new requirements of industrial abnormal defect detection [1]. In the field of drug defect detection, considering the fuzziness and diversity of the defect images of capsules [2], Wang H et al. used Back Propagation (BP) network to identify the defect of abnormal matter in capsules. Jiang Yueqiu et al. designed an automatic detection software for image processing and recognition technology, which uses Support Vector Machine (SVM) algorithm tod. KS P et al. designed a detection system

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 68–75, 2023.
https://doi.org/10.1007/978-981-99-1260-5_9

that uses Canny edge detection algorithm for image processing. Defects are identified as speckles by convoluting the image matrix [4]. Xiangyu Ma et al. designed an X-ray computed tomography (XRCT) analysis combined with a deep learning Convolutional Neural Networks (CNN) algorithm for the detection of internal cracks in tablets, which can significantly improve the accuracy of in-slice crack analysis [5]. Yang J et al. introduced a background weakening method based on image filtering for the surface defects of medicinal capsules under complex background, and proposed a YOLO-OL target detection method based on image contour extraction [6].

In the application of deep learning technology in the actual industrial scene, there is the problem that the number of crack samples for drug detection training is small; For the larger object detection model, there will be the problems of slow reasoning speed and high delay in the CPU of embedded device based on ARM architecture. Joseph Redmon et al. proposed the YOLO algorithm [7] in 2016, which realizes the classification and positioning of target detection with neural network, and has been widely used in the field of target detection. After the development of YOLO algorithm from v1 to v4 [7–10], the YOLOv5 network structure launched in 2020 has the advantages of small volume, fast speed and high precision, and is easy to be implemented in Pytorch compilation software with mature development ecology. The YOLOv5 network structure is simple in deployment with strong hardware adaptability. Therefore, this paper intends to improve the YOLOv5 algorithm for drug crack detection and to propose an improved YOLOv5 drug crack detection algorithm, with which the detection accuracy and speed can be improved to meet the design requirements of the crack detection system based on ARM architecture.

2 Related Work

2.1 YOLOv5 Algorithm Structure

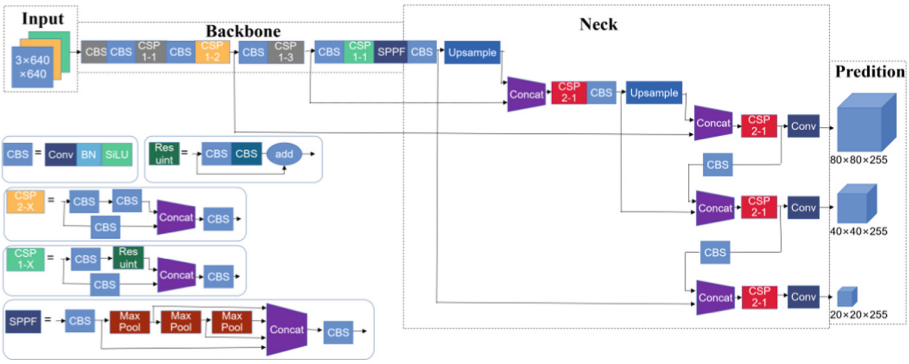


Fig. 1. Structure diagram of YOLOv5 model

The specific model structure of YOLOv5 is shown in Fig. 1. YOLOv5 algorithm model is mainly composed of four parts: input, backbone, neck and detection. The input

layer fills or scales the input image so that its size becomes 3*640*640. I t normalizes the image data and converts it into a floating-point number between 0 and 1. The backbone network realizes the extraction of target features. The neck network collects target features, which is composed of FPN + PAN structure to further extract and fuse the image feature information output from the backbone network. The detection layer classifies and locates the feature information output from the neck network, so as to output the classification probability, confidence, bounding box and other information of the detection target.

2.2 Involution

In order to solve the problems that Spatial-agnostic [11] weakens the ability of convolution kernels to adapt to different visual patterns in different spatial locations, and the redundancy between channels in Channel-specific [12] limits the flexibility of convolution kernels for different channels, Duo Li et al. proposed involution [13]. Compared with standard convolution, the involution has symmetric reverse characteristics and shares the involution kernel in the channel dimension to reduce the redundancy of the kernel.

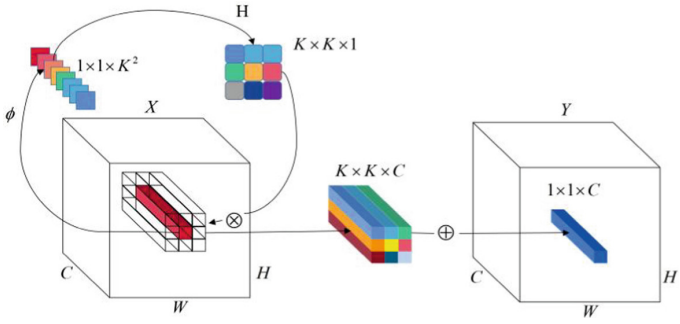


Fig. 2. Schematic illustration of involution

The involution is shown in Fig. 2. Its core $\boldsymbol{H} \in \boldsymbol{R}^{\boldsymbol{H} \times \boldsymbol{W} \times \boldsymbol{K} \times \boldsymbol{K} \times \boldsymbol{G}}$ is for the pixel $\boldsymbol{X_{ij}} \in \boldsymbol{R}^{\boldsymbol{C}}$ located in the corresponding coordinate $(\mathbf{i}, \mathbf{j})$, and $\boldsymbol{G}$ is the number of groups sharing the same involution operator core for each group. The input data can be multiplicatively added by involution to map its output characteristics. The definition of involution is as follows.

$$Y_{i,j,k} = \sum_{(u,v) \in K} H_{i,j,u+[K/2],v+[K/2],[kG/C]} X_{i+u,j+v,k} \tag{2.1}$$

To make the inner convolution operator as simple as possible, first provide a design space for the kernel generation function $\boldsymbol{\phi}$, and then carry out some effective design examples for practical use. Therefore, select a single pixel $\boldsymbol{X_{i,j}}$ to cross all inner convolution operator kernels $\boldsymbol{H_{i,j}}$, and each pixel $\boldsymbol{\Psi_{i,j}}$ has a kernel generation function $\boldsymbol{\phi}$, the formula is as follows.

$$H_{i,j} = \varphi(X_{i,j}) = W_1 \sigma(W_0 X_{i,j}) \tag{2.2}$$

In the formula $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$ represent two linear transformations, which together constitute the bottleneck structure. The dimension of the intermediate channel is controlled by the reduced order ratio $\boldsymbol{r}$, and $\boldsymbol{\sigma}$ represents the nonlinear activation function of the two linear transformations after normalization.

2.3 Varifocal Loss

Inspired by focal loss [14], Haoyang Zhang et al. designed a varifocal loss [15] for training dense detectors, which can predict the IOU aware classification score (IACS). Varifocal loss solves the problem of class imbalance when returning to continuous IACS in the process of intensive detector training by using the weighting idea of focal loss for reference. Varifocal loss is based on binary cross entropy loss, and its definition is as follows.

$$VFL(p,q) = \begin{cases} -q(q\log(p) + (1-q)\log(1-p)) & q > 0 \\ -\alpha p^{\gamma}\log(1-p) & q = 0 \end{cases} \tag{2.3}$$

Here $\boldsymbol{p}$ is the predicted IACS and $\boldsymbol{q}$ is the target score. For a front scenic spot, the $\boldsymbol{q}$ value of the ground truth category is set as the IOU (gt_IOU) of the prediction frame and the ground truth. For a background point, the target q is 0 for all categories. Varifocal loss scales the loss by a coefficient of $\boldsymbol{p^{\gamma}}$, which will only reduce the loss contribution of negative samples ($\boldsymbol{q} = 0$), but not reduce that of positive samples ($\boldsymbol{q} > 0$). This is because compared with negative samples, positive samples are extremely rare, so this learning signal should be retained. At the same time, inspired by the Program for International Student Assessment, the training goal $\boldsymbol{q}$ is used to weight the positive samples.In addition, in order to balance the loss of positive samples and negative samples, an adjustable scaling factor $\boldsymbol{\alpha}$ is added to the negative loss term.

3 Lightweight YOLOv5 Detection Algorithm

3.1 Data Enhancement

YOLOv5 adopts mosaic enhancement in data enhancement, which can expand the data set and enhance the recognition ability of target objects. Mixup enhancement is to randomly select two samples from all samples, read their sample label data, get a new sample label data after fusion calculation, and then add the data to the samples for training. Mixup has low computational overhead and can be implemented in a few.

lines of code, and can improve the image classification accuracy by 1%.

$$\begin{aligned} \tilde{x} &= \lambda x_i + (1-\lambda)x_j, \quad \textit{where } x_i, x_j \textit{ are raw input vectors} \\ \tilde{y} &= \lambda y_i + (1-\lambda)y_j, \quad \textit{where } y_i, y_j \textit{ are one} - \textit{hot label encodings} \end{aligned} \tag{3.1}$$

where $\boldsymbol{\lambda} \epsilon [0, 1]$. $(\mathbf{x_i}, \mathbf{y_i})$ and $(\mathbf{x_j}, \mathbf{y_j})$ are two samples randomly selected from the training sample data in Formula 3.1. In order to maximize GPU training and enrich data features, this paper uses mosaic and mixup algorithm to enhance data.

3.2 Fusing Involution

In this paper, the involution module is integrated into the backbone network area of the network model. The structure comparison is shown in Fig. 3.

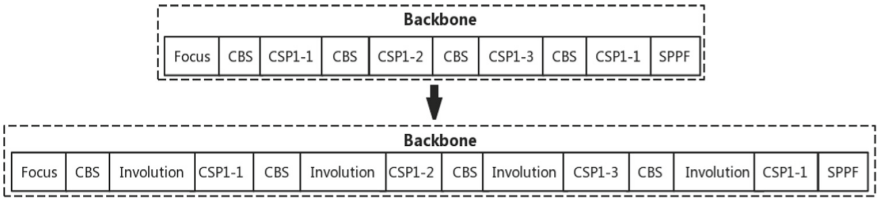


Fig. 3. Structure comparison before and after fusion

The fusion of involution can reduce the amount of parameters and calculations. The kernel of the involution is a method similar to the attention mechanism SE, 1×1 convolution for channel compression $= (\boldsymbol{C/r})$. Channel expansion to group number. $\boldsymbol{G} \times \boldsymbol{K} \times \boldsymbol{K}$. The parameter quantity of convolution is $\boldsymbol{K} \times \boldsymbol{K} \times \boldsymbol{C} \times \boldsymbol{C}$, then the parameter quantity is reduced. Although the parameters of the fused network are reduced through channel sharing, the more characteristic channels sharing a convolution kernel, the lower the performance. However, the performance does not degrade after the number of characteristic channels sharing a convolution kernel increases from 4 to 16. For comprehensive performance, space and computational efficiency, every 16 feature channels are divided into a group, and a convolution kernel is shared within the group. While the kernel is generated, using multiple linear transformation matrices to compress the channel can reduce the number of parameters, but it will reduce the performance. In this paper, two linear transformations are used to compress the channel first and then restore it, where the compression ratio is $\boldsymbol{r} = 4$.

3.3 Improving Loss Function

YOLOv5 network uses the Cross Entropy Loss function as the loss function of confidence and category. When the Cross Entropy Loss function is used for gradient descent, it only pays attention to the accuracy of correct label prediction, but does not pay attention to the differences of other incorrect labels. Therefore, this paper will use varifocal loss to improve the loss function of YOLOv5 to monitor and optimize the detection performance of the network. The formula is as follows.

$$\begin{aligned} Loss = & \frac{1}{N_{pos}} \sum_{i} \sum_{c} VFL(p_{c,i}, q_{c,i}) \\ & + \frac{\lambda_0}{N_{pos}} \sum_{i} q_{c^*,i} L_{bbox}(bbox_i^{'}, bbox_i^*) \\ & + \frac{\lambda_1}{N_{pos}} \sum_{i} q_{c^*,i} L_{bbox}(bbox_i, bbox_i^*) \end{aligned} \tag{3.2}$$

where $p_{c,i}$ and $q_{c,i}$ are the predicted IACS and target IACS of the position on the characteristic map of each level of FPN respectively. L_{bbox} is the GIoU loss, and $bbox_i'$, $bbox_i$, and $bbox_i^*$ respectively represent the initial, refined, and ground truth bounding box. Taking training target $q_{c^*,i}$ as the weight L_{bbox}, λ_0 and λ_1 are the super parameters used to balance the two sub losses, which are set to 1.5 and 2.0 respectively. N_{pos} is the expected number of points used to normalize the total loss.

4 Experimental Analysis

4.1 Experimental Data Set

Since there is no public data set for tablet cracks and there are few samples, a total of 700 lobes with different color backgrounds, different types and different numbers are taken as the data set of this experiment. The collected data sets are labeled with Labelimg, and xml files recording category information and location information are generated. Then the training set and test set are divided according to the ratio of 9:1, of which 630 images are the training set and 70 images are the test set.

4.2 Experimental Parameters

Operating parameters of the PC of the experiment: Windows10 operating system, Intel Core i5-9300 h CPU processor with main frequency of 2.4 GHz, GTX1650 GPU processor, and 8G DDR4 memory; Pytorch machine learning framework; The whole model is trained through 200 epochs, the batch is 32, and the image training size is 640*640, and the initial learning rate is 0.01. Operating parameters of Raspberry Pi: Raspberry PI OS (64 bit) operating system, ARM cortex-a72 CPU processor with main frequency of 1.5 GHz, and 4 GB LPDDR4 memory.

4.3 Experimental Result

The YOLOv5s and YOLOv5n models before and after improvement were trained and the crack detection results were output. The comparison of the experimental results is shown in Table 1.

Table 1. Index comparison table of cracks detection dataset

Model	m AP_{50}	Training time	PC inference time	Raspberry Pi 4B inference time
YOLOv5s	99.5%	13h	13 ms	1.945s
Lightweight YOLOv5s	99.5%	10h	8 ms	0.552s
YOLOv5n	99.5%	11h	8 ms	0.797s
Lightweight YOLOv5n	99.5%	9h	7 ms	0.299s

Compared with the YOLOv5 algorithm model, the training time and reasoning time of the PC end of the lightweight YOLOv5 algorithm are reduced when the detection accuracy remains the same, and the reasoning time of applying the lightweight YOLOv5 algorithm to the Raspberry Pi is also reduced. It is shown that the performance of this algorithm is greatly improved.

(a) Original drawing (b) YOLOv5 test results (c) Improved YOLOv5 test results

Fig. 4. Comparison chart of cracks detection effect

The comparison of tablet crack detection results is shown in Fig. 4. It can be concluded that although the confidence of the improved algorithm is slightly lower than that of the original algorithm, there are no false detection and missed detection. Its accuracy is high and the detection speed is fast, so the detection effect is still good.

Table 2. Detection model inference time comparison table in Raspberry Pi

Model	Raspberry Pi 4B inference time
320(image training size).pt	0.199 s
320.onnx	0.15 s
320-sim.onnx	0.086 s

In order to continue to improve the reasoning speed of the Raspberry Pi crack detection, this experiment reduces the image training size of the lightweight YOLOv5n algorithm model to 320*320 retrain, then convert the pt model to onnx model, simplify the onnx model using Python's onnxsim package, and test the models in the above steps on the Raspberry Pi. The comparison of experimental results is shown in Table 2. After conversion and simplification, the reasoning time of Raspberry Pi crack detection model is reduced to 86 ms, which can meet the needs of crack detection in drug tablet pressing production.

5 Conclusion

In this paper, aiming at the shortage of the current real-time crack detection algorithm for tablets, an improved crack detection algorithm for tablets is proposed by using

data enhancement, involution and varifocal loss to improve YOLOv5.The experimental results show that the improved algorithm improves the detection speed while maintaining high detection accuracy. Through model transformation and simplification, the detection speed is improved in the detection system designed with Raspberry Pi as the core, which can meet the needs of crack detection in tablet pressing production.

Acknowledgment. The special fund for innovation and Entrepreneurship of Postgraduates of Inner Mongolia University (11200–121024).

References

1. Cui, Y., Liu, Z., Lian, S.: A survey on unsupervised industrial anomaly detection algorithms (2022). arXiv preprint arXiv:2204.11161
2. Wang, H., Liu, X., Chen, Y.: Detection of capsule foreign matter defect based on BP neural network. In: 2014 IEEE International Conference on Granular Computing (GrC), pp. 325–328. IEEE (2014)
3. Yueqiu, J., Shuang, M., Hongwei, G., Fei, X.: Research on defect detection technology of tablets in aluminum plastic package. Open Autom. Control Syst. J. **6**(1), 940–951 (2014)
4. Puneeth, K.S., Sahay, A., Shreya, G.H., et al.: Designing an affordable system for early defect detection using image processing. Adv. Transdisciplinary Eng. **9**, 121–128 (2019)
5. Ma, X., et al.: Deep learning convolutional neural networks for pharmaceutical tablet defect detection. Microscopy and Microanalysis, 26(S2) (2020)
6. Yang, J., Wang, Y., Wang, Z., et al.: Using stochastic gradient descent and deep learning to defect detection for medicinal hollow capsule. In: 2021 China Automation Congress (CAC), pp. 4926–4933. IEEE (2021)
7. Redmon, J.,Divvala, S., Girshick, R., et al.: You only look once: unified, real-time object detection. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–10. Las Vegas (2016)
8. Redmon, J.,Farhadi, A.: YOLO9000: better, faster, stronger. IEEE Conference on Computer Vision and Pattern Recognition, 1–9 (2017)
9. Redmon, J., Farhadi, A.: YOLOv3: anincremental improvement (2018). arXiv:1804.02767
10. Bochkovskiy, A.,Wang, C.Y., Liao, H.Y.M.: YOLOv4: optimal speed and accuracy of object detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
11. Zhang, R.: Making convolutional networks shiftinvariant again. In: ICML (2019)
12. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: ICLR (2015)
13. Li, D.,et al.: Involution: inverting the inherence of convolution for visual recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 12321–12330 (2021)
14. Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: ICCV (2017)
15. Zhang, H., Wang, Y., Dayoub, F., Sunderhauf, N.:.VarifocalNet: an IoU-aware dense object detector. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8514–8523 (2021)

Research on Classification Model of Visual Function Abnormality for Auxiliary Diagnosis Based on Machine Learning

Bao-zhu Han, Ying Tong(✉), Xing Zhang, and Lin-yun Ma

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
tongying2334@163.com

Abstract. In order to help ophthalmology clinical staff make correct judgments on patients' visual function abnormalities in a shorter period of time, and to alleviate to a certain extent the pressure of consultation due to the mismatch of medical resources, this paper presents a study on the design of a classification model based on machine learning to assist in diagnosing visual function abnormalities. Using visual function data from major optometric medical centers, a predictive classification model was developed using six algorithms (K-nearest neighbor algorithm, decision tree, random forest, support vector machine, plain Bayes, and XGBoost), and the accuracy, precision, recall, and F1 parameters were calculated to compare the training effect of the model under each algorithm, and the data set was validated using cross-validation. The results showed that the models trained with random forest could classify the "set" and "adjustment" labels with clinical applicability.

Keywords: Visual Function Abnormalities · Machine Learning · Random Forest · SMOTE · Boosting

1 Introduction

According to statistics, the number of myopia in China has exceeded 500 million in 2020, and the overall myopia rate among children and adolescents is 52.7%, including 14.3% among 6-year-old children, 35.6% among elementary school students, 71.1% among junior high school students, and 80.5% among high school students [1].The low age of myopia is becoming more and more serious and has become a major problem that plagues families, schools and even national development. However, there are only about 40,000 ophthalmologists in China and the number of consultations is rising, and professional medical resources are very scarce to match the extensive social demand.

In recent years, AI technology has rapidly developed, and it achieves a similar response to the judgment of things as human intelligence by simulating the way of human thinking and recognition. The application of artificial intelligence technology in the field of ophthalmology is relatively common, but less research has been done in ophthalmic optometry. Abnormalities in binocular visual function are one of the important factors that accelerate the growth of myopia and affect visual quality, and most

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 76–84, 2023.
https://doi.org/10.1007/978-981-99-1260-5_10

of the visual function tests rely on human judgment due to its theoretical complexity and the lack of popular inspection equipment, and require a high professional level of clinical staff. There are still many difficulties in the diagnosis of visual function abnormalities.In this paper, we use machine learning technology to study the classification problem of visual function abnormalities in assembly and regulation, and establish a classification model to assist in the diagnosis of visual disorders. This paper aims to establish a classification model to assist in the diagnosis of visual disorders [2].

2 Algorithm Processing Process

The processing process of the machine learning-based visual function abnormality classification algorithm can be divided into six steps. In the first step, a database of patients with abnormal visual function is collected and established, and the characteristic attributes of patients are observed and compared; in the second step, the single-factor influence of characteristic attributes is understood, non-essential attributes are removed, and appropriate characteristic values are screened; in the third step, different machine learning models are used for classification prediction, and the classification effect of each algorithm is observed by confusion matrix; in the fourth step, cross-validation is used to evaluate the comparative model In the fifth step, the SOMTE algorithm is selected to oversample the data set, and the best performing algorithm is used as the base evaluator, and then the sample data is trained by multiple sampling based on the Boosting principle to further optimize the model. In the sixth step, the system interface is visualized using the Pyside2 graphical design tool.The flow chart of the specific data processing algorithm is shown in Fig. 1.

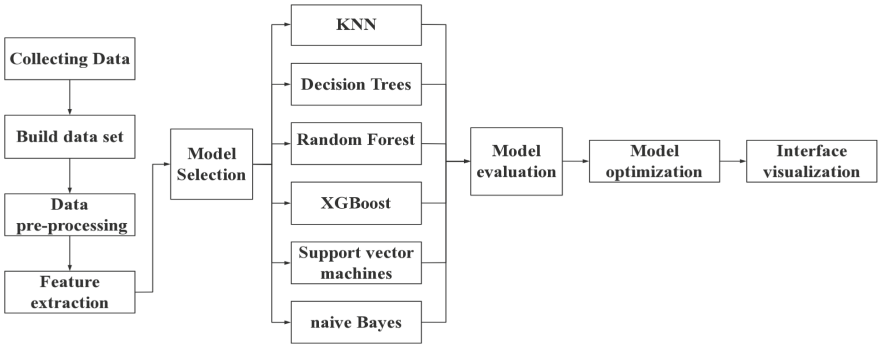


Fig 1. Flow chart of data processing algorithm

3 Data and Methods

3.1 General Information

Common binocular visual function abnormalities include: under-regulation, over-regulation; under-assembly, over-assembly; under-dispersion, over-dispersion; and

abnormal eye movements [3]. Internationally, binocular visual function abnormalities are already an important part of optometry, and the ophthalmologist or optometrist will make an analysis based on the specific conditions of the patient's eyes and treat the symptoms to achieve a more comfortable visual condition.

This paper focuses on the classification of abnormalities in regulation and assembly of visual function. The data set used was obtained from the consultation data of optometric outpatients in major ophthalmology hospitals, totaling 6,368 cases, and all patients' sensitive information was removed before analyzing the data. All patients met the diagnostic criteria for abnormal visual function behavior. The data sample labels were categorized according to the following Table 1.

Table 1. Sample label statistics and classification of raw data

Sample Category	Exception Category	Number of set samples	Adjusting the number of samples
0	Normal	2699	4322
1	Not enough	3295	1847
2	Excessive	391	217

The original datasets has 43 feature attributes and 2 labels, among which, two labels are "set" and "regulation", both of which are divided into three categories of abnormal categories, and are represented by numbers, with 0 indicating "normal The abnormality categories of both are divided into three categories and are represented by numbers, with 0 indicating "normal", 1 indicating "insufficient", and 2 indicating "excessive". Since the original data had many features and attributes and all of them were specialized terms, before pre-processing the data, through communication with professional ophthalmologists, it was determined that there were actually only 9 attributes related to "collection" and "adjustment" abnormalities. Therefore, the non-essential features were removed, and some of the data were selected and presented as shown in Fig. 2 below.

	Sex	Age	BI/BO	BI/BO.1	NRA	PRA	BCC	AC / A	Worth-4-Dot (far)	Set	Adjustment
0	1	32	6.0	4.5	+3	-1.25	+0.5	2.5	4	0	1
1	0	37	-5.0	-3.0	+1.75	-1.75	+0	3.0	4	1	1
2	1	14	-4.5	-7.0	+2.5	-1.75	+25	3.0	5	0	0
3	1	14	0.5	6.5	+2	-1.00	+0.75	2.0	4	0	1
4	0	14	-2.0	-8.0	+2.5	-0.75	+0.5	4.0	5	0	1

Fig 2. Original partial sample data

3.2 Data Pre-processing

After the initial screening of the characteristic attributes, the data are pre-processed. To detect the existence of the same data in the sample information, duplicate information

is deleted; some abnormal values in the data sample are referred to the professional physician to evaluate whether to use; for noisy data, the method of direct deletion or replacement using the average is chosen; for some data with its possibility and meaning, it is chosen to retain.

The type of missing data is divided into two categories, a 0 value and a null value. If the data itself is a 0 value, it is a valid value, and if it is a null value, it needs further processing. In this paper, we use the function of pandas library for line-by-line detection, and for missing samples, we need to combine the characteristics of the features themselves to deal with the situation. For the two attributes of "near horizontal skew" and "far horizontal skew", since the parameter is divided into positive and negative values, it is not good to fill with the average or median directly, and the actual missing values are small, so the method of direct deletion is used; for "NRA" and "PRA" are filled with the mean value, and the mean value of the "NRA" column is 1.90 and the mean value of the "PRA" column is −1. The mean value of "NRA" column is 1.90 and the mean value of "PRA" column is −1.23, which are replaced to the missing values of the corresponding columns respectively; the characteristic attribute of "BCC" column is similar to BI/BO, and the missing samples are also deleted; "AC/A". The mean values of "AC/A" and "Worth-4-dot (far)" are 2.6 and 4.0, respectively, so they are filled to the corresponding missing items. After the initial screening of the data is completed, the data types need to be converted and normalized. The data type conversion ensures the uniformity of the value types, and the normalization process eliminates the influence of dimensionality among the attributes. After the key features were selected empirically by the professional physicians in the first stage, the data were then processed for feature extraction based on the relationship between each feature and the labeled samples. A decision tree algorithm was used to evaluate the contribution of each feature, as shown in Fig. 3 and Fig. 4.

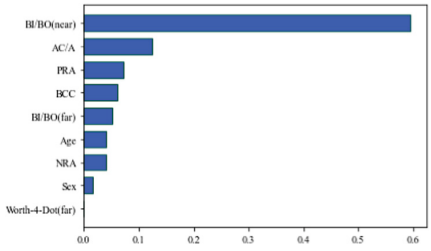


Fig. 3. Importance of each feature under the "Set" tab

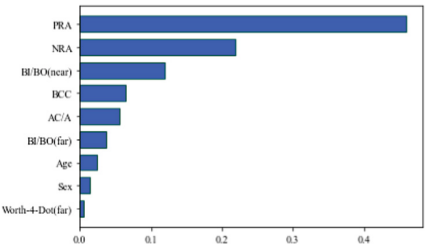


Fig. 4. Importance of each feature under the "adjustment" tab

After the above pre-processing of the data, the samples of the data set were reduced to 3983, and the distribution of each data sample is shown in Table 2. From the above statistics, it can be seen that the data samples of pooling and conditioning have very little distribution in the "excessive" category, and the data set has a serious imbalance problem.

3.3 Model Training

The data set was divided into training and test sets in the ratio of 7:3, and 70% of the random data in each classification was used for model training, and the remaining 30% of the data was used as the test set. In this study, the following six machine learning algorithms were used to build the visual function abnormality detection models: K-nearest neighbor algorithm, decision tree, random forest, support vector machine, plain Bayes and XGBoost, respectively, to understand the classification effect of each class of labels by drawing confusion matrix, and the accuracy of the models was evaluated by using cross-validation to compare the classification performance of different machine learning on the test set.

3.4 Model Optimization

3.4.1 Data Oversampling Processing Based on SMOTE Algorithm

From Table 2, we can see that the number of "excessive" samples in the datasets is too small, and the model learns the labeled features poorly during the training process, which affects the overall learning effect of the model. Due to the imbalance of samples in the data itself, the generalization ability of the model will be reduced. Therefore, this paper proposes to use the SMOTE algorithm to over-sample the data set and combine the three sample categories of the training set into new samples in the ratio of 1:1:1, so that the number of samples of each type of label is equal.

SMOTE algorithm is not simply a copy and paste of the original data, but by analyzing the characteristics of the lesser class of samples, randomly selecting a sample from the neighboring data, and then randomly selecting a point in the middle of the concatenation of the two samples as a new lesser sample, and constantly having put back in the few samples to randomly repeat the extraction of new samples to constitute a new data set until the number of sample labels reaches a balance. The processing by SMOTE algorithm can reduce the probability of model over-fitting to a great extent.

3.4.2 Multiple Sampling Method Based on Boosting Algorithm

After solving the problem of data imbalance, this paper proposes a new optimization algorithm in which the machine learning algorithm with the best performance is used as the base evaluator, and then trains the data three times using the Boosting-based principle

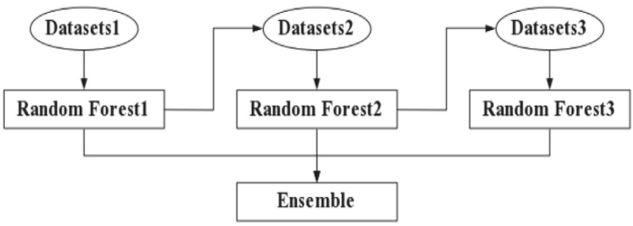


Fig 5. Flow chart of Boosting algorithm

to sample the data, integrates the three trained models, and arrives at the final result by voting.The specific algorithm flow is shown in Fig. 5.

Since random forest performs best on both "set" and "adjustment" labels, random forest is used as the base learner for all three training sessions. First, the first modeling is performed using the data from training set 1. Secondly, 50% of the correct samples and 50% of the incorrect samples are taken from the first learning to form a new data sample, i.e., training set 2, and then modeled again, and this is repeated twice. Finally, the models from the three training sessions are integrated. This method allows difficult samples to receive more attention in classification, accumulate learning experience, and thus improve the accuracy of the model.

3.5 Model Evaluation

In this study, a Pandas database based on Python programming software was used for statistical analysis of the data, combined with the matplotlib plotting library for graph plotting. A 10-fold cross-validation was used to evaluate the prediction performance of the model; after the modeling was completed, the confusion matrix of the test set was automatically generated and used to visually evaluate the classification effect of each sample label; the performance of the model was evaluated by calculating the accuracy, precision, recall and F1 score.

4 Results and Presentation

4.1 Performance Comparison of Six Machine Learning Algorithms

The performance of the classification model based on the test set was evaluated by calculating the accuracy, precision, recall and F1 score of the model under each algorithm from the prediction results of the confusion matrix. The metrics obtained from modeling the six machine learning algorithms are shown in Table 3 and Table 4.

From the data in the above table, we can see that the models built by Random Forest and XGBoost algorithms perform better in terms of accuracy and precision in the "pooled" model. Combined with the performance of the labels on the model, the "normal" and "insufficient" samples perform better, but the performance of the "excessive" samples has a recall rate of only 50%, which leads to a decrease in the overall recall rate. However, for the XGBoost algorithm, the classification of the "excessive" samples is worse than that of the random forest, and the performance of the recall rate is slightly worse. Combining the statistics in the table, the best performing model is the random forest. Looking at Table 4, we can see that in the "conditioning" model, again the random forest performs best in all metrics.

4.2 Results and Analysis of Model Optimization

Using the new datasets sampled by the SMOTE algorithm and trained with the random forest algorithm that performs best on the original data, the parameters of each indicator on the model for the "aggregation" and "conditioning" labels are shown in Table 5.

Table 2. Classification of data samples

Sample Category	Exception Category	Number of set samples	Adjustment the number of samples
0	Normal	1884	2986
1	Not enough	1851	955
2	Excessive	251	135

Table 3. Metrics for each algorithm model ("Set" tab)

Algorithm category	Accuracy	Accuracy rate	Recall Rate	F1 Score
KNN	0.765	0.578	0.544	0.557
Decision Trees	0.860	0.651	0.647	0.648
Random Forest	0.865	0.887	0.656	0.778
XGBoost	0.865	0.825	0.681	0.746
Support vector machines	0.820	0.602	0.601	0.602
Plain Bayesian	0.789	0.626	0.547	0.584

Table 4. Model metrics for each algorithm ("adjustment" tab)

Algorithm category	Accuracy	Accuracy rate	Recall Rate	F1 Score
KNN	0.792	0.740	0.707	0.724
Decision Trees	0.801	0.728	0.744	0.736
Random Forest	0.830	0.803	0.758	0.780
XGBoost	0.822	0.806	0.757	0.782
Support vector machines	0.816	0.784	0.688	0.733
Plain Bayesian	0.786	0.738	0.729	0.734

Table 5. Parameters of each indicator of random forest model

Label Name	Accuracy	Accuracy rate	Recall Rate	F1 Score
Assemblies	0.841	0.734	0.837	0.782
Adjustment	0.871	0.723	0.801	0.760

As shown in Table 4, comparing the parameters of the model before and after the data processing, we can see that the "set" label and the "adjustment" label have improved the model in terms of accuracy and recall rate, but the performance of the accuracy

rate and F1 score is more general. Therefore, the model was further optimized by using the Boosting principle based on multiple sampling of data as mentioned above. The indicators of the “pooling” and “conditioning” labeling model after optimization are shown in Table 6.

Table 6. Parameters of each index of the integrated model

Label Name	Accuracy	Accuracy rate	Recall Rate	F1 Score
Assemblies	0.876	0.868	0.823	0.845
Adjustment	0.911	0.889	0.886	0.887

Comparing Table 5 and Table 6, it can be seen that the integrated model with random forest as the base evaluator, which was sampled three times and trained three times, showed a significant improvement in accuracy. The accuracy of the test set obtained by cross-validation is slightly higher than that of the training set, and the models do not produce any over-fitting phenomenon. The integrated model achieves better classification results on both the "set" and "adjustment" labels.

4.3 Operation Effect Demonstration

Since the ultimate goal of this study is to make the model practical, after the model was built, the pyside2 [4] graphical design tool was used to develop an interface for model use, which facilitates the optometrist to enter patient information and view the classification results predicted by the model.

The interface design of the model, which is mainly divided into two parts. The left interface is the information entry area, where the doctor enters the patient’s characteristics according to the rules of the system and clicks the "Detect" button to start the detection. The right interface is the feedback area of the classification results, where the system makes AI calculations on the entered information and displays the judgment results in the result area within a short time.

5 Conclusion

This paper proposes and designs a classification model to assist ophthalmology clinical staff in diagnosing visual function abnormalities, with an accuracy of 0.876 and 0.911 for the classification of abnormalities in the “Set” label and “adjustment” label, respectively. The classification of visual function abnormalities can help optometrists to diagnose and treat abnormalities quickly, with convenient feature input, fast computation speed, and high accuracy, which can be used in clinical applications. At present, it has been tried in a small range of treatment institutions and has achieved good trial feedback.However, since the number of cases in the “excessive” and “scattered” categories of the two abnormality labels is relatively small, there is a need to collect more sufficient data samples in the future and continue to update and optimize the machine learning model

to further improve the accuracy of the prediction function. The accuracy of the prediction function should be further improved.

In summary, there are rules for the classification and detection of visual function abnormalities, which can be identified and predicted based on the obvious manifestations of characteristic abnormalities, and the improved and optimized model achieves better performance in terms of accuracy and system performance. This study provides a good datasets for the classification of visual function abnormalities using a mechanistic approach, and is expected to provide a decision basis for ophthalmology clinical staff to assist in the diagnosis of visual function abnormalities.

References

1. Department of Publicity. Transcript of the July 13, 2021 press conference of the National Health and Wellness Commission [EB/OL] http://www.nhc.gov.cn/xcs/s3574/202107/2fef24a3b77246fc9fb36dc8943af700.shtml
2. Jiewen, W.: AI technology empowered eye health-optometric new starting point. Vitreous Enamel Optical **49**(05), 44–52 (2021)
3. Zongyuan, G., Wanji, H., Lie, J., et al.: New advances in algorithms for ophthalmic artificial intelligence. J. Shandong Univ. (Med. Ed.) **58**(11), 17–23 (2020)
4. Ying, Z.: Effect of visual function training on visual acuity of children with strabismus and amblyopia. Medical Equipment **33**(24), 122–123 (2020)
5. Hu, H., Wu, J.P.: Design and implementation of a graphical configuration tool for Samba. Electron. Design Eng. **30**(04), 180–183+188 (2022)

A Hybrid Architecture of 5G Private Network Solution for Urban Rail Transportation Systems

Jianming Zhou[1], Haoxian Jing[2(✉)], Lei Ning[2], and Kailei Pang[1]

[1] China United Network Communications Corporation ShenZhen Branch, Shenzhen, China

[2] College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China

201902010223@stumail.sztu.edu.cn

Abstract. Recently, urban rail transportation is constantly developing to intelligent urban rail based on the Internet of Things, artificial intelligence, and high-speed communication. Long Term Evolution for Urban Rail Transportation System (LTE-URTS) is a communication system for urban transportation service, which carries the essential services by urban rail with 4G LTE. However, with the continuous increase of new demands, LTE also faces the shortcomings of insufficient bandwidth and high construction costs. As an advanced solution, networking through 5G can bring better performance to satisfy multiple demands of the modern Urban Rail Transportation System (URTS) than LTE-URTS. This article provides an overview of the characteristics of 5G and LTE-URTS. In addition, we propose a new network architecture solution based on 5G for the URTS with field test results, which have better performance compared to LTE-URTS.

Keywords: 5G · Urban Rail Transportation System (URTS) · Long Term Evolution (LTE) · Multi-access Edge Computing (MEC) · Network Slice (NS)

1 Introduction

Urban rail transportation plays a significant role in modern society as one of the fundamental choices for people to travel and goods to transport. More and more countries would like to develop more intelligent rail transportation to satisfy multiple needs nowadays. Communication-Based Train Control (CBTC) is the most widely used train control system, but it has many problems [1]. It is necessary not only to solve the existing problems but also to empower some new directions

This work was sponsored by Scientific Research Capacity Improvement Project From Guangdong Province (No.2021ZDJS109) , and SZTU Experimental Equipmental Equipment Development Foundation (No.JSZZ202102007).

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 85–94, 2023.
https://doi.org/10.1007/978-981-99-1260-5_11

of rail transportation technologies like the train autonomous circumambulate system (TACS). As a novel generation of mobile communication technology, 5G has provided several outstanding enhancements. With these high-performance enhancements, 5G has lower transmission delay, higher average network throughput rate, and more application scenarios [2,3]. 3GPP has defined three typical application scenarios of 5G:

- eMBB (enhanced Mobile BroadBand): 5G will provide users with better Internet services with its high rate and other characteristics. The mobile network will have a better rate and a more stable transmission capacity. It is suitable for real-time video surveillance systems and mobile live streaming platforms [4].
- mMTC (massive Machine Type of Communication): With the continuous increase of devices on the network, people are no longer satisfied with human-to-human or human-to-machine communication. The Internet of Things(IoT) is an information carrier based on generalized networks, which provides a method for all independently addressed devices to communicate. Low Power Wide Area Networks(LPWAN) is a transmission network for IoT applications with wide coverage and low-power Dissipation. As a representative of LPWAN, Narrow Band Internet of Things(NB-IoT) is a low-power solution in this scenario [5]. 5G enables many IoT devices to access the network with a very low-power dissipation [6,7].
- uRLLC (ultra-Reliable Low-Latency Communications): Nowadays, many scenarios such as driverless driving, industrial robots, and telemedicine require extremely low latency and high reliability. 5G uses network slice, edge computing, and other modes to ensure the rate and quality of communication [8–11].

eMBB can provide low latency and HD video services, such as real-time video surveillance and onboard equipment live broadcast for the URT. mMTC enables rail transportation to perform better by deploying more Rail IoT (RIoT) along the train line. uRLLC allows trains to realize services like autonomous driving. In addition, in order to better support eMBB and uRLLC, 5G introduces New Radio (NR) air interface. It drastically increases the wireless network capacity of 5G and brings new opportunities to URTS [12].

In this paper, we briefly describe some key technologies in 5G in Sect. 2. Then, We highlight wireless communication solution with 4G LTE and list its advantages and disadvantages, we also propose a solution with 5G in Sect. 3.

2 The Key Characteristics of 5G

In this section, we briefly list some critical technologies of 5G

2.1 Massive MIMO

Multiple-Input-Multiple-Output (MIMO) is a multi-antenna technology in which both the transmitter and receiver have an antenna group composed of multiple antenna elements. A communication system can use MIMO to achieve

space diversity and space multiplexing. Space diversity uses multiple antennas to transmit or receive a data stream to enhance communication quality. Space multiplexing is the simultaneous transfer of multiple data streams to the terminal to improve data rate. It is beneficial for a system to increase the whole network throughput, quality, and capacity. Many wireless communication technologies apply MIMO, and MIMO was also evolving to form what is now called massive MIMO [13]. Combining with beamforming, massive MIMO can enhance the overall performance of 5G systems [14].

2.2 Network Slice

5G Network Slice (NS) [15] is an idea for catering to three significant scenarios in 5G: eMBB, mMTC, and uRLLC with just building one 5G network in an area. NS divides a physical 5G network into several end-to-end logic networks. A 5G network NS is a collection of service networks with their resources and functions. 5G NS enables operators to configure the relevant information of the service by slicing, which realizes the effective utilization and management of 5G network resources and isolates critical services to reduce data delay and jitter.

3 5G and Transportation

As a typical public URTS, subways and trains have a lower price and larger transportation volume. However, due to construction costs, road planning, and other factors, the coverage of the railway station is not sufficient in some districts in most cases. Many commuters must take a relatively long walk to reach the station they want and even need to transfer by other traffic tools to arrive at the final destination from the station [16]. In this section, we will introduce an innovative small volumes URTS with 5G.

3.1 Small Volumes URTS

The system uses guideway rubber-tired trams that are suitable for building the extension of metropolitan subways, trunk lines for small-medium cities, travel lines, and airport express. The system has excellent environmental adaptability, a small footprint, a small turning radius, and is suitable for collaborative construction with buildings (Fig. 1). The advantages are as follows:

- As the capillaries of the city: connect to main stations and get through to the last mile.
- A relatively low construction costs with a small transportation volume.
- High security, high intelligence, low energy consumption.

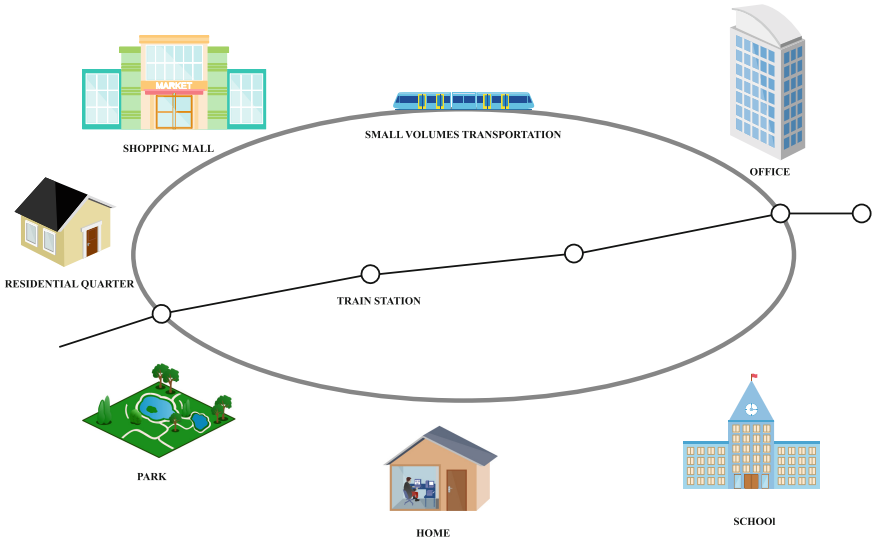


Fig. 1. The applicaiton scenario of small volumes URTS.

3.1.1 LTE-URTS

LTE-URTS is a system, especially for general operations of urban rail transportation, which undertakes many missions, such as CBTC, Passenger Information System (PIS), Video Surveillance System, and Cluster scheduling [17]. Figure 2 illustrates the LTE-URTS demand standards for rail transportation.

3.1.2 Network Solution for LTE-URTS

The original solution is to use LTE to build a private network to carry network services. However, resources in the conventional licensed spectrum are increasingly scarce, which is hard to apply for industry-specific spectrum in 1.4 and 1.8 GHz. In addition, the cost of constructing multiple private networks is high, and the maintenance of wireless networks is complex due to the different systems in different countries. It is necessary to put some data streams into the unlicensed spectrum. Operators can deploy LTE under both licensed and unlicensed spectrum using LTE-U and do not need to change any network structure, enhancing the network capacity and bringing a better experience to users [18]. The solution uses LTE-U to support operations like CBTC, PIS, and Video Surveillance System:

- Service Carrying (two separate networks A and B): Net A is responsible for communicating, and net B undertakes the missions like communication, PIS, Video Surveillance System, and cluster administrating.
- Redundant Coverage: The two-layer network covers the same district simultaneously.
- Redundant Device: Provide core services with board-level redundancy and end-to-end master-slave configuration.

Service Type	Uplink Bandwidth	Downlink Bandwidth	Delay	Packet Loss Probability	Service Statement
GOA3/4 CBTC	0.5Mbps	0.5Mbps	≤150ms	≤1%	Driverless
PIS emergency text	0	0.01Mbps	≤150ms	≤1%	OCC(operation control center) send emergency text to the train
Video Surveillance System	1-6Mbps	0	≤500ms	≤1%	OCC obtains the video of the passenger compartment in the vehicle, according to the two-way 2048K.
Center Video for PIS	0	2-6Mbps	≤500ms	≤1%	OCC transmits live video programs to the train, according to the line 6144K.
Cluster Scheduling(Audio)	0.1Mbps	0.1Mbps	≤150ms	≤1%	Staff for voice transmission
Cluster Scheduling(Video)	1Mbps	1Mbps	≤500ms	≤1%	Staff for video transmission
General	2.6-7.6Mbps	3.61-7.61Mbps			

Fig. 2. The standard of traffic demand for LTE-URTS.

- Unlicensed: Free to apply.
- Wide spectrum: Rich spectrum resources.
- High Landscape Integration: Compared with WLAN, the number of trackside device is reduced by 2/3.

Nevertheless, due to deployment in the unlicensed spectrum, LTE-U will likely interfere with other unlicensed spectrum technologies, such as Wi-Fi. It still needs to use some solutions to avoid the landscape.

3.2 5G Solution

Nowadays, passengers have higher requirements for the train ride experience. PIS is no longer just showing passengers simple information like the arrival time of the vehicle but also provides more information with high real-time demand, such as live broadcasts of events, train status, and even alarm information under emergency conditions. In modern society, surfing the Internet is the primary demand for people, and the train needs to provide a high-speed Internet environment for passengers [19]. In addition, it is necessary to monitor the vehicle's state through the real-time video monitoring system to ensure the safety of passengers. These requirements require ample bandwidth, low latency, and stability, which is difficult to achieve under the LTE scheme, while 5G is a better solution. Figure 3 illustrates the requirements analysis of Rail Transportation in the wireless application scenario with 5G.

Service Type	Priority	Uplink Bandwidth	Downlink Bandwidth	Delay	Packet Loss Probability	Concurrent terminals
Control System	Foremost	1Mbps	1Mbps	≤25ms	≤1%	2
Cluster Scheduling(audio)	Foremost	1Mbps	1Mbps	≤150ms	≤1%	7
Emergency video	High	2-6Mbps	2-6Mbps	≤150ms	≤1%	1
Video Surveillance System	Medium	16—60Mbps		≤500ms	≤1%	4
PIS	High		8—10Mbps	≤500ms	≤1%	1
Wi-Fi for passengers	Low					2
Cluster Scheduling(video)	High	6—12Mbps	6—12Mbps	≤500ms	≤1%	5
General		26-80Mbps	18-30Mbps			

Fig. 3. The requirement analysis of URTS in the wireless application scenario with 5G

3.2.1 Overall Architecture

The system achieves full coverage through 5G, with a 5G base station every 1.2 km along the track. In addition, there is going to deploy 5G room sub-equipment to satisfy the demand for 5G of the station. All the base stations are in charge of transmitting data to the carrier transmission network and bringing public data onto the Internet. MEC nodes in the operator center use User Plane Function (UPF) to offload the data and connect PIS, Video Surveillance System, and cluster data to the private cloud to ensure data security. There are three major parts to the operator transmission network: access, transmission, and MEC. The access network is responsible for establishing the wireless link. The network transmission completes the wired connection part of the access network to the core network and the cloud rail data center. MEC realizes data offload to ensure the security of operation-related data of tracks.

3.2.2 Network Architecture

Customer premises equipment (CPE) and the base station mainly constitute the Communication System Between Train And Ground. All data pass through the vehicular network and access the CPE, CPE will establish a link with the base station to satisfy the wireless communication between train and ground. The vehicular network is built using a vehicular ring network, which deploys a CPE active and standby redundancy in the front and rear of the vehicle. It makes a strategy on the vehicle server to ensure that the backup CPE can communicate with the ground when the communication of the primary CPE is abnormal. The vehicular AP, camera and PIS are connected to the CPE through the vehicular

ring network to complete communication with the ground. Due to the occlusion of the platform area, it is necessary to add an additional 5G room sub-equipment to complete the platform's coverage, and the platform's passengers can access the Internet through the 5G network.

3.2.3 System Functions

- 5G back transmission for PIS: With the help of 5G and the system's specific network architecture, the train can provide users with 4K video services and real-time train information to improve their travel experience further.
- Real-time Video Surveillance System: It is necessary to build a real-time Video Surveillance System to monitor the state of trains, which can drastically enhance the quality of train management and the safety of the train and passengers. The traditional practice of onboard Video Surveillance System systems uses Wi-Fi for wireless transmission or a hard disk for data copy while the train arrives at the station. This method does not have real-time performance, which leaves security risks for the operation of the train. However, using 5G makes it easy for a monitoring system to transmit real-time video with low latency because of the characteristic of large bandwidth.
- 5G for onboard networking: Because of the development of mobile phones, passengers now have more demands while traveling onboard, such as watching high-quality videos online and playing games. However, the traditional Communication System Between Train And Ground can not provide a wireless network to meet the needs of passengers for a high-speed network limited by bandwidth performance. Passengers can use 5G for high-speed networking through the network architecture in this new system.
- 5G public network cluster: It is common to use the wireless cluster for train scheduling. However, the traditional narrowband group needs an independent network with a small bandwidth and can not fully meet the scenario for train scheduling and personal communication. Using 5G can provide a new clustering idea for rail transportation.

Figure 4 is the schematic of the overall architecture (HQVSS: HD Video Surveillance System, 5G NSS: 5G Network Scheduling System, 5G HSSS: 5G High-Speed Storage System). Figure 5 illustrates the 5G network slice solution for URTS.

4 Field Test Results

In order to test the performance of the transportation system, we use 5G mobile phones to do 5G SA upload with the system deployed by a dual-operator networking in different frequencies. The band of 5G is 3.5 GHz, and the center frequency point is 3509.76 MHz, which is different from other public networks in the testing district. Table 1 and Table 2 illustrates the statistic of the test.

The field test result proves that the proposed 5G network solutions of the transportation line can reduce interference while guaranteeing a smooth handover sequence and a stable service rate.

Table 1. The statistics of the uplink performance

Index	Max	Min	Average
RSRP (dBm)	−44.00	−103.69	−72.91
SINR (dB)	37.25	−8.31	16.65
RSRQ (dB)	−9.19	−20.63	−10.72
UL Throughput (Mbps)	291.58	5.9	153.35

Table 2. The statistics of the field networking results

Index	Value
Total Distance (km)	4.86
Total Duration (s)	1643.25
NR Access Delay (ms)	152.859
NR Switching Success Rate (%)	100.00
NR Handoff Control Layer Delay (ms)	19.448
Failure Probability (%)	0.00
Total Uploaded Data (MB)	30298.36
Total Uploaded Time Cost (s)	1635.50

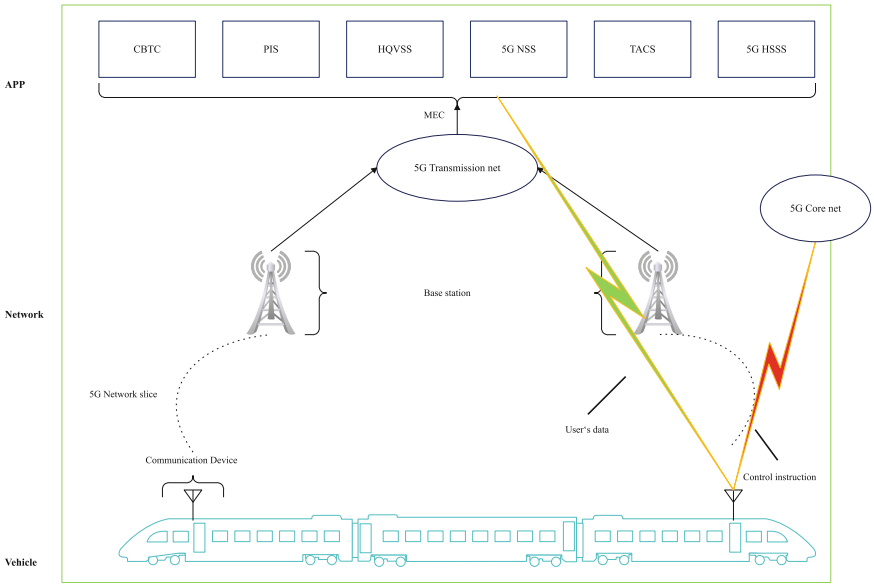


Fig. 4. The overall architecture of the small volumes URTS.

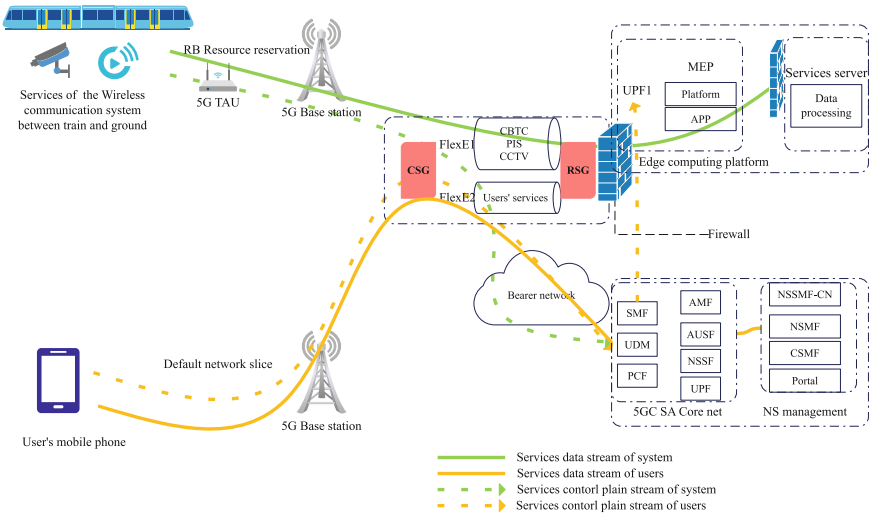


Fig. 5. The diagram of the 5G network slice solution for URTS.

5 Conclusion

It is necessary to develop intelligent URTS which can significantly improve the passenger travel experience. In the early stages of commercial applications of 5G, scenarios like uRLLC and mMTC are not mature, and 5G complements LTE-URTS as an additional technology, such as eMBB. It is believed that with the constant maturity of 5G in commercial scenarios, 5G can bring more vitality to the URTS in the future.

References

1. Zhao, J., Liu, J., Yang, L., Ai, B., Ni, S.: Future 5G-oriented system for urban rail transit: opportunities and challenges. China Commun. **18**(2), 1–12 (2021)
2. Navarro-Ortiz, J., et al.: A survey on 5G usage scenarios and traffic models. IEEE Commun. Surv. Tutorials **22**(2), 905–929 (2020)
3. Pisarov, J., Mester, G.: The impact of 5G technology on life in 21st century. IPSI BgD Trans. Adv. Res. (TAR) **16**(2), 11–14 (2020)
4. Abdullah, D.M., Ameen, S.Y.: Enhanced mobile broadband (embb): a review. J. Inf. Technol. Informat. **1**(1), 13–19 (2021)
5. Gbadamosi, S.A., Hancke, G.P., Abu-Mahfouz, A.M.: Building upon NB-IoT networks: a roadmap towards 5G new radio networks. IEEE Access **8**, 188641–188672 (2020)
6. Chettri, L., Bera, R.: A comprehensive survey on internet of things (IoT) toward 5G wireless systems. IEEE Internet Things J. **7**(1), 16–32 (2019)
7. Wijethilaka, S., Liyanage, M.: Survey on network slicing for internet of things realization in 5G networks. IEEE Commun. Surv. Tutorials **23**(2), 957–994 (2021)

8. Addad, R.A., Taleb, T., Flinck, H., Bagaa, M., Dutra, D.: Network slice mobility in next generation mobile systems: challenges and potential solutions. IEEE Network **34**(1), 84–93 (2020)
9. Kekki, S., et al.: Mec in 5G networks. ETSI White Paper **28**(28), 1–28 (2018)
10. Hassan, N., Alvin Yau, K.-L., Wu, C.: Edge computing in 5G: a review. IEEE Access **7**, 127276–127289 (2019)
11. Pham, Q.-V., et al.: A survey of multi-access edge computing in 5G and beyond: fundamentals, technology integration, and state-of-the-art. IEEE Access **8**, 116974–117017 (2020)
12. Fodor, G., et al.: 5G new radio for automotive, rail, and air transport. IEEE Commun. Mag. **59**(7), 22–28 (2021)
13. Lu, L., Li, G.Y., Swindlehurst, A.L., Ashikhmin, A., Zhang, R.: An overview of massive MIMO: benefits and challenges. IEEE J. Selected Topics Signal Process. **8**(5), 742–758 (2014)
14. Rao, L., Pant, M., Malviya, L., Parmar, A., Charhate, S.V.: 5G beamforming techniques for the coverage of intended directions in modern wireless communication: in-depth review. Int. J. Microwave Wireless Technol. **13**(10), 1039–1062 (2021)
15. Zhang, S.: An overview of network slicing for 5G. IEEE Wireless Commun. **26**(3), 111–117 (2019)
16. Lunke, E.B.: Commuters satisfaction with public transport. J. Transp. Health **16**, 100842 (2020)
17. Tang, T., Dai, K., Zhang, Y., Zhao, H., Jiang, H.: Field test results analysis in urban rail transit train ground communication systems of integrated service using LTE-M. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pp. 2017–2021. IEEE (2016)
18. Zhang, R., Wang, M., Cai, L.X., Zheng, Z., Shen, X., Xie, L.-L.: LTE-unlicensed: the future of spectrum aggregation for cellular networks. IEEE Wireless Commun. **22**(3), 150–159 (2015)
19. Yuan, T., ZhiWei, H.: Research and application of 5G collaborate carrying for metro passenger information system. In: Proceedings of the 5th International Conference on Electrical Engineering and Information Technologies for Rail Transportation (EITRT) 2021: Rail Transportation Information Processing and Operational Management Technologies, vol. 867, p. 316. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-9909-2_35

Fault Diagnosis of the Power Transformer Based on PSO-SVM

Xinxin Zhou[1], Zhanshuang Liu[1], Zhigang Shi[1], Le Ma[1], Hang Du[1], and Di Han[2(✉)]

[1] State Grid Qinghai Extra High Voltage Company, Xining 810021, China
[2] Advanced Smart Electric Information Technology, Hangzhou 310000, China
handi202111@163.com

Abstract. Transformers play an important role in power systems. Dissolved gases analysis (DGA) has been widely used in the transformer fault diagnosis. A novel fault diagnosis method using support vector machine (SVM) with particle swarm optimization (PSO) algorithm is developed for transformer in this paper. To enhance the ability and performance of SVM, the particle swarm algorithm is used to optimize the parameters of SVM in this study. The proposed model is named PSO-SVM. Then, the optimized model was applied to identify and classify the faults. The diagnostic accuracy of the proposed PSO-SVM model reached 82%. Results demonstrate that the superior performance of the PSO-SVM, compared with the ordinary support vector machine classifier.

Keywords: Transformer · fault diagnosis · particle swarm optimization · support vector machine

1 Introduction

Transformers are one of the most important hubs in the transmission and conversion of electrical energy in the power grid, so it is important to ensure their safe operation. To reduce the operation and maintenance costs of transformer stations, various new fault diagnosis technologies are constantly applied to transformer. The structure of power transformer is complex and the safety hazards are many and not easy to find. To find out the potential pitfalls of transformers quickly and accurately, the most used diagnostic method is Dissolved Gases Analysis (DGA) [1–3]. The idea behind of DGA is based on the concentration of dissolved gases in transformer oil to make a diagnosis and prediction.

Recently, since artificial neural networks (ANNs) can provide effective feature representations from data, thus it has received much attentions in the field of fault diagnosis [4]. Mohamed et al. presented an artificial neural network (ANN)-based scheme to detect and locate the fault for power transformers. [5]. In the study of Sun et al. [6], to reduce the training time and seek global minima, an efficient back propagation with adaptive learning rate and momentum coefficient algorithm was proposed. However, there are some limitations for deep learning-based power transformer fault diagnosis methods,

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 95–102, 2023.
https://doi.org/10.1007/978-981-99-1260-5_12

such as high computation cost, sensitivity to noise, effect of different environmental conditions [7, 8].

Unlike ANNs, support vector machine (SVM) which was developed by Vapnik [9] based on statistical learning theory, is a popular and widely used machine learning method. SVM can successfully deal with the overfitting problems, low convergence rate and local optimal solutions in ANN. Therefore, SVM has been successfully applied to transformer fault diagnosis [10, 11, 15]. However, the classification performance of SVM will degrade with the inappropriate selections of penalty factor and kernel function. In view of the limitations of the current DGA based transformer fault diagnosis algorithm, the particle swarm optimization (PSO) algorithm and SVM classification algorithm are combined to offer optimal SVM model for fault diagnosis in this work.

2 Brief Review of SVM and PSO

2.1 SVM

The Support vector machine (SVM) is commonly used as a generalized classifier for the classification of data. The principle is to divide the training data set to construct a hyperplane so that the data collected from different categories are as far away from the hyperplane as possible [9]. As shown in Fig. 1, The 2-dimensional dimensional training sample set is $\{(x_n, y_n), n = 1, 2 \cdots, e\}$, and the circles and triangles in Fig. 1 represent different classes of sample points with a total of e.

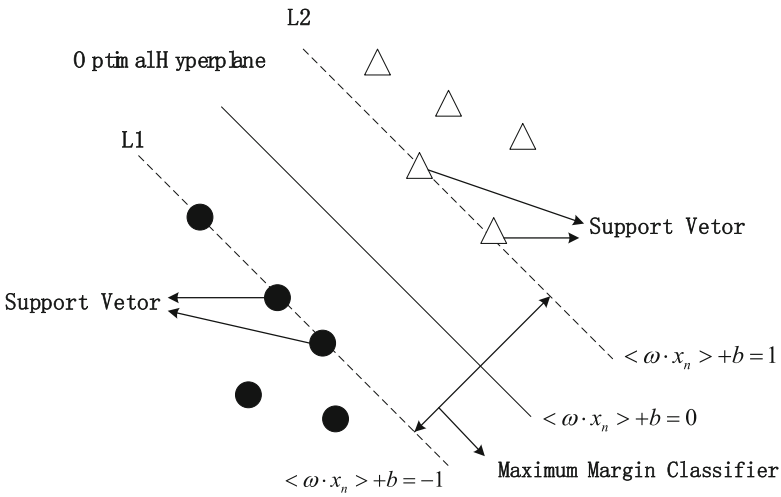


Fig. 1. Optimal classification hyperplane in SVM.

As shown in Fig. 1, two classes of sample points are separated by the optimal classification hyperplane $\omega \cdot x_1 + b = 0$.x_1 is the point on the hyperplane, ω is the hyperplane normal vector, b is a constant term, and L_1 and L_2 are the sample planes with the shortest linear distance and parallel to the optimal classification hyperplane, respectively. For

the linear indistinguishable problem, a nonnegative relaxation factor is introduced to improve the learning ability of the support vector machine with a minimum objective function of,

$$m\phi(\omega) = \|\omega\|^2/2 + C \cdot \sum\nolimits_{n=1}^{e} \xi_n \tag{1}$$

where C is the error penalty factor to determine the acceptable level of the classification error of the sample data. The popular Gaussian radial basis kernel function is employed as the kernel function of the support vector machine, and the decision function of the support vector machine model is,

$$f(x) = \operatorname{sgn}\left[\sum\nolimits_{n=1}^{e} y_n \lambda_n K(x, x_n) + b\right] \tag{2}$$

$$K(x, x_n) = \exp\left(-\frac{\|x - x_n\|^2}{2\sigma^2}\right) \tag{3}$$

$$g = \frac{1}{2\sigma^2} \tag{4}$$

where $K(x, x_n)$ is denoted to represent the inner product of the test sample and the n th training sample after mapping to the feature space. λ_n is the Lagrangian factor and d is the kernel function bandwidth.

2.2 PSO

The PSO [12–14] is a popular heuristic optimization algorithm, since it has fewer parameters and fast convergence speed. In PSO, the number of particles, initial speed, and fitness are defined in a specific search space. The best position in the set space is iteratively searched through individual and global mechanisms to update the position. The PSO algorithm is briefly elaborated as follows.

In a N-dimensional search space, a population $X = (X_1, X_2, \ldots X_m)$ is composed of m particles. And $X_i = (x_{i1}, x_{i2}, \ldots x_{im})$ denotes the position of the ith particle in the population in space. Denote $p_i = (p_{i1}, p_{i2}, \ldots p_{im})$ as the best position passed by the ith particle. By denoting the corresponding flight speed by $V = (v_{i1}, v_{i2}, \ldots, v_{im})$, and the best position through the whole group passes by $p_g = \left(p_{g1}, p_{g2}, \ldots, p_{gm}\right)$, then

$$v_{id}(t+1) = v_{id}(t) + c_1 r_1 (p_{id}(t) - x_{id}(t)) + c_2 r_2 \left(p_{gd}(t) - x_{id}(t)\right) \tag{5}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{6}$$

where $d = 1, 2, 3 \ldots n$ is the population dimension, $i = 1, 2, 3 \ldots m$ is the population size, t is the current iteration number, c_1 and c_2 are the learning factors, r_1, r_2 are the random values which obey uniformly distribution between (0,1). $v_{id} \in [-v_{\max}, v_{\max}]$, maximum speed $v_{\max} = k \cdot x_{\max}$ $(0.1 \le k \le 1)$.

3 Transformer Fault Diagnosis Method Based on PSO-SVM

The data set of the collected dissolved gas data is first pre-processed and divided into the two parts including training set and test set. The DGA data is normalized as,

$$x_{in} = \frac{x_i - x_{min}}{xmin_{max}} \tag{7}$$

where x_i is the ith sample data of each indicator, and x_{max} and x_{min} are the maximum and minimum values of each indicator, respectively. The six states of transformer are coded. Then, the conventional SVM and the PSO-SVM are trained using the training data set, respectively. Finally, the effectiveness of the SVM and PSO-SVM models is evaluated using the test data set.

Two parameters of SVM including the penalty factor C and the variance of the kernel function g, are to be determined for data classification. Since the dissolved gas data are nonlinear indistinguishable, a Gaussian kernel is introduced to map the training data to a high-dimensional space. The parameters C and g play an important role in the classification accuracy of the SVM. In the proposed PSO-SVM, the optimal C and g are derived using PSO. The procedure of the proposed PSO-SVM is described in Fig. 2. The details of the algorithm are as follows,

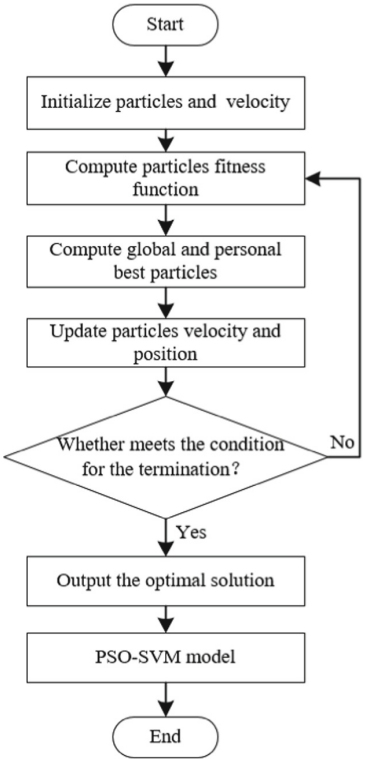


Fig. 2. The flowchart of PSO-SVM algorithm.

Step 1. Determine the fitness function for PSO, and use the accuracy of fault diagnosis of the training dataset as the fitness function.

Step 2. Calculate the fitness value of each particle as the individual extremum of the current particle and the largest individual extremum as the global extremum.

Step 3. Update the velocity and position of the selected particles and determine whether the current state meets the end condition. If it satisfies, end. If not, continue to update the particle position and velocity.

Step 4. Update the individual poles and global poles until the end of the search, and output the current search result penalty factor C and the kernel function parameter g to form the two-dimensional parameter search space of the particle.

4 Case Study

To assess the ability and effectiveness of the proposed PSO-SVM, 188 dissolved gas content samples were collected under for 6 fault scenarios for experiments. The data set contains 5 types of dissolved gas content, some of which are shown in Table 1. For the convenience of experiments, the fault types are coded for labelling, as shown in Table 2. The data set is divided into training and test sets for model training and testing the performance of the proposed model, respectively. The test set and the training set account for 50% of the samples.

Experiments were conducted using Matlab. In the model, two learning factor parameters are set as $C_1 = 1.7$ and $C_2 = 1.7$ in the PSO algorithm, respectively. The number of iterations of the PSO-SVM algorithm is set as 50, with the number of particles pop = 20. Figure 3 displays the iterative optimization search process of PSO-SVM. From the data in Fig. 3, accompanying with the increase in the number of iterations, the adaptation degree also becomes larger continuously and reach the maximum at the 42nd time. The fault diagnosis accuracy reaches 86.17%.

Table 1. Description of transformer fault data.

No. Sample	Dissolved gas content					Fault Type
	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2	
1	568	26.5	6.9	2.1	0	Partial discharge
2	3433	180	34.4	3.6	0.4	Partial discharge
3	2083	85.6	18.4	2.6	0	Partial discharge
4	568	26.5	6.9	2.1	0.1	Partial discharge
5	361	29.56	2.74	25.69	197.01	Low energy discharge
…	…	…	…	…	…	…
93	32	41.6	10	120	2.6	High Temperature Overheating
94	43	50.9	11.6	127	2.9	High Temperature Overheating

Table 2. Codes for transformer fault type.

Fault type	Partial discharge	Low energy discharge	High energy discharge	Normal state	Medium to low temperature overheating	High temperature overheating
Fault code	0	1	2	3	4	5

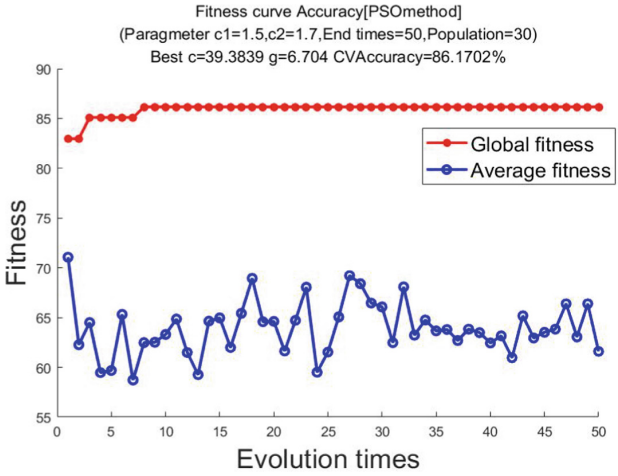
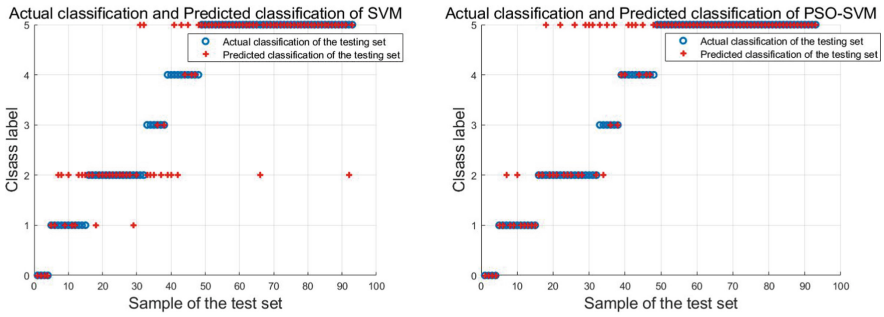


Fig. 3. The iterative optimization search process of PSO-SVM.

To show the advantages of the PSO-SVM model, the comparison study with SVM classifier is carried out. The test samples are normalized and fed to the two classifiers. The classification results are shown in Fig. 4(a) and Fig. 4(b), where the symbols " + " are the predefined label values and the symbols "o" are the actual classification results. From Fig. 4, it can be observed that the classification results of the proposed PSO-SVM model are more promising and reliable, as can be seen from Fig. 4. Table 3 lists the diagnostic accuracy of the comparable methods.

As plotted in Fig. 4 and Table 3, the accuracy of the PSO-SVM model algorithm is 81.72%, which is a significant improvement compared to the accuracy of 75.26% of the traditional SVM.



(a) (b)

Fig. 4. Comparison results of different transformer fault diagnosis.

Table 3. Comparison results of different transformer fault diagnosis.

Method	Training accuracy(%)	Test accuracy(%)
PSO-SVM	94.68	81.72
SVM	90.42	75.26

5 Conclusion

A new transformer fault diagnosis method based on the combination of PSO and SVM is proposed in this work. The issues of low accuracy, weak learning ability of the traditional SVM algorithm are addressed by introducing PSO. Simulation experiments are conducted using actual DGA data from a real transformer, and the experimental results illustrate that the proposed PSO-SVM model has a comprehensive correct diagnosis result of 81.72%, which verifies the accuracy and practicality of the model.

Acknowledgements. This work was supported by the key technology project of the State Grid Corporation of China under Grant 522821200090.

References

1. Seifeddine, S., Khmais, B., Abdelkader, C.: Power transformer fault diagnosis based on dissolved gas analysis by artificial neural network. In: 2012 First International Conference on Renewable Energies and Vehicular Technology, pp. 230–236. IEEE (2012). DOI:https://doi.org/10.1109/REVET.2012.6195276
2. Bacha, K., Souahlia, S., Gossa, M.: Power transformer fault diagnosis based on dissolved gas analysis by support vector machine. Electric Power Syst. Res. **83**(1), 73–79 (2012). https://doi.org/10.1016/j.epsr.2011.09.012
3. Ghoneim, S., Shoush, K.A.: Diagnostic tool for transformer fault detection based on dissolved gas analysis. Adv. Electr. Eng. Syst. (3), 152–156 (2014). DOI:https://doi.org/10.9790/1676-09532026

4. Zhang, Y., Ding, X., Liu, Y., et al.: An artificial neural network approach to transformer fault diagnosis. IEEE Trans. Power Delivery **11**(4), 1836–1841 (1996). https://doi.org/10.1109/61.544265
5. Mohamed, E.A., Abdelaziz, A.Y., Mostafa, A.S.: A neural network-based scheme for fault diagnosis of power transformers. Electric Power Syst. Res. **75**(1), 29–39 (2005). https://doi.org/10.1016/j.epsr.2004.11.013
6. Sun, Y., Zhang, S., Miao, C., et al.: Improved BP neural network for transformer fault diagnosis. J. China Univ. Min. Technol. **17**(1), 138–142 (2007). https://doi.org/10.1016/S1006-1266(07)60029-7
7. Afrasiabi, M., Afrasiabi, S., Parang, B., et al.: Power transformers internal fault diagnosis based on deep convolutional neural networks. J. Intell. Fuzzy Syst. **37**(1), 1165–1179 (2019). https://doi.org/10.3233/JIFS-182615
8. Kim, S., Jo, S.H., Kim, W., et al.: A semi-supervised autoencoder with an auxiliary task (SAAT) for power transformer fault diagnosis using dissolved gas analysis. IEEE Access **8**, 178295–178310 (2020). https://doi.org/10.1109/ACCESS.2020.3027830
9. Vapnik, V.N.: The nature of statistical learning theory. Springer, New York (1995). https://doi.org/10.1007/978-1-4757-2440-0
10. Fei, S., Zhang, X.: Fault diagnosis of power transformer based on support vector machine with genetic algorithm. Expert Syst. Appl. **36**(8), 11352–11357 (2009). https://doi.org/10.1016/j.eswa.2009.03.022
11. Kari, T., Gao, W., Zhao, D., et al.: Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm. IET Gener. Transm. Distrib. **12**(21), 5672–5680 (2018). https://doi.org/10.1049/iet-gtd.2018.5482
12. Tang, W.H., Goulermas, J.Y., Wu, Q.H., et al.: A probabilistic classifier for transformer dissolved gas analysis with a particle swarm optimizer. IEEE Trans. Power Delivery **23**(2), 751–759 (2008). https://doi.org/10.1109/TPWRD.2008.915812
13. Zhang, L., Yuan, J.: Fault diagnosis of power transformers using kernel based extreme learning machine with particle swarm optimization. Appl. Math. Inf. Sci. **9**(2), 1003 (2015). https://doi.org/10.12785/amis/090251
14. Cao, Z., Chen, C., Chen, Y., et al.: Application of particle swarm optimization algorithm in power transformer fault diagnosis. J. Phys. Conf. Ser. IOP Publishing, **1624**(4), 042022 (2020). https://doi.org/10.1088/1742-6596/1624/4/042022
15. Liu, T., Wang, Z.: Design of power transformer fault diagnosis model based on support vector machine. International symposium on intelligent ubiquitous computing and education. IEEE **2009**, 137–140 (2009). https://doi.org/10.1109/IUCE.2009.59

Data Acquisition and Performance Analysis on Discharge Process of Wide Temperature Resistant Battery

Jinmao Chen, Wanli Xu, Xudong Wang, Zixu Zhao, and Chunhua Xiong(✉)

Institute of System Engineering, Academy of Military Science, Beijing 100300, China
energyxch@163.com

Abstract. For lithium ion battery discharge performance of different problems in different ambient temperature, with a company of a new wide temperature resistant ternary lithium ion battery as test object, established a new method of testing, set up to −40 °C, −25 °C, 25 °C and 55 °C four different temperature conditions, with high and low temperature testing system for the ternary lithium battery charging and discharging 1 c to charge and discharge test, Recorded and processed data. The results show that the energy density of the battery reaches 244.12Wh/kg at room temperature of 25 °C. At −25 °C and −40 °C, the average energy density decreases to 77.47% and 50.24% of that at room temperature, respectively. The energy density decreases to 97.35% of the normal temperature condition after the high temperature charge retention and capacity recovery at 55 °C.

Keywords: Wide temperature resistant battery · Lithium ion battery · Energy density · The depth of discharge

1 Introduction

As a comprehensive battery system at present, lithium ion battery has the advantages of small volume, high specific energy, no memory effect, light weight, no pollution and high cycle life, and has been widely used in hybrid vehicles, information technology, aerospace and other aspects, with great demand. However, lithium-ion batteries are greatly affected by ambient temperature, and their optimal operating temperature range is generally 15–35 °C. When the temperature is below 0 °C, the performance of batteries will be greatly reduced [1–3]. Gao Chen et al. conducted discharge tests on lithium-ion batteries at low temperature and compared them with normal temperature. The results showed that the discharge performance of lithium-ion batteries decreased significantly at −15°C, and the capacity at 1C discharge was only 34.1% of that at normal temperature [4]. However, when the battery is in a high temperature environment, the side reaction rate in the battery will increase along with the decomposition of the electrolyte, affecting the service life of the battery [5].

In order to study the discharge characteristics of a new type of lithium-ion battery with wide temperature resistance at different temperatures, various comparative tests of the

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 103–110, 2023.
https://doi.org/10.1007/978-981-99-1260-5_13

discharge depth and energy density of the lithium-ion battery were carried out at 55 °C, 25 °C, −25 °C and −40 °C, respectively, to investigate the influence of temperature on the discharge performance of the battery with wide temperature resistance.

2 Experimental Methods

This part introduces the test device and data acquisition platform, battery to be tested and test process respectively.

2.1 Test Device and Data Acquisition Platform

The structure of the high and low temperature charge and discharge test system is shown in Fig. 1. The battery charge and discharge test equipment in the figure is the energy recovery battery test system Chroma 17020, which can simultaneously test voltage, current, energy, capacity and temperature, etc., the highest voltage is 20 V, the maximum current is 400 A, the test accuracy is 0.001; The ultra-low temperature environment simulation test chamber is SDJ710FA high and low temperature humid heat chamber, the highest temperature is 150 °C, the lowest temperature is −70 °C, the accuracy is ±1 °C.

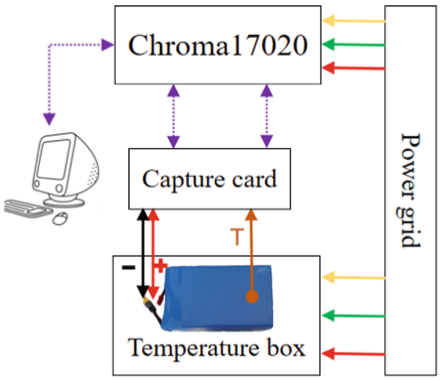


Fig. 1. Schematic diagram of high and low temperature test system

2.2 Test Data Collection Object

The experimental object is a new type of ternary lithium battery with width and temperature resistance, and its main technical parameters are shown in Table 1.

2.3 Test Process and Data Collection Content

The battery is connected to the test interface and sampling interface of the battery charge and discharge test equipment through the special battery fixture and sensor, and different

Table 1. Technical parameters

Project	Parameter
Nominal capacity	10.8 Ah
Nominal voltage	3.7 V
Mass	170 g ± 1 g
Charging cut-off voltage	4.4 V
Charging cut-off current	0.5 A
Maximum charge rate	1C
Discharge cut-off voltage	2.5 V

rates of discharge experiments are carried out in the high and low temperature humid heat chamber. The discharge voltage, charge and discharge current and battery temperature can be detected during the test.

Three identical batteries were selected for three tests of low temperature (−25 °C low temperature, −40 °C ultra-low temperature) and high temperature (55°C high temperature). The test process and data collection content are shown in Table 2 and Table 3.

Table 2. Test process and data collection content (low temperature)

Steps	Temperature	Experiment content	Data acquisition
1	25 °C	1C Discharge to 2.5 V, then stand for 1h	—
2		1C Charge to 4.4 V, then charge at constant voltage to 0.05C and stand for 1 h	Voltage, current, temperature
3		1C Discharge to 2.5 V, then stand for 1 h	Voltage, current, temperature
4		1C Charge to 4.4 V, then charge at constant voltage to 0.05C and stand for 1 h	Voltage, current, temperature
5	−25 °C/−40 °C	Stand for 12 h	temperature
6		1C Discharge to 2.5 V, then stand for 1 h	Voltage, current, temperature
7		1C Charge to 4.4V, then charge at constant voltage to 0.05C and stand for 1h	Voltage, current, temperature
8		1C Discharge to 2.5 V	Voltage, current, temperature

Table 3. Test process and data collection content (high temperature)

Steps	Temperature	Experiment content	Data acquisition
1	25°C	1C Discharge to 2.5 V, then stand for 1 h	—
2		1C Charge to 4.4 V, then charge at constant voltage to 0.05C and stand for 1 h	Voltage, current, temperature
3		1C Discharge to 2.5 V, then stand for 1 h	Voltage, current, temperature
4		1C Charge to 4.4 V, then charge at constant voltage to 0.05C and stand for 1 h	Voltage, current, temperature
5	55°C	Charge retention at high temperature for 7d	temperature
6	25°C	stand for 1 h	temperature
7		1C Discharge to 2.5 V, then stand for 1 h	Voltage, current, temperature
8		1C Charge to 4.4 V, then charge at constant voltage to 0.05C and stand for 1 h	Voltage, current, temperature
9		1C Discharge to 2.5 V	Voltage, current, temperature

3 Data Processing and Analysis

In order to study the effect of charging and discharging temperature for the depth of discharge, the experiment system, we use Fig. 1 respectively according to low temperature and ultra-low temperature and high temperature testing conditions for battery charging and discharging test, and in the process of test data collection and analysis, the condition of different charge and discharge and temperature was obtained to devolve power voltage and discharge depth, the relationship between the low temperature set by shown in Fig. 2.

3.1 Discharge Performance Data Processing and Analysis of Batteries at Low Temperature

According to the charge and discharge curve at room temperature, when the battery is discharged at room temperature, the discharge voltage will experience three stages: rapid decline at first, steady decline at last, and rapid decline at last. However, under the conditions of low charge and ultra-low charge, there is a phenomenon of voltage rise and then decrease between the two stages of rapid decline and gradual decline of discharge voltage, while the phenomenon of low temperature charge and discharge is small. Among them, the discharge voltage of normal temperature charge and ultra low temperature discharge has the largest rebound range, up to 0.36 V.

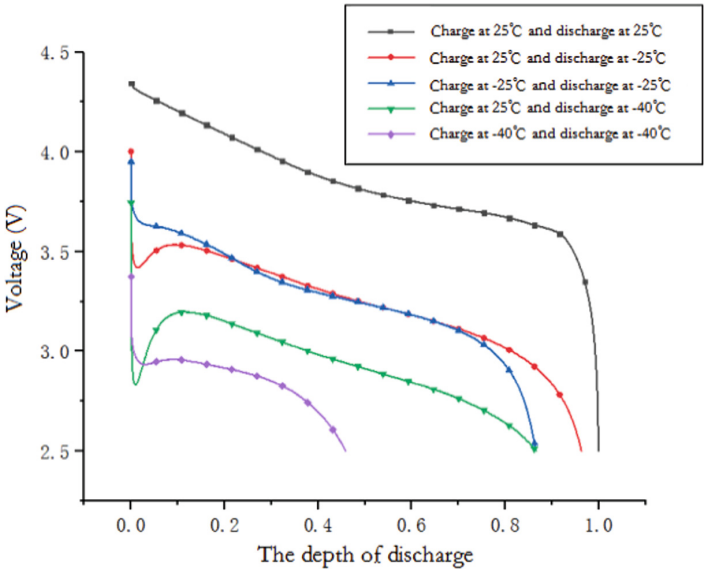


Fig. 2. Relationship between discharge depth and discharge voltage at different low temperature and charging and discharging conditions

The reason for this phenomenon is mainly related to the utilization rate of active substances inside the battery [6–8]. Due to the low temperature discharge of the battery, the internal active substances can not be fully used, the electrode polarization is serious, the battery internal resistance is large, so the initial discharge voltage of the battery drops rapidly. With the discharge, due to the large internal resistance of the battery, a lot of heat is generated inside the battery, so that the battery temperature rises rapidly, so that the active substance of the battery is activated, so the battery discharge voltage begins to rise. As the temperature of the battery rises, the internal resistance of the battery starts to fall, and the heat generated decreases, and the rate of temperature rise decreases, resulting in a decrease in the discharge voltage of the battery. As the charge at normal temperature is greater than that at (ultra) low temperature, the initial value of discharge voltage, voltage drop range and voltage rebound range of normal temperature charge and (ultra) low temperature discharge are all greater than that of (ultra) low temperature charge and discharge.

3.2 Discharge Performance Data Processing and Analysis of Battery Under High Temperature Conditions

Since the treatment method of high temperature condition is different from that of low temperature condition in the test, it is discussed separately, and the battery charge and discharge test is still carried out with the experimental system in Fig. 1 and data collection is carried out. The relationship between battery energy density before and after high temperature treatment is obtained, as shown in Fig. 3.

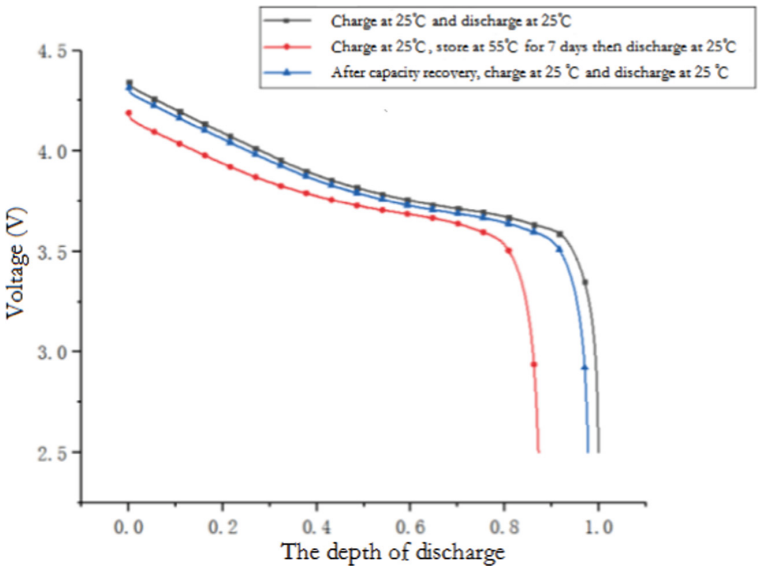


Fig. 3. Relationship between discharge voltage and discharge depth before and after high temperature treatment

The experimental results show that the charge retention rate and capacity recovery rate of the battery are 87.20% and 97.78%, respectively. The reason for the difference between the values of charge retention curve and capacity recovery curve is that during the storage at 55 °C for 7 days, the battery not only suffers from permanent loss due to high temperature, but also has self-discharge phenomenon, which leads to the reduction of electric quantity and the reduction of initial discharge voltage and discharge depth during subsequent discharge at room temperature. The difference between the capacity recovery curve and the original state is the permanent loss of battery capacity caused by high temperature of 55 °C [9–11].

3.3 Processing and Analysis of Energy Density Data

In order to study the influence of different temperatures and charging and discharging conditions on battery energy density, the experimental system in Fig. 1 was used to test the energy density under the above different conditions, and the energy density under different temperatures and charging and discharging conditions was obtained, as shown in Fig. 4.

The experimental results shown that both high temperature and low temperature reduce the energy density. When the low temperature test was conducted, the lower the temperature, the lower the energy density was, and the energy density of normal temperature charge and (ultra) low discharge was higher than that of (ultra) low temperature charge and discharge. The energy density of normal temperature charge and low temperature discharge was reduced by 19.04%, 26.02%, 33.74% and 65.78%, respectively,

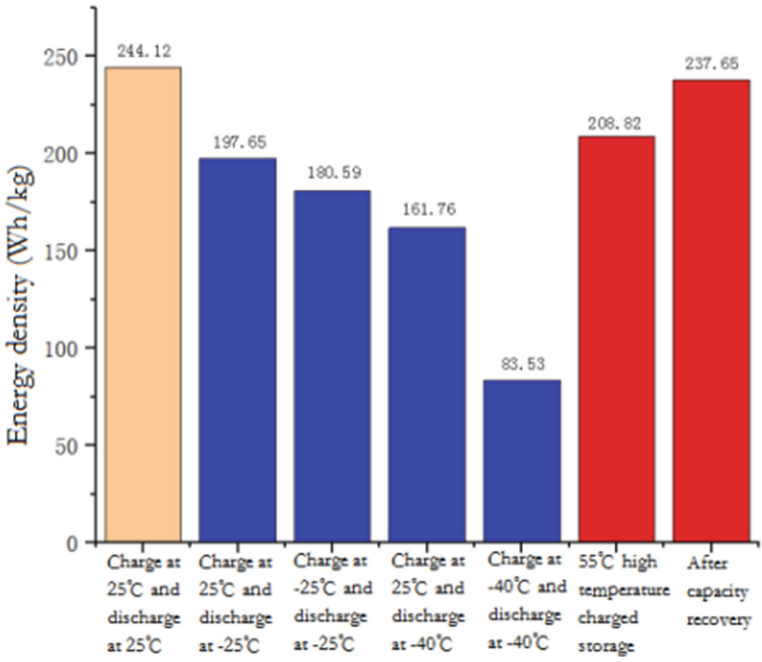


Fig. 4. Energy density under different temperature and charging and discharging conditions

compared with that of normal temperature charge and low temperature discharge. During the high temperature test, the energy density after charging and discharging at high temperature is reduced by 14.46% compared with that at normal temperature, and the energy density after capacity recovery is reduced by 2.65%.

4 Conclusion

Based on the above analysis of the battery discharge data, it can be seen that: At low temperature, the discharge depth and energy density decrease with the decrease of temperature, and the decrease is relatively gentle, only in the ultra-low temperature charge and discharge test, the discharge depth and energy density will decrease significantly compared with the conventional lithium ion battery at −15 °C, the low temperature tolerance of this battery is improved significantly. After 7 days of high temperature charge retention and capacity recovery, the charge retention rate and capacity recovery rate of the battery are 87.20% and 97.78%, and the loss is less. Overall good tolerance to wide temperature conditions.

References

1. Kaijing, B., Yaguo, C., Xianqing, P.: Research progress of low temperature lithium ion battery. Battery **49**(05), 435–439 (2019)
2. Zhenzhen, W., Mi, G., Shidi, T., Jiaxin, P.: Influence factors of ultra-low temperature discharge performance of lithium ion battery. Battery **48**(04), 262–266 (2018)
3. Zhiguo, L., Chengning, Z., Junqiu, L., Guangchong, F., Zhewei, L.: Research on low temperature performance of lithium ion battery for electric vehicle. Automot. Eng. **35**(10), 927–933 (2013)
4. Chen, G., Bixiong, H., Xiao, Y., Ying, W.: Study on low temperature characteristics of lithium titanate battery. J. Shanghai Univ. Eng. Sci. Technol. **33**(01), 21–25 (2019)
5. Zhang, L., Yang, Y., Cui, C., Zhang, Y., Qiao, Y.: Effect of ambient temperature on discharge performance of lithium ion battery. J. Chin. Power Sources :1–11
6. Shixi, Z., Shuangtao, G., Jianwei, Z., Song, Y., Cewen, N.: Research progress on low temperature characteristics of lithium ion battery. J. Chin. Ceram. Soc. **44**(01), 19–28 (2016)
7. Xie, X., Xie, J., Xia, B.: Study on low temperature charging and discharging performance of lithium ion batteries. Chem. World (10), 581–583+614 (2008)
8. Liu, R.: Study on Discharge characteristics and Capacity Estimation of Lithium Battery at low temperature. Southwest University of Science and Technology (2022)
9. Xing, H.: Study on influencing factors and mechanism of high temperature storage performance of lithium ion battery. Fudan University (2014)
10. Bin, Y., Teng, G., Liu, X., Chen, W., Cai, Y.: Degradation mechanism of high temperature storage performance of lithium iron phosphate battery. Power Supply Technol. **42**(07), 955–958 (2018)
11. Liu, Y.: Study on storage aging mechanism and life extension of lithium ion battery under high charge. National University of Defense Technology (2018)

Research and Design of Harmonic Measurement Instrument for High-Voltage Transmission Lines Based on Field Strength Method

Weiliang Han[1], Peidong Zhuang[1(✉)], Lisha Ma[1], and Xuping Wu[2]

[1] College of Electronics Engineering, Heilongjiang University, Harbin 150080, Heilongjiang, People's Republic of China
zhuangpeidong@163.com

[2] College of Electromechanical and Automotive Engineering, Tianshui Normal University, Tianshui 741001, Gansu, People's Republic of China

Abstract. Along with the development of new energy technology, more and more power equipment such as inverters are used in power grids. The power injected into the grid by inverters contains a large number of harmonics, which seriously affects the quality of power emitted from public outlets. At present, inside the substation, voltage harmonics of high-voltage transmission lines can only be measured indirectly with the help of CVT/PT cabinets; outside the substation, the measurement of voltage harmonics of high-voltage lines is not possible due to the absence of CVT/PT cabinets installed. Therefore, a harmonic measuring instrument based on the field strength method is designed to detect the voltage harmonic content of the power grid more conveniently and accurately.The voltage harmonic signal is obtained directly through the field strength method principle, and the FFT technology is used to perform harmonic analysis on the collected grid voltage signal, and the isolation technology is used to make the measured voltage level between 10 kV and 220 kV. After the actual test, the measurement accuracy meets the B-level measurement accuracy stipulated by GB/T19862-2016 standard, which solves the problems of traditional measurement and improves the utilization of electric power transportation more.

Keyword: Eelectric field coupling · Harmonic detection · Wireless radio frequency · Field strength method

1 Field Strength Method for Obtaining Harmonic Signals

The electric field of the transmission line is a low-frequency alternating electric field, the wavelength is much larger than the geometry of the measuring instrument, so the calculation of its electric field can be done with the help of electromagnetic field theory, using the mirror method equivalent charge [1]. The electric field strength is expressed as a function [2].

$$E = E_x \sin(\omega t + \varphi_x)x_0 + E_y \sin(\omega t + \varphi_y)y_0 \tag{1}$$

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 111–119, 2023.
https://doi.org/10.1007/978-981-99-1260-5_14

Style (1): there are $\varphi_x \neq \varphi_y$, that is, the electric field intensity synthesized at any point is a rotating elliptical field. In practice, when considering a single-phase conductor and neglecting the effect of the charge generated on the other two phases of the conductor, the field strength is determined by the line's own charge only, i.e. $E_i \approx \frac{Q_i}{2\pi\varepsilon R_i}$.

The error in the calculation is usually no more than 0.5%. Therefore, a point in space $P(x_i, y)$ is chosen to be at a point with the same horizontal coordinate as that wire $x - i$, see Fig. 1.

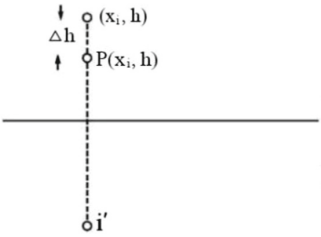


Fig. 1. Vertical distribution of the field strength of the conductors

$$E_{px} = \frac{Q}{2\pi\varepsilon}\left[\frac{x^2}{x^2 + (h-y)^2} - \frac{x^2}{x^2 + (h+y)^2}\right] \tag{2}$$

$$E_{py} = \frac{Q}{2\pi\varepsilon}\left[\frac{h-y}{x^2 + (h-y)^2} - \frac{h+y}{x^2 + (h+y)^2}\right] \tag{3}$$

When $x = 0$, there are $E_{px} = 0$, $E_{py} = \frac{Q}{2\pi\varepsilon}\left[\frac{1}{(h-y)} - \frac{1}{(h+y)}\right]$.

Suppose the voltage transient value of the high-voltage transmission line is $u_i = u_1\cos(\omega t + \varphi) + u_2\cos(3\omega t)$, note that the vertical distance between the point P and the wire is $\Delta h = h - y \geq R$, according to the principle of superposition can be obtained at the point $P(0, y)$ the field strength transient value is

$$\begin{cases} E_{px} = 0 \\ E_{py} = \dfrac{u_1}{\ln\frac{2h}{R}}\left(\dfrac{1}{\Delta h} - \dfrac{1}{2y + \Delta h}\right)\cos(\omega t + \varphi) + \dfrac{u_2}{\ln\frac{2h}{R}}\left(\dfrac{1}{\Delta h} - \dfrac{1}{2y + \Delta h}\right)\cos(3\omega t) \\ \left|E_p\right| = \sqrt{E_{px}^2 + E_{py}^2} = \sqrt{\frac{1}{\ln\frac{2h}{R}}\left(\frac{1}{\Delta h} - \frac{1}{2y+\Delta h}\right)^2(u_1\cos(\omega t + \varphi) + u_2\cos(3\omega t))} \end{cases} \tag{4}$$

There are two points in the electric field, assuming that the vertical distance between the two points and the wire is Δh_1, Δh_2, and the ground is zero potential, the potential of P_1 is expressed as

$$U_1 = \int_{h-\Delta h_1}^{h} E_p d\Delta h = \frac{1}{\ln\frac{2h}{R}}(u_1\cos(\omega t + \varphi) + u_2\cos(3\omega t))\ln\frac{h}{h - \Delta h_1} \tag{5}$$

P_2 The electric potential of

$$U_2 = \int_{h-\Delta h_2}^{h} E_p d\Delta h = \frac{1}{\ln \frac{2h}{R}} (u_1 \cos(\omega t + \varphi) + u_2 \cos(3\omega t)) \ln \frac{h}{h - \Delta h_2} \tag{6}$$

P_1 The voltage difference between P_2 and

$$U = U_1 - U_2 = \frac{1}{\ln \frac{2h}{R}} (u_1 \cos(\omega t + \varphi) + u_2 \cos(3\omega t)) \ln \frac{h - \Delta h_1}{h - \Delta h_2} \tag{7}$$

Let $U_m = \frac{1}{\ln \frac{2h}{R}} \ln \frac{h-\Delta h_1}{h-\Delta h_2}$,

then $U = U_m[u_1 \cos(\omega t + \varphi) + u_2 \cos(3\omega t)]$.

Equation (7) shows that in the space of two points and the wire distance to maintain a constant situation, the two points between the voltage difference signal amplitude value and the corresponding high-voltage line voltage signal is proportional to the relationship. Therefore, the field strength method is used to obtain the voltage difference signal at two points in the vertical direction near the high-voltage line, and then analyze the harmonic content of the high-voltage line.

2 Design of Harmonic Detector

High voltage transmission line harmonic detector consists of: a harmonic measurement sensor and a handheld receiver device, which transmits data via radio frequency between the detection sensor and the handheld receiver device.

2.1 Design of Harmonic Measurement Sensor

The principle block diagram of the harmonic measurement sensor is shown in Fig. 2.

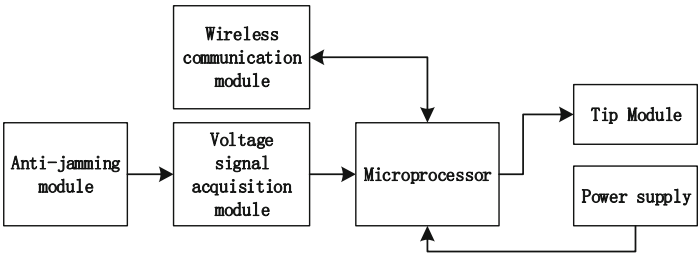


Fig. 2. Block diagram of the principle structure of the harmonic measurement sensor

2.1.1 Anti-interference Module

The interference problem of high-voltage strong magnetic field is solved by using shielding and high-frequency modulation technology [3], and the communication distance under high-voltage strong magnetic field conditions reaches more than 20 m. Using the principle of equalizing voltage, a circular shield is used to solve the problem of harmonic detection sensor tip discharge of portable high-voltage line harmonic detector.

2.1.2 Voltage Signal Acquisition Module

Bring the harmonic measurement sensor electrode close to the high-voltage transmission line with a handheld insulated operating rod, and the voltage signal acquisition circuit design is shown in Fig. 3.

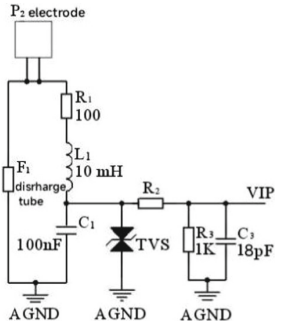


Fig. 3. Voltage signal acquisition circuit

If the harmonic measurement sensor is put into the electric field, according to the electric field theory, a certain amount of charge will be induced on the sensor electrode, Causes stray capacitance between the electrode conductor of the harmonic measurement sensor and the high-voltage transmission line.

The capacitors C_1,C_2 and C_3 form a series circuit, according to the principle of series impedance divider, thus the voltage signal consistent with the harmonic content of the high-voltage transmission line can be coupled at both ends of the voltage divider capacitor C_1. When the electrode touches or approaches the wire, a voltage signal with the same frequency and phase as the line signal will be induced at both ends of the electrode, and then a voltage signal of moderate amplitude will be obtained through the voltage divider capacitor for subsequent circuit acquisition and processing.

The equivalent circuit for acquiring the alternating voltage signal in Fig. 3 is shown in Fig. 4. The capacitor $C_{\_G}$ represents the equivalent capacitance generated by the electrode, and the capacitor $C_{\_F}$ represents the voltage divider capacitance of $C_1 \sim C_3$ in Fig. 3.

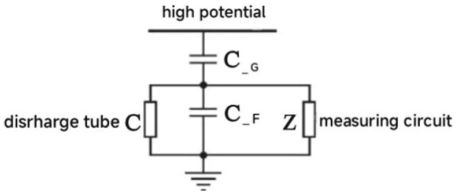


Fig. 4. Equivalent circuit for alternating voltage signal acquisition

2.1.3 Anti-aliasing Filter Circuit

Considering the standard FFT transformation requires the number of sample points per cycle to be 2^n, the processing speed of the chip and the maximum time interval of the conversion of the A/D module, $f_s = 6400$ Hz is selected. Identified as CFA-based 6th-order Butterworth low-pass filter[4], two third-order filters are designed from Fig. 5, and then two sections of the third-order low-pass filter are cascaded to obtain the 6th-order Butterworth low-pass filter.

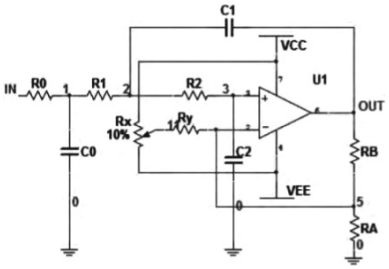


Fig. 5. Third-order low-pass filter improvement circuit improvement circuit

2.1.4 Wireless Communication Module

Using electric field coupled wireless RF transmission of harmonic information, the flow block diagram is shown in Fig. 6. Electric field coupled no RF technology to alternating electric field as a carrier of energy wireless transmission, first by the original side and vice side of a variety of plate components, and then with a number of coupling capacitors to increase the channel of energy transmission.

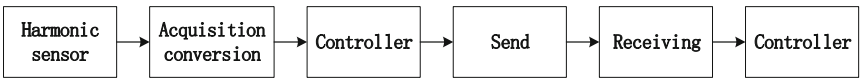


Fig. 6. Wireless RF flow block diagram

2.1.5 Prompt Module

Using two two-color light-emitting diodes and a buzzer, when the power is on an indicator light green, when coupled to the electric field signal two indicators light green, when the electric field signal is too strong two indicators light red, buzzer continuous sound.

2.2 Design of Handheld Receiver Device

The handheld receiver device mainly completes the functions of data reception, analysis and calculation, display of results, storage and query. The structure is illustrated in Fig. 7.

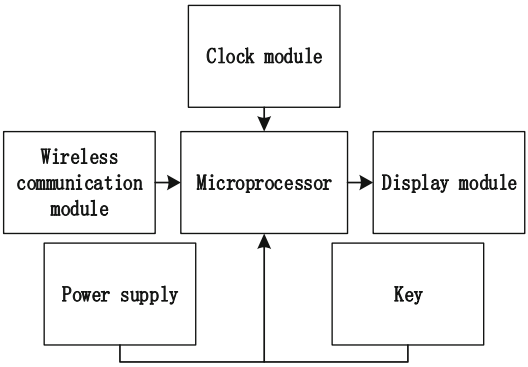


Fig. 7. Block diagram of the principle structure of the handheld receiver

3 Software Implementation Methods for Systems

3.1 Analysis of Voltage Harmonics Algorithms

The article's calculation of the harmonic components is the use of the Fourier algorithm[5]. The basic principle of the FFT algorithm is to decompose the DFT of a long sequence into the DFT of a shorter sequence one at a time.This system uses the time-extracted FFT algorithm, i.e. the base 2- FFT transformation.

3.2 Components of a Harmonic Detection System

The system mainly consists of a voltage acquisition module, waveform change circuit, A/D conversion circuit, microcontroller, etc. The structure diagram of the system is shown in Fig. 8.

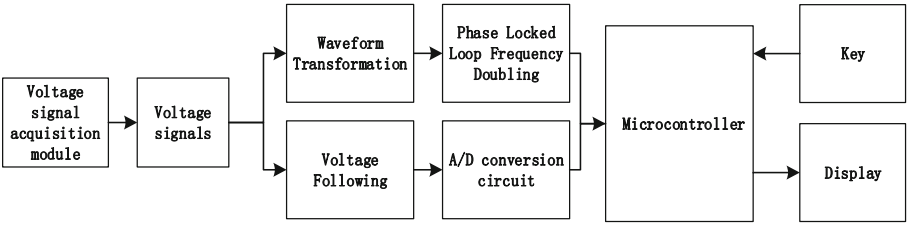


Fig. 8. System architecture diagram

The working principle is that the voltage signal is collected by the voltage acquisition module, then the signal is sent to the AD574 to complete the A/D conversion, the sampled data is read by the microcontroller and the voltage data is saved. Once a cycle of 32 samples has been taken, the harmonic value is obtained by the FFT algorithm, and the harmonic value is displayed according to the keypad.

3.3 Implementation of a Harmonic Detection System

The main function of the system is data acquisition and processing. The main control flow diagram is shown in Fig. 9.

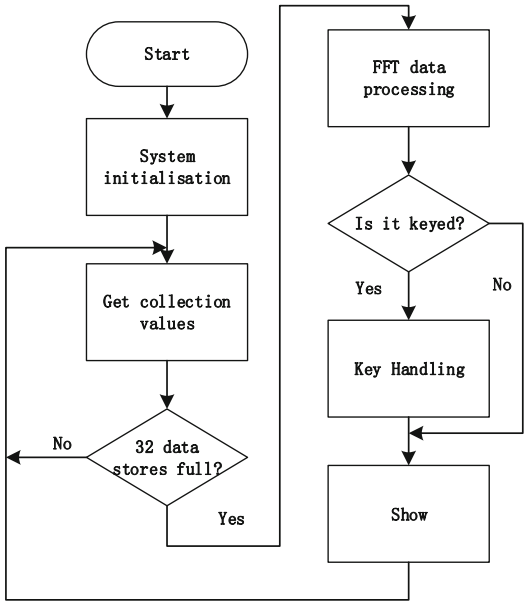


Fig. 9. Main control flow chart

4 Comparison of Experimental Data

Multifunctional three-phase electrical measuring instrument calibration device, the results of which are shown in Table 1.

Table 1. Odd harmonic measurement data

Data	Harmonic content/%	Measurement of harmonic content %	Absolute error
3	15	15.11	0.11
5	10	9.89	-0.11
7	5	5.02	0.02
9	3	3.17	0.17
11	3	2.98	-0.02
13	2	2.07	0.07
15	1	1.25	0.25

Generally only odd harmonics are analysed in the study of harmonics in power grids, because odd harmonics cause more harm than even harmonics.Therefore only the individual odd harmonic components are covered in the analysis of the data in the table above. When the data are analysed on this basis, there are large individual deviations, but none of the absolute errors exceed 0.5%, the errors being caused by the lack of strict synchronous acquisition. The accuracy was measured experimentally at 0.5 level, which meets the requirements of the national standard.

5 Conclusion

Through a combination of theoretical analysis, hardware design and experimental measurement, the field strength method and FFT are proposed to achieve the detection of voltage harmonics of high-voltage transmission lines, and a harmonic measuring instrument based on the field strength method is also designed, and the results show that the harmonic measuring instrument can measure the harmonic content of high-voltage transmission lines more accurately by comparing the measurement data of this instrument with the actual measurement results. The harmonic detector overcomes the traditional harmonic detection only for the station transmission line voltage harmonics detection, measurement accuracy is low and other problems, easy to use, not only for the substation station harmonics, but also for the substation station outside the harmonic condition can also be detected, has important significance for understanding the transmission line harmonics, has a very good promotion and application value.

References

1. Zhang, J., Chen, Z., Zhou, Z., Qian, J.: Research and design of a new voltage harmonic detection method for high-voltage transmission lines. High Volt. Electron. **12**, 92–97 (2012)
2. Zhang, G., Yang, W., Lian, H., Zeng, W., Zhang, J., Zhang, T.: Research and design of a harmonic detection method for high-voltage transmission lines based on electric field coupling method. Electr. Measur. Instrum. **21**, 18–23 (2020)
3. Yang, F., Zheng, Y., Zhang, X.: Effect of shield wire on the frequency electric field under ultra-high voltage transmission lines. High Volt. Electr. **47**, 64–68 (2011)
4. Dong, J.N., Chi, S.H., Liu, W., Zhao, B.S.: Design of sixth-order Butterworth low-pass filter based on CFA. Electron. Technol. **09**, 56–59 (2019)
5. Xueqin, T., Ziyi, Q., Xiong, J.: A software implementation method for AC voltage harmonic detection system. Sci. Technol. Innovat. Appl. **07**, 44–45 (2016)

Study of Smart Home Environment Monitoring System Based on Cloud Platform and Android

Shaopeng Yu(✉), Liyuan Dong, and Fengyuan Pang

College of Information and Intelligent Engineering, Tianjin Renai College, Jinghai District, Tianjin 301636, China
globely@163.com

Abstract. A smart home environment monitoring system based on cloud platform and Android client is proposed to improve the convenience of people's daily life. The multi-sensor node information fusion method is used to analyze the output characteristics of data in different indoor environments, which is on the basis of the back propagation neural network. The intelligent gateway is constructed by sensors, singlechip and wireless modules, which is connected with the cloud platform by LoRa and WiFi, and then, the wireless communication and remote control are realized. The environmental data and change curve are displayed by App and the real-time air quality index is given. The test results show that the accuracy of CO_2 and PM2.5 concentrations is up to 98% and the changing value can be displayed to users with curves, so that the environmental conditions can be understood conveniently and timely.

Keywords: Sensor · Neural network · Cloud platform · Android

1 Introduction

With the rapid development of high-tech equipment and internet of things (IoT) technology in recent years, the intelligence becomes a new technological revolution for people's living environment [1, 2]. The modern home has regarded the diverse information, safe, comfort and convenience as an ideal goal, a smart system can meet the needs of basic living conditions, which is combined with network communication, intelligent perception, embedded systems, cloud technology and other advanced technologies organically [3]. It has been applied to food, clothing, housing and transportation, so that people can enjoy the great improvement and wonderful experience of life, which is brought by scientific and technological achievements [4–6].

The traditional monitoring system has some problems, which are difficult to obtain the composite data and control the indoor environmental parameters [7]. Based on the background, an embedded home environment monitoring system combined with multi-sensor fusion technology and cloud platform is designed, and the wireless collection of indoor dynamic environmental information and the remote control of household equipment can be achieved. In order to improve the collection capacity of indoor environment data, the multi-sensor node information fusion method is used to analyze the output

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 120–127, 2023.
https://doi.org/10.1007/978-981-99-1260-5_15

characteristics of data in different indoor environments, which is based on the back propagation neural network, and the software processing of fuser and fuzzy inference can be realized by MATLAB. The detection results can be obtained by multi-level criteria, the accuracy of CO_2 and PM2.5 concentrations will be conspicuously improved.

2 Data Collection and Processing

According to the theory of information fusion, the multi-sensor node is used to acquire the output characteristics of data in different indoor environments. The collected sensor signals are preprocessed by using a single chip compute to complete the information layer fusion of the system, and then the preprocessed signals are sent to the singlechip. The design of neural network fusion device and fuzzy inference system is completed based on MATLAB, and the feature fusion of the system is completed through the processing of the underlying data by the upper computer.

2.1 Modeling Variables and Training Samples

The home environment is closely related to people's life and directly affects the health of their families, so it is necessary to monitor the home environment in real time [8]. For the analysis of environmental impact factors, the modeling variables are selected as follows.

The general indoor temperature range is 18 °C–26 °C and the air humidity range is 49–51. If it exceeds this range, people will have discomfort symptoms, so the temperature and humidity is selected as the model input. The light pollution will affect people's vision and body, which will cause pressure on psychology and physical health, so it is chosen as the input too. When the CO_2 concentration is too high, the ability of haemoglobin to deliver oxygen will be weakened greatly and the human body will suffer from hypoxia, so it is an important input of the model. PM2.5 is a fine particle with a diameter of less than 2.5 microns in the air. Once the human body is inhaled, the deeper it enters the respiratory tract, the more serious the impact on the ventilation function of the lung, and then the people will be in a state of hypoxia. Therefore, the content of PM2.5 is selected as the last input. In short, temperature, humidity, illumination, CO_2 concentration and PM2.5 content are five inputs of the system network model.

The inputs determine the number of neurons in the input layer, so there are five input neurons. The output value is the indoor environment level value, so there is one output neuron. The Gaussian function is used as the membership function, which is the input the value of neuron variable. The number of input neurons is 5, meanwhile, each Gaussian membership function has 2 parameters and the number of each membership functions variable is 3, which is correspond to the limit interval value of each input quantity. Thus, the number of precursor network parameters in the model is $5 \times 3 \times 2 = 30$. The subsequent network has a constant term, so the number of subsequent network parameters is $(5 + 1) \times 30 = 180$. Therefore, the total number of parameters in the network is $30 + 180 = 210$. In order to achieve the best training effect, it is best to make the training sample 3–5 times the network parameters, so 1000 training samples is taken.

2.2 Data Processing Based on Neural Network

The neural network has strong nonlinear mapping ability, which can realize self-learning and self-adaptive, so that when there are a large number of errors in the samples, the results will not be impacted seriously. The learning and training are core part of neural network computing, and the steepest descent method is used in this paper, the specific steps are as follows.

1. The network is initialized to assign the weights and the thresholds of each unit in the interval (−1,1), meanwhile, the error function e is selected and the calculation accuracy value ε and the maximum learning time M are given.
2. A group of samples is selected randomly as inputs and outputs. $X_k = [x_1^k, x_2^k, x_3^k]$ is the input of group k and $Y_k = \left[y_1^k, y_2^k, y_3^k\right]$ is the relative output.
3. The hidden layer is calculated. The input is

$$A_i = \sum\nolimits_{i=1}^{3} (\omega_{ij} x_i - \theta_j) \tag{2.1}$$

   The output is

$$b_i = \frac{1}{1 + e^{-A_i}} \tag{2.2}$$

   ω_{ij} is weight and θ_j is threshold.
4. The output layer is calculated. The input is

$$L_j = \sum\nolimits_{j=1}^{10} (v_{jt} b_j - \gamma_t) \tag{2.3}$$

   The output is

$$c_t = \frac{1}{1 - e^{-L_j}} \tag{2.4}$$

   v_{jt} is weight and γ_t is threshold.
5. The unit error between the actual output and the target output is calculated.

$$d_t^k = (y_t^k - c_t) c_t (1 - c_t) \tag{2.5}$$

6. The unit error of middle layer is calculated.

$$e_j^k = (\sum\nolimits_{t=1}^{3} d_t^k v_{jt}) b_j (1 - b_j) \tag{2.6}$$

7. The hidden layer is revised. The revised weight is

$$w_{ij}(N+1) = w_{ij}(N) + \alpha[(1-\eta) e_j x_i^k + \eta e_j x_1^{k-1}] \tag{2.7}$$

   The revised threshold is

$$\theta_j(N+1) = \theta_j(N) + \alpha[(1-\eta) e_j^k + \eta e_j^{k-1}] \tag{2.8}$$

   N is training time, η is learning rate and α is attenuation coefficient.

8. The output layer is revised. The revised weight is

$$v_{jt}(N+1) = v_{jt}(N) + \alpha d_t b_j \tag{2.9}$$

The revised threshold is

$$\gamma_j(N+1) = \gamma_j(N) + \alpha d_t \tag{2.10}$$

9. The weights and the thresholds are continuously updated in order to complete the error learning until the algorithm ends, which meets the preset accuracy or the maximum learning time. The calculation results are shown in Fig. 1.

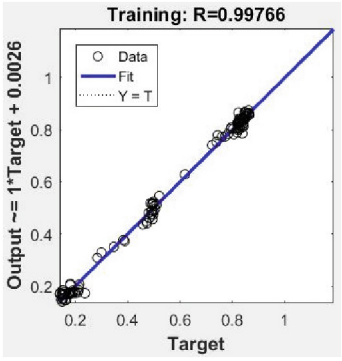


Fig. 1. Results of neural network algorithm

3 Structure of Monitoring System

An intelligent gateway is constructed by STM32F103RCT6 chip and ESP8266 wireless WiFi module. A monitoring network is established by LoRa technology and the information channel with Baidu cloud platform is built to obtain home environment data, and then the hardware control function can be realized. The MQTT protocol is used to subscription, publishing topics and data transmission. The Android devices can be accessed to cloud platform and a two-way link between mobile devices and cloud is built by the device cloud App. The subscription and forwarding of messages can be completed.

3.1 Hardware Design

The environmental parameters are detected by multiple sensors, which can form a wireless star network with LoRa coordinator, so the collection and transmission of environmental data can be realized. The error correction of processing data can be achieved by coordinator through STM32. The intelligent gateway is constructed by WiFi module and PC, and the data is uploaded to the cloud platform by router. The user can log in

to the device cloud to realize remote monitoring through an Android App. Meanwhile, the App also can be used to realize remote control of home appliances, which the control instructions are sent to the intelligent gateway by the cloud server, and then sent to the terminal node through coordinator. Finally, the instructions are sent to. The overall structure of the smart home environment monitoring system is shown in Fig. 2.

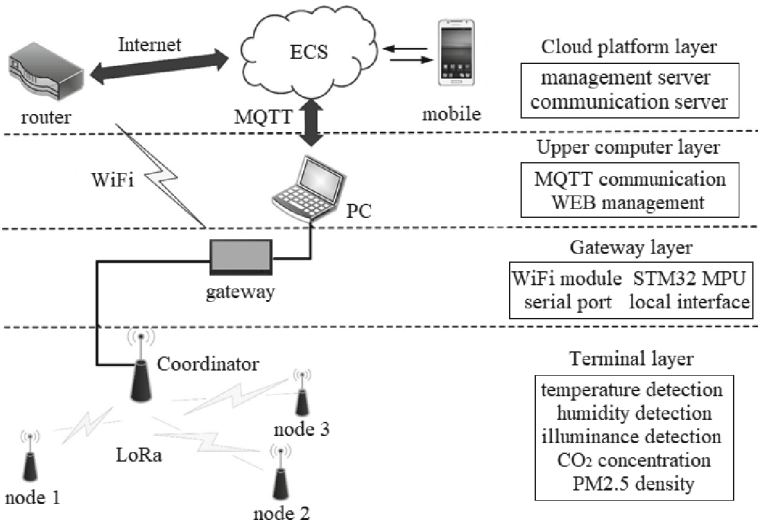


Fig. 2. System structure diagram

The hardware of monitoring system is composed of temperature and humidity sensor, light intensity sensor, CO_2 concentration sensor, PM2.5 content sensor and coordinator, which is shown in Fig. 3. The four sensors and the coordinator form a star shaped LoRa wireless sensor network, which constitutes the acquisition layer and transmission layer of the system. The main devices used in the system include smart home environment monitoring system, PC connected to LAN, wireless router and Android mobile phone equipped with client App. The LoRa is used to configure and interconnect the sensors and coordinator devices in the home internal network, and then the network is connected with the embedded home gateway by serial port. The gateway is connected with the home internal router through the network cable interface, and the mobile phone is connected with the home router through WiFi.

3.2 IoT Cloud Platform and Mobile Terminal

The IoT core suite of the internet of things is used to create a cloud platform. Click the instance name and enter the details page, and then the device list can be used to create and manage devices. Check the connection information of IoT Core, Device key, Device Secret, etc. and confirm whether the message sending and communication with IoT core can be realized by the MQTT application client. The rule engine is used to flow the device messages to the business server, and Baidu cloud platform will flow the data

Fig. 3. System hardware composition

reported by the network to Baidu Kafka, and then the data in Kafka can be consumed for business processing by business application server. The specific design process is shown in Fig. 4.

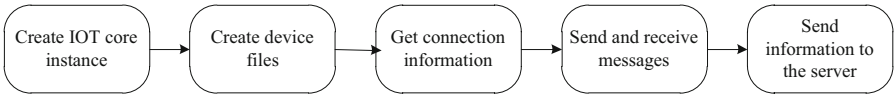


Fig. 4. Cloud platform design process

The developers of Android platform have submitted a large number of UI components to the internet, which include layout manager, TextView and its subclasses, ImageView and its subclasses, AdapterView and its subclasses, ProgressBar and its subclasses, dialog boxes and others, which can be used to create the interface to achieve the needs. This pattern can be used independently or nested with each other to meet a good spatial layout effect.

The linear distribution means that the information display part in the interface is arranged along the horizontal or vertical direction. If it is arranged along the horizontal direction, only one component can be placed in each column. If it is arranged vertically, each row can be installed with the same component. The android App designed in this paper adopts LinearLayout, so the users can access the server from client and obtain the latest environmental detection data from database.

3.3 Experimental Test and Analysis

In order to verify the practicability of the system, the acquisition node is placed indoors and the App is installed on the smart phone. After opening the App, the specific values of

indoor environmental factors can be observed clearly and the value of air quality index can be calculated, which is shown in Fig. 5. When the value of the detected environmental factors exceeds the standard value, the system will immediately push an alarm message to the user's phone, which could remind the user to start the corresponding air purification program to purify the room. The test results have proved that the accuracy of CO_2 concentrations and PM2.5 content measured by the system is up to 98%. Meanwhile, the data changes can be displayed to users with curves, so that users can easily and timely understand the indoor environmental conditions.

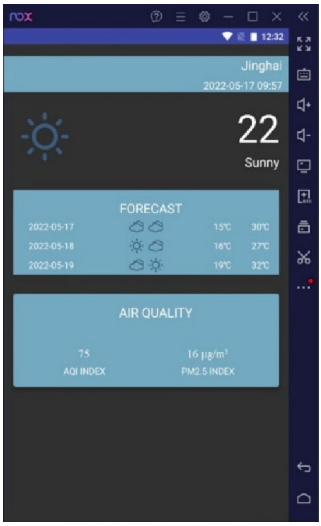


Fig. 5. Android client and test results

4 Conclusion

The smart home environment monitoring system combined with multi-sensor fusion technology and cloud platform is designed, and then the wireless collection of indoor dynamic environmental information and the remote control of household equipment can be realized. The multi-sensor information fusion method based on the back propagation neural network is used to analyze the output characteristics of data in different indoor environments, which can improve the accuracy of the monitoring system effectively, so as to provide the accurate environmental information to the users. The intelligent network is built by coordinator and wireless module, meanwhile, the mobile communication and remote control are realized by cloud platform and App. The detection results can be obtained by multi-level criteria, the accuracy of CO_2 and PM2.5 concentrations is up to 98%. This system has the advantages of convenient deployment, simple operation, low cost and small error, which will have a broad market development prospect.

Acknowledgements. This work was supported by Tianjin Research Innovation Project for Postgraduate Students under grant No. ZX21014.

References

1. Wang, X., Guo, Y., Ban, J., et al.: Driver emotion recognition of multiple-ECG feature fusion based on BP network and D-S evidence. IET Intell. Transp. Syst. (2020). https://doi.org/10.1049/iet-its.2019.0499
2. Duoont, M.F., Elbourne, A., Cozzolino, D., et al.: Chemometrics for environmental monitoring: a review. Anal. Meth. (2020), https://doi.org/10.1039/D0AY01389G
3. Al-Dabbous, A.N., Kumar, P., Khan, A.R.: Prediction of airborne nanoparticles at roadside location using a feed–forward artificial neural network. Atmos. Pollut. Res. (2017). https://doi.org/10.1016/j.apr.2016.11.004
4. Kong, F., Zhou, Y., Chen, G.: Multimedia data fusion method based on wireless sensor network in intelligent transportation system. Multimedia Tools Appl. **79**(47–48), 35195–35207 (2019). https://doi.org/10.1007/s11042-019-7614-4
5. Athira, V., Geetha, P., Vinayakumar, R., et al.: Deep air net: applying recurrent networks for air quality prediction. Procedia Comput. Sci. (2018). https://doi.org/10.1016/j.procs.2018.05.068
6. Lin, W., Xibin, A.: Risk assessment of knowledge fusion in an innovation ecosystem based on a GA-BP neural network. Cogn. Syst. Res. (2020). https://doi.org/10.1016/j.cogsys. 2020.12.006
7. Lingbao, K., Xing, P., Yao, C. et al.: Multi-sensor measurement and data fusion technology for manufacturing process monitoring: a literature review. Int. J. Extreme Manuf. (2020). https://doi.org/10.1088/2631-7990/ab7ae6
8. Wang, J., Yu, Q.: A Dynamic multi-sensor data fusion approach based on evidence theory and WOWA operator. Appl. Intell. **50**(11), 3837–3851 (2020). https://doi.org/10.1007/s10489-020-01739-8

PFANet: A Network Improved by PSPNet for Semantic Segmentation of Street Scenes

Jiangwei Ge[1,2], Yan Li[1,2(✉)], Mengfei Jiu[1,2], Zichen Cheng[1,2], and Jingwei Zhang[1,2]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
wdxyliy@163.com

[2] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China

Abstract. Streetscape scene understanding is an important task for semantic segmentation. In this paper, we propose a PFANet improved based on PSPNet for Street Semantic Segmentation. For further improving the feature fusion capability of our model, we added a feature fusion module to PSPNet to incorporate features of different dimensions. Meanwhile, we introduce an attention mechanism, using a combination of spatial attention mechanism and channel attention mechanism attention module to enhance the contextual dependencies of local features and the spatial interdependencies of features. We experimented on the Cityscapes dataset, and achieved a mIoU score of 80.38% on Cityscapes validation dataset which improved 0.68% than PSPNet.

Keywords: Semantic segmentation · Feature Fusion · Attentional Mechanisms

1 Introduction

Semantic segmentation is aim to classify each pixel and give it a corresponding category label in an image, which is one of the important fundamental tasks of computer vision. It has potential applications in many fields such as autonomous driving, medical imaging, robotics sensing, etc. And street semantic segmentation is a fundamental application of semantic segmentation in the field of autonomous driving which is a challenging task. In street scenes, the same category can be significantly different in different images due to shooting angles, occlusions and other factors, which can lead to excessive differences between the same category and lead to misclassification problems. Meanwhile, complex categories are prone to background and other neighboring pixels when extracting features.

To segment images, traditional semantic segmentation mainly applies thresholding [1], edge detection [2], pixel clustering [3], etc. However, the segmentation effect of these methods is not good enough and requires manual feature extraction, which is inefficient. With the emergence and development of deep learning-based semantic segmentation methods, image features can be learned and extracted by neural network, and no longer need manual extraction, and the segmentation effect and efficiency have been greatly

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 128–134, 2023.
https://doi.org/10.1007/978-981-99-1260-5_16

improved. Currently, semantic segmentation networks basically use encoder-decoder architecture, where CNN is commonly used in the encoder part to extract features and the decoder side produces predicted results. To address this problem, zhao et al. propose a pyramid scene parsing network (PSPNet) [4] to exploit the global scene category cues, and chen et al. [5–7] use atrous convolutions to increase the receptive field and obtain multi-scale contextual information while maintaining resolution and reducing computational effort. The use of attention mechanisms has also brought great improvements to semantic segmentation, and there are generally two types of channel attention and spatial attention. Channel attention [8, 9] models the interdependence between channels of its features by assigning different weights to different channels, while spatial attention [10, 11] enhances spatial correlation by adding corresponding weights to different spatial locations. And there is also a combination of channel attention and spatial attention [12–15] to make up for each other's deficiencies.

In this work, we improve PSPNet and use it for street view semantic segmentation. We add a feature fusion module that fuses low-dimensional features and high-dimensional features to PSPNet to improve the feature fusion capability. We also introduce an attention mechanism, that an attention module combining channel attention and spatial attention added to the network to strengthen the channel dependency and contextual dependencies between features.

2 Related Work

Semantic segmentation Full convolutional networks (FCNs) [15] is the pioneer of semantic segmentation based on deep learning, which replaces fully connected layers with convolutional layers and achieves pixel-to-pixel semantic segmentation in an end-to-end manner, but still exists the problem of rough segmentation effect. UNet [16] is modified from FCN, which uses a U-shaped structure to stitch together low-level features and deep-level features by channel to achieve contextual fusion, and has achieved better results in medical images. In order to combine context information effectively, PSPNet [4] designs a pyramid pooling module to combine multi-scale features. Deeplabv2 [5] proposes an atrous spatial pyramid pooling (ASPP) module to enhance the model's ability to recognize the same object of different sizes, which consist of atrous convolutions with different dilated rates and spatial pyramid pooling. Deeplabv3 [6] arranges the ASPP module in a parallel manner, which adds batch normalization, a global pooling layer and a bilinear interpolation upsampling layer to the module. Deeplabv3 + [7] introduces an encoder-decoder architecture which employs DeepLabv3 as an encoder to improve the image segmentation effect.

Attention mechanism Attention mechanism first appeared in the field of NLP to enhance attention to important information while ignoring unimportant information, and now have a wide range of applications in visual tasks. Squeeze-and-Excitation Networks (SENet) [8] utilizes global information to explicitly model the interdependencies between feature channels to enhance the representational ability of the network. ECANet [9] improved on SENet [8] by using one-dimensional convolution to effectively capture cross-channel interactions while avoiding dimensionality degradation. NLNet [10] designs a non-local block which uses non-local operations to capture long-range dependencies and construct dense spatial feature maps. GCNet [11] combines NLNet [10]

and SENet [8] to design a lightweight module to model global contextual dependencies. BAM [12] and CBAM [13] both applies both spatial and channel attention to aggregate contextual information. BAM [12] places the attention module in the bottleneck to save network parameters and computational overhead, and CBAM [13] combines average pooling and maximum pooling to aggregate information of feature map.

3 Method

The architecture of our network is given in Fig. 1. Our network improves on PSPNet by adding a feature fusion module and an attention module. We use ResNet [14] as the backbone to extract image features, and send all the extracted feature maps to the feature fusion module for feature fusion, while the last feature map generated by resnet is input into pyramid pooling module (PPM). And then fed the fused feature map into the attention module to enhance the contextual relationships and local dependencies of the features. Finally, it is concatenated with the multiscale features generated by PPM and then a convolutional layer is passed to generate the segmentation map.

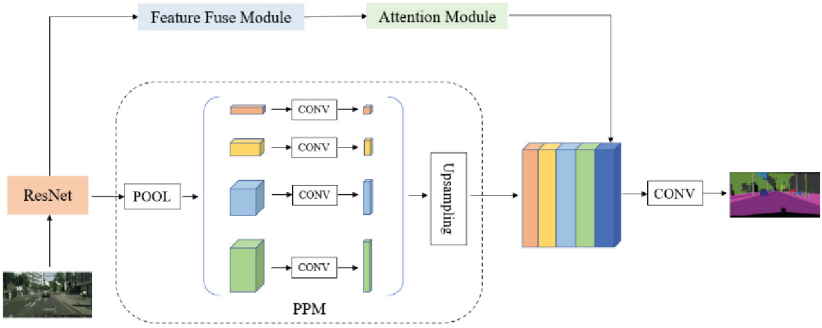


Fig. 1. Overview of the PFANet.

3.1 Feature Fuse Module

For taking full advantage of the low-dimensional features, we designed a feature fusion module which similar to the UNet architecture. As illustrated in Fig. 2, feature maps F1, F2, F3, F4 are generated by extracting features from backbone. F2 is generated by F1 through a reslayer consisted from a certain number of bottlenecks, and F3, F4 are obtained in the same way. The high-dimensional feature map uses 1×1 convolutional layer to adjust the number of channels to be consistent with the low-dimensional feature map, while maintaining the same scale as the low-dimensional feature map by a bilinear upsampling layer. In contrast to UNet, we use element-wise summation to fuse the high-dimensional features with the low-dimensional features, and finally output the fused feature map after adjusting the channels by a 3×3 convolutional layer.

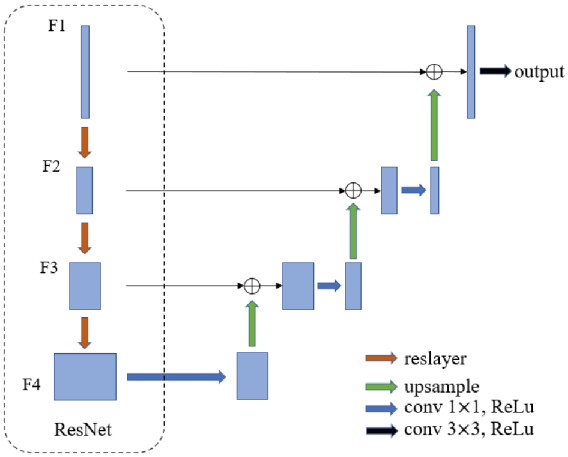


Fig. 2. The architecture of Feature Fuse Module. The left part of the module is part of the backbone resnet, and F1, F2, F3, F4 correspond to 4 feature maps of different channel dimensions extracted by resnet. Bold arrows with different colors indicate different operations or the layers of the network that feature maps go through.

3.2 Attention Module

As illustrated in Fig. 3, our attention module includes channel attention and spatial attention, which is a combination of spatial attention of BAM [12] and channel attention of CBAM [13]. In the channel attention branch, FC1 layer and FC2 layer both consist of a 1 × 1 convolutional layer and activation functions ReLU, and the reduction rate r is used to control the reduction multiplier of the channels. We apply the same average pooling and maximum pooling to obtain contextual information as CBAM. Denoting the convolutional layer in FC1 as C1 and the convolutional layer in FC2 as C2, the channel attention is computed as:

$$\begin{aligned} \mathrm{F_c} &= \mathrm{FC2}\big(\mathrm{FC1}\big(F_{avg}\big)\big) + \mathrm{FC2}(\mathrm{FC1}(F_{\max})) \\ &= \sigma\big(\mathrm{C2}\big(\delta\big(\mathrm{C1}\big(F_{avg}\big)\big)\big)\big) + \sigma(\mathrm{C2}(\delta(\mathrm{C1}(F_{\max})))) \end{aligned} \tag{1}$$

where F denotes input feature, α and δ indicate the ReLU function, F_{avg} and $F_{\max}$ respectively correspond to the feature maps after average pooling and max pooling.

In the spatial attention branch, we share the FC1 layer with the channel attention branch to compress the number of channels. In contrast to BAM, we use 3 atrous convolutions with different dilation rates connected in parallel to enlarge the receptive field while avoiding the gridding effect. And then obtaining a $\mathrm{R}^{1\times\mathrm{H}\times\mathrm{W}}$ spatial attention map $\mathrm{F_s}$ by a 1 × 1 convolution layer, which is computed as:

$$\mathrm{F}_s = f\left(f_{a1}(\mathrm{FC1(F)}) + f_{a2}(\mathrm{FC1(F)}) + f_{a3}(\mathrm{FC1(F)})\right) \tag{2}$$

where f_{a1}, f_{a2}, f_{a3} denote 3 atrous different convolutions, and f denotes 1 × 1 convolution operation. Finally, the channel attention map and the spatial attention map are multiplied point by point with the input feature map respectively, and then adding the two feature maps generated by this operation point by point to obtain the output feature.

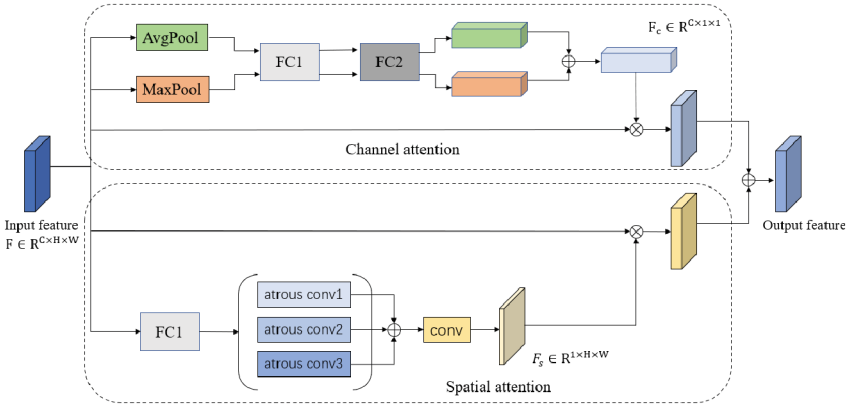


Fig. 3. The details of our Attention Module. It consists of two branches: channel attention and spatial attention. FC1 compress feature map channels to C/r, where r is the reduction ratio. FC2 has the same composition as FC1, but is used to resume channels to C. Atrous conv1, atrous conv2, atrous conv3 are 3 atrous convolutional layers with different dilation rates d1, d2, d3 that connected in parallel. In the experiments, we set the reduction rate r to 16 and the dilation rates [d1, d2, d3] to [1, 3, 5].

4 Experiment

4.1 Dataset

We mainly evaluate our model on the Cityscapes dataset [17]. Cityscapes is a high-resolution urban scenes dataset for semantic segmentation, which has 19 categories of densely annotated pixels. It contains 5000 finely annotated images of size 2048 × 1024, of which 2975 are used for training, 500 for validation, and 1525 for testing. Also 20,000 coarsely annotated images are provided for model training. We primarily test our model on the validation set.

4.2 Implementation Details

We use the publicly available codebase mmsegmentation [18] and trained our model with distributed training on a server with 2 RTX 3090. We apply data augmentation to the images following the standard settings of mmsegmentation such as random resize with ratio range in 0.5 to 2, random cropping the image to 512 × 1024 and random flip. For backbone, we apply ResNetV1c-101 [19] which employs a dilation network strategy in stage 3 and stage 4. We set the batch size to 4 and the total iteration to 60K. We also use the SGD optimizer, whose initial learning rate set to 0.02, momentum set to 0.9, and weight decay set to 0.0005. Meanwhile, adopting a "poly" learning policy with power of 0.9. As with PSPNet, we also employ an auxiliary loss, which is a cross-entropy loss like the main branch loss, and set the auxiliary loss weight to 0.4. And the mean Intersection over Union (mIoU) is applied as the evaluation metric. We take PSPNet as the baseline and compare the results of our method with it and other representative methods on the Cityscapes validation set. The results are given in Table 1, where our method performs better than several other methods.

Table 1. Segmentation results on the Cityscapes validation set

Method	Backbone	mIoU
DeepLabV3 [6]	Dilated-ResNet-101	78.5
DeepLabV3 + [7]	Dilated-Xception-71	79.6
UNet + + [20]	ResNet-101	75.5
HRNetV2 [21]	HRNetV2-W40	80.2
PSPNet [4]	Dilated-ResNet-101	79.7
Our method	ResNetV1c-101	**80.38**

4.3 Ablation Studies

For demonstrating that our designed feature fusion module and attention module are effective, we put these two modules into the same position in the network separately for experiments. As shown in Table 2, both modules are helpful in enhancing the performance of the network, with the feature fusion module improving by 0.03 and the attention module improving by 0.51 which is more compared to the feature fusion module. Performance with different modules using in baseline.

Table 2. FFM denotes Feature Fusion Module, and AM denotes Attention Module.

FM	AM	mIoU
✓	×	79.73
×	✓	80.21
✓	✓	80.38

5 Conclusion

We propose PFANet, a semantic segmentation partition network for street scenes, which is modified from PSPNet. To make full use of low-dimensional features, we add a feature fusion module to the network combining low-dimensional features and high-dimensional features. We also introduce an attention module which fuses channel attention and spatial attention to enhance the network's ability to combine context dependencies. We experiment on the Cityscapes validation set and our results improve 0.68 over baseline.

References

1. Davis, L.S., Rosenfeld, A., Weszka, J.S.: Region extraction by averaging and thresholding. IEEE Trans. Syst. Man Cybern. **SMC-5**(3), 383–388 (1975). https://doi.org/10.1109/TSMC.1975.5408419

2. Senthilkumaran, N., Rajesh, R.: Edge detection techniques for image segmentation– a survey of soft computing approaches. Int. J. Recent Trends Eng. **1**(2) (2009)
3. Ng, H.P., Ong, S.H., Foong, K.W.C., Goh, P.S., Nowinski, W.L.: Medical image segmentation using k-means clustering and improved watershed algorithm. In: 2006 IEEE Southwest Symposium on Image Analysis and Interpretation, Denver, CO, pp. 61–65 (2006)
4. Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J.: Pyramid scene parsing network. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2881–2890 (2017)
5. Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. CoRR, abs/1606.00915 (2016)
6. Chen, L.-C., Papandreou, G., Schroff, F., Adam, H.: Rethinking atrous convolution for semantic image segmentation. CoRR, abs/1706.05587 (2017)
7. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 833–851. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_49
8. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
9. Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q.: Eca-net: efficient channel attention for deep convolutional neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
10. Wang, X., Ross, G., Abhinav, G., He, K.: Non-local neural networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
11. Cao, Y., Xu, J., Lin, S., Wei, F., Hu, H.: Gcnet: non-local networks meet squeeze-excitation networks and beyond. In: ICCV Workshops (2019)
12. Park, J., Woo, S., Lee, J.-Y., Kweon, I.S.: Bam: bottleneck attention module. In: British Machine Vision Conference (BMVC) (2018)
13. Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: CBAM: convolutional block attention module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 3–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_1
14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: CVPR (2016)
15. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: CVPR, pp. 3431–3440 (2015)
16. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
17. Cordts, M., et al.: The cityscapes dataset for semantic urban scene understanding (2016)
18. MMSegmentation Contributors. MMSegmentation: Openmmlab semantic segmentation toolbox and benchmark (2020). https://github.com/open-mmlab/mmsegmentation
19. He, T., et al.: Bag of tricks for image classification with convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2019)
20. Zhou, Z., RahmanSiddiquee, M.M., Tajbakhsh, N., Liang, J.: Unet++: a nested u-net architecture for medical image segmentation. In: Stoyanov, Danail, et al. (eds.) DLMIA/ML-CDS -2018. LNCS, vol. 11045, pp. 3–11. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_1
21. Sun, K., et al.: High-resolution representations for labeling pixels and regions arXiv:1904.04514 (2019)

A Multi-band Frequency Conversion Scheme Based on Single Optical Frequency Comb for High-Throughput Satellite Applications

Bin Wu(✉), Chaoyue Zheng, and Qingchun Zhao

School of Computer and Communication Engineering, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China
wubin@neuq.edu.cn

Abstract. A flexible and efficient multi-band frequency conversion scheme is proposed and demonstrated based on single optical frequency comb. The simulation results indicate that the scheme can realize up and down cross-band frequency conversion simultaneously. The received 10 GHz signal (X-band) can be down-converted to S, C bands and up-converted to Ku, Ka, U and V band. In addition, it is also verified that a 28 GHz (Ka-band) input signal can be down-converted and up-converted to different satellite communication bands simultaneously. Moreover, the influence of the DC bias drift of MZM1 on the power of the frequency-converted signal is also investigated. The research shows that the system can adapt the DC drifting. The proposed scheme is proven to be reconfigurable, flexible and easy to implement for high-throughput satellite applications.

Keywords: Microwave photonics · Optical frequency comb · Multi-band frequency conversion

1 Introduction

With the increasing demands for high speed, large capacity and wide instantaneous bandwidths in high-throughput satellite communication, RF systems are driven to work at higher frequency and larger bandwidths nowadays [1]. Thus, RF repeaters are driven to convert signals among C, Ku, Ka and even V bands to mitigate the frequency spectrum congestion and orbital resource depletion problems [2]. Therefore, satellite repeaters with good performance play an important role in multi-band communication and have been extensively studied in recent years. However, traditional multi-band RF system faces with many electronic bottlenecks such as heavy weight, large size, low bandwidth, high power consumption and severe electro-magnetic interference (EMI) [3]. Fortunately, microwave photonics with its inherent advantages of wide bandwidth, low loss, small size and immunity to electromagnetic interference, can provide functions in microwave systems that are very complex or even impossible to implement directly in the radiofrequency (RF) domain [4]. Compared with conventional electronic mixers, photonics-based mixers can achieve frequency mixing within multiple frequency

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 135–142, 2023.
https://doi.org/10.1007/978-981-99-1260-5_17

bands and have the advantages of wide frequency coverage, broad instantaneous bandwidths, small frequency-dependent loss, and immunity to electromagnetic interference [5]. As one of the important research contents of microwave photonics, the optical frequency comb has received extensive attention since its invention. Most of the existing frequency conversion schemes can only realize the conversion of one single frequency. It is precisely because of the emergence of the optical frequency comb that multi-band frequency conversion can be realized. In recent years, many methods based on the OFC have been proposed to realize frequency conversion, which include single-OFC [6–9] and dual-OFCs [10–12]. In [10], the frequency conversion scheme composed of two OFCs with different center frequencies and WDM is proposed, which achieves frequency conversion in different channels to avoid harmonic interference. However, it is difficult to control the stability of two OFCs simultaneously and the use of WDM limits the flexibility of the system. In [12], the scheme based on two OFCs can only realize down frequency conversion. In contrast, the scheme based on a single OFC is more flexible and easy to control. However, most of the existing schemes based on single OFC cover less frequency bands. In [9], the scheme based on a single OFC can cover more frequency bands. Nevertheless, the center frequency offset of the local oscillator OFC needs to be large so that the obtained signal after frequency conversion covers more frequency bands. In this scheme, higher frequency of the RF source is required, and the system structure is more complicated. By contrast, since the scheme proposed in this paper adopts the idea of spectrum expansion, the OFC with a wider spectrum can be realized when the frequency of the RF source is relatively small.

On the basis of fully studying the advantages and disadvantages of the existing schemes, we propose a novel multi-band frequency conversion scheme based on a single OFC. The frequency conversion system is mainly composed of one DPMZM and two MZMs. In this scheme, a 10 GHz input microwave signal can be converted to 4 GHz (S band), 8 GHz (C band), 16 GHz (Ku band), 20 GHz (K band), 28, 32, 40 GHz (Ka band), 44 GHz (U band), and 52, 64 GHz (V band) simultaneously. By adjusting the spacing of OFC, the 28 GHz received microwave signal of Ka band can be converted to 10 GHz (X band), 14, 18 GHz (Ku band), 22, 26 GHz (K band), 30, 34, 38 GHz (Ka band), 42, 46 GHz (U band), and 56 GHz (V band) at the same time. The results show that the system has low structural complexity, can be reconfigured, and can simultaneously realize multi-band up-conversion and down-conversion with high efficiency.

2 Principle

The schematic diagram of the proposed multi-band frequency conversion system is shown in Fig. 1. It mainly consists of a laser diode, a dual-parallel Mach-Zehnder modulator (DPMZM), two Mach-Zehnder modulators (MZM) and a photodetector (PD).

A continuous light wave $E_{in}(t) = E_c \exp(\mathrm{j}\omega_c t)$ emitted by LD is divided into two branches by the optical power splitter. In the upper part, DPMZM cascade MZM2 to form an optical frequency comb generator. Firstly, one light wave is employed as the optical carrier of the input RF signal, and the RF signal is converted to the optical domain through a DPMZM. The DPMZM is an integrated modulator, which consists of two sub-MZMs and one main MZM. The MZM-x is driven by the RF1 signal to operate at the

maximum transmission point (MATP) to suppress odd-order optical sidebands and the MZM-y only transmits optical carrier without RF signal applied. By adjusting the value of *VDC*, a 180° phase difference is introduced between the output light waves of MZM-x and MZM-y.

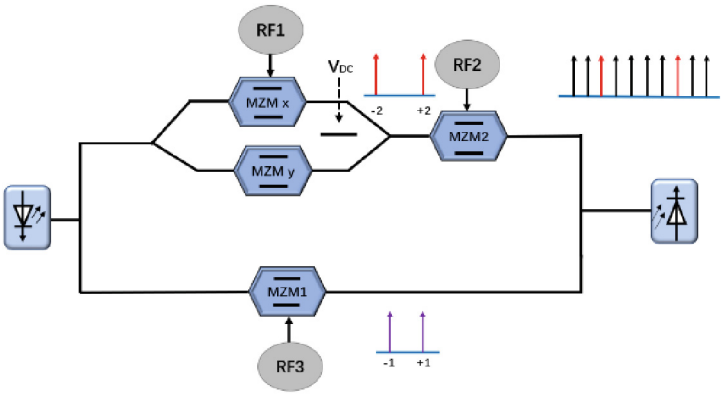


Fig. 1. Schematic diagram of multi-band frequency conversion scheme. LD: laser diode, MZM: Mach-Zehnder modulator, DPMZM: dual-parallel Mach-Zehnder modulator, RF: radio frequency signal, PD: photodiode.

The electrical field of input signal to the DPMZM is:

$$\frac{\sqrt{2}}{2}E_{in}(t) = \frac{\sqrt{2}}{2}E_c\exp(\mathrm{j}\omega_c t) \tag{1}$$

where *Ec* and ωc are the amplitude and angular frequency of the continuous light wave, respectively.

Because of MZM-x is operate at the MATP, the output electrical field of MZM-x can be expressed as [13]: $E_{out-x}(t) = \sqrt{2}/4 \times E_{in}(t)\cos[m\cos(\omega_{RF1}t)]$, where $m = \pi V_{RF1}/2V_{\pi-x}$ is the modulation index of MZM-x, while *VRF1* and $\omega RF1$ are the amplitude and angular frequency of the RF1 driving signal, $V\pi$-*x* is the half-wave voltage of MZM-x. And the electrical field at the output of the MZM-y can be simply expressed as $E_{out-y}(t) = \sqrt{2}/4 \times E_{in}(t)\cos(\varphi_y/2)$, where $\varphi_y = \pi V_{bias-y}/V_{\pi-y}$ is the optical phase difference between the two branches of the MZM-y. $V\pi$-*y* and *Vbias-y* are the half-wave voltage and the DC-bias voltage of MZM-y respectively.

By adjusting the bias voltage of *VDC*, the phase of the output of two sub MZMs is opposite. At the output of the main MZM, the electrical field can be simply expressed in the form of Bessel's function:

$$E_{out}(t) = \sqrt{2}/4 \times E_c\exp(j\omega_c t) \times \left\{J_0(m) + (-1)\cos(\varphi_y/2) + 2\sum_{n=1}^{n\to\infty}(-1)^n J_{2n}(m)\cos[2n\cos(\omega_{RF1}t)]\right\} \tag{2}$$

By adjusting the value of *Vbias-y*, the condition that $\cos(\varphi_y/2) = J_0(m)$ can be easily realized, the optical carriers transmitting in MZM-x and MZM-y will offset each other entirely.

$$E_{out}(t) = \sqrt{2}/4 \times E_c \exp(j\omega_c t) \times 2 \sum_{n=1}^{n\to\infty} (-1)^n J_{2n}(m) \cos[2n\cos(\omega_{RF1} t)] \tag{3}$$

When a moderately small modulation index m is taken, high-order optical sidebands can be ignored. The output electrical field of DPMZM can be simply expressed as:

$$E_{out-DPMZM}(t) = \sqrt{2}/4 \times E_c \exp\left[j(\omega_c t \pm 2\omega_{RF1} t)\right] J_{2n}(m) \tag{4}$$

The spectrum expansion is realized by the idea of canceling the carrier wave, which alleviates the limited bandwidth problem of the modulator in a certain extent. Then, the $\pm$ 2 order optical sideband outputs by the DPMZM are used as the input optical carrier of the MZM2. MZM2 works in push-pull mode, and its output can be simply expressed as:

$$\begin{aligned}
E_{out-MZM2}(t) =& \frac{1}{2} E_{out-DPMZM}(t) \exp\left[j\pi \frac{V_{DC2} + V_{RF2}\sin(\omega_{RF2} t + \varphi_{RF2})}{V_\pi} \right] \\
&+ \frac{1}{2} E_{out-DPMZM}(t) \exp\left[-j\pi \frac{V_{DC2} + V_{RF2}\sin(\omega_{RF2} t + \varphi_{RF2})}{V_\pi} \right] \\
=& \frac{1}{2} E_{out-DPMZM}(t) \exp(jn\omega_{RF2} t) \sum_{n=-\infty}^{n=+\infty} J_n\left(\frac{\pi V_{RF2}}{V_\pi} \right) \left[\exp\left(jn\varphi_{RF2} + j\frac{\pi V_{DC2}}{V_\pi} \right) \right. \\
&\left. + (-1)^n \exp\left(jn\varphi_{RF2} - j\frac{\pi V_{DC2}}{V_\pi} \right) \right] \\
=& \frac{1}{2} E_{out-DPMZM}(t) \exp(jn\omega_{RF2} t) \sum_{n=-\infty}^{n=+\infty} J_n(m_2) \exp(jn\varphi_{RF2}) \left[\exp(j\theta) \right. \\
&\left. + (-1)^n \exp(-j\theta) \right]
\end{aligned} \tag{5}$$

where *VRF2* and ωRF2 are the amplitude and angular frequency of the RF2 signal, $V\pi$ is the half-wave voltage of MZM2, $m2 = \pi VRF2 / V\pi$ is the modulation index of MZM2, *VDC2* is the DC-bias voltage of MZM2, $\theta = \pi V_{DC2}/V_\pi$ is phase change due to DC bias voltage.

By setting reasonable parameters, when $|E_0| = |E_{\pm 1}|$ and $4\omega RF1 = 3\omega RF2$ is satisfied, we can get 6 comb lines with good flatness. In addition, when $|E_0| = |E_{\pm 1}| = |E_{\pm 2}|$ and $4\omega RF1 = 5\omega RF2$ is satisfied, we can get 10 comb lines with good flatness. The optical frequency comb generator has good reconfigurability, and the comb line spacing can be flexibly tuned by changing the frequency of the RF signal. Then in the lower way light wave passing through the MZM1, which is operate at minimum transmission point to realize the CS-DSB, even-order sidebands are suppressed. When a moderately small modulation depth is taken, high-order optical sidebands can be ignored and only the ± 1 order optical sidebands is retained. Finally, two optical signals generated by upper and lower way are combined by optical coupler (OC), the multi-band frequency conversion signals are generated by frequency beating through a photodetector.

3 Simulation Results and Analysis

To demonstrate the proposed multi-band frequency conversion scheme, a simulation based on the structure of Fig. 1 is performed via commercial software OptiSystem15.0. In this scheme, a light wave emitted by LD with a center frequency of 193.1 THz, the linewidth of 10 MHz, and the power of 15 dBm is divided into two paths by a 3 dB optical power splitter. The half-wave voltage of the modulators in this paper is 4V. In the upper way, a DPMZM cascade a MZM to form an OFC generator. In the lower way, a MZM operate at minimum transmission point, only the ±1 order optical sidebands are output when a small modulation index is taken. By setting the parameters reasonably, the OFC generator can generate 6-line and 10-line flat OFCs, as shown in Table 1.

Table 1. The values of the parameters in OFC generator

	ωRF	*VRF1* (V)	*VRF2* (V)	*Vbias-x* (V)	*Vbias-y* (V)	*Vbias-2* (V)
6-Line OFC	4*ω*RF1 = 3*ω*RF2	1.834	0.637	0	2.54	3.4
10-Line OFC	4*ω*RF1 = 5*ω*RF2	1.834	2.344	0	2.54	1.274

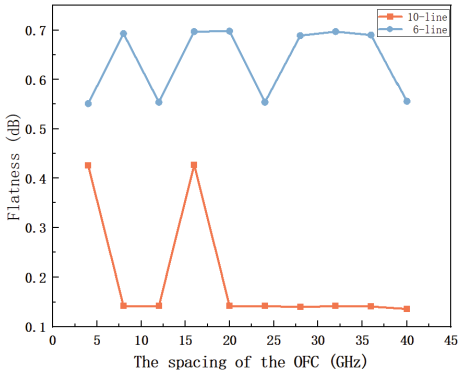


Fig. 2. Schematic diagram of the relationship between OFC flatness and comb line spacing

In addition, the spacing of OFC can be flexibly tuned by adjusting the frequency of the RF driving signal. We researched the flatness relationship of 6-line and 10-line OFCs with the same comb line spacing, the result is shown in Fig. 2. It can be easily seen that the flatness of the 10-line OFC is better than that of the 6-line OFC. And the flatness of the 10-line OFC tends to 0.14 dB as the OFC spacing increases, which has good flatness performance. Therefore, this paper will implement multi-band frequency conversion based on 10-line OFC.

In the simulation, a X band RF signal with frequency of 10 GHz is used to drive MZM1, and the RF voltage and DC bias of MZM1 are set to 1V and 4V respectively to

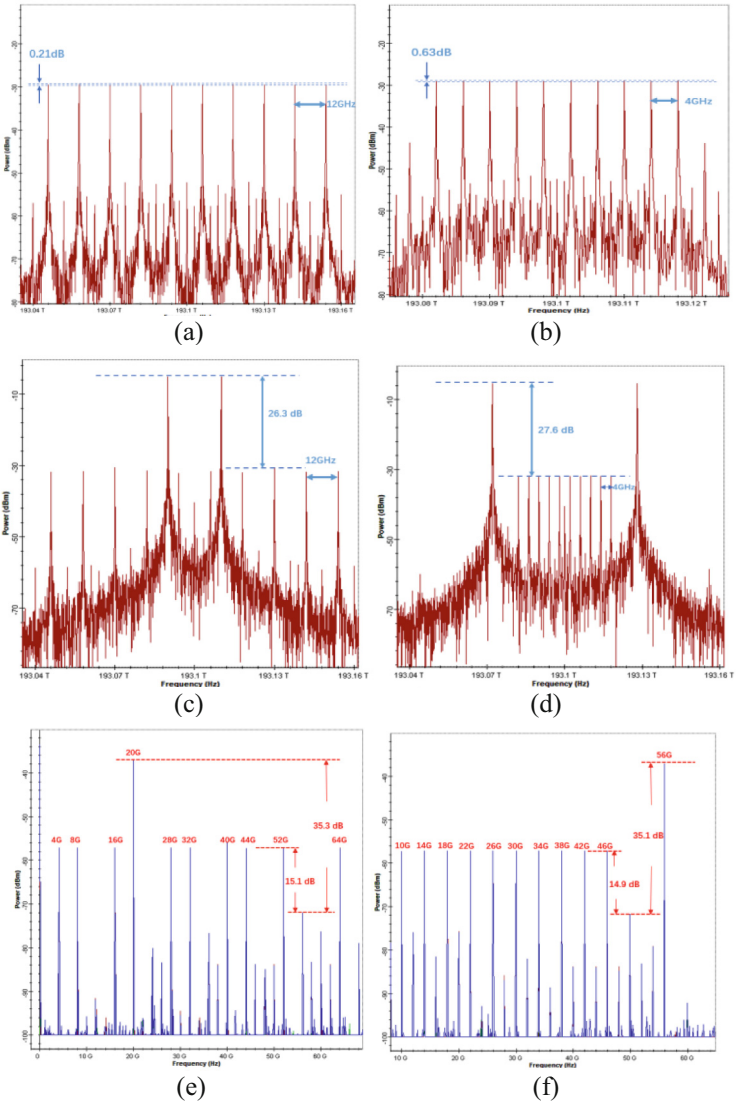
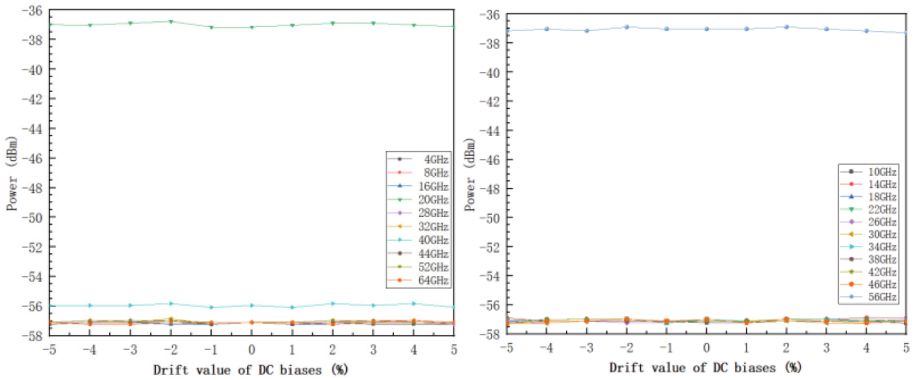


Fig. 3. The optical spectrum of 10-line OFC (a)(b) and combined signals (c)(d), and the electrical spectrum of converted signals (e)(f)

produce ±1 order optical sidebands. The frequency of RF1 and RF2 are set to 15 GHz and 12 GHz respectively to generate a 10-line OFC with frequency spacing of 12 GHz. A received X band frequency of 10 GHz can realize up and down cross-band frequency conversion simultaneously, which cover from S to V band. In Fig. 3(a), a 10-line OFC with flatness of 0.21 dB and frequency spacing of 12 GHz is produced. In Fig. 3(c), the power of 1 order sideband that is produced by RF3 signal is 26.3 dB higher than the 10-line OFC. In Fig. 3(e), the multi-band frequency conversion signals are achieved.

When a Ka band RF signal with frequency of 28 GHz is used to drive MZM1 and the frequency of RF1 and RF2 are set to 5 GHz and 4 GHz respectively, another multi-band frequency conversion result is obtained, as shown in Fig. 3 (b)(d)(f).

Moreover, how the DC drifting of MZM1 affects the power of converted frequencies is also investigated in this paper. We set the drift of DC biases of MZM1 between −5 and 5%. From Fig. 4, we can see that the power fluctuation of converted frequencies is always kept within 0.4 dB, indicating that the system has a good ability to deal with DC drift. In Fig. 4(a), the maximum power fluctuation of a converted 20 GHz signal is only 0.39 dB, which varies from −37.16 to −36.77 dB. The maximum power fluctuation of other converted frequencies is always kept within 0.26 dB. In Fig. 4(b), the maximum power fluctuation of a converted 56 GHz signal is only 0.39 dB, and the fluctuation range is −37.29 to −36.90 dB. The maximum power fluctuation of other converted frequencies is always kept within 0.30 dB. It can be seen that the system can well cope with the DC drifting.



(a) (b)

Fig. 4. Power fluctuation of converted frequencies when the drift of DC biases varies from −5 to 5%

4 Conclusion

A flexible and efficient OFC-based multi-band frequency conversion scheme without any optical filter is proposed in this paper. The proposed scheme proves to be reconfigurable, flexible and easily implemented. It has the advantages of realizing up and down frequency conversion simultaneously and covering wide bands. Meanwhile, the system can cope well with DC bias drift. The system has potential application value to high-throughput satellite transponders in the future.

Acknowledgement. This work is supported in part by the National Science Foundation of China (NSFC) under Grant 62101113, China; the Natural Science Foundation of Hebei Province (F2020501035); the Fundamental Research Funds for the Central Universities (N2023008).

References

1. Tavik, G.C., et al.: The advanced multifunction RF concept. IEEE Trans. Microw. Theory Tech. **53**(3), 1009–1020 (2005)
2. Panagopoulos, A.D., Arapoglou, P.D.M., Cottis, P.G.: Satellite communications at Ku, Ka, and V bands: propagation impair-ments and mitigation techniques. IEEE Commun. Surv Tuts **6**(3), 2–14 (2004)
3. Yao, J.: Microwave photonics. J. Lightwave Technol. **27**(3), 314–335 (2009)
4. Capmany, J., Novak, D.: Microwave photonics combines two worlds. Nat. Photonics **1**(6), 319–330 (2007)
5. Photonics-Based Microwave Frequency Mixing: Methodology and Applications. Laser Photonics Rev. **14**(1), 1800350.1–1800350.25 (2020). https://doi.org/10.1002/lpor.201800350
6. Lin, T., Zhao, S., Zheng, Q., Zhu, Z., Li, X., Qu, K.: Photonic microwave multi-band frequency conversion based on a dp-qpsk modulator for satellite communication. Opt. Rev. **24**(3), 1–8 (2017)
7. Wu, B., Yin, H., Liu, A., Ji, X., Zhao, Q.: A multi-band transparent transponder based on reconfigurable optical frequency combs for crossconnection of inter-satellite and satellite-ground. IEEE Photonics J. **PP**(99), 1–1 (2018)
8. Li, H., Zhao, S., Lin, T., Zhang, K., Jiang, W., Wang, G., et al.: Photonic phase shifter with full tunable range and multi-band frequency-conversion feature based on a pdm-dpmzm. Opt. Rev. **26**(6), 681–692 (2019)
9. Li, X., Liu, X., Li, H., Wei, Y.: Multi-band Frequency Conversion Scheme Employing Single Optical Frequency Comb. In: Asia Communications and Photonics Conference (2020)
10. Yang, X., Xu, K., Yin, J., et al.: Optical frequency comb based multi-band microwave frequency conversion for satellite applications. Optics express (2014)
11. Tao, L., Zhao, S., Zhu, Z., Xuan, L., Qu, K.: Photonic microwave multi-band frequency conversion scheme based on dual-OFCs for satellite communication. International Conference on Optical Communications & Networks. IEEE (2016)
12. Ji, Y., Li, H., Tian, M., Gong, C., Wei, Y.: Microwave Photonics Down Conversion Based on Optical Frequency Comb. In: 2020 IEEE 5th Optoelectronics Global Conference (OGC). IEEE (2020)
13. Zhou, F., et al.: Photonic generation of frequency quadrupling signal for millimeter-wave communication. Optics Commun. **304**, 71–74 (2013)

Energy-Efficient UAV Trajectory Plan for 6G Networks

Peng Qin[1,2(✉)], Xue Wu[1,2], Xiongwen Zhao[1,2], and Honghao Zhao[1,2]

[1] State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China
qinpeng@ncepu.edu.cn

[2] Hebei Key Laboratory of Power Internet of Things Technology, North China Electric Power University, Baoding 071003, China

Abstract. Leveraging UAVs for access and Altitude Platform Stations (HAPSs) for data backhaul to construct the Air-Ground Integrated Network (AGIN) is a feasible solution to achieve seamless network coverage for remote IoT devices in future 6G era. However, the limited battery of IoT terminals and constrained onboard energy storage of UAVs make system energy-efficiency becomes a new concern. To cope with the challenge, we propose a C-NOMA AGIN model for remote area. Then, we investigate the UAV trajectory plan problem for maximizing system energy efficiency (EE). We provide the solution to obtain the near-optimal UAV trajectory and flight speed. Results prove that the proposed approach is superior to others in terms of EE.

Keywords: Trajectory plan · Air-ground integrated 6G network (AGIN 6G) · NOMA · Energy Efficiency (EE)

1 Introduction

Although the fifth generation (5G) mobile communication can meet the demand of Internet of Things (IoT) applications in hot spots, there is still urgent need for more economical and efficient network coverage in remote districts. Meanwhile, to build the space-air-ground integrated network (SAGIN) is proposed in the vision of the sixth generation (6G) mobile communication, which can achieve seamless network coverage and ubiquitous terminal access [1].

This work was supported in part by the National Natural Science Foundation of China under Grant 62201212, the Natural Science Foundation of Hebei Province under Grant F2022502017, the Zhejiang Key Laboratory of Medical Electronics and Digital Health under Grant MEDH202215, the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources under Grant LAPS21018, the Fundamental Research Funds for the Central Universities under Grant 2021MS002, and the National Key Research and Development Program of China under Grant 2020YFB1806000.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 143–151, 2023.
https://doi.org/10.1007/978-981-99-1260-5_18

The SAGIN usually consists of the space-based segment, the air-based segment and the ground segment [2]. Air-based unmanned aerial vehicle (UAV) network has more advantages, such as lower constructive cost, higher line-of-sight (LoS) capability and flexible deployment characteristics, which can supply access for IoT terminals in remote districts, high altitude platform station (HAPS) ensures much wider coverage [3]. Paper [4] proposed the UAV-assisted SAGIN model, which mainly analyzed the performance of UAV-assisted network, and through network optimization, the throughput and spectral efficiency of the system were improved. It can be seen that heterogeneous AGIN has become an effective scheme to solve the network coverage problem and improve the system performance, which provides an important reference for the study of this paper.

Related studies had been carried out on NOMA technology [5, 6]. Since traditional orthogonal multiple access (OMA)-based technology is no longer adequate, non-OMA (NOMA) that can recover information using successive interference cancellation (SIC) at the receiving end and achieve better spectrum efficiency, has become the focus [7]. Paper [5] proposed a NOMA technology combined with orthogonal frequency division multiplexing (OFDM) technology, which reduced the bit error rate and improved the system performance based on cyclic stationary equalization technology. However, as the increasing of users' number, the complexity of SIC greatly increases. Therefore, by dividing terminals into multiple clusters, clustered-NOMA (C-NOMA) is an effective solution to both improve system throughput while balancing spectrum efficiency and terminal complexity.

In heterogeneous networks, it is a significant concern to plan UAV trajectory to enhance the system performance considering the constrained on-board energy storage of UAV [8, 9, 10]. Paper [9] proposed an integrated air-ground network system supporting XAPS, which maximized the uplink reachable rate of the system through spectrum, power and position optimization of UAVs based on matching theory. Paper [10] proposed a UAV-assisted space-ground remote IoT network architecture, optimized the deployment of UAVs through successive convex approximation (SCA) method, and solved the energy-saving communication problem between UAVs and ground terminals. The above

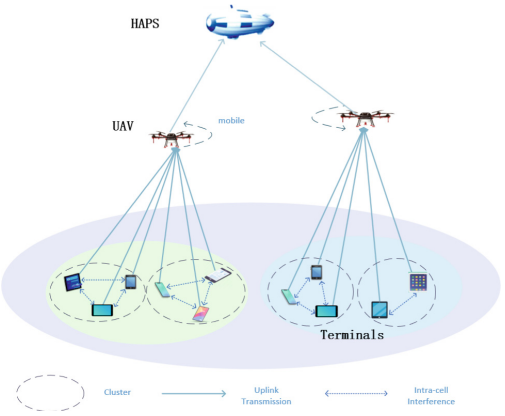


Fig. 1. C-NOMA-enabled AGIN 6G system model.

work mostly only considered communication energy consumption, however less likely to involve flight energy consumption, which have a greater effect on the capabilities of the UAV communication. So considering the flight trajectory and speed control, the propulsion energy consumption model of UAV is introduced to solve the optimization problem of trajectory and velocity of UAV, and achieve high EE and wide coverage of the system.

To cope with the above challenges for future remote IoT applications, we first put forward a C-NOMA AGIN model in this paper, including three segments of UAVs for massive access, HAPS for backhaul, and remote ground terminals. C-NOMA technology has stronger processing power and can achieve a balance between spectrum efficiency and complexity. Then, considering the finite battery of IoT terminals and limited energy storage on UAVs [11-13], we investigate the trajectory plan problem for maximizing system energy efficiency (EE). Solution is provided according to SCA and we obtain the near-optimal UAV trajectory. Simulations demonstrate that our approach is superior to others in terms of EE.

2 System Model

As shown in Fig. 1, we propose the C-NOMA-enabled AGIN 6G system model, including $\mathcal{U} = \{U_1, U_2, \ldots, U_M\}$ UAVs with flight altitude $\hat{h}_m$ for access, which share a channel with bandwidth of W, one HAPS with altitude H_A for backhaul, and $\mathcal{D} = \{D_1, D_2, \ldots, D_N\}$ terminal devices sharing a channel with bandwidth of W. The frequency resource is evenly divided into M blocks so that each UAV occupies different resource block without interference, and each block is further divided into $\mathcal{K} = \{S_1, S_2, \ldots, S_K\}$ subchannels. UAV serves its associated IoT devices using C-NOMA technology by dividing terminals into clusters. The flight period involves $\mathcal{T} = \{T_1, T_2, \ldots, T_L\}$ time slots, each of which is ΔT, and UAV flies back to the starting point when each cycle ends. Let $\mathbf{q}_m(t) = (x_m(t), y_m(t))$, $\mathbf{v}_m(t) \triangleq \dot{\mathbf{q}}_m(t)$, and $\mathbf{ac}_m(\mathrm{t}) \triangleq \ddot{\mathbf{q}}_m(t)$ respectively denote the location, speed and acceleration of the m th UAV in slot t. The position of the n th terminal is represented by $\mathbf{D}_n = (x_n, y_n, 0)$. Communication between terminal and UAV, UAV and HAPS is dominated by LoS links, and channel fading follows Rice fading [4, 10].

2.1 Communication Model

The distance between UAV m and user n in slot t is.

$$d_{n,m}(t) = \sqrt{\hat{h}_m^{\,2} + \|\mathbf{q}_m(t) - \mathbf{D}_n\|^2}, \forall n \in \mathcal{D}, m \in \mathcal{U}, t \in \mathcal{T}. \tag{1}$$

Then, the channel power gain occupying the k th subchannel can be expressed as.

$$h_{n,m,k}(t) = \frac{G_m G_n \beta_0}{d_{n,m}^2(t)}, \forall n \in \mathcal{D}, m \in \mathcal{U}, k \in \mathcal{S}, t \in \mathcal{T}, \tag{2}$$

where G_m and G_n are the directional antenna gains of UAV and user, respectively. β_0 represents the channel power gain at the reference distance $d_0 = 1\mathrm{m}$. Since HAPS is far

from UAVs, the distance can be replaced by H_A. Therefore, the channel gain from UAV m to HAPS is calculated by.

$$h_m = \frac{G_m G_2 \beta_0}{{H_A}^2}, \forall m \in \mathcal{U}, \tag{3}$$

where G_2 is the directional antenna gain of HAPS.

Suppose that in the t th slot, the transmission power from the n th user to the m th UAV is $P_{n,m,k}(t) \in \mathbf{P}^{\mathrm{D}\to\mathrm{U}}$, and the transmission power from the m th UAV to the HAPS is $P_m(t) \in \mathbf{P}^{\mathrm{U}\to\mathrm{H}}$. UAV trajectory, velocity and acceleration are defined as $\mathbf{Q} = \{\mathbf{q}_m(t), \forall m \in \mathcal{U}, t \in \mathcal{T}\}$, $\mathbf{V} = \{\mathbf{v}_m(t), \forall m \in \mathcal{U}, t \in \mathcal{T}\}$, $\mathbf{AC} = \{\mathbf{ac}_m(\mathrm{t}), \forall m \in \mathcal{U}, t \in \mathcal{T}\}$, respectively. The channel selection indicator between terminal and UAV is $\mathbf{A} = \left\{a_{n,m,k}(t), \forall n \in \mathcal{D}, m \in \mathcal{U}, k \in \mathcal{K}, t \in \mathcal{T}\right\}$, i.e., if terminal n connects with UAV m through the k th subchannel in the t th time slot, $a_{n,m,k}(t) = 1$; otherwise $a_{n,m,k}(t)$=0.

Then, the SINR between UAV m and terminal n occupying the k th subchannel can be calculated by

$$\gamma_{n,m,k}(t) = \frac{P_{n,m,k}(t) h_{n,m,k}(t)}{I_{n,m,k}(t) + N_0}, \tag{4}$$

where N_0 represents the noise power, $I_{n,m,k}(t) = \sum_{j=n+1}^{N} a_{j,m,k}(t) P_{j,m,k}(t) h_{j,m,k}(t)$ represents interference from other users in the same channel. SIC technology can be used to decode signals from terminal devices occupying the same sub-channel.

Likewise, the SINR between UAV m and HAPS can be calculated by.

$$\gamma_m(t) = \frac{P_m(t) h_m}{N_0}. \tag{5}$$

In the t th time slot, the uplink rate of user n to HAPS can be denoted by.

$$R_{n,m,k}(t) = \frac{W}{MK} log_2\left(1 + \frac{\gamma_{n,m,k}(t)\gamma_m(t)}{1 + \gamma_{n,m,k}(t) + \gamma_m(t)}\right). \tag{6}$$

The system capacity can be expressed as.

$$R = \frac{W}{MK} \sum_{t=1}^{L} \sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{\mathrm{n}=1}^{N} a_{n,m,k}(t) log_2\left(1 + \frac{\gamma_{n,m,k}(t)\gamma_m(t)}{1 + \gamma_{n,m,k}(t) + \gamma_m(t)}\right). \tag{7}$$

2.2 Power Consumption Model

UAV power consumption includes two parts. One is the power associated with communication, including radiation, signal processing, and power generated by other circuits. The other is propulsion power consumption to maintain flight and maneuverability [14]. Fixed-wing UAVs used in this paper normally have no sudden deceleration and do not require engine to generate reverse thrust for the forward movement. Therefore, propulsion power consumption is a function of trajectory $\mathbf{q}_m(t)$ denoted by

$$P = \sum_{t=1}^{L} (P_d(t) + P_k(t)), \tag{8}$$

where

$P_d(t) = \sum_{m=1}^{M}\left(c_1\|\mathbf{v}_m(t)\|^3 + \frac{c_2}{\|\mathbf{v}_m(t)\|}\left(1+\frac{\|\mathbf{ac}_m(t)\|^2-\frac{\left(\mathbf{ac}_m^T(t)\mathbf{v}_m(t)\right)^2}{\|\mathbf{v}_m(t)\|^2}}{g^2}\right)\right)$ represents the power required for overcoming drag, g denotes gravity, $P_k(t) = \sum_{m=1}^{M}\Delta k_m/\Delta T$, and $\Delta k_m = \frac{1}{2}m_m^{'}\left(\|\mathbf{v}_m(L)\|^2 - \|\mathbf{v}_m(1)\|^2\right)$ denotes the kinetic energy of flight in a cycle. $m_m^{'}$ is quality of UAV m, and c_1, c_2 are two parameters related to the weight of aircraft, air density, etc.

Then, the system EE is denoted by

$$\lambda = \frac{R}{P} \tag{9}$$

3 Problem Formulation

The objective is to maximize the uplink system EE for remote IoT terminals in AGIN by designing UAV trajectory plan and speed control, formulated as

$$\mathbf{P}0: \max_{\mathbf{Q},\mathbf{V}} \lambda$$

$$C1: \sum_{m=1}^{M}\sum_{k=1}^{K} a_{n,m,k}(t)R_{n,m,k}(t) \geq R_{min}(t) \tag{10}$$

$$C2: \sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} a_{n,m,k}(t)R_{n,m,k}(t) \leq R_{max}^{H}(t) \tag{11}$$

$$C3: V_{min} \leq \|\mathbf{v}_m(t)\| \leq V_{max} \tag{12}$$

$$C4: \|\mathbf{ac}_m(t)\| \leq ac_{max} \tag{13}$$

Constraint (10) guarantees the system QoS. (11) represents the capacity constraint of the system with R_{max}^{H}, which guarantees that the total data transmitted cannot exceed the system capacity. (12), (13) denotes the speed and acceleration constraint of UAV, respectively.

4 Algorithm Analysis

The UAV trajectory, velocity and acceleration are all continuous variables, making **P**0 difficult to be solved directly. Therefore, we propose the algorithm for solving it based on SCA. The algorithm starts from line 3 with iteration index w. Line 4–5 achieves the best trajectory $(\mathbf{q}_m^w(t))^*$, speed $(\mathbf{v}_m^w(t))^*$, and acceleration $(\mathbf{ac}_m^w(t))^*$ for the current loop. It continues until the increase of the objective value is below a threshold ε. Details are shown below.

The denominator's upper bound of the objective function λ obtained from (8), the lower bound of system EE is

$$\lambda \geq \lambda_{lb} = \frac{R}{\sum_{t=1}^{L}\sum_{m=1}^{M}\left(c_1\|\mathbf{v}_m(t)\|^3 + \frac{c_2}{\|\mathbf{v}_m(t)\|}\left(1+\frac{\|\mathbf{ac}_m(t)\|^2}{g^2}\right) + \Delta k_m/\Delta T\right)}. \quad (14)$$

We can obtain the sub-optimal result based on first-order and second-order Taylor expansion:

$$\mathbf{v}_m(t+1) = \mathbf{v}_m(t) + \mathbf{ac}_m(t)\Delta T \quad (15)$$

$$\mathbf{q}_m(t+1) = \mathbf{q}_m(t) + \mathbf{v}_m(t)\Delta T + \frac{1}{2}\mathbf{ac}_m(t)\Delta T^2. \quad (16)$$

Notice that constraints (10)–(12) are non-convex. Both numerator and denominator of **P**0 are non-convex making the objective function non-convex. Therefore, we introduce the relaxation variable $\{\tau_m(t)\}$ and restate **P**0 as

$$\mathbf{P}1: \underset{\mathbf{Q},\mathbf{V}}{max} \frac{R}{\sum_{t=1}^{L}\sum_{m=1}^{M}\left(c_1\|\mathbf{v}_m(t)\|^3 + \frac{c_2}{\tau_m(t)} + \frac{c_2\|\mathbf{ac}_m(t)\|^2}{g^2\tau_m(t)} + \Delta k_m/\Delta T\right)}$$

$$\text{s.t.} C1 - C4$$

$$C5: \tau_m(t) \geq V_{min} \quad (17)$$

$$C6: \|\mathbf{v}_m(t)\|^2 \geq \tau_m(t)^2. \quad (18)$$

By this reformulation, the denominator of the objective function is convex for $\{\mathbf{v}_m(t), \mathbf{ac}_m(t), \tau_m(t)\}$, but new non-convex constraint (18) is created. To solve this non-convex constraint, we apply a locally convex approximation. With regard to any local point $\mathbf{v}_m^r(t)$ obtained in the r th iteration, we have

$$\|\mathbf{v}_m(t)\|^2 \geq \|\mathbf{v}_m^r(t)\|^2 + 2(\mathbf{v}_m^r(t))^T(\mathbf{v}_m(t) - \mathbf{v}_m^r(t)) \triangleq \psi_{lb}(\mathbf{v}_m(t)). \quad (19)$$

Thus, we define new constraints,

$$\psi_{lb}(\mathbf{v}_m(t)) \geq {\tau_m}^2(t). \quad (20)$$

Then, for the sake of solving the non-concavity of the numerator, it is transformed into

$$R = \frac{W}{MK}\sum_{t=1}^{L}\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} a_{n,m,k}(t)\left(\widehat{\gamma}_{n,m,k}(t) - \widehat{\gamma}'_{n,m,k}(t)\right) \quad (21)$$

$$\widehat{\gamma}_{n,m,k}(t) = log_2\left(\Omega_{n,m,k}(t) + P_{n,m,k}(t)h_{n,m,k}(t)P_m(t)h_m + \Theta_m(t)\right) \quad (22)$$

$$\widehat{\gamma}'_{n,m,k}(t) = log_2\big(\Omega_{n,m,k}(t) + \Theta_m(t)\big) \tag{23}$$

where $\Theta_m(t) = N_0^2 + N_0 P_m(t) h_m$

$$\Omega_{n,m,k}(t) = N_0 \sum_{j=n+1}^{N} \rho_{j,m,k}(t) h_{n,m,k}(t) + N_0 P_{n,m,k}(t) h_{n,m,k}(t) \sum_{j=n+1}^{N} \rho_{j,m,k}(t) h_{j,m,k}(t).$$

By using SCA method, the nonconvex problem is transformed into convex optimization problem. Owing to the space limitation, the derivation details are not shown, and similar derivation can be found in [6]. Finally, the numerator is concave, the denominator is convex, and all constraints are convex, making **P**0 transformed into a standard convex problem. Thus, using CVX can solve this problem.

Algorithm EE maximum SCA-based UAV Trajectory optimization algorithm

1: **Input:** $(\mathbf{q}_m(t))^{\mathbf{0}}, \varepsilon, w$
2: **Output:** $\mathbf{Q}^*, \mathbf{V}^*$
3: **repeat**
4: settle problem (**P0**) by using CVX to obtain $(\mathbf{q}_m^w(t))^*, (\mathbf{v}_m^w(t))^*$.
5: **update** $\mathbf{q}_m^{w+1}(t) = (\mathbf{q}_m^w(t))^*, \mathbf{v}_m^{w+1}(t) = (\mathbf{v}_m^w(t))^*$.
6: **update** iteration index w=w+1
7: **until** $|(\lambda'_{lb})^w - (\lambda'_{lb})^{w-1}| < \varepsilon$.
8: **return** $\mathbf{Q}^*, \mathbf{V}^*$.

5 Simulation Results

In this section, we make a comparison between the proposed algorithm and random trajectory method (Random) and OMA method using MATLAB. We consider 3 UAVs and 20 terminals randomly distributed in an area of 200 m × 120 m, each NOMA cluster contains two terminals.

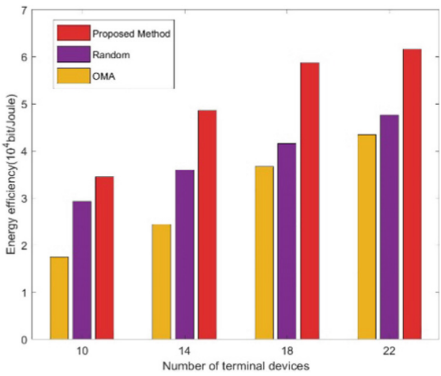


Fig. 2. The system EE for different benchmark algorithms.

Figure 2 shows the system EE for different benchmark algorithms. It is shown that as the number of users increases, so does energy efficiency. Moreover, compared with

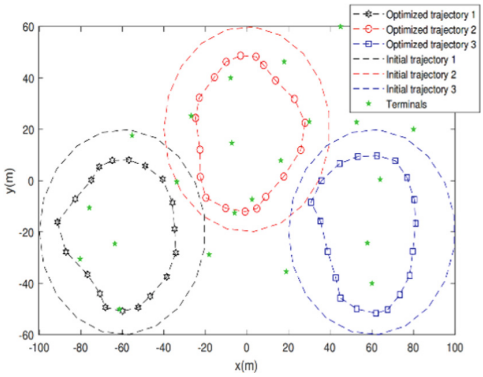


Fig. 3. Optimized trajectories for all UAVs.

OMA method and random trajectory method, our proposed method is always superior in terms of EE.

Figure 3 shows the optimized flight trajectories for three UAVs. It is shown that with the optimized trajectory, the UAV can get close to each terminal device as much as possible, so as to achieve better system EE and provide superior QoS.

6 Conclusion

In this paper, a C-NOMA-based heterogeneous AGIN network is proposed and we formulate the problem of maximizing the system EE by trajectory optimization. We design an algorithm to achieve the maximum system EE. Both theoretical analysis and results show that the proposed method is superior to others in terms of EE.

References

1. Li, B., Fei, Z., Zhang, Y.: UAV communications for 5G and beyond: recent advances and future trends. IEEE Internet Things J. **6**(2), 2241–2263 (2019)
2. Liu, J., Zhao, X., Qin, P., Geng, S., Meng, S.: Joint dynamic task offloading and resource scheduling for WPT enabled space-air-ground power internet of things. IEEE Trans. Network Sci. Eng. **9**(2), 660–677, March 2022
3. Qin, P., Fu, Y., Zhao, X., Wu, K., Liu, J., Wang, M.: Optimal task offloading and resource allocation for C-NOMA heterogeneous air-ground integrated power-IoT networks. IEEE Trans. Wireless Commun. (2022). https://doi.org/10.1109/TWC.2022.3175472
4. Wang, J., Jiang, C., Wei, Z., Pan, C., Zhang, H., Ren, Y.: Joint UAV hovering altitude and power control for space-air-ground IoT networks. IEEE Internet Things J. **6**(2), 1741–1753 (2019)
5. Datta, J., Lin, H.: Detection of uplink NOMA systems using joint SIC and cyclic FRESH filtering. In: 27th Wireless and Optical Communication Conference (WOCC), pp. 1–4 (2018)
6. Li, Y., Zhang, H., Long, K., Jiang, C., Gzuizani, M.: Joint resource allocation and trajectory optimization with QoS in UAV-based NOMA wireless networks. IEEE Trans. Wireless Commun. **20**(10), 6343–6355 (2021)

7. Kim, J., Kim, J., Park, S.: Joint design of power control and SIC decoding order for max-min fairness optimization in uplink NOMA systems. In: International Conference on Information Networking (ICOIN), pp. 339–342 (2021)
8. Qin, P., Zhu, Y., Zhao, X., Feng, X., Liu, J., Zhou, Z.: Joint 3D-location planning and resource allocation for XAPS-enabled C-NOMA in 6G heterogeneous internet of things. IEEE Trans. Veh. Technol. **70**(10), 10594–10609 (2021)
9. Zeng, Y., Zhang, R.: Energy-efficient UAV communication with trajectory optimization. IEEE Trans. Wirel. Commun. **16**(6), 3747–3760, June. 2017
10. Na, Z., Liu, Y., Shi, J., Liu, C., Gao, Z.: UAV-supported clustered NOMA for 6G-enabled internet of things: trajectory planning and resource allocation. IEEE Internet Things J. **8**(20), 15041–15048 (2021)
11. Qin, P., Fu, Y., Tang, G., Zhao, X., Geng, S.: Learning based energy efficient task offloading for vehicular collaborative edge computing. IEEE Trans. Veh. Technol. (2022). https://doi.org/10.1109/TVT.2022.3171344
12. Qin, P., Fu, Y., Feng, X., Zhao, X., Wang, S., Zhou, Z.: Energy efficient resource allocation for parked cars-based cellular-V2V heterogeneous networks. IEEE Internet Things J. **9**(4), 3046–3061 (2022)
13. Qin, P., Wang, M., Zhao, X., Geng, S.: Content service oriented resource allocation for space-air-ground integrated 6g networks: a three-sided cyclic matching approach. IEEE Internet Things J. (2022). https://doi.org/10.1109/JIOT.2022.3203793
14. Wu, X., Wei, Z., Cheng, Z., Zhang, X.: Joint optimization of UAV trajectory and user scheduling based on NOMA technology. In: IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6 (2020)

Survey on Decoding Schemes of the Polar Code Using the Deep Learning

Jida Yang[1] and Yuan Liu[1,2(✉)]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
xieyp@dlut.edu.cn

[2] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China

Abstract. In the era of 5G Internet of everything, communications between different machines have strict requirements on latency and reliability. Due to the large amount of computations of the decoding algorithm of the polar code, the decoding time is too long, which cannot meet the low-latency system requirements. Deep learning has the ability of parallel computing, associative memory, self-organization and self-adaptation. Therefore, it is of great theoretical significance to apply deep learning to the polar code decoding algorithm. This paper reviews the development of the deep learning, and indicate that deep learning algorithms have been widely used in various fields. The traditional decoding method of the polar code is introduced, and the advantages and disadvantages of different decoding methods are compared.

Keywords: Polar code · Deep learning · Channel decoding

1 Introduction

The decoding algorithm of the polar code has a large amount of computation, which leads to a long decoding time. Deep learnings with the ability of parallel operation, associative memory, self-organization and self-adaptation have advanced tremendously and have been applied to many fields [1, 2]. Therefore, it is of practical significance to study the decoding method of polar codes based on neural network. This paper firstly introduces the development of deep learning, describes the traditional decoding methods of polar codes, discusses the advantages and disadvantages of each method. Secondly, we describe the traditional decoding methods of polar codes. Finally, the future research directions are analyzed.

2 The Development Process of Deep Learnings

Deep learning is a type of machine learning (ML) that can mimic how humans acquire certain types of knowledge. Deep learning is an essential element of data science and

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 152–157, 2023.
https://doi.org/10.1007/978-981-99-1260-5_19

can be considered a way to automate predictive analytics [3]. Deep learning makes this process faster and easier. Traditional machine learning algorithms are linear. However, deep learning algorithms are stacked in different hierarchies to simulate a mathematical model of the central nervous system [4].

In 1943, McCulloch formed a manual network by mathematically extensively connecting processing units. This network is used to simulate a simple structure and function that approximates the human brain nervous system [5]. This is the starting point for the development of neural networks. This model has implemented simple logic for class neurons. In 1969, the American mathematician Minsky proved that the perceptron is essentially a linear model. Only linear classification problems can be handled, not xor problems [6]. As a result, many scholars have given up on further research. This has also led to a nearly 20-year stagnation in the study of neural networks.

In 1982, Professor Hopfield proposed the Hopfield network model, which is a single-layer feedback neural network, explaining the form of recurrent neural network [7]. The emergence of the Hopfield network breaks the application limitations of the perceptron. The model is a recurrent neural network that can implement associative memory functions. Hinton invented the belief propagation (BP) algorithm for multilayer perceptron (MLP) in 1986. It effectively solves the problem of nonlinear classification and learning [8]. BP neural network is a multi-layer feedforward network algorithm trained by error back-propagation [9]. The emergence of the BP algorithm caused the second upsurge of neural networks. BP neural network has been widely used because of its strong operability, strong adaptability, and good effect of approximating nonlinear systems [10].

Deep belief network (DBN) was first proposed by Hinton [11]. A deep belief network is a structure of multiple restricted Boltzmann machines (RBMs) stacked layer by layer. The network is trained by layer-by-layer greedy algorithm to find the optimal weight value, and the RBMs is equivalent to initializing the optimal weight [12–15].

With the development of computer networks, the algorithm of neural networks has also made great breakthroughs. Artificial neural networks have been widely used in various fields. Scientists use algorithms to simulate the structure of the human brain. Predicting the unknown based on what is known makes this deep learning a huge advantage in real life [16, 17]. This also brings renewed confidence in the application of deep learning in the field of communications. Researchers began to study the application of deep learning in communication systems.

3 Traditional Decoding Algorithm for Polar Codes

In 2008, E. Arikan et al. first proposed the concept of channel polarization [18]. According to the channel polarization theory, a corresponding coding method—polar code is proposed. It is strictly proved from the mathematical formula that the polar code can reach the Shannon limit in a specific channel. Since the birth of the polar code, many scientific research institutions at home and abroad have invested huge R&D efforts for further testing, evaluation and optimization thanks to its huge potential. in order to be better applied to the actual system. In 2016, polar code was selected by the Third Generation Partnership (3GPP) of the International Organization for Mobile Communication

Standardization as the channel coding method for the control channel in the 5G eMBB scenario.

Currently, Arikan proposed the Successive Cancellation (SC) decoding algorithm [19], which is a unique decoding algorithm for polar codes in terms of the decoding performance of polar codes. The SC decoding algorithm has low computational complexity, only $O(N\log N)$, where N is the code length. When the SC algorithm decodes each bit, only the path with the largest transition probability is reserved, and the bit is judged directly. However, as the code length increases, the complexity also increases, which in turn leads to a larger decoding delay. In the SC decoding process, it advances bit by bit. Therefore, the decoding end has high delay and low throughput in the process of hardware implementation. And in practice, the code length cannot be "long enough". The polarization phenomenon is not complete enough, once a wrong bit decision occurs. It is easy to cause errors to propagate. Therefore, for medium and short codes, SC decoding still has certain deficiencies.

In this case, the Successive Cancellation List (SCL) decoding algorithm was proposed [20]. SCL can effectively solve the problem of SC algorithm. The decoding is improved on its basis. When a certain bit is judged, multiple decoding paths are allowed to exist and the next bit is decoded. So instead of making a hard judgment on it right now. SCL decoding is a breadth-first search algorithm. By first expanding and then pruning, the leaf node is finally reached. In the decoding process, multiple decoding paths are allowed to exist, which greatly increases the decoding performance and improves the correctness of the decoding result. But as the number of paths increases, the complexity of the algorithm will also increase [21, 22].

In order to reduce the complexity, a continuous cancellation list decoding scheme is proposed [23, 24]. Cyclic redun-dancy check (CRC) auxiliary polarity codes are used in the SCL decoding scheme. CRC-SCL reduces decoding latency by dividing the entire bit message block into multiple message submodules, all of which are equal in length, and then each message submodule is processed by a CRC encoder. In this way, the waiting time can be reduced and the stored memory can be freed. However, the spatial complexity of the SCL decoder remains the same. Because only the stored memory is freed, the CRC-SCL decoding scheme can effectively reduce the spatial complexity and time complexity.

After Arikan proposed the SC decoding algorithm in the literature, research scholars have proposed some high-performance decoding algorithms based on the SC decoding algorithm, such as SCL, CRC-SCL and other improved decoding algorithms. However, these decoding methods are serial decoding algorithms with high latency and low throughput rate. The BP decoding algorithm is a new parallel decoding method, which performs parallel decoding by multiplexing the received data [25, 26]. The BP decoding has a low latency and low throughput rate. BP decoding is a two-way information transmission decoding compared to SC algorithms. The SC decoding algorithm is a one-way information transmission decoding algorithm. At the same time, the BP decoding algorithm has a high computational complexity. In order to ensure high throughput and improve bit error rate, different structures or different inputs can be executed in parallel [27, 28].

From the above analysis of the current research status of the mainstream decoding algorithms of polar codes, it is clear that the decoding algorithms of polar codes still have sequential decoding leading to large time delays, which cannot meet the requirements of low latency tasks. In view of the characteristics of deep learning algorithm information processing summarized in the first part of this paper. Researchers have introduced the much-anticipated deep learning algorithm into polar codes to find a balance between decoding performance and decoding latency.

4 Neural Network-Based Polar Code Decoding Method

In 2017, Gmber proved that neural networks can learn a form of decoding algorithm and proposed the use of neural networks as polar decoders for the difficult task of polar decoding and called them deep learning decoders [29, 30]. To investigate the problem of deep learning in decoders with long code lengths, Xu et al. [31] proposed to divide long codes into several sub-codes and train the decoder for each sub-code separately. Deep learning applied to communication systems is a qualitative change, and its simplicity and efficiency have accomplished again the value of research in this area.

Then, the paper [32] proposed an improved Convolutional Neural Network (CNN) decoding algorithm for polar codes by modifying the loss function of CNN [33]. The method simulates the decoding characteristics of polar codes by multiplying the value of the loss function of the decoding result during the training process by the corresponding weight determined by the bit error rate performance. This method can reduce the bit error rate without increasing the computational complexity and achieve performance improvement. Among them, CNN are usually applied in computer vision, image, and language processing [34, 35]. Since the signals received in the communication field are one-dimensional vectors. It is necessary to use one-dimensional convolution, which is divided into three types: full convolution, same convolution, and vaild convolution.

CNN perform poorly in processing sequences of before-and-after correlations. So recurrent Neural Network (RNN) was proposed [36] to solve the problem of sequences of anteroposterior correlations. RNN are also known as recurrent neural networks. Ability to send the output value back into itself or into other neurons. This structure can effectively help it obtain information before and after [37]. The codewords of the polarized code are related to each other, and polar code decoding can be achieved by using RNN. The input data of the prior likelihood ratio and the received likelihood ratio as the input data of the RNN is used as 1 in n out recurrent neural network. The introduction of weights in the process of BP decoding and the use of different weights in the left and right iterations can effectively improve the decoding performance [38].

5 Conclusions

This paper introduces the development of deep learning and the traditional decoding methods of polar codes, which are the SC decoding, the SCL decoding, the CRC-SCL decoding, and the BP decoding, and analyzes each decoding method as well as its advantages and disadvantages. Furthermore, we study the deep learning based polar code decoding methods including the CNN and the RNN. The current deep learning based

polar code decoding scheme of polar code still needs improvement. Further research is needed to use more complex neural networks for decoding.

Acknowledgement. This work was supported in part by the National Natural Science Foundation of China under Grant 61801327, in part by the Doctor Fund of Tianjin Normal University under Grant 043135202-XB1711, in part by the Natural Science Foundation of Tianjin City under Grant 18JCYBJC86400, and in part by the Tianjin Higher Education Creative Team Funds Program.

References

1. Arel, I., Rose, D.C., Karnowski, T.P.: Deep machine learning-a new frontier in artificial intelligence research. IEEE Comput. Intell. Mag. **5**(4), 13–18 (2010)
2. Bengio, Y.: Learning deep architectures for AI. Found. Trends® Mach. Learn. **2**(1), 1–127 (2009)
3. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
4. Sanyal, H., Agrawal, R.: Latest trends in machine learning & deep learning techniques and their applications. Int. Res. Anal. J **14**(1), 348–353 (2018)
5. Landahl, H.D., McCulloch, W.S., Pitts, W.: A statistical consequence of the logical calculus of nervous nets. Bull. Math. Biophys. **5**(4), 135–137 (1943)
6. Minsky, M.L., Papert, S.A.: Perceptrons: expanded edition (1988)
7. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. **79**(8), 2554–2558 (1982)
8. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Nature **323**(6088), 533–536 (1986)
9. Zhuang, Y., Lu, R.: Multi-parameter grain analysis model based on BP neural network. In: 2018 International Conference on Transportation & Logistics, Information & Communication, Smart City (TLICSC 2018), pp. 387–392. Atlantis Press (2018)
10. Achille, A., Soatto, S.: Information dropout: learning optimal representations through noisy computation. IEEE Trans. Pattern Anal. Mach. Intell. **40**(12), 2897–2905 (2018)
11. Hinton, G.E., Osindero, S., Teh, Y.W.: A fast learning algorithm for deep belief nets. Neural Comput. **18**(7), 1527–1554 (2006)
12. Fischer, A., Igel, C.: An introduction to restricted Boltzmann machines. In: Iberoamerican congress on pattern recognition, pp. 14–36. Springer, Heidelberg (2012)
13. Carreira-Perpinan, M.A., Hinton, G.: On contrastive divergence learning. In: International Workshop on Artificial Intelligence and Statistics, pp. 33–40. PMLR (2005)
14. Hinton, G.E.: Training products of experts by minimizing contrastive divergence. Neural Comput. **14**(8), 1771–1800 (2002)
15. Mnih, V., Larochelle, H., Hinton, G.E.: Conditional restricted Boltzmann machines for structured output prediction. arXiv preprint arXiv:1202.3748 (2012)
16. Yu, D., Deng, L.: Deep learning and its applications to signal and information processing [exploratory dsp]. IEEE Signal Process. Mag. **28**(1), 145–154 (2010)
17. Deng, L., Li, X.: Machine learning paradigms for speech recognition: an overview. IEEE Trans. Audio Speech Lang. Process. **21**(5), 1060–1089 (2013)
18. Arikan, E.: Channel polarization: a method for constructing capacity achieving codes. In: 2008 IEEE International Symposium on Information Theory. Toronto Canada, 1173–1177 (2008)
19. Arikan, E.: Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. IEEE Trans. Inf. Theory **55**(7), 3051–3073 (2009)

20. Cao, C., Fei, Z., Yuan, J., Kuang, J.: Low complexity list successive cancellation decoding of polar codes. IET Commun. **8**(17), 3145–3149 (2014)
21. Tal, I., Vardy, A.: List decoding of polar codes. IEEE Trans. Inf. Theory **61**(5), 2213–2226 (2015)
22. Chen, K., Niu, K., Lin, J.R.: List successive cancellation decoding of polar codes. Electron. Lett. **48**(9), 500–501 (2012)
23. Chiu, M.C., Wu, W.D.: Reduced-complexity SCL decoding of multi-CRC-aided polar codes. arXiv preprint arXiv:1609.08813 (2016)
24. Li, B., Shen, H., Tse, D.: An adaptive successive cancellation list decoder for polar codes with cyclic redundancy check. IEEE Commun. Lett. **16**(12), 2044–2047 (2012)
25. Yuan, B., Parhi, K.K.: Architecture optimizations for BP polar decoders. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 2654–2658. IEEE (2013)
26. Zhang, Y., Liu, A., Pan, X., Ye, Z., Gong, C.: A modified belief propagation polar decoder. IEEE Commun. Lett. **18**(7), 1091–1094 (2014)
27. Arlı, A.Ç., Gazi, O.: Noise-aided belief propagation list decoding of polar codes. IEEE Commun. Lett. **23**(8), 1285–1288 (2019)
28. Elkelesh, A., Ebada, M., Cammerer, S., Ten Brink, S.: Belief propagation list decoding of polar codes. IEEE Commun. Lett. **22**(8), 1536–1539 (2018)
29. O'Shea, T.J., Karra, K., Clancy, T.C.: Learning to communicate: Channel auto-encoders, domain specific regularizers, and attention. In: 2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 223–228. IEEE (2016)
30. Gruber, T., Cammerer, S., Hoydis, J., & ten Brink, S.: On deep learning-based channel decoding. In: 2017 51st Annual Conference on Information Sciences and Systems (CISS), pp. 1–6. IEEE (2017)
31. Xu, W., Wu, Z., Ueng, Y. L., You, X., Zhang, C.: Improved polar decoder based on deep learning. In 2017 IEEE International workshop on signal processing systems (SiPS), pp. 1–6. IEEE (2017)
32. 汤佳杰. Polar 码译码算法研究 (2020)
33. Rao, W., Liu, Z., Huang, L., Sun, J., Dai, L.: CNN-SC decoder for polar codes under correlated noise channels. In: 2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT), pp. 748–751. IEEE (2020)
34. Puhalanthi, N., Lin, D.-T.: Effective multiple person recognition in random video sequences using a convolutional neural network. Multimed. Tools Appl. **79**(15–16), 11125–11141 (2019). https://doi.org/10.1007/s11042-019-7323-z
35. Liu, R., Wei, S., Zhao, Y., Yang, Y.: Indexing of the CNN features for the large scale image search. Multimed. Tools Appl. **77**(24), 32107–32131 (2018). https://doi.org/10.1007/s11042-018-6210-3
36. Liao, Z., Carneiro, G.: A deep convolutional neural network module that promotes competition of multiple-size filters. Pattern Recogn. **71**, 94–105 (2017)
37. Ye, B., Feng, G., Cui, A., Li, M.: Learning question similarity with recurrent neural networks. In: 2017 IEEE International Conference on Big Knowledge (ICBK), pp. 111–118. IEEE (2017)
38. Neculoiu, P., Versteegh, M., Rotaru, M.: Learning text similarity with siamese recurrent networks. In: Proceedings of the 1st Workshop on Representation Learning for NLP, pp. 148–157 (2016)

Shadow Removal Network with Error Map Prediction

Haiyang Liu and Yongping Xie(✉)

Dalian University of Technology, Dalian 116081, China
xieyp@dlut.edu.cn

Abstract. The existence of shadows affects many computer vision tasks. Recovering image information in shadowed regions by removing shadows has been proven to improve the performance of various tasks. However, shadow removal is a challenging task. The complex presence pattern makes it difficult for us to com succinct pletely shadows. In this paper, we study the problem of shadow removal, which aims to obtain traceless shadow removal results and preserve the image information under the original shadow area. To more adapt to the shadow removal task in complex backgrounds flexibly, we propose an Error Map mechanism to pre-estimate the feature error generated by the fusion network and use it to guide the refinement work to obtain better shadow removal results. In addition, we redesigned the higher performance network structure. We evaluate the performance of our method on multiple large-scale shadow removal datasets, our method achieves better performance than the original scheme.

Keywords: Shadow removal · Exposure fusion · Error map

1 Introduction

Shadows are dark areas in natural scenes where light is completely or partially blocked, which is very common in natural images and may bring challenges to existing computer vision tasks [9–13], such as object detection, object tracking, object recognition and semantic segmentation, etc. Recovering image information in shadowed regions by removing shadows has been a long-existing research problem [14–17]. With the advent of advanced deep networks and large-scale annotated datasets, data-driven approaches using deep learning models have achieved remarkable performance in shadow removal [18–23].

To address the challenges of the shadow removal task, in [22], the authors take into account the often overlooked physical properties of shadows. Shadow casting degrades image quality as color and lighting degrade, and overexposing shadow images is an effective way to improve image quality. Using shadow lighting modeling and deep learning, the authors use a simplified physical lighting model to define the mapping between shadow and non-shadow pixels. Furthermore, Fu [1] et al. used a more flexible auto-exposure fusion network for shadow removal from a single image. Multiple overexposed images are generated by compensating for shadow areas at different exposure

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 158–166, 2023.
https://doi.org/10.1007/978-981-99-1260-5_20

levels, and through a shadow-aware fusion network, the multiple overexposed images and the original shadow image are "intelligently" fused to remove shadows.

However, in the event of multiple shadows in complex backgrounds, the auto-exposure fusion network does not propose a reasonable refinement scheme. The method results in a large number of differences within the shadow areas, and the complex fusion calculation also causes severe distortion in the non-shadow areas, as shown in Fig. 1. To more adapt to the shadow removal task in complex backgrounds flexibly, we propose an Error Map mechanism to pre-estimate the feature error generated by the fusion network and use it to guide the refinement work to obtain better refinement results. In addition, we redesigned the higher-performance network structure. We evaluate the performance of our method on multiple large-scale shadow removal datasets, and our method achieves state-of-the-art performance compared to the original scheme.

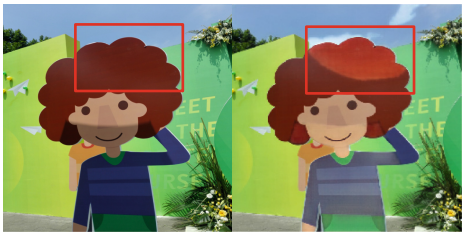


Fig. 1. The original scheme causes obvious distortion in the non-shadowed areas and obvious traces of the shadow's edge. Input shadow image (left) Auto-Exposure Fusion output(right).

2 Related Work

Exposure Fusion. The recent shadow removal work is mainly based on the physical shadow model, through the method of exposure fusion to remove shadows. Inspired by earlier work, Le et al. [22] propose a framework that outputs physics-inspired shadow lighting models and shadow matte images. The algorithm aims to compute a fused weight map for each image and fuse the inputs via a weighted sum operation sequence. Fu et al. [1] employ a multi-exposure fusion approach to remove shadows. The method generates multiple over-exposed shadow images to compensate for color and lighting degradation in shadow areas and "smartly" fuses them to obtain shadow-free images. This paper builds on the multiple exposure algorithm and redesigns a high-performance network that is more suitable for the algorithm principle.

Error Map. Error Map is often used in failure detection to represent the gap between the output and the true image. Xia et al. [5] used the prediction head of the predicted Error Map to detect anomalous segmentation failures in semantic segmentation scenarios. They used the feature distance metric to generate a binary pixel-level Error Map to represent segmentation anomalies. Lin et al. [4] applied Error Map in the field of video matting to provide error location information to the refinement network. In this paper, we design a novel Error Map mechanism to pre-estimate the feature error of the fusion

network, perform pixel-level classification of shadow images, and provide refinement guidance to the refinement network.

3 Methods

Auto-exposure fusion uses paired training data for the shadow removal task. This method models the shadow removal problem with a new approach, namely the problem of auto-exposure fusion of paired shadow and shadow-free images. However, in actual tests, there are often problems such as obvious shadow boundaries and partial distortion of non-shadow areas.

We believe that the aforementioned problems are caused by the multi-exposure fusion algorithm introducing larger errors and lacking a reasonable refinement scheme. Therefore, we solve this problem from two aspects.

First, we design a pixel-level Error Map to represent the error degree of the fusion network and use it to guide the refinement work to improve the system robustness. At the same time, we redesigned the network structure of the multi-exposure fusion algorithm for better performance and reduced feature errors. Our overall network structure is shown in Fig. 2.

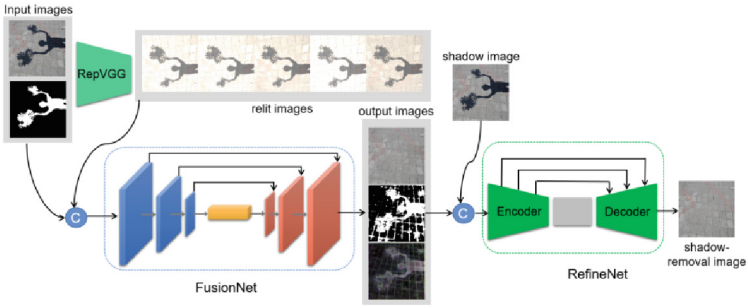


Fig. 2. Overall network structure

3.1 Error Map

Since the shadow is a dark area that changes continuously, the interior and the edge transition zone of the shadow will show different lighting characteristics. Additionally, the interior of the shadow will have internal inconsistencies due to the properties and reflectivity of the object. Obtaining seamless shadow-free images from complex backgrounds is still a major challenge.

In the fusion network, due to the complexity of computing the fusion kernel of multiple exposure images, color distortion or artifacts are generated in non-shadow areas. Therefore, methods to maintain the consistency of the lighting and color of the shadow area and the non-shadow area while preserving the original information of the image are needed.

We design an alternative output for the fusion network and calculate a pixel-level Error Map to represent the feature distance of the fusion network output. Our framework can perform pixel-level classification of shadow images under different backgrounds and use the information to guide subsequent refinement work to obtain higher-quality shadow-free images and improve system robustness.

We use a simple yet effective feature distance metric to compute the Error Map, whose supervised value E^* is expressed as the difference between the ground-truth and the predicted output of the fusion network: $E^* = I^* - I$. To more adapt to shadow refinement operations in complex backgrounds flexibly, we compute the RGB version and Mask version of the image at the same time, which are used to provide color, luminance distortion information, and pixel position information to the refinement network. According to the directly predicted Error Map, the Mask is defined as:

$$E_m = \begin{cases} 1 & E_{RGB} > error \\ 0 & |E_{RGB}| <= error \\ -1 & E_{RGB} < -error \end{cases} \tag{1}$$

In this formula, E_{RGB} is the average of the Error Map directly output by the fusion network in the channel dimension, and the threshold error is 0.02.

We use L_2 distance to optimize E and supervision values E^*:

$$L_{pix}(E, E^*) = \|E^* - E\|_2 \tag{2}$$

3.2 Network Design

ParamNet. Our goal is to train a DNN to adaptively estimate exposure parameters. To obtain better performance indicators, we chose Rep-VGG Plus [2], which has performed well in recent years, as our parameter prediction network, and made appropriate modifications to suit our algorithm. Since Rep-VGG Plus adopts the idea of parameter reconstruction, training and testing Rep-VGG Plus use different equivalent models and we can obtain excellent exposure image parameters at a small time cost.

FusionNet. Our purpose is to design an efficient fusion network to accurately estimate the fusion kernel, allowing the network to assign lower weights to non-shadow areas and higher weights to shadow areas. Our fusion network is a fully convolutional encoder-decoder network that adopts the U-Net [25] structure to extract effective features and perform feature fusion. Due to the image inpainting work has pixel sensitivity and the behavior of fusion networks has segmentation properties. We referred to the DeepLabV3 [3] and DeepLabV3 + [24] architectures in the semantic segmentation task, where the Atrous Spatial Pyramid Pooling (ASPP) module plays a key role. Our network structure uses ResNet-50 as the backbone, adding the ASPP module following DeepLabV3 to enlarge the feature receptive field of the model.

RefineNet. By adjusting the RGB version of the Error Map, we mapped values into the (–1, 1) interval. We feed the mapped RGB version and the Mask version of the Error Map into the refinement network for pixel-level refinement. At the same time, we add

CA + SA attention mechanism to the down-sampling module of U-Net [25]. We replace the channel attention module in CBAM [26] with the CA [27] module, which, in tandem with the spatial attention module, constitutes a new attention mechanism to improve the performance of our model.

4 Experiments

4.1 Network Architecture and Implementation Details

ParamNet adopts the RepVGGplus-L2pse model and reads the pre-trained model under ImageNet. We add a fully connected layer to modify the output class to 6 and add a convolutional layer to the input to change the number of channels to 3. FusionNet adopts an encoder-decoder architecture, including ResNet-50, ASPP, and a bilinear upsampling decoder. The ASPP module follows the definition in the DeepLabV3 model and reads the pre-trained model under Pascal VOC. RefineNet adopts the U-Net-128 architecture and adds a CA + SA attention module after the 3rd and 4th downsampling layers. Upsampling adopts bilinear upsampling instead of deconvolution operation to eliminate the checkerboard effect.

The input image size for our experiments is 256 × 256, the parameter prediction interpolation step size is set to 0.05, and the error threshold is set to 0.02. We use the Adam iterator to train our model. The base learning rate is set to 0.0001 in the first 100 epochs and afterwards we apply a linear decay strategy to reduce it to 0. We used the GTX2080 (8 G memory) platform and trained 400 epochs under the Pytorch training framework.

4.2 Experimental Evaluation

Based on real scene fitting considerations, we first evaluate our method on the test set of the SRD [18] dataset. To better evaluate the effectiveness of our method, according to the up-to-date shadow removal evaluation scheme, we uniformly calculate the RMSE, PSNR, and SSIM of shadow areas, non-shadow areas, and the whole image. All shadow removal results are rescaled to 256 × 256 for comparison with the ground-truth image at that size.

Table 1. Quantitative results with the original method on the SRD dataset

Method	Shadow			Non-Shadow			All		
	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑
Input	35.195	19.305	0.845	6.753	27.755	0.889	18.200	17.887	0.795
Auto-Exposure Fusion	8.560	33.446	0.925	5.222	31.357	0.867	5.976	28.663	0.864
Ours	**7.900**	**34.406**	**0.933**	**4.222**	**32.894**	**0.894**	**5.117**	**29.805**	**0.885**

Table 2. Quantitative results with the original method on the ISTD + dataset

Method	Shadow			Non-Shadow			All		
	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑
Input	32.044	20.774	0.931	3.795	33.843	**0.894**	13.805	20.184	0.878
Auto-Exposure Fusion	6.938	36.835	0.972	3.722	**34.742**	0.893	4.349	32.036	0.931
Ours	**6.258**	**37.300**	**0.972**	**3.670**	34.683	0.890	**4.180**	**32.154**	**0.931**

Table 3. Quantitative results with the original method on the ISTD dataset

Method	Shadow		
	RMSE↓	PSNR↑	SSIM↑
Input	7.833	20.774	0.931
Auto-Exposure Fusion	7.833	35.464	0.971
Ours	**7.139**	**36.306**	**0.971**

As shown in Table 1, we compare our model with the original Auto-Exposure Fusion method. We generated results with the Auto-Exposure Fusion method using their official code [link?]. The first row shows the RMSE, PSNR, and SSIM values of the input image. The results show that our method achieves the best shadow removal in both shadowed and non-shadowed regions.

To verify the generalizability of our method, we also report the shadow removal performance of our proposed method on the adjusted ISTD (ISTD+) [19] dataset as well as the original ISTD [22] dataset. Likewise, our method achieves the best results in shaded areas. The shadow removal results output by the Auto-Exposure Fusion method overfits the ground-truth image in non-shadow areas, resulting in large-area distortion. Our method effectively avoids this issue and better preserves the original non-shadow area image information.

Figure 3 shows the comparison between the shadow removal results of our method and the original scheme. It can be clearly seen that the original method has serious color and brightness distortion and obvious shadow boundaries.

Ablation Experiment. We perform an ablation study on the ISTD + dataset to evaluate the contribution of each step of our proposed method. As shown in Table 4, we provide RefineNet with different Error Map versions. From the quantitative analysis results, we can see that our final proposed method can effectively improve shadow areas.

Fig. 3. Comparison of results between the original method, our method and ground truth on the SRD and ISTD + dataset

Table 4. Ablation experiment results on the ISTD + dataset

Method\RMSE	Shadow	Non-Shadow	All
Input	32.044	3.795	13.805
Auto-Exposure Fusion	6.938	3.722	4.349
FusionNet	6.626	**3.597**	4.230
RefineNet-E_{m0-1}	6.504	3.634	4.212
RefineNet-E_{RGB}	6.454	3.624	4.188
RefineNet-E_{m0-1} + E_{RGB}	6.427	3.649	4.195
RefineNet-E_m + E_{RGB}	**6.258**	3.670	**4.180**

5 Conclusion

In this paper, we construct a new and robust image shadow removal model and propose a new Error Map mechanism to obtain traceless shadow-free images. Our work continues to build on the multi-exposure shadow removal idea of Fu et al. [1] and redesigns a high-performance shadow removal network that is more suitable for the task. To improve the robustness of shadow removal, we design a pixel-level Error Map to represent the error degree of the fusion network, convert image classification into a pixel-level classification for shadows in different backgrounds, predict the pixel-level Error Map, and use it to guide the refinement work. We quantitatively and qualitatively evaluate the effectiveness of our method on the SRD, ISTD, and ISTD + datasets. The results show that our method improves the results of the shadow area output by the FusionNet network, and at the

same time makes the transition between the non-shadow area and the shadow area smoother, so that the overall image visual effect is better. In the future, we plan to fuse the shadow detection task to facilitate our shadow removal task and consider solving more challenging video shadow removal tasks.

References

1. Fu, L., et al.: Auto-exposure fusion for single-image shadow removal. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2021)
2. Ding, X., Zhang, X., Ma, N., Han, J., Ding, G., Sun, J.: RepVGG: making VGG-style ConvNets great again. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2021)
3. Chen, L.-C., Papandreou, G., Schroff, F., Adam, H.: Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587 (2017)
4. Lin, S., Ryabtsev, A., Sengupta, S., Curless, B., Seitz, S., Kemelmacher-Shlizerman, I.: Real-time high-resolution background matting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2021)
5. Xia, Y., Yi, Z., Liu, F., Shen, W., Yuille, A.L.: Synthesize then compare: detecting failures and anomalies for semantic segmentation. In: Proceedings of the European Conference on Computer Vision (2020)
6. Finlayson, G., Drew, M., Lu, C.: Entropy minimization for shadow removal. International Journal of Computer Vision (2009)
7. Yago Vicente, T.F., Hoai, M., Samaras, D.: Leave-one-out kernel optimization for shadow detection and removal. IEEE Trans. Pattern Anal. Mach. Intell. **40**(3), 682–695 (2018)
8. Yago Vicente, T.F., Hou, L., Yu, C.-P., Hoai, M., Samaras, D.: Large-scale training of shadow detectors with noisily-annotated shadow examples. In: Proceedings of the European Conference on Computer Vision (2016)
9. Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detecting moving objects, ghosts, and shadows in video streams. IEEE Trans. Pattern Anal. Mach. Intell. **25**(10), 1337–1342 (2003)
10. Jung, C.R.: Efficient background subtraction and shadow removal for monochromatic video sequences. IEEE Trans. Multimed. **11**(3), 571–577 (2009)
11. Nadimi, S., Bhanu, B.: Physical models for moving shadow and object detection in video. IEEE Trans. Pattern Anal. Mach. Intell. **26**(8), 1079–1087 (2004)
12. Sanin, A., Sanderson, C., Lovell, B.C.: Improved shadow removal for robust person tracking in surveillance scenarios. In: International Conference on Pattern Recognition, pp. 141–144. IEEE (2010)
13. Zhang, W., Zhao, X., Morvan, J.-M., Chen, L.: Improving shadow suppression for illumination robust face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **41**(3), 611–624 (2018)
14. Le, H., Nguyen, V., Yu, C.-P., Samaras, D.: Geodesic distance histogram feature for video segmentation. In: ACCV, pp. 275–290. Springer (2016)
15. Nan, S., Zhang, Y., Tian, S., Yan, Y., Miao, X.: Shadow detection and removal for occluded object information recovery in urban high-resolution panchromatic satellite images. IEEE J. Sel. Top. Appl. Earth Observations Remote Sens. **9**(6), 2568–2582 (2016)
16. Le, H., Yu, C.-P., Zelinsky, G., Samaras, D.: Co-localization with category-consistent features and geodesic distance propagation. In: International Conference Computer Vision Workshop, pp. 1103–1112 (2017)
17. Müller, T., Erdnüeß, B.: Brightness correc-tion and shadow removal for video change detection with uavs. In: Autonomous Systems: Sensors, Processing, and Security for V ehicles and Infrastructure 2019, vol. 1009, p. 1100906. International Society for Optics and Photonics (2019)

18. Qu, L., Tian, J., He, S., Tang, Y., Lau, R.W.H.: Deshadownet: a multi-context embedding deep network for shadow removal. In: IEEE Conference on Computer Vision and Pattern Recognition (2017)
19. Wang, J., Li, X., Yang, J.: Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In: IEEE Conference on Computer Vision and Pattern Recognition (2018)
20. Hu, X., Fu, C.-W., Zhu, L., Qin, J., Heng, P.-A.: Direction-aware spatial context features for shadow detection and removal. IEEE Trans. Pattern Anal. Mach. Intell. (2019)
21. Hu, X., Jiang, Y., Fu, C.-W., Heng, P.-A.: Mask-shadowgan: Learning to remove shadows from unpaired data. In: International Conference Computer Vision (2019)
22. Le, H., Samaras, D.: Shadow removal via shadow image decomposition. In: International Conference Computer Vision (2019)
23. Le, H., Samaras, D.: From shadow segmentation to shadow removal. In: Proceedings of the European Conference on Computer Vision (2020)
24. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision (2018)
25. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional Networks for Biomedical Image Segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015)
26. Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: CBAM: convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (2018)
27. Hou, Q., Zhou, D., Feng, J.: Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2021)

FFP-MVSNet: Feature Fusion Based Patchmatch for Multi-view Stereo

Xing Luo and Yongping Xie(✉)

Dalian University of Technology, Dalian 116081, China
xieyp@dlut.edu.cn

Abstract. With the development of technology, 3D reconstruction has been widely used in many fields. In this paper, we propose a learnable 3D reconstruction method using a cascaded Patchmatch approach to form a new network. By introducing a dual-channel attention module, point clouds reconstruction has been improved in accuracy and completion This network has high computational speed and low memory requirements, which allows it to handle higher-resolution images. The network is more suitable for running on resource-constrained devices than competitors that employ 3D cost volume regularization. We introduce the feature fusion module to an end-to-end trainable framework for the first time. The weight parameters of the multi-scale network output can be adaptively learned in each calculation, which can reduce the feature dispersion caused by the multi-scale output. This method has good performance on DTU.

Keywords: 3D reconstruction · Feature fusion · Attention mechanism

1 Introduction

Multi-view stereo (MVS) is the process of turning multiple images from different views of an object into a three-dimensional scene. Although MVS has been a hot topic in the field of computer vision research, it is still a challenge for people at present. This is due to the fact that there are still many unsolved problems in practice, such as poor reconstruction in occlusion, illumination changes, and untextured regions [1].

Convolutional neural network (CNN) has achieved good results in computer vision, so people consider using CNN to solve the problem of 3D reconstruction. In fact, many learning-based methods [2] almost surpass all traditional methods [3] on the MVS benchmark. Currently, most learn-based MVS methods construct a 3D cost volume, regularize it with 3D CNN, and finally get the depth map by regression. Since the regularization process of cost volume takes a long time and occupies a large amount of memory, some methods [4, 5] consider calculating cost volume and depth map simultaneously at low resolution, which affects the accuracy of reconstruction. Some traditional MVS methods give up the idea of building cost volume and adopt Patchmatch [6] algorithm to achieve iterative propagation of depth map, which is to propagate the correct parallax plane in all original parallax planes to other pixels in the same parallax plane. This approach has lower losses and provides new ideas for deep learning-based MVS.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 167–174, 2023.
https://doi.org/10.1007/978-981-99-1260-5_21

In this paper, we add an attention mechanism and a feature fusion module to the traditional Patchmatch to improve accuracy. Furthermore, we propose a robust training strategy that makes the selection of views random, to improve the robustness of the network model. We verified the validity of our approach on DTU datasets, and the results show that our network achieves competitive performance compared to most learning-based approaches.

2 Related Work

Traditional MVS. Traditional MVS methods can be divided into four categories: point clouds reconstruction [7], voxel reconstruction [8], depth map reconstruction [9] and Patch-based reconstruction. Point clouds are easier to manipulate when geometrically transforming and deforming. However, points in the point cloud lack connectivity and will lack information on the surface of objects. The most intuitive feeling is that the reconstructed surface is uneven. The disadvantage of voxel reconstruction is that it is resource intensive. Once you want to improve the density of the final point cloud, you can only improve the resolution of the two-dimensional image, which leads to the calculation amount in the final calculation by the order of three. In contrast, depth maps are the most flexible representation. It decouples the complex MVS problem into a relatively small per-view depth map estimation problem. Among these 3D reconstruction methods, depth map is more concise and flexible, which can be regarded as a special representation of point cloud. In recent studies, the representation based on depth map has achieved good results.

Learning-Based Stereo. Many works in recent years [3, 4] are based on planar scanning stereo [11], using depth maps to reconstruct scenes. CasMVSNet [12] proposes a cascaded MVS network to construct a cost volume from coarse precision to fine precision. Compared with the method of average construction of costal body at the beginning, these two methods greatly reduce the memory limitation caused by image resolution or the large size of costal body. UCS-Net [13] proposed an adaptive structure, which gradually refines the depth interval and range of each source view through its encoded geometric information and context information CVP-MVSNet [14] proposed a new multi-level pyramid structure to complete the iteration and optimization of the cost volume. Fast-MVSNet [15] proposes a new sparse denser framework, and uses the new Gauss Newton layer to refine the depth map.

3 Method

We propose a method for multi-view 3D reconstruction using self-supervised methods. It consists of four parts: a feature extraction network, a feature fusion module, patch matching and a refinement module [16]. Experimental results show that our method has faster processing speed. Even though this method performs better than the traditional method on the MVS benchmark, there is still a gap in performance for some networks based on the 3D cost volume.

For the aforementioned problems, we infer that although Patch-based matching can greatly improve the efficiency, a large number of points will be matched incorrectly once a matching error occurs. Thereby it can affect the accuracy and the overall reconstruction quality. Therefore, we will solve this problem from two aspects: 1) We introduce a feature fusion mechanism to reduce the problem of inaccurate matching of subsequent patches caused by the loss of information between different layers during multi-scale output of the feature pyramid network. 2) We introduce a dual-attention mechanism to guide the model to filter irrelevant features, enhance effective features, and perform more effective feature extraction.

The model structure we designed is shown in Fig. 1. We introduce a dual-attention mechanism at the input. Feature enhancement is performed by pre-coding the spatial information and channel information of the source image. A feature fusion network is added to the multi-scale output of the feature pyramid, which adaptively fuses the outputs of different scales before performing subsequent patch matching, we named the network FFP-MVSNet (Feature fusion based Patchmatch for multi-view stereo).

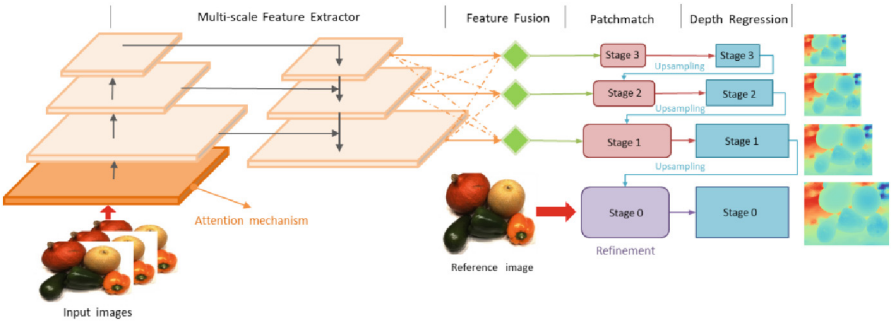


Fig. 1. Structure of FFP-MVSNet

Feature Fusion Module. The Feature Pyramid Network forms multi-level feature maps through continuous downsampling of convolution kernel. One of the main disadvantages of the feature pyramid is that the features are different at different scales. Specifically, large goals are usually associated with upper-level features, while small goals are associated with lower-level features. Therefore, if an image contains both large and small objects, the inconsistencies between features at different levels will interfere with the training, thus affecting the training results and failing to fully play the role of the feature pyramid network.

To address the above issues, we propose an adaptive feature fusion module, which enables the network to automatically filter features at other levels through training, and only retain consistent features for matching. This method includes two steps: feature rescaling and feature fusion.

Feature Rescaling. For the features of a certain layer, we adjust the features of other layers to the corresponding resolution. We denote the resolution features of the l^{th} stage ($l \in \{1, 2, 3\}$) as x^l. For stage l, we resize the feature map of another stage n ($n \neq l$) to

the size of feature x^n to make it the same resolution as x^l. Since the feature extraction network we adopt has three stages of output, the feature maps of each stage have different resolutions and different channels. Therefore, we need to formulate corresponding up-sampling and down-sampling methods for each scale. For up-sampling, we first use 1 $\times$ 1 convolution layer to adjust the channel number of features to the channel number of layer l, and then use bilinear interpolation to improve the resolution. For 1/2 scale downsampling, we use a 3 $\times$ 3 convolution layer with stride 2 to modify the number of channels and resolution. For 1/4 scale, we first add a maximum pooling layer, scale to 1/2, and then use a 3 $\times$ 3 convolution layer with stride 2 to scale 1/4. Through the above method, the output of all levels can have the same resolution, creating conditions for the subsequent fusion.

Feature Fusion. Based on the above feature maps of the same resolution obtained at each stage, different levels of adaptive fusion are carried out for each pixel position in the maps. That is, if certain features present inconsistent information in the same location, they will be filtered out in the fusion, while others Traits are dominated by more discriminating cues.

Let $x_{ij}^{n\to k}$ represent the feature vector at position (i, j) in the feature graph adjusted from level n to level k. When k is fixed and n is adjusted, the fusion strategy proposed is shown in the following formula:

$$y_{ij}^k = a_{ij}^k \cdot x_{ij}^{1\to k} + b_{ij}^k \cdot x_{ij}^{2\to k} + c_{ij}^k \cdot x_{ij}^{3\to k} \tag{1}$$

where y_{ij}^k represents the feature vector of the output feature map y^k at the position (i, j) on the k^{th} layer. a_{ij}^k, b_{ij}^k and c_{ij}^k are the weights obtained by adaptive learning from the feature graph of the three levels to layer k. a_{ij}^k, b_{ij}^k and c_{ij}^k as weights are simple numeric variables that can be shared across all channels by each level of the feature graph. Inspired by [17], we let $a_{ij}^k + b_{ij}^k + c_{ij}^k = 1$ and $a_{ij}^k, b_{ij}^k, c_{ij}^k \in [0, 1]$, and define:

$$a_{ij}^k = \frac{e^{\lambda_{a_{ij}}^k}}{e^{\lambda_{a_{ij}}^k} + e^{\lambda_{b_{ij}}^k} + e^{\lambda_{c_{ij}}^k}} \tag{2}$$

Here a_{ij}^k, b_{ij}^k and c_{ij}^k refer to three different levels of output. And $\lambda_{a_{ij}}^k$, $\lambda_{b_{ij}}^k$ and $\lambda_{c_{ij}}^k$ are expressions of the three exponents of the denominator of softmax function respectively. We use 1 $\times$ 1 convolution layers to change the output of the three layers into the same number of channels, which is convenient for the subsequent weight calculation.

Attention Mechanism. In recent years, attention mechanisms have been shown to be useful in various computer vision tasks. We propose a new mechanism for attention. We encode the location information of the objects in the image into the channel, so that richer information can be extracted. Our attention mechanism can be divided into two parts: channel attention and spatial attention. In this way, the spatial information of the image can be introduced into the attention map and the long-term dependence on features can be established through channel attention. Specifically, the spatial attention mechanism uses

two one-dimensional global pooling operations to compress input features vertically and horizontally, respectively. After the convolution layer, the feature graphs of vertical and horizontal directions respectively generate corresponding attention graphs, and each attention graph retains the feature content of specific spatial direction. Thus, channel attention incorporates the location information of the image. Then, the two attention maps are applied to the input feature map by multiplication. The obtained results are then passed through the channel attention to obtain a feature map with more significant feature information, which helps the model to enhance the effective features and improve the feature extraction ability of the model. The following experimental results will show that our proposed new attention mechanism is effective.

Loss function L_{tatal} considers the loss in the depth estimation of all stages. The formula is as follows:

$$L_{total} = \sum_{k=1}^{3} \sum_{i=1}^{n_k} L_i^k + L_{ref} \tag{3}$$

The formula is divided into two parts. The first part is the loss of each stage of patchmatch, while the last part is the loss of depth map output of the refined network.

4 Experiments

Datasets. In this experiment, we evaluated our work on the DTU dataset and analyzed the role of each new component through ablation experiments. This dataset has 124 different scenes, each with a total of 49 viewing angles and seven different brightness levels. DTU data sets are indoor data sets that are specially photographed and processed for MVS. According to the ratio in [18], the data set is divided into training, testing and verification set.

Implementation Details. The framework we used is Pytorch, and the training data is the training set divided by the DTU dataset. In the training, the number of input images N is five, and the resolution is 640 × 512. We set the number of iterations of Patchmatch in three stages as 2, 2, 1 respectively. For initialization, we set the number of assumed planes to 48. In the propagation stage, we set the value of pixels in the propagation neighborhood as 16, 8, 8. All parameter values can be adjusted according to experimental needs. We trained on a Nvidia GTX 3090 GPU with batchsize set to 4 and epoch set to 12. After getting the corresponding depth map after the training, we will fill the point cloud of the depth map containing RGB-D to get the final point cloud.

Ablation Study. In this section, we show how different components affect the results of the experiment. When evaluating the results, the resolution of each input image is 1600 × 1200, and the sampling depth assumption range is fixed as [425 mm, 935 mm].

As shown in Table 1, according to the scores obtained from the experiment, our method is significantly superior to other methods in completeness and overall quality. The result of point cloud reconstruction is shown in Fig. 2. It can be seen that the point cloud results reconstructed by our solution are more complete and more detailed. Table 2

shows how the attention module and feature extraction network are added to improve the model evaluation metrics, respectively. Figure 3 shows the influence of each component in the network on the point cloud reconstruction results. In addition, we also evaluated the number of input views N to evaluate the performance of the model when N = 2, 3, 5, 6. As shown in Table 3, using more views can improve the performance, The performance of the model reaches the best when N = 5.

Table 1. Quantitative results of different methods (lower is better).

Method	Acc.	Comp.	Overall.
Gipuma	**0.283**	0.873	0.578
MVSNet	0.396	0.527	0.462
R-MVSNet	0.383	0.452	0.427
Fast-MVSNet	0.336	0.403	0.370
Point-MVSNet	0.342	0.411	0.376
CasMVSNet	0.325	0.385	0.355
Ours	0.427	**0.257**	**0.342**

Table 2. The impact of each component on the DTU evaluation set

Method	Acc.	Comp.	Overall.
Base	0.450	0.290	0.370
Attention	0.435	0.274	0.355
Feature Fusion	0.436	0.265	0.351
Both	0.427	0.257	0.342

Table 3. Ablation study of the number of input views N

N	Acc.	Comp.	Overall.
2	0.439	0.312	0.375
3	0.429	0.265	0.352
5	0.427	0.257	0.342
6	0.429	0.258	0.344

Fig. 2. Qualitative comparison of point clouds reconstruction results

Fig. 3. The impact of each component on the result of reconstruction

5 Conclusion

We propose a novel learning-based FFP-MVSNet for 3D scene reconstruction, which augments a deep feature-based attention module and a feature fusion module. FFP-MVSNet inherits the advantages of Patchmatch, which naturally has lower memory requirements. A large number of experimental results on the DTU evaluation set show that FFP-MVSNet has good generalization performance and competitiveness compared with many advanced methods. Although FFP-MVSNet makes MVS based on learning more efficient and provides a new idea for people to carry out 3D reconstruction, there are still some shortcomings, such as low accuracy, which need to be solved one by one in the future.

References

1. Arno, K., Jaesik, P., Qianyi, Z., Vladlen, K.: Tanks and temples: Benchmarking large-scale scene reconstruction. ACM Trans. Graph. (2017)
2. Thoms, S., Johannes, L.S., Silvano, G., Torsten, S., Konrad, S., Marc, P., Andreas, G.: A multi-view stereo benchmark with highresolution images and multi-camera videos. In Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
3. Rui, C., Songfang, H., Jing, X., Hao, S.: Point-based multi-view stereo network. In International Conference on Computer Vision (ICCV) (2019)

4. Keyang, L., Tao, G., Lili, J., Haipeng, H., Yawei, L.: P-MVSNet: Learning patch-wise matching confidence aggregation for multi-view stereo. In: International Conference on Computer Vision (ICCV) (2019)
5. Silvano, G., Katrin, L., Konrad, S.: Massively parallel multiview stereopsis by surface normal diffusion. In: International Conference on Computer Vision (ICCV) (2015)
6. Qingshan, X., Wenbing, T.: Learning inverse depth regression for multi-view stereo with correlation cost volume. In: AAAI (2020)
7. Yao, Y., Zixin, L., Shiwei, L., Tian, F., Long, Q.: MVSNet: depth inference for unstructured multiview stereo. In: European Conference on Computer Vision (ECCV) (2018)
8. Connelly, B., Eli, S., Adam, F., Dan, B.G.: PatchMatch: a randomized correspondence algorithm for structural image editing. ACM Transactions on Graphics (SIGGRAPH) (2009)
9. Alex, K., Hayk, M., Saumitro, D., Peter, H., Ryan, K., Abraham, B., Adam, B.: End-to-end learning of geometry and context for deep stereo regression. In: International Conference on Computer Vision (ICCV) (2017)
10. Jia-Ren, C., Yong-Sheng, C.: Pyramid stereo matching network. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
11. Haofei, X., Juyong, Z.: AANet: adaptive aggregation network for efficient stereo matching. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
12. Abhishek, K., Christian, H., Jitendra, M.: Learning a multi-view stereo machine. In: Advances in Neural Information Processing Systems (2017)
13. Robert, C.: A space-sweep approach to true multi-image matching. In: Conference on Computer Vision and Pattern Recognition (CVPR) (1996)
14. Xiaodong, G., Zhiwen, F., Siyu, Z., Zuozhuo, D., Feitong, T., Ping, T.: Cascade cost volume for high-resolution multi-view stereo and stereo matching. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
15. Shuo, C., et al.: Deep stereo using adaptive thin volume representation with uncertainty awareness. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
16. Jiayu, Y., Wei, M., Jose, M.A., Miaomiao, L.: Cost volume pyramid based depth inference for multi-view stereo. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
17. Zehao, Y., Shenghua, G.: Fast-MVSNet: sparse-todense multi-view stereo with learned propagation and gaussnewton refinement. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
18. Hui, T.-W., Loy, C.C., Tang, X.: Depth map super-resolution by deep multi-scale guidance. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9907, pp. 353–369. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46487-9_22

Sentiment Analysis Based on CNN and BILSTM Model

Hailong You, Jie Yang(✉), and Zou Xu

College of Electronic Engineering, Heilongjiang University, Harbin 150080, China
guitaryang@126.com

Abstract. Aiming at the problems that the traditional convolutional neural network is easy to cause gradient disappearance and gradient explosion, and it is easy to ignore the context syntax and semantics, this paper proposes a method to integrate CNN and BILSTM. CNN is good at local feature extraction but cannot obtain the meaning of the context. In order to obtain the time information of the data, BILSTM solves the disadvantage that LSTM cannot process the information in reverse, and is used to extract the local structure of the data. In this paper, the local feature representation extracted by CNN is introduced into the emotional feature representation of BILSTM module, which effectively enhances the ability of BILSTM to capture emotional semantic information. Experimental results show that the ensemble model outperforms the two separate models. We are also able to achieve very high accuracy compared to previous work.

Keywords: Emotion analysis · BILSTM · CNN · Text categorization

1 Introduction

With the continuous improvement of hardware technology, the development of deep learning has ushered in new opportunities [1]. Sentiment analysis technology based on deep learning has a good performance in prediction accuracy and efficiency.

Sentiment analysis is also known as opinion mining [2]. The main task of sentiment analysis is to identify positive or negative opinions from text or comment text [3]. The main challenge for computers is to understand all of these points, and in 2015, Tang et al. [4]. Combined multiple deep learning algorithms to exploit the strengths of different algorithms, and sentence representations exploit underlying emotions [5].

The deep learning algorithm model learns simple features close to the input layer through a multi-layered neural network structure [6], and the higher layer derives more complex features based on the information obtained by the lower layer, and uses neurons to carry and transmit information [7], forming a powerful hierarchical structure Feature representation can predict text context, obtain connections between words, and unify the structure of input text [8]. In 2006, Hinton proposed a neural network with multiple hidden layers. This model has excellent feature self-learning ability [9], which made deep learning set off a wave of artificial intelligence in academia and industry [10]. Since 2013,

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 175–182, 2023.
https://doi.org/10.1007/978-981-99-1260-5_22

methods based on deep learning have continuously promoted the development of related tasks in the field of natural language processing.

In recent years, deep neural network models represented by convolutional neural networks and recurrent neural networks have been used to deal with sentiment analysis tasks [11]. In 2010, Mikolov et al. applied RNN to sentiment classification task, However, due to the complexity of the model and the large amount of calculation, the effect is not optimal. In 2012, Sunderm et al. proposed a three-layer LSTM mode [12]. In 2015, Tas to give full play to the advantages of different algorithms. Sentence representation uses CNN and LSTM, document representation uses Gated Recurrent Neural Networks [13], which significantly improve the predictors of sentiment analysis.

This paper attempts to improve the accuracy of sentiment analysis using the ensemble of CNN and Bidirectional LSTM (BILSTM) network [14], and is tested on the publicly available Cornell dataset published by Cornell University and the IMDB dataset published by Stanford University [15]. The structure of the algorithm block diagram of this paper is shown in Fig. 1.

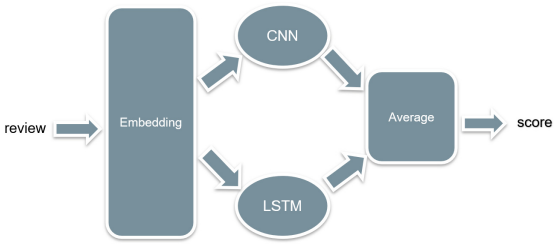


Fig. 1. The block diagram of the propose ensemble model.

2 LSTM Network Model Structure

LSTM is a recurrent network gating algorithm. The cell unit is the core of LSTM.

It is precisely because of the existence of LSTM cells that information can flow through the entire RNN model. LSTM controls the addition and deletion of intracellular information through the gate structure. LSTM consists of three control gates: input gate, forget gate, and output gate. The input gate discriminates incoming information and selectively receives information. The forget gate determines how much information is in the unit state at the previous moment from the current moment. Elimination, the output gate performs nonlinear changes to the current state through the activation function to obtain the output. The internal architecture of a single LSTM is shown in Fig. 2.

The relationship between the input gate, forget gate and output gate is as follows:

$$f_t = \sigma\left(w_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{1}$$

$$i_t = \sigma\left(w_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{2}$$

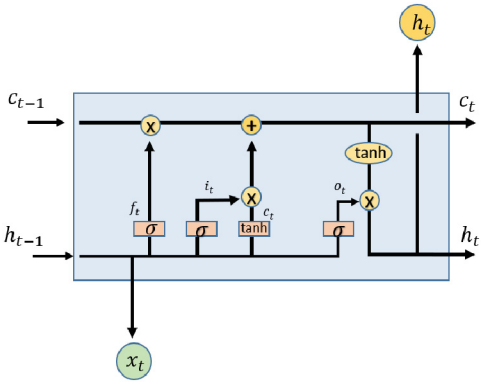


Fig. 2. The architecture of a standard LSTM module.

$$o_t = \sigma\left(w_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{3}$$

$$\tilde{c}_t = \tanh\left[w_c \cdot \left[h_{t-1}, x_t\right] + b_c\right] \tag{4}$$

$$c_t = f_t \cdot c_{t-1} + i_o \cdot \tilde{c}_t \tag{5}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{6}$$

In the above formula, it W_f represents the weight matrix of the forget gate, $[h_{t-1}, x_t]$ makes the vectors connect to each other, and then becomes a new vector larger than the previous one, b_f the bias term is σ expressed in the formula, and he sigmoid function is expressed in the formula. $\tilde{c}_t$ Represents the state of the current input unit, c_t represents the unit state at the current moment, represents the unit state c_{t-1} at the previous moment, and f_t represents the forget gate. The outputs control the current output, a vector between (–1, 1) is calculated by the tanh layer, and the LSTM unit output is obtained by multiplying it with the sigmoid function, where o_t is the ratio of the output information.

For many applications, we are interested in the flow of information in both directions. BILSTM is Bi-directional Long Short-Term Memory, which is a forward LSTM backward LSTM. When inputting, the input information will be provided to two LSTMs at the same time, which continues the advantages of LSTM, solves the gradient problem and LSTM can only perform forward learning. The problem of inability to do backward learning. The architecture of the BILSTM model proposed in this paper is shown in Fig. 3.

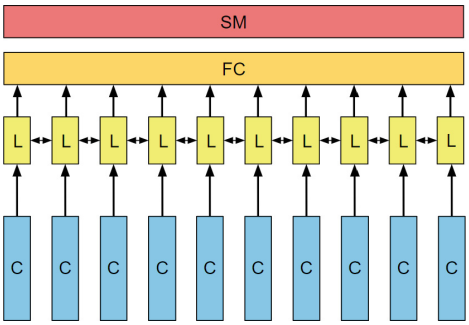


Fig. 3. The architecture of a standard LSTM module.

3 The CNN Model Architecture

The convolutional neural network includes two typical operations, convolution and pooling. The convolution operation enables the parameters of the network to be shared to reduce the scale of the network, and has the function of feature extraction. The pooling operation reduces the parameters and makes the network have certain differences. Denaturation, enabling screening of features.

The output of the convolution operation is only connected to part of the input. When the amount of data is huge, this sparse interaction can greatly reduce the amounts of parameters and speed up the operation. In addition, the sparse interaction can better capture local features and make the operation more efficient. The convolution operation is defined as a formula.

$$S(t) = (x * w)(t) \tag{7}$$

Where x is called input, w is called kernel function, also called convolution kernel, and S is called feature map.

CNN usually embeds different word sentences or paragraphs in one place to form a two-dimensional array, and then passes through the convolutional filter window into a representation of a new array. The new array will then be pooled (usually max-pooling) so that pooled features from different filters are connected to each other and form a hidden representation. These representations are followed by one (or more) fully connected layers to make the final prediction.

In this paper, we use a pre-trained model for word embeddings, then go through a convolutional neural network with 4 filters of size (1, 2, 3, 4), each containing 100 filters Map features. The hidden representation is followed by two fully connected layers, which are then fed into the softmax classifier. The overall architecture of the CNN model proposed in this paper is shown in Fig. 4.

4 Overall Model

LSTM model and the CNN model is good, and both perform well in sentiment analysis. CNN is good at extracting local phrase features of text and classifying according to local features, LSTM is good at extracting sequence features of text and classifying according to the extracted sequence features. This paper proposes a CNN-BILSTM sentiment analysis model based on the difference in the feature extraction methods of the two models. By combining the local phrase feature information extracted by CNN and the sequence features extracted by BILSTM, the performance of the model is further improved. In this paper, the average probability score of the model is used as the final prediction. The overall architecture of our proposed model is shown in Fig. 4.

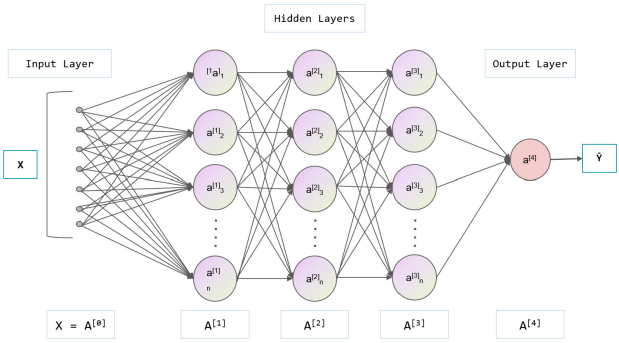


Fig. 4. The general architecture CNN based text classification models.

5 Dataset Introduction

A Cornell Dataset

This corpus consists of movie reviews, including 1,000 positive and negative attitudes; 5,331 sentences each with positive and negative polarities, and 5,000 sentences each with subjective and objective labels. The corpus can be applied to various granularity, such as word, sentence and discourse level sentiment analysis research.

B IMDB Dataset

The IMDB public movie review dataset is provided by Stanford University, with binary sentiment labels, where the "0" label represents a negative sentiment tendency, and the "1" label represents a positive sentiment tendency, with a total of 50,000 pieces of data. The test set and the training set respectively contain 12500 positive and negative samples.

The model is finally run on NVIDIA RTX 3080Ti GPU, Dropout is set to 0.5, activation function adopts sigmoid, and Learning rate is set to 0.001. The loss function uses cross-entropy validation. The optimization function uses ADAM.

6 Model Performance Comparison

In order to verify the superiority of CNN-BILSTM in text sentiment analysis, the models participating in the comparative experiments will be tested on two datasets and tested on the labeled test set, with Accuracy and Loss values as evaluation standard, taking into account the time cost. Experiments were carried out many times, the parameters were adjusted repeatedly, and finally the model data with the best performance was used. The network models involved in the comparison mainly include: LSTM, bidirectional LSTM (BILSTM), convolutional LSTM network (CNN-LSTM), and convolutional bidirectional LSTM network (CNN-BILSTM).

Table 1. Comparative experimental results (IMDB dataset)

model	Accuracy	loss	Time/epoch
LSTM	87.64%	47.37%	349.58
BILSTM	88.02%	37.92%	1109.27
CNN-LSTM	88.68%	28.07%	389.24
CNN-BILSTM	89.24%	28.61%	85.89

Table 2. Comparative experimental results (Cornell dataset)

model	Accuracy	loss	Time/epoch
LSTM	77.68%	59.18%	9.565
BILSTM	78.26%	50.58%	8.659
CNN-LSTM	77.88%	55.81%	8.988
CNN-BILSTM	78.32%	52.76%	7.788

(1) It can be clearly seen from Table 1, Table 2 and Fig. 5 and Fig. 6 that a separate LSTM has obvious shortcomings in capturing contextual information. BILSTM is better than LSTM in extracting "above" information, and its efficiency improved. In contrast, CNN-LSTM comprehensively considers all the information of the context, and the LSTM model combined with CNN has better classification results, which shows that the features extracted by the convolutional layer of CNN can improve the processing of the LSTM layer. Helpful. The accuracy rate and loss rate of the network model combined with BILSTM are generally better than the neural network model using LSTM, which shows that in sentiment analysis tasks, contextual information is necessary to determine the positive and negative tendency of a sentence.

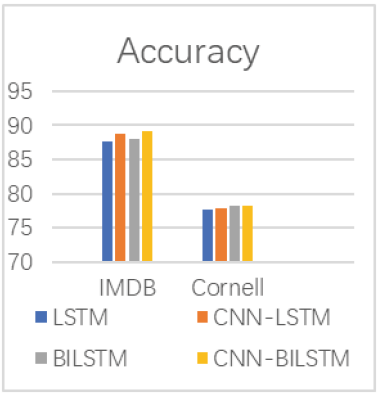


Fig. 5. Model accuracy comparison

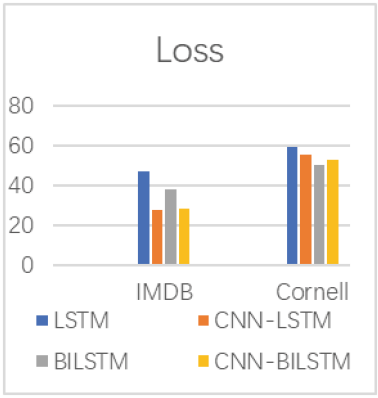


Fig. 6. Model loss comparison

(2) Data set size comparison. Figures 5 and 6 clearly show the comparison of the training performance of the neural network model on the two datasets. The same model trained and tested on the IMDB dataset has an average accuracy of about 13% higher and an average loss rate of about 25% lower than the results of training and testing on the Cornell dataset. The difference is very obvious, reflecting that when the data set is small, the generalization ability of the network model will be quite different, and the ability of the deep neural network model cannot be well utilized. And when the dataset is small, the advantages and disadvantages of the models cannot be well reflected. At this time, CNN is a very good choice. Its fast calculation speed and low cost make it the best choice for processing small-scale texts.

7 Conclusion

In this study, we propose a sentiment analysis framework based on BILSTM and CNN models. Feed the embedded words into CNN and BILSTM models for prediction. The prediction scores of the LSTM and CNN models are then averaged as the final prediction. After experimental research, we found that compared with LSTM, BILSTM, and CNN-LSTM models, the CNN-BILSTM integrated model has a certain improvement in performance. In future work, we will jointly train BILSTM and CNN models to further improve their performance gains relative to a single model. We believe that the prediction accuracy of deep learning-based text processing tasks will be improved by applying this model.

References

1. Subrahmanian, V.S., et al.: The DARPA twitter bot challenge. Computer **49**(6), 38–46 (2016)
2. Hu, M., Liu, B.: Mining opinion features in customer reviews. In: AAAI, vol. 4, no. 4, pp. 755–760 (2004)

3. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge, UK (2016)
4. Li, X., Chen, M., Nie, F.: Locality adaptive discriminant analysis. In: IJCAI 2017, pp. 2201–2207 (2017)
5. Dai, A., Le, Q.V.: Semi-supervised sequence learning. Adv. Neural Inf. Process. Syst. 3079–3087 (2015)
6. Kim, Y.: Convolutional neural networks for sentence classification. In: EMNLP (2014)
7. Bahdanau, D., Cho, K., Oshua Bengio, Y.: Neural machine translation by jointly learning to align andtranslate. In: ICLR (2015)
8. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C.: SSD: Single shot multibox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9905, pp. 21–37. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_2
9. Dai, A.M., Le, Q.V.: Semi-supervised sequence learning. Adv. Neural Inf. Process. Syst. 3079–3087 (2015)
10. Guo, J., Fan, Y., Ai, Q., Croft, W.B.: A deep relevance matching model for ad-hoc retrieval. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, pp. 55–64. ACM (2016)
11. Xie, Q., Dai, Z., Hovy, E., Luong, M.T., Le, Q.V.: Unsupervised data augmentation. arXiv preprint arXiv:1904.12848 (2019)
12. Yang, Z., Dai, Z., Salakhutdinov, R., Cohen, W.W.: Breaking the softmax bottleneck: A high-rank rnn language model. arXiv preprint arXiv:1711.03953 (2017)
13. Zhang, X., Zhao, J., LeCun, Y.: Character-level convolutional networks for text classification. Adv. Neural Inf. Process. Syst. 649–657 (2015)
14. Howard, J., Ruder, S.: Universal language model fine-tuning for text classification. arXiv preprint arXiv:1801.06146 (2018)
15. He, K., Gkioxari, G., Dollr, P., Girshick, R.: Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969 (2017)

A Novel Angle Measuring Method for 2D Digital Array Radars

Anqi Chen(✉), Yuzhu Zhu, Haihui Zhang, Shiming Wu, and Xu Liu

AVIC Aeronautical Radio Electronics Research Institute, Shanghai 200000, China
chloepjk@163.com

Abstract. Based on the idea of monopulse angle measurement and array signal processing, this paper proposed a novel angle measuring method for 2D-DAR. Firstly, the window function method is used for beamforming. Then, the matched filter is used for pulse compression to obtain the range dimension information of the target signal. Finally, the single beat signal of the target location extracted from the target signal is used to accurately measure the Angle based on DFT and angle rotation technology. Simulation results verified that the proposed algorithm is effective and feasible.

Keywords: Digital Array Radar · Array Signal Processing · DFT · DOA

1 Introduction

Conventional monopulse angle measurement is a mature technology of radar. By using a special antenna feed, it can generate four beams with only one monopulse. And the sum beam, azimuth difference beam and elevation difference beam are formed by a mixer at the antenna or RF front end [1–3]. As there are thousands of antenna elements in 2D Digital Array Radar (2D-DAR), microwave networks cannot be used to form monopulse sum beam and difference beam, and the sum beam and difference beam can only be formed by the digital signal processing method, which is called digital sum difference angle measurement [4]. In the field of wireless communication, the author of [5] proposed a uniform planar array (UPA) angle estimation technology based on 2D Discrete Fourier transform (2D-DFT) and angle rotation technology by using the physical characteristics of UPA. The purpose of angle rotation technology is to increase the accuracy of angle estimation [6–8]. Based on the idea of monopulse angle measurement and array signal processing, a novel angle measuring method of 2D-DAR is proposed in the radar field. Firstly, the window function method is used for beamforming. Then, the matched filter is used for pulse compression to obtain the range dimension information of the target signal. And then, the single beat signal of the target location extracted from the target signal is used to accurately measure the Angle based on DFT and angle rotation technology. Finally, the computer simulation results are given and its performance is analyzed.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 183–189, 2023.
https://doi.org/10.1007/978-981-99-1260-5_23

2 2D Digital Array Radar and Digital Beamforming Process

With the development of digital technology, digital technology not only makes the phase shift method variable but also enables the antenna pattern to be formed digitally, which is called digital beamforming technology [9]. The following is a brief introduction to 2D-DAR and digital beamforming process.

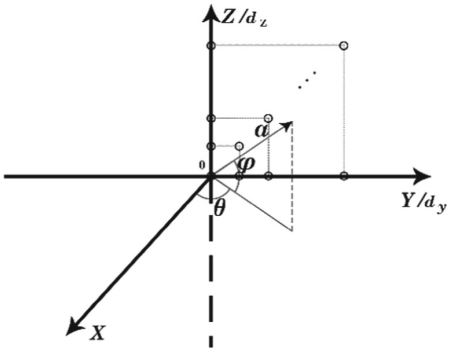


Fig. 1. 2D array antenna coordinate diagram

Figure 1 shows the coordinate system diagram of the 2D-DAR antenna with M rows and N columns vertically placed. The x-axis is the normal line of the array, and the yoz-plane is the plane where the radar array is located. The coordinate center is set as shown in the figure, then the y-axis and z-axis coordinate of each array element is

$$\begin{cases} Y = [0, 1, \ldots, N-1]^T d_y \\ Z = [0, 1, \ldots, M-1]^T d_z \end{cases} \tag{1}$$

where d_y and d_z are the spacing of elements in horizontal and vertical dimensions, respectively.

In the far-field condition, the distance of the target is much larger than the aperture of the antenna array, and the electromagnetic wave returned by the target can be regarded as a planar electromagnetic wave relative to the array. Assuming that $\boldsymbol{r} = \begin{bmatrix} x\, y\, z \end{bmatrix}^T$ is the coordinate vector of any array element in the antenna array, $\boldsymbol{\alpha} = [cos\theta cos\varphi\ sin\theta cos\varphi\ sin\varphi]^T$ is the propagation direction vector of target echo, where θ and φ are azimuth angle and elevation angle, respectively. Then the far point of the coordinate is $\tau = \boldsymbol{r}^T \boldsymbol{\alpha}/c$, c is the propagation speed of electromagnetic wave. Thus, the steering matrix of the array can be obtained as

$$\begin{cases} A(\theta, \varphi) = \boldsymbol{a}_z(\theta, \varphi) \cdot \left[\boldsymbol{a}_y(\theta, \varphi)\right]^T \\ \boldsymbol{a}_z(\theta, \varphi) = \left[1 e^{j\frac{2\pi}{\lambda} d_z \sin\theta \cos\varphi} \ldots e^{j\frac{2\pi(N-1)}{\lambda} d_z \sin\theta \cos\varphi}\right]^T \\ \boldsymbol{a}_y(\theta, \varphi) = \left[1 e^{j\frac{2\pi}{\lambda} d_y \sin\varphi} \ldots e^{j\frac{2\pi(M-1)}{\lambda} d_y \sin\varphi}\right]^T \end{cases} \tag{2}$$

The first and the second equation of (2) are the directional steering vectors of row and column antennas, respectively. Where λ is the wavelength. In the digital beamforming

processing, the weighting matrix of the whole array is

$$\boldsymbol{W}(\theta,\varphi)=\boldsymbol{A}(\theta,\varphi)\odot\left[\boldsymbol{w}_z\boldsymbol{w}_y^T\right]=\left[\boldsymbol{a}_z(\theta,\varphi)\odot\boldsymbol{w}_z\right]\cdot\left[\boldsymbol{a}_y(\theta,\varphi)\odot\boldsymbol{w}_y\right]^T \tag{3}$$

Where, $\boldsymbol{w}_y$ and $\boldsymbol{w}_z$ are window function vectors controlling the elevation and azimuth-dimension sidelobe level respectively, and $\odot$ represents the dot product.

3 Beamforming Process Based on the Window Function

In antenna design, Taylor distribution is generally used for sum distribution, and Bayliss distribution is mainly used for different distribution. Bayliss distribution is a classical difference distribution. It inverts the phases of the elements on the left and right sides or the upper and lower sides of the array to form azimuth or pitch difference beams, and simultaneously, the sidelobe level of the difference beam is reduced. In Eq. (3), the weighting matrix of sum beam is expressed as

$$\boldsymbol{W}_{sum}(\theta,\varphi)=\left[\boldsymbol{a}_z(\theta,\varphi)\odot\boldsymbol{w}_z^{sum}\right]\cdot\left[\boldsymbol{a}_y(\theta,\varphi)\odot\boldsymbol{w}_y^{sum}\right]^T \tag{4}$$

where, $\boldsymbol{w}_z^{sum}$ and $\boldsymbol{w}_y^{sum}$ are Taylor sum distribution of azimuth and elevation dimension, respectively [6]. The distribution of Taylor and is determined by Eq. (5)

$$w(i)=1+2\sum\nolimits_{m=1}^{\overline{n}-1}F_m\cos\left(m\pi\cdot\frac{2i-K-1}{K-1}\right) \tag{5}$$

where, $1\le i\le K$, $F_m=\frac{(-1)^{m+1}}{2\prod_{n-1}\left(1-\left(\frac{m}{n}\right)^2\right)}\cdot\prod_{n=1}^{\overline{n}-1}\left[1-\frac{m^2}{\delta^2\left(A^2+(n-0.5)^2\right)}\right]$, $\overline{n}$ is the number of constant sidelobe levels, K is the number of distribution points. $\delta=\frac{\overline{n}}{\sqrt{A^2+(\overline{n}-0.5)^2}}$, $A=\pi^{-1}\cosh^{-1}\xi$, ξ is sidelobe ratio. Refer to the references [10] for the Bayliss difference distribution.

4 Angle Measuring Method for DAR Based on DFT and Angle Rotation

The most common transmitted signal waveform of pulse compression radar is linear frequency modulation (LFM) signal. The transmitted signal of radar can be expressed as

$$\boldsymbol{s}(t)=\boldsymbol{a}(t)e^{j2\pi f_c t} \tag{6}$$

Where, $\boldsymbol{a}(t)=e^{j2\pi\frac{\mu}{2}t^2}$ represents the complex envelope function of the transmitted pulse signal $s(t)$, f_c is the operating frequency, and μ is the frequency modulation slope. Its time-domain waveform and amplitude-frequency characteristics are shown in the Fig. 2.

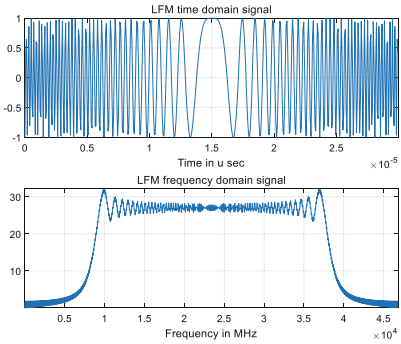


Fig. 2. LFM time-domain signal and frequency-domain signal

In theoretical analysis, the time-domain impulse response of the matched filter of the corresponding $s(t)$ is

$$\boldsymbol{h}(t) = \boldsymbol{s}^*(-t) \tag{7}$$

The impulse response in the frequency-domain is $\boldsymbol{H}(f)$, which is obtained by frequency transformation of $\boldsymbol{h}(t)$.

For the target to be detected at the distance R, the reflected echo signal can be expressed as

$$\boldsymbol{s}_r(t) = A_0 a(t-\tau) e^{j2\pi(f_c+f_d)t} \tag{8}$$

Where, $A_0 = \gamma\sqrt{\lambda^2/(4\pi)^2 R^4}$ represents the amplitude fluctuation of the target echo signal, and γ is a random variable describing the reflection characteristics of the target, which depends on the selected target fluctuation model. τ is the two-way delay of the target echo signal, that is, $\tau = 2R/c.f_d$ is the Doppler shift generated by the point target relative to the radar motion. When the reflected signal reaches the two-dimensional array antenna, the signal received by the array antenna can be expressed as

$$\boldsymbol{Y}_{echo} = \boldsymbol{A}(\theta, \varphi)\boldsymbol{s}_r(t) + \boldsymbol{N} \tag{9}$$

Where, $\boldsymbol{N}$ is Complex Gaussian white noise matrix. The received signal $\boldsymbol{Y}_{echo}$ is filtered by $\boldsymbol{W}_{sum}(\theta, \varphi)$ weighting matrix to get the sum beam s_{sum}

$$\boldsymbol{s}_{sum} = [\boldsymbol{W}_{sum}(\theta, \varphi)]^H \boldsymbol{Y}_{echo} \tag{10}$$

Then, the pulse compressed signal $\boldsymbol{s}_{sum}^{pc}$ is obtained by matching filtering $\boldsymbol{s}_{sum}$. The time-domain waveform of $\boldsymbol{s}_{sum}^{pc}$ is shown in the Fig. 3.

Take the point's information with the most concentrated energy and convert it into an $M \times N$ matrix, which is the single beat signal $\boldsymbol{Y}_{ti}^{\boldsymbol{echo}}$. Then the Angle is measured accurately based on DFT and Angle rotation technology for $\boldsymbol{Y}_{ti}^{\boldsymbol{echo}}$. Firstly, the two standard DFT matrices $\boldsymbol{F}_{\mathrm{M}}$ and $\boldsymbol{F}_{\mathrm{N}}$ for M and N points are defined, where $\boldsymbol{F}_{\mathrm{M}}$ is the $M \times M$ DFT matrix, and the element can be expressed as $[\boldsymbol{F}_{\mathrm{M}}]_{p,q} = \frac{1}{\sqrt{M}} e^{-j\frac{2\pi}{M}pq}, p, q =$

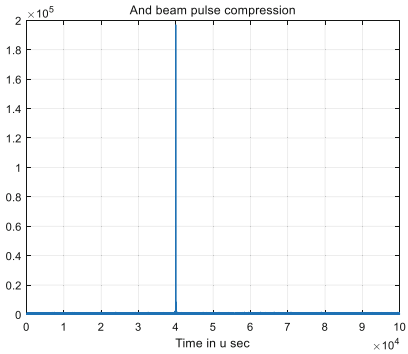


Fig. 3. Time-domain signal s_{sum}^{pc} after pulse compression

$0, 1, \ldots M-1$, $\boldsymbol{F}_{\mathrm{N}}$ is $N \times N$ DFT matrix, the element can be expressed as $[\boldsymbol{F}_{\mathrm{N}}]_{p',q'} = \frac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}p'q'}, p', q' = 0, 1, \ldots N-1$.

Then, the signal $\tilde{\boldsymbol{Y}}_{ti}^{\boldsymbol{echo}}$ processed by DFT and angular rotation technology can be expressed as

$$\tilde{\boldsymbol{Y}}_{ti}^{\boldsymbol{echo}} = \boldsymbol{F}_{\mathrm{M}} \boldsymbol{\Phi}\left(\psi_y\right) \boldsymbol{Y}_{ti}^{\boldsymbol{echo}} \boldsymbol{\Phi}(\psi_z) \boldsymbol{F}_{\mathrm{N}} \tag{11}$$

Where, $\boldsymbol{\Phi}\left(\psi_y\right) = diag\left\{\left[1, e^{j\psi_y}, \ldots, e^{j(M-1)\psi_y}\right]\right\}$, $\psi_y \in [-(\pi/M), \pi/M]$ spatial rotation matrix of azimuth dimension. $\boldsymbol{\Phi}(\psi_z) = diag\left\{\left[1, e^{j\psi_z}, \ldots, e^{j(N-1)\psi_z}\right]\right\}$, $\psi_z \in [-(\pi/N), \pi/N]$ is the spatial rotation matrix of elevation dimension. When the optimal spatial rotation angle is found, the energy and power can be more concentrated on a small number of DFT points, thereby reducing energy leakage and obtaining more accurate DOA estimation. Denoted $\{p_0, q_0\}$ as the optimal spatial signature information, denoted ψ_z and ψ_y as the optimal rotation angle of z-axis and y-axis, respectively. Then the angle information of the target can be expressed as

$$\begin{cases} \hat{\varphi} = \sin^{-1}\left(\frac{\lambda q_0}{\mathrm{N}d} - \frac{\lambda \psi_y}{2\pi d}\right) \\ \hat{\theta} = \sin^{-1}\left(\left(\frac{\lambda p_0}{\mathrm{M}d} - \frac{\lambda \psi_x}{2\pi d}\right) / \cos\hat{\varphi}\right) \end{cases} \tag{12}$$

5 Numerical Simulation and Discussions

The effectiveness of the proposed method is verified by numerical simulation. The 2D-DRA is equipped with 32×32 and 64×64 UPA, and the antenna interval is $d = \lambda/2$, and the search times of angle rotation are set as $N_{ro} = 10$ and $N_{ro} = 30$. The linear range of the target is within the detectable range of the 2D-DAR, and the azimuth angle and elevation angle are within the detectable angle of the 2D-DAR. Definition of signal noise ratio (SNR), that is $SNR = 10\log_{10} P_1/\sigma^2$. The performance metric of angle

measurement is the mean square error of the average single angle ¼ be expressed as

$$\mathrm{ARMSE} = \frac{1}{2}\left(\sqrt{\frac{\left\|\hat{\theta}-\theta\right\|^2}{\|\theta\|^2}+\frac{\left\|\hat{\varphi}-\varphi\right\|^2}{\|\varphi\|^2}}\right) \tag{13}$$

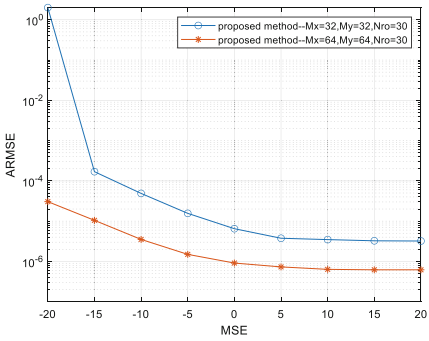


Fig. 4. Comparison ARMSE performances of the proposed method for different antenna numbers with SNR.

The Fig. 4 shows the curve of ARMSE changing with SNR under the array antenna configuration of 32×32 and 64×64 the number of angle rotation search $N_{ro} = 30$, and different array antenna numbers.

It can be seen from the Fig. 4 that under the same search times of angle rotation, the larger the number of array antennas, the better the angular measurement performance of the algorithm.

The Fig. 5 shows the curve of ARMSE changing with SNR under different angle rotation search times when the array antenna configurations are 32×32 and 64×64, respectively. The angle rotation search times are set as $N_{ro} = 10$ and $N_{ro} = 20$.

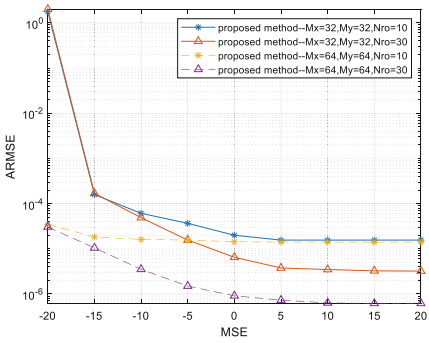


Fig. 5. Comparison ARMSE performances of the proposed method for different angler rotation search times and different antenna with SNR.

As can be seen from the Fig. 5, when the number of array antennas is the same, under low SNR and different search times of Angle rotation, the angular measurement performance of the algorithm is the same. Under high SNR, the greater the search times of Angle rotation, the better the angular measurement performance of the algorithm.

6 Conclusion

In this paper, the angle measurement of 2D-DAR is studied. Firstly, the window function method is used for beamforming. Then, the matched filter is used for pulse compression to obtain the range dimension information of the target signal. Finally, the single beat signal of the target location extracted from the target signal is used to accurately measure the Angle based on DFT and angle rotation technology. Computer simulation results demonstrate the effectiveness of the proposed algorithm.

References

1. Cantrell, B., De Graaf, J., Willwerth, F., et al.: Development of a digital array radar (DAR). IEEE Aerosp. Electron. Syst. Mag. **17**(3), 22–27 (2002)
2. Yue, X., Wan, W., Ning, B., et al.: An active phased array radar in China. Nature Astron. **6**(5), 619 (2022)
3. Harger, M., Conway, M.D., Thomas, H., et al.: Fully digital phased array development for next generation weather radar. In: 2021 IEEE Radar Conference (RadarConf21), pp. 1–6 ((2021))
4. Yueli, L., Meng'en, M., Chonghui, Z., et al.: Forward-looking Imaging via doppler estimates of sum-difference measurements in scanning monopulse radar. J. Radars, **10**(1), 131–142 (2021)
5. Fan, D., Gao, F., Huawei, L., et al.: Angle domain channel estimation in hybrid mmwave massive MIMO systems. IEEE Trans. Wireless Commun. **17**(12), 8165–8179 (2018)
6. Xie, H., Gao, F., Jin, S., et al.: Channel estimation for TDD/FDD massive MIMO systems with channel covariance computing. IEEE Trans. Wireless Commun. **17**(6), 4206–4218 (2018)
7. Fan, D., Gao, F., Wang, G., et al.: A 2D-DFT based channel estimation scheme in indoor 60 ghz communication systems with large-scale multiple-antenna. In: IEEE Vehicular Technology Conference, pp. 1–5 (2016)
8. Xie, H., Gao, F., Zhang, S., et al.: UL/DL channel estimation for TDD/FDD massive MIMO systems using DFT and angle reciprocity. In: IEEE Vehicular Technology Conference, pp. 1–5 (2016)
9. Shunjun, W., Xiaochun, M., et al.: Radar Signal Processing and Data Processing Technology. Publishing House of Electronics Industry, Beijing (2008)
10. Yang, B.B.: An angle measuring method using sum and difference beams for 2D digital array radars. RADAR & ECM **34**(3), 6–11 (2014)

Blockchain-Driven Revocable Ciphertext-Policy Attribute-Based Encryption for Public Cloud Data Sharing

Zhaoqian Zhang, Jianbiao Zhang(✉), Shuang Kang, and Zheng Li

Faculty of Information Technology, Beijing Key Laboratory of Trusted Computing, Beijing University of Technology, Beijing 100124, China
zjb@bjut.edu.cn

Abstract. Sharing data through the public cloud has turned into a popular way to exploit the data value today, and ciphertext-policy attribute-based encryption (CP-ABE) can protect data confidentiality while improving sharing efficiency. In CP-ABE, attribute change raises the issue of revocation, which is solved by a third party in the existing schemes. However, it cannot achieve timely and certain revocation. Thus, this paper proposes a blockchain-driven revocable CP-ABE scheme. We design a re-writable blockchain to realize attribute revocation, and the ciphertext no longer needs to be updated. Attribute authority only needs to update an *update factor* bound to the attribute and the private keys for non-revoked users, while the revoked users cannot recover the outdated ciphertext. Moreover, our scheme is selectively secure and collusion resistant. Compared with other schemes, the performance of our scheme is satisfactory.

Keywords: Data sharing · ciphertext-policy attribute-based encryption · access control · blockchain

1 Introduction

Nowadays, more and more individuals and organizations share their sensitive data through the public cloud server because of higher performance, greater flexibility, and better customer support. Although data can be shared anytime and anywhere, the lax security mechanism of the public cloud server increases the risk of data leakage, and the data owners lose complete control over the data. Thus, to prevent data leakage, data is often encrypted before sharing. Ciphertext-policy attribute-based encryption (CP-ABE) [1,2] is a promising cryptographic mechanism to protect data confidentiality while improving sharing efficiency. Data owners encrypt the data with the policy formed by a set of attributes, and data users whose attributes satisfy the policy can decrypt the ciphertext.

J. Zhang—This work was supported by the Natural Science Foundation of Beijing Municipality under Grant M21039.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 190–197, 2023.
https://doi.org/10.1007/978-981-99-1260-5_24

However, the access revocation due to the attribute changes hinders the application of CP-ABE in public cloud data sharing.

In CP-ABE, the decryption is only related to the user's attribute set. When a user's attribute set changes, he should not have access rights corresponding to the previous attribute set. If a user's privileges are not revoked on time, it can lead to unauthorized access, which causes data leakage. For example, in the "Patient-Hospital" scenario, a physician who no longer holds the "Senior" title due to medical malpractice cannot access Personal Health Records (PHRs) that only "Senior" physicians can access. Revocation in CP-ABE includes two types, user revocation, and attribute revocation. User revocation [3,4] is coarse-grained, which revokes the user's all access rights. However, it is not appropriate for many scenarios. For example, a physician should not have all access rights revoked arbitrarily when his attributes change. On the contrary, attribute revocation [5] revokes only the user's relevant attribute, not all privileges. Compared to user revocation, fine-grained attribute revocation has more flexibility.

Revocation often requires updating the ciphertext and private key. For the ciphertext update, the existing schemes realize it through a third party [6,7], which faces a severe trust crisis and centralization risk and loses the autonomy of revocation. On the one hand, third parties are lazy. They may wait for an idle time to perform the update to save computing power; on the other hand, when a third party becomes untrustworthy due to attacks, it may be unable to update or refuse to update. Relying on a third party can not guarantee the timeliness and certainty of revocation, resulting in unauthorized access.

This paper proposes a revocable CP-ABE based on blockchain to provide timely and certain attribute revocation. To ensure certainty, we employ a re-writable blockchain to update the *update factor* UF bound to the attribute, and the ciphertext does not need to update. Since the blockchain no longer contains the old version UF, the revoked user cannot recover the outdated ciphertext. To ensure timeliness, we set the highest priority for the transaction RT used for update, and UF will be updated in the next consensus phase. Besides, we strictly analyze the security and performance of the scheme and thoroughly compare it with other similar schemes, which shows our advantages.

2 Revocable CP-ABE Based on Re-writable Blockchain

2.1 System Model

The system has five entities, depicted in Fig. 1, including the cloud server, the data owner, the data user, the attribute authority, and the blockchain.

- **Cloud Server (CS)**
  stores the ciphertext that does not need to be updated. Suppose the CS is untrusted.
- **Data Owner (DO)** specifies a access policy and creates the ciphertext which will be uploaded to the CS.

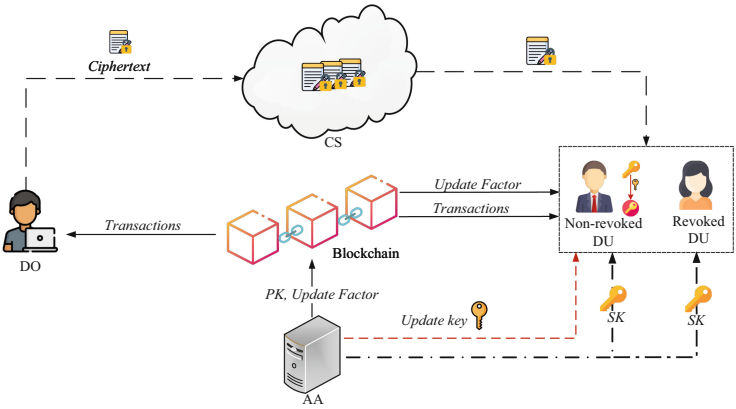


Fig. 1. System model.

- **Data User (DU)** decrypts the ciphertext downloaded from the CS. When a revocation event occurs, the non-revoked users update the private keys using the update keys from the attribute authority.
- **Attribute Authority (AA)** initiates the system and distributes the private keys for the DUs. Once an attribute revocation occurs, AA first updates the revocation list, then publishes the new *update factor* of the attribute in the blockchain and distributes update keys for the non-revoked users. AA is assumed to be fully trusted.
- **Blockchain** records and updates the *update factors*. Besides, it also records the public key of the system.

2.2 Chameleon Hash Function

CH function is a trapdoor collision-resistant hash function that contains a key pair (CH_{sk}, CH_{pk}) and the following algorithms. Only the one who holds the trapdoor CH_{sk} can obtain a hash collision to update the hashed data.

- $CH\_KeyGen(1^{\kappa})$. The algorithm takes as input a security parameter κ, outputs the private key CH_{sk} used as a trapdoor and the public key CH_{pk} used for hash.
- $CH\_Hash(m, r, CH_{pk})$. The algorithm takes as input a message m, a random number r and the CH public key CH_{pk}, outputs chameleon hash value $\mathcal{H}$.
- $CH\_Update((m, r), m^{'}, CH_{sk})$. The algorithm takes as input the previous message m, random number r corresponding to m, a new message $m^{'}$, and the CH private key CH_{sk}, outputs an updated random number $r^{'}$.
- $CH\_Verify((m, r), (m^{'}, r^{'}), CH_{pk})$. The algorithm takes as input the previous message m and its random number r, new message $m^{'}$ and its random number $r^{'}$, and the CH public key CH_{pk}, and check $CH\_Hash(m, r, CH_{pk}) \stackrel{?}{=} CH\_Hash(m^{'}, r^{'}, CH_{pk})$, if equal, output 1, otherwise output 0.

2.3 Concrete Construction

– **System Setup**$(1^\kappa, U)$. AA executes the algorithm to initialize the system. It takes as input security parameter κ and the number of attributes U. Firstly, AA chooses a group $\mathbb{G}$ of prime order p, and a generator g. Then AA randomly selects $\alpha, a, \beta, \tau, c \in \mathbb{Z}_p$, and for each attribute, it randomly selects $A_1, A_2, ..., A_U \in \mathbb{Z}_p$. Besides, it sets $CH_{sk} = c, CH_{pk} = g^c$.
The public key PK is

$$g, e(g,g)^\alpha, g^a, g^\beta, g^\tau, A_1, ..., A_U, CH_{pk}.$$

For the attribute group AG_i consists of the DUs holding A_i, chooses $v_i, \epsilon_i \in \mathbb{Z}_p$ randomly, where v_i can be considered as the version number of the attribute A_i, and ϵ_i is the random number used to update v_i.
The master secret key MSK is

$$g^\alpha, g^{v_1}, ..., g^{v_U}, CH_{sk}.$$

The *update factors* are

$$UF_1 = (\frac{v_1 - \tau}{\beta}, \epsilon_1), ..., UF_U = (\frac{v_U - \tau}{\beta}, \epsilon_U).$$

– **Encryption**$(PK, (M, \rho), \mathcal{M})$. The algorithm is executed by DO. It takes as input the public key PK, the LSSS access structure (M, ρ) generated by a policy, and a message $\mathcal{M}$. Firstly, DO selects a vector $v = (s, v_2, v_3, ..., v_n) \in \mathbb{Z}_p$ randomly. For $i \in [1, \ell]$, DO computes $\lambda_i = M_i v^T$, where M_i is the i-th row of the matrix M. Besides, DO randomly selects $r_1, ..., r_\ell \in \mathbb{Z}_p$. Finally, the ciphertext CT is

$$C = \mathcal{M} e(g,g)^{\alpha s}, C^{'} = g^s,$$
$$\forall i \in [1, \ell], C_i = g^{a\lambda_i} g^{A_{\rho(i)} r_i \tau}, D_i = g^{-r_i}, R_i = g^{\beta r_i A_{\rho(i)}}.$$

– **Key Generation**(MSK, S). The algorithm is executed by AA. It takes as input the master secret key MSK and DU's attribute set S. For each DU, AA selects a random $t \in \mathbb{Z}_p$. The private key SK is

$$K = g^\alpha g^{at}, L = g^t, \forall x \in S, K_x = g^{A_x t v_x}.$$

– **Decryption**(CT, SK). The algorithm is executed by DU. It takes as input the ciphertext CT and DU's secret key SK. Suppose S satisfies the policy, so that defining $I \subset \{1, ..., \ell\}$ as $I = \{i : \rho(i) \in S\}$. Firstly, DU needs to find a set of constants $\omega_i \in \mathbb{Z}_p$ such that $\omega M = (1, 0, ..., 0)$, thus $\sum_{i \in I} \omega_i \lambda_i = s$. Then, DU computes

$$e(g,g)^{\alpha s} = \frac{e(C^{'}, K)}{\prod_{i \in I} (e(C_i \cdot R_i^{\frac{v_i - \tau}{\beta}}, L) e(D_i, K_{\rho(i)}))^{\omega_i}} \quad (1)$$

Finally, DU can get $\mathcal{M}$ from C.

2.4 Re-writable Blockchain Design

We update the *update factor* UF through the re-writable blockchain. There are two kinds of transactions in the blockchain, Normal Transaction (NT), and Re-writable Transaction (RT). NT is immutable and is used to record common data, such as public keys, etc. RT is re-writable, and the old RT is updated by the new RT. In the transactions, G is the transaction hash function, H is the data hash function, and both are collision-resistant. $\mathcal{H}_i$ is the chameleon hash value of UF_i. The original RT_{og} is generated by DO, and the subsequent RT_{up}s can only created and published by AA who owns the trapdoor CH_{sk}. The detailed structures of two transactions are shown as follows.

$$
\begin{aligned}
NT &= (ID, Data, sign(anyone), G(Data, sign), timestamp) \\
RT_{og} &= (ID, RT.ID, (UF_i, \mathcal{H}_i), sign(DO), G(\mathcal{H}_i, sign), timestamp) \\
RT_{up} &= (ID, RT.ID, (UF_i, \mathcal{H}_i), sign(AA), G(\mathcal{H}_i, sign), timestamp)
\end{aligned}
$$

In the transaction RT, $RT.ID$ denotes the target transaction that RT will update, and the $RT.ID$ of the RT_{og} is *null*. We design a backtracking update to avoid the revoked users from getting the old UF in the ledger. As shown in Fig. 2, when a new RT_{up} is received, then all the target transactions associated with it will be updated in sequence in the order of $RT.ID$.

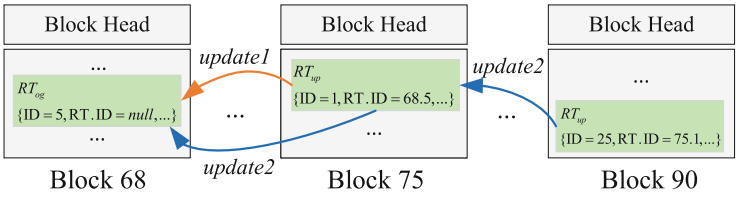


Fig. 2. Backtracking update.

To update the *update factor* UF timely, we set the priority of the transaction RT to the highest. Besides, after receiving a new block, full nodes perform regular block validation, while the consensus nodes need to republish the RTs whose timestamps are earlier than the newest block until they are written to the subsequent new block. Since the transactions can not be tampered with except for RT, no user can reject the update unless he no longer uses the blockchain ledger.

2.5 Attribute Revocation

The attribute revocation in our scheme does not need to update the ciphertext. If a DU no longer holds the attribute A_u, AA only needs to update the UF_u and the non-revoked DUs' SK.

To update the UF_u, AA randomly selects $v_u^{'} \in \mathbb{Z}_p$, and computes $\frac{v_u^{'}-\tau}{\beta}$. Then AA executes CH_Update to find a $\epsilon_u^{'}$ to make the hash value of $\frac{v_u-\tau}{\beta}$ unchanged. AA first computes

$$\epsilon_u^{'} = \frac{\epsilon_u \cdot CH_{sk} - H(\frac{v_u^{'}-\tau}{\beta}) + H(\frac{v_u-\tau}{\beta})}{CH_{sk}} \tag{2}$$

Finally, $UF_u^{'} = (\frac{v_u^{'}-\tau}{\beta}, \epsilon_u^{'})$, and AA publishes the following transaction to the blockchain.

$$RT_{up} = (ID, RT.ID, (UF_u^{'}, \mathcal{H}_u), sign(AA), G(\mathcal{H}_u, sign), timestamp)$$

When the above RT_{up} is received, consensus nodes execute CH_Verify. It first obtains $UF_u, \mathcal{H}_u$ from the target RT to be updated, then checks $\mathcal{H}_u \stackrel{?}{=} CH_{pk}{}^{\epsilon_u^{'}} \cdot g^{H(\frac{v_u^{'}-\tau}{\beta})}$.

If the above equation holds, the RT_{up} will be packed into a new block in the next consensus round.

To update the SK, AA computes $UK_u = g^{A_u t(v_u^{'}-v_u)}$ and sends it to non-revoked DU in the attribute group AG_u. DU updates SK as $K, L, K_u^{'} = K_u \cdot UK_u, \forall x \in S \backslash u, K_x^{'} = K_x$.

The above updates for the *update factor* UF and the private keys achieve fine-grained attribute revocation. The revoked user cannot get UK or outdated ciphertext to realize unauthorized access.

3 Security Analysis

Our scheme has the following security properties.

- **Confidentiality**. Our scheme can achieve selective security under the q-parallel BDHE assumption. There is no polynomial-time adversary has the non-negligible advantage to break our scheme selectively by a challenge access structure. The proof is similar to the Waters' scheme [2].
- **Forward Security**. Newly joined users get the new version of the *update factor* UF by synchronizing the blockchain ledger. If their attribute set matches the policy, they are able to decrypt the ciphertexts published in a previous time, which are already updated.
- **Backward Security**. Users who are revoked are unable to update his SK after attribute revocation. AA updates private keys only for non-revoked users, and the only way for a revoked user to update his SK is by getting the UK_i from a corrupted non-revoked user. However, each $UK_i = g^{A_i t(v_i^{'}-v_i)}$ is associated with t which is different from each other.
- **Collusion resistant**. The private key for each user contains a t related to the user, and different users cannot combine their SKs to decrypt the ciphertexts.

4 Performance Analysis

We evaluate the computation overhead of our scheme and make a comprehensive comparison with Guo et al. [8] and Waters' scheme [2] which we are mainly based on. The experiments were performed on a desktop computer with Ubuntu 18.04 LTS, Intel(R) Core(TM) i5-4590 CPU@ 3.30 GHz with 4 GB RAM and the Java paring-based cryptography library 2.0.0, and we used the Type A curve over the field $\mathbb{F}_q$ for some prime $q = 3 \mod 4$. For each experiment, we used the mean of 50 replicates, and the message $\mathcal{M}$ is randomly selected in $\mathbb{G}_{\mathbb{T}}$.

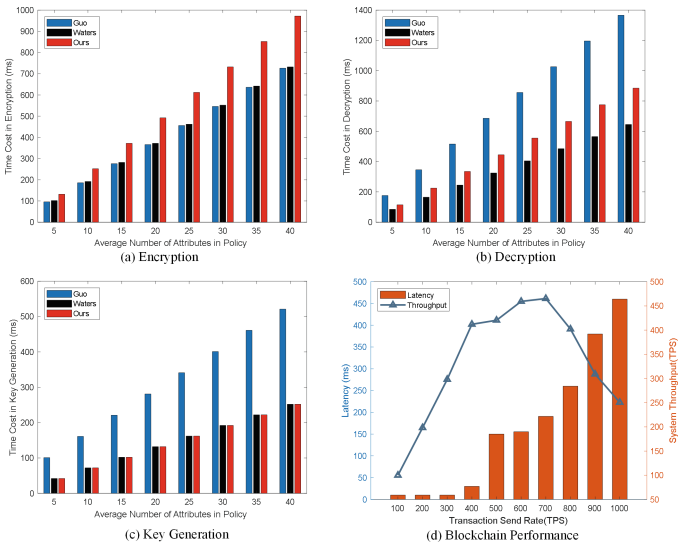


Fig. 3. Performance comparison.

The encryption cost of our scheme is the largest. This is because we add a parameter R_i in the ciphertext for each attribute in the policy to help DU get the latest ciphertext component C_i. Compared to Waters' scheme, ours requires only one additional exponentiation operation for each attribute in the encryption and decryption algorithm, while the computational overhead in key generation is the same for both schemes and much smaller than Guo's scheme.

Transaction confirmation time (latency) and throughput of the blockchain reflect the efficiency of the update of UF. We chose Hyperleder Fabric v2.2 as the blockchain platform and tested blockchain performance using Hyperledger Caliper[1], a performance benchmark tool capable of measuring different blockchain platforms. The Block size is 200 and the block interval time is 1s. We conduct tests under the Raft consensus mechanism, and the result is shown in Fig. 3(d). System throughput peaks at a transaction send rate of 700 and

[1] Hyperledger caliper. https://www.hyperledger.org/projects/caliper.

decreases thereafter. This is because as the number of transactions increases in the blockchain network, more communication time is needed between nodes, and the system throughput decreases. When the transaction number reaches the preset value during the block interval time, the consensus mechanism will batch the transactions into a new block each time. Therefore, a blockchain with an efficient consensus mechanism enables ciphertext updates in a timely manner; more importantly, it can achieve deterministic updates.

5 Conclusion

In this paper, we proposed a revocable CP-ABE scheme based on blockchain to achieve timely and certain attribute revocation. We discarded the third party and designed a re-writable blockchain based on the chameleon hash function to perform the revocation. To be precise, in addition to the private keys, only the *update factor* UF needs to be updated, and it can be updated within a consensus time. The revoked users cannot recover the outdated ciphertext because the old UF is rewritten by the new UF. Besides, our scheme is proved to be selectively secure and collusion resistant, and the computation overhead of our scheme is affordable compared to other schemes.

References

1. Bethencourt, J., Sahai, A., Waters, B.: Ciphertext-policy attribute-based encryption. In: 2007 IEEE Symposium on Security and Privacy (SP 2007), pp. 321–334 (2007). https://doi.org/10.1109/SP.2007.11
2. Waters, B.: Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization. In: International Workshop on Public Key Cryptography, pp. 53–70 (2011)
3. Yu, K., Tan, L., Aloqaily, M., Yang, H., Jararweh, Y.: Blockchain-enhanced data sharing with traceable and direct revocation in IIoT. IEEE Trans. Industr. Inform. **17**(11), 7669–7678 (2021)
4. Li, Q., Xia, B., Huang, H., Zhang, Y., Zhang, T.: TRAC: traceable and revocable access control scheme for mHealth in 5G-enabled IIoT. IEEE Trans. Industr. Inform. **18**(5), 3437–3448 (2022). https://doi.org/10.1109/TII.2021.3109090
5. Hur, J., Noh, D.K.: Attribute-based access control with efficient revocation in data outsourcing systems. IEEE Trans. Parallel Distrib. Syst. **22**(7), 1214–1221 (2010)
6. Zhong, H., Zhou, Y., Zhang, Q., Xu, Y., Cui, J.: An efficient and outsourcing-supported attribute-based access control scheme for edge-enabled smart healthcare. Futur. Gener. Comput. Syst. **115**, 486–496 (2021)
7. Yang, Y., Shi, R.-H., Li, K., Wu, Z., Wang, S.: Multiple access control scheme for EHRs combining edge computing with smart contracts. Futur. Gener. Comput. Syst. **129**, 453–463 (2022)
8. Guo, R., Yang, G., Shi, H., Zhang, Y., Zheng, D.: O^3-R-CP-ABE: an efficient and revocable attribute-based encryption scheme in the cloud-assisted IoMT system. IEEE Internet Things J. **8**(11), 8949–8963 (2021)

D2D-Aided Multicasting with Multi-armed Bandit Based Clustering

Zhiping Wu(✉)

The Third Research Institute of Ministry of Public Security, Shanghai 201204, China
zhiping.wu@aliyun.com

Abstract. With a growing demand for high-quality mobile live video streaming services, cellular networks are facing an unprecedented challenge of dramatically increased mobile data usage. Recently, device-to-device (D2D) aided multicasting has been introduced and considered as one of the promising techniques to tackle such imminent mobile live video streaming dilemma of cellular networks by enabling direct data transmissions among users in proximity. The clustering strategy is essential to the D2D-aided multicasting in cellular networks. To address this issue, a novel clustering policy is presented in this paper, under which a BS simply groups the associated users based on their received sounding reference signal (SRS) strength and selects the users with the largest average received SRSs as the cluster head in each group to maximize the multicasting performance. Further, by applying the idea of multi-armed bandits, the multicasting latency induced by the process of clustering and cluster head selection can be significantly reduced. We analyze the performance of the D2D-aided multicasting network. The effectiveness of the proposed multi-armed bandit based clustering strategy is demonstrated through extensive simulations.

Keywords: D2D aided multicasting · Multi-armed bandit based clustering · Stochastic geometry · coverage probability · spatial throughput

1 Introduction

Mobile live video streaming has attracted tremendous interest over recent years due to the dramatically increased demand on live E-sports, live commerce, and "in real life" (IRL) streaming. Particularly, the mobile users of live video streaming in China reached 560 million in March, 2020, an increase of 163 million over that in December, 2018. The fast growing traffic induced by the mobile live video streaming has placed a significant burden on cellular networks. Thus, how to satisfy the quality-of-service (QoS) requirements of mobile live video streaming while alleviate the potential on-air congestions at the base stations (BSs) has become one of the major technical challenges for cellular network designers.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 198–205, 2023.
https://doi.org/10.1007/978-981-99-1260-5_25

Multicasting is considered as an effective approach to enhance the delivery performance of mobile live video streaming. Compared with the conventional unicasting, multicasting can satisfy multiple requesters concurrently and thereby significantly improves the spectrum efficiency of the cellular network. Tremendous efforts have been put into the study of multicasting of mobile live video streaming [1–9]. In [1], Wang *et al.* studied the scalable video coding based multicasting and evaluate the economic value of the system via a game-theoretic approach. In [2], Christopoulos *et al.* addressed frame-based precoding problem in the forward link of a broadband multibeam satellite system by means of the multicast-aware user scheduling policy. In [3], Zhou *et al.* investigated the multicast over a large-scale mobile ad hoc network and derived the delay constrained capacity. In [4], Li *et al.* considered the multicasting of the hard deadline constrained prioritized multimedia streams and proposed a novel random network coding scheme to maximize the network performance. In [5], Sim *et al.* investigated the opportunistic multicast beamforming to maximizes the minimum SNR of the selected subset of users. In [6], Zhang *et al.* considered applying the technologies of millimeter wave and non-orthogonal multiple access to enhance the multicast performance. In [7], Bejerano *et al.* presented a novel dynamic monitoring system architecture to ensure the high quality of service of the multimedia multicasting. In [8], Xu *et al.* developed a novel mathematical framework to analyze the multi-view video transmissions by exploiting both natural multicast opportunities and view synthesis-enabled multicast opportunities. In [9], Araniti *et al.* investigated the MBSFN area formation algorithm in 5G NR networks to enhance the multicasting performance of the SVC traffic.

Recently, the technique of device to device (D2D) communication has been widely studied and considered as a promising approach to enhance the performance of cellular networks by offloading the traffics from the overloaded BSs. Particularly, D2D communication enables the direct transmission between users in proximity, which thereby can significantly improve the spectrum and power efficiency of the system. Due to its advantages, several works have introduced the D2D communication into the multicasting scenarios to guarantee the quality of service of video transmissions [10–12].

One of the important issues in D2D aided multicasting is clustering. With proper clustering strategy, the multicasting performance can be further enhanced due to the boosting of the received signal to interference power ratio. It is worth noting that in the above mentioned works on D2D aided multicasting [10–12], the issue of clustering was either ignored [11] or not well addressed [10,12]. It is also worth noting that though clustering can dramatically improve the network throughput, the delay induced by the process of D2D clustering may significantly deteriorate the QoS of the mobile live video transmissions, especially for large amount of users.

Recently, the emerging machine learning framework has been applied in the study of wireless communication and shows great potentials to enhance the network performance [14–16]. Particularly, in [13], Bagaria *et al.* proposed a novel multi-armed bandit based Med-dit algorithm to determine the exact medoid of a data set with high probability. Compared with C_n^k evaluations required by the

naive method for a data set with n points, the Med-dit algorithm only needs $O(n \log n)$ computations, which thereby can significant enhance the mediod searching efficiency. In this paper, motivated by [13], we investigate the multi-armed bandit based clustering algorithm in D2D-aided multicasting networks to reduce the processing delay induced by the cluster heads searching. The main contributions of this paper are summarized as follows:

- We propose a novel multi-armed bandit based clustering strategy in D2D-aided multicasting network to simultaneously enhance the received SIR of users and reduce the processing delay induced by the D2D cluster heads searching. Particularly, under the proposed clustering strategy, the cluster head is selected to maximize the average received signal power at all users in the respective cluster. Further, by applying the method of multi-armed bandit, the cluster head of each cluster is determined based on the confidence interval of the estimated (rather than the exact) average received sounding reference signal (SRS) power broadcasted by the D2D users.
- We develop a stochastic geometry framework to analyze the performance of D2D-aided multicasting networks under the proposed multi-armed bandit based clustering strategy.
- Extensive simulation results are presented under a variety of network settings to validate the proposed multi-armed bandit based clustering strategy.

The remainder of this paper is organized as follows. In Sect. 2, we describe the system model and present the proposed multi-armed bandit based clustering strategy. In Sect. 3, we evaluate the performance of the D2D-aided multicasting network. Section 4 validates the proposed multi-armed bandit based clustering strategy. Finally, the conclusions are drawn in Sect. 5.

2 Network Model

We consider a D2D-aided multicasting network, where the BSs and users are spatially distributed as homogeneous Poisson point processes (HPPPs) with density λ_b and λ_u, respectively. We assume that for each BS, the associated users are grouped into K clusters, where K follows Poisson distribution with mean give by ξ_c. In the process of mobile live video streams delivery, it is designed that the BSs first offload the multimedia traffics to the selected cluster heads, and then the cluster heads multicast the received data to other users in the respective clusters. To simply the analysis, it is assumed that the delivery of mobile live video streams from the BSs to the selected cluster heads have no transmission error.

The power gain $l(d)$ of the wireless channel over a distance of d is given by $l(d) = hd^{-\alpha}$, where h denotes the Rayleigh fading power coefficient with unit mean, and $\alpha \geq 2$ denotes the path-loss exponent. Further, for the successful reception of the multicasted stream at users via D2D communication, the SIR target is given by θ_u.

To simultaneously enhance the received SIR of users and reduce the processing delay induced by the D2D cluster heads searching, a novel multi-armed

bandit based clustering strategy is proposed, which is described in detail as follows.

Multi-armed Bandit Based Clustering Strategy: For the initialization of D2D clustering, in each cell, we first randomly select K users as the cluster heads for the K clusters. Then, the selected K cluster heads are scheduled in different time slots to broadcast the SRS signal with power P_u for D2D discovery. Upon receiving the SRS signals sent from the K cluster heads, a user selects the k-th cluster head as its associated cluster head if the respective SRS signal power is larger than others. With the initialized K clusters, the users in each cluster are then scheduled to broadcast the SRS signal with power P_u for the determination of the actual cluster head which maximize the average received signal power of mobile live video streaming. Without loss of generality, we denote D_k^i as the i-th user in the k-th cluster to send the SRS signal. Further, let $S_k^i(j)$ denote SRS signal (sent by D_k^i) power received at the j-th user in the k-th cluster. For the first iteration, upon receiving the SRS of D_k^1, each user is designed to upload the SRS signal strength information $S_k^1(j)$ to the BS. By assuming $S_k^1(j)$ are independent and σ-sub-Gaussian, the BS then builds a $(1-\delta)$-confidence interval $[\hat{\mu}_k^j(1) - C_k^j(1), \hat{\mu}_k^j(1) + C_k^j(1)]$ of the average received SRS signal strength for each users in the k-th cluster and finds the user with the largest upper bound of the confidence interval denoted by $\hat{D}_k$, where

$$\hat{\mu}_k^j(t) = \frac{1}{T_k^j(t)} \sum S_k^i(j) \tag{1}$$

with $T_k^j(t)$ denoting the number of uploaded SRS signal strength information by the j-th user in the k-th cluster upto time t, and[1]

$$C_k^j(t) = \sqrt{\frac{2\sigma^2 \log \frac{2}{\delta}}{T_k^j(t)}}. \tag{2}$$

Further, the BS randomly selects a user in $\widetilde{\mathcal{D}}_k$ to sent the SRS and require only $\hat{D}_k$ to upload the received SRS signal strength information, where $\widetilde{\mathcal{D}}_k$ denotes the set of users which have not broadcast the SRS signal in the k-th cluster. Then, the BS updates the $(1-\delta)$-confidence interval for the average received SRS strength. The process continues until the lower bound of the confidence interval of a user is larger than the upper bounds of the confidence interval of all the other users. Then, the respective user is selected as the new cluster head. With the updated cluster heads, users are grouped again based on the received SRS signals sent from the K cluster heads. By repeating the same procedure, the whole process ends if and only if the updated cluster heads remain the same.

In the next section, under the proposed multi-armed bandit based clustering strategy, we analyze the performance of the D2D-aided multicasting network.

[1] It is worth noting that given n users in the k-th cluster, if $T_k^j(t) \geq n-1$, $C_k^j(t) = 0$.

3 Network Performance

In this section, we analyze the performance of the proposed multi-armed bandit based clustering strategy. We first characterize the number of uploading times M of the received SRS signal strength per cell, which is provided in the following theorem.

Theorem 1. *For D2D-aided multicasting networks, under the proposed multi-armed bandit based clustering strategy, the uploading times M of the received SRS signal strength per cell is upper bounded by*

$$M \leq \sum_{i\in[n]} \left(\frac{24\sigma^2}{\Delta_i^2} \log n \wedge 2n \right). \tag{3}$$

We further characterize the coverage probability of the D2D-aided multicasting networks in the following theorem.

Theorem 2. *Under the proposed multi-armed bandit based clustering strategy, the coverage probability $\mathcal{C}$ of the D2D-aided multicasting network is given by*

$$\begin{aligned}
\mathcal{C}_n = & \int_0^{R_d} \frac{2\lambda_g^n \pi l \cdot e^{-\lambda_g^n \pi l^2}}{1 - e^{-\lambda_g^n \pi R_d^2}} \cdot \exp\left\{ -\frac{2\pi^2 \theta_u^{\frac{2}{\alpha}} l^2 \overline{\lambda}_g^n}{\alpha \sin\left(\frac{2\pi}{\alpha}\right)} \right\} \\
& \times \exp\left\{ -2\pi\lambda_g^n \int_l^\infty \frac{1}{1 + \frac{u^\alpha}{\theta_u l^\alpha}} u\, du \right\} \\
& \times \exp\left\{ \frac{2\pi}{\alpha} \overline{\lambda}_g^n \left(\frac{P_u}{\mathcal{I}_{th}} \right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha}\right) \right\} \\
& \times \exp\left\{ -2\pi\overline{\lambda}_g^n \int_0^\infty \frac{\frac{u^\alpha}{\theta_u l^\alpha}}{1 + \frac{u^\alpha}{\theta_u l^\alpha}} \cdot \frac{e^{-\frac{\theta_u \mathcal{I}_{th} l^\alpha}{P_u}}}{e^{\frac{\mathcal{I}_{th} u^\alpha}{P_u}}} u\, du \right\} dl,
\end{aligned} \tag{4}$$

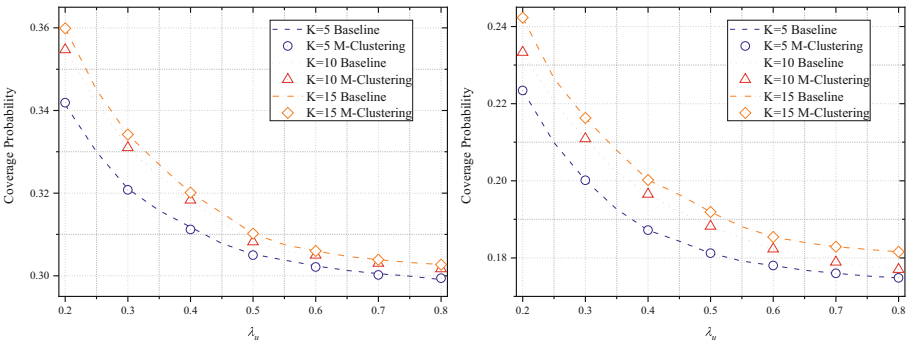
where $\overline{\lambda}_g^n = \lambda_g^a - \lambda_g^n$.

In the next section, under the proposed multi-armed bandit based clustering strategy, we analyze the performance of the D2D-aided multicasting network.

4 Numerical Results

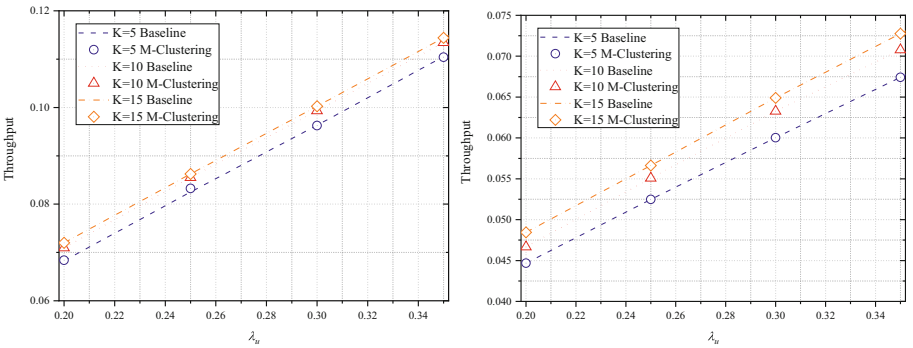
In this section, simulation results are presented to evaluate the effectiveness of the proposed multi-armed bandit based clustering strategy. For default setting, we have $\lambda_b = 0.001$, $\alpha = 4$, $K = 10$.

Figure 1 shows the coverage performance of the D2D-aided multicasting network, when $\lambda_g = 0.005$, and 0.01, respectively. It is observed that the proposed multi-armed bandit based clustering strategy has the same performance on the coverage probability as the baseline strategy, i.e., the naive clustering approach.



(a) $\theta_u = 0.5$ (b) $\theta_u = 0.1$

Fig. 1. Coverage probability versus the density of users λ_u.



(a) $\theta_u = 0.5$ (b) $\theta_u = 1$

Fig. 2. Spatial throughput versus the density of users λ_u.

It is also observed that the coverage probability of D2D-aided multicasting network is a decreasing function with respect to λ_u. Further, it is observed that the coverage probability of D2D-aided multicasting network is an increasing function with respect to the number of clusters.

Figure 2 shows the spatial throughput of the D2D-aided multicasting network, when $\lambda_g = 0.005$, and 0.01, respectively. Similarly as that in Fig. 1, it is observed that the proposed multi-armed bandit based clustering strategy has the same performance on the spatial throughput as the baseline strategy. Further, it is observed that the spatial throughput of D2D-aided multicasting network is an increasing function with respect to the number of clusters. However, different from Fig. 1 the spatial throughput of D2D-aided multicasting network is an increasing function with respect to λ_u.

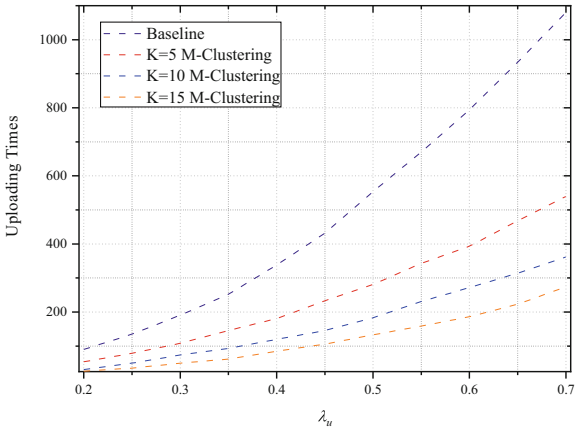


Fig. 3. Uploading times versus the density of users λ_u.

Figure 3 demonstrates the average number of uploaded times per cell to complete the clustering, where the baseline denotes the naive clustering policy. It is observed that the uploading times of SRS signal information strength decrease with the number of clusters.

5 Conclusion

In this paper, we proposed a novel multi-armed bandit based clustering strategy in D2D networks to enhance the overall performance. Particularly, with the proposed strategy, in each cell, the respective BS simply selects the D2D users with the largest average received sounding reference signals (SRSs) sent from other users as the cluster heads to maximize the multicast performance. Further, by applying the idea of multi-armed bandits, the multicasting latency induced by the process of clustering can be significantly reduced. We analyze the performance of the D2D-aided multicasting network. The effectiveness of the proposed multi-armed bandit based clustering strategy is demonstrated through extensive simulations.

References

1. Wang, C., Chen, Y., Wei, H., Liu, K.J.R.: Scalable video multicasting: a stochastic game approach with optimal pricing. IEEE Trans. Wirel. Commun. **14**(5), 2353–2367 (2015)
2. Christopoulos, D., Chatzinotas, S., Ottersten, B.: Multicast multigroup precoding and user scheduling for frame-based satellite communications. IEEE Trans. Wirel. Commun. **14**(9), 4695–4707 (2015)
3. Zhou, S., Ying, L.: On delay constrained multicast capacity of large-scale mobile ad hoc networks. IEEE Trans. Inf. Theory **61**(10), 5643–5655 (2015)

4. Li, H., Li, B., Tran, T.T., Sicker, D.C.: Transmission schemes for multicasting hard deadline constrained prioritized data in wireless multimedia streaming. IEEE Trans. Wirel. Commun. **15**(3), 1631–1641 (2016)
5. Sim, G.H., Widmer, J.: Finite horizon opportunistic multicast beamforming. IEEE Trans. Wirel. Commun. **16**(3), 1452–1465 (2017)
6. Zhang, Z., Ma, Z., Xiao, Y., Xiao, M., Karagiannidis, G.K., Fan, P.: Non-orthogonal multiple access for cooperative multicast millimeter wave wireless networks. IEEE J. Sel. Areas Commun. **35**(8), 1794–1808 (2017)
7. Bejerano, Y., et al.: DyMo: dynamic monitoring of large-scale LTE-multicast systems. IEEE/ACM Trans. Network. **27**(1), 258–271 (2019)
8. Xu, W., Cui, Y., Liu, Z.: Optimal multi-view video transmission in multiuser wireless networks by exploiting natural and view synthesis-enabled multicast opportunities. IEEE Trans. Commun. **68**(3), 1494–1507 (2020)
9. Araniti, G., Rinaldi, F., Scopelliti, P., Molinaro, A., Iera, A.: A dynamic MBSFN area formation algorithm for multicast service delivery in 5G NR networks. IEEE Trans. Wirel. Commun. **19**(2), 808–821 (2020)
10. Chen, Y., He, S., Hou, F., Shi, Z., Chen, J.: An efficient incentive mechanism for device-to-device multicast communication in cellular networks. IEEE Trans. Wirel. Communi. **17**(12), 7922–7935 (2018)
11. Niu, Y., Liu, Y., Li, Y., Chen, X., Zhong, Z., Han, Z.: Device-to-device communications enabled energy efficient multicast scheduling in mmWave small cells. IEEE Trans. Commun. **66**(3), 1093–1109 (2018)
12. Santana, T.V., Combes, R., Kobayashi, M.: Performance analysis of device-to-device aided multicasting in general network topologies. IEEE Trans. Commun. **68**(1), 137–149 (2020)
13. Bagaria, V., Kamath, G.M., Ntranos, V., Zhang, M.J., Tse, D.: Medoids in almost linear time via multi-armed bandits. arXiv preprint arXiv:1711.00817, November 2017
14. Zhao, Z., Feng, C., Yang, H.H., Luo, X.: Federated-learning-enabled intelligent fog radio access networks: fundamental theory, key techniques, and future trends. IEEE Wirel. Commun. **27**(2), 22–28 (2020)
15. Jiang, C., Zhang, H., Ren, Y., Han, Z., Chen, K., Hanzo, L.: Machine learning paradigms for next-generation wireless networks. IEEE Wirel. Commun. **24**(2), 98–105 (2017)
16. Zhu, G., Liu, D., Du, Y., You, C., Zhang, J., Huang, K.: Toward an intelligent edge: wireless communication meets machine learning. IEEE Commun. Mag. **58**(1), 19–25 (2020)

Collaborative Sparse Unmixing Network Based on Deep Algorithm Unrolling

Yuhan Zheng(✉), Mengyue Chen, Donghao Shi, and Guanglong Ren

College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

zheng_yuhan@nuaa.edu.cn

Abstract. Hyperspectral unmixing, as an important preprocessing technique for hyperspectral processing, has attracted much attention from researchers in recent years. In this paper, an efficient sparse unmixing network is proposed for the sparse property of mixed pixels' abundance matrix in hyperspectral images. The method is based on the CLSUnSAL algorithm, and transfers the idea of alternating iterations to the neural network, namely deep algorithm unrolling. Specifically, we stack two update blocks for variables update, in which we set multiple learnable parameters to increase the flexibility of the network ensuring better unmixing results. This scheme combines the interpretability of traditional unmixing with the efficiency of deep learning. The experimental results show that the network can obtain better unmixing accuracy than other traditional methods based on sparse unmixing and is suitable for hyperspectral data with high signal-to-noise ratio.

Keywords: Hyperspectral unmixing · ADMM · Algorithm unrolling · Deep learning

1 Introduction

Hyperspectral remote sensing records various substances observed in the field of view with complete spectral curves. The most important advantage of hyperspectral images (HSIs) is that they are characterized by high spectral resolution. However, on account of the complex distribution of ground materials and the low spatial resolution of remote sensing instruments, pixels in HSIs generally contain a variety of materials. The purpose of hyperspectral unmixing (HU) technology is to extract the spectral features of various basic substances that constitute the pixels, and obtain the proportion of these substances.

In the past decade, traditional hyperspectral unmixing algorithms have achieved great success. Some scholars introduced the idea of sparse representation into the HU model, using the known spectral library as the endmember set. Optimization problems based on sparse unmixing can be transformed into convex optimization problems, among which the most representative one is the SUnSAL [1] algorithm. When solving the optimization problem, the non-negativity constraint (ANC) and the sum-to-one constraint (ASC) of abundance coefficients are added to the optimization problem, and we utilize the alternating direction method of multipliers (ADMM) to decompose the optimization

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 206–212, 2023.
https://doi.org/10.1007/978-981-99-1260-5_26

problem into several simple sub-optimization problems. On this basis, scholars have successively made various improvements to SUnSAL, such as CLSUnSAL [2], and SUnSAL-TV [3].

With the advancement of artificial intelligence, deep learning has gradually begun to be applied in hyperspectral unmixing. The network of autoencoders is suitable for solving the unmixing task. Specifically, the unmixing task is divided into two parts, the first step is to estimate the abundance through the encoder, and then input the result into the decoder to get the reconstruction of the original image. Based on such theories, unmixing networks of autoencoder architectures such as uDAS [4] and DAEN [5] have been successively proposed. In addition, using the convolutional layer instead of the fully connected layer can extract rich spatial information. Therefore, spectral-spatial unmixing network had been proposed, such as CNNAEU [6]. However, there is a fatal obstacle with deep learning, that is, the lack of interpretability, and it is difficult to ensure that the final results conform to the physical meaning. Therefore, some scholars put forward the idea of using the traditional model to drive the network, such as unfolding algorithm into neural network [7].

In this paper, inspired by algorithm unrolling, we propose an unmixing network based on CLSUnSAL to solve the optimization problem. Benefit from the function of updating variables through gradient descent in deep learning, the network proposed in this paper only stack two iterative blocks, which can obtain good unmixing results while reducing the computational consumption.

2 The Unmixing Model

Linear mixture model (LMM) is the most commonly used spectral mixture model in HSIs sparse unmixing model. The LMM assumes that the spectrum in a mixed pixel is the weighted sum of the endmembers contained in the pixel. The expression of unmixing model based on the LMM is as follows:

$$Y = AX + N \tag{1}$$

where $Y \in \mathbf{R}^{L\times N}$ represents the observed values of N pixels in L bands, $A \in \mathbf{R}^{L\times M}$ represents the spectral library, which incorporates the reflection values of M different endmembers in L bands, $X \in \mathbf{R}^{M\times L}$ is the abundance coefficient matrix, and N is the additive noise matrix.

Based on this model, an optimization problem is proposed:

$$\min_{X} \|Y - AX\|_F^2 + \lambda \sum\nolimits_i \|X_i\|_2 \qquad s.t.\ X \geq 0 \tag{2}$$

in which X_i represents the i-th column of the abundance matrix, and $\sum_i \|X_i\|_2$ is an $L_{2,1}$ norm, which is used to normalize the sparsity of the rows of the abundance matrix. To solve this optimization problem, the CLSUnSAL algorithm is proposed. Expression (2) can be adjusted to:

$$\min_{U,V_1,V_2,V_3} \frac{1}{2}\|Y - V_1\|_F^2 + \lambda\|V_2\|_{2,1} + l_{\{R+\}}(V_3) \quad s.t.\ V_1 = AU, V_2 = U, V_3 = U \tag{3}$$

Then, the Lagrange multiplier is introduced, and the constraint term is added to the objective function as a quadratic penalty function to obtain the augmented Lagrange function:

$$L(U, V, D) = \frac{1}{2}\|V_1 - Y\|_F^2 + \lambda\|V_2\|_{2,1} + l_{\{R+\}}(V_3) + \frac{\mu}{2}\|AU - V_1 - D_1\|_F^2 + \frac{\mu}{2}\|U - V_2 - D_2\|_F^2 + \frac{\mu}{2}\|U - V_3 - D_3\|_F^2 \tag{4}$$

where μ is the penalty factor. When the penalty factor is selected appropriately, the optimal solution can be better approximated.

According to ADMM, we first solve the parameter U, and then use U to solve each single regular term and update corresponding variable in turn [2].

3 The Proposed Unmixing Network

In this part, we are going to unfold algorithm into neural network layers. For each different variable update, we use the different network layer.

- U update:

$$U^{(k+1)} = B^{T(k)}(V_1^{(k)} + D_1^{(k)}) + C^{T(k)}(V_2^{(k)} + D_2^{(k)} + V_3^{(k)} + D_3^{(k)}) \tag{5}$$

where

$$B^{T(0)} = (A^T A + 2I)^{-1}A^T \quad C^{T(0)} = (A^T A + 2I)^{-1} \tag{6}$$

We can think of (5) as the addition of the two convolutional layers' outputs and use (6) to initialize the weights of the convolutional layers. The network structure is shown in the Fig. 1(a).

- V_1 update:

$$V_1^{(k+1)} = \alpha^{(k)}Y + \beta^{(k)}(AU^{(k+1)} - D_1^{(k)}) \tag{7}$$

where

$$\alpha^{(0)} = \frac{1}{1+\mu} \qquad \beta^{(0)} = \frac{\mu}{1+\mu} \tag{8}$$

Unlike the update of U, here α and β are scalars, so just set μ as a learnable parameter. The network structure is shown in the Fig. 1(b).

- V_2 update:

$$V_{2,r}^{(k+1)} = \frac{M_r^{(k)}}{M_r^{(k)} + \rho^{(k)}} \cdot U^{(k)} \tag{9}$$

where

$$\rho^{(0)} = \frac{\lambda}{\mu} \qquad M_r^{(k)} = \max(\left\|U_r^{(k)} - D_{2,r}^{(k)}\right\|_F - \rho^{(k)}, 0) \tag{10}$$

Set ρ as a learnable parameter here. The max expression is implemented using the ReLU activation function, and the network structure is shown in the Fig. 1(c).

- V_3 update:

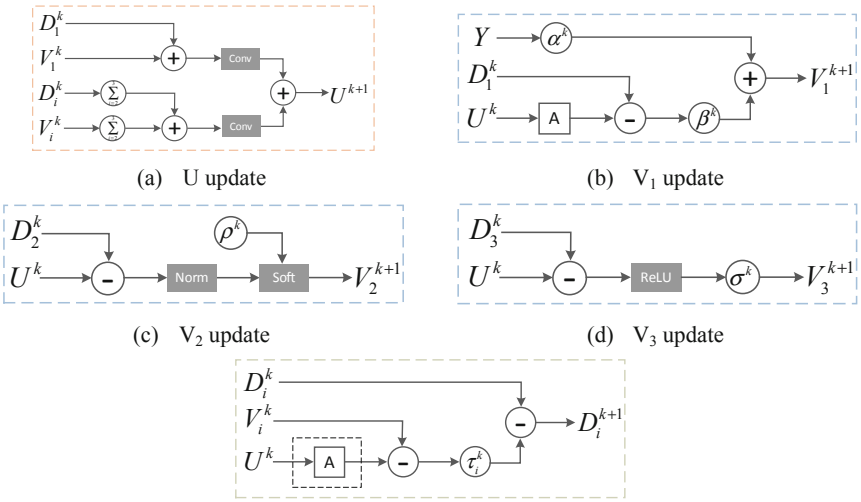
$$V_3^{(k+1)} = \sigma^{(k)} \mathrm{ReLU}(U^{(k)} - D_3^{(k)}) \tag{11}$$

In order to make the network more flexible, we introduce a learnable parameter σ to multiply the result obtained by ReLU. The Fig. 1(d) shows the network structure.

- D update:

$$\begin{aligned} D_1^{(k+1)} &= D_1^{(k)} - \tau_1^{(k)}(AU^{(k+1)} - V_1^{(k+1)}) \\ D_i^{(k+1)} &= D_i^{(k)} - \tau_i^{(k)}(U^{(k+1)} - V_i^{(k+1)}) \quad i = 2, 3 \end{aligned} \tag{12}$$

The update of D_1 is slightly different from D_2 and D_1. τ_i is the introduced learnable parameter. The network structure is shown in the Fig. 1(e).



(a) U update (b) V_1 update

(c) V_2 update (d) V_3 update

(e) D update

Fig. 1. Update network for each variable

After designing the variables update layers, we can stack these network layers to build a complete algorithm unfolding network, and we select V_3 whose result satisfies the non-negative abundance constraint as the output. Figure 2 shows the updates of U, V, and D form an iterative block. Weighing the unmixing result and the amount of computation, we finally choose to stack two iterative blocks and perform the third iteration on V_3 alone. In order to satisfy ASC, we perform the following calculation on V_3, and the output is used as the result of abundance estimation:

$$\hat{X}_{(\cdot,j)} = \frac{V_{3,(\cdot,j)}}{\sum_i V_{3,(i,j)}} \tag{13}$$

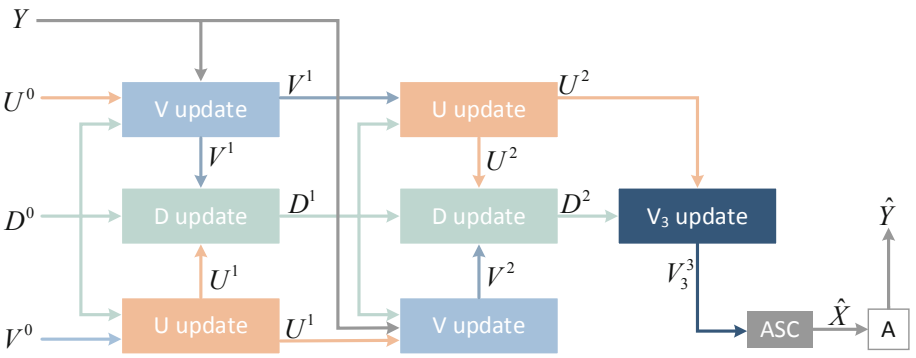


Fig. 2. The proposed unmixing network structure

Because implementing unsupervised learning, we need to reconstruct the original image at the end. The reconstructed image $\hat{Y}$ can be obtained by multiplying $\hat{X}$ and A. We use the cross-entropy function, which measures the similarity of the data, as the loss function for the entire network:

$$L = -\frac{1}{L}\frac{1}{N}\sum_{i=0}^{L}\sum_{j=0}^{N}[Y_{(i,j)}\log\hat{Y}_{(i,j)} + (1 - Y_{(i,j)})\log(1 - \hat{Y}_{(i,j)})] \tag{14}$$

4 Experimental Results

In order to evaluate the performance of our network, we employ a synthetic hyperspectral data with 9 endmembers, and four different methods based on sparse unmixing were chosen for comparison, including SMP, SUnSAL, SUnSAL-TV and CLSUnSAL. In the simulation data experiments, we use the signal reconstruction error as the evaluation standard of the algorithm performance, which is expressed as: $E_{SRE} = E(\|X\|_F^2) \Big/ E(\left\|X - \hat{X}\right\|_F^2)$, in which X represents the original abundance coefficient matrix, and $\hat{X}$ represents the reconstructed abundance coefficient matrix.

Table 1. SRE(dB) values of abundance under 10 dB, 20 dB, 30 dB Gaussian noise

SNR (dB)	SUNSAL	SUNSAL_TV	SMP	CLSUNSAL	Proposed
10 dB	2.8701	**5.9708**	4.2108	3.0319	5.519
20 dB	8.2555	12.6439	**17.9381**	8.4423	15.2833
30 dB	15.7865	20.2733	22.7212	16.0853	**28.5972**

From Table 1, we observe that our proposed network has obtained better unmixing results under different Gaussian noise conditions. Especially at 30 dB SNR, the SRE of

our proposed method is much higher than other sparse unmixing methods. At 10 dB and 20 dB SNR, our proposed method also achieves the second-best results, and is not far behind the first-place method. We note that the results of our method are much better than the CLSUnSAL method, which illustrates the feasibility of the algorithm unfolding and also demonstrates the powerful adaptive learning capabilities of deep networks.

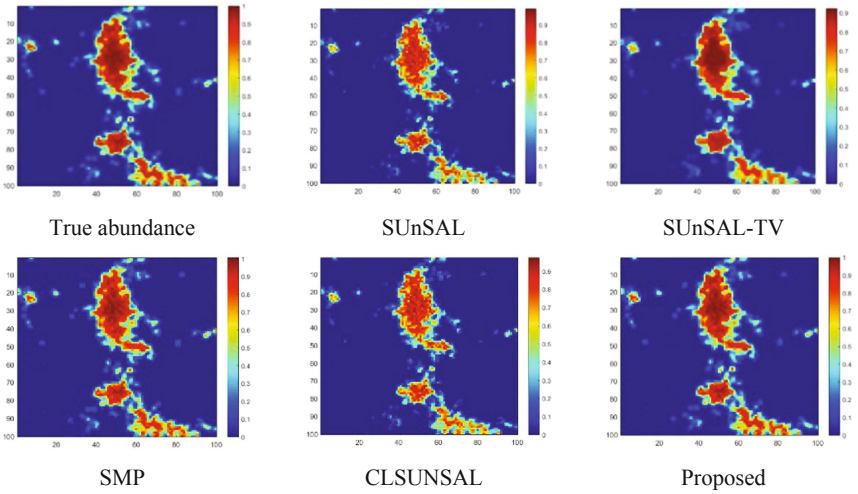


Fig. 3. Abundance map of 9^{th} endmember under 30 dB Gaussian noise

The Fig. 3 shows the reconstructed abundance results of the five methods under 30 dB SNR. We can intuitively see that the abundance result output by our proposed network is closer to the real abundance map than other methods. With respect to the experimental results, we explain that our network not only ensures the physical meaning of the results, but also leverages the powerful feature extraction function of deep learning. The network based on algorithm unrolling is an interpretable neural network. Thanks to the back propagation of the network, our parameters can be updated adaptively, while the parameters in the traditional methods are fixed and lack of flexibility.

5 Conclusion

Based on the algorithm of CLSUnSAL, this paper unfolds the traditional algorithm into the neural network. On the one hand, this method has the interpretability of traditional algorithm, and on the other hand, it takes full advantage of the powerful feature extraction ability of deep learning. In comparative experiments, we also demonstrate that the proposed method has a strong unmixing ability.

References

1. Bioucas-Dias, J., Figueiredo, M.: Alternating direction algorithms for constrained sparse regression: application to hyperspectral unmixing. In: 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp. 1–4 (2010)

2. Iordache, M., Bioucas-Dias, J., Plaza, A.: Collaborative sparse unmixing of hyperspectral data. In: 2012 IEEE International Geoscience and Remote Sensing Symposium, pp. 7488–7491 (2012)
3. Iordache, M., Bioucas-Dias, J., Plaza, A.: Total variation spatial regularization for sparse hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. **50**(11), 4484–4502 (2012)
4. Su, Y., Li, J., Plaza, A., et al.: DAEN: deep autoencoder networks for hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. **57**(7), 4309–4321 (2019)
5. Qu, Y., Qi, H.: uDAS: An untied denoising autoencoder with sparsity for spectral unmixing. IEEE Trans. Geosci. Remote Sens. **57**(3), 1698–1712 (2019)
6. Palsson, B., Ulfarsson, M., Sveinsson, J.: Convolutional autoencoder for spectral-spatial hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. **59**(1), 535–549 (2021)
7. Monga, V., Li, Y., Eldar, Y.: Algorithm unrolling: interpretable, efficient deep learning for signal and image processing. IEEE Sig. Process. Mag. **38**(2), 18–44 (2021)

Channel-Stationary Entropy Model for Multispectral Image Compression

Guanglong Ren(✉), Tongbo Cao, Fanqiang Kong, and Jiahui Tang

College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
rgl19981212@nuaa.edu.cn

Abstract. Multispectral image have a large number of complex features. The existing network models have good compression performance for multispectral image, but the encoding and decoding time is long. In order to extract latent statistical features to guide arithmetic coding and save compression time as much as possible, a multispectral image compression framework based on channel-stationary is proposed. The multi-spectral image first go through the feature extraction network to get the feature maps, they can be down-sampling and compressed and then input to the arithmetic encoder and entropy priori module respectively. The weight features obtained by the entropy prior module feed the aid into the image features of the autoregressive module for entropy coding. The auto-regressive module can realize decoding of images, thus saving the time required for encoding and decoding. In addition, the balance between image reconstruction quality and bit rate is achieved by using distortion optimization technology. Then the image is reconstructed by inverse quantization, up-sampling and deconvolution module. The experimental results show that the method is better than JPEG2000 under the condition of similar bit rate, the compression time is shorter, and the network time complexity of the method is higher.

Keywords: Multispectral image compression · Entropy coding · codec time · Rate distortion optimization

1 Introduction

Multispectral image is playing an increasingly important role in aerospace, military and other fields because of its rich spectral information and texture details of ground targets. However, high dimension data makes the multispectral image occupy a lot of memory.

In recent years, Ballé et al. [1] proposed a visible image compression technology based on neural network, and achieved good results. In 2018, this paper [2] further proposed entropy prior image compression technology based on variational autoencoder architecture. Later, Minnen et al. [3] proposed a context technology applied in the encoding process, which achieved better performance but doubled the time of encoding and decoding.

To address these issues, such as high time complexity, the method is proposed. This neural network utilizing channel-stationary consists of two parts: a feature extraction

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 213–219, 2023.
https://doi.org/10.1007/978-981-99-1260-5_27

network and an entropy modeling network. The feature extraction network is used for spatial feature extraction of multispectral image, abstracting high-dimensional features into feature maps. Entropy modeling estimates the point probability of the feature map for subsequent guidance of image entropy coding. In terms of coding, the residual block [7] is used to extract the spatial features of multispectral image. The output features of the entropy encoder side and the entropy modeler side are input to the stationary grouped encoder-decoder network. The output of Entropy modeling network is used to encode the grouped features in the arithmetic encoder according to the Gaussian model.

2 Related Work

Since multispectral image reflect the imaging characteristics of the same feature in different spectra, the imaging targets between different spectral bands are essentially the same, and there is no spatial position shift and time dimension change. Therefore, for 3D multispectral image, we can still use the principle of 2D image encoding for multi-spectral image encoding.

Learning-based image compression involves many fields of research, and recently, the entropy bottleneck-based autoencoder structure has achieved better performance in image compression.

For hyperprior approach in image compression, Ballé, et al. [1] proposed rate-distortion technology on the basis of variational autoencoder, and further used variational autoencoder to construct prior entropy model to guide image coding in [2]. Then, Minnen et al.[3] further improved the reverse process by adding an autoregressive module constructed by using context-adaptive models. Mentzer et al. [4] proposed a technique to navigate the rate-distortion trade-off for an image compression auto-encoder. The main idea is to directly model the entropy of the latent representation by using a context model. Klopp et al. [5] introduced a neural network based image coding model that utilizes a code-space predictor to reduce code length by modelling dependencies within the code. J. Lee et al.[6] proposed a contextual model for end-to-end optimized image compression to more accurately estimate the distribution of each latent representation, enabling an adaptive entropy model.

In these context- adaptive models, masked convolution takes into account local contextual information, which is highly conditioned in the causal context, and is exploited in series during the coding and decoding process. This not only takes up a lot of computer resources, but also is not conducive to multi-GPU parallel training. Moreover, the existence of the mask will greatly increase the real-time performance of image compression and decompression, which is not conducive to practical applications.

To address this problem, current solutions include a multi-level prior scheme [8] and a channel conditioning (CC) scheme [9]. Among them, channel-conditioning scheme is based on stationary processes. The channel conditioning model divides the latent tensor into N equal-sized Stationary along the channel dimension. Stationary slices take the point probability distribution from the entropy model and determine the entropy parameter of each slice previously decoded. Therefore, the CC model can be interpreted as an autoregression along the channel dimension, which correlates with the extracted spatial features to achieve sufficient abstraction of multispectral features for efficient

encoding. Because the idea of autoregression is applied on the basis of smooth process theory in information theory. This enables an autoregressive process within the channel dimension, enabling efficient encoding. In summary, we propose a multispectral image compression framework based on channel-Stationary.

3 Proposed Method

Based on the model in [6], we propose an image compression framework based on autoregressive coding model. The overview of the proposed network is shown in Fig. 1. The details of the forward network and the backward network are shown in Fig. 2. The forward network consists of four feature extraction modules: residual block, downsampling block and GDN block. After the multispectral image is input to the forward network, the latent feature $\hat{y}$ is output. On the one hand, $\hat{y}$ is sent into the entropy prior module for image entropy estimation; on the other hand, $\hat{y}$ can be divided into multiple parts equally along the channel dimension, and the mean and variance can be estimated through the autoregression module respectively. $\hat{y}$ and the mean and variance of the previous layer's predictions. Lastly, each part of the autoregressive code stream is integrated through the network, and the input is decoded to the network.

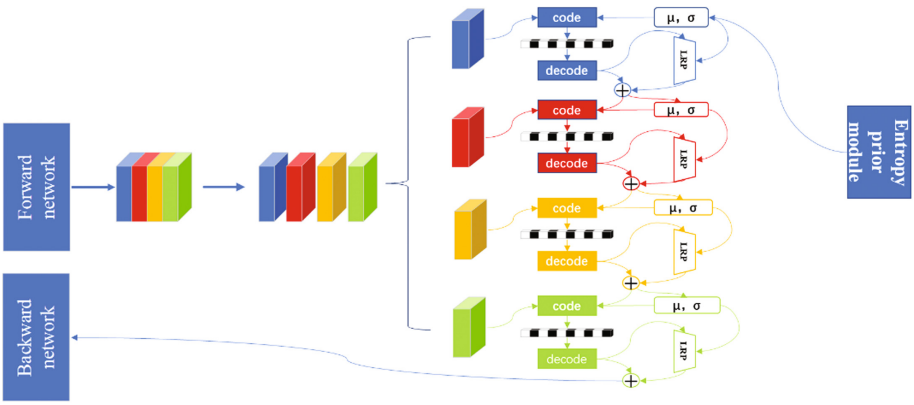


Fig. 1. The overview of the proposed network

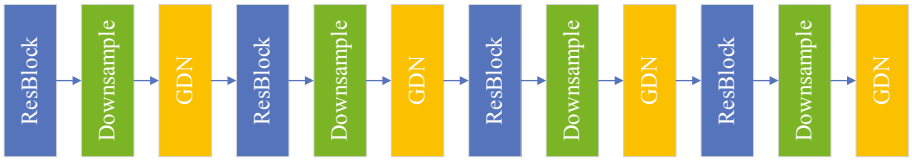


Fig. 2. The details of forward and backward networks

However, the image compression code rate directly affects the image reconstruction quality, and the two are contradictory. It is critical to optimize distortion and bitrate and

strike a balance between them. Ballé used a rate-distortion optimization (RDO) technique based on the Lagrangian multiplier method in the paper [1] to achieve a balance between bit rate and image quality.

3.1 Network Core Module

The backbone consists of a bunch of residual blocks and subsampling modules. Residual module has strong feature extraction ability to build deep network to prevent gradient disappearance. The residual module is shown as Fig. 3. Given that image compression is an image reconstruction task, using a depth network facilitates image reconstruction. Due to the computational complexity of the estimation of the entropy model, down-sampling can reduce the size of the feature map and tighten the distribution of characteristics to achieve more efficient estimation of entropy.

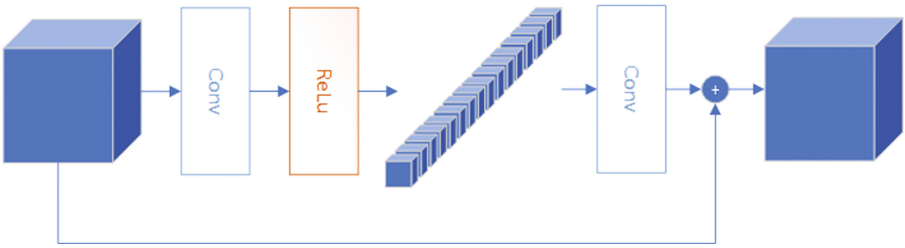


Fig. 3. The residual block used in our method

Generalized Divider Normalization (GDN) is a normalization layer that is beneficial for image reconstruction tasks. In Convolutional Neural Networks, Batch Normalization (BN) [10] is a common intermediate processing layer that processes an image into a feature map with mean 0 and standard deviation 1, which is close to Gaussian distribution. In addition, it can speed up training. The characteristic of the BN layer is that the mean and standard deviation of each batch of feature maps are different, so this is equivalent to adding noise to the feature maps. Therefore, the BN layer gives the model a stronger generalization ability. However, it is not friendly to generative models such as image super-resolution reconstruction, image formation, image denoising, and image compression, because the generated images are required to be as clear as possible, and noise should not be introduced. Therefore, the BN layer should not be used in these application scenarios. GDN is continuously differentiable and This paper only use form for this transformation is:

$$y_i = \frac{x_i}{(\beta^2 + \sum_j \gamma_j x_j^2)^{\frac{1}{2}}} \tag{1}$$

where $\theta = \{\alpha, \beta, \gamma\}$ are parameters. Where x_i is the input feature map of layer i, β_i and γ_i are the parameters to be learned. Are the parameters to be learned.

3.2 Rate Distortion Optimization

Rate-distortion theory is based on the basic viewpoints and methods of information theory to study the theory of data compression, also known as limited source distortion coding theory. It was proposed to solve the problem of describing the minimum bit rate within a certain distortion limit. That is, rate-distortion theory describes an optimized form of the optimal solution for compression ratio and image quality.

Rate-distortion optimization is a constrained optimization problem because of the contradiction between rate and distortion. The Lagrange multiplier method is an efficient algorithm for solving constrained optimization problems. Based on this idea, Ballé et al. used a rate-distortion optimization function based on Lagrange multipliers. A loss function of the weighted sum of bitrate and distortion is constructed to guide network learning and training.

Forasmuch as the large amount of multispectral image data, it is very important to discuss the reconstructed image quality based on low bit rate compression task. Therefore, the rate distortion loss function used in this article is as follows:

$$L = D + \lambda R \tag{2}$$

where D is the degree of distortion expressed by the MSE λ is a hyperparameter, and R is the bit rate. R can be expressed as crossentropy:

$$R = E_{x \sim p_x}[-\log_2 p_{\hat{y}}(Q(g_a(x; \phi_g)))] \tag{3}$$

where Q stands for the quantification function and $p_{\hat{y}}$ stands for the entropy model.

3.3 Implementation Details

In the proposed network, we first use residual blocks to extract image functionality, then we use convolution to reduce the image and adjust the network channel structure. All residual blocks are implemented with a convolution kernel of 3×3 size, and the number of channels is set to 192. The convolution block used for down-sampling and up-sampling has a convolution kernel size of 5×5 and a step size of 2, so down-sampling can be realized and more features can be extracted under the condition of ensuring a large receptive field.

In terms of learning rate setting, we adaptively adjust the learning rate by comparing the size of the model loss, and the initial learning rate is set to 0.00001.

In the design of the rate-distortion optimization parameter λ, we control the bit rate by changing the λ. When the bit rate is about 0.5, the λ is set to 0.001, when the bit rate is about 0.3, it is set to 0.0001, and when the bit rate is about 0.1, it is set to 0.00001.

3.4 Datasets

The multispectral image dataset used in this paper comes from the WorldView-3 satellite (https://resources.maxar.com/) with a spatial resolution of 0.3–0.4 m. In terms of the number of datasets, 8700 images of size 128×128 are randomly selected for training, and 14 images of size 512×512 are selected for testing, and the ratio of training set

to test set is about 600:1. The regions selected for the dataset include Rio de Janeiro, Brazil, Stockholm, Sweden, Tripoli, Libya, and Washington, D.C., USA. It is important to point out that in order to guarantee the validity of the dataset, there are no identical images between the training and test sets. All test images are representative and contain various scenes at different scales.

4 Results

4.1 Quality Criteria

In the field of image compression, the most widely used quality criteria is peak signal-to-noise ratio (PSNR). PSNR is obtained by normalizing the mean square error (MSE) of the image before and after compression. The PSNR is positively correlated with the bit rate, that is, the higher the bit rate, the higher the PSNR. We use the JPEG2000 algorithm and the context algorithm proposed by Minnen et al. to compare with the proposed algorithm.

4.2 Experimental Result Analysis

The experimental data set we use are 8-band multispectral image. PSNR is used to evaluate the quality of the restored image. In the mean time, the PSNR is positively correlated with the image quality. Since the bit rate is an important variable that needs to be controlled in the experiment, we first obtain the compression results under different bit rates by controlling the rate-distortion parameters. Then use these bit rates to control JPEG2000 for image compression. For the convenience of comparison, we obtained three representative experimental results, which are shown in Table 1.

Table 1. Average PSNR of two methods at different bit rate

	Rate	PSNR	Rate	PSNR	Rate	PSNR
JPEG2000	0.108	31.85	0.355	36.22	0.535	32.56
Context model		31.96		39.31		42.01
Proposed		**37.42**		**40.25**		**43.07**

As shown in Table 1, our approach is 6 dB better than JPEG2000 on average. In the case of increasing bit rate, the proposed method did not lose advantage. Meanwhile, the parallel technology used by the network reduces the image compression time to about 0.2 s and the decompression time to about 0.05 s. The compression technology using the context method has an encoding time of 16.7 s and a decoding time of 20.7 s. The time complexity of the proposed method is greatly reduced.

5 Conclusion

In this paper, we propose a channel-stationary multi-spectral image compression framework, which can achieve higher performance while ensuring shorter compression and decompression time. The experimental results show that the comprehensive performance of the proposed method is better than that of JPEG2000, which provides the research direction of channel stabilization technology.

References

1. Ballé, J., Laparra, V., Simoncelli, E.P.: End-to-end optimized image compression. In: 5th International Conference on Learning Representations, ICLR 2017 (2017)
2. Ballé, J., Minnen, D., Singh, S., Hwang, S.J., Johnston, N.: Variational image compression with a scale hyperprior (2018)
3. Minnen, D., Ballé, J., Toderici, G.D.: Joint autoregressive and hierarchical priors for learned image compression. In: Advances in Neural Information Processing Systems, vol. 31 (2018)
4. Mentzer, F., Agustsson, E., Tschannen, M., Timofte, R., Gool, L.V.: Conditional probability models for deep image compression. In: 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
5. Klopp, J.P., Wang, Y.-C.F., Chien, S.-Y., Chen, L.-G.: Learning a code-space predictor by exploiting intra-image-dependencies. In: British Machine Vision Conference (2018)
6. Lee, J., Cho, S., Beack, S.-K.: Context-adaptive entropy model for end-to-end optimized image compression. In: International Conference on Learning Representations (ICLR) (2019)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
8. Qian, Y., et al.: Learning accurate entropy model with global reference for image compression. In: International Conference on Learning Representations, September 2020
9. Minnen, D., Singh, S.: Channel-wise autoregressive entropy models for learned image compression. In: 2020 IEEE International Conference on Image Processing (ICIP), pp. 3339–3343. IEEE, October 2020
10. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International Conference on Machine Learning, pp. 448–456. PMLR, June 2015
11. Klopp, J.P., Wang, Y.-C.F., Chien, S.-Y., Chen, L.-G.: Learning a code-space predictor by exploiting intra-image-dependencies. In: British Machine Vision Conference (2018)

IOWA-Based Dynamic Trusted Assessment Model for System Operational State Behavior

Haoxiang Huang[1,2], Jianbiao Zhang[1,2(✉)], and Yanhui Liu[1,2]

[1] Faculty of Information Technology, Beijing University of Technology, Beijing, China
zjb@bjut.edu.cn
[2] Beijing Key Laboratory of Trusted Computing, Beijing, China

Abstract. In the era of "digital intelligence", the security of critical information systems is vitally important. We propose a trusted assessment model to guarantee the trusted of the system operation process, which is a three-dimensional trustworthiness assessment model. The model comprehensively considered entity trustworthiness, access control compliance trustworthiness, and the trusted of atomic behavior. Subsequently, the Induced Ordered Weighted Averaging (IOWA) is introduced to improve the accuracy and flexibility of the weight setting of the different metric values.

Keyword: critical information systems · dynamic · trusted assessment · system operation behavior

1 Introduction

Critical information systems, such as cloud computing and the Internet of Things (IoT) have been deeply applied in various fields of society with the rapid advancement of the digitalization process of society. However, as critical information systems shift to an open, dynamic, and shared service model, making it necessary to carry more complex and variable massive data processing requirements, they face more security vulnerabilities [1] and are more vulnerable to malicious attacks. In June 2022, China's National Computer Virus Emergency Response Center released a research report [2] that the United States used the "Acid Fox Platform" to continue Trojan horse attacks on hundreds of important information systems in China, and that the importance and urgency of trusted protection of critical information systems, as the core foundation for building a sovereign space in the cyber domain, are increasing. As early as 2011, the U.S. envisioned in [3] that strategic aspirations could be achieved by attacking other countries' critical infrastructure. Unfortunately, the above-mentioned scenario has been realized in the series of cyber attacks launched by Anonymous in 2022 on Russia's critical infrastructure, including transportation and defense systems, which marked the extension of the purpose of cyber attacks from profit acquisition to political appeals. Cyber attacks against critical information systems have grown to be a major threat factor that seriously affects national public security.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 220–227, 2023.
https://doi.org/10.1007/978-981-99-1260-5_28

Trusted computing, as a new technology, is widely considered can achieve information security technology from "soft security technology" to "hard security technology". Accordingly, there is a broad consensus based on active immune trusted computing to guarantee the trusted of the operation process of critical information systems. However, the current research on dynamic metrics is mainly from the perspective of engineering implementation [4, 5], and the objectivity and accuracy of trusted evaluation results are doubtful. In addition, theoretical models for dynamic metrics are only based on the trust perspective, which is not only inflexible but also inconsistent with the trustworthy nature of "entities behaving as expected". Generally, the challenges in building dynamic metric models are as follows: (1) The formulation of metric strategies and the setting of metric weights are affected by subjective experience, subjective decisions, limited rationality, lacking rigorous data analysis or theoretical support, and cannot effectively reflect the complexity and dynamism of changes in the system's operating state trust; (2) The existing behavior-based trust assessment models lack consideration of behavior contextual semantics, and also ignore the issue that the legality of the behavior of the same entity may change at different moments during the operation of the system, making the accuracy of trust assessment doubtful. (3) Trust is a dynamically changing value with a certain trend. Accordingly, neither a single metric nor the average value of multiple metrics can accurately portray the trend of trust changes in the operational system.

To tackle the aforementioned drawbacks, we propose a dynamic trusted metric model IB-DTASOS. The contributions can be summarized as follows:

- To address the dynamics and uncertainty of the operation process of complex systems, we propose a behavior-based dynamic metric model that integrates entity trustworthiness and access control compliance trustworthiness with system invocation as atomic behavior to establish a three-dimensional trustworthiness assessment system;
- Prediction of trusted expectations of system behavior based on Bayesian decision theory, combined with access control and entity trustworthiness for trusted assessment of behavioral context, improving the trustworthiness and accuracy of the metric model;
- The integration of the Induced Ordered Weighted Averaging (IOWA) operator to dynamically assign weights to the plausible expectations of system behavior at different moments ensures that the metric model conforms to the law of dynamic decay of trust over time, which improves the scientificity of the metric model.

2 Related Work

In recent years, some researchers combined Markov chains to build the dynamic metric model [7, 8], and hierarchically analyze and quantify the dynamic trustworthiness of the system. Li et al. [9] constructed a dynamic trust prediction model based on historical evidence window, direct trust tree (DTT) mechanism [10], and IOWA operator theory [11], which improves the problem of insufficient dynamic capability of traditional trust prediction models and weakens the defect that setting trust degree weights with a high subjectivity. However, the above-mentioned schemes [7–9] do not model the behavior and also are not able to strictly satisfy the trustworthy nature. Liu et al. [12] assessed the trust state comprehensively by measuring the physical environment and behavior of

nodes in the IoT environment, which can portray the dynamic change trend of trust more accurately from a multidimensional perspective with good dynamicity and adaptability. However, this scheme not only lacks consideration for the behavioral complexity issues, but the trust weight setting is more subjective. In [13, 14], by introducing Bayes theory, the posterior probability is used to continuously modify the prior probability for trust assessment of system operation state behavior, which has good dynamic adaptability and security. However, the described work [13, 14] lacks trustworthiness support and does not introduce the concept of trustworthiness. Zhang et al. [4] pointed out that the current dynamic metrics scheme lacks theoretical support of mathematical models and has low real-time metrics, and presented a software dynamic metrics model based on interference-free theory, which is a typical combination of theory and engineering. Huang et al. [15] proposed a dynamic metric model for the virtual machine startup process based on interference-free theory, which improves the deficiencies of insufficient dynamic protection of the virtual machine startup process and lack of unintended interference exclusion among multiple virtual domains. However, this work does not deeply consider the difference in the contextual semantics of system invocations, and there are limitations in the model's ability to resist malicious attacks. Gong et al. [16] designed a dynamic trustworthiness metric model in a remote proof scheme, introducing various functions, such as risk assessment function, feedback control function, etc. to dynamically evaluate node behavior. Wan [17] et al. proposed a trustworthiness evaluation scheme for internal behavior based on dynamic Bayesian networks for simple scenarios, which promoted the development of dynamic trust relationship evaluation research based on behavior analysis modeling.

3 IB-DTASOS Model

Trusted Computing Group (TCG) defines trusted based on the entity's behavior expectations [18], which means that the entity is only considered trustworthy if its behaviors are always in the expected way to achieve the expected goal. Shen proposed the active immune trustworthy computing based on the concept of the human immune system, that is, computing at the same time to perform security protection, using the password as a gene to implement identity recognition, state metrics, confidential storage, and other functions, timely identification of "own" and "non-self" components, to destroy and reject the harmful substances entering the body, which is equivalent to cultivating immunity for network information systems [19].

3.1 Abstract Description and Definition

The proposed IB-DTASOS model leverages system behavior to portray the trustworthiness of the operational state, and it is described using a three-tuple:

$$\text{IB - DTASOS} = <Be, En, Pc>$$

where Be, En, Pc denote the system behavior, the entity that is responsible for the perform specific behaviors, and access control compliance, respectively. IB-DTASOS is a three-dimensional trusted metric model for system operation, where entity trustworthiness is the basic, behavior trusted is the core, and access control compliance is the supplement.

Definition 1. *SC* denotes the system call: The interface between the user layer and the kernel layer is the majority implemented means by which the kernel provides system resource services to applications.

Definition 2. Atomic behavior: It consists of a system call, the subject that initiates the system call, and the object that acts on the system call.

$$Atomic\_B = (s, o, sc)$$

where s, o denote the subject and object, respectively. Among them, the subject in the system can be represented by the process identifier *PID*, while the object is already hidden in the semantics of the system call parameters. Therefore, the above can be simplified as: $sc_{p\_index}(param_1, param_2, \cdots, param_{n(0\leq n\leq 6)})$. Where *p_index* denotes the process PID number, $param_i$ indicates the system call parameter. Therefore, the atomic behavior can be reduced $Atomic\_B = sc_{p\_index}$.

3.2 Trusted Measurement Model

Assume that the probability that the atomic behavior sc_{p_index} meets the access control requirements is $p_{ac}(sc_{p\_index})$, then the probabilities that the subject and object of the atomic behavior are trusted are $p_E(s)$ and $p_E(o)$, respectively. Then we consider the probability that the atomic behavior is trustworthy:

$$\Pr(Atomic\_B) = p_{ac}(sc_{p\_index}) \wedge p_E(s) \wedge p_E(o) \tag{1}$$

A system behavior may be performed by multiple atomic behaviors together, and $\Pr(Atomic\_B)$ has only two assessment results, i.e., trusted or untrusted, and its probability distribution conforms to the binomial distribution. Assuming that the plausibility of the system behavior takes the value range [0, 1], as the conjugate distribution of the binomial distribution, the *Beta* distribution can better describe the plausible distribution of the system behavior after multiple atomic behaviors, then the probability density function of the system behavior is:

$$f(x|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} = \frac{1}{\mathrm{B}(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{2}$$

Among them, $\Gamma(\cdot)$ denotes gamma function, and $\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$. Therefore, Eq. (2) can be transformed into Eq. (3):

$$f(x|\alpha, \beta) = \frac{1}{\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx} x^{\alpha-1}(1-x)^{\beta-1} \tag{3}$$

Assume that a system behavior Be_i consists of n atomic behaviors, where the number of trusted atomic behaviors is α and the number of untrusted atomic behaviors is β, and $\alpha + \beta = n$. Then the trusted expectation of the system behavior is:

$$D(Be_i) = \int_0^1 xf(x)dx = \int_0^1 x \frac{1}{\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx} x^{\alpha-1}(1-x)^{\beta-1} \tag{4}$$

From the integration by parts: $D(Be_i) = \frac{\alpha}{\alpha+\beta}$.

Assume that at a certain time t, the sequence of monitored system behaviors is $(Be_1, Be_2, Be_3, \cdots, Be_n)$, $n \geq 1$, then the trusted measurement function of the system behavior at the time t is:

$$T_b(t) = \frac{1}{n} \sum_{i=1}^{n} [D(Be_i) \times \theta(i)] \tag{5}$$

where $\theta(i)$ is the reward and punishment factor, which is a dynamic function and can be set dynamically according to the system security policy. What calls for special attention is that the reward factor should be much larger than the punishment factor, so that the trusted changes of the system behavior follow the "fast-falling, slow-rising" characteristics.

Setting a suitable reward and punishment factor function can effectively improve the fault tolerance of the dynamic metric function and the trustworthiness of the assessment results, reducing the subjectivity of the behavioral trustworthiness assessment, that is, when the running system is attacked, the overall trustworthiness would have an obvious downward trend, while only the continuous trusted system behavior assessment results can improve the trusted of the system.

The operation process of the critical information system is complex and dynamic, so it is not possible to determine whether its operation process is trusted only by a limited number of metrics. Therefore, the measurement behavior should have low dispersion to effectively reflect the dynamic changes in the trustworthiness of the system operating. Additionally, from the perspective of sociology, trustworthiness has the space-time decay, i.e. the more distant a metric value is from the current moment, the less strongly it characterizes the plausible state of the system at the current moment. Therefore, trusted assessment of behavior is a dynamic, continuous, and complex process. In our model, we introduce the concept of IOWA to improve the accuracy and flexibility of the weight setting of the different metric values, where the metric time is the induction factor.

Assume using $V = (v_1, v_2, v_3, \cdots, v_n)$ represent the induction factor, and the trusted metric values are denoted by the $X = (x_1, x_2, x_3, \cdots x_n)$, then using the function f_w described the n-dimensional induced ordered weighted average operator generated by $V = (v_1, v_2, v_3, \cdots, v_n)$. Among them, $v - index(i)$ denotes the subscript of the i-th largest number after sorting $(v_1, v_2, v_3, \cdots, v_n)$ from largest to smallest.

$$f_w(< v_1, x_1 >, < v_2, x_2 >, \cdots, < v_n, x_n >) = \sum_{i=1}^{n} w_i x_{v-index(i)} \tag{6}$$

Assume that n measurements are made within the time window $T_w = [t_i, t_{i+1}]$, and the time-induced factor is added to the trusted measure values at different moments, expressed as follows:

$$T_b(t) = (T_{b1}^{(t_1)}, T_{b2}^{(t_2)}, T_{b3}^{(t_3)}, \cdots, T_{bn}^{(t_n)}) \tag{7}$$

Accordingly, the result of the integrated system metric within the time window T_w can be expressed as follows:

$$M_T(T_w) = f_w(T_{b1}^{(t_1)}, T_{b2}^{(t_2)}, T_{b3}^{(t_3)}, \cdots, T_{bn}^{(t_n)}) = \sum_{i=1}^{n} T_{bi}^{v-index(i)} \times \omega_i \tag{8}$$

where ω_i denotes the weighting coefficient of i th trusted metric value in the T_w, which represents the importance of the behavioral metric value compared to other moments.

According to the literature [20], in the method of obtaining an ordered weighted vector based on the maximum dispersion. $W = (w_1, w_2, w_3, \cdots w_n)^T$ should cover the following three conditions:

1 Maximize: $-\sum_{i=1}^{n} w_i \ln(w_i)$;

2 $\frac{1}{n-1}\sum_{i=1}^{n}(n-i)w_i = \alpha, 0 \le \alpha \le 1$;

3 $\sum_{i=1}^{n} w_i = 1, 0 \le w_i \le 1,\ i = 1, 2, \cdots n$;

Based on the principle of maximum dispersion [1], combined with 1) 2) 3), we can derive that:

$$\alpha = \frac{1}{n-1}\sum_{i=1}^{n}(n-i)w_i \tag{9}$$

$$w_1[(n-1)\alpha + 1 - nw_1]^n = ((n-1)\alpha)^{n-1}[((n-1)\alpha - n)w_1 + 1] \tag{10}$$

$$w_n = \frac{((n-1)\alpha - n)w_1 + 1}{(n-1)\alpha + 1 - nw_1} \tag{11}$$

$$\ln w_j = \frac{j-1}{n-1}\ln w_n + \frac{n-j}{n-1}\ln w_1 \Rightarrow w_j = \sqrt[n-1]{w_1^{n-j} w_n^{j-1}} \tag{12}$$

In practice, setting a suitable α, w_1 can be obtained by Eq. (10), then w_n also can be calculated according to Eq. (11), and thus the other weighted vector values can be calculated step by step according to Eq. (12). The process of the weighted vector w_i is given in Algorithm 1:

Algorithm 1: Calculating weights

Input: α , n

Output: $(w_1, w_2, w_3, \cdots w_n)$

BEGIN

  if n > 1

    Calculate w_1 depending on Eq. (10) ;

    Calculate w_n depending on Eq. (11) ;

    For j=2 to n-1 **do**

      $w_j = \sqrt[n-1]{w_1^{n-j} w_n^{j-1}}$

END

4 Conclusion

In this paper, we proposed an IOWA-Based Dynamic Trusted Assessment Model for System Operational State Behavior (IB-DTASOS) to assess the trusted of the runtime system. This scheme comprehensive considers the trusted of three factors: system behavior, the entity which performs the system behavior, and the compliance of access control. Additionally, we introduce the bayesian theory and IOWA to effectively improve the scientificity of the trusted assessment model.

References

1. Hallappanavar, Vijay L., Birje, M.N.: Trust management in cloud computing. In: Cloud Security: Concepts, Methodologies, Tools, and Applications. IGI Global, pp. 1686–1711 (2019)
2. https://www.cverc.org.cn/head/zhaiyao/news20220629-FoxAcid.htm
3. Richard Clarke, R.K.: Cyberwar: the next threat to national security and what to do about it. In: ECCO (2011)
4. Zhang, F., Xu, M.D., Chao, H.C., Zhang, C., Liu, X.L., Hu, F.N.: Real-time trust measurement of software: behavior trust analysis approach based on noninterference. Ruan Jian Xue Bao/J. Softw. **30**(8), 2268–2286 (2019)
5. Wang, Z., Zhuang, Y., Yan, Z.: Probe-based dynamic integrity measurement scheme for mobile devices using ARM trustzone. J. Chin. Comput. Syst. **42**(11), 2422–2427 (2021)
6. Jiang, R., Xin, Y., Chen, Z., Zhang, Y.: A medical big data access control model based on fuzzy trust prediction and regression analysis. Appl. Soft Comput. 108423 (2022)
7. Zhuang, L., Cai, M., Shen, C.: Trusted dynamic measurement based on interactive Markov chain. J. Comput. Res. Dev. **48**(08), 1464–1472 (2011)
8. Luo, X., Tang, Z., Zhao, Y.: Dynamic software reliability assessment based on Markov chain. Appl. Res. Comput. **32**(08), 2400–2405 (2015)
9. Li, X.Y., Gui, X.L.: Cognitive model of dynamic trust forecasting. J. Softw. **21**(1), 163–176 (2010)
10. Li, X., Gui, X., Zhao, J., Feng, D.: Novel scalable aggregation algorithm of feedback trust information. J. Xi'an Jiaotong Univ. **08**, 879–883 (2007)
11. Yager, R.R.: Induced aggregation operators. Fuzzy Sets Syst. **137**(1), 59–69 (2003)
12. Gong, B., Liu, J., Guo, S.: A trusted attestation scheme for data source of internet of things in smart city based on dynamic trust classification. IEEE Internet Things J. **8**(21), 16121–16141 (2020)
13. Yu, N., Wang, G.: Study on cross-domain access security based on trusted expectations. Appl. Res. Comput. **37**(11), 3406–3410+3416 (2020)
14. Zhao, B., He, J.: Bayes decision theory based risk minimization authorization mapping. J. Commun. **36**(S1), 157–161 (2015)
15. Huang, H.X., Zhang, J.B., Yuan, Y.L., Wang, X.: Research on trusted startup of virtual machine based on non-interference theory. Ruan Jian Xue Bao/J. Softw. (2021)
16. Gong, B., Wang, Y., Liu, X., Qi, F., Sun, Z.: A trusted attestation mechanism for the sensing nodes of Internet of Things based on dynamic trusted measurement. China Commun. **15**(2), 100–121 (2018). https://doi.org/10.1109/CC.2018.8300276
17. Liu, W., Ci, L., Liu, L.: Research on behavior trust based on Bayesian inference in trusted computing networks. In: 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), pp. 1134–1138 (2015). https://doi.org/10.1109/SmartCity.2015.222

18. Feng, D.G., Liu, J.B., Qin, Y., et al.: Trusted computing theory and technology in innovation-driven development. Sci. Sin. Inform. **50**, 1127–1147 (2020). https://doi.org/10.1360/SSI-2020-0096
19. Shen, C.X.: To create a positive cyberspace by safeguarding network security with active immune trusted computing 3.0. J. Inf. Secur. Res. **4**(04), 282–302 (2018)
20. Fullér, R., Majlender, P.: An analytic approach for obtaining maximal entropy OWA operator weights. Fuzzy Sets Syst. **124**(1), 53–57 (2001)

Multispectral Image Compression Based on Improved Residual Network

Jiahui Tang(✉), Kang Wang, Fanqiang Kong, and Guanglong Ren

College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
jiahui_tang@nuaa.edu.cn

Abstract. In view of the fact that multispectral image has extensive application prospects and contains rich information, how to extract features efficiently is crucial for multispectral image compression. This paper proposed a new compressed framework which utilizes ConvGRU and contextual transformer respectively as the backbone of the spectral attention and Cot block to capture the spectral and spatial features. Additionally, residual units are used to explore deeper features to reserve fine details of image. The image data is firstly fed into the encoder which is composed of aforementioned units and then the quantizer and entropy encoder to convert data into the bit stream. The rate distortion optimization is utilized to perfect the reconstructed image better. Our method is compared with JPEG2000 via PSNR. The experimental results demonstrate that the method outperforms JPEG2000.

Keywords: Multispectral image compression · Convolutional Grated Recurrent Unit · Contextual transformer

1 Introduction

Recent years have witnessed remarkable progress in the application of multispectral images such as precision agriculture, mineralogy, military reconnaissance, for its rich spectral and spatial information contained. Multispectral imaging produces three-dimensional data cubes (s_I, w_I, h_I) which indicate the number of spectral bands, the length and width of images respectively. Noting that large quantities of image data result in difficulties of image storage and transmission, hence how to compress multispectral image efficiently and effectively has become a research hotspot. There is great redundancy information in multispectral image, so extracting discriminative features which means capturing the spectral and spatial correlation can remove redundancy effectively and obtain higher quality reconstruction accordingly.

With the surge of deep learning, neural networks show great potential in image compression and achieve promising image reconstruction performances indeed. Toderici, et al. [1, 2] generated the bit streams of entropy encoding by making full use of the recurrent networks in the field of RIB compression. Balle et al. [3] firstly proposed a variable auto-encoder based on a hyperprior which efficiently mitigates the conditions

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 228–232, 2023.
https://doi.org/10.1007/978-981-99-1260-5_29

of incomplete spatial information for lossy image compression. However, most existing methods are susceptible to overlooking the extraction of the spectral correlation between adjacent bands, which are critical for the multispectral compression reconstruction. Considering this, we provide an optimization perspective to better address this problem: we treat multispectral image data as sequential input in spectral dimension so that dependencies between adjacent spectra can be mined deeply from the attention mechanism constructed by Convolution Grated Recurrent Unit. Inspired by the resurgence of transformer [4], contextual transformer is applied to acquire the global features of image as the complementarity of 2D convolutional operations for its capability of local features. This is conductive to reserving fine details and improving the quality of reconstruction. Furthermore, residual units are applied to extract deeper features. The experimental results demonstrate the proposed network has superior performance than JPEG2000.

2 Proposed Method

The input image is fed into the encoder which is illustrated in Fig. 1 to extract features through the spectral attention and Cot block simultaneously. Figure 2 shows that the residual unit consisted of two 2D convolution layers is utilized to explore deeper features for the purpose of further removing image redundancy. Then the extracted features are downscaled to compress the acquired features. The quantizer quantizes the acquired features to obtain the compressed stream by the lossless entropy codec. The decoder and encoder are symmetrical in structure. In terms of decoding in Fig. 3, the bit stream primarily converts into the image data through the entropy decoder, inverse quantization and feature extracted network. The more detailed structure of the spectral attention and Cot block will be illustrated as follows.

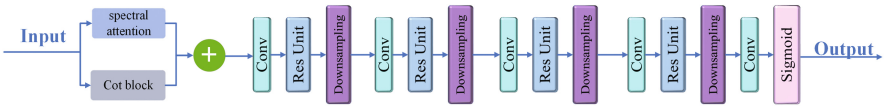


Fig. 1. The structure of encoder

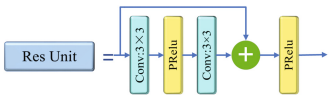


Fig. 2. The residual unit

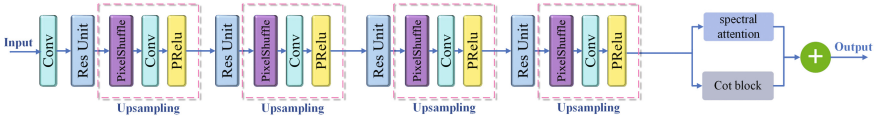


Fig. 3. The structure of decoder

2.1 Spectral Attention

In order to explore dependencies of adjacent spectral bands, more and more methods tried to make full use of RNN whose memory structures are capable of modeling the relations of spectral features. In this block, we employ ConvGRU because ConvGRU has few parameters but fast convergence. We denote $X_I \in R^{s_I \times w_I \times h_I}$ as the original multispectral image data, where s_I, w_I and h_I represent the number of spectral bands, width and height respectively. Figure 4 shows the structure of the block. X_I is split into s_I components along the spectral dimension and then integrated as a sequence $\left\{X_{I1} \in R^{1 \times w_I \times h_I}, \cdots, X_{It}, \cdots, X_{Is_I}\right\}, t \in \{1, 2, \cdots, s_I\}$. The sequential data is fed one by one to the ConvGRU layer in order to extract spectral information between adjacent spectra and generate the attention map Z_{Se}. Moreover, a global average pooling layer and the softmax function layer are also added to compose the backbone of the spectral attention. The details of the module can be described as

$$Z_{se} = f_{gap}\left(f_{Convgru}(X_I)\right) \tag{1}$$

$$\alpha_{se} = soft\max(Z_{se}) \tag{2}$$

$$\tilde{X}_I = X_I \odot \alpha_{se} \tag{3}$$

where $\odot$ indicates a element-wised product operation and α_{se} represents the attention weight generated by the softmax function $soft\max(\cdot)$. In addition, $f_{gap}(\cdot)$ as a spatial pooling layer is utilized to pool with the width and height of the image, and the $f_{Convgru}(\cdot)$ means the 3×3 ConvGRU layer where the time steps are set as s_I.

Compared with input, the output $\tilde{X}_I$ of the same size has obtained the long-dependencies spectral information through the attention module.

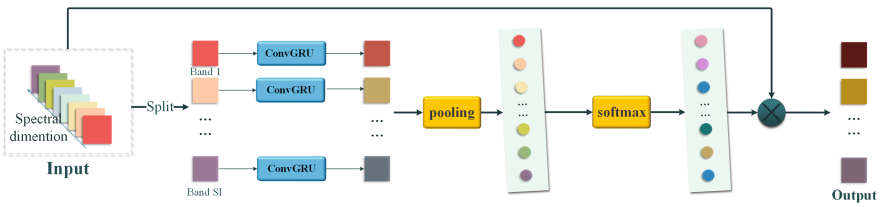


Fig. 4. The spectral attention

2.2 Cot Block

Vision Transformer and lots of its variants have been widely applied in the image processing field for their competitive results. However, a potential disadvantage of transformers is that they are consisted of multiheaded self-attention and the computational complexity increases exponentially as the size of the image increases. To tackle this

limitation, contextual transformer (Cot) is designed to reduce the cost of large quantities of self-attention.

Due to receptive field and shift-invariance, the convolution is capable of extracting the local features of image very well but ignores the global features instead. Considering this problem, cot block utilizes 3 × 3 convolution to encode the input keys of self-attention and then uses two consecutive 1 × 1 convolution, so it can model the global relation of pixels. As Fig. 5 shows, cot block is composed of 2D convolution and cot attention aforementioned hybrid network, which combines both advantages to extract local and global features simultaneously.

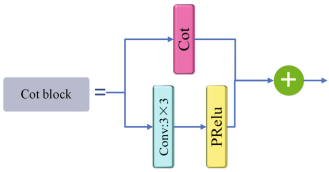


Fig. 5. The Cot block

2.3 Rate-Distortion Optimizer

On the account of interdependence of the bit rate and loss of distortion, we adopt the method of rate-distortion optimization so as to make the quality of the reconstructed image better. The loss function we utilize is as follows:

$$L = L_D + \lambda L_R \tag{4}$$

$$L_R = -E\left[\log_2 P_d(x)\right] \tag{5}$$

$$P_d = \int_{x-\frac{1}{2}}^{x+\frac{1}{2}} P_c(x)dx \tag{6}$$

where L_D,λ and L_R respectively denotes the distortion counted by MSE, the penalty weight and approximate calculation of the entropy, namely the bit rate. Furthermore, the $P_d(x)$ represents the probability of the image data.

3 Experiment

The dataset of our experiment comes from the 8-band WorldView-3 satellite, which is composed of almost 8700 images as the size of 128 × 128. The testing part has 14 images of size of 512 × 512 to verify performance of this novel compressed network.

PSNR is the performance indicator that we evaluate to measure the quality and distortion of reconstructed images. As Table 1 shows, the network we proposed in the paper is almost 6 dB higher than JPEG2000 at the same bit rate.

Table 1. The average PSNR in dB of two contrast methods at different bit rates

	Rate	PSNR	Rate	PSNR	Rate	PSNR
JPEG2000	0.505	34.58	0.420	33.7	0.140	29.81
Proposed		**42.64**		**40.11**		**37.74**

4 Conclusion

In this paper, we proposed a novel compressed network based on residual units which utilizes ConvGRU as the backbone of the spectral attention to explore the adjacent spectral dependences. The results demonstrate that the method outperforms JPEG2000 and shows ConvGRU and contextual transformer both have great abilities of feature extraction.

References

1. Toderici, G., O'Malley, S.M., Hwang, S.J., et al.: Variable rate image compression with recurrent neural networks. In: ICLR2016. CoRR abs/1511.06085:n.pag (2016)
2. Toderici, G., Vincent, D., Johnston, N., et al.: Full resolution image compression with recurrent neural networks. In: IEEE CVPR2017 1, 5435–5443 (2017). https://doi.org/10.1109/CVPR.2017.577
3. Ballé, J., Minnen, D., Singh, S., Hwang, S.J., Johnston, N.: Variational image compression with a scale hyperprior. In: ICLR20018 (2018). https://doi.org/10.48550/arXiv.1802.01436
4. Li, Y., Yao, T., Pan, Y., Mei, T.: Contextual transformer networks for visual recognition. In: CVPR2021 (2021). https://doi.org/10.48550/arXiv.2107.12292

Research on Multi-domain Data Transmission and Presentation Algorithm Based on WiFi Network

Wanqing Wu(✉), Baoguo Yu, He Dong, Qingwu Yi, and Chenglong He

State Key Laboratory of Satellite Navigation System and Equipment Technology, The 54Th Research Institute of CETC, Shijiazhuang 050081, China
2672624360@qq.com

Abstract. Aiming at the perception and visualization of multi-domain data in indoor and outdoor environments, we proposed a multi-domain data transmission and presentation method based on WiFi networking. First, the multi-domain data sensed by distributed nodes is preprocessed to remove abnormal data and convert it into a standard format. Then, the distributed nodes conduct WiFi networking, and transmit the multi-domain data to the fixed node, and the fixed node forwards the received data to the multi-domain data processing and display platform. Finally, the multi-domain data processing and display platform uses the graphing algorithm to visualize the multi-domain data. The experimental results show that the algorithm we proposed can effectively solve the problem of reliable transmission and mapping of multi-domain data in unknown environments.

Keywords: WiFi networking · Multi-domain data · Situational awareness · Heat map

1 Introduction

In recent years, with the rapid development of information technology, new challenges have been brought to situational awareness technology. In order to increase the probability of success in military operations it is crucial to predict multiple factors and to coordinate actions involving various elements [1]. Because a single perception field is highly related to the sensor environmental adaptability, detection performance, and the difference in field spatial distribution, situational awareness based on a single field has the problem of insufficient reliability. Therefore, situational awareness based on field map needs to be based on multiple domains. Through the interaction of multi-domain fields, the spatial distribution characteristics and sensor characteristics of different fields can be fully utilized, and it has stronger adaptability. With the development of information transmission [2], fusion [3], visualization [4] and other technologies, the connotation of situational information has been greatly expanded. WiFi is the most common global

This paper is supported by Development Fund of The 54th Research Institute of CETC (SXX21684X025, SXX22684X001).

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 233–241, 2023.
https://doi.org/10.1007/978-981-99-1260-5_30

wireless access technology, and the robust WiFi infrastructure can be used anywhere in the world for high-speed data transmission [5]. Data visualization is a technology that uses graphics and computer vision to express and analyze data from the visual expression level [6]. Data visualization can be used to display data more intuitively. At present, the transmission, aggregation and visualization of multi-domain data has become a research hotspot.

Our paper aims to carry out research on multi-domain data transmission and presentation technology based on WiFi networking to solve the perception and visualization of multi-domain data in indoor and outdoor environments. This paper studies distributed node data preprocessing, data transmission, multi-domain data interpolation and data visualization technology to solve the problem of reliable transmission and graphing of multi-domain data in unknown environments.

2 System Model

The paper proposed a multi-domain data transmission and presentation algorithm based on WiFi networking. Figure 1 shows the system model of this paper, which mainly includes distributed nodes, fixed nodes, transmission network and visual interface.

First, distributed nodes perceive surrounding multi-domain data. Then the fixed node collects data and transmits the multi-domain data to the background. Finally, the data visualization is performed in the background.

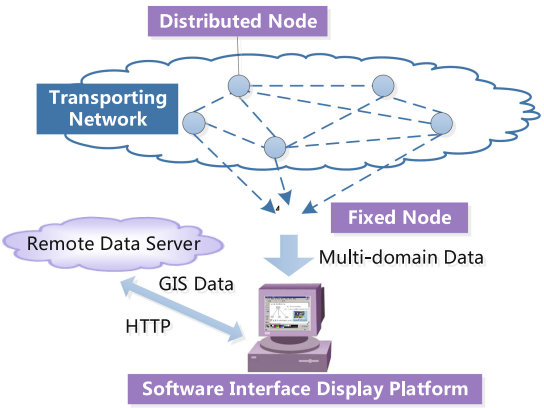


Fig. 1. System model

3 Proposed Method

3.1 Data Preprocessing

The node senses the electromagnetic, geomagnetic and other data of the surrounding environment through the sensors on board. The node preprocesses the acquired data, removes abnormal data and converts the data into a defined standard format.

When there is a significant difference between the data collected by the sensor node and the data at the previous sampling time, it is called abnormal data.

The sensor node samples every ΔT time, and at a certain sampling time f collects m different types of data, which can be regarded as a data point set $\mathbf{X}$ in an m-dimensional space, and the data set of the node in a sampling period can be regarded as a matrix $\mathbf{r}$:

$$\boldsymbol{r} = \begin{bmatrix} r_1(t_1) & r_1(t_2) & \cdots & r_1(t_n) \\ r_2(t_1) & r_2(t_2) & \cdots & r_2(t_n) \\ \vdots & \vdots & \vdots & \vdots \\ r_m(t_1) & r_m(t_2) & \cdots & r_m(t_n) \end{bmatrix} \tag{1}$$

Among them, $t_1, t_2, t_3, \ldots, t_n$ is the sampling time.

The sliding window model is used to observe the data flow sequence in the latest sampling time period. The method is to take a sliding window of length $|W|$ for the data flow, and divide the window into m blocks of size n, $B_1, B_2, \ldots, B_m$. If the sampling frequency is f times/minute, the window will be filled after mn/f minutes. Whenever a new time t_{new} data enters the sliding window, the relationship with the sampled data that will be replaced by the time t_{old} is shown in the following formula.

$$\mathrm{mod}(t_{new}, |W|) = \mathrm{mod}(t_{old}, |W|) \tag{2}$$

Among them, mod (a, b) represents the remainder function. The data at the new sampling time t_{new} is placed in the B_i block of the sliding window, and the distance calculation formula of the data point is as follows.

$$d(x_i, y_j) = \sqrt{\sum_{k=1}^{d} (x_{ik}, x_{jk})} \tag{3}$$

The similarity between the perceptual data x_i and x_j at sampling time i and sampling time j is $q(x_i, x_j) = 1/d(x_i, x_j)$. . The distance between two identical data is 0, and the similarity is 1. If the relative distance between the two data is infinite, the similarity tends to 0. It follows that the similarity of two sensor readings is always between (0,1].

Therefore, data similarity is an important criterion for measuring the similarity of two sensor readings, and it is also an important basis for measuring whether a data point is abnormal. Data points that are determined to be abnormal are directly eliminated to complete data preprocessing.

3.2 Data Transmission

Distributed nodes transmit multi-domain data in standard format to fixed nodes through WiFi networking, and fixed nodes transmit the collected multi-domain data to the multi-domain data processing and display platform by wire, using UDP transmission, and the transmission frequency is 1 Hz.

The nodes are networked, the fixed nodes enable the AP mode, other distributed nodes enable the Managed mode to connect to the AP, and the fixed nodes and the multi-domain data processing and display platform use network cables for wired connection.

Distributed nodes send multi-domain data in standard format after data preprocessing to fixed nodes through network transmission, and the sending frequency is 1 Hz. And receive the background command forwarded by the fixed node, determine whether the termination command is received, if so, terminate the data transmission, otherwise, continue to send the data.

The workflow of distributed nodes is shown in Fig. 1, and the workflow of fixed nodes is shown in Fig. 2.

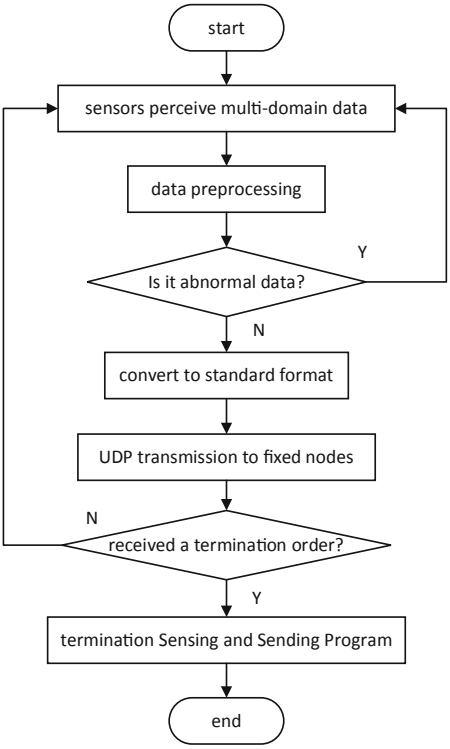


Fig. 2. Distributed node workflow

3.3 Multi-domain Data Interpolation

After the multi-domain data processing and display platform receives the data, it stores the data in the local database, and then uses the interpolation algorithm to fill the data (Fig. 3).

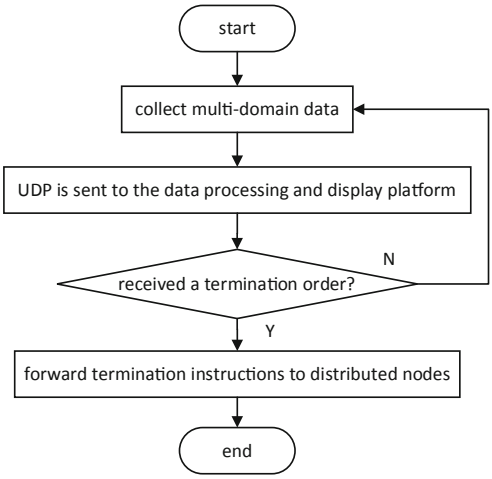


Fig. 3. Fixed node workflow

The variance and expectation are assumed to be constant at any point in the space.

$$\begin{aligned} E[z(x)] &= c \\ D[z(x)] &= \sigma^2 \end{aligned} \tag{4}$$

Based on the above assumptions, the value of any point can be composed of the average value of the space and the random deviation $R(x)$ of the point.

$$z(x) = E[z(x)] + R(x) = c + R(x) \tag{5}$$

Among them, $R(x)$ represents the random deviation at point x, and its variance is constant.

$$Var[R(x)] = \sigma^2 \tag{6}$$

The interpolation formula is:

$$\hat{z}_0 = \sum_{i=1}^{n} \lambda_i z_i \tag{7}$$

Among them, $\hat{z}_0$ is the estimated value at the point (x_0, y_0), that is $z_0 = z(x_0, y_0)$, λ_i is a set of optimal coefficients with the smallest error between the sampling point and the point to be estimated, and both $\min_{\lambda_i} Var\left(\hat{z}_0 - z_0\right)$ and unbiased conditions are satisfied, as long as λ_i is calculated, The estimated value of this point can be calculated by weighted summation, and the value of the regionalized variable at any point in space can be estimated by a linear combination of the known observation point values. Specific steps are as follows:

(1) According to the data of the sampling points, calculate the distance and experimental variogram values in pairs, group them according to the distance tolerance and angle tolerance, and then sum the average values of the semi-variances in the same group.

$$\gamma(h) = \frac{1}{2n}\sum_{i=1}^{n}[z(x) - z(x+h)]^2 \tag{8}$$

(2) According to the fitting model, the MFO algorithm is used to automatically fit the experimental variogram to obtain a curve that can estimate the variogram value at any point;
(3) Calculate the semivariance r_{ij} between all known points;
(4) For the unknown point z_0, calculate the variogram r_{i0} from it to all known points z_i, $i = 1, 2, 3, \ldots, n$.
(5) Solve the following equations to get the optimal coefficient λ_i, and then the weighted average can find the value of the estimated point.

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} & 1 \\ r_{21} & r_{22} & \cdots & r_{2n} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \\ -\phi \end{bmatrix} = \begin{bmatrix} r_{10} \\ r_{20} \\ \cdots \\ r_{n0} \\ 1 \end{bmatrix} \tag{9}$$

3.4 Data Visualization

The multi-domain data processing and display platform stores the processed data in the local database, and visualizes the data in the form of a heat map through a graphing algorithm.

After the multi-domain data processing and display platform receives the data, it first calls the data storage interface to store the data in the local database. The storage format is shown in Table 1. Then use the mapping algorithm to draw the data on the map in the form of heat map.

Table 1. Data storage format

DeviceId	Lng	Lat	H	GeoMag	ElecMag
id1	lng1	lat1	h1	geomag1	electmag1
id2	lng2	lat2	h2	geomag2	electmag2
…	…	…	…	…	…
idN	lngN	latN	hN	geomagN	electmagN

4 Implementation and Performance Analysis

The following is an experimental verification of our proposed multi-domain data transmission and presentation algorithm based on WiFi networking. Raspberry Pi 4B is used to carry sensors, the installation system is Ubuntu18.04, and the physical nodes are shown in the Fig. 4.

Fig. 4. Physical photos of distributed nodes

The algorithm proposed in Sect. 3 is tested and verified. The distributed nodes transmit the preprocessed sensory data to the WiFi network. Figure 5 shows the json format analysis of the transmitted data. The fixed node collects the data and forwards the data to the multi-domain data. Processing and presentation platform.


```
{
    "deviceId": 1,
    "lng": 114.857,
    "lat": 38.403,
    "h": 5.871,
    "geoMag": 2921.5682432556664,
    "magX": -1162.0,
    "magY": 226.0,
    "magZ": 2671.0,
    "electMag": 17.678
}
```


Fig. 5. The json format of the transmitted data

The multi-domain data processing platform visualizes the collected multi-domain data, and the display results are shown in the Fig. 6.

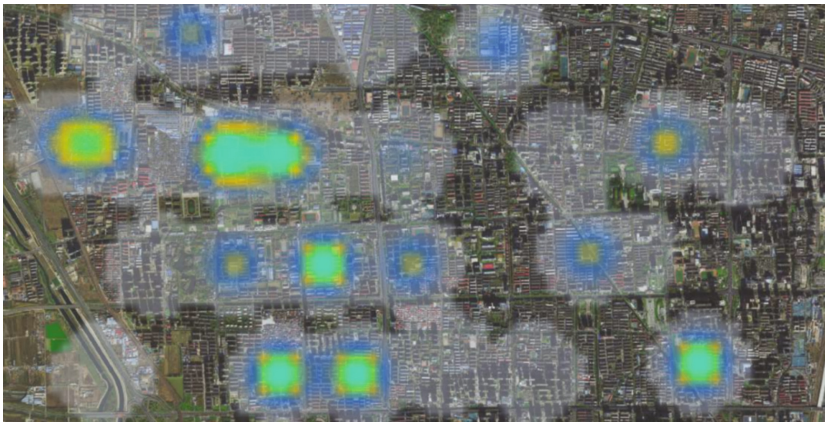


Fig. 6. Visualization of data

5 Conclusion

We proposed a multi-domain data transmission and presentation algorithm based on WiFi networking. It can be seen from the experimental results that our algorithm can effectively solve the problem of perception and visualization of multi-domain data in indoor and outdoor environments. Through WiFi networking, multi-domain data can be quickly transmitted and collected, and the multi-domain data processing platform can perform real-time visual display of multi-domain data. In summary, this algorithm can solve the problem of reliable transmission and mapping of multi-domain data in unknown environments.

References

1. Suchanski, M., Kaniewski, P., Romanik, J., Golan, E., Zubel, K.: Radio environment maps for military cognitive networks: deployment of sensors vs. map quality. In: 2019 International Conference on Military Communications and Information Systems (ICMCIS), pp. 1–6 (2019)
2. Hou, Y., Hu, Z.: Information transmission control technology of electronic information system based on IP network. In: 2021 International Conference on Networking, Communications and Information Technology (NetCIT), pp. 366–370 (2021)
3. Kim, J., Eo, Y., Pyeon, M., Park, J.: Real-time mapping technology for the multi-dimensional spatial information service. In: The 7th International Conference on Networked Computing and Advanced Information Management, pp. 280–283 (2011)
4. Cui, H., Lv, Z., Guo, Y., Zhang, R., Zhang, L.: A visualization system that ensures the localization of private geographic information. In: 2020 5th International Conference on Universal Village (UV), pp. 1–5 (2020)
5. Lee, S., Choi, H., Kim, T., Park, H.-S., Choi, J.K.: A Novel Energy-Conscious Access Point (eAP) System with Cross-Layer Design in Wi-Fi Networks for Reliable IoT Services. In: IEEE Access **10**, pp. 61228–61248
6. Liang, Z., NuErTaYi, B., Xu, Y., Zhang, F., Yang, J., Wang, J.: Application of multidimensional data visualization technology and WebGIS in lightweight visualization system design. In: 2021 9th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), pp. 1–4 (2021)

7. Sahu, A., et al.: Multi-source multi-domain data fusion for cyberattack detection in power systems. IEEE Access **9**, 119118–119138 (2021)
8. Hazarika, A., Barman, P., Talukdar, C., Dutta, L., Subasi, A., Bhuyan, M.: Real-time implementation of a multidomain feature fusion model using inherently available large sensor data. IEEE Trans. Industr. Inf. **15**(12), 6231–6239 (Dec.2019)
9. Li, X., He, J., Zhu, N., Hou, Z.: Collaborative filtering recommendation based on multi-domain semantic fusion. In: 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 255–261 (2020)
10. Ohta, M., Sasaki, J., Takahashi, S., Yamashita, K.: WiFi positioning system without AP locations for indoor evacuation guidance. In: 2015 IEEE 4th Global Conference on Consumer Electronics (GCCE), pp. 483–484 (2015)

Research on Spectrum Sensing Algorithm Based on Deep Neural Network

Xinqi Zhu[1], Xiaohua Zhang[2], and Shubin Wang[1(✉)]

[1] College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China
wangshubin@imu.edu.cn

[2] Department of Foreign Languages, Guizhou University of Commerce, Guiyang, China

Abstract. To solve the existing problems in spectrum sensing methods adopted in cognitive radio systems, such as poor sensing performance and easily being influenced dramatically by noise in low signal-to-noise ratio (SNR) environments, we propose a spectrum sensing algorithm based on deep neural network in this paper. First, we preprocess the received primary user (PU) signal with energy normalization, and then the normalized received signal is fed into the proposed Noisy Activation CNN-GRU-Network (NA-CGNet), which includes a one-dimensional convolutional neural network (1D CNN) and a gated recurrent unit (GRU) network with noisy activation function. NA-CGNet is able to extract spatially and temporally correlated features of the signal received by the time series. The deep neural network with noisy activation function can improve the anti-noise performance of the network, thus enhancing the spectrum sensing performance at low signal-to-noise ratio (SNR) level. Simulation results show that the proposed NA-CGNet model achieves a spectrum sensing accuracy of 0.7557 at SNR = -13dB, which is 0.0567 better than the existing DetectNet model, achieving a lower false alarm probability of P_f=7.06%.

Keywords: Cognitive radio · Deep neural network · Spectrum sensing · Noisy activation function

1 Introduction

Cognitive Radio (CR) has received much attention from academia and industry as a potential solution to the problem of wireless spectrum resource strain [1].The concept of CR is spectrum sharing and spectrum sensing is an important prerequisite for CR, where the user authorized to use the band is the Primary User (PU) and the unlicensed Secondary User (SU) is allowed to opportunistically occupy the free licensed spectrum resources of the PU to improve spectrum utilization [2]. With the advantages of low complexity and no a prior knowledge of PU signals, traditional energy detection method are easy to implement, but it is greatly affected by noise interference in a low SNR environment, which leads to a significant decrease in the accuracy of spectrum sensing [3]. Traditional methods are model-driven and it is difficult to obtain prior knowledge about signal or noise power, especially for non-cooperative communication [4]. Therefore, traditional

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 242–250, 2023.
https://doi.org/10.1007/978-981-99-1260-5_31

methods do not always meet the requirements for fast and accurate spectrum sensing in real communication environments.

Deep learning (DL) is able to explore the most appropriate functions and models in a data-driven mining manner that can improve signal detection performance [5, 6]. Deep learning based spectrum sensing have attracted increasingly research work, i.e., Zheng et al. [7] proposed a deep learning classification technique that uses the power spectrum of the received signal as an input to train a CNN model and normalizes the input signal to solve the problem of noise uncertainty. Xie et al. [8] utilized their historical knowledge of PU activity and used a CNN-based deep learning algorithm to learn PU activity patterns to achieve spectrum sensing (APASS). However, CNN-based deep neural networks are not sufficient to process PU signals because wireless spectrum data are time-series data with fixed time correlation, and long-term memory (LSTM) network is introduced into the DL network to extract total time correlation [9]. Gao et al. [10] combined LSTM and CNN to form DetectNet to extract total time correlation from received signal time series data to learn total dependencies and local features, which enables better spectrum sensing than LSTM and CNN structures separately.

The spectrum sensing method based on Deep Learning may focus on extracting various aspects of the input. However, existing CNNs are not suitable for temporal modeling and time series data analysis [11], and 1D CNNs are good at extracting local features from serial data to obtain sufficient locally relevant features [15]; LSTM and GRU are capable of temporal correlation extraction. However, GRU [12] has a simpler network structure than LSTM and is easier to train, which can largely improve the training efficiency of the network while achieving similar results with LSTM. To the best of the authors' knowledge, the temporal dependence of spectral data has not been modeled in the literature using GRU networks with noisy activation function (NAF) [13], to solve the problem that the existing methods are not robust to noise in a low signal-to-noise environment, the NAF is used in deep neural networks, and Noisy Activation CNN-GRU-Network (NA-CGNet) is proposed to deal with the non-cooperative spectrum sensing problem. Simulation results verify that the proposed algorithm outperforms the existing spectrum sensing methods in terms of spectrum sensing accuracy at low SNR and achieves low false alarm probability at the same time.

2 Related Work

2.1 System Model

According to the idle or busy state of the PU, spectrum sensing can be viewed as a binary classification issue, i.e., the presence or absence of PU. Therefore, the signal detection of the SU scenario can be modeled as the following binary hypothesis testing issue [14].

$$X(n) = \begin{cases} H_0 : U(n) \\ H_1 : h * S(n) + U(n) \end{cases} \tag{2.1}$$

where $X(n)$ is the n-th received signal of the SU, $U(n)$ is additive noise following the zero mean circularly symmetric complex Gaussian (CSCG) distribution with variance σ_w^2. $S(n)$ denotes the noise-free modulated signal transmitted by the PU, and h represents

the channel gain between the PU and the SU, where it is assumed that this channel is constant and follows Gaussian distribution mode. H_0 is used to assume that the PU is not present, and H_1 is used to assume that the PU is present.

Typically, two performance metrics for spectrum sensing are considered, namely the detection probability P_d and the false alarm probability P_f. If a deep learning-based approach is used to sense spectrum, the output of the activation layer is the probability of each classification. For this binary classification issue, determining $i = arg\max(P_i)$ is equivalent to comparing the output probability under hypothesis H_1 with a threshold value γ_D. Thus, the performance metric can be defined as

$$\begin{cases} P_d = P_r\left(z^1 > \gamma_D | H_1\right) \\ P_f = P_r\left(z^1 > \gamma_D | H_0\right) \end{cases} \tag{2.2}$$

where z^1 denotes the probability of output neurons labeled "1" in the presence of PU.

2.2 Noisy Activation Function

The activation function enhances the nonlinear fitting ability and nonlinear representation of the neural network. Due to the saturation phenomenon of the generally used nonlinear activation function itself, it can lead to gradient vanishing problems. Caglar et al. [13] proposed to use Noisy Activation Function (NAF) learn the scale of the noise to make the gradient more significant.

Gate-structured RNN are proposed to solve the long-term dependence of time series, it includes LSTM and GRU, which both use soft saturated nonlinear functions such as sigmoid and tanh to imitate the hard decisions in digital logic circuits. However, the saturation characteristics cause the gradient to vanish when crossing the gate, and they are only softly saturated, which leads to the problem that the gates cannot be fully opened or closed, so hard decisions cannot be realized.

The hard decision problem can be solved by using a hard saturated nonlinear function, taking the sigmoid and tanh function as an example, the first-order Taylor expansion near the 0 point which is constructed as a hard saturation function, it is as follows,

$$\begin{cases} sigmoid(\mathrm{x}) \approx \mathrm{u}^{\mathrm{s}}(\mathrm{x}) = 0.25x + 0.5 \\ tanh(\mathrm{x}) \approx \mathrm{u}^{\mathrm{t}}(\mathrm{x}) = x \end{cases} \tag{2.3}$$

Clipping the linear approximations result to,

$$\begin{cases} hard\_sigmoid(\mathrm{x}) = max(\min(\mathrm{u}^{\mathrm{s}}(\mathrm{x}), 1), 0) \\ hard\_tanh(\mathrm{x}) = max(\min\left(\mathrm{u}^{\mathrm{t}}(\mathrm{x}), 1\right), -1) \end{cases} \tag{2.4}$$

The use of a hard saturated nonlinear function aggravates the problem of vanishing gradient because the gradient is set to exactly zero at saturation rather than converging to zero. However, further stochastic exploration of the gradient can be facilitated by introducing noise into the activation function that varies with the degree of saturation. Let us consider a noisy activation function.

$$\phi(x, \xi) = \alpha \mathrm{h}(x) + (1 - \alpha)\mathrm{u}(x) + \mathrm{d}(x)\sigma(x)\epsilon \tag{2.5}$$

where $\alpha \mathrm{h}(x)$ is the nonlinear part, α is a constant hyperparameter that determines the noise and the direction of the slope, $h(x)$ represents the initial activation functions such as hard_sigmoid and hard_tanh, $(1-\alpha)\mathrm{u}(x)$ is the linear part, $\mathrm{d}(x)\sigma(x)\epsilon$ is the random part of the noise, $\mathrm{d}(x) = -sgn(x)sgn(1-\alpha)$ changes the direction of the noise, $\sigma(x) = c(g(p\Delta) - 0.5)^2$, and ϵ is the noise that generates the random variable.

3 Spectrum Sensing Based on Noisy Activation Deep Neural Network

In this paper, we provide a spectrum sensing method based on noisy activation deep neural network (NA-CGNet) for a SU in CR, where the SU does not need a prior knowledge about the PU or noise power but uses the raw received signal data to achieve spectrum sensing. The proposed algorithm flow is shown in Fig. 1, including offline training and online sensing. First, the received I/Q signals of the primary user are sampled and pre-processed with energy normalization, and then the normalized dataset is input to the NA-CGNet model proposed in this paper for network model training and validation to obtain a well-trained model. In online sensing detection, the SU received signal testset is input to the well-trained NA-CGNet model to obtain the corresponding probabilities under H_1 and H_0 to make a spectrum sensing decision.

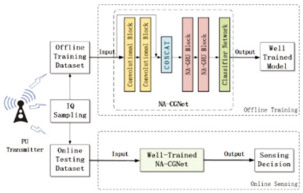


Fig. 1. Flow of the spectrum sensing algorithm proposed in this paper

Figure 2 shows the proposed NA-CGNet network structure, we treat the received signal as a $2 \times 128 \times 1$ image and process it first with a CNN-based model, 1D CNN is first used to extract spatially relevant features of the input time-series received signal, so as to obtain sufficient local features. Among them, the 1D CNN consists of two convolutional blocks, each of which includes a Conv1D layer, a regularized Dropout layer that prevents overfitting of the network and enhances the network generalization ability, and a modified line unit (ReLU) activation function that increases the nonlinearity between the layers of the network. Since functional information can be lost during network transmission, a cascade layer (CONCAT) is used to combine the input of the first convolutional block and the output of the second convolutional block in order to compensate for the loss of features.

1D CNN performs well in local feature extraction, but it lacks the ability to connect the long-term dependencies of the input time series. GRU is introduced for processing time series data to fully extract the temporal global correlation features. Compared with LSTM, GRU has a simpler network structure and is easier to train, which can largely

improve the training efficiency, and most importantly, GRU can achieve comparable results to LSTM. By adding normal noise to the zero-gradient part of the activation function the NA-GRU block, the added noise can grow according to the degree of saturation of the activation function, so that more noise points in the saturated state can be randomly explored toward the unsaturated state to make more attempts in the training convergence process, thus the network is noise-resistant at low SNR to achieve good sensing performance.

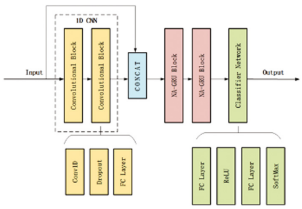


Fig. 2. Network structure of NA-CGNet

The signal processed by the NA-GRU block is fed to the final classification network, which consists of two FC layers and each FC layer maps the input features to a low-dimensional vector. As for the activation function, the first FC layer uses Relu and the last FC layer uses SoftMax to obtain the classification results. We map the output to a binary classification vector normalized by the SoftMax function, which represents the probability of the presence and absence of PU. The network parameters are optimized using the Adam optimizer, the classification cross-entropy is the loss function used, the initial learning rate is set to 0.0003, and the Dropout ratio is kept at 0.2 to obtain the best hyperparameters. The above network is called "NA-CGNet" and the hyperparameters confirmed by extensive cross-validation are detailed in Table 1.

Table 1. Hyperparameters of the proposed NA-CGNet

Hyperparameter	NA-CGNet
Filters per Conv layer	60
Filter size	10
Cells per GRU layer	128
Neurons per FC layer	128 & Sample length & 2
Initial learning rate	0.0003
Batch size	200

4 Experimental Analysis

4.1 Dataset Generation and Pre-processing

Based on the baseline dataset RadioML2016.10a [16], which is widely used in modulation identification tasks, eight digital modulation signals are generated as positive samples, and they consist of baseband I/Q signal vectors and common radio communication practical channel effects are took into account. According to the signal energy and the required signal-to-noise ratio level, the corresponding AWGN conforming to the CSCG distribution is generated as negative samples. The (SNR) varies from -20 dB to 5 dB in 1 dB increments. For each modulated signal type, 1000 modulated signal positive samples and 1000 noise samples are generated at each SNR. The whole dataset is divided into three different sets with a common split ratio of 3:1:1. The dataset parameters are detailed in Table 2 below.

Table 2. Dataset parameters

Parameter	Value
Modulation scheme	BPSK, QPSK, 8PSK, PAM4, QAM16, QAM64, GFSK, CPFSK
SNR range	−20 dB~5 dB in 1 dB increments
Sample length	64,128,256,512
Training samples	30000
Validation samples	10000
Testing samples	10000

Considering that the energy interference of the signal has minimal impact and the modulation structure of the signal can be better exposed, the received time-domain complex signal $X_{in} = [X_{real}; X_{imag}]$ data is pre-processed with energy normalization before the network training. It works as follows.

$$X_i^{norm} = \frac{X_i}{\sum (abs(X_i)^2)} \tag{4.1}$$

where X_i denotes the i-th sequence of the sampled signal and $abs(X_i)$ denotes the absolute value of X_i.

4.2 Experimental Configuration and Network Training

The experiment uses Tensorflow as the framework and the Python language for programming. The operating system is Ubuntu 18.04. GNU Radio is a free open source software development kit that provides signal processing modules to implement software radio and wireless communication research.

Considering the necessity of setting a constant false alarm probability P_f, a two-stage training strategy is used. In the first stage, convergence is achieved by training the

model with early stopping, so that both validation loss and accuracy are kept stable. A stopping interval of P_f is first set in the second stage, and stops when P_f falls into this interval. Applying a two-stage training strategy can control the detection performance to some extent by adjusting the preset stop intervals.

4.3 Simulation Results

In this section, extensive simulation results are provided to demonstrate the performance of the proposed model. For GFSK modulated signals with sample length 128, the sensing accuracy of dif-ferent neural network models CNN, LSTM, DetectNet and the proposed model NA-CGNet with different SNRs were compared as shown in Fig. 3, it can be seen that the NA-CGNet model proposed in this paper provides better sensing accuracy than other networks, especially the accuracy improvement of NA-CGNet is more obvious at low SNR level. The NA-CGNet model achieves a sensing accuracy P_d=0.5239 at SNR $= -15$ dB which is 0.042 better than the DetectNet model while ensuring a lower P_f=7.06%, and has the highest sensing accuracy P_d at SNR $= -20$ dB than other models. This is due to the proposed NA-CGNet model adding noisy activation function to the GRU block, which enables further stochastic exploration of the network in a saturated state, more robust to noise in a low SNR environment, and a corresponding improvement in sensing performance.

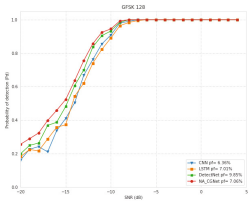


Fig. 3. Comparison of the sensing performance of different models

The generalization of the proposed NA-CGNet model is demonstrated by varying the modulation characteristics of the PU signal. Figure 4 shows the comparison of the sensing performance of NA-CGNet on eight different digital modulation schemes with sample length of 128 and SNRs ranging from -20 dB to 5 dB. It can be observed that the difference in sensing performance between the various modulated signals is negligible, which implies that the NA-CGNet model performance is insensitive to the modulation order. In addition, it is observed that the difference in sensing performance between the same type of modulation such as BPSK, 8PSK and QPSK is very small, which indicates the robustness of the proposed NA-CGNet model for different modulation schemes of radio signals.

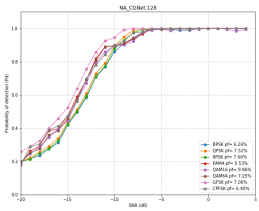


Fig. 4. Sensing performance of NA-CGNet model with different modulation schemes

5 Conclusion

In this paper, a deep neural network model for spectrum sensing in a low SNR environment is proposed. The 1D CNN and the GRU network with NAF respectively capture sufficient local features and global correlations of the input SU received signal, while the cascade layer helps to compensate for feature loss. Adding noise activation helps to improve the anti-noise performance of the network, which enables the network model to maintain good sensing performance at low SNR level. Simulation results show that the proposed model performs significantly better than CNN, LSTM, and DetectNet. in addition, the proposed NA-CGNet model is applicable to various modulation signal schemes, such as GFSK, BPSK, QPSK, QAM16, etc., and has good robustness to different modulation scheme signals. Finally, the proposed NA-CGNet model can achieve better spectrum sensing performance in low SNR environment which is shown to achieve high P_d and low P_f simultaneously.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (61761034), the Natural Science Foundation of Inner Mongolia (2020MS06022).

References

1. Arjoune, Y., Kaabouch, N.: A comprehensive survey on spectrum sensing in cognitive radio networks: recent advances, new challenges, and future research directions. Sensors **19**(1), 126 (2019)
2. Yucek, T., Arslan, H.: A survey of spectrum sensing algorithms for cognitive radio applications. IEEE Communications Surveys & Tutorials **11**(1), 116–130 (2009)
3. Upadhye, A., Saravanan, P., Chandra, S.S., Gurugopinath, S.: A survey on machine learning algorithms for applications in cognitive radio networks. In: 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT). pp. 01–6 (2021)
4. Awin, F., Abdel-Raheem, E., Tepe, K.: Blind spectrum sensing approaches for interweaved cognitive radio system: a tutorial and short course. IEEE Communications Surveys & Tutorials **21**(1), 238–259 (2019)
5. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
6. Schmidhuber, J.: Deep learning in neural networks: an overview. Neural Netw. **61**, 85–117 (2015)
7. Zheng, S., Chen, S., Qi, P., Zhou, H., Yang, X.: Spectrum sensing based on deep learning classification for cognitive radios. China Commun. **17**(2), 138–148 (2020)

8. Xie, J., Liu, C., Liang, Y.C., Fang, J.: Activity pattern aware spectrum sensing: a CNN-based deep learning approach. IEEE Commun. Lett. **23**(6), 1025–1028 (2019)
9. Soni, B., Patel, D.K., Lopez-Benitez, M.: Long short-term memory based spectrum sensing scheme for cognitive radio using primary activity statistics. IEEE Access. **8**, 97437–97451 (2020)
10. Gao, J., Yi, X., Zhong, C., Chen, X., Zhang, Z.: Deep learning for spectrum sensing. IEEE Wireless Commun. Lett. **8**(6), 1727–1730 (2019)
11. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
12. Cho, K., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. 2014 Jun 3 [cited 2022 May 14]. https://arxiv.org/abs/1406.1078v3
13. Gülçehre, Ç., Moczulski, M., Denil, M., Bengio, Y.: Noisy Activation Functions. ICML (2016)
14. Quan, Z., Cui, S., Sayed, A.H., Poor, H.V.: Optimal multiband joint detection for spectrum sensing in cognitive radio networks. IEEE Trans. Signal Process. **57**(3), 1128–1140 (2009)
15. Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D.J.: 1D convolutional neural networks and applications: a survey. Mech Syst Signal Proc. **151**, 107398 (2021)
16. O'Shea, T.J., West, N.: Radio machine learning dataset generation with GNU radio. In: Proceedings of the GNU Radio Conference [Internet]. 2016 [cited 2022 Jun 23]. https://xueshu.baidu.com/usercenter/paper/show?paperid=1b963c825ef5ed29011305f8305cc727&site=xueshu_se&hitarticle=1

Analysis and Optimization of Cogging Torque in Line-Start Permanent Magnet Synchronous Machine by Changing the Rotor Teeth and Slots

Jing He[1(✉)], Ronaldo Juanatas[2(✉)], and Jin-lin Huang[1,2]

[1] Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241002, Anhui, China
0121000098@ahcme.edu.cn
[2] Technological University of the Philippines, Manila, Philippines

Abstract. When the windings of permanent magnet machine (PMM) are not energized, the interaction between the permanent magnet and the armature core generates cogging torque. It will cause vibration, noise and speed fluctuation. As for line-start PMM, both the stator and rotor core are slotted, and the manufacturing tolerances have influenced on cogging torque. The analytical expression of cogging torque of line-start PMM is obtained by using the equivalent method of rotor permanent magnets magneto motive force (MMF). Based on the analytical expression, the influence of rotor slots processing technic, magnetic bridge manufacturing parameters as well as permanent magnets technology on cogging torque are analyzed. Furthermore, the optimum combination of these three factors is deduced to minimize the cogging torque of line-start PMM. The finite element method (FEM) calculation verifies that the proposed method can reduce the cogging torque of LSPMSM.

Keywords: Cogging torque · Manufacturing tolerances · Line-start PMSM · Optimization

1 Introduction

Self-starting permanent magnet synchronous motor (LSPMSM) is receiving more and more attention. The main reason is that, firstly, high-energy rare earth permanent magnet (PM) materials can provide high air-gap magnetic flux density, thus reducing copper loss and providing high power factor [1]. At the same time, LSPMSM has the self-starting ability and high efficiency, making it more suitable for pump sets, compressors and some energy-saving applications [2, 3].

The interaction between the permanent magnet motor and the stator slot will occur when the motor is unloaded, resulting in cogging torque. It will cause vibration, noise and speed fluctuation. For linear starting PMM, the slot and tooth parameters of stator and rotor will have a great impact on the cogging torque [4, 5]. In order to reduce the cogging torque, the influence on of the cogging torque for stator tooth width have been studied in [6–8]. In [9], the cogging torque have been weakened by the unequal stator

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 251–260, 2023.
https://doi.org/10.1007/978-981-99-1260-5_32

tooth width. However, the above research is limited to single-side slotted PMM, and there is little literature on the analysis of cogging torque of line-start PMSM.

The stator slot and rotor slot are existed in LSPMSM, and the distribution of effective air-gap length is more complex compared with the traditional PMSM when the relative position of the stator and rotor is changed. The traditional method of reduce cogging torque cannot apply to the LSPMSM directly. The distribution of effective air-gap of LSPMSM is obtained through the superimposed the effective air-gap of the stator side and rotor side, and then the analytical expression of cogging torque is obtained in [10, 11]. In addition, the generation mechanism of LSPMSM cogging torque is discussed. However, the expression used to analyse the cogging torque is extremely complex [12, 13].

Based on the energy method and Fourier decomposition method, the analytical expression of cogging torque of linear starting permanent magnet synchronous motor is derived in this paper, including rotor slot offset angle, magnetic bridge as well as permanent magnets width. Based on the analysis above, the optimum combination of these three parameters are proposed to minimize cogging torque.

2 Analytical Model of Lspmsm Cogging Torque

According to the mechanism of cogging torque, and the cogging torque can be followed by:

$$T_{\text{cog}} = -\frac{\partial W}{\partial \alpha} \tag{1}$$

where W is magnetic field energy, and α is relative position angle of the stator and rotor.

Suppose the magnetic permeability of the rotor core is infinite, and the magnetic field energy in the stator and rotor core can be neglected compared with the magnetic field energy in the air-gap. Therefore, the magnetic field energy can be given by:

$$W \approx W_{\text{airgap}} = \frac{1}{2\mu_0} \int_V B^2(\theta, \alpha) dV \tag{2}$$

where μ_0 is the air permeability, $B(\theta, \alpha)$ is the distribution of air-gap flux density along the circumferential direction as the position of the stator and rotor changes.

Here, the distribution of air-gap flux density can be followed by:

$$B(\theta, \alpha) = \mu_0 \frac{F(\theta)}{\delta(\theta, \alpha)} \tag{3}$$

where $\delta(\theta, \alpha)$ is the effective air-gap length of versus the position of stator and rotor core.

There are many kinds of magnetic circuit structures of LSPMSM, and the interior structure can improve the power density. Figure 1 shows the configuration of spoke-type LSPMSM. It can be seen that the structure is different from the traditional permanent magnet synchronous machines (PMSMs). Due to the rotor slot is existed in LSPMSM rotor core, and the distribution of effective air-gap length is more complex compared

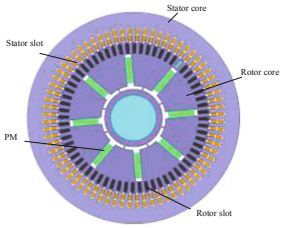


Fig. 1. Relative position of the stator and rotor

with traditional PMSMs. In order to reduce the difficulty of cogging torque analytical expression, and the rotor magnetomotive force (MMF) is represented by a distributed MMF.

Hence, the distribution of effective air-gap length is only related to the stator slots and teeth. The magnetic field energy can be followed by.

$$W \approx W_{\text{airgap}} = \frac{\mu_0}{2} \int_V F^2(\theta, \alpha) \frac{1}{\delta^2(\theta, \alpha)} dV \tag{4}$$

Furthermore, when the stator and rotor leakage flux is ignored, and the air-gap MMF $F(\theta)$ is produced rotor is shown in Fig. 2. As shown in Fig. 2, p is the pole pairs, τ_r is the distance of rotor tooth, θrt is the rotor tooth, F is the amplitude of air-gap MMF. As a result, the distribution map of $F^2(\theta)$ is shown in Fig. 3, and the Fourier expansion can be expressed as.

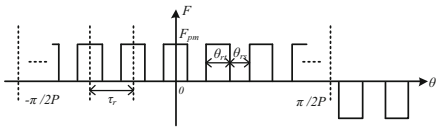


Fig. 2. Distribution of the equivalent MMF of rotor

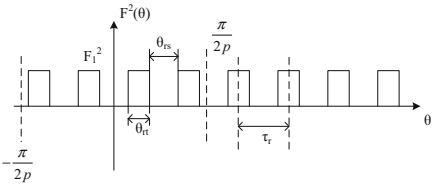


Fig. 3. Distribution of F2(θ)

$$F^2(\theta, \alpha) = F_0 + \sum_{n=1}^{\infty} F_n \cos 2np\theta \tag{5}$$

where, the Fourier decomposition coefficient F_0 can be expressed as

$$F_0 = \frac{(Q_2 - 1)pF^2}{\pi} \theta_{rt} \tag{6}$$

$$F_n = \begin{cases} (-1)^{n+1} \frac{2F_1^2}{n\pi} \sin np\theta_{ra}, & n \text{ is not an integer multiple of } Q_2 \\ (-1)^n \frac{2(Q_2-1)F_1^2}{n\pi} \sin np\theta_{ra}, & n \text{ is an integer multiple of } Q_2 \end{cases} \tag{7}$$

where Q_2 is number of rotor slots, p is pole paris.

Moreover, $\frac{1}{\delta^2(\theta,\alpha)}$ is only related to stator slot and tooth, and the Fourier expansion can be expressed as.

$$\frac{1}{\delta^2\theta}=G_0+\sum_{n=1}^{\infty}G_n\cos nQ_1\theta \tag{8}$$

$$\frac{1}{\delta^2\theta,\alpha}=G_0+\sum_{n=1}^{\infty}G_n\cos nQ_1(\theta+\alpha) \tag{9}$$

where Q_1 is number of stator slots, and the Fourier decomposition coefficient G0 can be expressed as.

$$G_0=\frac{Q_1\tau_s}{2\pi\delta^2} \tag{10}$$

$$G_n=\frac{2}{n\pi\delta^2}\sin\frac{nQ_1\tau_s}{2} \tag{11}$$

where τ_s is the stator tooth width.

At last, the LSPMSM cogging torque expression can be followed by.

$$T_{cog}=\frac{\pi Q_1 L_{Fe}}{4\mu_0}(R_2^2-R_1^2)\sum_{n=1}^{\infty}nG_nB_{r\frac{nQ_1}{Q_2}}\sin nQ_1\alpha \tag{12}$$

where L_{fe} is the armature core axial length, R_1 is the outer radius of the rotor, and R_2 is the inner radius of the stator, n is an integer which $nQ_1/2p$ is integer.

According to the formula 12, the number of rotor slot, rotor tooth width, air-gap length, he tooth width and slot width of the stator will affect Gn and the amplitude of the cogging torque. Based on this, this paper studies these parameters, and realizes the purpose of reducing the cogging torque by optimizing the parameters of stator tooth width, slot width and the number of stator slots.

3 Method to Reduce Cogging Torque

As the rotor tooth width of LSPMSM changes to θ_{ra}, the distribution map of $F^2(\theta)$ is shown in Fig. 4. So the amplitude of $F^2(\theta)$ is changed as.

$$F_1^2=\frac{\theta_{\mathrm{rs}}^2}{\theta_{ra}^2}F^2 \tag{13}$$

In addtion, the Fourier decomposition coefficient $F^2(\theta)$ is changed as.

$$F_0=\frac{(q_2-1)pF_1^2}{\pi}\theta_{ra} \tag{14}$$

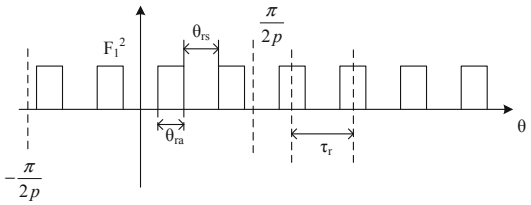


Fig. 4. Distribution of $F^2(\theta)$ as the rotor tooth is θ_{ra}

The cogging torque is related to $F_{nQ1}/2p$, and let Fourier decomposition coefficient is 0.

$$\sin\frac{t_0}{t_2}\frac{nQ_1}{Q_2}\pi = 0 \tag{15}$$

$$\frac{t_0}{t_2}\frac{nQ_1}{Q_2}\pi = k\pi,\ k \text{ is an integer} \tag{16}$$

It can be obtained that when the rotor tooth width and the rotor tooth pitch meet $t_2/t_0 = nQ_2/kQ_1$, the cogging torque can be reduce effectively. Take the case of a 72 stator slot- 8 pole spoke-type LSPMSM, and the influence of rotor parameters on cogging torque is studied. In Table 1, the main parameters of the machine are listed.

Meanwhile, the comparison of cogging torque waveforms of LSPMSM with initial model and optimized model are shown in Fig. 5. Obviously, the cogging torque is reduced as the rotor tooth is 5°.

Table 1. Main parameters of prototype machine

Item	Value	Item	Value
Rated power/kW	30	Core material	BAT1500
PM material	N35EH	Air-gap length/mm	1
Pole	8	PM width/mm	48
Stator slot Q_1	72	PM thickness/mm	12
Outer diameter of stator /mm	400	Pressure coefficient	0.97
Outer diameter of rotor /mm	283.6	Rotor inner diameter /mm	85
Axial length/mm	225	Rated voltage/V	380

Change the rotor slots. According to the air-gap flux density Fourier decomposition, and the rotor slot will make the harmonic amplitude of the rotor MMF increases. In addition, the air-gap flux density harmonic content can be changed as the rotor slot number increases, and the cogging torque must change. In order to study the number of rotor slots is influenced on the cogging torque, and the FEA model is established with different rotor slots.

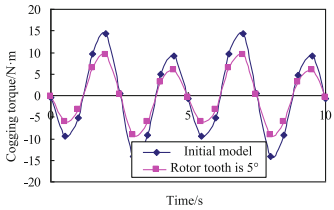


Fig. 5. Slot torque of different rotor tooth width

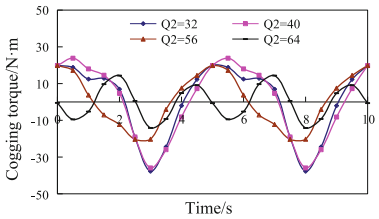


Fig. 6. Relationship between cogging torque and rotor slot number

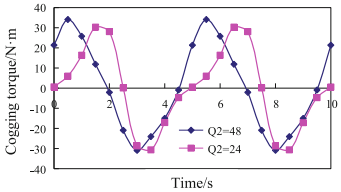


Fig. 7. Relationship between cogging torque and rotor slot number

Figure 6 and Fig. 7 shows the cogging torque versus number of rotor slot. It can be noted that the number of cycles and amplitude of cogging torque are changed with the number of rotor slots changes. The cogging torque of 32 rotor slot and 40 rotor slot is similar, and the cogging torque of 56 rotor slot is closed to 64 rotor slot. In order to further study the influence of rotor slot number on cogging torque, and the cogging torque of different rotor slots is shown in Table 2.

Table 2. Cogging torque with different rotor slot

Q2	Nq	Nq2p/Q2	COGGING TORQUE(N·m)
24	9	3	31
32	36	9	43
40	45	9	42
48	18	9	34
56	63	9	20
64	72	9	16

We could draw the following conclusions that: when the value of Nq2p/Q2 is larger, the amplitude of cogging torque is smaller. It is because of the value of Nq2p/Q2 is larger, and the flux density harmonic components of rotor tooth is lower.

Effect on the rotor tooth shape. The outer diameter of rotor is different from the inner diameter of stator when the shaped of rotor tooth is changed. There is an eccentric

distance between the rotor center and the stator center, and denoted as h_px, as shown in Fig. 8. It can be seen that the air-gap length is non-uniform, and the air-gap length changes with the change of rotor position. Therefore, the distribution of air-gap magnetic field along the circumference is changed, and the cogging torque is reduced.

The LSPMSM cogging torque versus eccentric distance are shown in Fig. 9. It can be seen that the cogging torque decreases with the increase of eccentric distance, and the amplitude of cogging torque from 9.5N·m reduce to 4.5N·m. However, when the eccentric distance reaches 15mm, the reduction of cogging torque is not obvious with the increase of eccentric distance. As a results, the eccentric distance is selected 15mm.

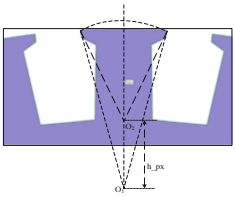


Fig. 8. The configuration of rotor tooth with eccentric

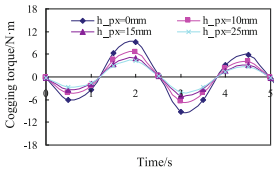


Fig. 9. The cogging torque versus eccentric distance

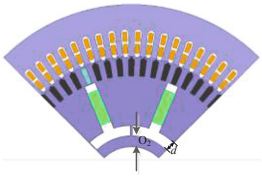


Fig. 10. Structure of magnetic bridge

Effect on the magnetic bridge size. The size of magnetic bridge can influence on the distribution of rotor magnetic field, so the cogging torque is effected on magnetic bridge size. The structure and size of magnetic bridge is shown in Fig. 10. The distance between magnetic bridge and shaft is defined as O_2, and the magnetic bridge width is defined as d.

Figure 11 shows the cogging torque and rated torque versus distance magnetic bridge and shaft. It can be seen that the cogging torque decreases first and then increases with the increase of distance between magnetic bridge and shaft. However, the rated torque keep constant with the increase of distance between magnetic bridge and shaft. Because of the air-gap flux density keep constant as the volume of PM is equal. Figure 12 shows the cogging torque and rated torque versus magnetic bridge width. It can be noted that the cogging torque decrease obviously first and then reduces with the increases of magnetic bridge width, and the rated torque increases and then decreases with the increase of magnetic bridge width.

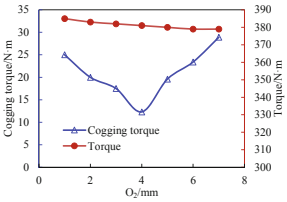


Fig. 11. The cogging torque versus eccentric distance

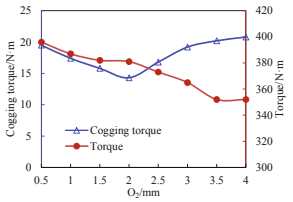


Fig. 12. The cogging torque versus eccentric distance

4 Performance of LSPMSM After Optimized

It can be seen from the above conclusion, and the cogging torque of LSPMSM can be weaken effectively for the rotor parameters are changed. However, the electromagnetic performance has been changed with the rotor parameters are changed. By using the time-step finite element method, the no-load back-EMF and rated electromagnetic torque of the prototype before and after optimization are calculated and compared, as shown in Table 3 and Table 4.

It can be noted that the amplitude of back-EMF no large change by adopting the measures of cogging torque reduction in this paper. In addition, the change of starting torque and starting-current of the prototype is also small. Therefore, the cogging torque reduction method in this paper can effectively weaken the cogging torque of the LSPMSMs, and will not have a great impact on the performance of the machine.

Table 3. Comparison of no-load performance of the prototype machine

Model	Back-EMF	THD	Cogging torque
Original Model	246.7 V	3.5%	16.2 N·m
Change rotor tooth	245.4 V	3.2%	10.5 N·m
Change rotor tooth shaped	235.6 V	2.7%	5.7 N·m
Unequal rotor tooth	237.2 V	3.0%	N·m

Table 4. Comparison of load performance of the prototype machine

Model	Stating-torque	Rated torque	Starting-current
Original Model	875 N·m	381 N·m	378 A
Change rotor tooth	889 N·m	376 N·m	387 A
Change rotor tooth shaped	905 N·m	372 N·m	399 A
Unequal rotor tooth	897 N·m	374 N·m	

5 Conclusions

LSPMSMs have been increased interest due to self-starting capability, high power factor and high efficiency. Cogging torque is a common problem to LSPMSMs, and it is one of the issues that should be paid attention to at the design stage. Meanwhile, LSPMSMs has rotor slots in the rotor core due to improve the self-starting capability, and the cogging torque will adversely the operation of LSPMSM. In order to comperhend the cogging torque of LSPMSM mechanism, and the analytical expression of the cogging torque with the changing the rotor tooth width, rotor slots and the rotor tooth shaped are deduced in this paper, then some conclusions can be obtained as follows:

(1) The interior PM rotor is equivalent to the MMF generated by the surface PM can simplify the cogging torque analytical model of LSPMSM.
(2) The number of rotor slots have influenced on the cogging torque of LSPMSM. When the larger $N_q 2p/Q_2$ is selected, the cogging torque of the LSPMSM can be reduced.
(3) Selecting appropriate rotor tooth width can effectively weaken the cogging torque of LSPMSM.
(4) The appropriate rotor eccentricity distance can further reduce the cogging torque of the LSPMSM.

References

1. Cheng, M., Hua, W., Zhang, G.: Overview of stator-permanent magnet brushless machines. IEEE Trans. Industr. Electron. **61**(8), 5087–5101 (2011)
2. Cheng, M., Zhang, G., Hua, W.: Overview of stator permanent magnet brushless machine systems and their key technologies. Proc. CSEE **34**(29), 5204–5220 (2014)
3. Liu, X.P., Zheng, AI.H., Wang, CH.: 3-D Finite element analysis and experiment study of a stator-separated axial flux-switching hybrid excitation synchronous machine, Transactions of china electro-technical society 27 (10) 2012, 110–113
4. Hwang, S., Eom, J., Jung, Y., Lee, D., Kang, B.: Various design techniques to reduce cogging torque by controlling energy variation in permanent magnet motors. IEEE Trans. Magn. **37**(4), 2806–2809 (2001)
5. Yongping, L., Yong, L.: Tooth-slot cogging torque and noise analysis of permanent magnet motors. In: Proceedings of the Fifth International Conference on Electrical Machines and Systems, Shenyang, pp. 860–862 (2001)
6. Li, J.T., Liu, Z.J., Jabbar, M.A.: Design optimization for cogging torque minimization using response surface methodology. IEEE Trans. Magn. **40**(2), 1176–1179 (2004)
7. Bianchini, C., Immovilli, F., Lorenzani, E., et al.: Review of design solutions for internal permanent-magnet machines cogging torque reduction. IEEE Trans. Magn. **48**(10), 2685–2693 (2001)
8. Islam, J., Svechkarenko, D., Chin, R., et al.: Cogging torque and vibration analysis of a direct-driven PM wind generator with concentrated and distributed windings. In: 2012 15th International Conference on Electrical Machines and Systems, Sapporo, pp. 1–6 (2012).
9. Xintong, J., Jingwei, X., Yong, L., et al.: Theoretical and simulation analysis of influences of stator tooth width on cogging torque of BLDC motors. IEEE Trans. Magn. **45**(10), 4601–4604 (2009)

10. Fei, W., Luk, P.C.K., Shen, J.: Torque analysis of permanent-magnet flux switching machines with rotor step skewing. IEEE Transactions on Magnetics 48(10), 2664–2673 (2012). https://doi.org/10.1109/TMAG.2012.2198223
11. Wang, D.H., Wang, X.H., Jung, S.Y.: Reduction on cogging torque in flux-switching permanent magnet machine by teeth notching schemes. IEEE Trans. Magn. **48**(11), 4228–4231 (2012). https://doi.org/10.1109/TMAG.2012.2200237
12. Hao, L., Lin, M.Y., Xu, D.: Cogging torque reduction in axial field flux-switching permanent magnet machines. Trans. China Electro-Tech. Soc. **30**(2), 21–26 (2015)
13. Hao, L., Lin, M.Y., Xu, D.: Cogging torque reduction of axial field flux-switching permanent magnet machine by adding magnetic bridge in stator tooth. IEEE Trans. Appl. Supercond. **24**(3), 503–506 (2014). https://doi.org/10.1109/TASC.2014.2302911

Beam Tracking Scheme Based on Dispersed Beams in mmWave Communication

Zhengqi Shen and Weixia Zou(✉)

Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
{szq2016,zwx0218}@bupt.edu.cn

Abstract. Benefiting from the large-scale antenna array used in millimeter wave massive MIMO systems, beams generated by beamforming technology have the characteristics of high directivity and array gain. However, due to the narrow width of the millimeter wave beam and the mobility of dynamic users such as vehicles, high-speed trains and drones, the communication beams between BS and UE will be misaligned and slight misalignment will result in severe gain loss, thus beam tracking is the key technology to solve this problem. We adopt Delay-Phase Precoding (DPP) structure in mmWave communication, which can generate frequency-dependent beams by jointly controlling the delay and phase, extending the dimensions to control beams. Thus the beam generated by MIMO-OFDM can be dispersed into a set of sub-beams and expand the coverage. Based on this we implement a new beam tracking scheme in mmWave communication, increasing the accuracy and convergent speed.

Keywords: Millimeter wave communication · Beam tracking · Delay-phase precoding · OFDM

1 Introduction

Millimeter wave communications have been exploited due to the extremely high frequency and abundant spectrum resources. Millimeter wave denotes the electromagnetic wave that has a wavelength range from 1–10 mm. It indicates that the frequency of the wave ranges from 30–300 GHz. Compared to Sub-6 GHz communications, millimeter waves might not be diffracted easily but the energy attenuation is more severe. In phased array antennas, beamforming technology can be used to generate highly directional beams, while millimeter-wave-based massive Multiple Input Multiple Output (MIMO) systems can generate narrower beams, which can effectively enhance array gain and compensate for severe attenuation. However, high directivity means narrow beam width. When the user moves, beam misalignment is more likely to occur, resulting in a sharp drop in array gain [1]. Therefore, beam tracking technology that can achieve real-time beam alignment is crucial.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 261–270, 2023.
https://doi.org/10.1007/978-981-99-1260-5_33

Beam training, also known as beam search, is a basic beam tracking method. For the sparsity and large dimension of mmWave channels, beam training can effectively obtain channel state information (CSI) [2]. Exhaustive search is the most common beam training method, which enumerates the transmit beams in the entire space and compares the received signal power to obtain the optimal

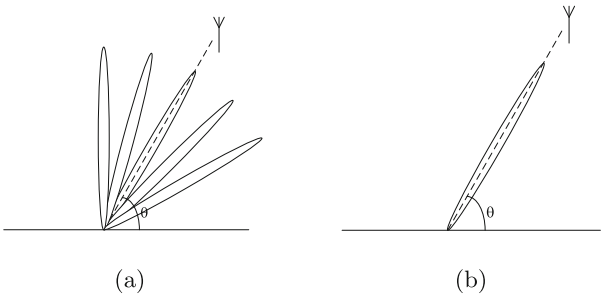


Fig. 1. Beamforming in THz communication. (a) Traditional beamforming with beam split effect. (b) Beam split effect eliminated with DPP structure.

beam. Hierarchical search divides the search range into multi-level. For example, IEEE 802.15.3C adopts a three-layer search strategy of quasi-omnidirectional level, sector level and beam level [3], while IEEE 802.11.ad adopts Sector Level Sweep (SLS) and Multiple sector ID capture (MIDC) two-layer strategy [4]. However, the high dynamic of the user will cause the channel to change rapidly, resulting in a high overhead for beam training. Therefore, various methods for beam tracking have been designed. Beam tracking based on Kalman Filter is a common method. [5] uses Extended Kalman Filter (EKF) for beam tracking, which has high tracking accuracy. Besides, [5] draws inspiration from radar communication and proposes beam tracking based on Auxiliary Beam Pairs (ABP), which can track multipath channels. With the rapid development of Artificial Intelligence (AI) technology in recent years, beam tracking based on reinforcement learning has also become popular. [6] uses the Multi-Armed Bandit (MAB) which models the behavior and state of the beam can effectively utilize the historical information in the tracking process. In this paper, inspired by the beam split effect in the terahertz (THz) band [7] and its solution [8], we study the millimeter-wave beam tracking algorithm under MIMO-OFDM systems. Beam split refers to the fact that in THz band, due to the ultra-high frequency and ultra-wide frequency band, the beam that is supposed to be pointed in a specific direction in the MIMO-OFDM system is split into a set of beams pointing in different directions with different sub-carrier frequencies (Fig. 1), resulting in a loss of array gain and a decrease in effective achievable sum rate [7].

Different from the traditional beam tracking algorithm, DPP structure introduces the dimension of delay as a new parameter and performs precoding in conjunction with phase, so that the antenna array can generate frequency-dependent

beams, while the traditional hybrid precoding structure can only generate frequency-independent beams. In the MIMO-OFDM system, this frequency-dependent characteristic can realize the manipulation of beams of different frequency sub-carriers. In addition to solving beam split, it can also freely control the degree of beam dispersion. [9] achieves simple beam tracking under this controlling idea, however, it's based on the beam split phenomenon in THz band. Although there is no beam split phenomenon in mmWave communication, through our reverse divergent thinking, we utilize DPP structure in mmWave MIMO-OFDM system to disperse the beam which points to a specific direction (Fig. 2), then implement the beam tracking scheme. Simulation results show the proposed algorithm has good performance in effective achievable sum rate and MSE of tracking result.

2 System Model

We consider the millimeter wave MIMO-OFDM communication system [10], and use the recently proposed DPP structure to implement beamforming. The transmitter is equipped with N antennas Uniform Linear Array (ULA) and N_{rf} Radio Frequency (RF) chains to serve single-antenna users. M subcarriers OFDM system is used over the bandwidth B.

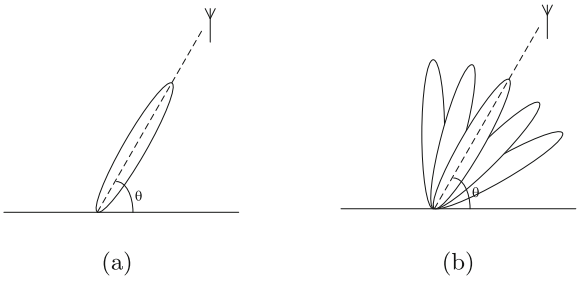


Fig. 2. Beamforming in mmWave communication. (a) Traditional beam generated by hybrid precoding structure. (b) Dispersed beams set generated by DPP structure.

2.1 Channel Model

Due to the sparsity of mmWave channels, we consider the extended Saleh-Valenzuela geometric channel model [11] in this paper. At subcarrier $m(m = 1, 2, \ldots, M)$, the channel vector between BS and UEs is denoted as

$$\boldsymbol{h}_m = \sum_{l=0}^{L-1} \boldsymbol{\beta}_m^{(l)} \mathbf{a_N}\left(\boldsymbol{\psi}_m^{(l)}\right), \tag{1}$$

L is path number. $\beta_m^{(l)} = g_m^{(l)} e^{-j\pi\tau_m^{(l)} f_m}$, $g_m^{(l)}$ and $\tau_m^{(l)}$ represent the path gain and delay under the $l{-}th$ path and the $m{-}th$ subcarrier respectively, f_m is the carrier

frequency where the $m-th$ subcarrier located in, the relation between center frequency f_c and f_m are $f_m = f_c + \frac{B}{M}\left(m - 1 - \frac{M-1}{2}\right)$. $\psi_m^{(l)}$ denotes the spatial direction under $l-th$ path and $m-th$ subcarrier. $\theta^{(l)}$ represents the related physical direction under $l-th$ path, $\psi_m^{(l)} = \frac{2d}{c}\sin\widetilde{\theta}^{(l)}$, d is antenna distance. For simplicity, denote $\boldsymbol{\theta}^{(l)} = \sin\widetilde{\theta}^{(l)}$ is used to represent the user physical direction. The array response vector $\mathrm{a_N}\left(\psi_m^{(l)}\right)$ is presented as

$$\mathrm{a_N}\left(\boldsymbol{\psi}_m^{(l)}\right) = \frac{1}{\sqrt{N}}\left[1, e^{j\pi\psi_m^{(l)}}, e^{j\pi 2\psi_m^{(l)}}, \cdots, e^{j\pi(N-1)\psi_m^{(l)}}\right]^T. \tag{2}$$

2.2 Delay-Phase Precoding Structure

The existing beam tracking schemes are typically based on analog beamforming or hybrid precoding structure, both are based on phased array, which adjusts the phase of each antenna to generate beams pointing in specific directions. However, these beams are frequency-independent and cannot fully utilize the characteristic that different subcarriers located in different frequencies in MIMO-OFDM system. In the DPP structure, a set of delay networks is introduced between the RF chains and the phase shift network [8], so that the beamforming vector is affected by the two parameters of delay and phase shift at the same time, generating frequency-dependent beams, and making full use of the frequencies of different subcarriers.

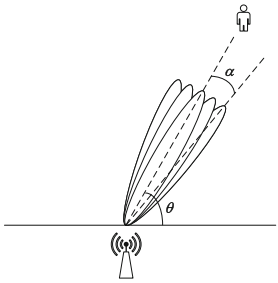


Fig. 3. DPP generated beams covering $\left[\theta_k^{(0)} - \alpha_k, \theta_k^{(0)} + \alpha_k\right]$ with $M = 5$.

In DPP structure, each RF chain is connected to K_d time-delayers (TDs), and those TDs are connected to all antenna elements. More specifically, each TD is linked to N/K_d antenna elements. The number of RF chains is equal to the number of users, so $N_{rf} = K$. Thus the received signal related to $m-th$ subcarrier is denoted as

$$\mathbf{y}_m = \mathbf{H}_m^H \mathbf{A}_m \mathbf{D}_m \mathbf{s} + \mathbf{n}, \tag{3}$$

the channel matrix between BS and K UEs is denoted as $\mathbf{H}_m = [\mathbf{h}_{1,m}, \mathbf{h}_{2,m}, \cdots, \mathbf{h}_{K,m}]$. $\mathbf{A}_m \in C^{N\times K}$ is the analog beamformer generated by

the TDs and PSs jointly, and $\mathbf{D}_m \in C^{K\times K}$ is the digital precoder with transmitting power constraint $\left\|\mathbf{A}_m\mathbf{D}_{m,[:,k]}\right\|_F \leq \rho$, in which ρ is the power for each user. n represents the Gaussian noise.

Because of the delay network which does not exist in the traditional hybrid precoding structure, the analog beamformer under the DPP structure consists of two parts, $\mathbf{A}_m = \mathbf{A}^{(\mathrm{PS})}\mathbf{A}_m^{(TD)}$. On the one hand, $A^{(\mathrm{PS})}$ is the traditional analog beamformer realized by PSs, which controls the center beam to point in a given direction. On the other hand,

$$\mathbf{A}_m^{(\mathrm{TD})} = \mathrm{diag}\left(\left[e^{-j2\pi f_m \mathrm{t}_1}, e^{-j2\pi f_m \mathrm{t}_2}, \cdots, e^{-j2\pi f_m \mathrm{t}_K}\right]\right), \tag{4}$$

is the analog beamformer realized by TDs, $\mathrm{t}_i \in C^{K_\mathrm{d}\times 1}$ is the delay generated by the K_d TDs linked to the $i-th$ RF chain.

Let the beamforming matrix $\mathbf{A}_m = [\mathbf{f}_{1,m}, \mathbf{f}_{2,m}, \cdots, \mathbf{f}_{K,m}]$, the $k-th$ column of $\mathbf{A}_m$ represents the beam pointing to the $k-th$ user under $m-th$ subcarrier, which is denoted as

$$\mathbf{f}_{k,m} = \mathbf{A}_k^{(\mathrm{PS})} e^{-j2\pi f_m t_k} \tag{5}$$

$\mathrm{A}_k^{(\mathrm{PS})}$ is the $k-th$ column of $\mathrm{A}^{(\mathrm{PS})}$. The decisive parameters of $\mathbf{f}_{k,m}$ is as below

$$\mathbf{A}_k^{(\mathrm{PS})} = \mathrm{blkdiag}\left(\mathbf{a}_P\left(\theta_k^{(0)} + (1-\xi_1)\,\alpha_k\right) e^{j\pi\left(P\left(\theta_k^{(0)}+(1-\xi_1)\alpha_k\right)+2s_k\right)\mathbf{p}^T(K_\mathrm{d})}\right) \tag{6}$$

$$\mathbf{t}_k = \frac{1}{f_c} s_k \mathbf{p}\left(K_\mathrm{d}\right), \tag{7}$$

$1/f_c$ is the period of subcarrier, and s_k is the number of periods that delayed via TDs, $\mathbf{p}\left(K_\mathrm{d}\right) = [0, 1, \cdots, K_\mathrm{d}-1]^T$, $\boldsymbol{\theta}_k^{(0)}$ is the physical direction of $k-th$ user under LOS path. Denotes $\xi_m = \frac{f_m}{f_c}$, according to the *Lemma 1* in [9], when we set $s_k = -\frac{P}{2}\left(\boldsymbol{\theta}_k^{(0)} + (1-\xi_m)\,\alpha_k + \frac{2\xi_M\xi_1\alpha_k}{\xi_M-\xi_1}\right)$, beams cover the range $\left[\theta_k^{(0)} - \alpha_k, \theta_k^{(0)} + \alpha_k\right]$ can be generated, each sub-beam points in the direction $\bar{\theta}_{k,m} = \theta_k^{(0)} + (1-\xi_1)\,\alpha_k + \frac{2\xi_M\xi_1(\xi_m-1)}{\xi_m(\xi_M-\xi_1)}\alpha_k$, which is illustrated in Fig. 3.

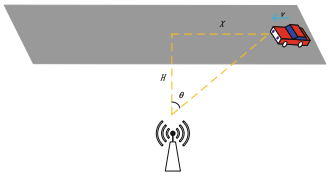


Fig. 4. Model of base station vehicle communication.

3 Beam Tracking Based on Dispersed Beams

In this section, a scenario where vehicles on the road communicating with BS is designed. As shown in Fig. 4, we assume that the distance between BS and the road is H, the vehicle moves forward at an estimated speed of V from X away, but there are limited fluctuations in the angle between the car and the base station during the process. The proposed algorithm uses the angle as the calculation unit, and only needs to calculate the boresight through the geometric distance in the scenario, then model the physical direction of the UE and the BS. We can divide one frame into T_{max} time slots at most, and the number of time slots T used to reach the tracking requirement is the training overhead.

Algorithm 1: Beam tracking scheme based on dispersed beams

Input: Road length X; Distance of BS and road H, Supposed vehicle linear speed V, Max vehicle base station angle fluctuation $\Delta\theta_k$, Total training slot T.

Initialization.

$\theta_{k,0}^{(0)} = \sin\left(\arctan\left(\frac{-X}{H}\right)\right)$

for $i = 1 : time$ **do**

  $\hat{\theta}_{k,i+1}^{(0)} = H\cot\left(\theta_{k,i}^{(0)}\right) + Vi$

  for $t = 1 : T$ **do**

    $\theta_{k,i+1,cen} = \hat{\theta}_{k,i+1}^{(0)} - \Delta\theta_k + \frac{(2t-1)\Delta\theta_k}{T}$

    $s_k^{(t)} = -\frac{P}{2}\left(\theta_{k,i+1,cen} + \frac{(1-\xi_1)\Delta\theta_k}{T} + \frac{2\xi_M\xi_1\Delta\theta_k}{(\xi_M-\xi_1)T}\right)$

    $\mathbf{A}_k^{\mathrm{s},(t)} = \mathrm{diag}\left(\mathbf{a}_P\left(\theta_{k,i+1,cen} + \frac{(1-\xi_1)\Delta\theta_k}{T}\right)e^{j\pi\left(P\left(\theta_{k,i+1,cen}+\frac{(1-\xi_1)\Delta\theta_k}{T}\right)+2s_k^{(t)}\right)\mathbf{p}(K_\mathrm{d})}\right)$

    $\mathbf{t}_k = s_k^{(t)} T_c \mathbf{p}\left(K_\mathrm{d}\right)$

    $\mathbf{f}_{k,m}^{(t)} = \mathbf{A}_k^{\mathrm{s},(t)} e^{-j2\pi f_m \mathbf{t}_k^{(t)}}$

    $\mathbf{A}_m^{(t)} = \left[\mathbf{f}_{1,m}^{(t)}, \mathbf{f}_{2,m}^{(t)}, \cdots, \mathbf{f}_{K,m}^{(t)}\right]$

    $\mathbf{Y}_{m,t} = \mathbf{H}_m^H \mathbf{A}_m^{(t)} \boldsymbol{X}_m^{(t)} + \mathbf{N}^{(t)}$

  end

  $(t,m)_k = \underset{t\in 1,2,\cdots,T, m\in 1,2,\cdots,M}{\arg\max} \left\|\mathbf{Y}_{m,t,[k,i]}\right\|_2^2$

  $\theta_{k,i+1}^{(0)} = \theta_{k,i+1,\ \mathrm{cen}} - \Delta\theta_k + \frac{(2t_k-1)\Delta\theta_k}{T} + (1-\xi_1)\frac{\Delta\theta_k}{T} + \frac{2\xi_M\xi_1\left(\xi_{m_k}-1\right)}{\xi_{m_k}(\xi_M-\xi_1)T}\Delta\theta$

  Output: Physical direction $\theta_{k,i+1}^{(0)}$.

end

Assuming the physical direction of the $k-th$ user at $i-th$ frame is $\theta_{k,i}^{(0)}$, and the maximum possible range of motion of the $k-th$ user is $\Delta\theta_k$, and $\theta_{k,i+1}^{(0)} \in \left[\theta_{k,i}^{(0)} - \Delta\theta_k, \theta_{k,i}^{(0)} + \Delta\theta_k\right]$, our target is to calculate the physical direction of the user in the next frame $\theta_{k,i+1}^{(0)}$ within T_{max} time slots in the $i-th$ frame. We divide the angular range of $2\Delta\theta_k$ into T equal parts as the range that needs to be tracked for each time slot, that is, let $\alpha_k = \frac{\Delta\theta_k}{T}$. Then it can be calculated that in $t-th$ time slot at $i-th$ frame the dispersion beam set covers the range

$\left[\theta_{k,i}^{(0)} - \Delta\theta_k + \frac{(2t-2)\Delta\theta_k}{T}, \theta_{k,i}^{(0)} - \Delta\theta_k + \frac{2t\Delta\theta_k}{T}\right]$, whose center angle is $\theta_{k,i+1,cen}^{(0)} = \theta_{k,i}^{(0)} - \Delta\theta_k + \frac{(2t-1)\Delta\theta_k}{T}$, so the angle pointed by each sub-beam is represented as $\theta_{k,i}^{(0)} - \Delta\theta_k + \frac{(2t-1)\Delta\theta_k}{T} + (1-\xi_1)\frac{\Delta\theta_k}{T} + \frac{2\xi_M\xi_1(\xi_m-1)}{\xi_m(\xi_M-\xi_1)T}\Delta\theta_k$. The above beam sets are sequentially transmitted in all T time slots, then the corresponding received signal strengths are obtained, and the effective achievable sum rate is calculated.

The pseudocode of the proposed beam tracking scheme is summarized in *Algorithm* 1. In step 2, we calculate the physical direction at the $0-th$ frame through the initial relative positions H and X of the vehicle and BS. In step 4, the estimated angle at the next frame is calculated from the expected speed and current direction of the vehicle. In step 5-13, the beams are generated via DPP structure according to the estimated angle and the maximum angle fluctuation within T time slots and the received signals are calculated. In step 14, $(t,m)_k$ corresponding to the maximum received signal strength is obtained, and the label indicates the direction pointed by the subcarrier m_k in the $t_k - th$ time slot is the tracking result. Eventually we obtain the expected angle $\theta_{k,i+1}^{(0)}$ based on the label $(t,m)_k$ in step 15.

4 Simulation Result

In this section, the simulation result of the Mean Square Error (MSE) and Achievable Sum Rate (ASR) performance of the proposed algorithm are illustrated. The simulation parameters are set as follows: $\mathrm{N} = 128, \mathrm{M} = 32$, $\mathrm{K} = 4, f_c = 10\,\mathrm{Ghz}, \mathrm{B} = 1\,\mathrm{Ghz}, \mathrm{L} = 1$, all the simulation results are the average of 100 Monte Carlo realizations. We designed 4 different routes based on the scenario shown in Fig. 4, these routes are the trajectories of different users. A Cartesian coordinate system is established with the base station as the origin, and the starting points of the four routes are: $[(-200,100),(-400,50),(50,-100),(100,300)]$, and the expected vehicle speed is represented by vectors as: $[(32,0),(24,0),(0,16),(0,48)]$.

In Fig. 5, we compare MSE with the exhaustive search [12] and extended Kalman filter algorithms [13]. It can be seen from the result that the MSE of the traditional exhaustive search algorithm often fluctuates wildly, which indicates that the tracking error occurs frequently. Besides, the magnitude is larger than the other two algorithms. The tracking result of the EKF-based algorithm is very stable due to the complicated calculation process, but the MSE is still much higher than the proposed algorithm in this paper. Our proposed tracking algorithm based on dispersed beams has slight fluctuations in general, but there is no large fluctuation in adjacent moments, which shows the overall stability is relatively stable. Moreover, the MSE has the smallest order of magnitude, which means the best performance in these algorithms. In addition, we analyze the ASR

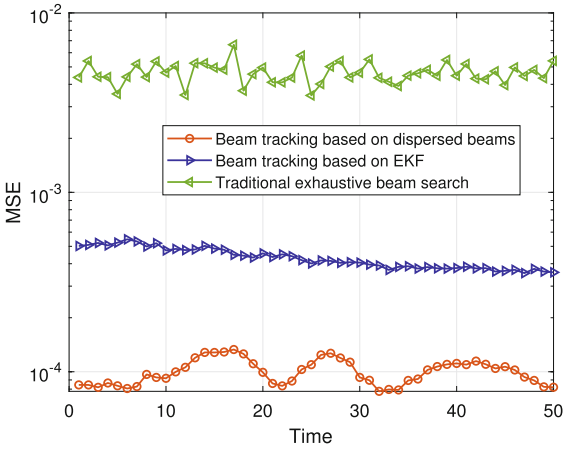
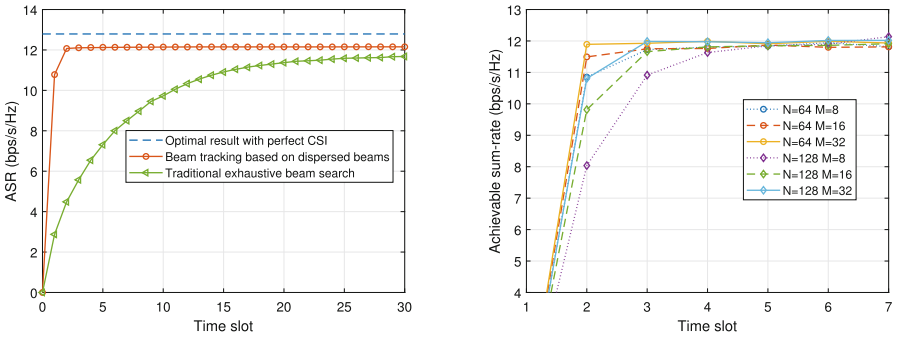


Fig. 5. MSE of different algorithms.

of traditional exhaustive search and the proposed algorithm, simulation results are shown in Fig. 6. The straight line at the top is the ASR that can be achieved by digital precoding to transmit the beam with perfect CSI, which means the real position of the vehicle as the beamforming target. This is equivalent to the theoretical upper bounds that beam tracking algorithms can achieve, thus it's an optimal result. The proposed algorithm only needs 4 time slots to reach the near-optimal result, then become stable in the future time slots. The traditional exhaustive search algorithm needs over 20 time slots to reach an acceptable ASR, and the stability is not good as the proposed algorithm.

Finally, we compare the ASR with different transmitter antenna numbers and subcarrier numbers, the results are shown in Fig. 6. In this simulation, we carried out a combined experiment on $N = [64, 128]$ and $M = [8, 16, 32]$. It's explicit to see that when $\mathrm{N} = 64, \mathrm{M} = 32$ it takes 2 time slots to converge and when $\mathrm{N} = 128, \mathrm{M} = 8$ it takes 4. This result indicates that when the ratio of N and M is larger, the algorithm reaches convergence faster. In massive MIMO system, the width of the beam only relates to the number of antenna elements. More antenna elements can generate narrower beams. When the coverage of the dispersed beam set to be generated for each time slot is set, narrower beams may lead to larger inter-beam spacing, which may result in inaccurate tracking results. In addition, since the number of subcarriers determines the number of sub-beams generated by the DPP structure, when the beam width and angle coverage is fixed, larger number of subcarriers means greater density, eventually we can get better tracking result with higher ratio.



(a) Different algorithms with N=128 M=32. (b) Different N and M in proposed algorithm.

Fig. 6. ASR simulation results in different situations.

5 Conclusion

In this paper, we proposed the beam tracking scheme based on the dispersed beams inspired by the solution of beam split effect in THz system. We use the DPP structure and apply it in mmWave communication by reverse thinking, dispersing the non-split mmWave beams to cover a specific range and implementing a beam tracking scheme. The simulation results show that the proposed algorithm has excellent MSE performance and can quickly converge to a near-optimal ASR. We also exploited that when the ratio of N and M is larger, the convergence is faster.

Acknowledgement. This work was supported by NSFC (No. 61971063), Key Lab of Universal Wireless Communications, MOE Beijing University of Posts and Telecommunications.

References

1. He, S., et al.: A Survey of millimeter-wave communication: physical-layer technology specifications and enabling transmission technologies. Proc. IEEE **109**(10), 1666–1705 (2021). https://doi.org/10.1109/JPROC.2021.3107494
2. Zhang, J., Masouros, C.: Beam drift in millimeter wave links: beamwidth tradeoffs and learning based optimization. IEEE Trans. Commun. **69**(10), 6661–6674 (2021). https://doi.org/10.1109/TCOMM.2021.3088896
3. IEEE Standards 802.15. 3cTM-Part 15.3: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal Area Networks (WPANs): Amendment 2: Millimeter-wave based Alternative Physical Layer Extension. IEEE Computer Society, New York (2009)
4. IEEE P802. 11adTM/D0. 1-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications-Amendment 6: Enhancements for Very High Throughput in the 60 GHz Band. New York, USA, IEEE (2010)

5. Liu, F., Zhao, P., Wang, Z.: EKF-based beam tracking for mmWave MIMO systems. IEEE Commun. Lett. **23**(12), 2390–2393 (2019). https://doi.org/10.1109/LCOMM.2019.2940660
6. Zhang, J., Huang, Y., Zhou, Y., You, X.: Beam alignment and tracking for millimeter wave communications via bandit learning. IEEE Trans. Commun. **68**(9), 5519–5533 (2020). https://doi.org/10.1109/TCOMM.2020.2988256
7. Sarieddeen, H., Alouini, M.-S., Al-Naffouri, T.Y.: An overview of signal processing techniques for Terahertz communications. Proc. IEEE **109**(10), 1628–1665 (2021). https://doi.org/10.1109/JPROC.2021.3100811
8. Dai, L., Tan, J., Chen, Z., Vincent Poor, H.: Delay-phase precoding for wideband THz Massive Mimo. In: IEEE Transactions on Wireless Communications. https://doi.org/10.1109/TWC.2022.3157315
9. Tan, J., Dai, L.: Wideband beam tracking in THZ massive MIMO systems. IEEE J. Sel. Areas Commun. **39**(6), 1693–1710 (2021). https://doi.org/10.1109/JSAC.2021.3071817
10. Uthansakul, P., Bialkowski, M.E.:Multipath signal effect on the capacity of MIMO, MIMO-OFDM and spread MIMO-OFDM. In: 15th International Conference on Microwaves, Radar and Wireless Communications (IEEE Cat. No.04EX824), vol. 3, pp. 989–992 (2004). https://doi.org/10.1109/MIKON.2004.1358536
11. García-Loygorri, J.M., Briso, C., Yuste, A.P.: Saleh-Valenzuela modelization and clustering information for a mmWave intra-train channel. In: 2019 13th European Conference on Antennas and Propagation (EuCAP), pp. 1–5 (2019)
12. Liu, C., Zhao, L., Li, M., Yang, L.: Adaptive beam search for initial beam alignment in millimetre-wave communications. IEEE Trans. Veh. Technol. **71**(6), 6801–6806 (2022). https://doi.org/10.1109/TVT.2022.3164533
13. Wang, C., Chen, Y., Lu, Z., Wen, X., Wang, Z., Wang, L.: FC-BET: a fast consecutive beam tracking scheme for mmWave vehicular communications. IEEE Wirel. Commun. Netw. Conf. (WCNC) **2021**, 1–6 (2021). https://doi.org/10.1109/WCNC49053.2021.9417595

Hierarchical Attention Networks for Multispectral Image Compression

Xinran Huang(✉), Kang Wang, Yuxin Meng, and Guanglong Ren

College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

547144022@qq.com

Abstract. Aiming at the different i mportance a nd p ixel s ize o f feature information at different p ositions o f m ultispectral i mages, t his paper proposes an improved ConvNext architecture for multispectral image compression based on swin transformer, and designs the corresponding decoding network while adding multi-scale channel attention. The multispectral image obtains hierarchically mapped feature information through forward coding network, and then obtains the binary code stream through quantization and entropy coding. The reverse decoding network recovers the original image through hierarchical structure and upsampling, and the rate distortion optimization balances the reconstructed image quality and bit rate in end-to-end overall architecture. The experimental results show that at the same bit rate, the PSNR of the proposed method is better than that of the existing JPEG2000.

Keywords: Multispectral image compression · Improved ConvNext · Multiscale channel attention · Rate distortion optimization

1 Introduction

Multispectral images are different from visible images, which need a lot of resources in transmission and storage. Multispectral images bring rich spatial information and spectral information. Due to the redundant information in space and spectrum, such as a large number of identical pixels in an image with sky or grass as the background, compression is required before transmission or storage. In the process of compression, the high frequency detail information with little human perception can be appropriately ignored.

Traditional image compression can be divided into lossless compression and lossy compression. In addition to scenes with high image quality requirements such as medical imaging and fingerprint acquisition, lossy compression is applied in most scenes. JPEG standard based on discrete cosine transform and JPEG2000 standard based on discrete wavelet transform are commonly used image compression methods in transform coding. Similar to transform coding architecture, we use neural network to replace mapping transformation and extract the spatial-spectral feature of multispectral image. The feature information is then quantized and sent to entropy encoder, and the recovery process

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 271–277, 2023.
https://doi.org/10.1007/978-981-99-1260-5_34

is basically symmetrical to the whole forward process. This paper designs an end-to-end multispectral image compression architecture based on convolutional neural network. Following the hierarchical mapping architecture of ConvNext, the feature extraction of forward network is divided into four stages. At each stage, we design ImConvNext module and introduce multi-scale channel attention (MSCA) to extract channel dimension features.

2 Related Work

Since deep learning is applied to image compression, there have been a large number of compression architectures based on convolutional neural network. The convolutional layer, pooling layer, normalization layer and nonlinear activation layer in these architectures are used to extract features of the image and strengthen the representation ability of network to generate feature maps. He et al. [1] proposed a deep residual network, which makes it possible to train hundreds or thousands of layers of networks. The network ensures that when the depth of the model is deepening, there will be no degradation of deep convolutional neural networks. He et al. proved in one of the experiments that the degradation phenomenon was not related to over-fitting. Gradient disappearance is not the reason for network degradation. Ioffe et al. [2] proposed that batch normalization standardizes each channel of the same batch size to control the input modulus of each layer of the network, which effectively solves the problem of gradient disappearance. Balduzzi et al. [3] believed that the deeper the network layer is, the worse the correlation between gradients is, and the gradient update will be randomly disturbed, while the residual network is proved to effectively reduce the attenuation of gradient correlation. Nowadays, the design of residual block in deep network is very important. Skip connection can dynamically adjust the complexity of the model.

Dosovitskiy et al. [4] proposed Vision transformer, which does not use neural network. The paper proves that the application of Vision transformer (ViT) in the field of computer vision can achieve good results. The structure of Swin Transformer is inspired by ViT and improves the original architecture. The layered mapping of neural network is used and multi-head self attention is transformed into window-based multi-head self attention (W-MSA). Because the pixel resolution is much larger than the text, and the computational complexity of multi-head self attention is quadratically related to image size. W-MSA merges patches to a fixed window, and calculates attention in each window, which greatly reduces the computational complexity. Liu [5] returns convolution neural network based on the overall design of Swin Transformer, using convolution blocks instead of attention, increasing the receptive field by hierarchical mapping and downsampling before each layer, obtaining context information and global semantic information. This paper improves the ConvNext architecture, proposes the ConvNext residual structure based on multi-scale channel attention [6], and designs the corresponding reverse decoding network.

3 Proposed Method

The overall framework of this paper is an end-to-end multispectral compression based on deep convolution network. Compared with traditional methods, endto-end image

compression can be jointly optimized. The specific compression process is shown in Figure 1.

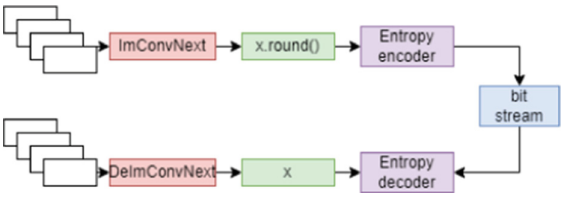


Fig. 1. Overall compression framework

Among them, ImConvNext is the forward coding network, and DeImConvNext is the decoding network. The width and height of image is reduced to one eighth of the original size after three down-sampling operations. After feature extraction, quantization and entropy coding are in turn. The reverse reconstruction image process is basically symmetrical with the whole forward process, finally t he i mage t hat r etains i mportant p ixel i nformation a fter t he p rocess of decompression is generated.

3.1 ImConvNext and DeImConvNext Networks

The main body of the ImConvNext network is composed of UpChannel module, down-sampling module and ImConvNext block. The main body of the DeImConvNext network is composed of DownChannel module, upsampling module and ImConvNext block. The overall structure is shown in Fig. 2.

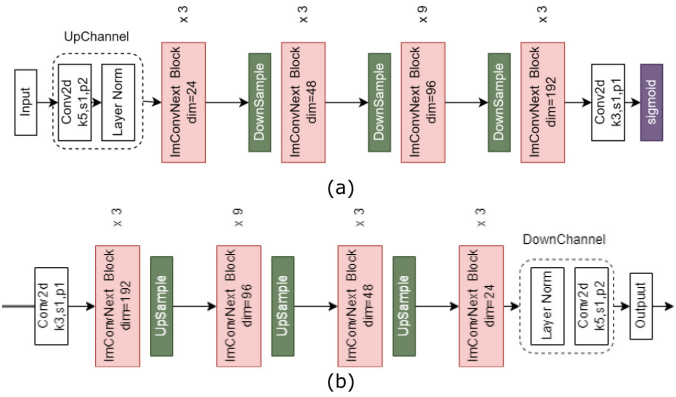


Fig. 2. Codec Network: (a) Forward Network (b) Reverse Network

The forward coding network is shown in Fig. 2(a). The down-sampling operation is performed between ImConvNext Blocks of different dimensions. The input image is processed through four ImConvNext blocks of different dimensions and down-sampling

operation to obtain the extracted spatial attention feature and multi-scale channel attention feature. The overall framework of the forward network is based on the Swin Transformer architecture, which has achieved good results in visual tasks. For visual tasks, multi-scale features are very important. Forward network has four stages. The first stage is composed of DownChannel module and ImConvNext block. The second to fourth stages are composed of DownSample block and ImConvNext block. Each stage maps the image hierarchically to obtain multi-scale intermediate feature map. ImConvNext block continues to replace the transformer block with a convolution operation, adding multiscale channel attention residual block while reducing computational complexity and achieving the same attention effects as transformer. The multi-scale intermediate feature information is realized by the DownSample block, which increases the receptive field of convolution kernel, so that each feature passing through this module focuses on different sizes of the object, obtains a wider range of context information, and then extracts global and local attention features from the context information. The UpChannel block increases the channel of input image while maintaining the size of output feature map consistent with the original image, which facilitates the multi-scale channel attention calculation and fusion in ImConvNext block.

The ImConvNext block adopts residual structure, and this module extracts attention features in the spatial dimension and channel dimension, the specific structure is shown in Fig. 3(a). Drop path is a regularization method, which is similar to the Dropout idea. The effect is that the sub-paths of multi-branch structure in deep learning model are randomly deleted, which can prevent overfitting, improve generalization ability and overcome the problem of network degradation. Depthwise convolution is a special form of group convolution, that is, the number of groups is equal to the number of channels. This technology was mainly used in MobileNet, a lightweight network, to reduce the amount of calculation. In this paper, each convolution kernel of depthwise convolution processes one channel separately. This form is similar to the self-attention mechanism, in which spatial information is mixed weighting within a single channel. In image compression, the adjacent spatial locations of different parts of same object or different objects with similar semantics will affect the quality of reconstructed image. According to local inductive bias of convolutional neural network, replacing self-attention with sliding window can save the resources of visual tasks.

After extracting spatial attention maps, the network learns the importance of spatial information in each spectral segment of the image, and then extracts the attention maps of channels in spectral dimensions after layer normalization of the feature information of spatial dimensions, as shown in Fig. 3(b). The left branch compresses the two-dimensional features of each channel into a real number through global average pooling (gloavgpool), and then increases the channel dimension of feature vector by four times through pointwise convolution. After activation function, original channel dimension is restored through pointwise convolution. Similarly, the right branch constructs the correlation between channels by two pointwise convolutions. Compared with one pointwise convolution, the complex correlation between channels can be better fitted while adding nonlinearity. In this paper, global channel maps and local channel maps extracted from two branches are fused to generate normalized weight between 0 and 1 through broadcast mechanism and sigmoid function. The weight value generated by each

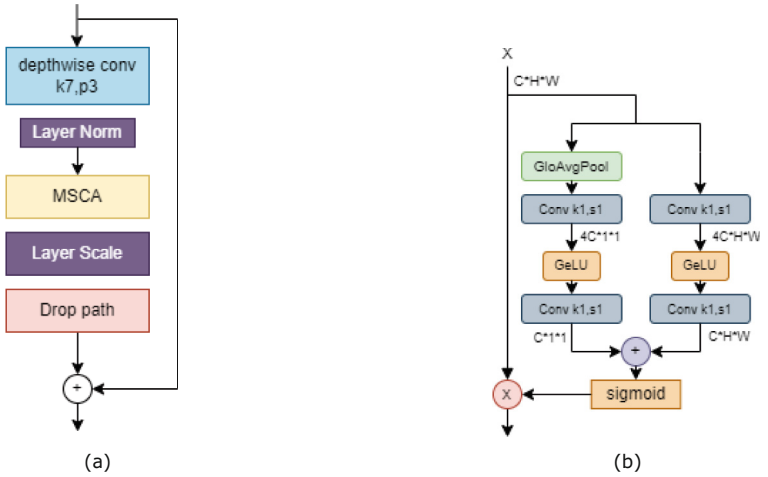


Fig. 3. ImConvNext Block structure: (a) ImConvNext overall structure (b) MSCA

channel is multiplied by original input feature map to obtain channel attention maps. Different spectral segments of multispectral images contain different types of pixels, such as visible spectrum and infrared spectrum, which focus on different feature details. The MSCA module adaptively learns feature weights in the network, obtains the importance of each channel in feature map, and assigns larger weight values to channels containing more key information of image. Through the whole ImConvNext block, spatial and channel attention maps were fully extracted. The decoding network DeImConvNext is shown in Fig. 2(b), its overall architecture is basically symmetric with forward network. The intermediate feature data output from forward network goes through four stages successively after dimensionality increase. In the first three stages, the image spatial size is gradually restored through ImConvNext block and upsampling module. In the fourth stage, DownChannel module reduces the dimension of feature channel to the size of original image channel. The downsampling and upsampling modules in forward and reverse networks are shown in Fig. 4. After each downsampling, the height and width of feature map become 1/2 of that before input, and the number of channels becomes 2 times of that before input.

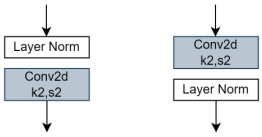


Fig. 4. Down-sampling and up-sampling

3.2 Rate Distortion Optimization

The bit rate and PSNR generated in test stage can be further improved through rate distortion optimization to achieve a relatively large PSNR value at a relatively small bit rate, so that the quality of restored image can be maintained while the compression performance is good, and the distortion caused by compression can be reduced. The loss function is as below:

$$L = L_D + L_R \tag{1}$$

where L_D represents distortion loss and L_R represents bit rate. The purpose of rate distortion optimization is to balance entropy and distortion.

3.3 Implementation Details

The ImConvNext blocks of each stage in forward network stack 3, 3, 9, 3 successively, and the reverse network stack 3, 9, 3, 3 successively. Since the more intermediate feature channels are, the more information is retained. Considering final compression rate, the forward network reduces the dimension of semantic feature with the size of 192 × H × W through convolution with kernel size of 3 and padding size of 1 after the fourth stage, and obtains the intermediate data size 48 × $H/8$ × $W/8$.

4 Results

We use PSNR as the evaluation index of image compression effect, as shown in Table 1. At the approximate bit rate, the proposed method achieves better image quality than JPEG2000.

Table 1. PSNR (dB) results comparison on 8-band test set

	rate	PSNR	rate	PSNR	rate	PSNR
JPEG2000	0.260	33.17	0.337	34.52	0.456	35.62
Proposed	0.260	37.60	0.363	38.47	0.456	39.27

5 Conclusion

In this paper, we propose a multispectral image compression framework based on spatial attention and multi-scale channel attention, which extracts features at different scales through hierarchical mapping structures. The experimental results show that the PSNR of proposed method is better than JPEG2000. The architecture in this paper sufficiently extract feature information for image compression.

References

1. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, pp. 770–778 (2016)
2. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd International Conference on Machine Learning, ser. JMLR Workshop and Conference Proceedings, F. R. Bach and D. M. Blei, Eds., vol. 37. JMLR.org, pp. 448–456 (2015)
3. Balduzzi, D., Frean, M., Leary, L., Lewis, J.P., Ma, K.W., McWilliams, B.: The shattered gradients problem: if resnets are the answer, then what is the question?. In: Proceedings of the 34th International Conference on Machine Learning, ser. Proceedings of Machine Learning Research, D. Precup and Y. W. Teh, Eds., vol. 70. PMLR, pp. 342–350 (2017)
4. Dosovitskiy, A., et al.: An image is worth 16x16 words: transformers for image recognition at scale. In: 9th International Conference on Learning Representations. OpenReview.net (2021)
5. Liu, Z., Mao, H., Wu, C., Feichtenhofer, C., Darrell, T., Xie, S.: A convnet for the 2020s. CoRR, vol. abs/2201.03545 (2022)
6. Dai, Y., Gieseke, F., Oehmcke, S., Wu, Y., Barnard, K.: Attentional feature fusion. CoRR, vol. abs/2009.14082 (2020)

Directional Modulation Design Based on Generalized Recurrent Neural Network

Ting Xu[1], Bo Zhang[1(✉)], Baoju Zhang[1(✉)], Wei Wang[2], Bo Wang[2], Lin Mu[3], Xiaonan Zhao[1], and Cheng Wang[1]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
xt2398325087@163.com, b.zhangintj@tjnu.edu.cn, wdxyzbj@163.com
[2] Sitonholy (Tianjin) Technology Co., Ltd., Tianjin, China
[3] Tianjin Vocational Institute, Tianjin, China

Abstract. This paper proposes a generalized regression neural network (GRNN) based directional modulation (DM) method. Taking the uniform linear array as an example, the GRNN network is constructed and the weighted vector of the array is obtained to realize the required DM function. The results show the effectiveness of the proposed design.

Keywords: Directional modulation · uniform linear array · generalized regression neural network

1 Introduction

Since the same constellation mapping is transmitted in all spatial angles, it is possible for eavesdroppers in undesired directions to decode the transmitted signal. To solve this problem, directional modulation (DM) technique was introduced. The distortion of the constellation generated by DM and its safety advantages were explained in [1–3]. For reconfigured antenna array designs, DM modulation can be achieved by on-off control of each symbol [4]. Multi-carrier based phased antenna array design for directional modulation was studied in [5], followed by a combination of DM and polarisation design in [6]. The bit error rate (BER) performance of a system based on a two-antenna array was studied using the DM technique for eight phase shift keying modulation in [7]. A more systematic pattern synthesis approach was presented in [8], followed by a time modulation technique for DM to form a four-dimensional (4-D) antenna array in [9]. However, the traditional directional modulation cannot perform fast processing and requires a large amount of calculation, which cannot meet the high requirements of modern wireless communication for real-time performance.

In the paper, we propose a General Regression Neural Network (GRNN) based DM design. GRNN is an artificial neural network model based on nonlinear

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 278–284, 2023.
https://doi.org/10.1007/978-981-99-1260-5_35

regression theory and it has good nonlinear approximation performance and is highly fault-tolerant and robust [10,11].

The paper is structured as follows. DM design is reviewed in Sect. 2. In Sect. 3, the GRNN based DM design is proposed. Design examples are provided in Sect. 4 and conclusions are drawn in Sect. 5.

2 Review of DM Design Based on Narrowband Linear Antenna Arrays

2.1 ULA-Based Narrowband Beamforming

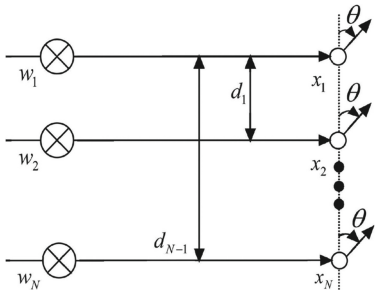


Fig. 1. Narrowband transmit beamforming structure

The narrowband linear array used for transmit beamforming is shown in Fig. 1, where N equally spaced antennas with a spacing of d_1 are formed and the emission direction $\theta \subseteq [0, 180]$. The weighted vector for the whole structure can be represented by

$$\mathbf{w} = [w_1, w_2, ..., w_N]^T. \tag{1}$$

The steering vectors $\mathbf{s}(\omega, \theta)$ is given by

$$\mathbf{s}(\omega, \theta) = [1, e^{j\omega dcos\theta/c}, e^{j\omega 2dcos\theta/c}, ..., e^{j\omega(N-1)dcos\theta/c}]^T, \tag{2}$$

where $\{\cdot\}^T$ is the transpose expression and c is the transmission speed. Then, the beam response of the array can be given by

$$p(\theta) = \mathbf{w}^H \mathbf{s}(\omega, \theta), \tag{3}$$

where $\{\cdot\}^H$ represents the Hermitian transpose expression.

2.2 DM Design with an Interference Constraint

The goal of DM design is to find appropriate weight coefficients for all symbols to achieve the resultant main beams pointing to the desired direction(s) with given phase shifts, but random phase shifts with a low magnitude level in other directions. Then, for the m-th symbol ($m = 1, 2, \ldots, M$), the weight vector is given by

$$\mathbf{w}_m = [w_{(m,1)}, ..., w_{(m,N)}]^T, m = 1, 2, \ldots, M. \tag{4}$$

Similarly, the corresponding beam response for the transmission angle θ can be represented by $p_m(\theta)$. Here, the transmission angle θ can be differentiated into three parts: the interest direction(s), the interference direction(s), and the rest of ranges. Without loss of generality, we assume r responses in the mainlobe directions, q responses in interference directions and $R - r - q$ responses in sidelobe regions. Then, for the m-th symbol we have

$$\begin{aligned}
\mathbf{p}_{m,ML} &= [\mathbf{p}_m(\omega, \theta_0), \mathbf{p}_m(\omega, \theta_1), ..., \mathbf{p}_m(\omega, \theta_{r-1})], \\
\mathbf{p}_{m,IL} &= [\mathbf{p}_m(\omega, \theta_r), \mathbf{p}_m(\omega, \theta_{r+1}), ..., \mathbf{p}_m(\omega, \theta_{r+q-1})], \\
\mathbf{p}_{m,SL} &= [\mathbf{p}_m(\omega, \theta_{r+q}), \mathbf{p}_m(\omega, \theta_{r+q+1}), ..., \mathbf{p}_m(\omega, \theta_{R-1})].
\end{aligned} \tag{5}$$

Similarly, the steering vector in the mainlobe, interference and sidelobe ranges can be given by

$$\begin{aligned}
\mathbf{S}_{ML} &= [\mathbf{S}(\omega, \theta_0), \mathbf{S}(\omega, \theta_1), ..., \mathbf{S}(\omega, \theta_{r-1})], \\
\mathbf{S}_{IL} &= [\mathbf{S}(\omega, \theta_r), \mathbf{S}(\omega, \theta_{r+1}), ..., \mathbf{S}(\omega, \theta_{r+q-1})], \\
\mathbf{S}_{SL} &= [\mathbf{S}(\omega, \theta_{r+q}), \mathbf{S}(\omega, \theta_{r+q+1}), ..., \mathbf{S}(\omega, \theta_{R-1})].
\end{aligned} \tag{6}$$

Then, the weight coefficient for the m-th symbol can be obtained by

$$\begin{aligned}
min \quad & \| \mathbf{p}_{m,SL} - \mathbf{w}_m^H \mathbf{S}_{SL} \|_2 \\
subject\ to \quad & \mathbf{w}_m^H \mathbf{S}_{ML} = \mathbf{p}_{m,ML} \\
& |\mathbf{w}_m^H \mathbf{S}_{IL}| <= 0.001,
\end{aligned} \tag{7}$$

where the objective function ensures the minimisation between the expected sidelobe response and the designed one (where $||\cdot||_2$ represents the l_2 norm). The equality constraint keeps the mainlobe designed response the same as the expected one, and the inequality constraint is used to suppress the interference signal [12,13].

3 Generalized Regression Neural Network

The network structure of GRNN is shown in Fig. 2, which is a four-layer structure with an input layer, a pattern layer, a summation layer and an output layer.

(1) Input layer: the number of neurons is the dimension of input, represented by L. Here, the mainlobe and interference directions are used as input variables

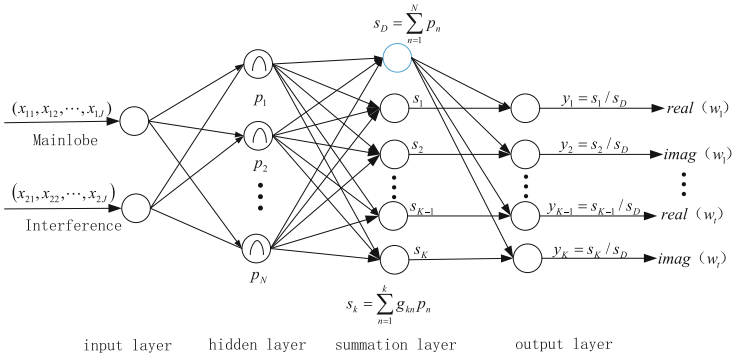


Fig. 2. GRNN network structure

($L = 2$). For input testing samples, the matrix $\mathbf{X}$ is represented by $L \times J$, where J is the number of testing samples.

(2) Hidden layer: the number of neurons in this layer is N. The dimension of training samples $\mathbf{C}$ is $L \times N$, where N is the number of training samples. We define matrix $\mathbf{D}$

$$\mathbf{D}(n,j) = (\mathbf{X}(:,j) - \mathbf{C}(:,n))^T (\mathbf{X}(:,j) - \mathbf{C}(:,n)) \quad j = 1,2,\cdots,J; n = 1,2,\cdots,N \tag{8}$$

Here, we use Gaussian activation kernel function. Then, the output $N \times J$ matrix $\mathbf{P}$ of the hidden layer is given by

$$\mathbf{P}(n,j) = exp[-\frac{\mathbf{D}(n,j)}{2\sigma^2}] \tag{9}$$

We set vector $\mathbf{p}_n = \mathbf{P}(n,:)$. The threshold of the hidden layer neurons is given by

$$b = \frac{1}{2\sigma^2} = \frac{\sqrt{-log(a)}}{spread} = \frac{0.8326}{spread} \tag{10}$$

where a is the Gaussian amplitude, $spread$ represents the expansion speed, and σ is called the smoothing factor.

(3) Summation layer: There are two types of summations neurons, namely, $\mathbf{s}_D$ and $\mathbf{s}_k$, which given in Eq. (11). The $\mathbf{s}_D$ summation neuron computes the sum of the unweighted outputs of the hidden layer, while the $\mathbf{s}_k$ summation neuron computes the sum of the weighted outputs of the hidden layer [14]. We set the training output samples matrix as $\mathbf{G}$ ($K \times N$), where K is the dimension of output. The weight between the n-th hidden neuron and the k-th summation neuron is set to be g_{kn}. g_{kn} represents the element of the K-th row and the n-th column of the matrix G

$$\mathbf{s}_D = \sum_{n=1}^{N} \mathbf{p}_n, \quad \mathbf{s}_k = \sum_{n=1}^{N} g_{kn}\mathbf{p}_n, \quad k = 1,2,...,K \tag{11}$$

(4) Output layer: The number of neurons is K. Here, we set $K = 2t$, where t is the number of complex weighting coefficients. The output of network is given by

$$\hat{\mathbf{Y}}(\mathbf{X}) = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_K]^T = [\frac{\mathbf{s}_1}{\mathbf{s}_D}, \frac{\mathbf{s}_2}{\mathbf{s}_D}, ..., \frac{\mathbf{s}_K}{\mathbf{s}_D}]^T \tag{12}$$

4 Design Examples

Here, we design the mainlobe $\theta_{ML} \in [30°, 150°]$ with the step size 5°. The sidelobe region is $\theta_{SL} \in [0°, \theta_{ML} - 5°] \cup [\theta_{ML} + 5°, 180°]$, and the interference direction is $\theta_{IL} = [2°, 174°]$ with the step size 2°. The desired response is a value of one (magnitude) with corresponding phase shift at the mainlobe (quadrature PSK) and a value of 0.1 (magnitude) with random phase shifts over the sidelobe regions. We obtain DM results based on an 8-element ULA with half-wavelength spacing using the method in Eq. (7), and generate the samples with mainlobe and interference as input variables and output as weight vector.

The testing and training GRNN network does not require repeated training, and the number of neurons in each layer is determined by the training samples. For the 8-element ULA, the dimension of complex weight coefficients is 1×8, and these coefficients are divided into real and imaginary part. Therefore, we have $K = 16$. For each symbol, the number of samples is 2065. The sample is divided and normalized according to the ratio of the training set to the testing $(4 : 1)$. Then, the matrix dimension of $\mathbf{C}$ is 2×1652, $\mathbf{X}$ is 2×413, $\mathbf{G}$ is 1652×16. Bringing the matrix into the formula (12) can calculate the network output.

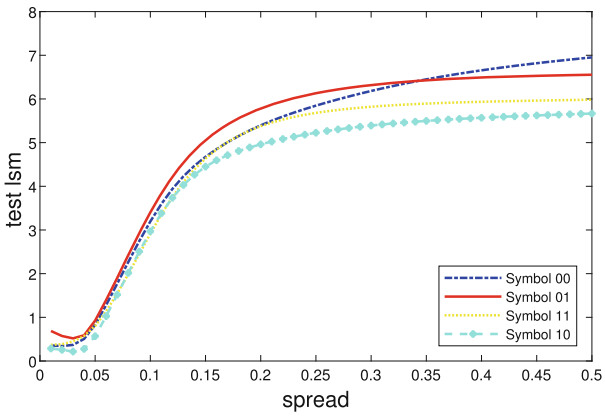


Fig. 3. LSM of different spreads with different symbols in the GRNN model

In GRNN network, the *spread* determines the predictive ability. Here, we set *spread* as a variable, which is increased from 0.1 to 4 with the step size 0.01. As shown in Fig. 3, through the simulation experiment on the *spread* value for four symbols, we can see that the GRNN model maintains a good prediction

accuracy when the optimal spread is between 0 and 0.05, where the least mean square (LSM) for testing samples is low. But after that, the LSM continues to rise. Here, the optimal *spread* and minimum LSM obtained are shown in Table 1, where we can see that the best *spread* value is 0.01 for symbols $00, 01, 11$, and 0.03 for symbol 10. Here, we also give magnitude and phase responses to show the effectiveness of the proposed design. We assume the desired and interference direction is set to 60° and 100° as a testing sample. As shown in Fig. 4, we can see that based on the optimised GRNN network, all mainlobes point to the desired direction with very low magnitude levels in other directions. The phase of these four symbols follows the standard QPSK modulation in the desired direction but random in other ranges.

Table 1. Spread and LSM of four-symbols

Symbol	Spread value	Least squares mean
00	0.01	0.6557932
01	0.01	0.3460669
11	0.01	0.1164718
10	0.03	0.3609596

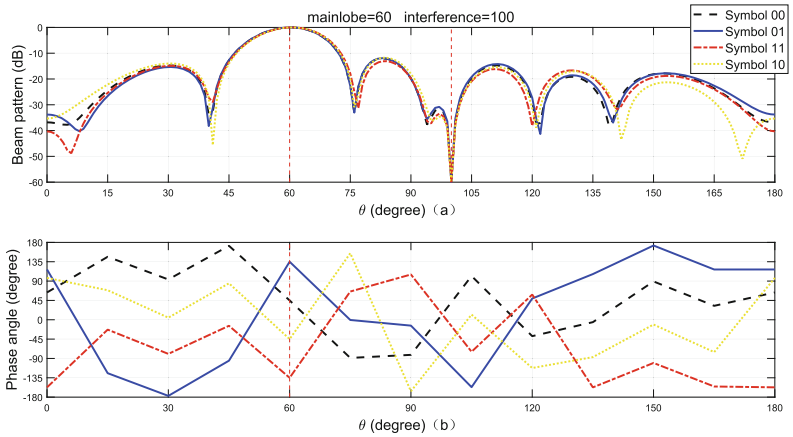


Fig. 4. Design results for four-symbols based on the 8-element ULA by using GRNN.

5 Conclusions

In this paper, generalized regression neural network based direction modulation design is proposed, which realizes the mapping from the desired and interference

directions to the array weight vectors. Through the trained GRNN, the array weights that satisfy DM can be obtained.

Acknowledgments. The work was supported by the National Natural Science Foundation of China (62101383) and Hedong District Science and Technology Project (TJHD-JBGS-2022-08).

References

1. Daly, M.P., Bernhard, J.T.: Directional modulation technique for phased arrays. IEEE Trans. Antennas Propag. **57**(9), 2633–2640 (2009)
2. Ding, Y., Fusco, V.: A vector approach for the analysis and synthesis of directional modulation transmitters. IEEE Trans. Antennas Propag. **62**(1), 361–370 (2014)
3. Daly, M.P., Bernhard, J.T.: Beamsteering in pattern reconfigurable arrays using directional modulation. IEEE Trans. Antennas Propag. **58**(7), 2259–2265 (2010)
4. Hong, T., Song, M.Z., Liu, Y.: Dual-beam directional modulation technique for physical-layer secure communication. IEEE Antennas Wirel. Propag. Lett. **10**, 1417–1420 (2011)
5. Zhang, B., Liu, W., Li, Q.: Multi-carrier waveform design for directional modulation under peak to average power ratio constraint. IEEE Access **7**, 37528–37535 (2019)
6. Zhang, B., Liu, W., Lan, X.: Orthogonally polarized dual-channel directional modulation based on crossed-dipole arrays. IEEE Access **7**, 198–206 (2019)
7. Shi, H.Z., Tennant, A.: Enhancing the security of communication via directly modulated antenna arrays. IET Microwaves Antennas Propag. **7**(8), 606–611 (2013)
8. Ding, Y., Fusco, V.: Directional modulation transmitter radiation pattern considerations. IET Microwaves Antennas Propag. **7**(15), 1201–1206 (2013)
9. Zhu, Q.J., Yang, S.W., Yao, R.L., Nie, Z.P.: Directional modulation based on 4-D antenna arrays. IEEE Trans. Antennas Propag. **62**(2), 621–628 (2014)
10. Xu, X.T., Sun, X.Y., Li, C.H.: Prediction of hardenability of sae8620h automobile gear steel based on grnn neural network. Yunnan Metallurgy **49**(5), 108–111 (2020)
11. Zhou, S.Y., Huang, W.N., Li, C.C.: A method for beamforming based on RBF neural network. Computer Simul. **37**(10), 159–163 (2020)
12. Hayward, S.D.: Effects of motion on adaptive arrays. IEEE Proceed. Radar Sonar Navig. **144**(1), 15–20 (1997)
13. Shu, B.W.: Research on array digital beamforming technology. National University of Defense Technology, Changsha (2006)
14. Wang, Y., Peng, H.: Underwater acoustic source localization using generalized regression neural network. Acoust. Soc. Am. **143**(4), 2321–2331 (2018)
15. Li, C.: Research on wide area backup protection system based on multi-information fusion (2014)

Target Parameter Estimation Method Based on Information Theory

Kong Xiaolong[1], Xu Dazhuan[1(✉)], Hua Boyu[1], and Bao Junwei[2]

[1] College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China
{xiao_long_kong,xudazhuan,byhua}@nuaa.edu.cn
[2] College of Physics, Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China
broadenway@nuaa.edu.cn

Abstract. As the target detection environment becomes more and more complicated, passive radar has gradually become the focus of researchers' attention because of its advantage of strong detection concealment. In this paper, we propose the parameter information estimation method for passive radar by spatial information theory. First, the concept of spatial information is proposed through the Shannon information theory, secondly, the theoretical formula and closed expression for spatial information are derived from the definition of mutual information, finally, the entropy error (EE) in the passive radar is defined to further evaluate the proposed parameter estimation method. We focus on the theoretical derivation of the spatial information quantity, the concept of EE and its relation to the mean squared error (MSE). Then, to demonstrate the performance of the proposed method in passive radar, extensive numerical simulation work is carried out. The results indicate that radar distance information can be quantified by bits, and the EE method is a better choice in low and medium the signal-to-noise ratio (SNR) conditions compared with the traditional MSE method. In this paper, the basic target parameter estimation problem of passive radar is solved, and the spatial information theory system can be further developed and improved in the future.

Keywords: Information theory · spatial information · Passive radar · Entropy error

1 Introduction

A radar detection system can obtain spatial information of the target from the received signal, such as distance information, direction information and scattering information. Therefore, it plays important role in the national defense and the national people's livelihood sectors [1, 2].

This work was supported in part by the Key Project of Aeronautical Science Foundation of China under Grant NO. 2020Z073009001, in part by the open research fund of National Mobile Communications Research Laboratory, Southeast University, No. 2022D04, in part by the Future Network Scientific Research Fund Project, No. FNSRFP-2021-YB-04.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 285–292, 2023.
https://doi.org/10.1007/978-981-99-1260-5_36

Shannon's information theory has been applied in communication systems for many years, while the first application in radar systems was in the 1950s. Woodard adopted the inverse probability principle to research radar distance information problems [3]. However, the limitation of the inverse probability principle also led to their work being unable to be extended to multi-target situations. Bell first applied mutual information measurements to the waveform design of radar detection systems in the 1980s [4]. Paul proposed a radar target scheduling algorithm based on the mutual information criterion in [5]. However, these relevant research results mainly focus on the radar waveform design, and the radar detection information acquisition process is rarely involved [6]. In recent years, the parameter estimation method of active radar and sensor array are proposed in [7–10]. With the target detection environment becoming more and more complicated, passive radar has been widely utilized due to the advantage of strong detection concealment [11]. However, the parameter information acquisition process in passive radar has not been studied yet. Therefore, a target parameter estimation method for passive radar is needed.

In this paper, the information acquisition process of the passive radar through the spatial information theory is proposed. At first, according to the commonly passive radar model, the conditional probability density function (PDF) of the received energy under the known normalized time delay condition is obtained. Then, the mutual information between the target distance and the received energy in the passive radar is defined as the distance information, and the theoretical bound of distance information is derived. Finally, the theoretical bound of the EE method is given, and corresponding simulations are conducted.

2 Parameter Estimation Method

2.1 The Passive Radar Model

Passive radar is also an information acquisition system that relies on passively receiving microwave energy scattered by the target to track and locate the target [12]. To facilitate discussion, we consider a single-antenna passive radar and only one target source in the observation domain. A typical passive radar detection system is shown in Fig. 1.

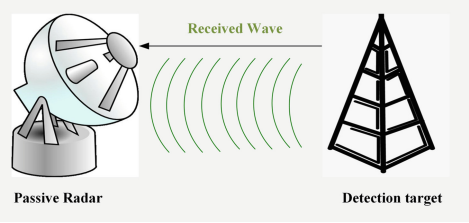


Fig. 1. A typical passive radar detection system

The receiving end of the passive radar only receives the signal scattered by the target, and it is assumed that the scattered signal of the target is always present in the observation interval. Therefore, the received signal can be expressed as

$$z(t) = su(t - \tau) + w(t) \tag{1}$$

where s is the complex scattering coefficient of the target, τ denotes the target propagation delay, $u(t\text{-}\tau)$ represents the moment the target appears, and $w(t)$ denotes the complex Gaussian noise.

According to the Nyquist sampling theorem, the received signal $z(t)$ can be sampled at rate B. The discrete form of (1) can be described as

$$z(n) = \alpha e^{j\varphi} u(n - x) + w(n) \tag{2}$$

where $x = B\tau$ represents the normalized time delay, $N = TB$ is the time-bandwidth product, n denotes the integral point in [0, N-1], α is constant that denotes the modulus, φ is the phase and distributed uniformly in [0, 2Π], $u(n\text{-}x)$ represents the time delay of complex scattering signal.

The passive radar receives microwave energy from the target to track and locate the target. Therefore, the energy form of the received signal is used to discuss the distance information problem. The energy form of (2) can be written as

$$r(n) = |z(n)|^2 = \sum_{0}^{N-1} \left\{ \left| \alpha e^{j\varphi} u(n - x) \right|^2 + |w(n)|^2 + 2\mathrm{Re}\left[\alpha e^{j\varphi} u(n - x) \cdot w(n) \right] \right\}. \tag{3}$$

The variable $u(n\text{-}x)$ and the variable $w(n)$ in (3) are discrete, random and independent of each other. The central limit theorem states that if the random variable under study is the sum of many independent and uniform random variables, then its distribution will approximate a normal distribution. Therefore, the Gaussian approximation method is adopted to solve the energy form of the received signal. The mean and variance in (3) can be calculated as

$$E(r(n)) = \alpha^2(N - x) + NN_0 \tag{4}$$

$$D(r(n)) = NN_0^2 + 2\alpha^2 N_0(N - x) \tag{5}$$

The conditional energy PDF of the received signal under the given time delay condition can be approximated as

$$p(\mathbf{r}|x) = \frac{\exp\left\{ -\frac{\left(\mathbf{r}/N_0 - \left(\rho^2(N-x)+N\right)\right)^2}{2\left(N+2\rho^2(N-x)\right)} \right\}}{\sqrt{2\pi\left(N + 2\rho^2(N - x)\right)}} \tag{6}$$

where $\rho^2 = \alpha^2/N_0$ represents the SNR.

According to the Shannon information theory, distance information can be defined as follow. Suppose $p(x)$ is the PDF of the target distance X, $p(\boldsymbol{r}|x)$ is the conditional energy

PDF of the received signal $\boldsymbol{R}$, for which the target distance is known. The target distance information is defined as the mutual information $I(\boldsymbol{R}; X)$ obtained from the energy of the received signal about the target distance, i.e.,

$$I(\mathbf{R}; X) = E\left[\log \frac{p(\mathbf{r}|x)}{p(\mathbf{r})}\right] \tag{7}$$

3 Parameter Estimation Method

The main purpose of the proposed parameter estimation method is to estimate the distance information of the target. In this part, Bayes' theorem is introduced to address the target distance information estimation problem [13]. Therefore, the *a posteriori* PDF of the target distance information can be obtained

$$p(x|\mathbf{r}) = \frac{1}{\delta} \frac{\exp\left\{-\frac{\left(\mathbf{r}/N_0 - \left(\rho^2(N-x)+N\right)\right)^2}{2\left(N+2\rho^2(N-x)\right)}\right\}}{\sqrt{2\pi\left(N+2\rho^2(N-x)\right)}} \tag{8}$$

where $\delta = \int p(x|\mathbf{r})dx$ denotes the regularized factor and it is a constant.

According to the definition of mutual information, the theoretical formula of target distance information can be expressed as

$$\begin{aligned} I(\mathbf{R}; X) &= h(X) - h(X|\mathbf{R}) \\ &= \log N - E_r\left[-\int_0^{N-1} p(x|\mathbf{r}) \log p(x|\mathbf{r})dx\right] \end{aligned} \tag{9}$$

The formula (9) is very complicated, and another simpler form is derived. It is assumed that the target is located at x_0, and the energy of the received signal $r_E(n)$ can be expressed as

$$r_E(n) = \alpha e^{j\varphi_0} u(n - x_0) + w_0(n) \tag{10}$$

Substituting the new expectation and new variance in (10) into (8), the energy PDF of the received signal can be rewritten as

$$p(x|\mathbf{r}_E) = \frac{1}{\delta_E} \frac{\exp\left\{-\frac{\left[\rho^2(x-x_0)\right]^2}{2\left(N+2\rho^2(N-x_0)\right)}\right\}}{\sqrt{2\pi\left(N+2\rho^2(N-x_0)\right)}} \tag{11}$$

where δ_E denotes the regularized factor and it is also a constant.

Therefore, the target distance information in (9) can be rewritten as

$$I(\mathbf{R}_E; X) = \log N - E_{r_E}\left[-\int_0^{N-1} p(x|\mathbf{r}_E) \log p(x|\mathbf{r}_E)dx\right] \tag{12}$$

4 Performance Evaluation

In this section, the performance evaluation for the proposed parameter estimation method in passive radar is presented. Evaluation indexes such as target distance information theoretical upper bound, Cramér-Rao bound (CRB) and EE method are introduced. The detailed results and discussion can be seen in the next section.

To derive the target distance information theoretical bound in the passive radar, (11) under the high SNR conditions can be approximately expressed as

$$\begin{aligned} p(x|\mathbf{r}_E) &= \frac{\exp\left\{-\frac{1}{2}\frac{\left(\rho^2\right)^2(x-x_0)^2}{N+2\rho^2(N-x_0)}\right\}}{\int_0^{N-1}\exp\left\{-\frac{1}{2}\frac{\left(\rho^2\right)^2(x-x_0)^2}{N+2\rho^2(N-x_0)}\right\}dx} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right) \end{aligned} \tag{13}$$

where $\sigma^2 = \left(\rho^2/N\right)^{-1}$.

The theoretical bound of the target distance information under the high SNR conditions can be calculated as

$$I(\mathbf{R}_E; X) \le \log N - \frac{1}{2}\log\left(2\pi e\left(\rho^2/N\right)^{-1}\right) = \log\sqrt{\frac{N\rho^2}{2\pi e}} \tag{14}$$

The CRB is the lower bound of the MSE evaluation method in the unbiased parameter estimation method, which provides a theoretical basis for the comparison of the performance of various parameter estimation methods [14]. It can be described as

$$\sigma_{CRB}^2 = -\frac{1}{E\left[\frac{\partial^2[\ln p(\mathbf{r}_E|x)]}{\partial x^2}\right]} = \left(\rho^2/N\right)^{-1} \tag{15}$$

It is found that the MSE evaluation method cannot reflect the real situation well, because it has certain limitations under the low and medium SNR conditions [15]. The EE method is a parameter evaluation index defined from the perspective of information theory, which can better reflect the performance of the radar parameter estimation system. The lower bound of the EE method can be obtained as

$$\sigma_{EE}^2 \ge \frac{2^{2\log\sqrt{\frac{2\pi eN}{\rho^2}}}}{2\pi e} = \left(\rho^2/N\right)^{-1} \tag{16}$$

5 Results and Discussion

To further analyze the proposed parameter estimation method performance, the evaluation indexes are evaluated qualitatively and quantitatively. All numerical simulation processes are performed by MATLAB R2016a.

The target distance information and it's upper bound under the three different time-bandwidth product conditions are shown in Fig. 2. We can further understand the passive radar detection process from the perspective of information theory. Under the low SNR conditions, the noise interference is relatively large, it is almost impossible to find the target and the amount of the target distance information also is very small. As the SNR value increases, the change in target distance information is divided into two important stages, the target detection stage, and the target tracking stage. In Fig. 2, curves near the marker ① represents the target detection stage, the target distance information rapidly increases, curves near the marker ② denotes the target tracking stage, and the target distance information gradually coincides with the upper bound of the target distance information.

The comparison results among the EE method, the MSE method and the CRB are shown in Fig. 3. Both the error of the EE method and the MSE method decreases with the SNR value increase and gradually tends to CRB. Figure 3 also illustrates that the error of the MSE method and the error of the EE method gradually approximate the CRB under high SNR conditions, however, under the low and medium SNR conditions, compared with the MSE method, the EE method has a smaller error. Therefore, this part can draw an important conclusion, the EE method as a new evaluation index is more intuitive and accurate than the MSE method.

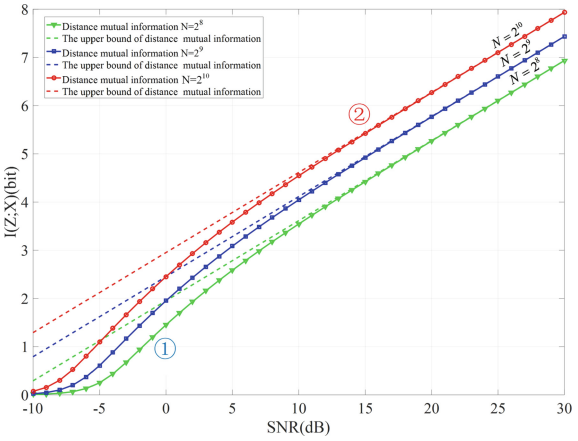


Fig. 2. Target distance information and target distance information upper bound

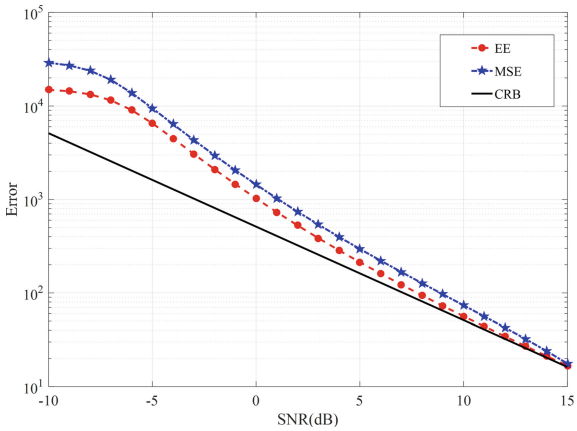


Fig. 3. Comparison results between the EE method, the MSE method and the CRB

6 Conclusion

In this paper, a target parameter estimation method and corresponding performance evaluation index for passive radar are proposed to describe the target distance information acquisition process. The passive radar detection system is assumed to be equivalent to a communication system through Shannon's information theory. The target distance information is also obtained by introducing mutual information, and it can be quantified by bits. Then, evaluation metrics are used to validate the performance of the proposed parameter estimation method. The numerical simulation results show that the EE method is a better choice in low and medium SNR conditions compared with the traditional MSE method. Finally, the proposed parameter estimation method can be used in multiple target cases in the future study.

References

1. Lu, J., Liu, F., Sun, J., Liu, Q., Miao, Y.: Joint estimation of target parameters and system deviations in MIMO radar with widely separated antennas on moving platforms. IEEE Trans. Aeros. Electron. Syst. **57**(5), 3015–3028 (2021)
2. Ustalli, N., Pastina, D., Lombardo, P.: Target motion parameters estimation in forward scatter radar. IEEE Trans. Aeros. Electron. Syst. **56**(1), 226–248 (2020)
3. Woodward, P., Davies, I.: A theory of radar information. Dublin Phil. Mag. J. Sci. **41**(321), 1001–1017 (1950)
4. Bell, M.R.: Information theory and radar waveform design. IEEE Trans. Inf. Theory **39**(5), 1578–1597 (1993)
5. Paul, B., Bliss, D.W.: The Constant Information Radar. Entropy-SWITZ. **18**(9), 338 (2016)
6. Zhu, Q., et al.: Geometry-Based Stochastic Line-of-Sight Probability Model for A2G Channels Under Urban Scenarios. IEEE Trans. Antennas Propag. **70**(7), 5784–5794 (2022)
7. Xu, D.Z., Yan, X., Xu, S.K., Luo, H., Liu, J.W., Zhang, X.F.: Spatial information theory of sensor array and its application in performance evaluation. IET Commun. **13**(15), 2304–2312 (2019)

8. Luo, H., Xu, D.Z., Tu, W.L., Bao, J.W.: Closed-Form asymptotic approximation of targets range information in radar detection systems. IEEE Access **8**, 105561–105570 (2020)
9. Zhu, Q.M., et al.: A novel 3D non-stationary wireless MIMO channel simulator and hardware emulator. IEEE Trans. Commun. **66**(9), 3865–3878 (2018)
10. Zhu, Q.M., Zhao, Z.K., Mao, K., Chen, X.M., Liu, W.Q., Wu, Q.H.: "A real-time hardware emulator for 3d non-stationary u2v channels," *IEEE Trans.* Circuits Syst. I: Reg. Papers **68**(9), 3951–3964 (2021)
11. Moscardini, C., et al.: Batches algorithm for passive radar: a theoretical analysis. IEEE Trans. Aeros. Electron. Syst. **51**(2), 1475–1487 (2015)
12. Blasone, G.P., Colone, F., Lombardo, P., Wojaczek, P., Cristallini, D.: Passive radar STAP detection and DOA estimation under antenna calibration errors. IEEE Trans. Aeros. Electron. Syst. **57**(5), 2725–2742 (2021)
13. Yang, J., Yang, Y.: Sparse Bayesian DOA estimation using hierarchical synthesis lasso priors for off-grid signals. IEEE Trans. Signal Process **68**, 872–884 (2020)
14. Marzetta, T.L.: A simple derivation of the constrained multiple parameter Cramer-Rao bound. IEEE Trans. Signal Process. **41**(6), 2247–2249 (1993)
15. Fisher, R.A.: on the mathematical foundations of theoretical statistics. Phil. Trans. Roy. Soc. **222**, 309 (1922)

A Modified Subspace Tracking Algorithm for Space-Time Adaptive Processing in Inhomogeneous Environment

Duo Lu, Ying Zhang, and Zhang-Xin Chen(✉)

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
zhying@uestc.edu.cn, 344256963@qq.com

Abstract. For non-side-looking airborne radar, Doppler frequency from clutter is varied with range and has inhomogeneous characteristics. The strong inhomogeneous results will make the performance of space-time adaptive processing (STAP) decline seriously. To obtain better clutter suppression performance in inhomogeneous clutter environment, a modified subspace tracking algorithm was presented by adding subspace constraint in this paper. Its performances were verified by simulations.

Keywords: STAP · inhomogeneous clutter · subspace tracking · subspace constraint

1 Introduction

In airborne radar system, space-time adaptive processing (STAP) is well known for its ability to suppress strong and Doppler spreading clutter and improve radar target detection capabilities [1]. To apply STAP technique, a large number of independent and identically distributed (IID) data samples [2] are required and the operation of high-dimensional matrix inversion will be involved. Added to that, the clutter spectrum of side-looking airborne radar is distributed along the diagonals of angular-Doppler plane, but for non-side-looking applications it is distributed along elliptic curves in the angular-Doppler plane and shows strong clutter inhomogeneous characteristics. The detection area is closer to radar operation range, the inhomogeneous characteristics will be stronger. Thus the clutter suppression performance will be seriously affected.

To improve the clutter suppression performance of STAP in inhomogeneous environment, many clutter compensation methods had been proposed such as Doppler compensation (DC) [3], angle-Doppler compensation (ADC) [4], derivative-based updating (DBU) [5] and registration-based compensation (RBC) [6]. The basic idea behind them is to compensate the inhomogeneous of clutter spectrum caused by non-side-looking array and reduce the correlation between clutter spectrum and distance. The compensation are performed from

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 293–300, 2023.
https://doi.org/10.1007/978-981-99-1260-5_37

the data-field with the assumption that the clutter spectrum of all samples has the same structure as the detected range bin.

More attractive viewpoint comes from signal subspace tracking and the projection approximation subspace tracking algorithm (PAST) [7] is one of the representative methods. Unfortunately, the performance of PAST are not very stable, especially in the inhomogeneous clutter environment. From the perspective of similarity constraint, a modified subspace tracking algorithm for STAP is proposed and its performance will be verified in inhomogeneous environment.

2 Signal Model and Problem Formulation

Without loss of generality, the airborne radar system under consideration is equipped with a uniform linear array (ULA). The altitude of platform is denoted by h, the velocity is a constant denoted by v. The ULA has N antenna elements with a half-wavelength inner spacing d. M pulses are transmitted at a constant pulse repetition frequency (PRF) f_r during a coherent processing interval (CPI). θ_p is the angle between direction perpendicular to bore-sight of ULA and the moving direction of platform called the crab angle. The transmitter emits a pulse and the receiver collects samples of the echos, the time slot corresponding to a given range is called a range bin. Airborne radar detection can be described as a binary hypothesis as

$$\begin{cases} H_0 : x = x_c + n \\ H_1 : x = x_t + x_c + n \end{cases} \tag{1}$$

where x_t, x_c are the returned signals from radar targets and the clutter respectively, n includes all system thermal noise. The clutter is often modeled as N_c independent clutter patches which are uniformly distributed in azimuth angles on a given range cell under test (CUT). Each clutter patch can be described by the azimuth angle θ_i and the elevation angle φ_i of the lth range bin, *i.e.*,

$$x_c = \sum_{i=1}^{N_c} \sigma_i v_{(fsi,fdi)} \tag{2}$$

where σ_i denotes the random complex amplitude corresponding to the ith clutter patch. The spatial-temporal steering vector of the ith clutter patch $v_{(f_{si},f_{di})} = v_{(f_{si})} \otimes v_{(f_{di})}$, where $\otimes$ represents the Kronecker product and

$$v_{(f_{si})} = [1, e^{-jf_{si}}, \ldots, e^{-j(N-1)f_{si}}]^T, \tag{3}$$

$$v_{(f_{di})} = [1, e^{-jf_{di}}, \ldots, e^{-j(M-1)f_{di}}]^T, \tag{4}$$

represent the spatial steering vector and the temporal steering vector respectively. The spatial frequency f_{si} and the Doppler frequency f_{di} of the ith clutter

patch are respectively expressed as

$$f_{si} = \frac{d \cos\theta_i \cos\varphi_i}{\lambda} \tag{5}$$

$$f_{di} = \frac{2v \cos(\theta_i + \theta_p) \cos\varphi_i}{\lambda} \tag{6}$$

Expression (6) can be rewritten as,

$$\left(\frac{f_r}{2f_{dm}}\right)^2 \left(\frac{2f_{di}}{f_r}\right)^2 + \cos^2\psi_i - \frac{f_r}{f_{dm}} \frac{2f_{di}}{f_r} \cos\theta_p \cos\psi_i = \sin^2\theta_p \cos^2\varphi_i \tag{7}$$

where $f_{dm} = \frac{2v}{\lambda}$ is the maximum Doppler frequency, and $\cos\psi_i = \cos\theta_i \cos\varphi_i$. Then we can see that, when $\theta_p = 0°$,the spectrum of the clutter will be distributed along diagonal in the $\cos\psi_i - \frac{2f_{di}}{f_r}$ plane. On the contrary, when $0^\circ < \theta_p < 90^\circ$, the spectrum of the clutter will be distributed along oblique ellipses, when $\theta_p = 90°$,the curves are positive ellipses. Overall, when $\theta_p \neq 0°$, the clutter spectrum will be widened with the increase of distance as shown in Fig. 1.

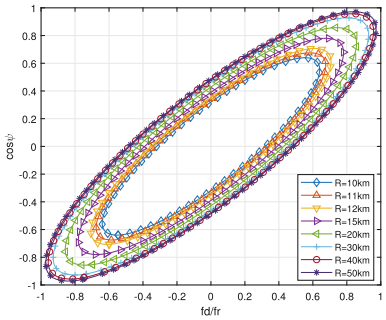


Fig. 1. The spectrum of the clutter at different distances of $\theta_p = 30°$

In conventional STAP framework, the optimum adaptive weighting vector can be expressed by

$$w_{opt} = \frac{R^{-1} v_{(fst,fdt)}}{v^H_{(fst,fdt)} R^{-1} v_{(fst,fdt)}} \tag{8}$$

where $R = R_c + R_n$ is sum of clutter and noise covariance matrix (CNCM) and usually estimated by training samples. If the samples were captured under IID condition, the estimate $\widehat{R} = \frac{1}{L} \sum_{l=1}^{L} x_l x_l^H$. L is the number of training samples and $x_l = x_c + n$ is the space-time snapshot of the clutter-plus-noise. It can

be seen from Fig. 1 that when the array is non-side-looking, the CNCM estimated from inhomogeneous training samples will be mismatched and can not accurately describe the clutter distribution characteristics of a range bin. Thus STAP performance will be degraded.

The clutter subspace U_c can be obtained by eigenvalue decomposition (EVD) of R. Suppose that the rank of clutter covariance matrix is r and can be estimated by [2], *i.e.*, $r \approx N + \frac{2v}{df_r}(M-1)$. R can be decomposed as

$$R = U_c \Lambda_c U_c^H + U_n U_n^H \tag{9}$$

where $\Lambda_c = diag(\delta_1, \cdots, \delta_r)$ consisted by r principal eigenvalues of R and $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_r$. U_n denotes the noise subspace spanned by other eigenvectors. Therefore the clutter subspace satisfies $\text{span}(R) = U_c$, implying the CNCM is of low rank. Therefore, after obtaining CNCM by sample matrix inversion (SMI) method, the STAP weight vector with EVD can be obtained by,

$$w_{EVD} = \frac{\left(I - \widehat{U}_c \widehat{U}_c^H\right) v_{(f_{st}, f_{dt})}}{v_{(f_{st}, f_{dt})}^H \left(I - \widehat{U}_c \widehat{U}_c^H\right) v_{(f_{st}, f_{dt})}} \tag{10}$$

The EVD-based algorithm effectively reducing required number of sample from $2NM$ to $2r$. However, because the computational complexity of the EVD-based method remains $O\left[(NM)^3\right]$, it is impractical for real-time processing. Therefore, subspace tracking algorithms with iterative process are more attractive.

3 STAP with Modified Subspace Tracking Algorithm

The clutter subspace can be obtained by minimizing the following minimization problem

$$\min_W J_{PAST}\left[W(L)\right] = \sum_{l=1}^{L} \rho^{L-l} \left\| x(l) - W(L) W^H(L) x(l) \right\|_2^2 \tag{11}$$

where $0 < \rho \leq 1$ is the forgetting factor intended to increase the adaptability to the inhomogeneous environment and improve the convergence speed. $W(L)$ is an $n \times r (r \leq n)$ orthonormal subspace basis to be spanned as the clutter subspace U_c. In order to solve (11) by an iterative method, it is assumed that the orthogonal subspace basis changes slowly $W(l) \simeq W(l-1)$ during the iteration. As a result $W^H(l)x(l)$ can be effectively approximated by $W^H(l-1)x(l)$. The optimization problem can be written as

$$\min_W J_{PAST}\left[W(L)\right] = \sum_{l=1}^{L} \rho^{L-l} \left\| x(l) - W(L) y(l) \right\|_2^2 \tag{12}$$

where $y(l) = W^H(l-1)x(l)$. However, when the crab angle $\theta_p \neq 0°$, the clutter spectrum distribution is distance dependence and the echo $x(l)$ is non-stationary. Thus the performance of PAST algorithm will be severely degraded and clutter subspace cannot be obtained accurately.

The purpose of clutter compensation algorithms is to make the clutter spectrum have the same structure as the detection range bin. Therefore, the clutter data close to the detection range bin data was selected as benchmark $\widehat{x}$ and the clutter subspace obtained during each iteration should be constrained like

$$\min_W J_{MPAST}\left[W(L)\right] = \sum_{l=1}^{L} \rho^{L-l} \left\|x(l) - W(L)y(l)\right\|_2^2 + \rho^{L-l} \left\|\widehat{x} - W(L)\widehat{y}(L)\right\|_2^2 \tag{13}$$

Let $\nabla J_{MPAST}\left[W(L)\right] = 0$, thus

$$W(L) = G(L)T^{-1}(L) \tag{14}$$

where, $G(L) = R_{xy}(L) + \widehat{R}_{xy}(L)$ and $T(L) = R_{yy}(L) + \widehat{R}_{yy}(L)$, the auto-correlation matrix $R_{xx} = E\left\{xx^H\right\}$ is assumed to be positive definite. $G(L)$ and $T(L)$ can be calculated by iterative processes as

$$G(L) = \rho G(L-1) + x(L)y^H(L) + \widehat{x}\widehat{y}^H(L) \tag{15}$$

$$T(L) = \rho T(L-1) + y(L)y^H(L) + \widehat{y}(L)\widehat{y}^H(L) \tag{16}$$

Using the inverse lemma of the matrix, we have

$$T^{-1}(L) = \frac{1}{\rho}T^{-1}(L-1)\left[I_r - \widehat{y}(L)\widehat{g}(L)\right]\left[I_r - y(L)g(L)\right] \tag{17}$$

where

$$g(L) = \frac{y^H(L)T^{-1}(L-1)\left[I_r - \widehat{y}(L)\widehat{g}(L)\right]}{\rho + y^H(L)T^{-1}(L-1)\left[I_r - \widehat{y}(L)\widehat{g}(L)\right]y(L)} \tag{18}$$

$$\widehat{g}(L) = \frac{\widehat{y}^H(L)T^{-1}(L-1)}{\rho + \widehat{y}^H(L)T^{-1}(L-1)\widehat{y}(L)} \tag{19}$$

Let $\Theta(L) = T^{-1}(L)$, the subspace update would be

$$\begin{aligned} W(L) = W(L-1) &+ W(L-1)\left[\widehat{y}(L)\widehat{g}(L)y(L)g(L) - \widehat{y}(L)\widehat{g}(L) - y(L)g(L)\right] \\ &+ \left[x(L)y^H(L) + \widehat{x}\widehat{y}^H(L)\right]\Theta(L) \end{aligned} \tag{20}$$

$\Theta(L)$ is updated by (17), (18) and (19).

4 Simulations

In this section, simulations are performed to verify the clutter suppression performance of MPAST in inhomogeneous clutter environment. The improvement factor (IF) is often adopted as the metric for comparison, *i.e.*,

$$IF = \frac{\left|w^H v_{(f_{st},f_{dt})}\right|^2 tr(R)}{w^H v_{(f_{st},f_{dt})} w} \tag{21}$$

Assume the echo is corrupted by additive white Gaussian noise (AWGN) and the clutter to noise ratio $CNR = 40\,\text{dB}$. The crab angle $\theta_p = 60°$, the number of spatial elements $N = 8$ and the number of temporal pulses $M = 8$, the inhomogeneous clutter is caused by Doppler expansion. The simulation parameters are listed in Table 1. In subsequent experiments, unless otherwise specified, the clutter rank was set to $r = 15$.

Table 1. Simulation parameters [8]

Parameter	Value
Number of spatial elements	8
Number of temporal pulses in a CPI	8
Radar frequency	450 MHZ
Platform velocity	200 m/s
Height of platform	12 km
Channel spacing	$\lambda/2$
Pulse repetition frequency	1200 HZ
Range sampling frequency	5 MHZ
Clutter-to-noise ratio (CNR)	40 dB

The IF performance of MPAST with ρ versus the number of training samples are first investigated and the results are shown in Fig. 2(a). The value of ρ is set to 0.9, 0.8, 0.7, 0.6, 0.5 and 0.4 respectively. It is found that the forgetting factor ρ can not be too small, because the effect of the past input data becomes insignificant, the error in the estimation of clutter subspace would be increased. Similarly, ρ can not be too large, because more information from previous input data used for estimation of clutter subspace would lead to much deviation in the inhomogeneous environment. Based on the simulation results, ρ equals to 0.6 is selected in the following processing.

The IF performance of MPAST varied with different rank is shown in the Fig. 2(b). According to the Reed-Mallett-Brennan (RMB) rule [2], all these methods can obtain a high IF capability when the selected rank r is larger than 15. It can be seen from this figure that MPAST and EVD algorithms comply with

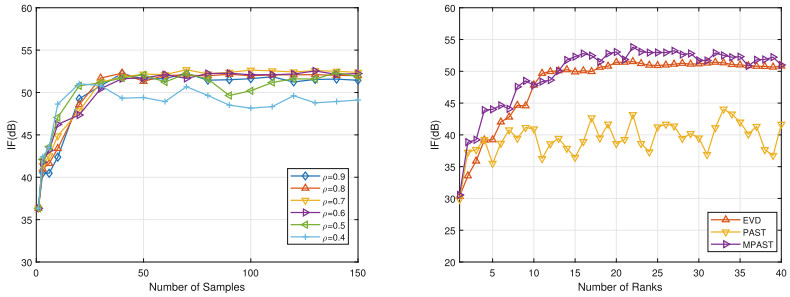


Fig. 2. (a) IF performance of MPAST with different forgetting factors ρ versus the number of training samples. (b) IF performance of EVD, PAST and MPAST versus the number of selected rank.

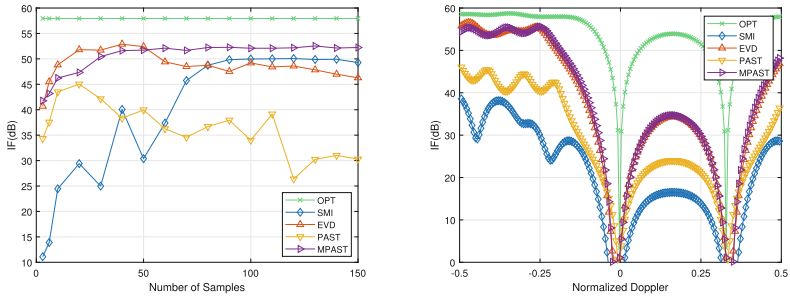


Fig. 3. (a) IF performance of SMI, EVD, PAST, and MPAST versus the number of training samples. (b) Performance of SMI, EVD, PAST, and MPAST based on 60 training samples.

RMB criterion, while PAST oscillates due to data heterogeneity. Based on the simulation results, r equals to 15 is selected in the following processing.

Figure 3(a) shows the IF performances of SMI, EVD, PAST and MPAST as the number of training samples increases. It can be seen that in the inhomogeneous clutter environment, the PAST algorithm fails with the increase of inhomogeneous training samples. On the contrary, MPAST algorithm can effectively improve the subspace tracking performance with more training samples in inhomogeneous clutter environment. In addition, it is superior to SMI and EVD algorithms when the number of non-stationary samples increases. Figure 3(b) shows the results when 60 training samples from 10 km range bins to 11.5 km range bins are used. The advantages of MPAST performance can be intuitively seen.

In particular, this method can be applied with some compensation method, *i.e.*, clutter data can be compensated simply before MPAST algorithm.

5 Conclusion

In this paper, a modified subspace tracking algorithm based on PAST for space-time adaptive processing in inhomogeneous clutter environment is proposed. In inhomogeneous clutter environment, the performance of the proposed method is similar to clutter compensation and can achieve better performance than direct SMI and EVD. The iterative process provides a tailored computation complexity. In addition, the proposed method can be combined with the clutter compensation methods to obtain better clutter suppression performance.

Acknowledgements. The authors much appreciate the supports from team members Mr. Jun-Hao Cui and Prof. Jing Liang.

References

1. Brennan, L.E., Reed, L.S.: Theory of adaptive radar. IEEE Trans. Aerospace Electr. Syst. AES **9**(2), 237–252 (1973)
2. Reed, I.S., Mallett, J.D., Brennan, L.E.: Rapid convergence rate in adaptive arrays. IEEE Trans. Aerospace Electr. Syst. AES **10**(6), 853–863 (1974)
3. Kreyenkamp, O., Klemm, R.: Doppler compensation in forward-looking STAP radar. IEE Proceed. Radar Sonar Navigation **148**(5), 253–258 (2001)
4. Leatherwood, D.A., Melvin, W.L.: Adaptive processing in a nonstationary space-borne environment. In: 2003 IEEE Aerospace Conference Proceedings. IEEE (2003)
5. Zatman, M.: The properties of adaptive algorithms with time varying weights. Sensor Array and Multichannel Signal Processing Workshop (2000)
6. Lapierre, F.D., Ries, P., Verly, J.G.: Foundation for mitigating range dependence in radar space-time adaptive processing. IET Radar Sonar? Navigation **3**(1), 18–29 (2008)
7. Yang, B.: Projection approximation subspace tracking. IEEE Trans. Signal Processing **43**(1), 95–107 (1995)
8. Cui, J.-H., Chen, Z.-X.: A modified subspace tracking algorithm based on the constrained projection approximation for space-time adaptive processing. https://assets.researchsquare.com/files/rs-1676126/v1_covered.pdf?c=1654189102

Integrated Recognition of Multi-class Battlefield Signals Based on Multi-resolution Analysis

Hong Hu, Yue-Feng Li, and Zhang-Xin Chen(✉)

University of Electronic Science and Technology of China, Chengdu 611731, China
{honghu,liyuefeng00}@std.uestc.edu.cn,
zhangxinchen@uestc.edu.cn

Abstract. In order to solve the problem of accurately acquiring data signals propagated in complex battlefield environments, this paper designs a classification decision tree based on support vector machine (SVM) algorithm and uses a multi-resolution analysis method based on the wavelet transform to extract the detailed features of the signals, which compose a collection of energy features to conduct network training. The results show that the support vector machine decision tree classifier trained by the multi-resolution analysis method can extract features for recognition from the simulated battlefield environment with noise and interference.

Keywords: Battlefield signal · multi-resolution analysis · wavelet transform · support vector machine · decision tree

1 Introduction

The battlefield situation is changing rapidly, in which the radio signals carrying various information may a key role. Among these systems, tactical data chain, TACAN navigation signals and recognition Friend-or-Foe signals are frequently utilized.

The recognition of battlefield signals is the most critical part for targeted information interception and information countermeasures. At present, there are few domestic studies on the comprehensive recognition of battlefield signals, which only cover the recognition of single class of battlefield signals, for example, Bin-gang Zhu [1] detects Link4A signals based on instantaneous features and delayed correlation method, Tang Suhua [2] detects Link11 based on combined parameters, Yu-cai Pang [3] identifies TACAN signals using parameter estimation and pattern matching algorithm, Wei-ke Li [4] identifies Mode 5 signals based on time domain and coding features to identify Mode 5 signals for enemy recognition, etc. Therefore, few scholars have considered the coexistence of these types of signals in complex battlefield environments, and our study is necessary.

In this paper, we use the time-frequency analysis method of wavelet transform to obtain the detailed features of the signal as the classification principle, then

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 301–308, 2023.
https://doi.org/10.1007/978-981-99-1260-5_38

use the multi-resolution analysis method to extract the energy features, finally design the support vector machine training to obtain the classification decision tree according to the obtained features. The simulation experiments confirm that the proposed method is helpful to obtain stable and accurate recognition rate in the presence of low signal-to-noise ratio and interference.

2 Analysis of Battlefield Signals

In this paper, the modulation method, frequency band range and other technical regimes of the studied battlefield signals are shown in Table 1.

Table 1. Five types of battlefield signal technology system table

Name	Modulation	Frequency band	Code rate
Link4A	2FSK	255–400 MHz	5 kbps
Link11	$\pi/4$-DQPSK	2–30 MHZ and 225–400 MHz	1364 kbps and 2250 kbps
Link16	MSK	960–1215 MHz	25.8 kbps–238 kbps
TACAN	PAM	962–1213 MHz	7155 pulses/sec
IFF Mode 5	MSK	1030 MHz(Ask) or 1090 MHz (Answer)	16 Mbps

Link4A data chain uses single frequency time division multiple access technology, 2FSK modulation, and works in the UHF band of 255–400 MHz at a code element rate of 5 kbps, with an RF channel spacing of 25 KHz in the frequency band. Link11 data chain is a high capacity data chain working in the HF and UHF bands, using standard defined message types and formats, with 16 audio parallel DQPSK modulation; using two different code element rates, 1364 kbps and 2250 kbps respectively. Link16 data chain works in the L-band of 960–1215 MHz with code element rates ranging from 25.8 kbps–238 kbps, using MSK modulation with continuous phase and direct spread spectrum sequence, frequency hopping, time hopping, channel coding, R-S error correction coding and other anti-jamming measures [5]. TACAN signal is a pulse sent in the air by the ground beacon station 15 Hz 135 Hz compound modulation of the rotating field, using pulse envelope modulation, also working in the L-band of 962–1213 MHz, with a transmission rate of 7155 pulses/sec [3]. The Mode 5 signal for enemy recognition also uses MSK modulation, and the interrogation and answer signals are heterodyne, at 1030 MHz and 1090 MHz, respectively, also in the L-band, with a data rate of 16 Mbps.

3 Recognition Principle

3.1 Wavelet Transform

The key idea of time-frequency analysis method is a series of orthogonal or non-orthogonal basis functions and signal as inner product, the value obtained reflects the degree of similarity between that basis function and the signal . Wavelet transform is a kind of time-frequency analysis method, using finite-length scalable transform of wavelet basis, with the signal adaptively changing the width of the time-frequency window. It is an ideal tool for signal processing and feature extraction, and its transformation formula is

$$WT(\alpha, \tau) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{\infty} s(t) * \Psi(\frac{t-\tau}{\alpha}) dt \tag{1}$$

In the above equation,$s(t)$ is the input signal, and $\Psi(\frac{t-\tau}{\alpha})$is the wavelet function, where τ is the time shift and α is the scale, which respectively control the frequency shift and scaling of the wavelet function.

3.2 Multi-resolution Analysis

Multi-resolution analysis, also known as multi-scale analysis, is a theory created by S. Mallat in his study of image processing problems,it establishes a link between wavelet transforms and digital filters, whose structure tree is shown in Fig. 1.

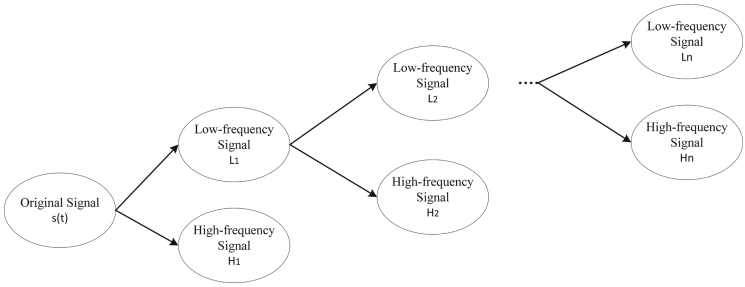


Fig. 1. Structural tree diagram for multi-resolution analysis

In the figure, the resolution analysis decomposes the original signal s(t) at different resolution levels n to obtain two sub-signals, the high-frequency signal H_n and the low-frequency signal L_n. The low-frequency signal L_n is obtained by low-pass filter after wavelet decomposition, which reflects the contour information of the signal, while the high-frequency signal H_n is obtained by high-pass filter, which reflects the detail information of the signal and contains noise and mutation information. The higher the number of resolution steps, the more

details of the high frequency signal and the smoother the low frequency signal will be, but since the high frequency signal contains noise, the higher the number of steps, the more details occupied by the noise [7], so in practice, the number of resolution levels should be decided according to the details reflected by the high frequency signal.

3.3 Feature Extraction

Assuming that the input original signal s(t) is in the space $L^2(R)$ of the time domain and its signal energy is defined as the square of the Frobenius parametrization, then

$$E[s(t)] = ||s(t)||_F^2 = \int |s(t)|^2 dt \tag{2}$$

Wavelet energy is the energy of the detailed signal obtained after multi-resolution analysis of the original signal, when the wavelet basis function satisfies orthogonality, which satisfies Rayleigh's energy conservation theorem [6], that is

$$||s(t)||_F^2 = \sum_i^n |WT(i,k)|^2 \tag{3}$$

where n is the total number of resolution levels for multi-resolution analysis, i is the number of a particular resolution level, and k is the window translation value.

The wavelet energy at a single resolution level can be expressed as

$$E_i = \sum_k |L_i(k)|^2 \tag{4}$$

Then, the energy of the signal can be expressed as

$$E[s(t)] = \sum_i E_i = \sum_i \sum_k |L_i(k)|^2 \tag{5}$$

As a result, the energy characteristics of the signal are obtained using multi-resolution analysis methods, and thus the detailed information of the signal is fully recorded for subsequent support vector machine training.

3.4 Recognition Process

The specific process of identifying battlefield signals based on wavelet transform is as follows [8], and its flow block diagram is shown in Fig. 2:

1) Capturing the original signal using receiving equipment such as radar, antennas, etc.;
2) Filter, down-conversion and extraction to obtain the digital baseband signal;
3) Obtain the energy characteristics of the signal using the multi-resolution analysis method using equation (5);
4) Select the training sample set for training the support vector machine to obtain the SVM decision tree classifier;
5) Recognition and classification.

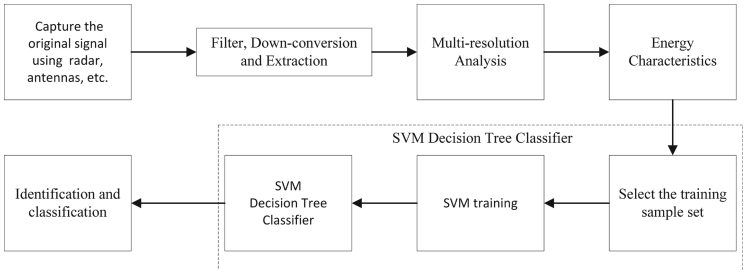
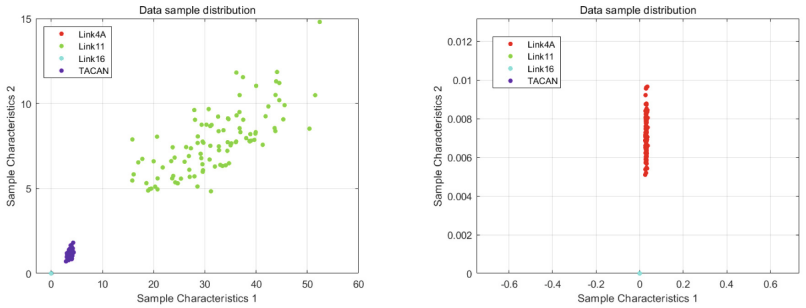


Fig. 2. Recognition process of SVM decision tree classifier based on multi-resolution analysis

4 Design of SVM Decision Tree Classifier

In this paper, we study five classes of battlefield signals, which is a multi-classification design problem. Although SVM is a typical two-class classifier, i.e., it only classifies objects into positive and parent classes, we can use the "one class to the rest" decision tree approach to solve one two-class classification problem at a time. That is, each recognition treats one class of battlefield signals as the positive class and the rest as the negative class, thus obtaining a two-class classifier, so that we can use five two-class classifiers to achieve all recognition.



(a) Original view of the detail features of the four types

(b) Enlarged view of the detail features of Link4A and Link16

Fig. 3. Detailed feature map of the four types of battlefield signals

As we can see in the analysis above, the four types of signals, Link4A, Link11, Link16 and TACAN, use different modulation methods, so the differences are large and the extracted signal detail features are shown in Fig. 3. In the figure, the coordinate sample features represent the wavelet transform values of the high frequency signals extracted at a certain resolution level. Figure 3(a) can clearly show the detailed features of Link11, TACAN, and Link16, but there is

confusion between the features of Link16 and Link4A, so we enlarge the feature values of these two signals in the Fig. 3(b) to show that the distinction between them is still obvious, about four orders of magnitude different.

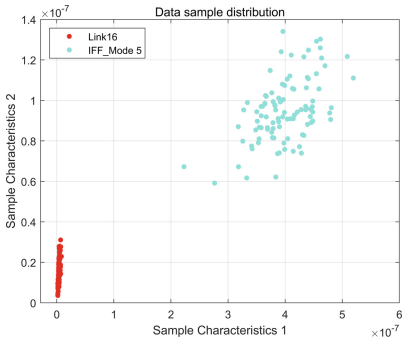


Fig. 4. Detail feature diagram of Link16 and IFF Mode 5

However, Link16 and IFF Mode 5 both use MSK modulation, so the difference is little. The extracted signal detail features are shown in Fig. 4, and it can be seen that the feature values of both are in the same order of magnitude, although there are differences but the differences are difficult to use to distinguish. Therefore, the training process of SVM, if both Link16 and IFF Mode 5 are put into the training sample set, the recognition effect of the final obtained classifier is poor. The training recognition process of decision tree classifier in this paper is shown in Fig. 5.

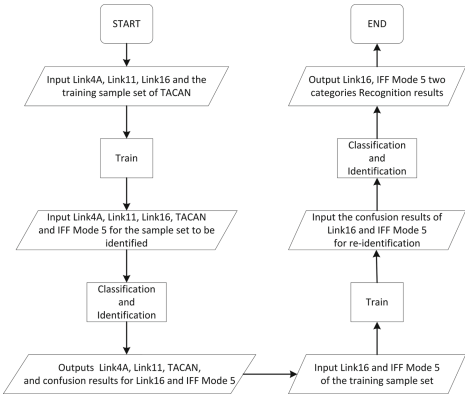
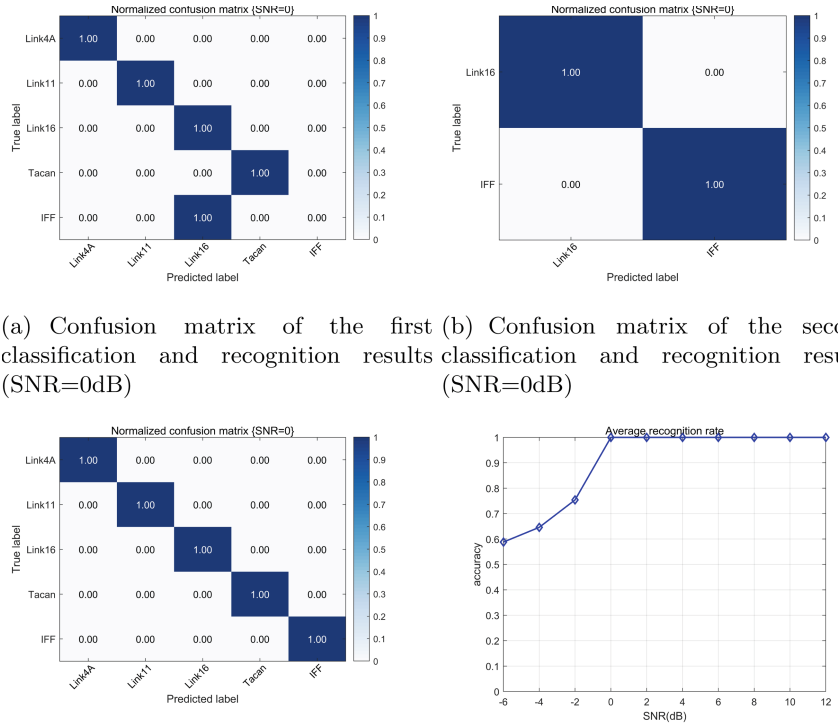


Fig. 5. SVM decision tree classifier training recognition flow chart

5 Simulations

In this paper, five classes of battlefield signals, Link4A, Link11, Link16, TACAN, and IFF Mode 5, were generated using random digital sequences based on the data in Table 1 with the effects of additive Gaussian noise and Rayleigh fading channels. For the simulation, 1100 samples were generated for each class of signals, of which 1000 samples were used as the training set and the remaining 100 samples were used as the test set. Then, the energy features of the samples are extracted according to the signal recognition process in Fig. 2, and the training and step-by-step recognition of the support trainer decision tree classifier is carried out according to the procedure shown in Fig. 5. Finally, the simulation results is obtained as shown in Fig. 6.



(a) Confusion matrix of the first classification and recognition results (SNR=0dB)

(b) Confusion matrix of the second classification and recognition results (SNR=0dB)

(c) Confusion matrix of overall classification and recognition results (SNR=0dB)

(d) Recognition rate of SVM decision tree classifier with different signal-to-noise ratios

Fig. 6. Simulation results

Figure 6 (a) shows the recognition results of the model obtained by inputting only four classes such as Link4A, Link11, Link16 and TACAN as the sample training set in the first training, and it can be seen that all the battlefield signals

can be accurately recognized except Link16 and IFF Mode 5 which are confused as one class. Figure 6 (b) shows the recognition results obtained by inputting two classes of signals such as Link16 and IFF Mode 5 as the sample training set for the second training, and inputting Link16 and IFF Mode 5 that were confused in the first recognition. Figure 6 (c) shows the overall recognition results after two classifications of recognition, and it can be seen that the recognition model designed in this paper has a high classification accuracy, and the average recognition rate under different signal-to-noise ratios is shown in Fig. 6(d).

It can be seen that the recognition effect of the classifier designed in this paper is very stable in different SNR environments, and the recognition rate is higher than 99% at low SNRs such as 0 dB.

6 Conclusions

This paper designs a unique method to train the recognition model based on the multi-resolution analysis method by wavelet transform. It can extract the detailed features of signals, and achieves the accurate recognition of five types of battlefield signals. The simulation experiments prove that the multi-resolution analysis method can extract the features of the signal with low signal-to-noise ratio (0 dB) and Rayleigh fading interference, and the proposed SVM decision tree possesses a good recognition effect.

References

1. Zhu, B.-G., Peng, S.-R., Su, Y.-W., Tang, T.: Automatic identification method of link4a. In: 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing, vol. 2, pp. 239–243. IEEE (2009)
2. Tang, S., Yin, H., Xu, P.: A method for modulating recognition with combined parameters. J. Data Acquisit. Process. **18**(1), 27–31 (2003)
3. Pang, Y.-C., Zhang, C.-Z., Dong, H., Liao, G.: Reconnaissance receiving and processing technique for Tacan signals. Dianxun Jishu/ Telecommun. Eng. **51**(10), 67–70 (2011)
4. Weike, L.: Identification of mark xiia iff signals based on time domain and coding characteristics. J. Ordnance Equip. Eng. **37**(7), 153–157 (2016)
5. Zheng, W., Jin, H., Liu, Y., Yu, Q.: Analysis and research on TTNT data link. In: 2017 5th International Conference on Frontiers of Manufacturing Science and Measuring Technology (FMSMT 2017), pp. 661–666. Atlantis Press (2017)
6. Wu, H., Xiao, X.-Y., Shen, R.-J.: Lightning strike and fault identification by the wavelet energy spectrum and neural network method. Gaodianya Jishu/ High Volt. Eng. **33**(10), 64–68 (2007)
7. Xiangkui, W.: ECG baseline drift depress algorithm based on multi-resolution analysis. Computer Eng. Design **29**(13), 3482–3484 (2008)
8. Liu, G., Xu, J.: Neural network recognition algorithm of breath sounds based on SVM. J. Commun. **35**(10), 218–222 (2014)

Research on Adaptive Cancellation Technology

Chao Wang(✉), Wen Pan, Jun Lu, Xin Hao, and Lei Guo

Beijing Aerospace Long March Vehicle Research Institute, Nan Da Hong Men Road, No.1, Beijing 100076, China
87750576@qq.com

Abstract. An adaptive cancellation technology in the mode of simultaneous receiving and transmitting is studied. The performance of complex VS-LMS algorithm under different interference patterns and parameters is analyzed. An electronic cancellation system is designed. A training storage method is proposed, which can shorten the convergence time of the filter. Simulation results show that this method can effectively realize interference cancellation, and the cancellation ratio can reach 25 dB.

Keywords: Adaptive filter · Adaptive cancellation · LMS

1 Introduction

In many electronic systems, the receiving antenna and transmitting antenna are very close. The system is prone to self-excitation. In order to solve this problem, many adaptive interference cancellation methods have been proposed. The least mean square (LMS) algorithm is the most commonly used method [1, 2].

Since LMS algorithm is proposed, mangy researchers have studied various structures and adaptation algorithm for filter.In 2015, the University of Electronic Science and Technology of China realized echo cancellation in the jamming environment, using DLMS algorithm. The maximum cancellation ratio is 15 dB [3].

In this paper, an adaptive echo cancellation method based on VS-LMS algorithm is proposed, and the performance of the algorithm is analyzed.

Since the convergence speed is the key index of the cancellation system, in order to improve the convergence speed, a training-storage method is proposed in this paper. Firstly, the system is trained according to the first signal data, and the convergent filter coefficients are obtained and stored. When the next signal comes, the stored filter coefficients are directly used as the starting coefficients of the system iteration.

2 Adaptive Cancellation Technology

Adaptive filter is shown as Fig. 1.The reference signal J0 is input into the adaptive filter, and the output signal of the filter is y(n), which cancels the received signal d(n) from the receiving antenna. The target of cancelling is to remove the J1 component from d (n), and the error signal e (n) is fed back to the filter, which is used as an auto-adjusting

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 309–315, 2023.
https://doi.org/10.1007/978-981-99-1260-5_39

parameter to adjust the next filter factor. The error signal e (n) is minimized by iteration, and the noise signal J1 is suppressed to the maximum extent. Ideally, the remaining error signal contains only signal S.

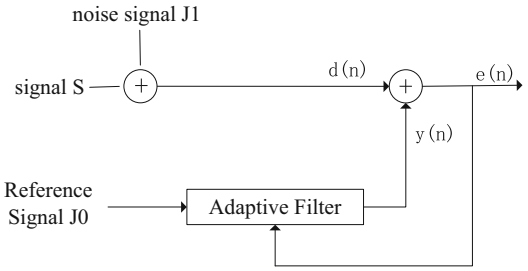


Fig. 1. Basic principles of adaptive filter control

In the filter design, minimum mean square error criterion is used. In the cancellation system, it is hoped that the algorithm will have large step-size factor to accelerate convergence in the initial stage and reduce steady-state error in the later state. In order to solve this contradiction further, a variety of variable step size LMS algorithms based on a fixed step size are proposed.

Since most of the electronic systems work in I/O mode, this paper uses VS-DLMS algorithm with complex number structure [4].

The VS-LMS algorithm can be given in the form of three basic relationships

$$\begin{aligned} \mathrm{y}_I(n) &= W_I^T(n) \times D_I(n) + W_Q^T(n) \times D_Q(n) \\ \mathrm{y}_Q(n) &= W_I^T(n) \times D_Q(n) - W_Q^T(n) \times D_I(n) \end{aligned} \tag{1}$$

$$\begin{aligned} D_I(n) &= S_I(n) + J_I(n) \\ D_Q(n) &= S_Q(n) + J_Q(n) \\ e_I(n) &= D_I(n) - y_I(n) \\ e_Q(n) &= D_Q(n) - y_Q(n) \end{aligned} \tag{2}$$

$$\begin{aligned} u_I(n) &= u_0 \times \left\{1 - e^{-\alpha \times |e_I(n)|^2}\right\} \\ u_Q(n) &= u_0 \times \left\{1 - e^{-\alpha \times |e_Q(n)|^2}\right\} \end{aligned} \tag{3}$$

$$\begin{aligned} W_I(n+1) &= W_I(n) + u_I(n) \times \big(e_I(n) \times D_I(n) - e_Q(n) \times D_Q(n)\big) \\ W_Q(n+1) &= W_Q(n) + u_Q(n) \times \big(e_I(n) \times D_Q(n) - e_Q(n) \times D_I(n)\big) \end{aligned} \tag{4}$$

$\mathrm{y}_I(n)$ and $\mathrm{y}_Q(n)$ are the output of the filter. $W(n)$ is the filter coefficient at time n. $D(n)$ is the reference signal. $e_I(n)$ and $e_Q(n)$ are the error signal. u_0 is initial step size. $u_I(n)$ and $u_Q(n)$ are the step size at time n. α is a variable step factor.

2.1 Step-Size Factor

The step factor is proportional to the convergence rate of the adaptive filter. If the step size is too large, the filter will not converge and the function will fail. If the step size

is too small, the convergence speed is too slow, which affects the response speed of the system. The range of step factor u is $0 < u < 1/\lambda_{max}$, and λ_{max} is the maximum eigenvalue of the reference signal correlation matrix. In formula (3), two convergence factors are set for the data of channel I and Q respectively, so that the algorithm can independently adjust the step size of the two channels and reduce the cross coupling of the two channels noise.

In formulas (3), α is a variable step size factor. When α is large, the influence of the step is greater, which can accelerate the convergence. Therefore, on the premise of satisfying the convergence, α should choose as large a value as possible.

2.2 Filter Order

Since the reference signal autocorrelation matrix is positive definite, the maximum eigenvalue of the autocorrelation matrix is $\lambda_{max} \leq tr[R]$, where $tr[R]$ is the trace of R. [5, 6] So the step factor can be expressed as

$$0 < u < 1/tr[R] \tag{5}$$

Under the same data input, if the filter order becomes higher, the $tr[R]$ will be greater, the maximum value of the step factor will become smaller, and the convergence speed becomes slower. If the order too small, the system cannot converge to the optimal solution. Considering the performance of the algorithm and the complexity of the implementation, the following simulation uses an 8-order filter.

2.3 Simulation and Analysis

The cancellation performance of VS-LMS algorithm in electronic systems is simulated and analysed. The signal is set as single carrier or linear frequency modulation (LFM).The interference forms are white noise or forwarding signal. The signal to noise ratio is uniformly set to 25 dB.

Figure 2 shows the mean square deviation (MSD) of LFM signal and single carrier signal. The convergence speed and stability are same with two signal forms.

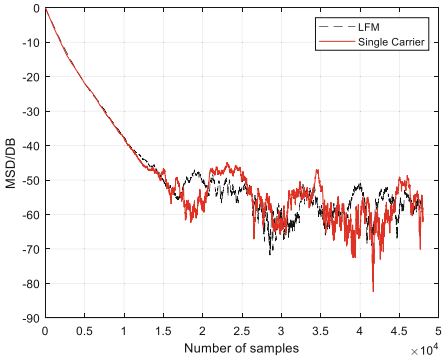


Fig. 2. MSD with white noise

Figure 3 shows the cancellation ratio curve. It can be seen that the cancellation ratio rises rapidly with the iteration, and finally reaches a stable state. The cancellation ratio of single carrier signal is above 22 dB. The cancellation ratio of LFM signal is lower, about 20 dB. This is because the frequency of the LFM is always changing, which requires a high convergence speed of the filter, resulting in a reduction in the cancellation ratio.

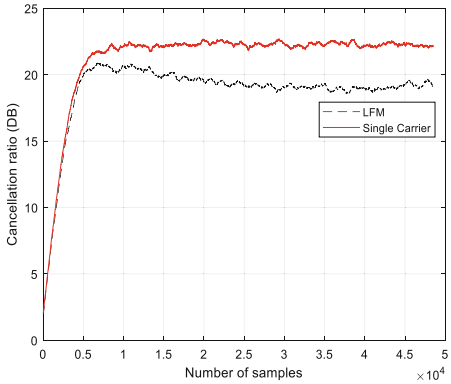


Fig. 3. Cancellation ratio with white noise

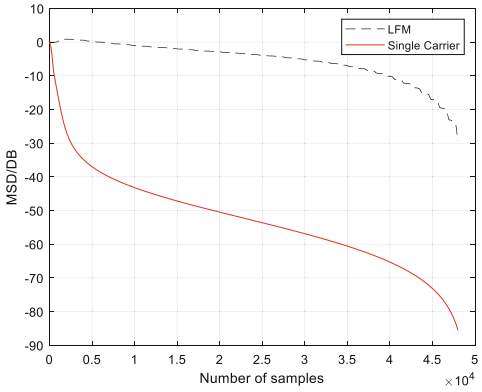


Fig. 4. MSD with forwarding signal

Figure 4 shows the MSD curve of LFM and single carrier signal. From the curve, the convergence rate of single carrier signal is significantly higher than that of LFM signal. The reason is that the LFM signal is always changing in amplitude and frequency. This also causes the cancellation ratio to fluctuate periodically, and the cancellation ratio is lower than that of single carrier signal.

Figure 5 shows the cancellation ratio curve. It can be seen that the cancellation ratio increases rapidly with the iteration. The cancellation ratio of single carrier signal is higher than LFM. The cancellation ratio of single carrier signal has a continuous

upward trend, which indicates that the residual energy will be smaller and smaller. In time domain, with the iteration of the filter, after entering the convergence interval, the output signal amplitude will gradually decrease, which indicates that the original signal S will be eliminated by the filter.

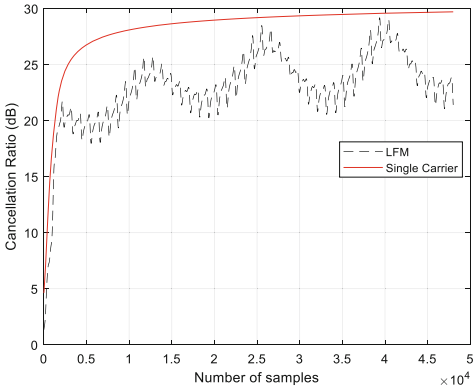


Fig. 5. Cancellation ratio with forwarding signal

The retransmitted signal is a copy of the original signal. They have high degree of correlation. The original signal and retransmitted signal J has the same frequency, only amplitude and phase are different. The new synthetic signal J is

$$J = \cos(f_0 t) + A\cos(f_0 t + \varphi) = \sqrt{1 + A^2}\cos(f_0 t + \theta) \tag{6}$$

The signal J is highly correlated with the original signal J_0. Therefore, when the signal J1 is related to the signal S. This is very disadvantageous to the cancellation system.

3 Training-Storage Mode Analysis

In order to solve this problem and save convergence time, a training-storage method is proposed in this paper. Firstly, the system receives the signal, obtains the optimal solution through training, and stores the filter coefficients. When the next signal arrives, the last stored coefficient is directly used for the single carrier signal filter, and the coefficient is no longer updated. For the LFM signal, because its frequency is changing all the time, the filter coefficient cannot be fixed. The stored coefficient is used as the initial value of the filter in the iteration. The signal processing flow is shown in Fig. 6 below.

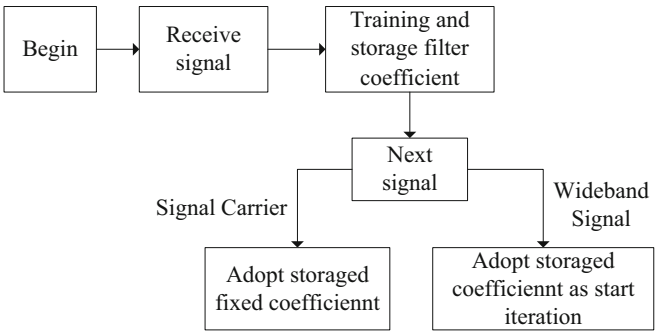


Fig. 6. Training - storage mode signal processing flow

The single carrier signal and LFM signal are simulated respectively. Cancellation ratio curve is shown in the Fig. 7 below. For a single carrier signal, the cancellation ratio quickly reaches a constant value. For LFM signals, the change of cancellation ratio is similar to that in Fig. 5, but the convergence speed is faster, and the cancellation ratio is higher than that in untrained mode.

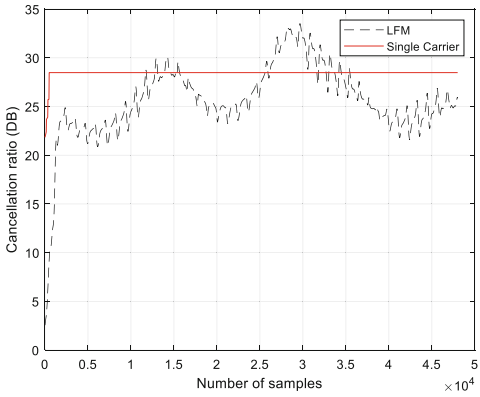


Fig. 7. Cancellation ratio in trained mode

4 Conclusions

This paper studies the application of VS-LMS adaptive algorithm in cancellation system. Simulation results show that VS-LMS algorithm is applicable to various signals and has good cancellation effect under different noise patterns. The proposed training storage method can effectively realize the cancellation of coherent signal, and can achieve better convergence speed and higher cancellation ratio, with high system stability.

References

1. Widrow, B., Stearns, S.: Adaptive Signal Processing. Prentice-Hall, Englewood Cliffs (1985)
2. Kabal, P.: The stability of adaptive minimum mean square error equalizers using delayed adjustment. IEEE Trans. Commun. **31**(3), 430–432 (I983)
3. Yuan, S., He, Z., Huang, D.: Frequency-domain equalization of channel inconsistency for wideband navigation anti-jamming receiver based on uniform circular array. In: 2014 The 3rd International Conference on Communications, Signal Processing, and Systems (2014)
4. Mao, Z., Wang, H.X., Song, H., etal.: Studies on improving the convergence ability of complex LMS algorithm. J. China Acad. Electron. Inf. Technol. **6**(3), 275–278 (2011)
5. Narayananrm, M.N.S.: Delay and summed adaptive noise waveforms for target matched radar detection. In: Proceedings of the 2013 22nd International Conference on Noise and Fluctuations. ICNF 2013, 24–28 June 2013, Montpellier, France. IEEE Computer Society (2013)
6. Schuldt, C., Lindstrom, F., Li, H., et al.: Adaptive filter length selection for acoustic echo cancellation. Sig. Process. **89**(6), 1185–1194 (2009)

Check for updates

Radiation Source Identification by Feature Fusion

Kaishuang Yu, Xiaoying Wang, and Zhangxin Chen(✉)

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
yukaishuang@std.uestc.edu.cn , zhangxinchen@uestc.edu.cn

Abstract. In order to solve the problem of accurately identifying the emitter in complex environments, time-frequency analysis is performed by Hilbert-Huang transform(HHT), integrated rectangular bispectrum is used to extract features from the received signal. The features extracted by the two methods are fused and a network training is conducted to make final decision. The simulation results show that the emitter individual recognition method based on feature fusion is more stable and has higher recognition accuracy in the case of low signal-to-noise ratio.

Keywords: Specific emitter identifying · Hilbert Huang transform · feature fusion · neural networks

1 Introduction

Radiation source identification, also known as radio frequency fingerprint identification or specific emitter identification (SEI), is the process to extract the characteristics of emitter by the received radio signal, then build the fingerprints of radiation sources and can make out the received signal belongs to which emitter. Due to the tolerance effect of the hardware circuit, radiation sources of the same batch produced by the same manufacturer have some differences, therefore, the radiation sources have unique fingerprint characteristics [1]. We can match the corresponding radiation source individuals according to the fingerprints. Viewpoints, to ensure the accuracy of radiation source identification and reliant results in practice, the fingerprints should be independent, stable, unique, and applicable.

To identify the radiation source with higher accuracy, there is no doubt that it is necessary to extract effective features from multiple. Therefore, the strategy of high-level feature fusion for identification of radiation sources is expected. Compared with a single feature, fused results would have more complete and richer information. In this paper, one-dimensional features and two-dimensional features are imported to two neural networks, and then the high-dimensional features extracted by the network are fused. Finally, the fully connected neural network is used for feature mapping, and the output layer uses the SoftMax function for classification and recognition. In the low signal-to-noise ratio environment (SNR<=10 dB), the recognition accuracy of the proposed method is higher than other three deep learning methods, and the recognition rate is more stable.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 316–323, 2023.
https://doi.org/10.1007/978-981-99-1260-5_40

2 Radiation Source Fingerprint Feature Extraction

2.1 Signal Model

The identification of the radiation source is first to extract the characteristics of the transmitter from the received signal. The modulated signal needs to be amplified by power amplifier before being transmitted through the antenna. Nonlinearity is the common characteristic of power amplifiers and often results in amplitude compression and amplitude phase conversion. For memoryless narrowband power amplifiers, the Taylor series model is usually used to describe it's nonlinear distortion. Assume that the modulated quadrature signal:

$$s(t) = \mathrm{Re}\{s_0(t)e^{j2\pi f_c t}\} \tag{1}$$

where $s_0(t)$ represents the baseband signal,f_c is the carrier frequency, Then the nonlinearly distorted signal output by the RF power amplifier can be modeled as:

$$X(t) = \sum_{i=1}^{L} {\lambda_i}^k.(s(t))^i \tag{2}$$

where L is the order of the polynomial, $\{{\lambda_1}^k, {\lambda_2}^k ... {\lambda_L}^k\}$ represent the nonlinear response coefficients of the k th radiation source, usually ${\lambda_1}^k = 1$.

In order to observe the nonlinearity described by the Tylor model intuitively, the sinusoidal signal is used as the RF transmission signal. The nonlinear response under different nonlinear parameters,is shown in Fig. 1.

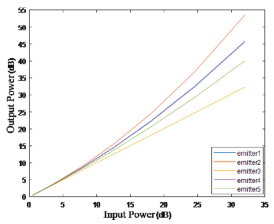


Fig. 1. Nonlinear response curves under different parameters as Table 1

It can be seen from the figure that the input and output power of the radiation source are different, and the nonlinear response between the radiation sources has a slight gap, and the gap becomes larger as the input power increased.

In a communication system, according to different propagation distances and communication methods, communication scenarios can be divided into single-hop mode and relay mode. The single-hop mode, as the name implies, is that the transmitted signal directly reaches the receiver after being transmitted through the channel and the received signal can be expressed as

$$\begin{aligned} R(t) &= \alpha.X(t) + n(t) \\ &= \alpha \sum_{i=1}^{L} {\lambda_i}^k.(s(t))^i + n(t) \end{aligned} \tag{3}$$

where,α is the fading coefficient of the channel,$n(t)$ represents channel noise, in general, it is considered to be additive white Gaussian noise. Under the Gaussian white noise channel, the channel fading coefficient α is often set as 1.

2.2 Hilbert Time Spectrogram

Time-frequency analysis by Hilbert-Huang transform (HHT) is called the Hilbert time-spectrogram. It is usually used for the analysis and processing of nonlinear and non-stationary signals [2]. HHT essentially includes two steps, one is empirical mode decomposition (EMD), the other is Hilbert transform. The former is a decomposition process. The received signal is decomposed into a set of intrinsic mode functions (IMFs) by EMD, and the Hilbert time spectrum is obtained by performing Hilbert transform on each IMFs [5].Assuming that the received signal is $x(t)$, EMD decomposition can be expressed as

$$x(t) = \sum_{i=1}^{N} c_i(t) + r_N(t) \tag{4}$$

where N is the EMD decomposition order,$c_i(t)$ is the i-th IMF component,$r_N(t)$ is the residual signal component. After EMD decomposition, the residual signal $r_N(t)$ usually is a constant or monotonic function, which is generally ignored. Hilbert transform is then performed on each IMF component to get the spectrum $\widehat{c_i}(t)$:

$$\widehat{c_i}(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{c_i(\tau)}{t - \tau} d\tau \tag{5}$$

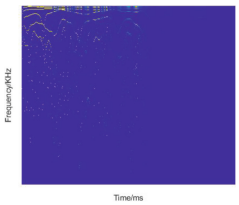


Fig. 2. HHT Spectrum of Radiation Source 1

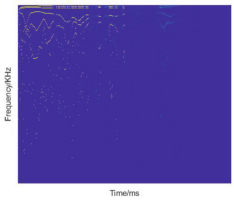


Fig. 3. HHT Spectrum of Radiation Source 2

The instantaneous amplitude and instantaneous frequency of each IMF can be derived from (7) and (8)

$$\theta_i(t) = \arctan \frac{\hat{c}_i(t)}{c_i(t)} \tag{6}$$

$$a_i(t) = \sqrt{c_i(t) + \hat{c}_i(t)} \tag{7}$$

$$\omega_i(t) = \frac{d\theta_i(t)}{dt} \tag{8}$$

The received signal can be expressed as:

$$r(t) \approx \mathrm{Re}\left\{\sum_{i=1}^{N} a_i(t) e^{j\int \omega_i(t)dt}\right\} \tag{9}$$

From Eq. (9), it can be known that based on the instantaneous amplitude and instantaneous frequency of each IMF component, the Hilbert time spectrogram $H(\omega, t)$ can be calculated.Figure 2 and 3 show the modulation diagram of the Hilbert time spectrums of two sources with QPSK. It can be seen that the time spectrum of different radiation sources are different but not obvious, so further analysis and processing are required.

3 High-Level Feature Fusion

A strategy of high-level feature fusion is proposed to contain more information shown in Fig. 4. It includes two branches,one is the residual unit structure, the other is the Efficient Channel Attention(ECA) module. As the first branch, the original time series and it's rectangular bispectral feature vector are the input of residual network(ResNet), for the other branch, the Hilbert time spectrum is the input to the ECANet module.

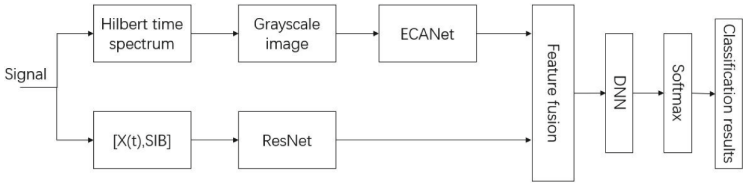


Fig. 4. Flowchart of feature fusion algorithm

The one-dimensional residual network designed in this experiment has a total of 11 layers [3], including four residual structural units, and each residual unit structure contains two one-dimensional convolutional layers. The neural network used in this paper is essentially a basic convolutional neural network. The ECA module is used in the convolutional network to capture effective feature channels and further improve the network recognition efficiency. First, the fingerprint feature information contained in the feature image is extracted through 4 layers of convolutional layers, and then the ECA module is used to highlight the effective fingerprint features. The specific network structure is:

The number of convolution kernels in the first and second convolutional layers is 64, and the convolution kernel sizes are 5×5 and 3×3 respectively; followed by a 2×2 max pooling layer with a stride of 2; The third and fourth layers are

both set with 128 3×3 convolution kernels, then add a maximum pooling layer, then insert the ECA module, then use global average pooling, and finally add a full connection layer and use the SoftMax function to implement individual classification.

4 Simulations and Analysis

4.1 Parameters

Before accessing the fully connected layer, the two feature vectors are spliced, and finally the fully connected layer is used for classification. In the experiment of this paper, the classical cross-entropy loss function is used as the classification loss cost function, the Adam optimization method is used, the initial learning rate is set to 0.0001, the maximum training round of the network is set to 100, and the method of Early Stopping is used.

In this paper, five kinds of communication radiation sources are considered in the experiment. According to reference [4], the third-order Taylor model is used for modeling. The nonlinear parameters of the five kinds of radiation sources are shown in Table 1:

Table 1. Nonlinear parameter settings of five communication radiation sources

Radiation source	Emitter1	Emitter2	Emitter3	Emitter4	Emitter5
λ_1	1	1	1	1	1
λ_2	0.5	0.08	0.01	0.01	0.6
λ_3	0.3	0.6	0.01	0.04	0.08

The transmitter adopts Quad-Phase Shift Keyed (QPSK) modulation mode, the symbol rate is 125kBaud, the sampling rate is 1000 MHz, and the carrier frequency is 420 KHz. The sampled signal samples are 100 symbol periods, and the randomly generated baseband symbol sequence is pulse-shaped through a root raised cosine filter with a roll-off factor of 0.35 and an oversampling ratio of 8. The simulation only considers the Gaussian white noise channel in the single-hop scenario, the signal-to-noise ratio range is 0 dB to 24 dB, the interval is 2 dB, and all communication radiation sources are different except for the nonlinear parameters, other parameters are kept the same. Each radiation source randomly generates 1400 samples under different signal-to-noise ratios, 1000 samples are used as training set, 200 samples are used as validation set, and 200 samples are used as test set. To avoid correlation between samples, each radiation source baseband symbol sequence is a randomly generated sequence.

Perform HHT time-frequency analysis on the received signal to obtain a time-spectrogram, quantize the HHT time-spectrogram of each sample signal according to the number of quantization bits 8, map the amplitude of the time-frequency spectrum to the pixel value of the grayscale image, and map The

range is [0,255]. Assuming that $H(\omega, t)$ represents the time spectrum, and $H_{i,j}$ represents the amplitude value of the element in the i-th row and the j-th column, the quantization calculation formula of the HHT time spectrum is expressed as:

$$G_{i,j} = \left\lfloor (2^8 - 1) \bullet \frac{H_{i,j}}{\max\{H_{i,j}\}} \right\rfloor \tag{10}$$

where $\lfloor \bullet \rfloor$ denotes rounded down, $G_{i,j}$ denotes the value of the (i,j)th pixel of the corresponding grayscale image, the size of the grayscale image is set to 224×224, then $G_{i,j} \in [0, 255]$, $i, j \in [1, 224]$ .For branch network 2, light-weight fusion and splicing of the original time series signal and the SIB feature vector is performed.

4.2 Identification Results

In order to verify the effectiveness of the feature fusion algorithm proposed in this paper under low signal-to-noise ratio. This paper uses the extracted fingerprint features and different deep neural networks to do simulation experiments for comparison.

Four different deep neural network algorithms are used to identify fingerprint features. The selection of fingerprint features is based on the Hilbert time spectrogram or the grayscale image of the original time series signal and the one-dimensional feature vector obtained by splicing with the extracted SIB feature vector.

Experiment 1: Individual identification of radiation sources based on the Hilbert time spectrum, using the grayscale image of the Hilbert time spectrum as the network input, and using ECA-Net for classification and identification [6].

Experiment 2: Individual identification of radiation sources based on one-dimensional feature sequences. The one-dimensional feature vector obtained by splicing the original time series signal and the extracted SIB feature vector is used as the network input, and the one-dimensional residual network is used for classification and identification.

Experiment 3: Using the method in the literature, using the EM^2 algorithm, the energy entropy is extracted from the Hilbert time spectrum, the first-order moment and the second-order moment constitute the feature vector, and the Support Vector Machine(SVM) classifier is used for individual classification and identification.

The experimental results are shown in Fig. 5. Under the same training and test sets, with different the signal-to-noise ratio(SNR), the recognition rate of the individual radiation source method based on feature fusion proposed in this paper is better than the other four comparisons method.Figure 6 shows the confusion matrix for the identification of individual radiation sources when the SNR is 16 dB. It can be seen that the recognition accuracy of the five types of radiation sources is above 95% at 16 dB, while the correct recognition rate of radiation source 3 is as high as 99%. Obviously radiation source 1 and radiation source 5 are easily confused, and radiation source 2 and radiation source 4 are easily confused. On the whole, under low SNR, the algorithm based on feature fusion

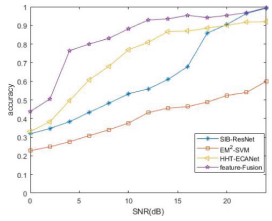


Fig. 5. Comparison of radiation source identification by different algorithms

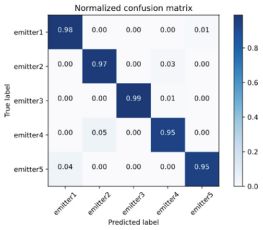


Fig. 6. Confusion matrix for radiator identification at 16 dB

can better identification of confusing radiation source signals, more stable and more accurate.

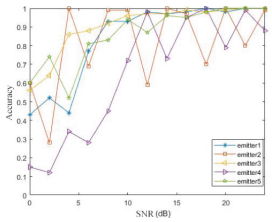


Fig. 7. Recognition of different radiation sources by feature fusion

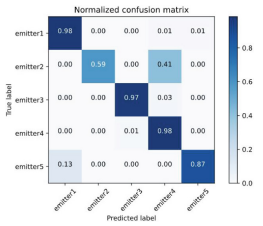


Fig. 8. Confusion matrix for radiator identification at 12 dB

Figure 7 shows the recognition effect of the feature fusion method on different radiation sources. We can find that with the increase of the signal-to-noise ratio, the recognition effect of some radiation sources will fluctuate up and down, which is not completely linear, especially radiation sources 2 and 4. From Fig. 8, at 12 dB, 41% of radiation sources 2 are identified as 4. It can be seen that radiation sources 2 and 4 are easily confused. The reason for this phenomenon has a certain relationship with the Taylor model parameter settings. Further research will be conducted on this highly confusing type of radiation source.

5 Conclusion

Aiming at improving the accuracy of identification of radiation source in complex environment, a method based on feature fusion was proposed in this paper. The extracted signal HHT time spectrum feature and rectangular bispectrum feature are used as the input of two neural networks respectively, then the high-dimensional features extracted by the network are fused, and finally the

full connected neural network is used for feature recognition. Experiments have proved the effectiveness of the proposed method in the identification of individual communication radiation sources, two other deep learning methods and the traditional EM^2-SVM method for comparison. Experiment results show that the individual identification method of radiation source based on feature fusion is more stable than the single branch network identification, the correct identification rate is also higher under low signal-to-noise ratio, and it is more suitable for the complex and unpredictable electromagnetic environment.

References

1. Wang, W., Sun, Z., Piao, S., Zhu, B., Ren, K.: Wireless physical-layer identification: modeling and validation. IEEE Trans. Inf. Forensics Secur. **11**(9), 2091–2106 (2016)
2. Zhang, J., Wang, F., Dobre, O.A., Zhong, Z.: Specific emitter identification via Hilbert-Huang transform in single-hop and relaying scenarios. IEEE Trans. Inf. Forensics Secur. **11**(6), 1192–1205 (2016)
3. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
4. Ding, L., Wang, S., Wang, F., Zhang, W.: Specific emitter identification via convolutional neural networks. IEEE Commun. Lett. **22**(12), 2591–2594 (2018)
5. Huang, N.E., et al.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceed. Royal Soc. London. Ser. A: Math. Phys. Eng. Sci. **454**(1971), 903–995 (1998)
6. Pan, Y., Yang, S., Peng, H., Li, T., Wang, W.: Specific emitter identification based on deep residual networks. IEEE Access **7**, 425–434 (2019)

Design and Trial Manufacture of Methanol Reforming Hydrogen Fuel Cell Power Supply

Youjie Zhou, Jinmao Chen, Chunhua Xiong, Yanli Sun, Lei Xu(✉), Long Huang, and Guang Hu

Institute of System Engineering, Academy of Military Science, Beijing 100300, China
energyxl2022@163.com

Abstract. The integrated technology of methanol-reforming hydrogen production and fuel cell power supply system is implemented in this paper based on the combination of theoretical analysis with critical technologies and systematic design research. The methanol reforming hydrogen fuel cell power supply is trial-manufactured and tested. The test results show that the rated output power of the prototype can reach 5kW, and the overload power is 110% of the rated power. In addition, the regular operation can be started continuously in a short time, and the long-term continuous and reliable operation of 100h can be realized. The research results provide a trustworthy theoretical basis and technical support for the development of reforming fuel cell power supply.

Keywords: Methanol reforming · Hydrogen generator · Fuel cell power supply

1 Introduction

Mobile power supplies are small-scale power generation devices often used as the self-contained power supply of remote areas and some factories. On the other hand, it is also used as an alternative power supply for enterprises to make up for the shortage of grid power supply or to deal with the power cut-off. The traditional mobile power supply is generally composed of diesel generator sets with relatively mature and widely used technics. However, the common problems still cannot be ignored, such as low energy efficiency, high noise, and high pollution [1, 2].

Compared with the traditional diesel generator sets, the fuel cell has gradually become a research hotspot due to its advantages in energy conversion rate, working noise, and environmental friendliness. With the depletion of petroleum energy and especially the increasing demand for energy in today's society, fuel cell power supply has a considerable development space [2].

Based on the technology of methanol reforming hydrogen production, a small-scale fuel cell power supply with an output voltage of AC220V and a rated output power of 5kW is trial produced. It is integrated with a methanol reforming hydrogen generator module, fuel cell power generation module, energy storage module, inverter module of DC/DC and DC/AC. The dynamic response characteristics, overload capacity, reliability, and continuous startup are tested.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 324–331, 2023.
https://doi.org/10.1007/978-981-99-1260-5_41

2 Scheme Design and Prototype Trial Production

2.1 Structure Design of Fuel Cell Power Supply

The methanol reforming hydrogen fuel cell power supply is composed of two parts: the methanol reforming hydrogen production module and the fuel cell power generation module. The methanol reforming hydrogen production module relies on methanol as fuel to produce hydrogen for the fuel cell system. The fuel cell power generation module is mainly composed of a buffer tank, fuel cell stack, DC / DC module, energy storage battery, DC / AC inverter, and control execution module. The overall design block diagram of the system is shown in Fig. 1.

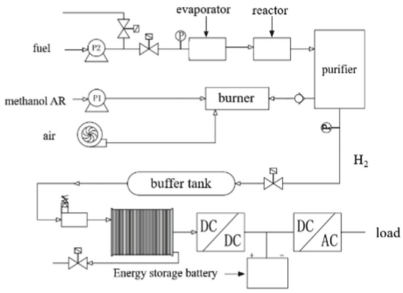


Fig. 1. Overall design diagram of fuel cell power supply based on methanol reforming hydrogen technology

The sub-modules are described as follows:

(1) Fuel cell stack.

The fuel cell stack is the core component of the fuel cell system, which converts the hydrogen generated by the methanol reforming hydrogen generator into electric energy. The voltage generated by the fuel cell stack varies widely with different loads. Since the rated output power of methanol reforming fuel cell power supply is 5kW, the rated output power of the fuel cell stack selected in this paper is 6kW. The specific parameters are given in the following table:

Table 1. The relationship between hydrogen production and fuel consumption

No	Parameters	Values
1	Rated Power	6 kW
2	Rated Voltage	69 V
3	Rated Current	86.9 A
4	Hydrogen Pressure	0.05 ~ 0.07 MPa

(2) Methanol reforming hydrogen generator.

The methanol reforming hydrogen generator is used as the hydrogen supply device for the fuel cell power supply. According to the rated power generation of the methanol fuel cell stack of 6kW, the flow of hydrogen produced by the hydrogen generator to meet the regular operation of the fuel cell power supply is calculated to be about $6m^3$ / h.

The specific calculation steps are as follows [3]:

Number of hydrogen ions produced per hour

$$N_{H^+} = \frac{P \cdot t}{V \cdot e^-} = \frac{6 \times 10^3 \times 3600}{0.85 \times 1.6 \times 10^{-19}} = 1.59 \times 10^{26}$$

The cathode undergoes an oxygen reduction reaction, and the anode undergoes a hydrogen oxidation reaction. The theoretical voltage can reach 1.23–0 = 1.23V. However, the actual output voltage is usually only 0.7 ~ 0.9V in practical applications due to overpotential, internal resistance, and mass transfer. Thus, 0.85V is taken here.

Hydrogen consumption

$$rate(n_{H_2}) = \frac{N_{H^+}}{2N_A} = \frac{1.59 \times 10^{26}}{2 \times 6.02 \times 10^{23}} = 131.9\,\text{mol/h}$$

$$rate(V_{H_2}) = \frac{rate(n_{H_2})}{V_m} = \frac{131.9}{22.4} = 5.89\,\text{m}^3/\text{h}$$

In the previous research, the research group designed and trial produced a methanol reforming hydrogen generator [4] with a hydrogen production capacity of 6m3/h according to the hydrogen consumption demand of fuel cell power supply. It is mainly composed of a regenerator, burner, evaporator, reactor, purifier, insulation layer, and control execution module. The general block diagram and basic structure diagram of the scheme are as follows:

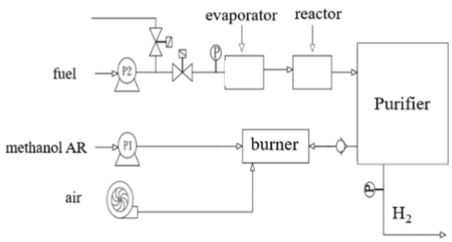


Fig. 2. General block diagram of methanol reforming system scheme

(3) Buffer tank.

The buffer tank is used to store and buffer reformed gas in order to reduce the disturbance and fluctuation of hydrogen pressure during stack pulse exhaust. The volume of the buffer tank is designed as 8L,

(4) DC/DC module.

The DC/DC module is used to convert the DC voltage of the fuel cell stack that changes in a wide range into a DC 48V voltage, which matches the voltage of the energy storage cell. The control board controls it in real-time to output the set power.

(5) Energy storage battery.

A 48V lithium battery is selected as the energy storage battery, which is used to provide the power supply when the system is started. It can provide parts of electric energy and form an electric hybrid power supply system with the stack and DC/DC.

(6) DC/AC Inverter.

The inverter converts 48V DC voltage into AC220V 50Hz AC in order to expand the applicable power supply range.

(7) Fuel cell control execution module.

The fuel cell control execution module includes a pressure reducer, a solenoid valve, a temperature sensor, a cooling fan, a control board, etc. The pressure reducer stabilizes the high hydrogen pressure generated by the methanol reforming hydrogen production system to the pressure range required by the stack. Under the control of the control board, each sensor actuator of the stack ensures that the fuel cell stack works within the set temperature range and load range.

At the same time, the control board controls the DC/DC to work within the appropriate load range. The stacks and lithium batteries are controlled to supply power in appropriate proportion. In addition, the control board can adjust the current to charge the lithium batteries when necessary to realize the energy management function.

2.2 Main Control Design Ideas

The control system of the methanol reforming hydrogen fuel cell power supply is able to identify multiple states such as standby, start, operation, stop and fault states. The conversion relationship of each state is shown in Fig. 3.

After the power is switched on, the system enters the standby state. Then, it transfers to the startup state while receiving the relevant instruction and prepares for regular operation. Meanwhile, the system judges whether all parameters meet the requirements of the startup. If startup timeout, the system will be determined as abnormal automatically and switch to the fault state; Otherwise, it will switch to the operation state.

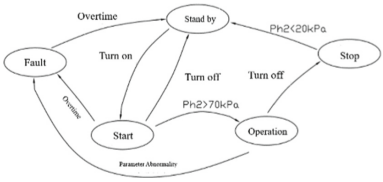


Fig. 3. State transition diagram of fuel cell power supply system

In the operation state, the system monitors all parameters at all times. If there is any abnormality, it will turn to the fault state. The shutdown command will turn to the stop state and perform the shutdown operation if the shutdown command is received.

In the stop state, the system judges all parameters. It enters the standby state if the conditions for entering the standby state are met.

After the system switches to the fault state due to abnormal parameters or timeout, the system will wait for a specific time. During this period, the system will not respond to the command. After the timeout, the system will automatically enter the standby state and can try to start again.

2.3 Prototype Trial Production

The following figure is a physical diagram of the 5kW fuel cell power supply prototype, including the methanol reforming hydrogen generator, fuel cell power generation module, energy storage module, DC/DC and DC/AC inverter module mentioned above. These modules are reasonably arranged and distributed in the customized cabin. The external dimension of the prototype is 900 × 1300 × 1300mm; the total weight is 640kg; the inverter output is 220V AC voltage.

Fig. 4. Methanol reforming to hydrogen fuel cell power supply prototype

3 Test and Result Analysis

3.1 Fuel Cell Dynamics

As shown in Fig. 5, the dynamic response performance of the fuel cell is increased by 3KW at position 1. At this time, the output power of the stack starts to increase at a specific rate after a small step. With the increase of the output power of the stack, the output power of the lithium battery decreases, and the current of the lithium battery decreases. The current of the lithium battery gradually decreases until it turns into the charging state with the continuous increase of the stack current. When the charging current limit is reached, the stack current stops increasing. At the time of position 2, the load decreases by 3KW, the stack current decreases at the same time, and the battery maintain the charging state to supplement the electric energy consumed by the battery after loading. By changing the load vastly, the system can respond to the load change in time to ensure the reliability and stability of the power supply.

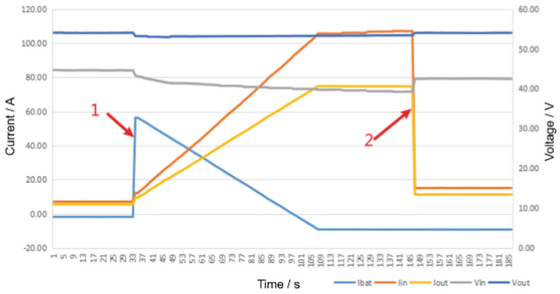


Fig. 5. Dynamic response performance of fuel cell

3.2 Reliability Test

The reliability test results of the fuel cell stack are shown in Fig. 6. Where the abscissa is the time, and the total operation time is 100 h. Discharge with a rated current of 80A and constant current. In the beginning, the stack voltage was 68.9V. At the end of the test, the stack voltage decreases to 67.2V. The voltage drop rate equals $1.7/68.9 \times 100\% = 2.47\%$。

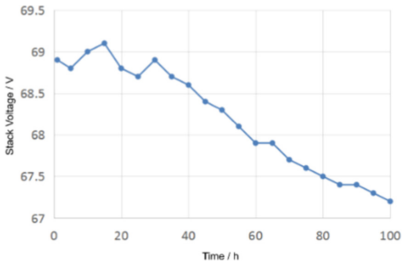


Fig. 6. Reliability test

3.3 Overload Capacity

Overload capacity test results are shown in Fig. 7. In the figure, time 1 corresponds to the completion of system startup and the gradual loading. It is loaded to full load at time 2 and operated for a while until time 3. It is loaded to 110% of the rated load, i.e., 5500W. The system is regular without any abnormality, and it operates for 10min.

3.4 Continuous Start Test

Figure 8 indicates the results of the continuous startup test, which shows that the system starts from the start of the cold engine to the start of the system at time 1. It lasts for 48 min. After startup, the system is loaded and operated. Then, it is loaded to full load gradually (time 2). The load is removed and the system is started again at time 3 after

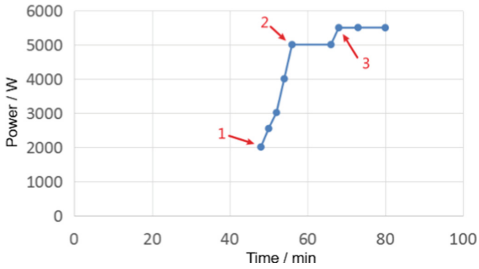


Fig. 7. Overload Capacity

the shutdown. Because it is in the hot engine state, the startup speed is fast. After the start, the system is loaded to full load gradually and operated to time 4. After shutdown, the system is started for the third time (time 5), The startup is normal, and the operation is regular. It can be started and operated normally three consecutive times.

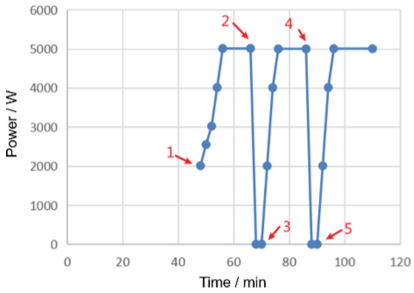


Fig. 8. Continuous Start Test

4 Conclusions

Based on the methanol reforming hydrogen production technology, a methanol reforming hydrogen generator module and a fuel cell power generation module are systematically designed and combined to produce a methanol reforming hydrogen fuel cell power supply. The test proves that the prototype installation is reasonable, and the performance is stable and reliable. The test results are summarized as follows:

(1) The methanol reforming fuel cell power supply can respond to the load change in time to ensure stable and reliable power output. The lithium battery and fuel cell in the system can adjust the power output according to the load change;

(2) The voltage reduction rate of the fuel cell stack shall not exceed 2.5% within 100h of continuous operation;

(3) The methanol reforming fuel cell power supply prototype can normally operate at 110% of the rated power for 10min without abnormality;

(4) The fuel cell power supply system for methanol reforming to hydrogen can be continuously started and normally operated in a short time.

References

1. Xuejun, M., Bo, W.: Diesel Generator Set and Selection. Cereal & Food Industry. (01), 47–49, 57 (2006)
2. Changsong, Q.: Fuel Cell Applications in the Communication Power Supply. Telecom Power Technol. **29**(01), 53–57 (2012)
3. Ronghui, X., Wei, Y.: Design of Proton Exchange Membrane Fuel Cell Station Systems. East China Electr. Power. **14**, 48–49 (2008)
4. Changbo, Lu., Lei, Xu., Yan, Qin., Guang, Hu., Weigui, Zhou., Youjie, Zhou.: Design and Trial Manufacture of Methanol Reforming Hydrogen Generator. Springer. Singapore, p. 491–497 (2022)

Multi Station Direct Location Algorithm of Narrowband Coherent Source Based on Spatial Smoothing

Heng Sun(✉), XinGuang Zhang, TianBo Liu, Jun Lu, Lei Guo, and Shuo Wang

Beijing Aerospace Long March Vehicle Research Institute, No. 1. Nan Da Hong Men Road, 100076 Beijing, China
sunheng_2012@163.com

Abstract. Aiming at the problem that the direct location algorithm has reduced or even failed to locate coherent source directly, coherent signals dierct location algorithm based on spatial smoothing is proposed. Which adopts the idea of spatial smoothing to realize the full rank of the received source covariance matrix and adopts "Z" word spatial smoothing extends the application range of this idea from linear array to area array, and realizes the direct location of multi station coherent source by using the technology of multi station noise subspace fusion and normalization. The simulation results show that the algorithm can achieve better resolution and direct location for narrow-band coherent source under the condition of long source distance and low signal-to-noise ratio, and under the configuration of linear array and area array, it has good positioning effect.

Keywords: Narrowband target · coherent signals · level direct positioning · multi stationcooperative center

1 Introduction

At present, passive location technology is mainly divided into two-step location method and direct location method. The two-step positioning method converts the received sampling signal into pulse description word, and uses the information of the time difference, frequency difference and direction finding to slove the target position. In the process of the transformation from sampling information to pulse description word, there is inevitably the loss of effective information. In order to make full use of effective information to improve positioning accuary, direct positioning method has been favored by the majority of researchers.

Reference 1 and 2 proposed a direct location algorithm, which directly uses the sampled data to estimate the position of the observation station, avoiding the information loss and error accumulation caused by parameter estimation. According to different solutions, direct location algorithms are divided into three categories: intelligent optimization, maximum likeihood and subspace [3]−[4].Among them, subspace class has low computational complexity and high resolution performance, and has strong application prospects [5]. However, when there is a coherent source, the signal array will

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 332–338, 2023.
https://doi.org/10.1007/978-981-99-1260-5_42

receive coherent signals in different directions, and the coherent signal will lead to the rank deficiency of the source covariance matrix, so that the signal eigenvector diverges to the noise subspace, resulting in the decline of resolution algorithm and matrix reconstruction algorithm. Typical decoherence methods are spatial smoothing algorithm and matrix reconstruction algorithm, but matrix reconstruction algorithm requires more array elements [6], and spatial smoothing technology has strong decoherence ability under uniform linear array and uniformarea array [7]. In this paper, the full rank of the received source covariance matrix is realized by using the idea of spatial smoothing, and the multi station cooperative coherent source direct location is realized by using the multi station noise subspace fusion and normalization processing technology.

1.1 Mathematica Model of Coherent Source Direct Location

Assuming that K independent narrowband far-field signals are incident on the array element of the array antenna from different directions, the received data vector $x(t)$ can be expressed as:

$$x(t) = \sum_{i=1}^{K} \alpha(\theta_i) s_i(t) + n(t) \tag{1}$$

where $\alpha(\theta_i)$ is the guidance vector of the array for the signal in the direction, $s_i(t)$ is the signal incident from the direction, and $n(t)$ is the random noise of the array.

$$a(\theta) = \left[\exp\left(i \frac{2\pi}{\lambda} d((n-1)\sin(\theta_i) - (m-1)\cos(\theta_i)) \right) \right] \tag{2}$$

where λ is the signal wavelength, and m and n are the row and column of the array element respectively.

Suppose the coordinates of the observation station are $sta(x_i\, y_i)$, $(i = 1 \cdots M)$ and the target coordinates are $Tar(x_{ej}\, y_{ej})$, $(j = 1 \cdots k)$, which can be obtained from Eq. 1. 2:

$$a(x_{ej}\, y_{ej}) = \exp\left(\frac{2\pi i}{\lambda} d\left((n-1) \cdot (y_i - ye_j) / |Sta - Tar| - (m-1) \cdot (x_i - xe_i) / |Sta - Tar| \right) \right) \tag{3}$$

Equation 1.3 shows the direct correlation between the received signal and the target position parameters, and the direct estimation of the source position is completed through the fusion processing of the received signal.

1.2 Multi Station Direct Algorithm Base on Spatial Spectrum Estimation

1.2.1 Target Location Estimation Model

The covariance matrix of array data is

$$R = E\left[XX^H \right] = AE\left[SS^H \right] A^H + \sigma^2 I = AR_s A^H + \sigma^2 I \tag{4}$$

Signal and noise are independent of each other. The data covariance matrix is decomposed into two parts related to signal and noise, where R_s is the covariance matrix of siganal and AR_sA^H is the signal part.

Feature decomposition of R:

$$R = U_S \sum sU_S^H + U_N \sum_N U_N^H \tag{5}$$

U_S is the signal subspace formed by the eigenvector corresponding to the large eigenvalue, and U_N is the noise subspace formed by the eigenvector corresponding to the small eigenvalue. Under ideal conditions, the signal subspace and noise subspace in the data space are orthogonal to each other, and the guidance vector in the signal subspace is also orthogonal to the noise subspace. Therefore, spectral estimation is achieved by minimizing the search:

$$P = \frac{1}{A^H(x\ y)\hat{U}_N\hat{U}_N^H A(x\ y)} \tag{6}$$

1.2.2 Normalization Processing

Normalize the spatial spectrum of the search area obtained by each station, record the maximum value P max of the spatial spectrum of the search area, and normalize the regional spatial spectrum value of each station according to the following formula.

$$Pm_k = 10 \cdot \lg\left(Pm_k \big/ P\max\right),\ k = 1 \cdots M \tag{7}$$

Sum the normalized search spatial spectrum of each station to obtain the spatial spectrum of the search area:

$$Psum = Pm_1 + Pm_2 + \cdots Pm_M \tag{8}$$

The target position can be obtained directly by calculating the value of the total spatial spectrum made and obtained by each station.

1.3 Uniform Area Array Division Decoherent Direct Location

A typical uniform array is shown in Fig. 1., 16 array elements are arranged into 4×4 array, the phase array element spacing of each row and column is d.Taking four element in each subarray as an example, the covariance matrices of the two subarrays are R_1 and R_2 respectively, and the spatially smooth covariance matrix is $\overline{R} = (R_1 + R_2)/2$. By replacing R in Eq. 5 with $\overline{R}$ and suing spatoal spectrum estimation and normalization technology, the direct location of coherent sources can be realized.

The number of elements N_i in the subarray($N_i = l^2$,l is the number of elements in one row of the subarray)shall be greater than or equal to the target number K. In Order to reduce the sacrifice of effective array aperture, the limiting conditions for the selection of array elements in one row of subarray are: $(l-1)^2 < K \leq l^2$,N is the total number of elements. Number of subarrays at this time is $P = \left(\sqrt{N} - l + 1\right)^2$.

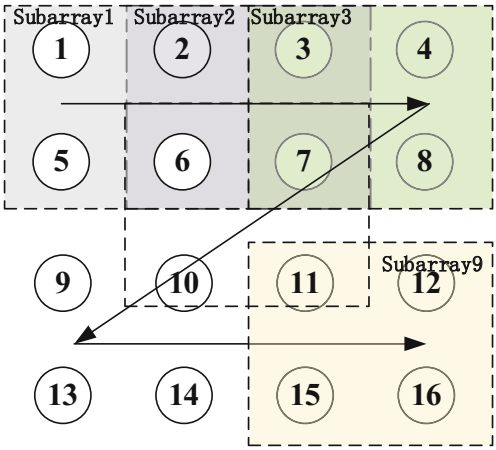


Fig. 1. Schematic diagram of uniform array segmentation

The positioning process of the algorithm is shown in the figure below:

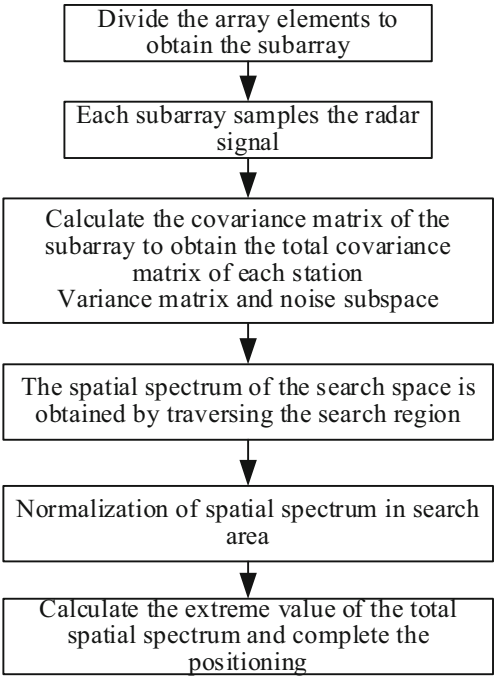


Fig. 2. Multi station direct location algorithm flow based on spatial smoothing

1.4 Simulation Experiment and Analysis

The simulation test mainly analyzes the performance of the direct location algorithm based on spatial smoothing from the aspects of phase source adaptability and the number of subarray elements.

The location setting of observation station is shown in Table 1

Table 1.1 Location of Station

Observation Station Number	***1***	***2***	***3***	***4***
X/km	600	200	0	100
Y/km	0	200	600	1000

The source location setting is shown in Table 2:

Table 2 Source Location

Source Number	***1***	***2***	***3***
X/km	10	7	30
Y/km	30	5	42

1.4.1 Adaptability to Coherent Sources

Three narrowband far-filed signals are respectively set to incident on the array, the signal center frequency is 2.7 GHz, the received signal -to-noise ratio of each channel is 20 dB, the number of antenna elements of each observation station is 16, and the array element spacing is $0.5/\lambda$. The subarray is divided into four array elements as shown in Fig. 1. Figure 3 shows the positioning results of classical direct positioning algorithm and multi station direct positioning algorithm based on spatial smoothing.

It can be seen from the simulation results that when there are coherent sources, the three source location regions in the spatial spectrum based on the classical direct location algorithm are fuzzy, the algorithm fails, and the target location can not be solved. The spatial spectrum of the direct location algorithm based on spatial smoothing has clear location regions of the three sources, has good decoherence ability, and can accurately reflect the real location of the source.

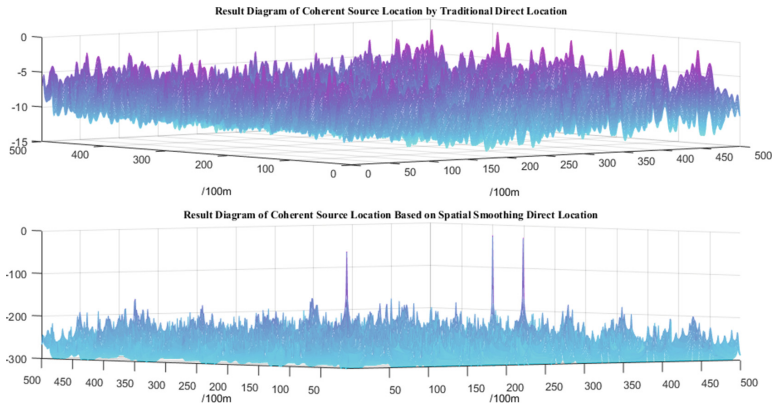


Fig. 3. The positioning results of classical direct positioning algorithm and multi station direct positioning algorithm based on spatial smoothing

1.4.2 Loction Performance Analysis of Different Segmentation Methods of Spatial Smoothings

Three narrowband far-filed signals are respectively set to incident on the array, the signal center frequency is 2.7 GHz, the received signal -to-noise ratio of each channel is 20 dB.The array antenna adopts the layout of 3×3 uniform area array, and the array element spacing is $0.5/\lambda$.The space of the array antenna is smoothly divided according to the division mode of 6-element 2-subarray and 4-element 4-subarray. Figure 4 shows the multi station positioning results under two segmentation modes.

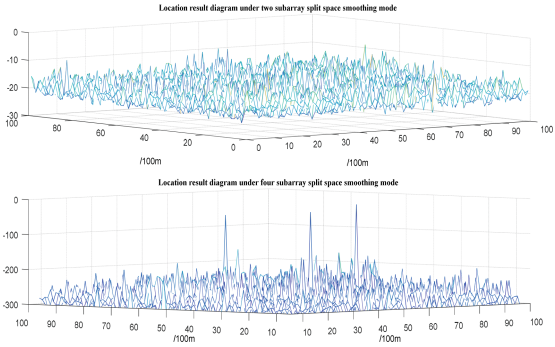


Fig. 4. Loction results of different segmentation methods of spatial smoothings

From the simulation results, the two subarray segmentation method can not distinguish the positioning region of three coherent sources, and does not have the ability to decoherent signals. The four subarray spatial segmentation method has clear positioning results for three coherent sources, has good decoherence ability, and can accurately

reflect the real position of the source. Therefore, the positioning of coherence sources, the number of subarrays must be greater than the number of coherent sources to accurately locate.

1.5 Simulation Experiment and Analysis

Based on the idea of DOA coherent source measurementally by uniform linear array forward smoothing, this paper deduces a multi base station direct location technology that can be applied to uniform area arry. This method averages the covariance matrix of each subarray through area array z-word smoothing wo make up for the rank deficiency of covariance matrix caused by coherent sources. Simulation results show that the algorithm has the ability of phase unwrapping and positioning. The number of decoherent signals corresponding to different subarray space division is different. When the number of subarrays is less than the number of decoherent signals, decoherence cannot be realized.

References

1. Weiss, A.J.: Direct position determination of narrowband radio frequency transmitters. IEEE Signal Process. Lett. **11**(5), 513–516 (2004)
2. Oispuu, M., Nickel, U.: Direct detection and position determination of multiple sources with intermittent emission. Signal Process. **90**(12), 3056–3064 (2010)
3. Guizhou, W., Fucheng, G., Min, z.:overview of signal direct positioning technology J. Radar , 9(6), 998–1013 (2020)
4. Linpeng, G., Changwen, Q., Qi, F., Yanbo, P., Cheng, C.: Research on coherent source direct location method based on decoherent music Ship Electron. Eng., 39(02), 52–55 (2019)
5. Foy, W.H.: Position-location solution by Taylor-series estimation .IEEE Trans. Aerosp. Electron. Syst., 1 AES-12, 187–194 (1976)
6. Schmidt, R.: Multiple emitter location and signal parameter estimation. Int. J. Eng. Res. 2(3), 276–280 (2013)
7. Guohui, Z., Zhike, D., Bin, H., Xiaoyan, L.: Simulation and implementation of subarray spatial smoothing algorithm based on uniform array. Comput. Simul. **34**(10), 252–256 (2017)

Fast Multi-target Sensing and Tracking for Directional Array-Deployed UAV Systems

Zhipeng Lin[1], Qiuming Zhu[1(✉)], Tiejun Lv[2(✉)], Weizhi Zhong[1], Jie Wang[1], Kai Mao[1], and Kun Liu[1]

[1] Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
{linlzp,zhuqiuming,zhongwz,bx2104906wangjie,maokai,lkhard}@nuaa.edu.cn
[2] School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
lvtiejun@bupt.edu.cn

Abstract. Accurate sensing and fast tracking are challenging for large-antenna unmanned aerial vehicles (UAVs), due to computational capacity limits. This paper presents a novel simultaneously multi-target sensing and tracking scheme for millimeter-wave (mmWave) large-antenna UAV systems. By exploiting the low-rank and high-directivity nature of mmWave multi-antenna channels, a new directional beamformer is designed for the UAV. As a result, the horizontal angles of arrival (AOAs) can be roughly estimated and the number of targets can be quickly identified. We also propose a low-complexity multi-target azimuth and elevation AOA estimation algorithm. The asymmetric discrete Fourier transform (A-DFT) is employed to reduce the dimension of the received signals. By using the elevation estimates for azimuth AOA estimation, the mismatching of the estimated parameters for different targets is avoided. Simulations show that the proposed scheme can accomplish the task of high-accuracy multi-target sensing and tracking in a short period of time.

Keywords: UAV · Large-antenna array · Directional beamforming · Simultaneously sensing and tracking

1 Introduction

Accurate target sensing and tracking have already become a necessity with the widespread usage of portable devices and the dramatic development of internet of vehicle networks [1–3]. Thanks to the high mobility and convenient deployment character of unmanned aerial vehicles (UAVs), using UAV technique is promising for achieving tasks of precise sensing and fast tracking [4]. However, most current UAV-based sensing methods do not consider the computational capacity limit of UAVs [5,6]. To accurately estimate channel parameters, such as the angles

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 339–346, 2023.
https://doi.org/10.1007/978-981-99-1260-5_43

of arrival (AOAs) and the power of the signals, these methods deploy high-directivity millimeter-wave (mmWave) large antennas to the UAVs and employ high-computational complexity parameter estimation algorithms at the UAVs. When the number of antenna elements is large and the dimension of parameters to be estimated is high, these method would spend a tremendous amount of time for parameter estimation. As a result, the task of real-time tracking cannot be accomplished in practice.

In this paper, we propose a novel three dimensional (3D) mmWave-based simultaneously multi-target sensing and tracking scheme for large antenna-deployed UAV systems. By reasonably assigning tasks, the scheme fully considers the computational capacity limit of UAVs and the time sensitivity of tracking tasks. By exploiting the low-rank and high-directivity nature of mmWave multi-antenna channels, a new directional beamforming structure is presented for the UAV, which can quickly estimate the rough horizontal AOAs and identify the number of moblie stations (MSs). We also propose a low-complexity multi-target azimuth and elevation AOA estimation algorithm. The asymmetric discrete Fourier transform (A-DFT) is employed to reduce the dimension of received signals. By using the elevation estimates for azimuth AOA estimation, the mismatching of the estimated parameters for different targets is also avoided.

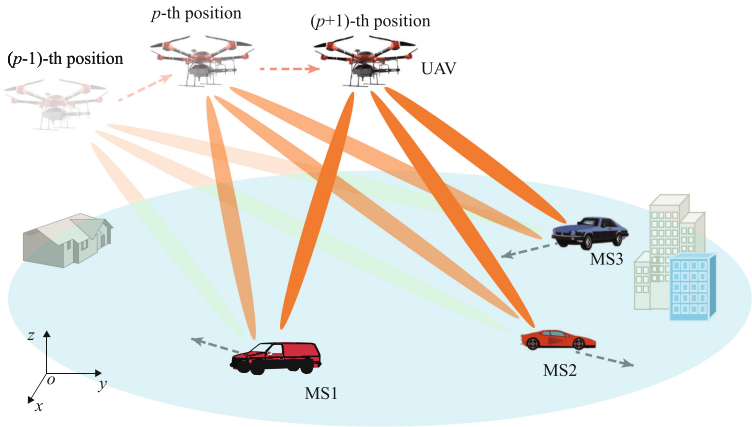


Fig. 1. The considered simultaneously multi-target sensing and tracking scenario.

2 System Model

We consider a 3D mmWave-based simultaneously multi-target sensing and tracking problem, as shown in Fig. 1. We assume that the channel state remains unchanged when the UAV receives M time-slot signals from the MSs at each position, and the UAV flying at constant speed and constant altitude moves Δ_P

distance for every receiving M signals. Consider that a mmWave uniform cylindrical array (UCyA) with N_{R} antennas is used at the UAV. The received signal at the m-th time slot of the p-th ($p = 0, 1, \ldots, P-1$) position is given by [7]

$$\mathbf{r}_{m,p} = \mathbf{H}_p x_{m,p} + \mathbf{n}_{m,p}, \tag{1}$$

where $\mathbf{H}_{m,p} \in \mathbb{C}^{N_{\mathrm{R}} \times 1}$, $\mathbf{n}_{m,p} \in \mathbb{C}^{N_{\mathrm{R}} \times 1}$, and $x_{m,p}$ denote the channel matrix, the Gaussian noise, and the transmitted signal at the m-th time slot of the p-th ($n = 0, 1, \ldots, M-1$) UAV position, respectively. Since the UAV flies in the sky, it is reasonable to assume that the signals can be transmitted from the MSs to the UAV directly without being blocked. The channel matrix, $\mathbf{H}_p$, can be expressed as

$$\mathbf{H}_p = \sum_{l=1}^{L} \beta_l e^{-j2\pi f \tau_l} \mathbf{a}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l}) \widetilde{P}_p, \tag{2}$$

where L is the number of paths; β_l is the complex amplitude of the l-th path; $\mathbf{a}(\phi_{\mathrm{V},p,l}, \phi_{\mathrm{H},p,l})$ is the array response vector with $\phi_{\mathrm{A},p,l}$ and $\phi_{\mathrm{E},p,l}$ being the azimuth and elevation AOAs of the l-th path; τ_l is the time delay of the l-th path and f is the frequency. In this paper, we assume that only the line-of-sight (LOS) signals can be transmitted from the MSs to the UAV, and thus, the number of paths is equal to the number of MSs. In the above equation,

$$\widetilde{P}_p = \exp\left(j\frac{2\pi}{c} f(P_{\mathrm{H},p} + P_{\mathrm{V},p})\right), \tag{3}$$

where $P_{\mathrm{V},p}$ and $P_{\mathrm{H},p}$ are the p-th distance of UAV from the origin decoupled between the vertical and horizontal planes, which can be expressed as

$$P_{\mathrm{H},p} = |\mathbf{p}_0| \cos\theta_0 + \Delta_P \cos\theta_1 + \cdots + \Delta_P \cos\theta_{p-1};\ P_{\mathrm{V},p} = H_0 \tag{4}$$

where H_0 is the flight altitude of UAV, $\mathbf{p}_0$ is the original position of UAV, and θ_p is the p-th flying direction of the UAV. To simplify, we can let $\mathbf{p}_0$ be the origin of coordinate, and thus, $\mathbf{P}_0 = \mathbf{0}$ and $\theta_0 = 0$.

The mmWave UCyA antenna array consists of N_{V} horizontal layers of UCAs, each having N_{H} antennas, so $N_{\mathrm{R}} = N_{\mathrm{V}} N_{\mathrm{H}}$. The receive array steering vector is constructed as

$$\mathbf{a}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l}) = \mathbf{a}_{\mathrm{H}}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l}) \otimes \mathbf{a}_{\mathrm{V}}(\phi_{\mathrm{E},p,l}), \tag{5}$$

where $\otimes$ denotes the Kronecker product,

$$[\mathbf{a}_{\mathrm{H}}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l})]_{n_{\mathrm{H}},1} = \frac{1}{\sqrt{N_{\mathrm{H}}}} \exp\left[j\frac{2\pi}{c} fr \sin(\phi_{\mathrm{E},p,l}) \cos(\phi_{\mathrm{A},p,l} - \varphi_{n_{\mathrm{H}}})\right]$$

$$[\mathbf{a}_{\mathrm{V}}(\phi_{\mathrm{E},p,l})]_{n_{\mathrm{V}},1} = \frac{1}{\sqrt{N_{\mathrm{V}}}} \exp\left(j\frac{2\pi}{c} fh(n_{\mathrm{V}} - 1) \cos(\phi_{\mathrm{E},p,l})\right)$$

are the array response vectors on the horizontal and vertical planes, respectively, with $n_{\mathrm{V}} = 1, 2, \ldots, N_{\mathrm{V}}$ and $n_{\mathrm{H}} = 1, 2, \ldots, N_{\mathrm{H}}$. r is the radius of each UCA; h is the vertical distance between any two adjacent UCAs; c is the speed of light; and $\varphi_{n_{\mathrm{H}}} = 2\pi(n_{\mathrm{H}} - 1)/N_{\mathrm{H}}$ is the difference between the central angles of the n_{H}-th antenna and the first antenna of each UCA.

3 Proposed Multi-target Sensing and Tracking Scheme

An illustration of the proposed multiple target positioning and tracking scheme is shown in Fig. 2, which contains six tasks corresponding to six stages. Specifically, at Stage 1, the UAV receives the signals of MSs from the directional beamformer. The structure of the directional beamformer will be presented in Sect. 4, and the received signals at this stage are used to plan the direction of UAV for multiple target tracking and to estimate the number of MSs. The signals received at Stage 2 are used to precisely estimate the AOAs of MS. Due to the computing capacity limit of UAV, only the low-complexity preliminary parameter estimation (Task 4) is processed by the UAV. After receiving the M-th time-slot signals, the UAV sends the received $(M-1)$ signals at Stage 2 to the BS and starts to fly in the planned direction for tracking. Accurate AOAs are estimated at BS (Task 6) by using high performance computers to reduce the time. Note that in this scheme, we assume that time used for preliminary and accurate parameter estimation (i.e., Tasks 4 and 6) is shorter than the time used to finish Tasks 2 and 3, respectively.

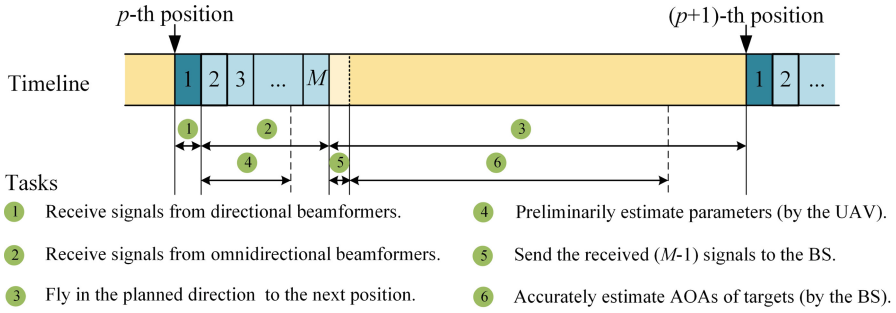


Fig. 2. The timeline and task specification of the proposed scheme.

4 Fast Preliminary Parameter Estimation

We first design a new directional beamforming, which can quickly accomplish the task of preliminary parameter estimation by the UAV. As discussed in Sect. 3, the received signals from the directional beamformer at Stage 1 are used for preliminary parameter estimation. The beamforming matrix is expressed as $\mathbf{W}_{\mathrm{D}} = \mathbf{W}_{\mathrm{H},D} \otimes \mathbf{W}_{\mathrm{V,D}}$, where $\mathbf{W}_{\mathrm{V},D}$ and $\mathbf{W}_{\mathrm{H},D}$ are decoupled beamforming matrices between the vertical and horizontal planes.

By applying $\mathbf{W}_{\mathrm{D}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ to the received signal $\mathbf{r}_m$, the output signal after directional beamforming (at Stage 1) can be expressed as

$$\begin{aligned}\mathbf{y}_{m,p,\mathrm{D}} &= \sum_{l=1}^{L} \tilde{X}_{m,p,l} \left(\mathbf{W}_{\mathrm{H},D}\mathbf{a}_{\mathrm{H}}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l})\right) \otimes \left(\mathbf{W}_{\mathrm{V,D}}\mathbf{a}_{\mathrm{V}}(\phi_{\mathrm{E},p,l})\right) + \mathbf{W}_{\mathrm{D}}\mathbf{n}_{m,p} \\ &= \sum_{l=1}^{L} \tilde{X}_{m,p,l} \left(\tilde{\mathbf{a}}_{\mathrm{H},D}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l}) \otimes \tilde{\mathbf{a}}_{\mathrm{V,D}}(\phi_{\mathrm{E},p,l})\right) + \mathbf{W}_{\mathrm{D}}\mathbf{n}_{m,p} \end{aligned} \tag{6}$$

where $\tilde{X}_{m,p,l} = \beta_l e^{-j2\pi f \tau_l} \widetilde{P}_p x_{m,p}$.

We evenly divide the horizontal angle space into γ parts, and make the beamformers of UCyA point in such γ directions. Therefore, $\mathbf{W}_{\mathrm{H}}$ can be expressed as

$$\mathbf{W}_{\mathrm{H},D} = \begin{bmatrix} \mathbf{B}_{K,1} & & \mathbf{0}_{N_{\mathrm{H}}-K,1} & & \\ \mathbf{0}_{K,1} & \mathbf{B}_{K,1} & & \mathbf{0}_{N_{\mathrm{H}}-2K} & \\ \vdots & \vdots & \ddots & \vdots & \\ & & \mathbf{0}_{N_{\mathrm{H}}-K,1} & & \mathbf{B}_{K,1} \end{bmatrix} \in \mathbb{C}^{\gamma \times N_{\mathrm{H}}}, \tag{7}$$

where $K = N_{\mathrm{H}}/\gamma$. In the above equation, $\mathbf{B}_{K,1} = [b_1, b_2, \ldots, b_K]$ is the beamforming weight coefficient matrix used to make the superimposed beams of K successive horizontal antenna elements of $\mathbf{a}_{\mathrm{H}}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l})$ point in a certain direction.

Since the horizontal angle space is evenly divided into γ parts, by detecting whether some of beams receive signals from the MSs, we can roughly estimate the horizontal AOAs of MSs. Assume $\Theta_{\mathrm{BH},p}$ is the set of the indexes of the beams receiving signals. The p-th flying direction of the UAV, θ_p, can be calculated by $\theta_p = \sum_{i \in \Theta_{\mathrm{BH},p}} \varphi_i$, where φ_i is the central direction of the i-th beam.

We proceed to design the vertical-plane beamforming $\mathbf{W}_{\mathrm{V,D}}$ to estimate the number of MSs by exploiting the sparsity of mmWave multi-antenna channels. The vertical-plane beamforming $\mathbf{W}_{\mathrm{V,D}}$ is constructed as

$$\mathbf{W}_{\mathrm{V,D}} = [\mathbf{U}_{\mathrm{D},1}, \mathbf{U}_{\mathrm{D},2}, \ldots, \mathbf{U}_{\mathrm{D},N_{\mathrm{V}}}] \in \mathbb{C}^{N_{\mathrm{V}} \times N_{\mathrm{V}}} \tag{8}$$

where

$$\mathbf{U}_{\mathrm{D},\widetilde{n}_{\mathrm{V}}} = [\exp(-j\frac{2\pi}{N_{\mathrm{V}}}\widetilde{n}_{\mathrm{V}}), \exp(-j\frac{4\pi}{N_{\mathrm{V}}}\widetilde{n}_{\mathrm{V}}), \ldots, \exp(-j2\pi\widetilde{n}_{\mathrm{V}})]^T \tag{9}$$

with $\widetilde{n}_{\mathrm{V}} = 1, 2, \ldots, N_{\mathrm{V}}$. Thus, we have

$$\begin{aligned}[\tilde{\mathbf{a}}_{\mathrm{V,D}}(\phi_{\mathrm{E},p,l})]_{\widetilde{n}_{\mathrm{V}},1} = \frac{1}{\sqrt{N_{\mathrm{V}}}} &\frac{\sin\left(\pi\widetilde{n}_{\mathrm{V}} - \pi f h N_{\mathrm{V}} \cos(\phi_{\mathrm{E},p,l})/c\right)}{\sin\left(\pi\widetilde{n}_{\mathrm{V}}/N_{\mathrm{V}} - \pi f h \cos(\phi_{\mathrm{E},p,l})/c\right)} \\ &\times \exp\left(-j\pi(N_{\mathrm{V}}-1)\left(\widetilde{n}_{\mathrm{V}}/N_{\mathrm{V}} - fh\cos(\phi_{\mathrm{E},p,l})/c\right)\right). \end{aligned} \tag{10}$$

From (10), we can see that when $\widetilde{n}_{\mathrm{V}} = \frac{N_{\mathrm{V}}}{c} fh\cos(\phi_{\mathrm{E},p,l})$, the largest modular of the elements in $\tilde{\mathbf{a}}_{\mathrm{V}}(\phi_{\mathrm{E},p,l})$ can be obtained. In other words, the power of the vertical-plane signal depends on the elevation AOAs. Therefore, we can calculate the power of the output signal decoupled on the vertical plane, and use the vertical beam selection method in [8] to estimate the number of MSs.

5 Accurate AOA Estimation for Multiple Targets

In this section, we propose a low-complexity accurate AOA estimation method based on the preliminary estimates obtained in Sect. 4. As discussed in Sect. 3, in this paper, we use the $(M-1)$ time-slot signals received from the omnidirectional beamforming to estimate the AOAs of MSs. As a result, the central symmetry structure of UCyA can be fully exploited for high-accuracy angle estimation.

We first apply A-DFT to the signals on the horizontal plane and then exploit the convergence property of the Bessel function to reduce the dimension of the signals. Define an signal transformation matrix $\mathbf{W}_{\mathrm{T}} = \mathbf{W}_{\mathrm{H},T} \otimes \mathbf{W}_{\mathrm{V,T}} \in \mathbb{C}^{(2Q+1)N_{\mathrm{V}} \times N_{\mathrm{R}}}$, where $\mathbf{W}_{\mathrm{V,T}} = \mathbf{I}_{N_{\mathrm{V}}}$. The element of $\mathbf{W}_{\mathrm{H},T} \in \mathbb{C}^{(2Q+1)\times N_{\mathrm{H}}}$ is $[\mathbf{W}_{\mathrm{H},T}]_{q+Q+1,n_{\mathrm{H}}} = e^{j2\pi(n_{\mathrm{H}}-1)Q/N_{\mathrm{H}}}$, where $q \in [-Q,Q] \cap \mathbb{Z}$ and $P = \lfloor 2\pi fr/c \rfloor$. Using the same way as (6), we can write the transformed signal as

$$\mathbf{y}_{m,p,\mathrm{T}} = \sum_{l=1}^{L} \tilde{X}_{m,p,l} \tilde{\mathbf{a}}_{\mathrm{T}}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l}) + \mathbf{W}_{\mathrm{D}} \mathbf{n}_{m,p}, \tag{11}$$

where $\tilde{\mathbf{a}}_{\mathrm{T}}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l}) = \tilde{\mathbf{a}}_{\mathrm{H},T}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l}) \otimes \mathbf{a}_{\mathrm{V,D}}(\phi_{\mathrm{E},p,l})$. According to the Theorem 1 in [8], exploiting the convergence property of the Bessel function, we have

$$[\tilde{\mathbf{a}}_{\mathrm{H},T}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l})]_{q+Q+1,1} \approx j^q J_q(\varpi_{p,l}) e^{-jq\phi_{\mathrm{A},p,l}}, \tag{12}$$

where $\varpi_{p,l} = \frac{2\pi}{c} fr \sin(\phi_{\mathrm{E},p,l})$ and $J_q(\varpi_{p,l})$ is the Bessel function of the first kind of order q.

By using the estimation of signal parameters via rotational invariance techniques (ESPRIT), the elevation AOA of the l-th MS can be estimated as

$$\hat{\phi}_{\mathrm{E},p,L} = \arccos\left[(jc \ln(\lambda_{\mathrm{V},p,l})) / (2\pi fh) \right], \tag{13}$$

where $\lambda_{\mathrm{V},p,l}$ is the eigenvalue of $\mathbf{\Psi}_p$.

We proceed to present a low-complexity developed multiple signal classification (MUSIC) algorithm to estimate the azimuth AOAs. The azimuth AOAs are estimated based on the orthogonality between the signal and noise subspaces, i.e., $\mathbf{E}_{\mathrm{s},p,\mathrm{H}}$ and $\mathbf{E}_{\mathrm{n},p,\mathrm{H}}$, of received signal vectors of each UCA. Substituting $\hat{\phi}_{\mathrm{E},p,L}$ in the MUSIC estimator, then the azimuth AOAs can be estimated by

$$\hat{\phi}_{\mathrm{A},p,l} = \arg \max_{\boldsymbol{\phi}_{\mathrm{A},p,l}} \left\| \mathbf{E}_{\mathrm{n},p,\mathrm{H}}^{H} \tilde{\mathbf{a}}_{\mathrm{H},T}(\phi_{\mathrm{A},p,l}, \phi_{\mathrm{E},p,l}) \right\|_{\mathrm{F}}^{-2}. \tag{14}$$

Note that in the conventional MUSIC algorithms, the above equation needs to be solved by search the spectrum peaks within the entire azimuth angle space, which is time-consuming. In fact, in our scheme, by using directional beamforming to preliminarily estimate the azimuth AOAs at Stage 1, we have known that the the azimuth AOAs of MSs only belongs to the coverage scopes of some directional beams on the horizontal plane. Therefore, we can only search the spectrum peaks within the coverage scopes of these directional beams to further reduce the time.

6 Complexity Analysis

In this section, we analyze the software complexity of the proposed proposed multi-target sensing and tracking scheme. The complexity of horizontal AOA estimation at Stage 1 is $O(\gamma N_{\mathrm{H}}^2)$. The FFT can be used to estimate the number of MSs, whose complexity is $O(N_{\mathrm{V}} \log N_{\mathrm{V}})$ while searching the vertical beams needs the complexity of $O(N_{\mathrm{V}})$. The computational complexity of performing SVD to the received signal matrix is $O(L^3 Q^3 N_{\mathrm{V}}^3)$. The complexity of elevation AOAs estimation is $O(QN_{\mathrm{V}} + L^3)$. Estimating azimuth AOAs with 1D peak search needs the complexity of $O(L^2 Q + Q^2 LD)$, where D is the size of the search dimension.

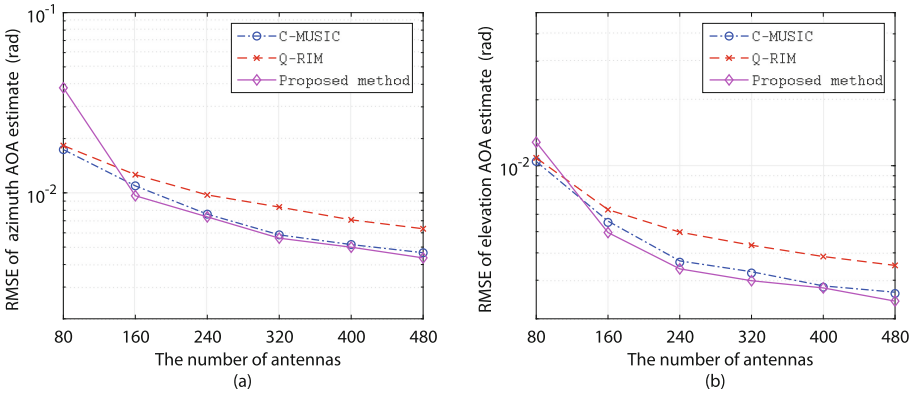


Fig. 3. The RMSE vs. the number of BS antennas for the estimation of AOAs. (a) Azimuth AOA; (b) Elevation AOA.

7 Simulation Results

In this section, simulation results are presented to demonstrate the AOA estimation performance of the proposed method. We set $f = 27$ GHz and assume that there are three MSs. The moving speeds of MSs are 1 m/s, 0.5 m/s, and 2 m/s. The UAV flies at an altitude of 20 m and moves 5 m every receiving $M = 50$ signals. The distance h between adjacent UCAs and the radius r of each UCA are 0.5λ and 2λ, respectively, where $\lambda = c/f$ is the wave length. Fig. 3 plots the root mean square errors (RMSEs) of the estimated azimuth and elevation AOAs with the increasing number of receive antennas. We compare the proposed method with conventional MUSIC (C-MUSIC) [10] and the quadric rotational invariance-based method (Q-RIM) [9]. As shown in Figs. 3(a) and 3(b), the proposedmethod is worse than C-MUSIC and Q-RIM in terms of AOA estimation, when the number of antennas is small. The reason is that under this conditions,

the proposed method may suffer from an inaccurate approximation in (12). However, as the number of antennas increases, the accuracy of the proposedscheme improves faster than C-MUSIC and Q-RIM, and quickly outperforms these two methods.

8 Conclusion

In this paper, the UAV-based simultaneously multi-target sensing and tracking scheme has been presented, which considers the computational capacity limit of UAVs and the time sensitivity of tracking tasks. The directional beamforming has been designed for the UAV to quickly estimate the horizontal AOAs of MSs and the number of MSs. The multi-target azimuth and elevation AOA estimation algorithm has also been proposed. Simulations have demonstrated the performance of the proposed scheme.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant 61827801, and in part by the Basic Scientific Research Project under Grant NS2022046.

References

1. Wang, Y., Su, Z., Zhang, N., Li, R.: Mobile wireless rechargeable UAV networks: challenges and solutions. IEEE Commun. Mag. **60**(3), 33–39 (2022)
2. Sayeed, A., Behdad, N.: Continuous aperture phased MIMO: basic theory and applications. In: Annual Allerton Conference on Communication, Control, and Computing (Allerton), pp. 1196-1203 (2010)
3. Lin, Z., Lv, T., Ni, W., Zhang, J.A., Liu, R.P.: Tensor-based multi-dimensional wideband channel estimation for mmWave hybrid cylindrical arrays. IEEE Trans. Wireless Commun. **68**(12), 7608–7622 (2020)
4. Zhu, Q., Li, H., Fu, Y., et al.: A novel 3D non-stationary wireless MIMO channel simulator and hardware emulator. IEEE Trans. Commun. **66**(9), 3865–3878 (2018)
5. Zhu, Q., Zhao, Y., Huang, Y., et al., : DEMO abstract: an UAV-based 3D spectrum real-time mapping system. In: INFOCOM WKSHPS, pp. 1-2 (2022)
6. Zhang, J., Xu, W., Gao, H., Pan, M., Han, Z., Zhang, P.: Codebook-based beam tracking for conformal array-enabled UAV mmWave networks. IEEE Internet of Things J. **8**(1), 244–261 (2021)
7. Shahmansoori, A., Garcia, G.E., Destino, G., Seco-Granados, G., Wymeersch, H.: Position and orientation estimation through millimeter-wave MIMO in 5G systems. IEEE Trans. Wireless Comm. **17**(3), 1822–1835 (2018)
8. Lin, Z., Lv, T., Ni, W., Zhang, J.A., Zeng, J., Liu, R.P.: Joint estimation of multipath angles and delays for millimeter-Wave cylindrical arrays with hybrid frontends. IEEE Trans. Wireless Commun. **20**(7), 4631–4645 (2021)
9. Guo, X., Wan, Q., Shen, X., Dou, H.: Low-complexity parameters estimator for multiple 2D domain incoherently distributed sources. Turk. J. Elect. Eng. Comput. Sci. **3**(19), 445–462 (2011)
10. Lin, Z., Lv, T., Mathiopoulos, P.T.: 3-D indoor positioning for millimeter-Wave massive MIMO systems. IEEE Trans. Commun. **66**(6), 2472–2486 (2018)

Design and Optimization of Sounding Sequence for Multi-antenna Air-to-ground Channel Measurements

Fuqiao Duan, Kai Mao, Qiuming Zhu(✉), Yanheng Qiu, Xiaomin Chen, and Zhipeng Lin

Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
{duanfuqiao,maokai,zhuqiuming,qiuyanheng,chenxm402,linlzp}@nuaa.edu.cn

Abstract. Unmanned aerial vehicle (UAV) has been widely used in different areas, where air-to-ground (A2G) channel characteristics are the fundamental to design reliable UAV to ground communication systems. For the channel measurement system with multiple transmitting antennas, this paper designs two sounding sequences based on the root index and nonlinear phase modulation, respectively. Due to the non-stationarity of A2G channels, the traditional time-division multiplex mode is not appropriate for the multi-antenna A2G channel measurement system anymore. The proposed sounding sequence can enable the measurement system to transmit the sounding sequence simultaneously but do not increase the complexity of data processing in the receiver. Meanwhile, the pros and cons of two sequences are analyzed. Finally, the effectiveness of two proposed methods is verified by the channel measurement in real A2G scenario.

Keywords: A2g channel measurements · Multiple antennas · Sounding sequence design · Nonlinear phase modulation

1 Introduction

Due to the advantages of low cost and strong maneuverability, unmanned aerial vehicles (UAVs) have been expected to play an important role in future communication networks [1–3]. Compared with terrestrial communication, the air-to-ground (A2G) communication is confronted with more challenges due to more obvious three-dimensional (3D) channel characteristics [4]. Channel measurement is considered as the most realistic way to obtain the characteristics of the A2G channel [5]. Therefore, it is of great significance to deeply study the A2G channel measurement.

Thanks to more flexible hardware structures, the time-domain channel measurement system based on sliding correlation technology has been widely applied in the A2G channel measurement [6–12]. So far, normally only one antenna is mounted on the UAV side due to the hardware complexity and limitation of

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 347–355, 2023.
https://doi.org/10.1007/978-981-99-1260-5_44

UAV payload [7–9]. Therefore, they do not need to consider the transmitting order of sounding sequences. With the development of multi-antenna communication and the increasing demand of angle estimation, some studies on A2G channel measurement also involved multiple antennas on the UAV side [11,12]. Literature [12] uses a time-reversal space-time block code-based sequence as a sounding sequence to extract the channels through which different transmitted signals pass. Note that it is the easiest way to transmit the sounding sequence from different transmitting antennas in time-division multiplex mode. However, it could cause some deviations of measured channels between different pair of antennas due to the non-stationarity of A2G channel.

To fill up this gap, this paper proposes two design methods of sounding sequence for multi-antenna A2G measurement system to enable simultaneous signal transmitting and simplify the data processing in the receiver as well. The proposed methods are also verified by the field measurements in real A2G scenario.

2 Channel Measurement System

The channel measurement system in this paper is divided into two parts, i.e., the UAV transmitting unit and the ground receiving unit as shown in Fig. 1. In the UAV transmitting unit, the signal processing module inside field programmable gate array (FPGA) is mainly used to generate the sounding sequence. In the ground receiving unit, the received data is processed in the software defined radio (SDR) platform, where the hardware real-time correlation and optimization algorithm are performed to extract channel impulse response (CIR). The industrial computer and the disk array are used to display and store the CIRs for subsequent analysis of channel characteristics, respectively. In addition, both the UAV transmitting unit and the ground receiving unit are equipped with GPS modules for recording the position and providing trigger signal. The overall hardware parameters of the measurement system are summarized as shown in Table 1.

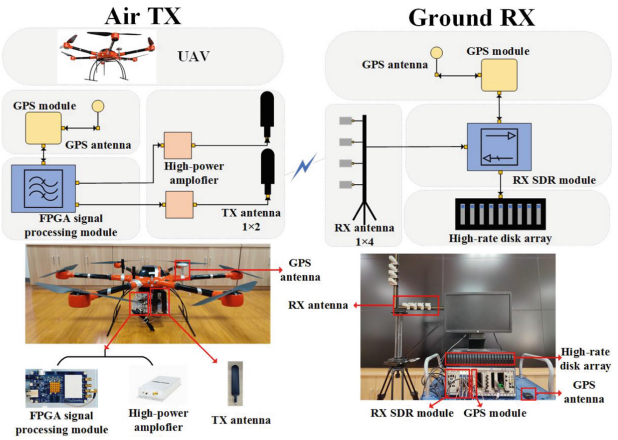


Fig. 1. Block diagram of multi-antenna measurement system.

Table 1. Summary of main components for the A2G channel measurement system.

Types	Features
RX/TX GPS module	L1 frequency (1575.42 MHz) / 1PPS(RMS) 20ns
FPGA signal processing module	70MHz ∼ 6GHz / Bandwidth 100MHz
High-power amplifier	600 ∼ 6000 MHz / Gain 42 dB
TX antenna	1400 ∼ 2700 MHz , 3200 ∼ 3900 MHz
RX antenna	2.4 GHz±50 MHz, 3.5 GHz±40 MHz
RX SDR module	200 ∼ 4400 MHz / Bandwidth 100MHz

3 Sounding Sequence Design for Multi-antenna Measurement System

3.1 Sliding-correlation-based Channel Measurement

In the multi-antenna A2G measurement system, the received signal of each receiving antenna can be seen as the superposition of signal from different pair of transceiver antenna, which can be expressed as

$$y(t) = x(t) * h_{M\times N}(t,\tau) \tag{1}$$

where N and M represent number of transmitting antennas and receiving antennas, respectively. $*$ represents convolution operation. $x(t) = [x_1(t)x_2(t)\cdots x_N(t)]^{\mathrm{T}}$ represents the vector of transmitted signal, $y(t) = [\tilde{y}_1(t)\tilde{y}_2(t)\cdots \tilde{y}_M(t)]^{\mathrm{T}}$ represents the vector of received signal. $h_{M\times N}(t,\tau)$ is a $M \times N$ matrix, and each matrix element corresponds to the channel impulse response between m th $(m = 1, 2, \ldots, M)$ receiving antenna and n th $(n = 1, 2, \ldots, N)$ transmitting antenna, which can be further expressed as

$$h_{m,n}(t,\tau) = \sum_{l=1}^{L(t)} a_l(t)\tilde{h}_{m,n,l}(t)\delta(\tau - \tau_l(t)) \tag{2}$$

where $L(t)$ represents the number of effective propagation multipaths, and $a_l(t)$ and $\tau_l(t)$ represent the amplitude and delay of the l th path, respectively. $\tilde{h}_{m,n,l}(t)$ represents the normalized channel complex fading.

In the sliding-correlation-based time-domain channel measurement system, the sliding correlation operation is performed to obtain the CIRs. The output of sliding correlation operation can be expressed as

$$r_{m,n}(t,\tau) = y_m(t) \otimes x_n(t) = \sum_{k=1}^{N} h_{m,k}(t,\tau) * x_k(t) \otimes x_n(t) \tag{3}$$

where $\otimes$ represents correlation operation. Note that when the same sounding $x_k(t)$ is transmitted in each transmitting antenna, it is really hard to extract the

CIR between different pair of transceiver antenna because $r_{m,n}(t,\tau)$ is the superposition of CIR between N transmitting antennas and mth receiving antenna. Therefore, it is necessary to design different sounding sequence with low cross-correlation for each transmitting antenna.

3.2 Sounding Sequence Based on Different Root Indexes

Note that the designed ZC sequences with different root indexes have low cross-correlation results, which is a good option for multi-antenna channel measurement system. Therefore, the sounding sequence for different transmitting antenna can be expressed as

$$z^{\mathrm{r}}[p] = \exp\left(-\frac{\mathrm{j}\pi\mu_r p(p+1)}{N_{\mathrm{ZC}}}\right) \tag{4}$$

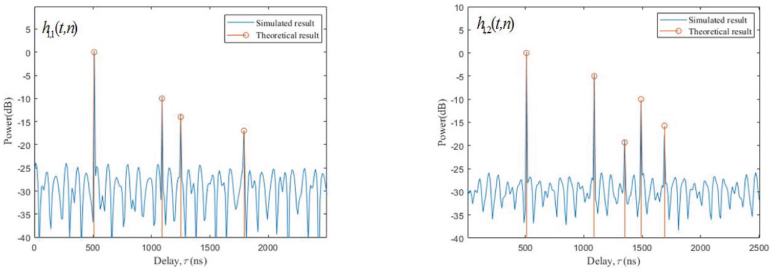
where $p = 0, 1, ..., N_{\mathrm{ZC}} - 1$, $\mu_r \in \{0, 1, ..., N_{\mathrm{ZC}} - 1\}$ is the root index of the ZC sequence, where N_{ZC} represents the sequence length. $z^{\mathrm{r}}[n]$ represents ZC sequences with different μ_r .Therefore, Eq. (3) can be rewritten as

$$r^{\mathrm{r}}_{m,n}(t,p) = y_m(p) \otimes z^{\mathrm{r}}_n(p) = N_{ZC} h_{m,n}(t,p) \tag{5}$$

So the CIR between different pair of transceiver antenna can be expressed as

$$h_{m,n}(t,p) = y_m(p) \otimes z^{\mathrm{r}}_n(p)/N_{ZC} \tag{6}$$

In order to verify the effectiveness of ZC sequences with different root indexes , the simulation is performed in a measurement system with dual transmitting antennas. The parameters of sounding sequence are set to $N_{\mathrm{ZC}} = 1023$, $\mu_1 = 1$, $\mu_2 = 2$.The simulation results are shown in Fig. 2. The test results in Fig. 2 (a) and (b) are obtained by performing the sliding correlation for received signal with the ZC sequences with $\mu_1 = 1$ and $\mu_2 = 2$, respectively. It can be seen that the CIRs between different pair of transceiver antenna can be well extracted, which are also consistent with the preset CIRs.



(a) (b)

Fig. 2. Sounding sequence verification based on different root indexes.

Moreover, the dynamic range of the channel measurement system is also an important index, which can be calculated by

$$D = \frac{\max(peak_l(r_{m,n}[t,p]))}{\max(peak(\{r_{m,n}[t,p]|r_{m,n}[t,p] \neq peak_l(r_{m,n}[t,p])\}))} \tag{7}$$

where $peak()$ refers to the peak value, $peak_l(r_{i,j}[t,n])$ represents the sliding correlation peak of the l th propagation path. Although ZC sequences with different indices have low correlation with each other, there is still have a certain impact on the dynamic range of the channel measurement system. The theoretical dynamic range of different root indexes can be expressed as

$$D_k^{\mathrm{r}} = \frac{\max\left(peak(z_n^r[p] \otimes z_n^r[p])\right.}{\max\left(peak(\sum_{i=1,i\neq n}^{N} z_i^r[p] \otimes z_n^r[p])\right)} \tag{8}$$

where $z_n^r[p]$ represents the ZC sequences with different root index transmitted by the n th transmitting antenna.

We simulate the theoretical dynamic range of sliding correlation by using ZC sequences with root indices $\mu_1 = 1$ and with different μ_2 from 2 to $N_{\mathrm{ZC}} = 1023$. The simulated results are shown in Fig. 3. The abscissa is the different values of μ_2, and the ordinate is the theoretical dynamic range. It can be seen that the theoretical dynamic range is changing along with different μ_2. Therefore, we should select appropriate root index for sounding sequences according to the results in Fig. 3 to ensure better dynamic range.

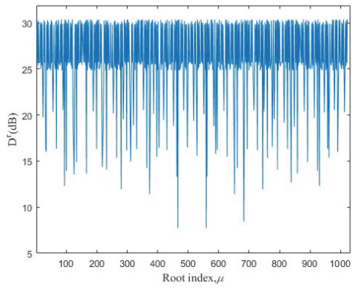


Fig. 3. Dynamic range of designed sounding sequence with different μ_2.

3.3 Sounding Sequence Based on Nonlinear Phase Modulation

In this section, another sounding sequence based on nonlinear phase modulation is designed to compensate the disadvantage of the method based on different root index. The sounding sequence based on nonlinear phase modulation can be expressed as

$$z^{\mathrm{pm}}[p] = \exp\left(-\mathrm{j}\pi\mu p(p+1)/N_{ZC} + \mathrm{j}n\varphi\right) \tag{9}$$

where $n = 0, 1, ..., N-1$ represents the index of different transmitting antennas, and φ represents the nonlinear phase, which can be written as

$$\varphi = -\pi\mu((2n+1)L + L^2)/N_{ZC} \tag{10}$$

where L represents the weights of phase modulation, which determines the value of the modulation phase. Therefore, L should be neither too large nor too small.

The output result of sliding correlation based on the phase modulation method can be expressed as

$$r_m^{\mathrm{pm}}(t,p) = y_m(t,p) \otimes z^{\mathrm{pm}}(p) = N_{ZC} \sum_{k=1}^{N} h_{m,k}(t, n + (k-1)L) \tag{11}$$

where $r_m^{\mathrm{pm}}(t,p)$ represents the convolution result of signal received by the m th receiving antenna, which includes all the channels that signals from different transmitting antennas experince, and can be expressed as

$$h_{m,n}(t,p) = r_m^{\mathrm{p}}(t,p)/N_{ZC}, p = (n-1)L(n-1)L + 1, ..., nL-1 \tag{12}$$

In order to further verify the effectiveness of proposed phase modulation method, similar as Sect. 3.1. The simulation parameters are set as $N_{\mathrm{ZC}} = 1023$, $L = 250$. The simulation results are shown in Fig. 4. It can be found that the four-path channel and the five-path channel are effectively separated and extracted.

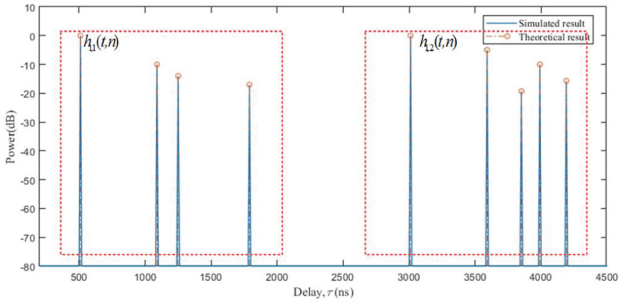


Fig. 4. Simulation of designed sounding sequence based on nonlinear phase modulation.

For the sounding sequence based on phase modulation, only the phases are changed but the values of the sequence remain unchanged. Therefore, the sliding correlation of ZC sequences based on phase modulation will not be affected by each other, which achieves good dynamic range. The dynamic range of the sounding sequence based on different root indexes is much larger, which can be visually found from the results in Fig. 2 and Fig. 4.

Regarding to two proposed sounding sequence design methods in this paper, they have their own pros and cons for different applications. For the sounding sequence based on different root indexes, only the root index needs to be changed for the sequence. Therefore, it is a great option for large-scale multi-antenna measurement system. For the sounding sequence based on nonlinear phase modulation, it has higher dynamic range but the length of sequence needs to be changed which would be too long for large-scale multi-antenna measurement system. Therefore, the proposed two methods can compensate each other.

4 Measurement Verification

In order to verify the effectiveness of the proposed two sounding sequence methods for multi-antenna design, this paper conducts the actual A2G measurement in a campus scenario, as shown in Fig. 5.

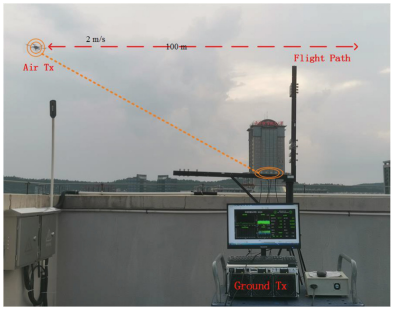
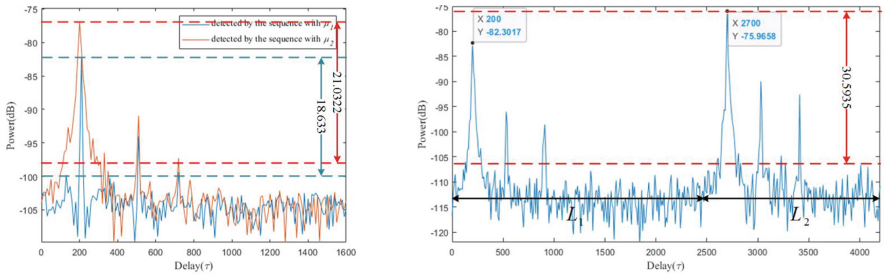


Fig. 5. A2G measurement in a campus scenario.

The actual measurement uses two transmitting antennas and four receiving antennas as a verification case. The parameters of sounding sequence based on the root index are set to $N_{\mathrm{ZC}} = 1023$, $\mu_1 = 1$, $\mu_2 = 2$. The parameters of sounding sequence based on phase modulation is set as $N_{\mathrm{ZC}} = 1023$, $L_1 = 0$, $L_2 = 250$. The measurement system are configured with the carrier frequency of 2.4 GHz and the sampling rate of 100 MHz. The measured results are shown in Fig. 6. Because the distance between the transmitting antennas is very close, the delay of the line of sight (LoS) path is basically the same. In Fig. 6(b), the LoS path of the two channels differs by 250 points, which is exactly the same as the setting parameter $L_2 - L_1$. It can be seen that both of two proposed methods can be used to effectively extract the CIRs when the sounding sequences are transmitted from different transmitting antennas simultaneously, and the sounding sequence based on phase modulation has larger dynamic range.



(a) Based on root index (b) Based on phase modulation

Fig. 6. Measured results of two different sounding sequences.

5 Conclusion

In this paper, two sounding sequences for A2G channel measurement system with multiple transmitting antennas have been proposed, i.e., sounding sequences based on different root index and nonlinear phase modulation. In addition, the effectiveness of the two methods has been verified by the actual A2G measurement. Both the simulated and measured results have shown that the proposed two methods can be used to extract CIRs between different pair of transceiver antenna. The sounding sequences based on different root index is suitable for the measurement system with a large number of transmitting antennas, and another method is suitable for a small number of antennas but has larger dynamic range. In the future, we will use the multi-antenna measurement sequence for MIMO channel measurement and co-channel measurement.

Acknowledgments. This work was supported in part by the National Key Scientific Instrument and Equipment Development Project under Grant No. 61827801, in part by the Key Technologies R&D Program of Jiangsu (Prospective and Key Technologies for Industry) under Grants BE2022067 and BE2022067-3., in part by Natural Science Foundation of Jiangsu Province, No. BK20211182, in part by Key Project of Aeronautical Science Foundation of China, No.2020Z07300 9001, in part by the Future Network Scientific Research Fund Project, No. FNSRFP-2021-YB-04.

References

1. Zhu, Q., Zhao, Z., Mao, K., Chen, X., Liu, W., Wu, Q.: A Real-Time Hardware Emulator for 3D Non-Stationary U2V Channels, IEEE Trans. Circ. Syst. Regul. Pap. **68**(9), pp. 3951–3964 (2021)
2. Q. Zhu et al.: Geometry-Based Stochastic Line-of-Sight Probability Model for A2G Channels Under Urban Scenarios. IEEE Trans. Antennas Propag., **70**(7), pp. 5784–5794 (2022)
3. Cui, Z., Briso-Rodrguez, C., Guan, K., Calvo-Ramrez, C., Ai, B., Zhong, Z.: Measurement-Based Modeling and Analysis of UAV Air-Ground Channels at 1 and 4 GHz. IEEE Antennas Wirel. Propag. Lett. **18**(9), pp. 1804-1808 (2019)

4. Mao, K., Zhu, Q., Song, M., Li, H., Ning, B., et al.: Machine Learning-Based 3D Channel Modeling for U2V mmWave Communications. IEEE Internet of Things J. early access 9(8), pp. 17592–17607 (2022)
5. Wang, C., Huang, J., Wang, H., Gao, X.,You, X., Hao, Y.: 6G Wireless Channel Measurements and Models: Trends and Challenges, IEEE Veh. Technol. Magazine, **15** (4), pp. 22–32 (2020)
6. Calvo-Ramrez, C., Gonzalez-Plaza, A., Briso, C., Prez-Yuste, A.: Wide band propagation measurements and modelling for low altitude UAVs, 12th European Conference on Antennas and Propagation (EuCAP 2018), pp. 1–5 (2018)
7. Schneckenburger, X., et al.: Measurement of the l-band air-to-ground channel for positioning applications, IEEE Trans. Aerosp. Electron. Syst. **52** (5), pp. 2281–2297 (2016)
8. Ning, B., et al.: A UAV-aided channel sounder for air-to-ground channel measurements, Phys. Commun. **47**, pp. 1874–4907 (2021)
9. Khawaja,W., Guvenc, I., Matolak, D.: UWB Channel Sounding and Modeling for UAV Air-to-Ground Propagation Channels, IEEE Global Communications Conference (GLOBECOM), pp. 1–7 (2016)
10. Cai, X., et al.: Low altitude UAV propagation channel modelling, European Conference on Antennas and Propagation (EUCAP), pp. 1443–1447 (2017)
11. Jiang, Y., Tiwari, A., Rachid, M., Daneshrad, B.: MIMO for airborne communications Industry Perspectives. IEEE Wirel. Commun., vol. 21, no. 5, pp. 4–6 (2014)
12. Rice, M., Saquib, M.: MIMO equalization for helicopter-to-ground communications, Military Communications Conference, pp. 501–506 (2011)

LoS Probability Prediction for A2G mmWave Communications by Using Ray-tracer Under Virtual Urban Scenarios

Yue Tian[1], Minghui Pang[1], Hongtao Duan[2], Bing Lv[2], Xiaomin Chen[1], Qiuming Zhu[1(✉)], and Boyu Hua[1]

[1] Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
{tian_yue,pangminghui,chenxm402,zhuqiuming,byhua}@nuaa.edu.cn
[2] The State Radio Monitoring Center, Beijing 100037, China
{duanht,lvbing}@srrc.org.cn

Abstract. Line-of-sight (LoS) probability is of great significance for reliable air-to-ground (A2G) millimeter wave (mmWave) communication. In this paper, a machine learning (ML)-based LoS probability prediction method for A2G mmWave communication is proposed. A fully connected neural network optimized by Bayesian optimization is utilized to predict LoS probability accurately with the type of scene, ground distance and vertical height between unmanned aerial vehicles (UAVs) and ground receivers as the inputs. For the difficulty of measured data acquisition, massive channel data is generated by ray-tracing (RT) simulation under the virtual scenarios reconstructed by statistical building characteristics. Simulation results show that the proposed method is consistent well with the RT simulation and verified the good performance of LoS probability prediction in four virtual urban scenarios by comparison with other existing models.

Keywords: LoS probability · A2G mmWave channel · Ray tracing · Machine learning · Bayesian optimization

1 Introduction

Unmanned aerial vehicles (UAVs) have been a significant component in 5G and beyond 5G wireless communication. Due to its flexibility of deployment, UAVs can be deployed as aerial base stations and access points, which can effectively improve the quality and coverage of air-to-ground (A2G) communication. The millimeter wave (mmWave) A2G communication with huge bandwidth can meet the surging demand of high data rate and transmission quality [1]. The three-dimensional mobility of UAVs and the propagation characteristics of mmWave signal bring new challenges to the A2G channel modeling [2].

Line-of-sight (LoS) probability has a significant impact on the performance of A2G channel modeling [3]. Urban scenario is the most common application

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 356–363, 2023.
https://doi.org/10.1007/978-981-99-1260-5_45

scenario in 5G and the LoS blockage of A2G channel is mainly involved in buildings [4]. Some standard models, such as International Telecommunication Union (ITU-R) M.2135–1 [5], the Third Generation Partnership Project (3GPP) TR 38.901 [6], 5G Channel Model (5GCM) [7], calculated the LoS probability based on massive data obtained from measurement campaigns. However, These empirical models requires high cost measurement and are only suitable for specific low-altitude environments. Several researches used the statistical geometry information of the built-up scenarios in urban environments to derive the LoS probability prediction model [8–11]. However, these analytical methods are limited by the complexity and inflexibility of theoretical derivation.

Recently, machine learning (ML) algorithms have been widely used in channel modeling and channel parameter estimation. ML methods can independently analyze data characteristics, learn the potential complex relationship between inputs and outputs, and predict the trend in data. Regarding related research on LoS probability prediction, the authors in [12] presented an artificial neural network (ANN)-based model to predict the indoors LoS probability. And a generative neural network consisting of a two-stage structure for mmWave channel modeling was proposed in [13], in which a classify neural network is used to model the link state probability. Due to the limited researches on ML-based LoS probability prediction, a ML-based LoS probability prediction for A2G mmWave communication by ray-tracing is developed. The LoS probability of different A2G links is obtained by ray-tracing (RT) simulation in virtual urban scenarios reconstructed by statistical building characteristics, which is used for a fully connected neural network (NN) optimized by Bayesian optimization (BO) training to achieve the accurate LoS probability prediction.

The rest of the paper is organized as follows. In Sect. 2, the ML-based LoS probability prediction of A2G mmWave communication is proposed. Section 3 introduces the RT data acquisition and optimized neural network training in detail. The simulation results and comparison with different existing models are given in Sect. 4. Finally, the conclusions are drawn in Sect. 5.

2 The Proposed LoS Probability Prediction

In this paper, we consider the A2G links between the UAVs (also the transmitters) and the ground receivers. The ground distance and vertical height between UAVs and receivers are denoted as d and h, respectively. The framework of the proposed LoS probability prediction is depicted in Fig. 1. It mainly includes RT data acquisition, neural network training and optimization.

Firstly, the virtual urban scenarios with statistical building characteristics are reconstructed due to its simplicity and generalization to more 5G application scenarios. The LoS probability between UAVs and receivers of different A2G links is obtained by RT simulation to meet the requirement of massive data. Then, A fully connected neural network is built with the input of d, h, scene type k and hyperparameter optimization is carried out by using BO to obtain the optimal prediction model. Finally, Validation is conducted to prove that utilizing

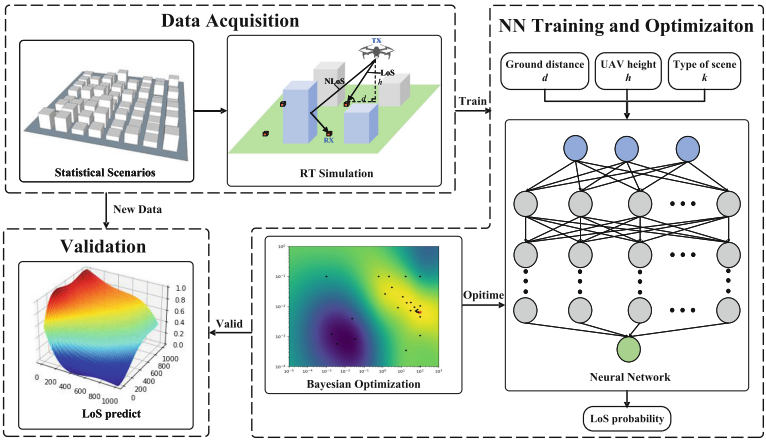


Fig. 1. The frame of the proposed LoS probability prediction.

the ability of the neural network can fit the complex non-linear relationship of the input feature and output target, and achieve the accurate prediction of LoS probability for A2G mmWave channel in statistical virtual scenarios.

3 Prediction Network Training and Optimization

3.1 Reconstructed Virtual Scenarios and Data Collection

The urban scenarios based on the statistical characters of buildings are reconstructed to get LoS probability data by RT simulation. In order to cover a wide range of building distribution in urban environments, ITU-R P.1410 [14] presented three statistics parameters $\{\alpha, \beta, \gamma\}$ to describe the geometrical statistics characteristic of the building deployment, where α denotes the ratio of building area to total land area, β denotes the mean number of buildings per square kilometer, γ denotes the parameter used to generate the Rayleigh distribution heights of the buildings, which can be expressed as

$$P(h) = \frac{h}{\gamma} \exp(-\frac{h^2}{2\gamma^2}). \tag{1}$$

The four generated statistical scenarios are selected as mentioned in [15]. For a brief representation, these scenarios limited to 500 m $\times$ 500 m region are depicted in Fig. 2. The buildings are supposed to be square and deployed in regular grid. The width of the buildings and streets denoted by $W = 1000\sqrt{\alpha/\beta}$ and $S = 1000/\sqrt{\beta} - W$ [16]. RT simulation is utilized in these reconstructed scenes to determine whether there is an LoS path in each A2G link. Each $P_{d,h}^{LoS}$ corresponding to different d and h is calculated by at least 500 data to ensure the accuracy, which is expressed as

$$P_{d,h}^{LoS} = \frac{N_{d,h}^{LoS}}{N_{d,h}^{ALL}}, N_{d,h}^{ALL} > 500 \tag{2}$$

where $N_{d,h}^{LoS}$ denotes the number of LoS paths in A2G links, $N_{d,h}^{ALL}$ denotes the number of all A2G links of d and h.

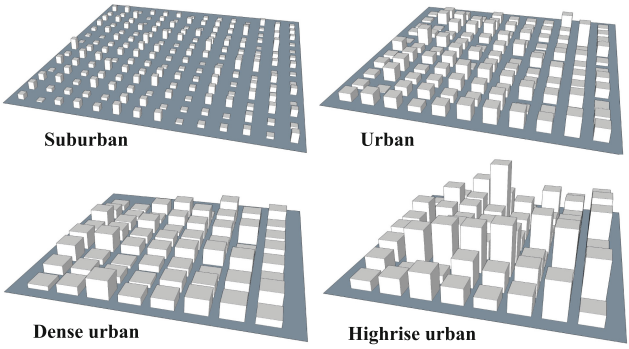


Fig. 2. An illustration of the four reconstructed scenarios.

3.2 NN-based LoS Probability Prediction Network

In order to predict LoS probability accurately by massive RT simulation data, a fully connected neural network shown in Fig. 3 is proposed. The dataset $\mathcal{D} = \{(x, y)\}$ is obtained by the RT method. The input vector of NN is denoted by $x = (d, h, k)$, where k is converted to 4-dimensional vector by using one-hot encoding. y is corresponding LoS probability, also the target output of NN. We split $\mathcal{D}$ into training set $\mathcal{D}_{train}$ and test set $\mathcal{D}_{test}$ by a ratio of 7: 3 for each scene category. The neural network is employed to fit the LoS probability model $P_{LoS}(x)$ on $\mathcal{D}_{train}$ and evaluate performance on $\mathcal{D}_{test}$.

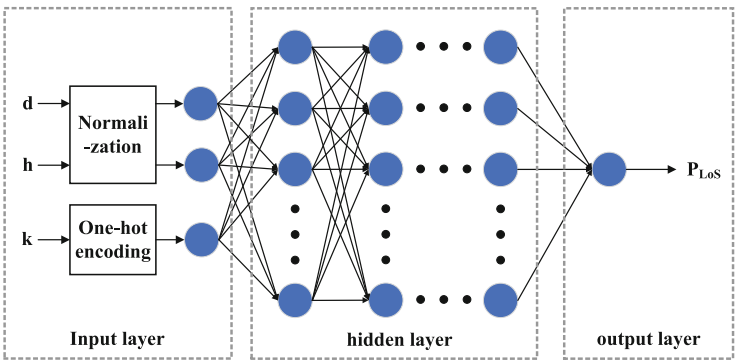


Fig. 3. The structure of the proposed fully connected NN.

For the proposed full connected NN, ReLU is used as the nonlinear activation function for each hidden layer and Sigmoid is used for output layer to limit the probability output to [0,1]. The loss function of this NN is defined by the mean

square error (MSE) between the prediction $P_{LoS}(x)$ and target output y, which can be expressed as

$$\mathcal{L}(x, y) = \mathbb{E}\left[(P_{LoS}(x) - y)^2\right] \tag{3}$$

when $(x, y) \in \mathcal{D}_{train}$, the NN minimizes $\mathcal{L}(x, y)$ by backpropagation algorithm to update the connection weights and bias. When $(x, y) \in \mathcal{D}_{test}$, $\mathcal{L}(x, y)$ is used for the evaluation of model accuracy and generalization on unseen dataset.

3.3 Hyperparameter Optimization of NN

The setting of hyperparameters before training has direct influence on the performance of neural network. Bayesian optimization can use prior information to find the best hyperparameters combination of machine learning algorithm considered as a complex function quickly with less cost. In this paper we use BO with Gaussian process prior to tune the three hyperparameters of the neural network, i.e., the number of layers, the number of nodes and the learning rate.

To employ BO first is to give the objective function and the bounded sets of hyperparameters first. We use the mean of the k-folded cross-validation loss function as the optimization objective function and split up the training set $\mathcal{D}_{train}$ into training set $\mathcal{D}_{train}^{'(i)}$ and validation set $\mathcal{D}_{valid}^{'(i)}$ in the process of optimization

$$\theta_{\text{opt}} = \arg\min_{\theta \in \Theta} (f(\theta)) = \arg\min_{\theta \in \Theta} (\frac{1}{k} \sum_{i=1}^{k} \mathcal{L}(\theta; \mathcal{D}_{train}^{'(i)}, \mathcal{D}_{valid}^{'(i)})) \tag{4}$$

where θ_{opt} denotes the optimal hyperparameter combination, Θ denotes the search space of the hyperparameters composed by the number of layers (1 to 5), the number of nodes (3 to 200) and learning rate (10–5 to 1), and k is set to 5. We use Gaussian process as the surrogate model to get a posterior approximating the objective function and use this posterior to construct the expected improvement (EI) acquisition function to select the next evaluation until the minimum is found. The optimization process iterates 50 epochs to find θ_{opt} and the hyperparameters of proposed NN are tabulated in Table 1.

Table 1. Hyperparameters of the Proposed NN

Hyperparameter	Value
Number of layers	4
Number of nodes	193
Learning rate	0.00109
Optimizer	Adam
Epochs	500
Batch size	512

4 Simulation and Validation

The simulation setup of employing ray tracing in urban environment is as follows. The size of ground terrain is 1 km × 1 km. Considering the simulation time and randomness of data acquisition, the receivers are set in grid deployment with the resolution of 10m, which are fixed 2 m above the ground. And the UAVs are placed randomly at every 5 m of the height in the range of 10 m to 1000 m m. The simulation is carried out in 28 GHz with bandwidth of 500 MHz and provide the LoS state of each A2G link in four virtual urban scenarios.

The proposed NN is trained by the RT data obtained by the simulation settings, then we get the NN-based model of LoS probability prediction for A2G mmWave communication in four urban environments and the MSE calculated on

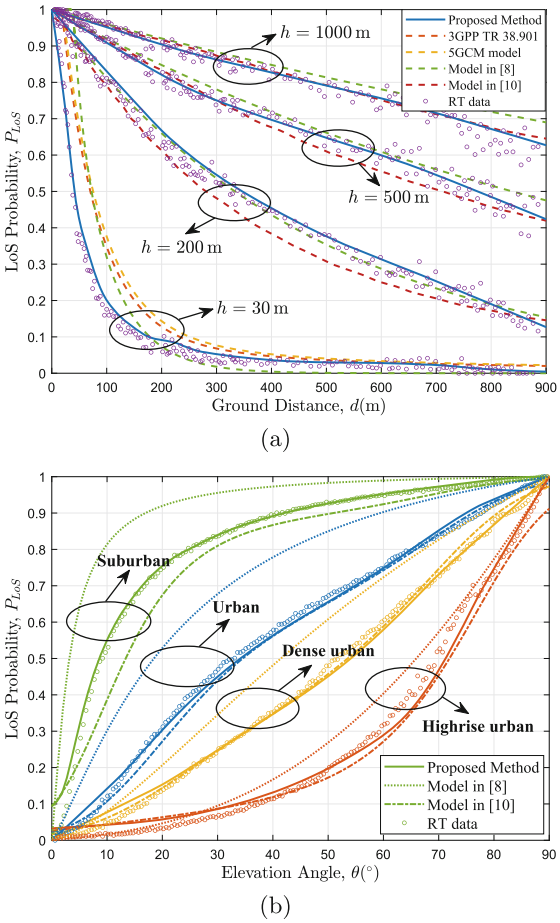


Fig. 4. Comparison with different models in urban scenarios. (a) LoS probability VS ground distance (b) LoS probability VS elevation angle.

$\mathcal{D}_{test}$ is 0.0003903. To verify the validity of the proposed method, we compare with RT simulation data generated randomly from Sect. 3.1, 3GPP standard model, 5GCM model, and models in [8] [10]. The LoS probability of different models with respect to d in four different height (30 m, 200 m, 500 m, 1000 m m) is shown in Fig. 4(a)

The simulation results in Fig. 4(a) shows that the proposed method matches well with the RT simulation and the prediction result with the same trend as models in [8,10] verifies the good performance and validity of the proposed network, while 3GPP and 5GCM standard models are close to each other at low altitude cases and the LoS probability is higher than RT simulation. This result is same as that in [11], as these standard models fail to capture the statistic characteristics of the building height and deployment. For a better representation and comparison in two dimensions, the elevation angle $\phi = arctan(h/d)$ is used to depict LoS probability in Fig. 4(b). From Fig. 4(b), we can get that these four statistical scenarios have different effects on LoS probability. With increase of the density and average height of buildings, the possibility of buildings blocking the LoS paths also increases. Compared with the existing A2G models in [8,10], the proposed approach shows the similar trends and has better consistency with RT simulation data in four statistics scenarios.

5 Conclusion

In this paper, we have proposed a LoS probability prediction based on ML method for A2G mmWave channel in urban environment. The proposed method has reconstructed four virtual urban scenarios according to the statistical building characteristics and RT simulation has been employed to obtain massive LoS probability data of each A2G link. A fully connected neural network optimized by Bayesian optimization has been constructed to predict the LoS probability with type of scene, ground distance and vertical height between UAVs and ground receivers. Simulation results have shown that the MSE of the proposed neural network on test set is 0.0003903, and compared with the standard models and other existing models, the proposed method can maintain better consistency with the RT simulation and has better performance in LoS probability prediction in four virtual urban scenarios.

Acknowledgment. This work was supported in part by the National Key Scientific Instrument and Equipment Development Project under Grant No. 61827801, in part by the Key Technologies R&D Program of Jiangsu (Prospective and Key Technologies for Industry) under Grants BE2022067 and BE2022067-3, in part by Natural Science Foundation of Jiangsu Province, No. BK20211182, in part by Key Project of Aeronautical Science Foundation of China, No. 2020Z0730-09001, in part by the Future Network Scientific Research Fund Project, No. FNSRFP-2021-YB-04.

References

1. Wang, X., Kong, L., Kong, F., et al.: Millimeter wave communication: a comprehensive survey. IEEE Commun. Surv. Tutor. **20**(3), 1616–1653 (2018)
2. Mao, K., Zhu, Q., Song, M., et al.: Machine learning-based 3D channel modeling for U2V mmWave communications. IEEE Internet Things J. (2022). https://doi.org/10.1109/JIOT.2022.3155773
3. Zhu, Q., Bai, F., Pang, M., et al.: Geometry-based stochastic line-of-sight probability model for A2G channels under urban scenarios. IEEE Trans. Antennas Propag. **70**(7), 5784–5794 (2022)
4. Khawaja, W., Guvenc, I., Matolak, D.W., et al.: A survey of air-to-ground propagation channel modeling for unmanned serial vehicles. IEEE Commun. Surv. Tutor. **21**(3), 2361–2391 (2019)
5. ITU-R: Guidelines for evaluation of radio interface technologies for IMT-advanced. REP. M.2135–1, (2009)
6. 3GPP: Study on channel model for frequencies from 0.5 to 100 GHz. Tech. Rep. TR 38.901 V15.0.0, (2018)
7. 5GCM: 5G Channel Model for Bands up to 100 GHz. Tech. Rep, (2016)
8. Al-Hourani, A., Kandeepan, S., Lardner, S.: Optimal LAP altitude for maximum coverage. IEEE Wirel. Commun. Lett. **3**(6), 569–572 (2014)
9. Al-Hourani, A.: On the probability of line-of-sight in urban environments. IEEE Wirel. Commun. Lett. **9**(8), 1178–1181 (2020)
10. Mohammed, I., Collings, I. B., Hanly, S. V.: Line of sight probability prediction for UAV communication. In: IEEE International Conference on Communications Workshops, pp. 1–6 (2021)
11. Gapeyenko, M., Moltchanov, D., Andreev, S., Heath, R.W.: Line-of-sight probability for mmWave-based UAV communications in 3D urban grid deployments. IEEE Trans. Wirel. Commun. **20**(10), 6566–6579 (2021)
12. Yang, W., Zhang, J., Zhang, J.: Machine learning based indoor line-of-sight probability prediction. In: Int. Symp. Antennas Propag., pp. 1–3 (2019)
13. Xia, W., Rangan, S., Mezzavilla, M., et al.: Millimeter wave channel modeling via generative neural networks. In: IEEE Globecom Workshops, pp. 1–6 (2020)
14. ITU-R: Propagation data and prediction methods required for the design of terrestrial broadband radio access systems operating in a frequency range from 3 to 60 GHz. Rec. P. 1410–5, (2012)
15. Holis, J., Pechac, P.: Elevation dependent shadowing model for mobile communications via high altitude platforms in built-up areas. IEEE Trans. Antennas Propag. **56**(4), 1078–1084 (2008)
16. Al-Hourani, A., Kandeepan, S., Jamalipour, A.: Modeling air-to-ground path loss for low altitude platforms in urban environments. In: IEEE Global Communications Conference, pp. 2898–2904 (2014)

Author Index

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
Q. Liang et al. (Eds.): CSPS 2022, LNEE 873, pp. 365–367, 2023.
https://doi.org/10.1007/978-981-99-1260-5