Lecture Notes in Electrical Engineering 1032

Wei Wang
Xin Liu
Zhenyu Na
Baoju Zhang *Editors*

Communications, Signal Processing, and Systems

Proceedings of the 12th International Conference on Communications, Signal Processing, and Systems: Volume 1

Springer

Lecture Notes in Electrical Engineering

Volume 1032

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. ****

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

Communications, Signal Processing, and Systems

Proceedings of the 12th International Conference on Communications, Signal Processing, and Systems: Volume 1

Springer

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

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

Xin Liu
Dalian University of Technology
Dalian, China

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

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-99-7539-6 ISBN 978-981-99-7505-1 (eBook)
https://doi.org/10.1007/978-981-99-7505-1

Contents

Text Recognition Method for Handwritten and Natural Scene Text Image: A Review

Ximin Sun, Jing Zhou, Mingda Wang, Jiang Chang, and Guoyu Ma

Abstract Text is an important medium for human communication, and it is also the most direct and clear form for people to understand and obtain information. Text recognition technology is currently used in many fields, such as intelligent transportation technology, automatic driving technology, blind assistance system, web image search system. In this paper, the text recognition method for handwritten and natural scene text image is concluded. The characteristic and the development prospect of text image has been reviewed.

1 Introduction

Text is an important medium for human communication, and it is also the most direct and clear form for people to understand and obtain information [1–3].

The development of automatic driving technology is closely related to text detection and recognition [1]. Vehicles can obtain road information through cameras, intelligently identify road signs and traffic symbols on the road ahead, and self-judge the most scientific and safe driving mode, bringing great convenience to people's lives. At the same time, it also has high requirements for the speed of the recognition model. The model needs to be able to detect and identify the text information in a very short time when driving and achieve true automatic driving by uploading it to the cloud (Fig. 1).

Although a large number of blind people believe that smartphones can interact with people, there are significant losses in many experiences, and these losses also have a lasting negative impact on their psychology. The natural scene text recognition technology can identify the fonts such as road signs and shops through the camera, and convert these text information into voice broadcasts, to facilitate people with visual impairment to obtain a better surrounding environment and improve their freedom of movement [2] (Table 1).

X. Sun (✉) · J. Zhou · M. Wang · J. Chang · G. Ma
State Grid Ecommerce Technology Co., Ltd., Tianjin, People's Republic of China
e-mail: sun_gw2023@163.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_1

Fig. 1 The application of text recognition

Table 1 The challenge of text recognition application

Application	Site	Challenge
Automatic driving technology	Outdoor	Fast speed and high accuracy
Blind assistant system	Outdoor or indoor	Complex environment
Network image search system	Computer screen	High accuracy

Web search is an important way for humans to access information, but it is still a huge challenge for machines to understand the content on web pages [3]. Currently, most of the searches are text searches. However, for some illiterate people or some complex fonts, there is no way to make a correct input of text, which has a certain impact on the text-based web search. If text in images can be intelligently extracted, the efficiency of the web search can be improved, and large manpower and material resources can be saved.

2 The Characteristic of Text Image

This paper mainly focuses on the recognition of two types of scene text images: handwritten text image and natural scene text image [4].

2.1 The Characteristic of Handwritten Text Image

Compared with printed document images with simple background and foreground, handwritten text images are more complicated [5]. In handwritten text images, the pattern and structure of text would be deformed due to different personal writing habits and fonts. Text can be difficult to separate because the color of the handwriting is similar to the background. It is also difficult to segment line image or word image. Therefore, traditional text recognition methods cannot effectively process these text images. Compared with ordinary printed document image, the text in handwritten text image has the following different image characteristics.

In handwritten text images, uncertain character direction and mutual connection between words and strokes often occur. It makes difficult to accurately segment the text into word images or character images, which results in poor end-to-end system performance. In addition, in the recognition of free handwritten text images, the segment the image into word images or character images is also a big difficulty to be overcome [6].

Different people's writing would cause the variation of the pattern and structure. In terms of letters and numbers, there are only 62 kinds and the writing is simple. The writing has obvious regional characteristics. Different people's writing habits would lead to the variation of the pattern structure, such as nonstandard and uneven strokes, unfixed relative position between strokes, stroke connection, character connection [7].

In handwritten text images, the text line is affected by the upper and lower text lines. For example, the connection between the upper line stroke and the current line stroke, etc. At the same time, in the application of handwritten text recognition, there are also inserted lines, smeared characters. In some cases, the difference between the noise and the normal text is small, which makes filtering algorithm very difficult [8].

2.2 The Characteristic of Natural Text Image

Compared with printed documents, text in natural scene text images has the following image characteristics [9].

The contents of printed documents and handwritten text images are mainly composed of characters. In order to facilitate readers' reading, the layout structure is manually designed and adopts a fixed layout format, which can easily and effectively segment text areas. In the natural scene image, the text is not the main body of the image, but a supplement to the image content, such as store name, road sign, commodity name, subtitle, logo, sensitive vocabulary. The text area may appear anywhere in the image. This adds many uncertainties to text detection. In the end-to-end recognition system, how to detect text regions is one of the research focuses [10].

The background of printed document and handwritten text image is relatively simple with single color, and the noise often appears outside the text area, so the noise can be filtered through the image preprocessing method. In natural scene images, text is often embedded in complex background images. Characters and background pixels are often mixed together. The color and brightness of the background are variable, and this change is irregular. This leads to low positioning accuracy of text detection results, while text recognition requires high alignment of text areas. In the end-to-end recognition system, how to precisely locate the text area is one of the research focuses [11].

Characters in printed documents and handwritten text images obtained by photoelectric scanning equipment will not undergo geometric deformation. Due to the change of visual angle, the influence of perspective effect, and the defects of the lens in the acquisition process, the captured image will have geometric distortion, which will cause changes in the shape of the text in the image, increasing the difficulty of recognition [12].

Most natural scene images are color images, and the pixels on the foreground and background are multivalued. This leads to the inapplicability of the traditional text detection method based on pixel to separate foreground and background. At the same time, the text image quality of the natural scene images is poor due to the influence of light, shadow, other obstacles and outdoor weather [13].

To sum up, there are great differences between complex scene text and printed document text, which bring great difficulties to scene text recognition [14–16].

3 The Text Recognition Method

3.1 Traditional Text Recognition Method Based on Morphology

Text recognition based on segmentation is a traditional algorithm, which uses morphological methods for text classification and recognition. The text is extracted and separated from its background recognized. Binarization plays a key role in the recognition method.

Chen et al. reported an adaptive image binarization algorithm, which can automatically adjust the window length and width according to the size of the text in the image to achieve the purpose of character segmentation, and then use OCR technology for text recognition [3]. Worf et al. proposed the Niblack algorithm which processed the image by binarization and then recognized the text in the document [1, 2]. Lienhart et al. used color clustering for text segmentation, clustering those text pixels in the image with the same color and segmentation. During the development of text recognition, the focus of research has gradually shifted to word-level recognition methods [4–6]. Feild et al. used language model to change the traditional color clustering method, which effectively solved the limitations of text recognition based

on single-character recognition [10]. Fraz et al. proposed a two-sided regression probability statistical model, which solved the problem caused by color gradient in image text to a certain extent [9]. The text recognition accuracy has been significantly improved.

3.2 Text Recognition Method Based on Deep Learning

CTC was first proposed by Graves et al. [11]. Shi et al. proposed CRNN text recognition algorithm, which has now become a widely used network in text recognition. CTC is commonly used after the Recurrent Neural Network (RNN) to align its input and output sequences [12]. Jadcrbcrg et al. and DicNct et al. also promoted the development of CNN network models in the field of text recognition [14]. Le et al. proposed the use of convolutional neural networks (CNN) for text recognition [13]. Bahdanau et al. proposed the Attention mechanism, which can imitate human attention and screen out the highest valuable information from the mass of information [17]. This method solves the problem that it is difficult to obtain the final reasonable vector representation when the RNN model inputs long sequences. Lee et al. proposed the R2AM model [18]. Cheng et al. proposed the FAN network to solve the inaccurate problem of the attention mechanism. Local supervision information is added to FAN in the Attention module, which promotes the alignment of attention features with real tag sequences [19]. Shi et al. combined Spatial Transformer Network with text sequence recognition network based on attention mechanism to adjust the text recognition input to the standard rectangular form, but the accuracy of this method is not very high [20, 21]. Liu et al. proposed to layer the attention mechanism to correct the deformation of each character separately [22]. Anuj Sharma presented a method to recognize online handwritten Gurmukhi characters [23]. Amit Kumar Gupta analyzed of back propagation of neural network method in the string recognition [24]. Ouchtati proposed an off line system for the recognition of the Arabic handwritten words of the Algerian departments [25]. Lin Meng proposed a novel approach to recognize the inscriptions by template matching [26]. Jyothi provided the comparative analysis of various wavelet transforms to recognize ancient Grantha script [27].

The detection and recognition system based on deep learning, when trained with huge data, usually has a significant improvement compared with traditional methods. However, methods based on deep learning mainly have three problems: (1) they need a lot of training data. (2) Most of them deal with horizontal or near horizontal texts. (3) These algorithms require a lot of computation. These three shortcomings may limit the promotion and application of this kind of algorithms.

4 Development Prospect for Text Recognition

Due to the complexity of the scene text, there are still many problems to be solved and improved. Therefore, future research will focus on the following aspects:

(1) Position text lines with different orientations. The existing algorithms have poor robustness to text lines with slant angles, so this direction can be considered.
(2) Recognition of Chinese characters. The scene text recognition technology in this paper is currently only for strings consisting of Chinese and English characters or numbers in scene images, and there is no in-depth research on the recognition of Chinese characters.
(3) The optical system is an important condition that determines the effectiveness of image processing, and improving the quality of images is the first problem to be solved.
(4) With the rapid development of high-performance hardware and the popularity of neural network-specific Accelerated chips on mobile terminals in recent years, computational complexity is no longer the most central.

Acknowledgements This work was sponsored by Research and application of key technologies of intelligent robot process automation (Grant No. 1500/2022-72002B).

References

1. Wolf C, Jolion JM (2004) Extraction and recognition of artificial text in multimedia documents. Formal Patt Anal Appl 6(4):309–326
2. Wilkinson MHF, Schut F (1998) Digital image analysis of microbes: imaging, morphometry, fluorometry and motility techniques and applications. In: Digital image analysis of microbes: imaging, morphometry, fluorometry, and motility techniques and applications. Wiley, London
3. Chen X (2004) Detecting and reading text in natural scenes. In: Proceedings of CVPR, vol 2
4. Lienhart R, Wernicke A (2002) Localizing and segmenting text in images and videos. IEEE Trans Circuits Syst Video Technol 12(4):256–268
5. Mancas-Thillou C, Gosselin B (2006) Spatial and color spaces combination for natural scene text extraction. In: 2006 International conference on image processing. IEEE
6. Mancas-Thillou C, Gosselin B (2007) Color text extraction with selective metric-based clustering. Comput Vis Image Underst 107(1–2):97–107
7. Mishra A, Alahari K, Jawahar CV (2012) Scene text recognition using higher order language priors. In: British machine vision conference, Surrey, UK, pp 1–11
8. Nvikova T, Barinova O, Lempitsky V et al (2012) Large-lexicon attribute-consistent text recognition in natural images. In: european conference on computer vision, Florence, Italy, pp 752–765
9. Fraz M, Sarfraz MS, Edirisinghe EA (2015) Exploiting colour information for better scene text detection and recognition. Int J Doc Anal Recogn 18(2):153–167
10. Feild JL, Learned-Miller EG (2013) Improving open-vocabulary scene text recognition. In: International conference on document analysis & recognition. IEEE
11. Graves A, Fernández S, Gomez F (2006) Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. In: International conference on machine learning. ACM

12. Shi B, Bai X, Yao C (2017) An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. IEEE Trans Pattern Anal Mach Intell 39(11):2298–2304
13. Lange S, Riedmiller M, Voigtlander A (2012) Autonomous reinforcement learning on raw visual input data in a real world application. In: international joint conference on neural networks. IEEE
14. Jaderberg M, Simonyan K, Vedaldi A et al (2014) Deep structured output learning for unconstrained text recognition. Eprint Arxiv 24(6):603–611
15. He P, Huang W, Qiao Y et al (2015) Reading scene text in deep convolutional sequences. 2015.
16. Yin F, Wu YC, Zhang XY et al (2017) Scene text recognition with sliding convolutional character models. arXiv preprint ar Xiv:1709.01727
17. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. ar Xiv preprint ar Xiv:1409.0473
18. Lee C-Y, Osindero S (2016) Recursive recurrent nets with attention modeling for OCR in the wild. In: IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, pp 2231–2239
19. Cheng Z, Bai F, Xu Y et al (2017) Focusing attention: Towards accurate text recognition in natural images. In: Proceedings of the IEEE international conference on computer vision, Venice, Italy, pp 5076–5084
20. Shi B, Wang X, Lyu P et al (2016) Robust scene text recognition with automatic rectification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, pp 4168–4176
21. Baoguang S, Mingkun Y, Xinggang W et al (2018) ASTER: an attentional scene text recognizer with flexible rectification. IEEE Trans Patt Anal Mach Intell 1–1
22. Liu Z, Li Y, Ren F et al Squeezed text: a real-time scene text recognition by binary convolutional encoder-decoder network. In: Thirty-second AAAI conference on artificial intelligence, New Orleans, LA, USA, pp 7194–7201
23. Meng L, Fujikawa Y, Ochiai A, Izumi T, Yamazaki K (2016) Recognition of oracular bone inscriptions using template matching. Int J Computer Theory Eng 8(1):53–57
24. Sharma A, Kumar R, Sharma RK (2009) Recognizing online handwritten Gurmukhi characters using comparison of small line segments. Int J Computer Theory Eng 1(2):131–135
25. Jyothi RL, Abdul RM (2017) Comparative analysis of wavelet transforms in the recognition of ancient Grantha Script. Int J Computer Theory Eng 9(4):235–241
26. Hosozawa K, Wijaya RH, Linh TD, Seya H, Arai M, Maekawa T, Mizutani K (2018) Int J Computer Theory Eng 10(5):170–174
27. Ouchtati S, Redjimi M, Bedda M (2014) Int J Computer Theory Eng 6(2):129–134

RSA-ABE: A Hybrid Encryption Approach for Medical Privacy Data with Cloud Storage

Yiheng Sun and Chenxu Li

Abstract Increasing sensitive medical data raises medical privacy concerns. Unauthorized access endangers patients. We propose an approach using ciphertext-policy attribute-based encryption (CP-ABE) and RSA to enable secure and controlled access to medical data based on user-defined access policies. Our approach enables: (1) Patient-defined fine-grained access control policies; (2) Secure "one-to-many" sharing with authorized users; and (3) Encrypted policy and data transmission. We generate CP-ABE keys and use socket programming to enable patient-user communication. The patient defines an attribute-based access policy. CP-ABE encrypts medical data under this policy. RSA encrypts the public key for transmission to users. Users submit attributes; If users' attributes satisfy the policy, the ciphertext can be decrypted, authenticating the users. Results show the hybrid scheme achieves secure, controlled medical data sharing through patient-defined access policies. Patients need not know accessing users in advance. Only authorized users related to a patient's condition access data.

Keywords Access control · CP-ABE · Hybrid encryption · Medical privacy

1 Introduction

First of all, we give a brief introduction to this paper from three aspects: background and related work, our solution and solution and contributions.

1.1 Background and Related Work

Cloud storage, as the further development of distributed computing, are widely used through the advantages of fine-grained price and high scalability. It provides users

Y. Sun (✉) · C. Li
Zhengzhou University, Zhengzhou, China
e-mail: yh_xsyx@163.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_2

with online storage services available anytime and anywhere, so that users can store local data into the cloud server. Thus, it facilitates people's life to a large extent. At the same time, it also has problems with information leaks, illegal access and so on. If medical privacy data is obtained by some illegal elements for illegal activities, it may cause great trouble to patients' life and even endanger the life of the owner. Therefore, the security of medical privacy data must be inseparable from access control.

Attribute-based Encryption assigns certain attributes to each legitimate participant. According to the Attribute set of the participant, the data owner formulates an access policy and encrypts the data. Only the data visitor whose attributes meet the access policy can decrypt the data. So, It can effectively protect medical privacy data.

Attribute encryption is derived from identity encryption (IBE), which was first proposed by Shamir in 1984 [1]. In 2020, Zheng et al. [2] proposed an attribute-based data sharing scheme that supports efficient revocation of users, allowing users to join, revoke and re-join efficiently. In 2021, Gao et al. [3] combined blockchain, CP-ABE and IPFS to propose a blockchain-based personal data secure sharing and privacy protection solution. In 2021, Hijawi et al. [4] proposed a lightweight KP-ABE scheme. In 2022, Li et al. [5] introduced a white-box traceable CP-ABE scheme that can solve the problems of user and authorization center key abuse.

These Attribute-based Encryption schemes use ABE encryption and decryption locally. However, data security is not guaranteed in the process of two-terminal communication transmission. These hybrid encryption schemes do not implement access control. Therefore, in order to remedy this defect, we propose a hybrid encryption approach: RSA-ABE. In this scheme, we not only uses ABE to implement access control of medical privacy data, but also uses RSA to encrypt ABE keys. Because this way can enhance the security of two-terminal communication. And in two-terminal communication, the mpk used for encryption needs to be transmitted through the two-terminal communication, which ensures the security of the mpk during transmission. This is different from the local use of ABE encryption and decryption.

1.2 Our Solution

The purpose of this paper is to design an access control approach, which not only satisfies the requirement that a person can specify his own access policy, namely fine-grained access, but also satisfies the requirement that users with legal permissions can access the resources they have the right to access, while illegal users or malicious users cannot access the protected resources. Medical privacy data access control based on attribute encryption, patients do not need to know in advance which medical staff can view their medical data, and in order to enhance the security of patients' health privacy, only medical staff related to the patient's condition can access the patient's medical data, that is, in addition to protecting the security of patients' medical data, One-to-many data sharing and flexible access control are also required. The data owner can precisely control the data user who decrypts the Ciphertext by

embedding the developed access policy in the ciphertext. With fine-grained access control. The patient only needs to formulate access policies for encrypted medical data according to the attributes of the authorizer. If the attributes of the user meet the access policies defined by the patient, the corresponding medical data can be obtained by decrypting the ciphertext.

1.3 Contributions

- CP-ABE is used to control access to medical privacy data, meet the "one-to-many" data sharing and flexible access control, and protect data security.
- Implement user-driven authorized access, where users can define their own access control policies and follow their own privacy preferences.
- Compared with traditional medical data access control, attribute-based access control is fine-grained, flexible, adaptable to the cloud environment, and does not require users to manage keys online.

2 Hybrid Encryption System

This paper uses hybrid encryption system (CP-ABE and RSA) to realize access control and protect data security. The process of hybrid encryption as Fig. 1.

Example: As shown in Fig. 1, Alice generates ABE keys (public parameters mpk and master key msk), then uses RSA to encrypt mpk to s_mpk. This ensures the security of mpk during transmission. Then, when Alice and Bob establish a

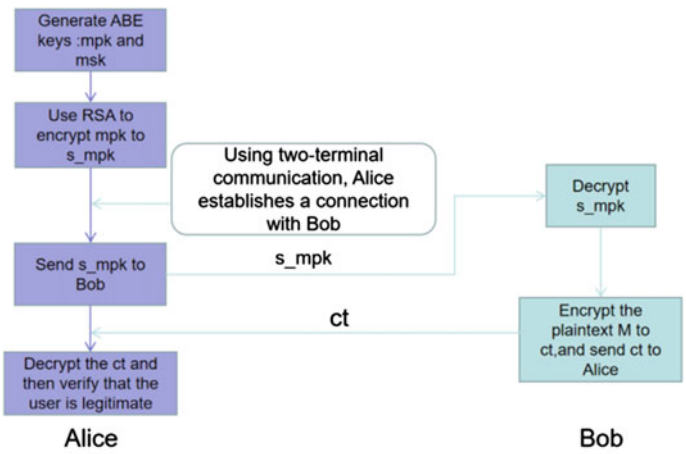


Fig. 1 The process of hybrid encryption

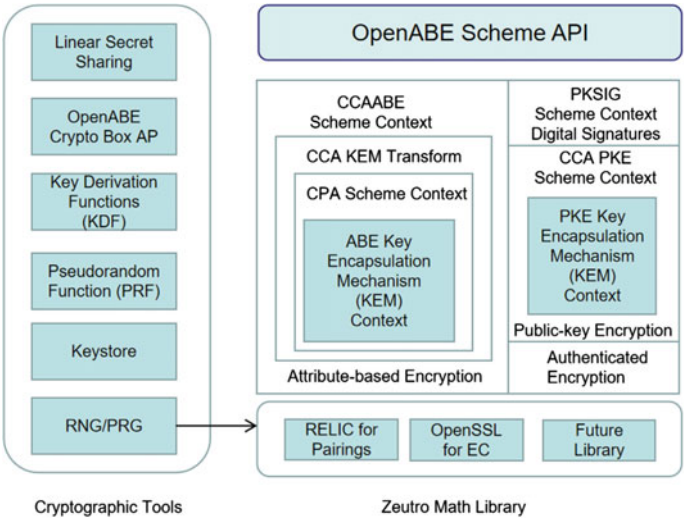


Fig. 2 OpenABE architecture diagram

connection using two-terminal communication, Alice sends s_mpk to Bob. After receiving s_mpk, Bob decrypt it. Moreover, Bob encrypts the plaintext M to CT, and send CT to Alice. After Alice receives the CT, she decrypts it. And then, Alice can verify the legitimacy of the user.

2.1 OpenABE

This article uses the attribute encryption and decryption library OpenABE. The architecture diagram is as Fig. 2.

In this paper, CP-ABE in OpenABE is used. The encryption and decryption process is as Fig. 3. After initializing the OpenABE library by constructing the Crypto Box context and generating domain parameters, you can perform key generation by specifying attributes, attribute lists, and access policies, then encrypt messages under a chosen access policy with the public key, and authorized users can decrypt the ciphertexts and recover the original messages using their private keys.

2.2 Two-Terminal Communication

As illustrated in Fig. 4, to achieve two-terminal communication between the Server and Client, two dedicated and non-interfering threads, one handling input and the other output, need to be implemented on both sides.

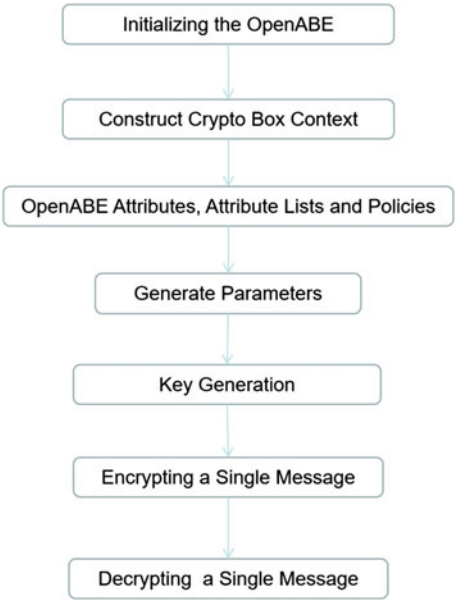


Fig. 3 The encryption and decryption process

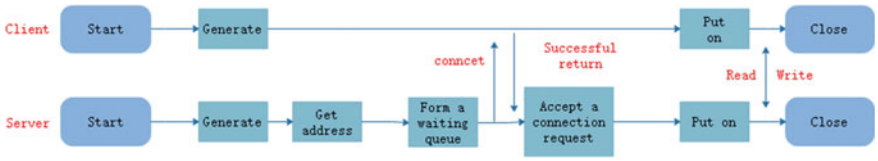


Fig. 4 The two-terminal communication process

3 Experiment and Results

This article adopts RSA-ABE: A Hybrid encryption Approach for Medical Privacy Data with Cloud Storage. This article not only uses ABE to implement access control of medical privacy data, but also uses RSA to encrypt ABE keys. Because this way can enhance the security of two-terminal communication. And in two-terminal communication, the mpk used for encryption needs to be transmitted through the two-terminal communication, which ensures the security of the mpk during transmission.

We use CP-ABE to realize access control, allowing patients to define the access control structure by themselves, and specifying people to access their medical privacy data, so that the control is fine-grained and flexible. In the process of public key transmission, the asymmetric key RSA is used for encryption to realize data confidentiality. One end defines the access control structure, and the other end uses attributes to decrypt access, which satisfies the characteristics of cloud storage.

```
CP-ABE key
mpk = AAAAFqpvyT/mYIB25dlWrNRvbe9L23ltcGsAAAHToQFBsgEEtLIBAANBC9Ko6QWZZ5qUvUbUvQ
tlFW+aBvVfOi9xdFe04gpIAd0x+o65uTTatNk4mIt5DN7EvYX6WeZ27Sp7v92e7EgdikyGk+CXK7j4Un
guaKk96xu+mLbT/LSSWnnSrcVvuQBKH7+3fhCab3wSDd4m66jy84KtgeN2F2jO0UwLfIp+ByfoRs7xBA
mpXhKZPqv/vfGrshCFT3UeSNeWUaskR2UMz3Z+iqvMjggdSmcy6I6+YJ9m4mjGtHbI4KDRfEEhMCGV+J
+2bJi/F9m0KT0rCj+GmY7A3u9pTDJ0PXywn6psBMKJiKoMY4lUs6x984sCfalijtW7LXxaWqRheIoRmP
KhAmcxoSSyoSEDCJ8DuANsYA0BVendQ+YOPs2fvu26oEIeNnGlprmaWiKhA2cxYaEksqEhAwiHbbpgVp
iOakaKS6gUiGhEqastgmDiHFiIDFte87GRoQJnMqFEs6FBAhKDU6W4dLXcmzu4KnEmHT/sExppqd5w7W
9Sh7igO1wiG5v6WMQJliRehnyKj+0PmBkp0wVyztkc32Rjxo8GWz6hAWuhJR0AAAAgywTxRr1gTii7Rx
e10Ytd2avi7DkMUaPdIQNbGffzk78=
```

Fig. 5 Mpk

1. First, generate mpk and msk at the A terminal. As shown in Fig. 5.
2. Then, use RSA to encrypt mpk on end A.
3. The connection between A terminal and B terminal is established. As shown in Fig. 6.
4. The encrypted data is sent to B terminal. As shown in Fig. 7.
5. After receiving the encrypted mpk, B terminal decrypts the mpk using RSA. Then, use CP-ABE to encrypt data. As shown in Fig. 8.
6. The encrypted data is sent to A terminal. As shown in Fig. 9.
7. A terminal decrypts and verifies whether the user is legitimate. The result is shown in Fig. 10.

In this experiment, first, the public and private keys of CP-ABE are randomly generated, and then stocket is used to establish the two-ended communication. After

```
server_fd = 3
lo ipv4 address = 127.0.0.1
ens33 ipv4 address = 192.168.22.136
ens34 ipv4 address = 192.168.22.138
ens35 ipv4 address = 192.168.22.137
lo ipv6 address ::1
ens33 ipv6 address fe80::20c:29ff:fe61:4e09
ens34 ipv6 address fe80::9f1e:a3:d4a8:78d5
ens35 ipv6 address fe80::8aa6:8166:d961:10f5
Bind success!
Server port number = 34316
Server ipv4 addr: 192.168.22.137
Listening ...
Client accepted: IP addr = 192.168.22.137, port = 41794
```

```
Input server's hostname/ipv4:  192.168.22.137
Input server's port number: 34316
server's official name = 192.168.22.137
        server address = 192.168.22.137
local socket ip addr = 192.168.22.137, port = 41794
```

Fig. 6 Establish a two-terminal connection

```
                                        server 192.168.22.137 >>>> nb7Jq3nehmXRa
gkhKg2ybiAxDgX4v19XQxZsAwj+ft8cdIDH66D51mjOf+kSrrbXSapNJ/ENp18Gb0wrQ1p0Z0WIUsWND
jhpCJSEEH8jCIpV1sXbbCWb6it8+4DaNAweNWdM1vCtmqfrTfXlI/K70z2EbvBa0uS0ViWm8HLJxQ3Wl
GeB3pso4EknfuZyHkngRq8IIFMUIkeqzvUtT8Tln/USlVEe7dKXhILlQqo0jvY/v9je13gz0DonL08Jj
ONmCmrAGrf4S4Gz2exIAUIcl2O6PjrB+/HzMjoM0/MPuWIKgjeRA8I5wflIXi2pmJ7Vlo7OYbg7WuGpp
N34RDOOeJW0tfXcC/DnuLOstS0N7Uk7MvXWWWm8LdYP9zLP3qKDqE8pGgcdNsA7ngQ6kT8jRDgh3BVme
6jzQhKtqRH+JWq8/vltmlZ389FgF3ghMyk4AwqdADFdSQgRqqvFIvJshED/HSw6buhXamDnBoDF7Wjv6
m//pFB/1AieR5l84ajzchHVmFsX/zVqQWSHM+2FZmH9fPU9N2jNE1C5iSyP9FcqWInf9gtkdstgd3B/l
uwpyv4WS/zlAiBHkjhYl5sjoExhCfBgPFz1N+vx5HzoWiioP7kQSX0js3rEthHg5Uv9UvFe4E1b1AByI
WSJXT7g+6h3Lg/IPOxG6Js3Y2rIqsh6pcTDP5FCb+gw4G2bAobyyX+1z7K8gnDWgmrX/cgQL6jQN1Beb
vSv6lpVJFsa3R0ViXBaV8JeaqVXd2xIdYgF6bPHbRR533s70wgqDOefGkSYG7iM3nnoPzuGf/gwXReeA
zGUUdyKY623/Mbwo71d2wLWFopo5JAALsxfUg1kGGGLDd2WNpmvFTqbAYeimyM9wV7IumaS+689hYyO/
XDNZIbLG9nADkMrM0XwJk3AG3fr96KDifbvqKDi7O3JAJRkFMZKz0lhnPODjk/BYkN5vYigbBwya44QW
zoucNOUJstm621I/JAihHjIO0vaQAgG4MQXR4gcHFbVbWifN1q
```

Fig. 7 Send to B terminal

Fig. 8 Encrypt data at the B terminal

Fig. 9 Send to A terminal

Fig. 10 Verify identity

encrypting the public key using RSA, it is sent to the patient using two-ended communication. The patient himself define access structure (id4254111988342 and man and age30 and day444 and chronicrespiratoryinfections).Obviously, it must have all the above properties to meet the access structure. Then, the patient encrypts the data using the access structure and the transmitted key. Encrypted data is transmitted to the end that Users need access to patients' medical data. When a user access to the patient's medical data, submit his own properties (| id4254111988342 | man | age30 | day444 | chronicrespiratoryinfections).Obviously, the set of properties meet the

patients themselves defined access structure, and the ciphertext can be decrypted, so the authentication passed (User qualification!).

Our hybrid encryption scheme can achieve the following goals:

First of all, patients can define their own access structure, and only the person designated by the patient can access the patient's medical data, enhancing the security of the patient's medical privacy.

Secondly, a patient's medical data can be accessed by multiple users, satisfying the "one-to-many" data sharing and flexible access control.

Finally: RSA is used to encrypt data during transmission to protect data security.

4 Conclusion

In this paper, to protect the privacy of medical data, we propose the RSA-ABE: a hybrid encryption approach for medical privacy data with cloud storage. In our construction we employ two-terminal communication and CP-ABE. The user can customize the access structure on one end. At the other end, Identity authentication can be performed through the transmitted data and attribute entered by a user for an access request. And RSA encryption is used during data transmission to protect data security. In the cloud storage environment, one end defines the access control structure, and the other end uses attributes to decrypt access, meeting the cloud storage characteristics. Therefore, this solution is feasible. In the future, this scheme can be applied to various industries, such as transportation, education, power, etc., and CP-ABE can also be improved to realize hierarchical control and encrypt transmitted data with other encryption algorithms.

References

1. Shamir A (1984) Identity-based cryptosystems and signature schemes. Adv Cryptol 21(2):47–53
2. Zheng D, Qin BD, Li YN et al (2020) Cloud-assisted attribute-based data sharing with efficient user revocation in the internet of things. IEEE Wirel Commun 27(3):18–23
3. Gao H, Ma Z, Luo S et al (2021) BSSPD: a blockchain-based security sharing scheme for personal data with fine-grained access control. Wireless Commun Mob Comput 1–20.
4. Hijawi U, Unal D, Hamila R et al (2021) Lightweight KPABEarchitecture enabled in mesh networked resource—constrained IoT devices. IEEE Access 9:5640–5650
5. Li JG, Zhang YC, Ning JT et al (2022) Attribute based encryption with privacy protection and accountability for CloudIoT. IEEE Trans Cloud Comput 10(2):762–773

Downlink Beamforming Technique for MU-MIMO-OFDM Systems

Roopesh Kumar Polaganga

Abstract Multiuser—Multi carrier systems like MU-MIMO-OFDM Systems has been widely known for offering maximum spatial diversity along with maximum spectral efficiency which is of high importance in wireless systems. We have analyzed the performance of equalization techniques like Maximum Ratio Combining (MRC) and Minimum Mean Square Error (MMSE) in this multi-carrier system and it is inferred that MMSE is the optimal equalization technique for MIMO-OFDM System. Beamforming and combining strategies are extensively used to harness spatial diversity gains in MU-MIMO-OFDM systems. This work also introduced the beamforming technique at the transmitter with STBC coding to further improve the performance of the system.

Keywords MU-MIMO-OFDM · Equalization · MRC · MMSE · Beamforming · STBC · ISI · CCI

1 Introduction

The growth of users and their demand has been increasing enormously in wireless communication and so the spectral efficiency has given the prime importance. Multi carrier system like Orthogonal Frequency Division Multiplexing (OFDM) solves this problem to a great extent by offering a very good spectral efficiency along with multiple access and interference rejection capabilities along with multipath propagation robustness. Its basic principle is the effective utilization of the orthogonal property of the signals. This reduces the Inter Symbol Interference (ISI) to a great extent which is the main constraint in achieving higher data rates. In OFDM transmission, the signal on a specific subcarrier is represented by their complex amplitudes and the channels are described by their transfer functions [1, 2].

Reliability is another important aspect in achieving higher data rates. This can be attained easily by employing technique like multiple-input multiple-output (MIMO).

R. K. Polaganga (✉)
The University of Texas at Arlington, Arlington, TX 76010, USA
e-mail: RoopeshKumar.Polaganga@Mavs.Uta.edu

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_3

This involves sending of same user information from more than one transmitting antennas to the receiving antennas to increase spatial diversity. Without additional bandwidth or increased transmit power, MIMO can significantly increase the data throughput and link range which in turn helps in realizing good error performance [3].

In a multiple carrier system, the problem of frequency selective fading is implicitly present which degrades the function of MIMO System [3]. This problem can be easily tackled with the help of OFDM in which any frequency-selective channel can be converted into parallel flat-fading channel which can mitigate the complexity [4–6]. Therefore, MIMO techniques with OFDM can be effectively used in non-flat fading channels.

MIMO techniques are classified into two groups: space time coding (STC) and space division multiplexing (SDM). STC increases the system's performance by coding over different transmitter branches, whereas SDM achieves a higher throughput by transmitting independent data streams on the branches simultaneously at the same carrier frequency [7]. Under STC scheme, there are two types: STBC-OFDM and SFBC-OFDM. For this work, SDM-OFDM scheme has been considered. As it requires a series of pseudo-inverse calculations of a channel matrix, its complexity is very high. Despite many fast and reduced-complexity algorithms proposed for SDM-OFDM systems, the hardware complexity is still unacceptable for handsets, in which a low hardware complexity design is an important issue [8, 9]. To obtain better performance, STBC-OFDM is also considered and implemented.

For a single user in proposed system, the corresponding base station communicates with several first-tier co-channel interfering users at same frequency and time slots which will impart co-channel interference (CCI) [10]. Incorporation of beamforming into the system will increase capacity and decrease receiver complexity [11]. To reduce CCI, it is evident to use Transmit Beamforming or Receiver Decoder. In terms of cost, transmitter optimization is preferred over receiver complexity. The spectral efficiency and error performance of a wireless system can be further improved if channel state information (CSI) is made available at the transmitter to certain extent.

Several amplitude and phase equalization techniques have been proposed for Multi Carrier systems, namely orthogonal restoring combining (ORC), equal gain combining (EGC), maximal ratio combining (MRC) and minimum mean square error (MMSE) [12]. MMSE per carrier performs better compared to all other schemes mentioned above [13]; however, it is also the most computationally complex to realize as it involves the matrix inversion operation of a large complex matrix. But in literature, many complexity reduction methods are proposed for MMSE channel estimator and DSCDMA detector [12]. This proposed scheme, however, is specific to the MMSE combiner/equalizer for MIMO-OFDM downlink.

Among the several methods proposed in transmit diversity; Alamouti space–time block coding is less complex with no CSI feedback requirement [14] and which can provide full rate and full diversity for any signal constellations [9]. The steps involved in STBC coding include the baseband modulated symbols passed through serial-to-parallel (S/P) converter which generates complex vectors which in turn

passed through the STBC encoder. This generates different sequences which are then given to their respective IFFT blocks and subsequently to multiple antennas [5].

2 System Model

For this work, downlink MU-MIMO-OFDM system (as shown in Fig. 1) has been considered with a base station equipped with M transmit antennas and with p subcarriers as mentioned in [10]. There are K geographically dispersed first tier co-channel users, each equipped with R receive antennas (same as the number of transmit antennas). The signal s(k) is then multiplied by a beamformer weight vector $w_u(k)$, where $w_u(k)$ is the beamforming vector for user u for the kth tone. Hence, the M × 1 signal vector for the kth tone is given by

$$X_{1:M,k} = \sum_{u=1}^{K} w_u(k) s_u(k) \tag{1}$$

Each column vector of X of size [1 × (N_{used} * N_{frame})] is the data vector to be transmitted over the m^{th} transmit antenna given by $X_m(k)$ (k = 0,1,2, … N − 1). Before being transmitted, the data vector is modulated by an Inverse Fast Fourier

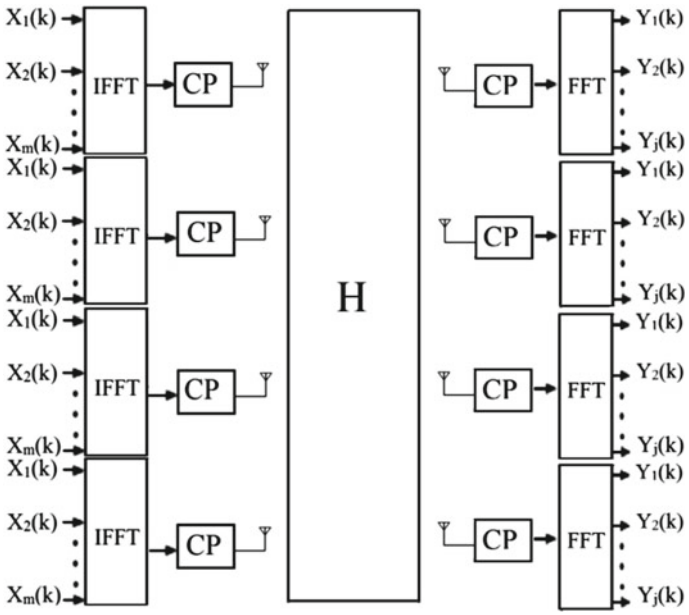


Fig. 1 MIMO-OFDM System model block diagram

Transform (IFFT) into an OFDM symbol vector $x_m(n)$. Then a cyclic prefix (CP) of length G is appended to $x_m(n)$ and the resultant vector is $x_m^g(n)$. This operation may be written as

$$\text{IFFT} : x_m(n) = (1/N) \sum_{k=1}^{N} X_m(k) \omega_N^{-(n-1)(k-1)} \tag{2}$$

where,

$\omega_N = e^{(-2\pi i)/N}$ is an Nth root of unity.

CP Inclusion:

$$x_m^g(n) = [x_{p-G}(n) \ldots x_{p-1}(n) x_0(n) \ldots x_{p-1}(n)]^T \tag{3}$$

The signal vector $x_m^g(n)$ is then transmitted through the m antennas over a frequency selective multi-user channel. To avoid inter-block interference (IBI), the guard interval is chosen to satisfy G = Nfft/4. Assuming that the channel impulse response is invariant during the entire block interval, the signal received at the jth antenna is given by

$$y_j^g(n) = \sum_{m=1}^{M} x_m^g(n) * H_{m,j}(p) + awgn_j(n) \tag{4}$$

where, '*' denotes convolution and $awgn_j(n)$ is additive white Gaussian noise (AWGN).

At the receiver, the CP is first removed and then an N-point discrete Fourier transform (DFT) is performed to yield the demodulated signal vector $Y_j(k)$. This operation may be written as

CP removal:

$$y_j(n) = [y_G^g(n) \ldots y_{G+1}^g(n) \ldots y_{p+G-1}^g(n)]^T \tag{5}$$

FFT:

$$Y_j(k) = \sum_{n=1}^{N} y_j(n) \omega_N^{(n-1)(k-1)} \tag{6}$$

The channel matrix $H_u(k)$ represents the frequency response of the channel for user u for the kth tone. Assuming the uth user employs R antennas, the R * M channel matrix for the kth tone can be written as:

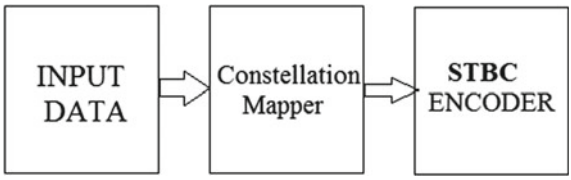


Fig. 2 STBC encoder block diagram

$$H_u(k) = \begin{pmatrix} h_u(k)^{(1,1)} & \cdots & h_u(k)^{(M,1)} \\ \vdots & \ddots & \vdots \\ h_u(k)^{(1,N)} & \cdots & h_u(k)^{(M,N)} \end{pmatrix} \tag{7}$$

where, $h_u(k)^{m,j}$ denote the channel gain between the mth transmit and jth receive antennas, for user u. Here, we assume that the receiver for user u has access to accurate CSI, H_u. However, given the reasons mentioned in earlier sections, system is allowed with imperfect CSI at the transmitter.

2.1 STBC Encoder

The STBC encoding system can be represented with Fig. 2. The input data is taken, and it is mapped using constellation mapper and then it is encoded. Before it is given to the STBC encoder, it is to be converted form serial-to-parallel. The output of STBC coder is in turn given to the IFFT after which it is converted from parallel-to-serial before transmitting. The reverse order is followed in the receiver side with STBC decoder.

3 Coded Beamforming

Basic OFDM System is realized, and error performance is obtained. MRC and MMSE techniques are implemented in MIMO-OFDM System, and the error performances are compared with the same channel conditions and parameters. MU-MIMO-OFDM system is realized, and Beamforming is incorporated in MU-MIMO-OFDM System along with MMSE to avoid further interference and subsequently to improve the performance [15, 16].

Among several coding techniques, STBC is the suitable technique for MIMO Systems. So, STBC coding technique is finally incorporated in the realized MU-MIMO-OFDM System (Beamformed) and it is compared with the non-coded system to obtain desired results.

4 Contributions

Since no effective equalization techniques have been proposed in specific to MIMO-OFDM systems, from [13] among different equalization techniques to combine signals from multiple paths, MMSE is the best Equalization technique for Multi-carrier System (MC-CDMA). In MIMO-OFDM System, MRC and MMSE equalization techniques were considered in [10] and [17] respectively. So, in this work, the performance of MMSE with MRC has been compared and further concluded that MMSE is comparatively better than MRC in MIMO-OFDM systems.

Transmit beamforming with STBC coding has been implemented and proposed that the above considered system (MU-MIMO-OFDM) performs better than other coding techniques. Simulation results also show that this transmitter-beamformed and coded MU-MIMO-OFDM system gives better performance than the non-coded system. Associated implementation complexity observations are also mentioned accordingly.

5 Simulation Results

Simulation is performed using MATLAB R2022b with BPSK modulation and the simulation parameters are shown in Table 1 below.

Figure 3 shown below compares the performance of single user MIMO-OFDM system with MRC equalization at the receiver for single antenna and multiple antenna cases. This shows that MIMO improves performance with the help of equalization/combining at the receiver.

Result 4 below shows the performance of single user MIMO-OFDM system with MMSE equalization at the receiver for multiple antenna system (2 antennas). This shows that MMSE equalization performs better than MRC for multi carrier system. However, there could be some computational complexity introduced which can be further explored in future studies (Fig. 4).

Then a MU-MIMO-OFDM system is considered with 6 users located in co-channels in a 7-cell environment. Each user and its corresponding BFs' are equipped

Table 1 Simulation parameters

Number of subcarriers	16
Channel	AWGN and Rayleigh fading
Modulation	BPSK
Antennas	$2 \times 2, 4 \times 4$
Equalization techniques	MRC, MMSE
Beamforming	Transmitter and receiver
Number of users	7
Coding technique	STBC

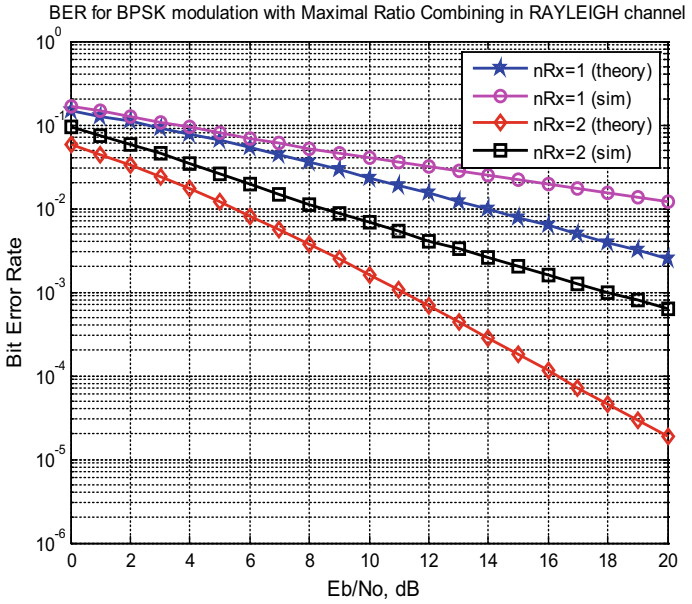


Fig. 3 BER Comparison of SISO and MIMO OFDM Systems with MRC equalization technique

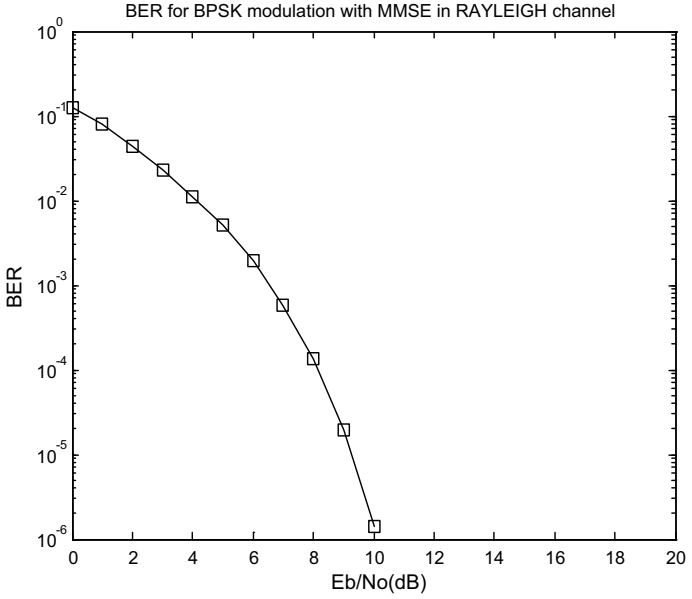


Fig. 4 BER plot for MIMO-OFDM System with MMSE Equalization

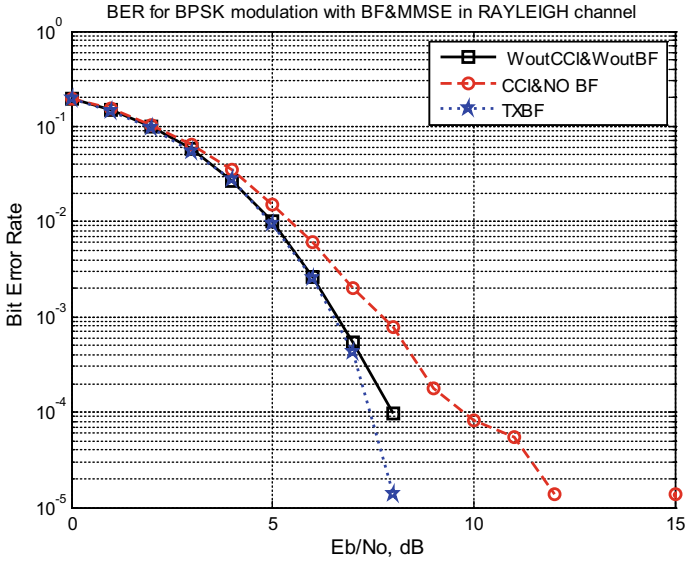


Fig. 5 BER plot for User-1 in MU-MIMO-OFDM System

with equal number of antennas and transmit beamforming is incorporated in the realized system. The following simulation results show that transmit beamforming gives optimum performance compared to the system without beamforming (Figs. 5, 6, 7, 8, 9, 10 and 11).

Finally, the same MU-MIMO-OFDM system is considered with 6 users located in co-channels in a 7-cell environment. Each user and its corresponding BSs' are equipped with equal number of antennas (4 antennas) and transmit beamforming is included. The following simulation results show that STBC coding with transmit beamforming gives better performance when compared to the system with transmit beamforming alone (as shown in Fig. 12). Note that the remaining users BER plot behaves in the same manner.

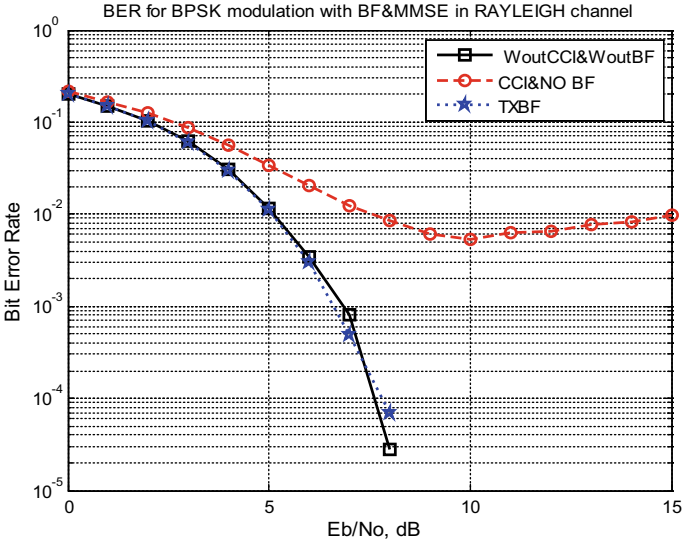


Fig. 6 BER plot for User-2 in MU-MIMO-OFDM system

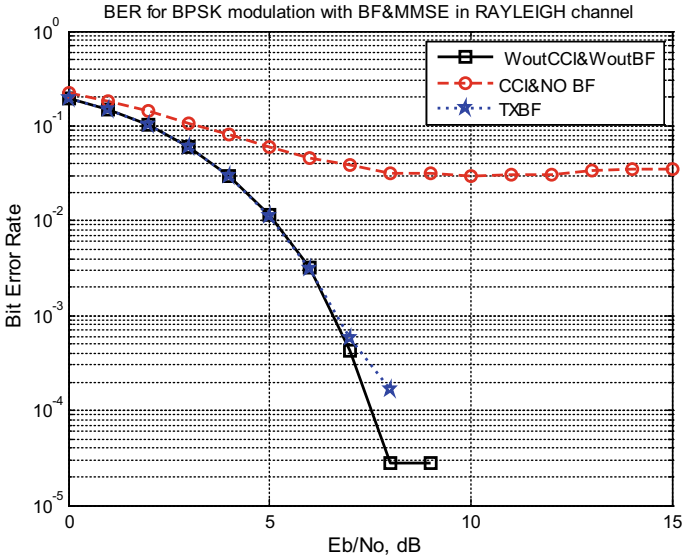


Fig. 7 BER plot for User-3 in MU-MIMO-OFDM system

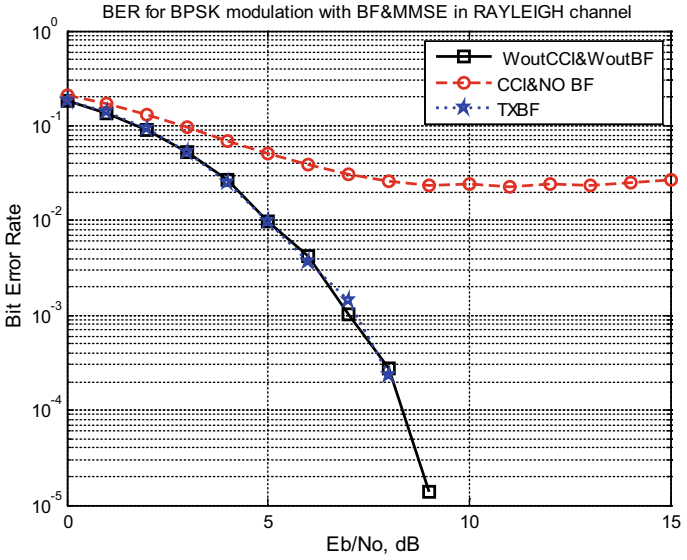


Fig. 8 BER plot for User-4 in MU-MIMO-OFDM system

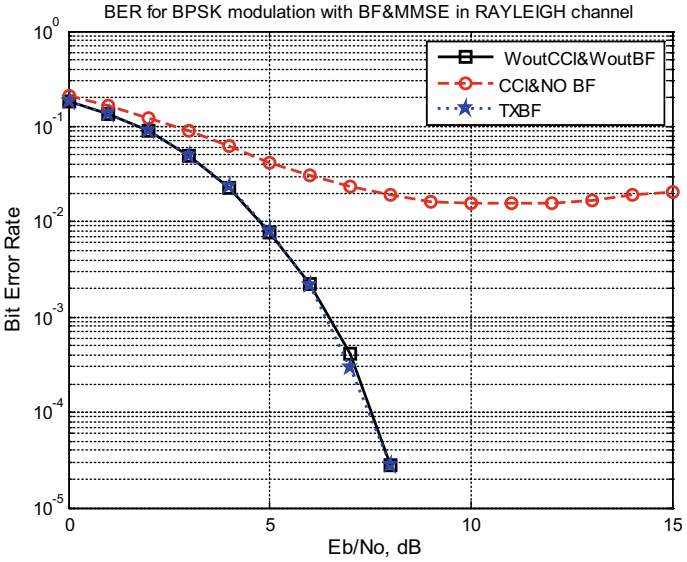


Fig. 9 BER plot for User-5 in MU-MIMO-OFDM system

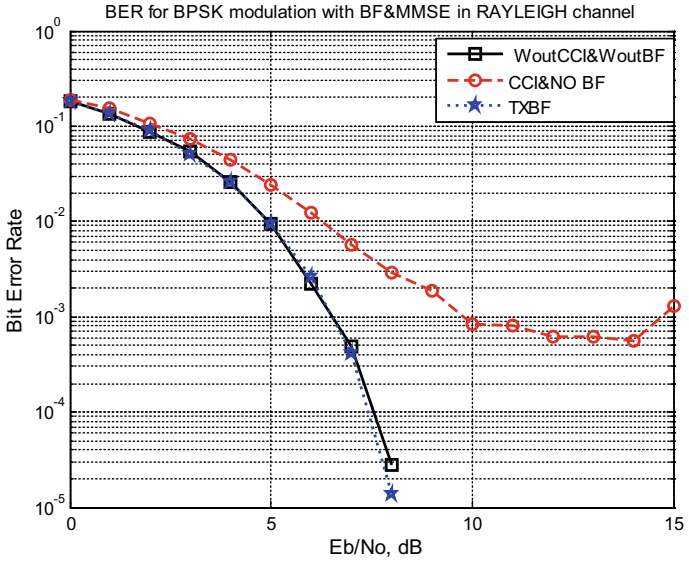


Fig. 10 BER plot for User-6 in MU-MIMO-OFDM system

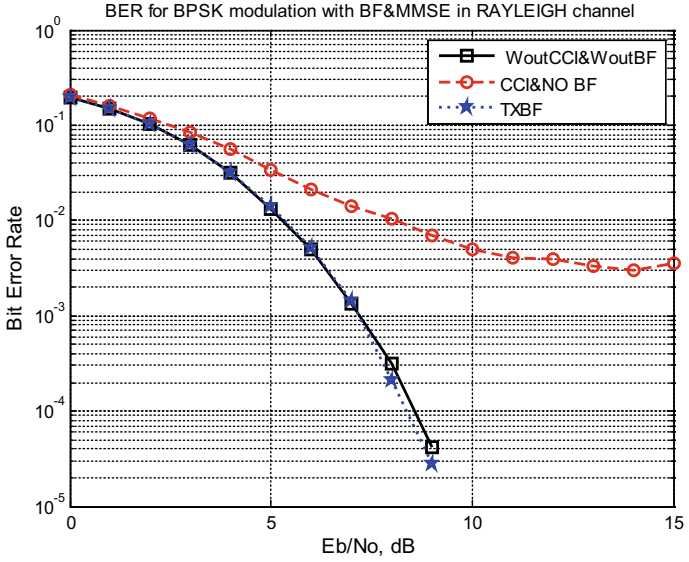


Fig. 11 BER plot for User-7 in MU-MIMO-OFDM system

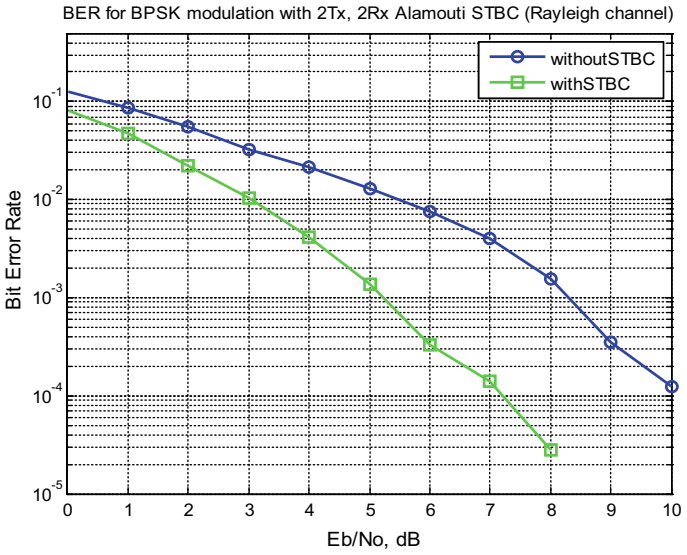


Fig. 12 BER comparison for a user in MU-MIMO-OFDM system with and without STBC coding

6 Conclusion

Simulation results infer that inclusion of MIMO in the existing OFDM system with an equalization technique at the receiver reduces the BER substantially and thus better reception is possible with high reliability. Apart from the inclusion of MIMO, proper selection of the equalization technique (MMSE instead of MRC) will provide better BER performance. In MMSE equalization technique, the predicted channel is inversed to nullify the effects of the channel in the received data which reduces the Bit Error Rate reasonably. But inverse matrix operations are complex in nature which increases the receiver complexity while decoding the received signal which in turn increases the cost of the end terminal which is not desirable. In literature, many complexity reduction algorithms are present to reduce the computations involved and moreover with the advent of VLSI chips these operations can be easily performed. Over and above the MIMO inclusion, beamforming at the transmitter along with STBC coding will give still better performance. But the bandwidth should be allocated for these redundant data which may reduce the carrier spacing of a particular user.

References

1. Gherekhloo S, Ardah K, Haardt M (2020) Hybrid beamforming design for downlink MU-MIMO-OFDM millimeter-wave systems. In: 2020 IEEE 11th sensor array and multichannel signal processing workshop (SAM), Hangzhou, China, pp 1–5. https://doi.org/10.1109/SAM48682.2020.9104405
2. Polaganga R, Liang Q (2015) Self-similarity and modeling of LTE/LTE-A data traffic. Measurement 75. https://doi.org/10.1016/j.measurement.2015.07.051
3. Kong N et al (2018) A downlink multiuser MIMO OFDM decomposition method in frequency selective fading channels. In: 2018 IEEE green energy and smart systems conference (IGESSC), Long Beach, CA, USA, pp 1–6. https://doi.org/10.1109/IGESC.2018.8745519
4. Kim YG, Beaulieu NC (2010) On MIMO beamforming systems using quantized feedback. IEEE Trans Commun 58(3):820–827
5. Lee S, Noh S, Jung Y, Kim J (2007) Efficient design of symbol detector for MIMO-OFDM based wireless LANs. In: IEEE International conference on telecommunications and Malaysia international conference on communications, Penang, Malaysia, 14–17 May 2007, pp 510–514
6. Wajid I, Gershman AB, Vorobyov SA, Karanouh YA (2007) Robust multi-antenna broadcasting with imperfect channel state information. In: IEEE International workshop on computational advances in multi-sensor adaptive processing, pp 213–216
7. Murakami T, Ujihara K, Takatori Y, Mizoguchi M, Maehara F (2016) Downlink multiuser MIMO-OFDM transmission using simple receive antenna selection. In: 2016 International symposium on antennas and propagation (ISAP), Okinawa, Japan, pp 604–605
8. Mizutani K, Matsumura T, Harada H (2017) A comprehensive study of universal time-domain windowed OFDM-based LTE downlink system. In: 2017 20th International symposium on wireless personal multimedia communications (WPMC), Bali, Indonesia, pp 28–34. https://doi.org/10.1109/WPMC.2017.8301823
9. Baing I, Jeoti V, Drieberg M (2011) A ZCMT precoding based STBC MIMO-OFDM system with reduced PARP. In: National postgraduate conference, Malaysia, 19–20 Sept 2011, pp 1–5
10. Sharma V, Lambotharan S (2006) Robust multiuser beamformers in MIMO-OFDM systems. In: 10th IEEE Singapore international conference on communication systems (ICCS), pp 1–5
11. Shen C, Fitz MP (2008) MIMO-OFDM beamforming for improved channel estimation. IEEE J Sel Areas Commun 26(6):948–959
12. Zhang K, Guan YL, Shi Q (2008) Complexity reduction for MC-CDMA with MMSE. IEEE Trans Veh Technol 57(3):1–4
13. Tamilarasan N, Nithyanandan L (2011) Performance of the MIMO-MC-CDMA system with MMSE equalization. Int J Recent Trends Eng Technol 6(2):223–226
14. Ahmed T, Mortuza Md., Ullah S (2012) BER performance analysis of a STBC encoded secured multiuser MIMO-OFDM wireless communication systems. Int J Hybrid Inform Technol 5(4):19–29
15. Nonaka N, Muraoka K, Suyama S, Okumura Y (2019) Two-step user selection algorithm for multi-user massive MIMO with hybrid beamforming. In: 2019 22nd International symposium on wireless personal multimedia communications (WPMC), Lisbon, Portugal, pp 1–5. https://doi.org/10.1109/WPMC48795.2019.9096171
16. Liu Y (2018) A modified approach on MMSE MIMO interference cancellation detection of free space optical communication. In: 2018 10th International conference on measuring technology and mechatronics automation (ICMTMA), Changsha, China, pp 178–181. https://doi.org/10.1109/ICMTMA.2018.00050
17. Vía J, Santamaría I, Elvira V, Eickhoff R (2010) A general criterion for analog Tx-Rx beamforming under OFDM transmissions. IEEE Trans Signal Process 58(4):2155–2167

CNN Based Resource Management for D2D Networks with Wireless Networks Virtualization

Kun Yang and Youyun Xu

Abstract The energy consumption of the system with wireless network virtualization will increase as the number of slices increase. Meanwhile, device-to-device (D2D) communication can reduce the load of base stations and improve spectrum efficiency, which applied in wireless virtualized networks can improve the system's energy efficiency. However, D2D communication can cause interference to cellular users, how to ensure the communication quality of different users while keeping the system energy consumption lower is noteworthy. In this paper, the problem of wireless resource virtualization allocation with D2D communication is formulated. We formalized the channel allocation and power control problem with the goal of system energy efficiency. We solve this problem at a lower computational complexity with convolutional neural network (CNN). Results show that faster speedups are obtained with lower losses compared to optimal results by the proposed scheme.

Keywords Wireless network virtualization · Device-to-device · Energy efficiency · Resource allocation · CNN

1 Introduction

In recent years, with the development of mobile smart terminals and wearable devices, the demand for large-scale access and system energy efficiency is increasing, so the spectrum allocation and energy efficiency issues in the fifth generation technologies have received extensive attention. Wireless network virtualization, as an emerging technology, optimizes spectrum allocation to greatly improves system performance

K. Yang · Y. Xu (✉)
School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
e-mail: yyxu@njupt.edu.cn

K. Yang
e-mail: 1021010313@njupt.edu.cn

National Engineering Research Center of Communications and Networking, Nanjing 210003, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_4

[1]. Meanwhile, D2D communication technology reduces the load on base stations and allows direct communication, which facilitates resource allocation and system energy efficiency optimization [2, 3]. In addition, artificial intelligence can make a quick and correct decision on the corresponding policy after a small amount of training time, which is used for traditional resource management in wireless networks with ultra-low time complexity.

A significant amount of publications have investigated the field of 5G network function virtualization, such as high-level architecture, isolation, mobility management, and resource allocation. In [4], the latest technologies of NFV are studied to design novel network frameworks. An approach based on genetic algorithm is proposed for efficient mapping of virtualized nodes in [5]. Reference [6] proposes a joint resource allocation algorithm for subcarrier allocation and power control to maximize the network rate. However, there are fewer papers that use the energy efficiency of the system as a metric. In addition, a lot of research have been conducted on resource allocation based on D2D communication in 5G technology. In the literature [7], an iterative power allocation algorithm is proposed to maximize the system energy consumption in D2D communication systems. Reference [8] maximizes the energy consumption in D2D communication by proposing a low-complexity and suboptimal algorithm. Reference [9] solves the power allocation problem with Dinkelbach method and Lagrange dual decomposition, and employs Q-learning to solve the relay selection problem.

Recently, many novel studies are related to D2D communication in virtualization scenarios, where the literature [10] maximizes the sum rate of D2D users while guaranteeing the performance of cellular users. Conference [11] builds on this by further considering the cost of wireless network providers. Reference [12] considers both the system rate and the delay as well as the requirements of different services. However, the system energy efficiency is also an worthy issue to further investigation under the complex network functions of 5G.

In this paper, we combine the above two possible scenarios and design a wireless resource virtualization problem to improve the energy efficiency based on D2D communication. We allocate the virtualized spectrum resources to the CUs of the system and use D2D pairs to share the spectrum resources of the CUs. The optimal system energy efficiency is achieved by controlling the power of CUs and D2D pairs while ensuring the minimum user demand. Inspired by [13], we propose an approximate convex optimization based algorithm to solve the system model. After that, we further reduce the computational complexity by using a convolutional neural network based on the spatial pyramidal pooling.

2 System Model

We assume an SDN & NFV-enabled virtualized RAN, where a D2D underlaying multi-cell cellular network with uplink scenario. As shown in Fig. 1, the system is managed by an infrastructure provider (InPr) that supports multiple wireless services

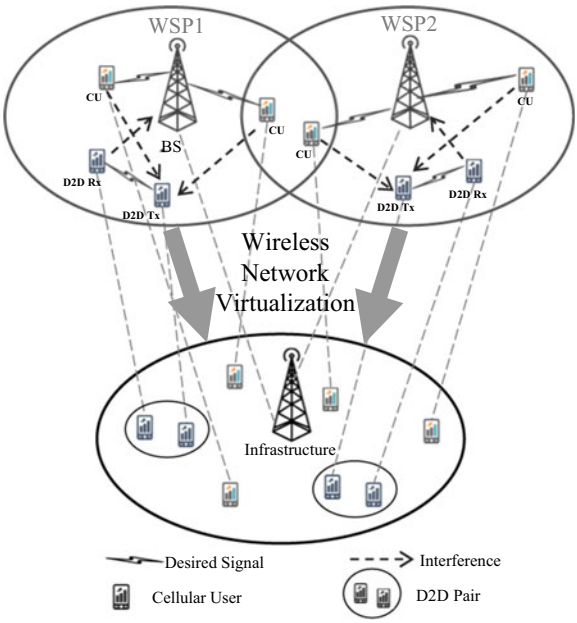


Fig. 1 System model

over a common physical network infrastructure. The RAN consists of N base stations belonging to the set $B = \{B_1, \ldots, B_N\}$, and contains M WSPs belonging to the set $M_{sp} = \{1, \ldots, M\}$. Each WSP m serves K_m users, which are deviced into the cellular users set C_m and D2D set D_m.

The spectrum band used in the network is divided into L non-overlapping orthogonal sub-channels with the same bandwidth B (kHz). Thus, for each time slot t, ℓ resource blocks can be divided for the slicing of WSPs. We agree in advance on an access ratio between SPs and INR as σ_m, and then set $\sum_{\sigma_m \in M} \sigma_m \leq 1$ to ensure fair competition among WSPs for each time slots. We consider that the spectrum reuse efficiency of the system is 1, that is, the spectrum resources of all base stations are the same. In addition, more D2D devices are required to allow D2D pairs for reusing resources blocks from the uplink to achieve higher spectrum efficiency. It is assumed that the BS has perfect channel state information (CSI) from all users (both CU and D2D link) belonging to all SPs. The wireless virtualization controller is responsible for collecting information from the WSPs and returning resource allocation results, to effectively coordinate inter-cell interference, which is then passed to all base stations. Furthermore, the channel is equally and independently distributed over different time slots.

3 Problem Formulation and Proposed Method

In the uplink scenario (see Fig. 1), a cellular user is interfered with four kinds of corresponding interferences when D2D pair use the same RB(s) simultaneously. The normalized achievable data rate R_c^m of CU C_c^m in WSP m is formulated as:

$$R_c^m = \log_2\left(1 + \frac{x_c^m P_c^l G_{c,B}^m}{H_1 + H_2 + H_3 + \sigma^2}\right) \tag{1}$$

$$H_1 = \sum_{d=1}^{D_m} P_{d,c}^m G_{d,B}^m, \quad H_2 = \sum_{m'=1}^{M} P_c^{m'} G_{c,B}^{m,m'}, \quad H_3 = \sum_{m'=1}^{M}\sum_{c=1}^{C_n} P_{c,d}^{m'} G_{c,B}^{m,m'} \tag{2}$$

where x_c^m represents the association relationship between cellular user $c \in C_m$ and RB $l \in L$, if RB l is allocated to user c, then $x_c^m = 1$, otherwise $x_c^m = 0$. In addition, P_c^m, $P_{d,c}^m$, $P_c^{m'}$, $P_{c,d}^{m'}$, σ^2 denote, respectively, the transmission power of CU, D2D pairs, and additive white Gaussian noise power. The normalized achievable data rate R_d^m of D2D pairs D_c^m in WSP m is formulated as:

$$R_d^m = \log_2\left(1 + \frac{x_d^m P_{c,d}^m G_{d,d}^m}{P_c^m G_{c,d}^m + H_1 + H_2 + H_3 + \sigma^2}\right) \tag{3}$$

$$H_1 = \sum_{d'=1}^{D_m} P_{d',c}^m G_{d',d}^m, \quad H_2 = \sum_{m'=1}^{M} P_c^{m'} G_{c,n}^{m,m'}, \quad H_3 = \sum_{m'=1}^{M}\sum_{c=1}^{c_n} P_{d,c}^{k'} G_{d,d}^{m,m'} \tag{4}$$

and, the all power consumption of CUs and D2D pairs can be expressed as:

$$P_c^m = \frac{1}{\mu} P_c^m + P_{cir} \quad P_d^m = \frac{1}{\mu} P_d^m + 2P_{cir} \tag{5}$$

where μ denotes the power amplifier efficiency factor, P_{cir} represents circuit power consumption of one device. As mentioned above, the optimization problem of energy can be resolved by Dinkbach method [14, 15] in this model as below,

$$\begin{aligned} P1: \min_{P_c^m, P_{d,c}^m} & \prod_{m=1}^{M}\prod_{c=1}^{C_m}\left(\frac{H_1 + H_2 + H_3 + \sigma^2}{f(P_c^m, P_{d,c}^m)}\right) \\ & * \prod_{m=1}^{M}\prod_{c=1}^{C_m}\prod_{d=1}^{D_m}\left(\frac{H_1 + H_2 + H_3 + P_c^m G_{c,d}^m + \sigma^2}{g(P_c^m, P_{d,c}^m)}\right) \\ & * \prod_{m=1}^{M}\prod_{c=1}^{C_m}\left(1 + \frac{\eta_{EE}^*}{\lambda} ln2 P_c^m\right) \\ & * \prod_{m=1}^{M}\prod_{c=1}^{C_m}\prod_{d=1}^{D_m}\left(1 + \frac{\eta_{EE}^*}{\lambda} ln2 P_{d,c}^m\right) \end{aligned} \tag{6}$$

subject to:

$$C1: \frac{\sum_{i=1}^{3} H_i + \sigma^2}{P_c^m G_{c,B}^m} \leqslant \frac{1}{2^{R_{min}^c} - 1} \quad \forall c \in C_m,\ m \in M_{sp}$$

$$C2: \prod_{m=1}^{M} \left(\frac{\sum_{i=1}^{3} H_i + P_c^m G_{c,d}^l + \sigma^2}{g(P_c^m, P_{d,c}^m)} \right) \leqslant \frac{1}{R_{min}^d} \quad \forall c \in C_m,\ m \in M_{sp}$$

$$C3: 0 \leqslant p_c^m \leqslant P_{max}, \quad \forall c \in C_m,\ m \in M_{sp}$$

$$C4: 0 \leqslant P_{d,c}^m \leqslant p_{max}, \quad \forall d \in D_m,\ c \in C_m,\ m \in M_{sp}$$

$$C5: 0 \leqslant \sum_{m=1}^{M} \sum_{c=1}^{C_m} P_{d,c}^l \leqslant P_{max}, \quad \forall d \in D_m,\ m \in M_{sp}$$

where $f(P_c^m, P_{d,c}^m)$ and $g(P_c^m, P_{d,c}^m)$ are the approximate monomial functions obtained by the condensation method. The objective function of the final obtained $P1$ is a geometric programming problem. Therefore, we can use the off-the-shelf convex optimization tool CVX toolbox to solve it.

3.1 Convolutional Neural Network Based Algorithm

CNN is used to solve resource allocation problem in this section. The computation of the above Dinkelbach-based geometric programming algorithm is very expensive, and difficult to be implemented in practice system due to the search space grows exponentially with the increase of the number of slice users. But training a convolutional neural network with little time beforehand, power control and channel assignment decisions can be made fast. Since the users at virtualization slicing are not fixed and the underlying CNN needs fixed input, we use a CNN with spatial pyramid pooling (CNN-SPP) to adapt the input channel gain data with dynamic value. The CNN-SPP architecture is given below [16].

Generally speaking, CNN consists of a series of convolutional layers, activation layers, pooling layers, fully connected layers and softmax classifiers. In Fig. 2, the basic idea of CNN is to fit the functional analytic of the classifier, where the test data can be classified directly according to the training of the existing data.

As mentioned above, the allocation of resources to the CUs and the D2D pairs is primarily determined by the condition of the channel gain, Therefore, we arrange the channel gains of all links in system into a three-dimensional matrix as input to the CNN. In the system model, there are have N base stations, and each WSP m has C_m CUs and D_m D2D pairs. Thus the shape of the three-dimensional matrix is $(N(C_m + D_m), N(D_m + 1), 1)$. The function of the convolutional layer is to further extract the features of the input data, we use two layers of filters in the convolutional layer to convolve with the data in turn to obtain the feature map composed of neurons. The role of pooling layer is to retain the main features while reducing the amount

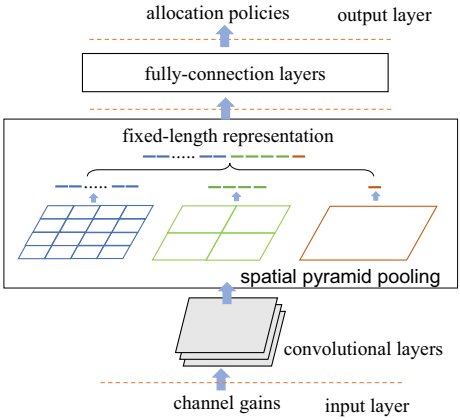


Fig. 2 Architecture of CNN-SPP

of data to reduce over-fitting, we use a spatial pyramid based pooling layer at the end of the convolutional layer not only to process the data at a faster rate, but also to eliminate the input limitation problem of CNN by using maximum pooling to get the maximum parameters and drop the rest. The fully connected layer performs weighted compression of the previously extracted features to derive a weight value equal to the number of categories classified, i.e., generating data that enables the execution of resource allocation policies. Due to the output of fully connected layer are not fixed, we implement zero padding strategy to increase the short output data, and use ReLU as the activation function to ensure getting appropriate results. The softmax classifier will calculate the gradient by calculating a single training data according to cross entropy of decision probability during CNN back propagation. The optimizer of the CNN is Adam, which has the advantages of adaptive learning rate and gentle convergence direction. In this article, we use the convex optimization algorithm to generate the training dataset. The output of our proposed CNN are power and channel allocation policies, and the input is channel gains. In order to solve the fixed input/output limitations of CNN, wo adopt the spatial pyramid based pooling and zero padding strategy to solve it. We use 80% of the dataset for training and remaining 20% for test the results of training. For the loss function of the network, we use the mean square error to evaluate the current estimation because it is more robust to larger errors and can tolerate smaller errors, which can be formulated as follows:

$$L(\hat{y}, y) = \sum_{i=1}^{n} \frac{1}{n} (y_i - \hat{y}_i)^2 \tag{7}$$

where n denotes the number of the neurons which are generated by the fully connected layer, y indicates the resource allocation vector obtained from training data, $\hat{y}$ indicates the predicted resource allocation vector generated by CNN. Besides, y_i and $\hat{y}_i$ represent the i-th element in y and $\hat{y}$.

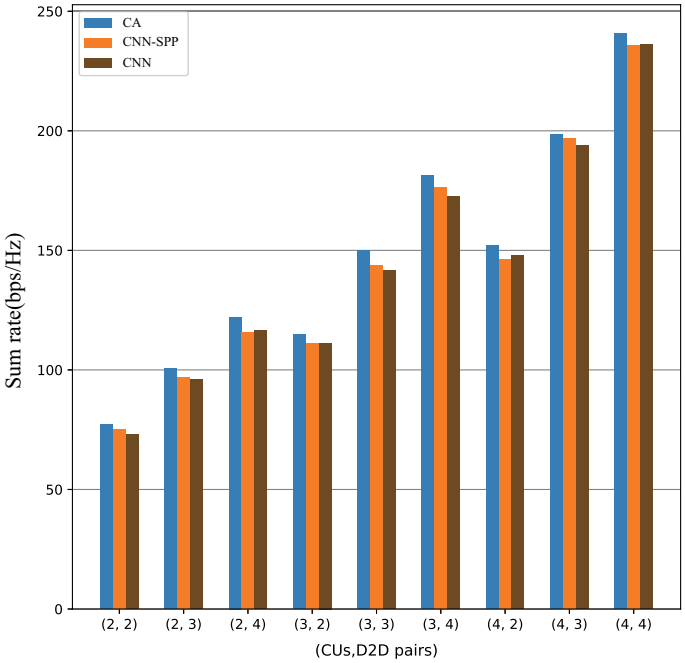


Fig. 3 System sum rate

4 Simulation Results and Analysis

4.1 Parameters

In this section, the system and convex approximation algorithm simulations are carried out by matlab software in this paper. In our case, the number of WSPs present simultaneously is set to 2 in a circular area of 500 m radius. Assuming that all CUs and D2D pairs are uniformly and independently distributed in the circular area, and the maximum power of both CUs and D2D pairs is set to 23 dB, and the minimum rate of D2D pairs is set to 3 bps/Hz, and the minimum rate requirement of CUs at least 20% of the initial rate. Noise power spectral density is -124 dBm/Hz. Path loss between BS and CUs is $128.1 + 37.6 \log_{10}(d)$, path loss between D2D pairs is $148 + 40 \log_{10}(d)$. For CNN based algorithm, we use tensorflow as the underlying keras deep learning toolkit to process the data with GUP. About the hyperparameter setting, the size of the filter of the convolutional layer is 16, the kernel shape is (2×2), the step size is 1, and the activation function is Relu, the size of the SPP layer is set to 1, 4 and 16, the size of the fully connected layer is set to 512.

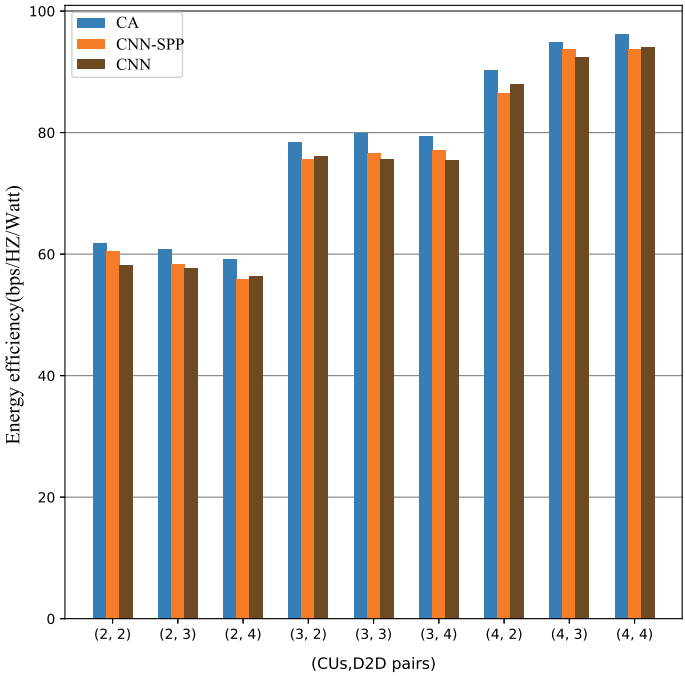


Fig. 4 System energy efficient

4.2 Results and Analysis

In Fig. 3, the total rate of the system increases with the increase of D2D pairs, due to D2D communication is characterized by the reuse of CU spectrum resources and direct communication. Meanwhile, due to the over-fitting of CNN, the sum rate is slightly lower than the convex approximation algorithm named as CA. Benefited from spatial pyramid pooling and zero padding strategy, CNN-SPP is superior to CNN and obtain the similar results compare to CA.

Figure 4 indicates that the energy efficient of our method increases with the number of slices for CUs becoming larger. We increase the number of slices for D2D pairs, however, the change of energy efficient is not obvious and even decrease. D2D pairs has more spectrum resources with the addition of CUs to be shared, even though the non-orthogonal resource sharing between CU and D2D can not avoid the co-channel interference between D2D pairs, the influence can be reduced by proper channel allocation and power control. So that the efficiency of the system is improved. That is, only increasing the D2D pair is not significant, due to the system sum rate increases at the same time the system power consumption is also increased, the D2D pair has no additional resources to share. The result obtained based on CNN-SPP is superior to CNN and slightly lower than the results of the convex approximation algorithm. But, it is clear that the energy efficient still achieves 97.81%.

Table 1 Computational time

Number of slices	Basic method (s)	CNN-SPP (ms)	CNN (ms)
(2, 2)	7.09	0.093	0.124
(2, 3)	8.18	0.123	0.161
(2, 4)	9.97	0.148	0.183
(3, 2)	8.25	0.121	0.175
(3, 3)	9.76	0.145	0.200
(3, 4)	14.04	0.173	0.237
(4, 2)	13.51	0.144	0.211
(4, 3)	16.46	0.169	0.254
(4, 4)	22.59	0.212	0.302

Finally, from Table 1, we find the calculation time of the optimal method is long. Because of the approximate convex optimization algorithm based on the Dinkle-bach's method, which requires multiple iterations and multiple geometric programming problems for each solution. At the same time, the search space increases exponentially and the time increases even more, with the number of slices increases. Although CNN specifically introduces the features of local connection and weight sharing to reduce the calculation of training, but CNN-SPP can effectively reduce parameters and complexity by adopting spatial pyramids pooling. In addition, it seems that there is little differences between CNN-SPP and CNN, but as said before, CNN must have fixed input and output. For our system model, increasing user slices need changing input, therefore, we need to train CNN model separately for each case. On the contrary, CNN-SPP is consist of spatial pyramids pooling and it allows input of any size, for changing slice users, we only need to train one model. Obviously, the train cost between CNN-SPP and CNN is enormous, especially in real communication systems where a large number of models need to be trained.

5 Conclusion

In this paper, we study resource allocation in wireless virtualized network, the D2D communication is proposed to improve energy efficient of the system. D2D pairs take full advantage of spectrum resources shared from cellular users. We formulated a joint channel resource allocation and power control problem to obtain system energy consumption maximization. To solve this NP-hard and non-convex problem, we use the approximate convex optimization algorithm to transform it into a convex optimization problem. On this basis, CNN based on spatial pyramid pooling is used for calculation. After the network is fully trained, it is able to select appropriate allocation strategy according to the input. Simulation results show that the sum rate and energy efficiency of the system are improved within wireless network virtual-

ization and D2D communication. In addition, the CNN-SPP method achieved close to optimal performance while having a much lower computational complexity by sacrificing some training time.

Acknowledgements This work is the results of the research project funded by the National Key Research and Development Program of China under Contract No. 2016YFE0200200.

References

1. Liang C, Yu FR (2015) Wireless network virtualization: a survey, some research issues and challenges. IEEE Commun Surv Tutor 17(1):358–380
2. Liu J, Zhang S, Kato N, Ujikawa H, Suzuki K (2015) Device-to-device communications for enhancing quality of experience in software defined multi-tier LTE-A networks. IEEE Netw 29(4):46–52
3. Lee J, Lee JH (2019) Performance analysis and resource allocation for cooperative D2D communication in cellular networks with multiple D2D pairs. IEEE Commun Lett 23(5):909–912
4. Anvith PV, Gunavathi N, Malarkodi B, Rebekka B (2019) A survey on network functions virtualization for telecom paradigm. In: 2019 TEQIP III sponsored international conference on microwave integrated circuits, photonics and wireless networks (IMICPW), pp 302–306
5. Nguyen K, Lu Q, Huang C (2021) Joint node-link embedding algorithm based on genetic algorithm in virtualization environment. In: 2021 IEEE 94th vehicular technology conference (VTC2021-Fall), pp 1–5
6. Linguaglossa L, Lange S, Pontarelli S, Rétvári G, Rossi D, Zinner T, Bifulco R, Jarschel M, Bianchi G (2019) Survey of performance acceleration techniques for network function virtualization. Proc IEEE 107(4):746–764
7. Lai W-K, Wang Y-C, Lin H-C, Li J-W (2020) Efficient resource allocation and power control for LTE-A D2D communication with pure D2D model. IEEE Trans Veh Technol 69(3):3202–3216
8. Li J, Wu D, Yue C, Yang Y, Wang M, Yuan F (2022) Energy-efficient transmit probability-power control for covert D2D communications with age of information constraints. IEEE Trans Veh Technol 71(9):9690–9704
9. Wang X, Jin T, Hu L, Qian Z (2020) Energy-efficient power allocation and Q-learning-based relay selection for relay-aided D2D communication. IEEE Trans Veh Technol 69(6):6452–6462
10. Moubayed A, Shami A, Lutfiyya H (2015) Wireless resource virtualization with device-to-device communication underlaying LTE network. IEEE Trans Broadcast 61(4):734–740
11. Cheng Y, Yang L, Zhu H (2017) Operator profit-aware wireless virtualization for device-to-device communications underlaying LTE networks. IEEE Access 5:11668–11676
12. Han Y, Tao X, Zhang X, Jia S (2020) Delay-aware resource management for multi-service coexisting LTE-D2D networks with wireless network virtualization. IEEE Trans Veh Technol 69(7):7339–7353
13. Xu D, Chen X, Wu C, Zhang S, Xu S, Cao S (2019) Energy-efficient subchannel and power allocation for HetNets based on convolutional neural network. In: 2019 IEEE 89th vehicular technology conference (VTC2019-Spring), pp 1–5
14. Boyd S, Kim S-J, Vandenberghe L, Hassibi A (2007) A tutorial on geometric programming. Optim Eng 8(1):67–127
15. Chi C-Y, Li W-C, Lin CH (2017) Convex optimization for signal processing and communications: from fundamentals to applications. CRC Press
16. He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell 37(9):1904–1916

Research on Key Technologies of Unmanned Combat Vehicle Early Warning Radar

Jinliang Dong, Lei Bian, and Yumeng Zhang

Abstract Aiming at the operational requirements of unmanned combat vehicles, this paper first introduces the operational advantages of unmanned aerial vehicles, and then analyzes the key technologies combined with the characteristics of unmanned aerial vehicles and gives solutions; Finally, the development of follow-up equipment is prospected, which can provide a good reference and help for the design of unmanned combat vehicle early warning radar.

Keywords Unmanned combat vehicle · Early warning radar · Array Integration · Classification identification

1 Introduction

Unmanned combat vehicles can be divided into various types according to their combat tasks. Unmanned combat vehicles with forward assault missions mainly carry machine guns, artillery, anti-tank missiles, etc.; Unmanned combat vehicles tasked with jamming and confrontation mainly carry microwave weapons, lasers, jamming equipment, etc.; Unmanned combat vehicles with the mission of reconnaissance and obstacle clearance are mainly equipped with photoelectric/infrared sensors, acoustic sensors, mechanical arms, etc.; Unmanned combat vehicles tasked with air defense must carry early warning radars, air defense antiaircraft guns, and air defense missiles [1].

As the military gap between major powers gradually narrows, wars will continue to be dominated by local wars, making it difficult to engage in direct military rivalries between major powers. Major power conflicts have evolved into confrontations between small countries that they support. In these local conflicts, the first round of strikes by one side of the war will be non-contact, precision guided weapon air strikes, and air defense operations are essential, just as unmanned combat vehicles are used

J. Dong (✉) · L. Bian · Y. Zhang
Nanjing Glarun Defense System Co, Nanjing, China
e-mail: falsejinliangdong@163.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_5

in other combat modes, The unmanned air defense vehicle will replace the currently manned air defense weapons with its unique advantages. Its main advantages are:

(1) The combat effectiveness of individual weapons and equipment is limited, and future operations will mainly focus on cluster operations. The regional joint defense, multi-layer interception, and multi-vehicle cooperative operations of unmanned air defense vehicles can achieve real-time information interaction and air intelligence information sharing, maximizing operational effectiveness.
(2) Unmanned air defense combat vehicles are suitable for 24/7 duty, without personnel involvement, and there will be no situations such as misoperation or delaying the aircraft. Moreover, equipment supply and energy supply can be managed in a data-based manner, enabling orderly operations and support.
(3) For autonomous unmanned air defense combat vehicles, the rapidity of battlefield deployment and deployment, complex battlefield situational awareness capabilities, multi target layered fire distribution, and rapid response capabilities to emergent targets are unmatched by personnel [2–4].

2 Key Technology Analysis

Unmanned air defense combat vehicles are mainly equipped with unmanned support and support teams to intercept and strike close range low altitude targets that appear and penetrate in the face of aircraft, supporting ground assault forces. As the focus of air defense weapons and equipment, early warning radar will also usher in the trend of unmanned development. Of course, due to the constraints of unmanned combat vehicle platforms, there is also a high demand for unmanned combat vehicle radar design, which requires lightweight, miniaturization, integrated design, and the ability to real-time target classification and recognition, and target detection in complex electromagnetic environments.

2.1 Integrated Technology of Active Phased Array Antenna

2.1.1 Technical Difficulties

Unmanned combat vehicle early warning radar is limited by the severe limitations of unmanned air defense vehicles on the volume and weight of radar loads, and generally uses X and above bands. During the design process of high band active phased array antenna arrays, there are many problems such as high antenna specification requirements, high integration, and difficult implementation. In addition to the current development of radar systems towards multi-function integration, conventional phased array antennas cannot meet the system requirements in terms of weight, volume, efficiency, reliability, and other aspects. Therefore, it is necessary to carry out

research on low-cost and highly integrated phased array antenna array design technology. Active phased array antenna arrays implemented using new technologies must have the following characteristics [5, 6]:

(1) High integration and lightweight. Limited by the vehicle platform, phased array antennas must occupy as little volume and weight as possible. Therefore, the antenna array needs to be designed and processed in an integrated antenna and feeder manner to reduce the number of interconnecting cables and connectors, as well as reduce volume and weight. At the same time, advanced heterogeneous integration technology, vertical interconnection technology, and advanced packaging technology are used to develop advanced highly integrated and efficient component integration modules, thereby further realizing array integration and lightweight.
(2) Low cost. Traditional discrete modular devices not only increase the volume and weight of radar, but also increase the cost of radar. The antenna array can use customized and highly integrated silicon based multifunctional chips in components, power supplies, and control systems on a large scale, effectively reducing the amount of module equipment and significantly reducing chip costs; At the same time, using advanced three-dimensional packaging technology to reduce dependence on packaging substrates can also significantly reduce chip packaging and integration costs.
(3) Scalability. Active phased array antenna integration can abandon the architectural idea of interconnecting functional modules in conventional arrays through cable components, and introduce the concept of microsystems for system level integration at a micro scale. It is possible to integrate a single functional circuit into a single chip, multiple functional circuits into a single heterogeneous chip, and components into SIP or SOC. Finally, advanced three-dimensional packaging technology and three-dimensional interconnection technology are used to replace conventional mechanical assembly and cable connection. The phased array integration technology based on microsystems can greatly improve the array integration degree, integrate more T/R channels and achieve more functions in limited space, and meet the requirements of multi-functional integrated radar systems for phased array antenna arrays.

2.1.2 Solutions

(1) Integrated design of components and antenna units

   (a) Using electromagnetic and circuit joint simulation to comprehensively consider the impact of cavity effects; Improve the stability coefficient of the entire link through reasonable gain allocation and cavity isolation;
   (b) By reasonably selecting the implementation form of the antenna unit, it can meet the requirements of the LTCC process, have good repeatability, stability, and sufficient mechanical strength, and meet the conditions for structural assembly and installation connection;

(c) Screening for amplitude and phase inconsistencies in microwave components and establishing semi-automatic assembly lines for the production and assembly of T/R components.

(2) Integrated Network Design

(a) Using ultra-thin LTCC multilayer hybrid digital analog wiring, the module uses LTCC boards as the circuit carrier, including passive circuits such as power splitters, microwave transmission lines, microstrip stripline conversion, and digital analog hybrid circuits such as power supplies, digital circuits, and drive circuits;
(b) Using a groove dug in the LTCC board, the microwave chip is installed in a shallow groove, coupled with an LTCC cover plate, and the integration between the two layers of LTCC is improved through ultrasonic hot pressing welding;
(c) LTCC multilayer boards and MCM technology are used to achieve highly integrated components. Precision assembly technologies are used, including the bonding and welding of various materials, chips, components, and LTCC boards, ultrasonic hot pressing welding between LTCC boards, welding and ball planting on the surface of LTCC boards, control of different welding temperatures, and gold wire bonding.

(3) Electromagnetic compatibility design of components

(a) Adjust the operating mode of the transceiver circuit, turn off the transmitter circuit during reception, provide maximum device isolation, and avoid the impact of the transmitter circuit on the receiver circuit;
(b) Using simulation software to guide wiring design;
(c) The microwave transmission line adopts a stripline system to isolate microwave circuits from digital circuits and low-frequency circuits;
(d) Simulate and calculate the cavity, and take measures such as adding partitions to reduce the cavity effect and mutual coupling;
(e) Reasonably distribute the gain links of the transmission circuit in the physical space, and conduct stability simulation analysis to ensure the stability and reliability of the transmission circuit;
(f) All vertical interconnections adopt coaxial like structures to reduce electromagnetic leakage at the interconnections.

2.2 Real Time Target Classification and Recognition Technology

2.2.1 Technical Difficulties

Due to the attitude sensitivity of the target, the RCS fluctuates greatly, affecting recognition performance. The classification of aerial targets depends on the modulation characteristics of the targets. However, the modulation characteristics of the targets have problems of instability and attitude angle sensitivity, and the modulation spectrum cannot be observed at some attitude angles. After obtaining target features, how to distinguish different targets to the maximum extent and improve the recognition rate is a problem to be solved in the design of the recognizer. When designing a recognizer for radar aerial target recognition, factors such as recognition performance, generalization ability, and recognition computational complexity should be considered.

2.2.2 Solutions

In order to utilize system resources as reasonably as possible, the system can use intelligent classification and recognition technology for real-time and rapid target classification, providing a basis for adaptive resource scheduling [7].

(1) Using hierarchical, multi feature, multi algorithm parallel, and sequential recognition, targets are classified and clustered against a type template library based on perceived motion characteristics and RCS characteristics, and key targets are subjected to fine feature perception and feature extraction.
(2) Build a deep recognition network with transfer learning capabilities to improve the ability to identify small sample targets. According to the results of target classification and threat assessment, decisions are adjusted based on the assessment results to achieve sequential fusion recognition of inter frame decisions and improve recognition robustness. The real-time target classification and recognition architecture is shown in Fig. 1.

A certain type of radar has achieved classification verification of fixed wing aircraft, propeller aircraft, and helicopters using RCS and JEM characteristics, with a recognition rate greater than 85%, as shown in Fig. 2.

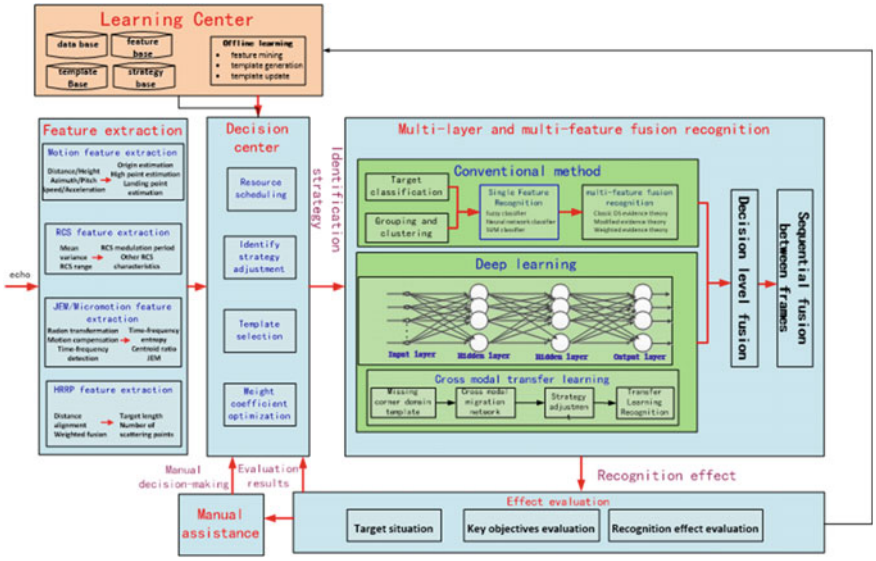


Fig. 1 Real time target classification and recognition architecture

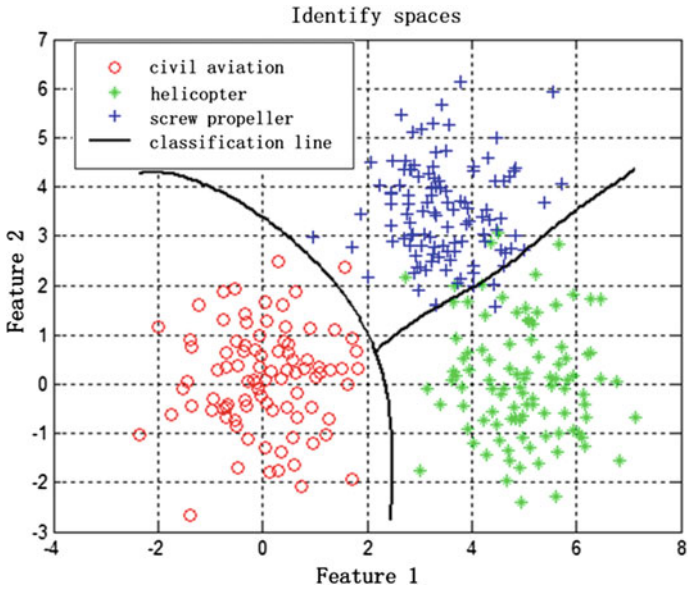


Fig. 2 Target classification and recognition features

2.3 Integrated Anti-interference Technology

2.3.1 Technical Difficulties

With the development of electronic countermeasures technology, the means of active jamming are becoming increasingly diverse and complex. Various electromagnetic radiation crisscross in the airspace, dynamically change in the time domain, densely overlap in the frequency domain, and fluctuate in the energy domain, seriously affecting the detection performance of radar systems. "Radar operators cannot simultaneously identify various types of interference for multiple simultaneous interferences, and therefore cannot implement corresponding targeted anti jamming measures. It is difficult to achieve satisfactory results by taking only a single type of anti jamming measures [8, 9]

2.3.2 Solutions

Comprehensive use of multiple spatial, temporal, and frequency domain means to refine anti-interference methods, using ultra-low sidelobe, adaptive nulling, and sidelobe blanking anti-interference techniques in the spatial domain, using large time bandwidth, low peak power, pulse compression, narrow pulse rejection, and interference suppression techniques based on waveform entropy in the time domain, and using frequency regulation, multiple complex signal modulation, and adaptive frequency agility anti-interference techniques in the frequency domain, It is possible to achieve the comprehensive anti-interference function of the radar as much as possible based on a limited amount of equipment [10].

Use the jamming environment to conduct comprehensive interception, identify multiple jamming types, use the expert intelligent anti-jamming decision-making system, call corresponding anti-jamming measures from the anti-jamming measures library to counter multiple jamming (Fig. 3)

Figure 4 shows the effect pictures before and after the comprehensive application of anti-interference measures. Compared with the display and control interfaces before and after the interference suppression measures take effect, the number of echoes decreases by 94.5%, the number of dots decreases by 92.5%, and the number of false tracks decreases by 100%. Through comprehensive interference suppression processing in space, time, and frequency domains, the number of jamming echoes, dots, and false tracks on the radar display and control interface are significantly reduced, and the comprehensive anti-interference measures have significant effects.

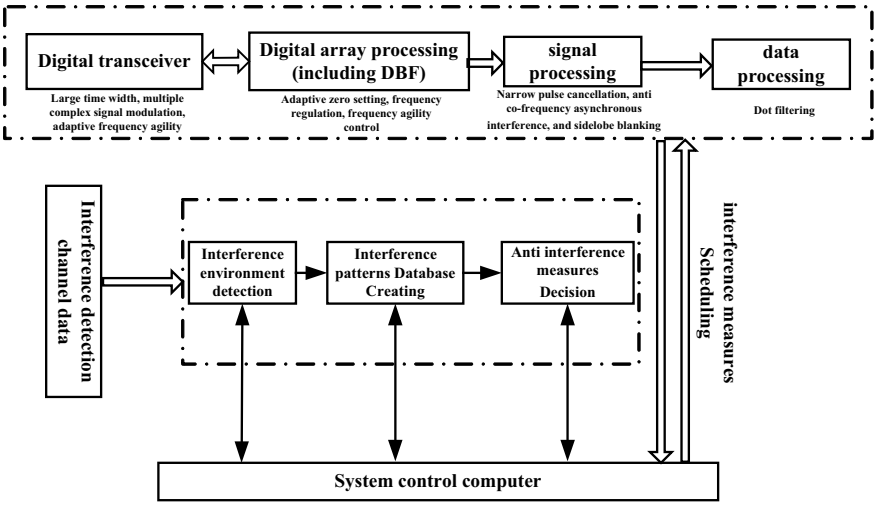
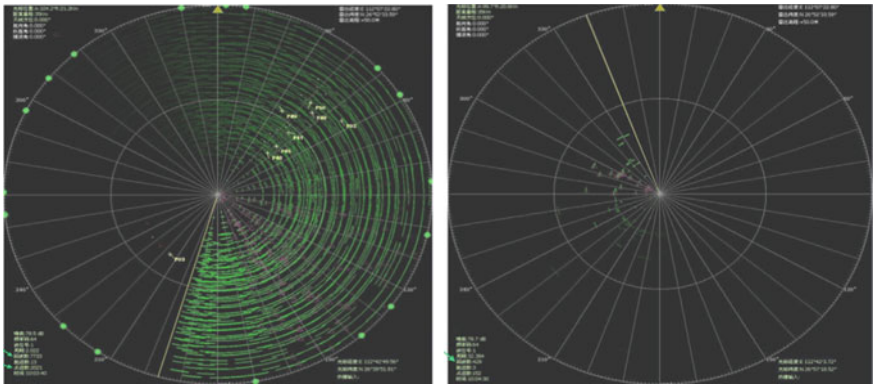


Fig. 3 Radar integrated anti-jamming process



(a) Before interference suppression (b) After interference suppression

Fig. 4 Effect diagram before and after interference suppression

3 Conclusion

In summary, unmanned air defense vehicles are an important component of ground unmanned combat forces, and early warning radar, as an important load of unmanned air defense vehicles, its performance will affect the operational effectiveness of the entire vehicle. With the further development of weapon technology, the future unmanned combat vehicle early warning radar will have the capabilities of automatic situational awareness, automatic optimization of operating parameters, intelligent

target detection and recognition, and conformal to combat vehicle platforms, which will also promote the further development of unmanned system technology.

References

1. Zhu C, Mao Y, Xu X (2019) Research on the application of intelligent unmanned combat vehicles in future warfare. Mod Defense Technol 47(2):6–11
2. Zhao M, Hao X, Jiang X (2019) Research on remote control data link overall technology of unmanned combat vehicle. Fire Control Command Control
3. Keller J (2020) General Dynamics to design unmanned combat vehicle for Army battlefield cargo. Military Aerosp Electron 8:31
4. Wang S, Zhang Y, Liao Z (2019) Building domain-specific knowledge graph for unmanned combat vehicle decision making under uncertainty. In: 2019 Chinese Automation Congress (CAC)
5. He Q, He H, Chen J (2015) System structure and design method of the high density integrated active phased array antennas. In: 2015 National Microwave and Millimeter Wave Conference
6. Liu X (2020) A low-cost millimeter wave Ka-band active phased arry antenna. Appl Electron Tech 46(12):1–4
7. Li M, Wu J, Zuo L et al (2018) Aircraft target classification and recognition algorithm based on measured data. J Electron Inf Technol 40(11):2606–2612
8. Liu Y, Wang C, Zheng G et al (2021) An active anti-jamming approach for frequency diverse array radar with adaptive weights. J Beijing Inst Technol 30(4):403–411
9. Wei W (2018) Study on modern radar anti-jamming technology under complex electromagnetic environment. Electronics World
10. Dong J, Zhang Y, Wang G et al (2021) Research on overall design of mobile phased array radar for field vehicle. In: The 10th International conference on communications, signal processing, and systems (CSPS)

Study of the Adaptive Bandwidth Communication Method Based on Software Defined Radio

Yin Long

Abstract Aiming at the problem of limited communication bandwidth between satellite and ground station, combining the communication scenario between satellite and ground station, the forward and backward link margin is analyzed, and a bandwidth adaptive communication method based on software defined radio is proposed, which dynamically adjusts the forward and backward link bandwidth according to the signal-to-noise ratio of the actual received signal from satellite and ground station. By this method, the waste of link resources is avoided, and the channel capacity is maximized. Finally, the transmission capacity in the measurement and control arc section of the ground station will be improved. The simulation results show that the transmission capacity of the proposed method is 2–3 times higher than that of traditional method.

Keywords Bandwidth · Adaptive · Signal-to-noise ratio · Transmission capacity · Software defined radio

1 Introduction

The communication system between the traditional satellite and the ground station is specified in advance according to the link budget results of the extremely bad conditions, to ensure that the satellite and the ground station can still establish and maintain the link normally even under the worst link conditions, and generally reserve the 3 dB link margin. In the actual scenario, only when the satellite just enters or leaves the elevation range of the ground station (generally 5°), and the weather is bad (rain, snow), the link margin will be reduced to 3 dB. In the rest of the time, the link margin between the satellite and the ground station is greater than 3 dB, and the link margin will increase significantly as the scene changes. Therefore, the traditional communication system is unchangeable. Although it can ensure the stable establishment and maintenance of the link in the whole measurement and

Y. Long (✉)
Beijing Institute of Spacecraft System Engineering, Beijing 100094, China
e-mail: ly24381@163.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_6

control arc of the ground station, there is a serious problem of link resource waste, and there is room for further optimization and improvement of channel capacity. At present, the mainstream satellite communication at home and abroad uses the DVB-S standard, which uses the fixed code modulation system, leaving a fixed link margin to compensate for channel fading, reducing the system data transmission rate and wasting satellite spectrum resources. In order to solve the above problems, DVB-S2 [1–6], the second generation DVB standard based on adaptive coding and modulation system, is proposed internationally to improve the transmission capacity of the system by adopting a combination of multiple channel coding and modulation schemes. However, the above standards lack the research on adaptive switching methods and engineering implementation details for coded modulation, and there are no in-orbit flight cases at home and abroad. In this paper, a bandwidth adaptive communication method is proposed. First, the satellite communication scenario is modeled. Through the evaluation and interaction of the signal-to-noise ratio of the forward and backward links, the bandwidth is adaptively adjusted to maximize the use of channel capacity and improve the transmission capacity within the arc of the TT&C station.

2 System Analysis

Considering the working scenario as shown in Fig. 1, the satellite S communicates with the ground station G in both directions, and the antenna beam pointing elevation of G is α, The height of S's orbit from the earth's surface is H km, the distance between S and G is D km, the earth's radius is R km, and the earth's center is O. According to geometric relationship

$$D^2 + 2\mathrm{R}D\sin\alpha - (2\mathrm{R} + H)H = 0 \tag{1}$$

It can be solved according to Formula 1

$$D = -\mathrm{R}\sin\alpha + \sqrt{(\mathrm{R}\sin\alpha)^2 + (2\mathrm{R} + H)H} \tag{2}$$

The free space attenuation of RF signal is specified as L dB, and the operating frequency is f MHz. The calculation result of free space attenuation L is shown in Formula 3

$$L = 20\lg D + 20\lg f + 32.45 \tag{3}$$

Substitute Formula 2 into Formula 3

$$L = 20\lg(-\mathrm{R}\sin\alpha + \sqrt{(\mathrm{R}\sin\alpha)^2 + (2\mathrm{R} + H)H}) + 20\lg f + 32.45 \tag{4}$$

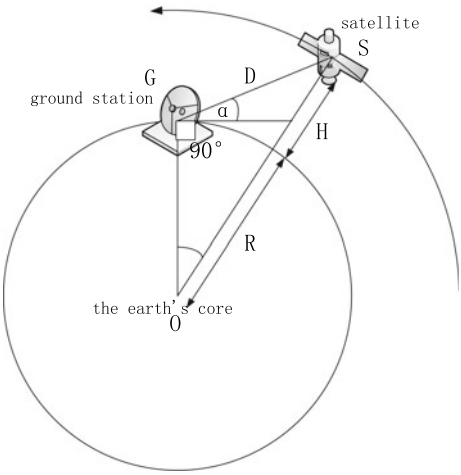


Fig. 1 Scenario of the communication between the satellite and ground station

The minimum elevation of the antenna beam direction of G is specified as α_{Min}, the orbital height of S from the earth's surface is H, and the distance between S and G is the maximum, which is recorded as D_{max}. It is specified that the transmission power of satellite and ground station is a constant value P in the whole arc section of observation. In order to ensure the normal establishment and maintenance of the link, the lowest elevation of the antenna beam entering or leaving G at S is required α. At the moment, the link margin M is greater than or equal to the minimum link margin M_{min} which is specified as 3 dB. As S enters G's α After min, D decreases gradually and M increases correspondingly. When S passes the zenith, D and M reach the minimum and maximum values at the same time, and then D increases gradually and M decreases gradually until S flies away from the lowest elevation of the antenna beam of G.

It is specified that the original link bandwidth is B_s, the carrier to noise ratio of the receiver's entrance signal is C/N0, the signal to noise ratio is Eb/N0, the carrier to noise ratio demodulation threshold of the receiver is $C/N0_{min}$, the signal to noise ratio demodulation threshold is $Eb/N0_{min}$, and the calculation formula of the link margin M is shown in Formula 5.

$$M = C/N0 - 10\lg B_s - Eb/N0_{\min} \tag{5}$$

The link margins at T1 and T2 are specified as M1 and M2 respectively. The calculation results of M1 and M2 are as follows

$$M1 = (C/N0)1 - 10\lg B_s - Eb/N0_{\min} \tag{6}$$

$$M2 = (C/N0)2 - 10\lg B_s - Eb/N0_{\min} \tag{7}$$

$$\Delta M = M1 - M2 = (C/N0)1 - (C/N0)2 = \Delta(C/N0) \tag{8}$$

$$\Delta(C/N0) = L1 - L2 = 20\lg D1 - 20\lg D2 = 20\lg \frac{D1}{D2} \tag{9}$$

$$\Delta M = 20\lg \frac{D1}{D2} \tag{10}$$

α_{Min} is specified as the minimum elevation of satellite S at the ground station, and M_{min} is specified as the minimum link margin. At this time, the satellite's transmission bandwidth is B_s, the operating frequency is f, and the low-order modulation system is adopted. The ground station G corresponds to the corresponding receiving system. According to Formula 10, the link margin M of the satellite changes with the distance between the satellite and the ground station, and the minimum margin is 3 dB at the farthest distance, and the maximum margin at the nearest distance.

The MPSK modulation system is specified for the link, and the bit error rate is specified as P_e which is required below 10^{-5}. The P_e calculation formula of MPSK system is as follows, where P_e represents the error rate, Eb/N0 represents the ratio of unit bit signal to noise, and M represents the modulation phase number of MPSK (Fig. 2).

$$P_e = \frac{erfc\left(\sqrt{\frac{Eb}{N0} * \log_2(M)} * \sin\left(\frac{\pi}{M}\right)\right)}{\log_2(M)} \tag{11}$$

3 System Design

In order to maximize the utilization of channel capacity, the threshold of receiver C/N0 is set to J1, J2 … Jn. If Ji ≤ C/N0 ≤ Ji + 1 (i = 1, 2 … n − 1), the transmission bandwidth B and modulation mode are updated accordingly, so that the link margin M is basically maintained at about 3 dB. According to Formula 11, the decision threshold J1, J2 … Jn of link carrier to noise ratio C/N0 is obtained, as shown in Table 1.

$$\frac{C}{N0} = \frac{Eb}{N0} + 10 * \lg B + \mathrm{M} \tag{12}$$

The calculation flow chart of C/N0 threshold is shown in Fig. 3.

It is specified that the backward link is from satellite to ground station, and the forward link is from ground station to satellite. The ground station and satellite update the MPSK modulation and demodulation mode of the backward link and the forward link respectively according to the estimated results of the link carrier-to-noise ratio.

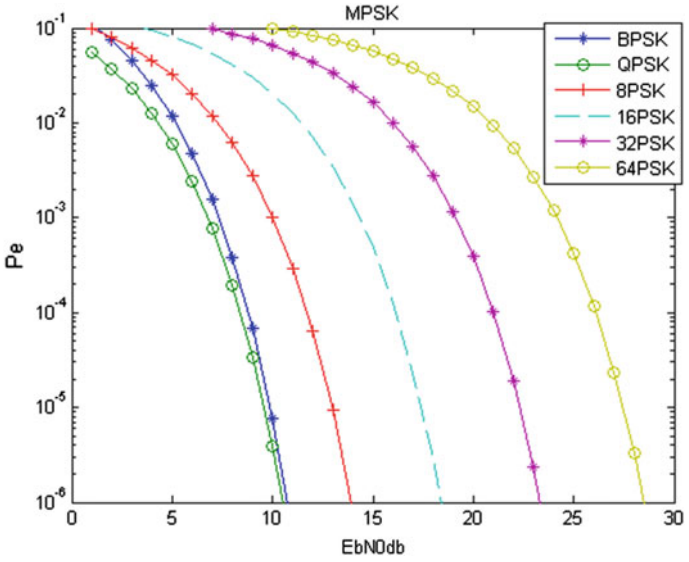


Fig. 2 Chart of Eb/N0 and bit error rate

Table 1 Criterion of the receiver modulation mode

Index	C/N0	Demodulation mode	Transmit bandwidth	Demodulation threshold Eb/$N0_{min}$ (dB)	P_e
1.	[C/N0 $_{min}$, J_1]	BPSK	B_s	10	7.74E−06
2.	[J1, J2]	QPSK	$2B_s$	10	3.87E−06
3.	[J2, J3]	8PSK	$4B_s$	13	9.42E−06
4.	[J3, J4]	16PSK	$8B_s$	18	2.93E−06
5.	[J4, J5]	32PSK	$16B_s$	23	2.39E−06
6.	[J5, J 6]	64PSK	$32B_s$	28	3.26E−06

For the return link, the ground station sends the switching command of MPSK modulation mode to the satellite according to the comparison result between the actual received signal carrier to noise ratio and the MPSK demodulation mode decision threshold, and switches to the corresponding MPSK demodulation mode at the same time. After receiving the command, the satellite will switch the corresponding MPSK modulation mode and send the command response message to the ground station. For the forward link, the satellite compares the carrier to noise ratio of the actual received signal with the MPSK demodulation mode decision threshold, switches its own MPSK demodulation mode according to the comparison result, and sends the MPSK modulation mode switching command to the ground station. After receiving the command, the ground station executes the MPSK modulation mode switch and

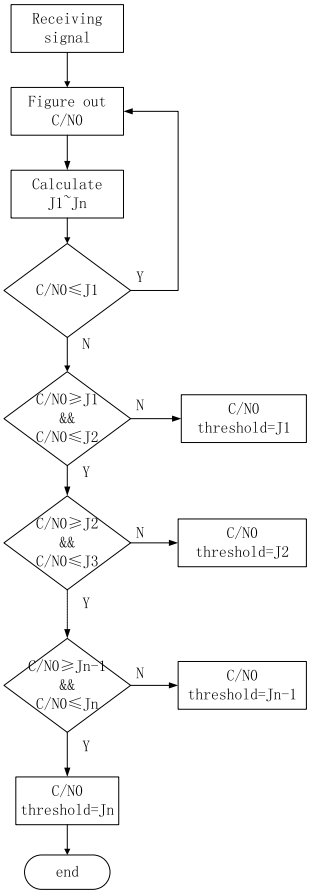


Fig. 3 Flow of C/N0 calculation

returns the command response message to the satellite. Information interaction is shown in Fig. 4.

4 Key Technologies

The key technology of bandwidth adaptive communication is software defined radio. The central idea of software defined radio [7–12] is to construct an open, standardized and modular general hardware platform, realize modulation and demodulation, encryption and decryption, communication protocol and other functions with software, and make broadband A/D and D/A converters as close as possible to the antenna. Software defined radio realizes multiple MPSK communication systems through software, and stores them in PROM or FLASH and other memories. According to the received MPSK switching instructions, it loads the corresponding

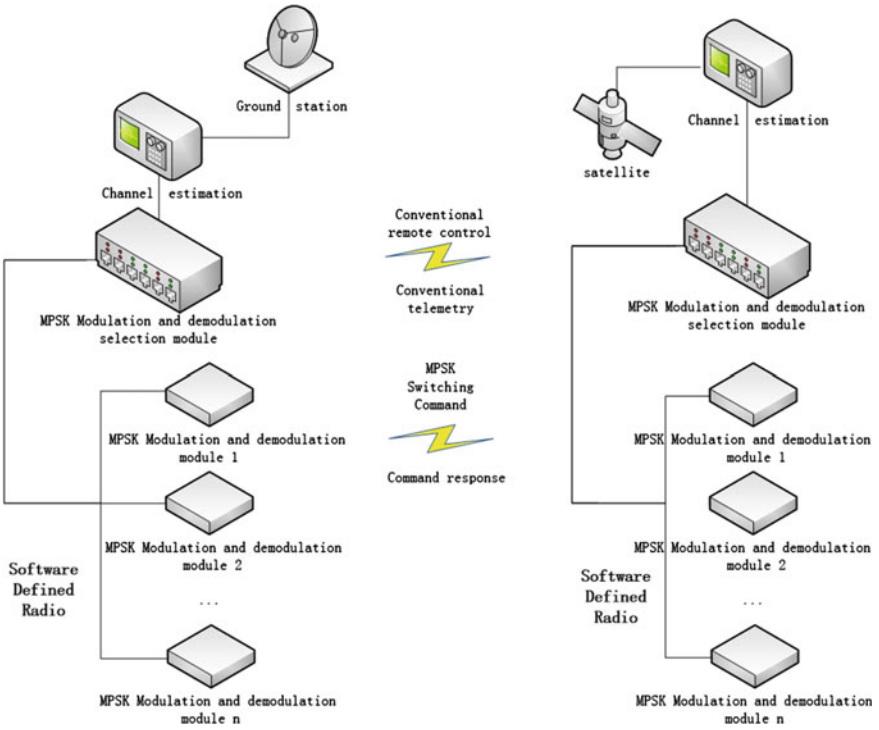


Fig. 4 Sketch map of the adaptive bandwidth communication method based on MPSK

MPSK communication code in real time to realize the mutual switching of multiple MPSK communication systems under the same hardware platform (Fig. 5).

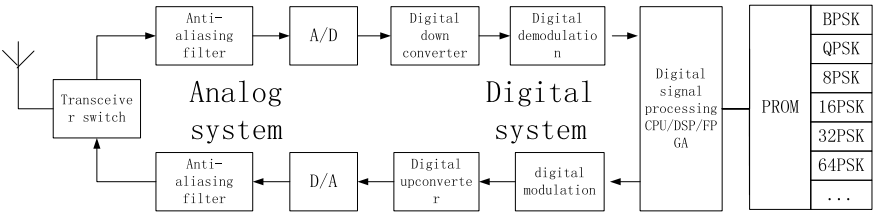


Fig. 5 System of software defined radio

5 Simulation Verification

The STK simulation scenario is set as follows. The six beams of the satellite orbit are shown in Table 2. The half-beam angle of the satellite is 70°. The station in South America is selected as the ground station, and the elevation of the ground station antenna is 5°.

The design initial bandwidth is 20kbps, the modulation mode is BPSK, the demodulation threshold Eb/N0 is 10 dB, and the C/N0 is 53 dBHz. The higher-order modulation mode is MPSK. The transmission bandwidth, demodulation threshold and C/N0 are shown in Tables 3 and 4.

The simulation results show that the bandwidth adaptive communication method is increased to 30,380,000/9,920,000 = 3.0625 times of the total transmission data of the traditional communication method in the measurement and control arc of the ground station.

Table 2 Satellite orbit

Simulation time	3 Nov 2019 11:00:00UTCG ~ 4 Nov 2019 11:00:00UTCG
Semi-major axis	6770.38 km
Eccentricity	0.000172979
Inclination	42.7836 deg
Perigee angle	48.0503 deg
RAAN	357.107 deg
Mean anomaly	175.408 deg

Table 3 Criterion of MPSK

Index	Modulation mode	Transmit bandwidth (kbps)	Demodulation threshold Eb/N0(dB)	P_e	C/N0 (dBHz)
1.	BPSK	20	10	7.74E−06	53.0103
2.	QPSK	40	10	3.87E−06	56.0206
3.	8PSK	80	13	9.42E−06	62.0309
4.	16PSK	160	18	2.93E−06	70.0412
5.	32PSK	320	23	2.39E−06	78.0515
6.	64PSK	640	28	3.26E−06	86.0618

Table 4 Comparison of the adaptive bandwidth communication method with the tradition way

Method	Arrival time	Outbound time	Modulation mode	Transmit bandwidth	Run time(s)	Total data transferred (bit)
Traditional communication method	2019.11.3 21:15:02	2019.11.3 21:23:18	BPSK	20kbps	496	9,920,000
Bandwidth adaptive communication method	2019.11.3 21:15:02	2019.11.3 21:23:18	8PSK	(1) 2019.11.3 21:15:28 ~ 2019.11.3 21:21:09, 80kbps (2) The rest of the time, 20kbps	(1) 341 (2) 155	30,380,000

6 Conclusion

In view of the fixed communication system of traditional communication methods and the lack of effective utilization of link margin, a bandwidth adaptive communication method is proposed. Compared with traditional communication methods, it increases the total amount of data transmission in the same communication arc section, and effectively improves the communication efficiency. In the future, research will be carried out on the bandwidth adaptation methods based on other modulation and coding methods, as well as the introduction of signal attenuation caused by atmospheric absorption, rainfall and other weather conditions into the system.

References

1. Tian M, Wang Z, Xu M (2020) DVB-S2 signal receiving and analysis based on cognitive radio. Comput Sci 47(4):226–232
2. Wu Q, Wang Z, Xie L et al (2022) Modulation recognition of DVB-S2 signal at low SNR. Mod Electron Tech 45(17): 65–69
3. Han X, Yang W (2021) Design and implementation of frame synchronization algorithm for DVB-S2. Comput Network 47(7):60–64
4. Lan Y, Yang H, Lin Y (2016) Efficient encoding architecture for LDPC code based on DVB-S2 standard J Electron Inf Technol 38(7):1781–1787
5. Xie TJ, Li B,Yang M et al (2019) LDPC decoder of high speed multi-rate DVB-S2 based on FPGA. J Northwestern Polytech Univ 37(2):299–307
6. Zhang Y, Yao C, Wang Z et al (2019) An optimal correlation interval frame synchronization method for DVB-S2 system. Telecommun Eng 59(8):912–917
7. Li X, Huang Z, Zhao W et al (2022) Design and implementation of a wide-band software defined radio. Telecommun Eng 62(7):898–903
8. Hou C, Zhao L, Sun D (2021) Design of a software defined radio standard conformance test system. Telecommun Eng 61(4):511-516
9. Hongliang D, Hao L, Zhiyuan Y et al (2020) Research on synchronization mechanism based on software defined radio. Comput Measur Control 28(8):228–232
10. Zhao W, He H, Deng C (2020) Exploration on security technology of software radio. Commun Technol 53(09):2297–2300
11. Zhao W, Huang W, Li X et al (2021) Software radio receiver based on ployphase decimation filter. Commun Technol 54(6):1514–1520
12. Zhuang Z, Ban T (2021) Design of an ADS-B signal receiver based on software defined radio. Telecommun Eng 61(7):833–838

Multi-spot Beam Wiener Model Channel Capacity Limit Analysis

Yumeng Zhang, Nan Ni, Qiyun Xuan, and Jinliang Dong

Abstract The purpose of multi-beam combined with full-frequency multiplexing technology is to eliminate inter-beam interference and greatly improve the system capacity. This paper will analyze the mechanism of multi-point beam channel capacity in comparison with the ground cellular mobile communication system. This chapter studies the optimal precoding algorithm that can achieve the capacity limit of any channel state and the practically applicable MMSE precoding algorithm with lower complexity. Simulate and analyze the capacity limit in different situations.

Keywords Spot beam · Optimal linear precoding algorithm · MMSE precoding algorithm

1 Introduction

In order to provide telecommunications services corresponding to large cities to remote areas, it is not practical to use terrestrial cellular networks for coverage. Today, high-throughput satellite systems are aimed at providing fixed and mobile personal communications services that cover the world. The future of mobile communications is also moving towards the seamless integration of satellite and terrestrial communications systems. The difference between the two systems will not be Re-exist [1].

Modern satellite communication systems use multi-point beams to cover the surface of the earth. Their beams are equivalent to the cellular concept in terrestrial communication systems, but their coverage is much larger than that of terrestrial mobile communication systems. However, the physical configuration of the satellite multi-beam system is very different from the terrestrial cellular system [2]. The most obvious difference is that all user information in the satellite system must be sent to a central location, that is, the satellite ground station instead of a local Base station. Compared with terrestrial mobile communications, the more obvious features of

Y. Zhang (✉) · N. Ni · Q. Xuan · J. Dong
Nanjing Glarun Defense System Co, Nanjing, China
e-mail: 616670674@qq.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_7

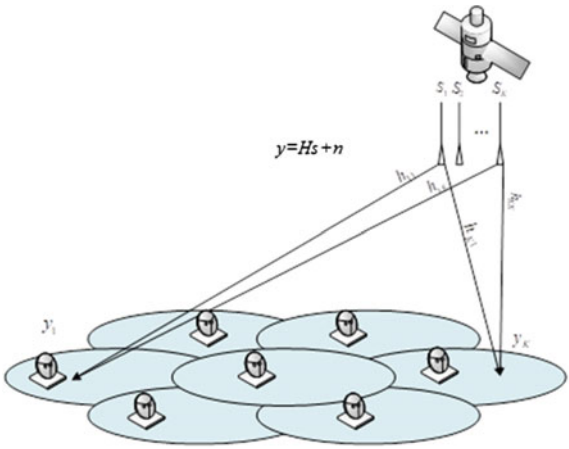


Fig. 1 Forward link downlink signal model

multi-beam satellite communications include: greater coverage, longer user signal transmission distance (causing more serious path loss), and satellite system antenna specifications are very different. Figure 1 shows the downlink signal model of the multi-beam satellite forward link.

Regarding the fading of multi-beam satellites, it is quite different from terrestrial cellular systems. Generally speaking, multi-beam satellites usually have only one antenna composed of multiple feeds. Compared with the communication distance between users and satellites, the distance between antenna feeds is negligible. This is the basic difference between multi-beam satellite fading channels and terrestrial cellular fading channels [3].

In this paper, the Wyner Gaussian access model is introduced. First, the capacity limit under non-fading channels is studied, and then the capacity limit values for the optimal linear precoding and the minimum mean square error precoding for the Wiener model and the multi-beam model are analyzed under the fading channel.

Note that in all formulas in this chapter, the logarithm to the base 2 is expressed, that is, the capacity unit is bit/s.

2 Multi-spot Beam Layout Model

Multi-point beam satellites form N beams to cover N different geographic areas, and each beam contains K users.

We assume that the system can perfectly achieve symbol and frame synchronization, so that without fading, at any given time interval l, the received signal of the multi-beam antenna can be modeled as N × 1 dimension by vector.

$$y(l) = Ax(l) + n(l) \tag{1}$$

Among them, A is N × NK dimension matrix, $n(l)$ is Gaussian noise, its mean is zero, and its variance is σ^2. $x(l)$ is the NK × 1 dimension user signal, $x_{i,k}(l)$ represents the signal of the kth user in the ith beam.

$$x(l) = \left(x_{1,1}(l), \ldots, x_{1,K}(l), \ldots, x_{N,1}(l), \ldots, x_{N,K}(l)\right)^t \tag{2}$$

Each user signal in any beam must meet the average power limit, $E\left[x_{n,k}(l)x_{n,k}^H(l)\right] \leq P$.

The N × NK dimension matrix A represents the influence of interference between beams, which can be expressed as:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,1} \\ a_{2,1} & a_{2,2} & \cdots & a_{1,1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{pmatrix} \tag{3}$$

where $a_{n,i} = \left(a_{n,i,1}, \ldots, a_{n,i,K}\right)$ is the gain of K users in the 1 × K-dimensional beam, which $0 \leq a_{n,i,k} \leq 1$ represents the path gain of beam n to user k in the ith beam, and the structure of A depends on many physical parameters. First, A depends on the topology of the system, that is, the spatial arrangement of the beam and the position of the user within the beam. Second, A depends on the radiation characteristics of the beam. The main lobe determines the isolation between adjacent satellite beams, and the side lobe characteristics are important for determining the isolation between the farther beams [4].

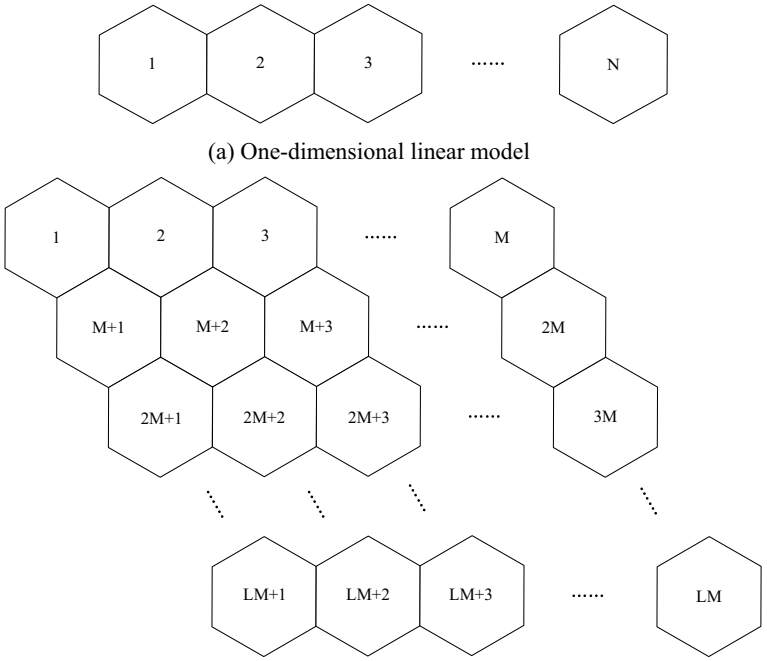
In this communication mode, we assume that K users in the beam are located near the center of the beam, so their received beam signals have the same gain coefficient. From the above analysis, we assume that $a_{i,n} = b_{i,n}1_K$ represents the effect of the n-th beam on the K users in the ith beam, so we can express the N × NK-dimensional gain matrix A as N × NK matrix B and 1 × K of all ones The kronecker product of the dimensional matrix, namely:

$$A = B \otimes 1_K \tag{4}$$

The matrix A depends on the beam coverage model. When modeling, only the adjacent beam interference is considered. The Wyner Gaussian cellular multiple access model, which we call the Wiener model, can be divided into one-dimensional Linear and two-dimensional hexagonal continuous beam arrangement (as shown in Fig. 2).

For the Wiener model, only the interference between adjacent beams is considered. For users in the designated beam, the interference signals of other beams received by them are α times of the transmitted signals of other beams, where $0 \leq \alpha \leq 1$.

For N linear beams, matrix B can be expressed as a tri-diagonal matrix when only adjacent beam interference is considered:



(a) One-dimensional linear model

(b) Two-dimensional model

Fig. 2 Wyner coverage model

$$T_{1D} = T_N \triangleq \begin{pmatrix} 1 & \alpha & & & & \\ \alpha & 1 & \alpha & & & \\ & \alpha & 1 & \ddots & & \\ & & \ddots & \ddots & \alpha & \\ & & & \alpha & 1 & \alpha \\ & & & & \alpha & 1 \end{pmatrix} \tag{5}$$

For the two-dimensional model of Wiener hexagonal arrangement, the number of beams N = LM, B can be expressed as a TBT (Toeplitz-block-Toeplitz) matrix composed of two sub-matrices of L × L and M × M dimensions [5].

$$T_{2D} = \begin{pmatrix} T_M & S_M & & & & \\ S_M^t & T_M & S_M & & & \\ & S_M^t & T_M & S_M & & \\ & & S_M^t & T_M & \ddots & \\ & & & \ddots & \ddots & S_M \\ & & & & S_M^t & T_M \end{pmatrix} \tag{6}$$

$$S_M = \alpha \begin{pmatrix} 1 & & & & \\ 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \end{pmatrix} \tag{7}$$

3 Channel Capacity Mechanism

3.1 Optimal Linear Precoding Algorithm

First, the capacity limit of the non-fading channel model is studied. We use the capacity obtained by the optimal linear precoding method as the capacity limit in any channel state. For the optimal linear precoding algorithm, the capacity limit of the non-fading channel can be expressed for:

$$C_{opt}(A;\gamma) = \max \frac{1}{N} \mathrm{I}(x;\, y|A) \tag{8}$$

Among them, the capacity limit value is the maximum capacity that can be reached by traversing the beam arrangement form and the user distribution form, but it is restricted by the average power. $\gamma = P/\sigma^2$ represents the signal-to-noise ratio (SNR), $\mathrm{I}(x;\, y|A)$ represents the amount of mutual information in the system. As we all know, the input–output mutual information of the Gaussian linear vector memoryless channel is given by the log determinant formula. Since each user's signal is independently transmitted under the constraint of average power, when the user sends an independently distributed Gaussian signal under the constraint of maximum average power, $E\left[xx^H\right] = PI_{NK}$. the standardized input and output mutual information will be maximized [6]. In the broadband transmission mode, the theoretical capacity limit can be expressed as:

$$C_{opt}(A;\gamma) = \frac{1}{N} \log\det\left(I_N + \gamma AA^H\right) = \frac{1}{N} \sum_{i=1}^{N} \log\left(1 + \gamma\lambda_i\left(AA^H\right)\right) \tag{9}$$

$\lambda_i\left(AA^H\right)$ represents the ith eigenvalue of AA^H.

For Wyner's one-dimensional linear model: $AA^H = KT_{1D}^2$, two-dimensional hexagonal model: $AA^H = KT_{2D}^2$. According to Wyner's early research, when intra-beam time division multiple access is used, the system capacity is the same as the broadband transmission scheme shown by Wyner. When the number of beams tends to infinity, the achievable capacity of a single beam no longer increases, that is, the capacity limit value can be obtained [7].

$\hat{\gamma} = K\gamma$ represents the total power of K users in a single beam. For N linear beams, the eigenvalues of T_{1D}:

$$\lambda_i(T_{1D}) = 1 + 2\alpha\cos\left(\frac{i\pi}{N+1}\right), \quad i = 1, \ldots, N \tag{10}$$

On this basis, the capacity limit expression of the one-dimensional linear beam arrangement:

$$\lim_{N\to\infty} C_{opt}(T_{1D} \otimes 1_K; \alpha, \gamma) = \int_0^1 \log\left(1 + \hat{\gamma}(1 + 2\alpha\cos 2\pi\theta)^2\right)d\theta \tag{11}$$

For a one-dimensional model applicable to a multi-beam satellite system, $\hat{T}_{1D}$ is, a model, its eigenvalues can be expressed as:

$$\lambda_i\left(\hat{T}_{1D}\right) = \frac{1-\alpha^2}{1 + 2\alpha\cos\left(\frac{i\pi+\varepsilon}{N+1}\right) + \alpha^2} \tag{12}$$

The system capacity can be expressed as:

$$C_{opt}\left(\hat{T}_{1D} \otimes 1_K; \alpha, \gamma\right) = \int_0^1 \log\left(1 + \hat{\gamma}\left[\frac{1-\alpha^2}{1 + 2\alpha\cos(\pi\theta) + \alpha^2}\right]^2\right)d\theta \tag{13}$$

When, $\alpha \to 1$ for any power in the beam, the system capacity of the multi-beam model is close to 0, because the multi-beam model considering all beam interference situations means that no matter how far apart the beams are, they will cause very serious mutual interference. This situation does not occur in the Wyner one-dimensional linear model, because each beam only interferes with the adjacent beam, resulting in T_{1D} full rank. Therefore, the multi-beam model is always limited in interference, and its ability to handle inter-beam interference is limited. When α is small, the performance of the multi-beam model and the Wyner model are basically the same.

$$\frac{1-\alpha^2}{1 + 2\alpha\cos(\pi\theta) + \alpha^2} \approx 1 - 2\alpha\cos(\pi\theta) + O\left(\alpha^2\right) \tag{14}$$

3.2 Minimum Mean Square Error Precoding Algorithm

The capacity limit of optimal linear precoding is a theoretical state. In practical applications, a MMSE precoding algorithm with lower complexity but better performance is proposed. In this section, we will derive its capacity limit and compare it with the system capacity difference obtained by the optimal linear precoding algorithm [8].

The NK × N-dimensional MMSE precoding matrix can be expressed as:

$$G = \gamma A^H \left(I_N + \gamma A A^H\right)^{-1} \tag{15}$$

The average MMSE can be expressed as:

$$mmse(A;\gamma) = P\left[\frac{1}{NK}\sum_{i=1}^{N}\frac{1}{1+\gamma\lambda_i\left(AA^H\right)} + 1 - \frac{1}{K}\right] \tag{16}$$

The system capacity with linear MMSE precoding is:

$$C_{mmse}(A;\gamma) = K\log\frac{P}{mmse(A;\gamma)} \tag{17}$$

$$C_{mmse}(A;\gamma) = -K\log\left(\frac{1}{NK}\sum_{i=1}^{N}\frac{1}{1+\gamma\lambda_i\left(AA^H\right)} + 1 - \frac{1}{K}\right) \tag{18}$$

Using the asymptotic equivalence of Toeplitz and circulant matrices, the system capacity of Wiener 1D and 2D models with MMSE precoding is given.

$$C_{mmse}(T_{1D}\otimes 1_K;\alpha,\gamma) = -K\log\left(\frac{P}{K}\int_0^1 \left(1+K\gamma(1+2\alpha\cos 2\pi\theta)^2\right)^{-1} \times d\theta + 1 - \frac{1}{K}\right) \tag{19}$$

$$C_{mmse}(T_{2D}\otimes 1_K;\alpha,\gamma) = -K\log\left(\frac{P}{K}\int_0^1\int_0^1 \left(1+K\gamma[1+2\alpha u(\theta_1,\theta_2)]^2\right)^{-1} \times d\theta_1 d\theta_2 + 1 - \frac{1}{K}\right) \tag{20}$$

The capacity of the one-dimensional linear model applied to the multi-beam satellite system using the MMSE precoding algorithm can be expressed as:

$$C_{mmse}\left(\hat{T}_{1D} \otimes 1_K; \alpha, \gamma\right) = -K \log\left(\frac{P}{K} \int_0^1 \left(1 + K\gamma\left[\frac{1-\alpha^2}{1+2\alpha\cos(\pi\theta)+\alpha^2}\right]^2\right)^{-1} \times d\theta + 1 - \frac{1}{K}\right) \tag{21}$$

4 Simulation and Performance Analysis

The precoding algorithm in this chapter adopts the optimal linear precoding algorithm and MMSE precoding. This section analyzes the achievable capacity of the two algorithms in different environments with high and low signal-to-noise ratio for non-fading channels. It aims to compare the potential limits of the two precoding algorithms and to analyze the difference between satellite communication systems and terrestrial cellular mobile systems affected by interference [9].

First, the capacity limit of the Wiener model and the multi-beam model using the optimal precoding algorithm and the MMSE precoding algorithm [10], respectively, is simulated and analyzed. The low signal-to-noise ratio environment is set to SNR = 0.15, and the high signal-to-noise ratio environment is set to SNR = 15. The number of beams formed is N = 50, and the channel bandwidth is 800 MHz.

Figure 3 compares the capacity limit of a single user per beam in the case of low signal-to-noise ratio for Wiener and multi-beam one-dimensional linear models. For these two models, when α is small, as α increases, the system capacity increases. The capacity of the model with $\alpha > 0.8$ gradually decreases with the increase of α. Therefore, the multi-beam model has an optimal level of inter-beam interference that maximizes capacity. Beyond this optimal level, the capacity will be reduced, that is, the system will be in a state where the interference between beams is limited.

Figure 4 shows the comparison of system capacity with high signal-to-noise ratio. Similarly, in the case where α is small, we find that the difference between the capacities of the two models with the same encoding method is very small, and both decrease as α increases. However, when the $\alpha > 0.4$ model uses different encoding methods, the capacity limit differs greatly. The capacity of Wiener's one-dimensional model starts to increase with the increase of α, while for the multi-beam model, the capacity continues to decline until it approaches zero. It is concluded that in the case of high signal-to-noise ratio, the multi-beam model is always interference-limited, and the best case is complete isolation between beams.

Figures 3 and 4 compare and analyze the performance of the two precoding algorithms. The capacity curve state of the two algorithms changes the same, but the capacity limit performance, the MMSE algorithm is always lower than the optimal precoding algorithm.

To sum up, we conclude that the Wiener model for cellular mobile communication systems can benefit from inter-beam interference, but the multi-beam model

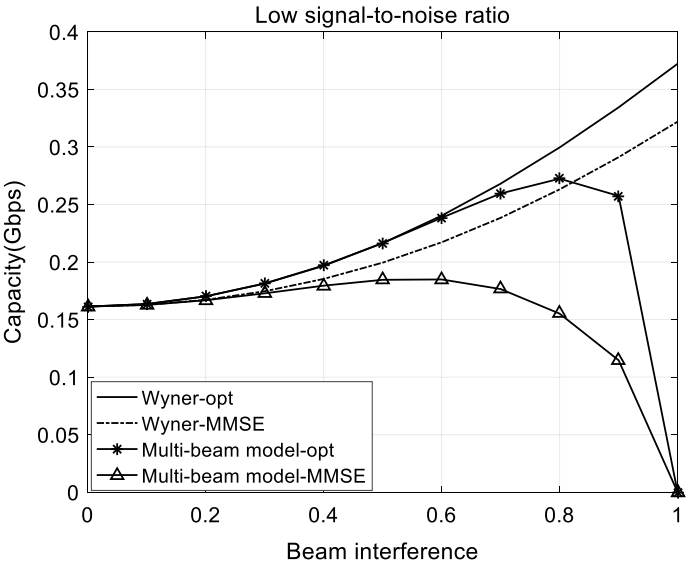


Fig. 3 Optimal linear precoding and MMSE precoding capacity for non-fading channels with low signal-to-noise ratio

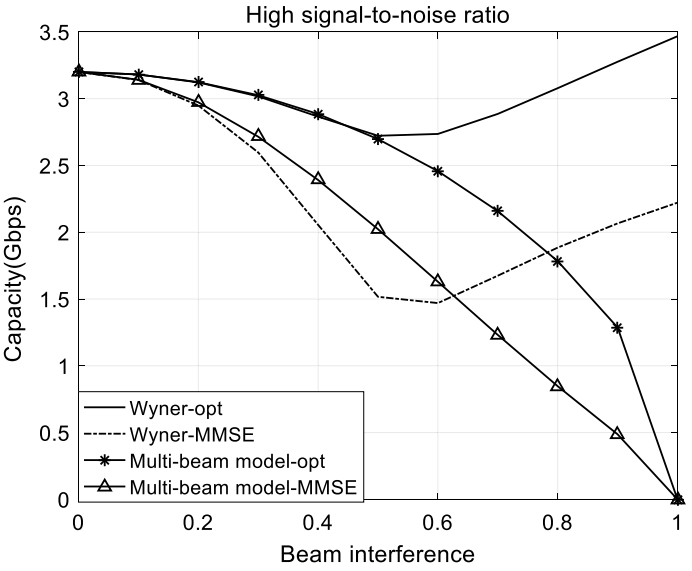


Fig. 4 Optimal linear precoding and MMSE precoding capacity for non-fading channels with high signal-to-noise ratio

for multi-beam satellite communication systems is always interference-limited, with interference levels exceeding a certain level, system performance Suddenly deteriorated. This also confirms the conclusion that the performance of the multi-beam satellite system is limited by severe beam interference. At the same time, the capacity performance of the optimal precoding algorithm is always better than the MMSE precoding algorithm.

We simulated and analyzed the two-dimensional hexagonal arrangement Wiener model using the optimal precoding algorithm and the MMSE precoding algorithm to achieve the capacity difference under high and low signal-to-noise ratio environments. The simulation analysis process also compared different channel fading conditions.

As can be seen in Fig. 5, similar to the Wiener one-dimensional model of cellular mobile communication systems, the optimal precoding algorithm and MMSE precoding algorithm in the case of low signal-to-noise ratio always achieve capacity performance from inter-beam interference. That is, the capacity increases as the beam interference level increases. However, it is different from the performance of the MMSE precoding algorithm in the Wiener one-dimensional model. With the increase of the beam interference level, the performance of the MMSE precoding algorithm is getting worse than the optimal precoding algorithm. In the case of high signal-to-noise ratio, it can be seen from Fig. 6 that the capacity performance trends of the two precoding algorithms in the two-dimensional model are the same and the optimal precoding algorithm performance is always better than the MMSE precoding algorithm under the same conditions. However, unlike the one-dimensional model, the channel attenuation of the two-dimensional model has a greater impact on capacity performance. The performance of the optimal precoding algorithm when the Rice factor is small is better than the performance of the MMSE precoding algorithm when there is no attenuation.

5 Conclusion

Through the relevant theoretical analysis and simulation verification in this paper, it can be seen that the Wiener model applicable to cellular mobile communication systems can benefit from inter-beam interference. The multi-beam model is always limited by interference, that is, the system performance deteriorates sharply after the interference level exceeds a certain level. At the same time, the capacity performance of the optimal precoding algorithm is always better than the MMSE precoding algorithm. It can be seen from the simulation of the Wiener model applicable to cellular mobile communication systems that the two-dimensional model is significantly more affected by the channel attenuation and the selection of precoding methods than the one-dimensional model.

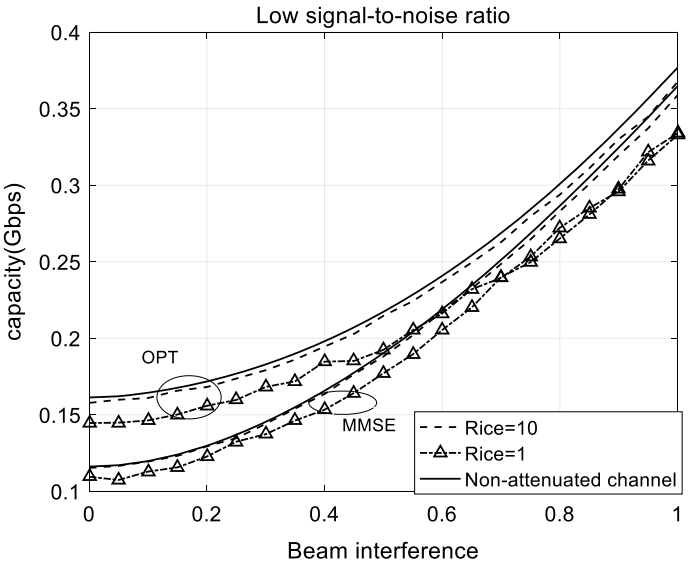


Fig. 5 Performance comparison of two precoding algorithms for Wiener 2D model with low signal-to-noise ratio

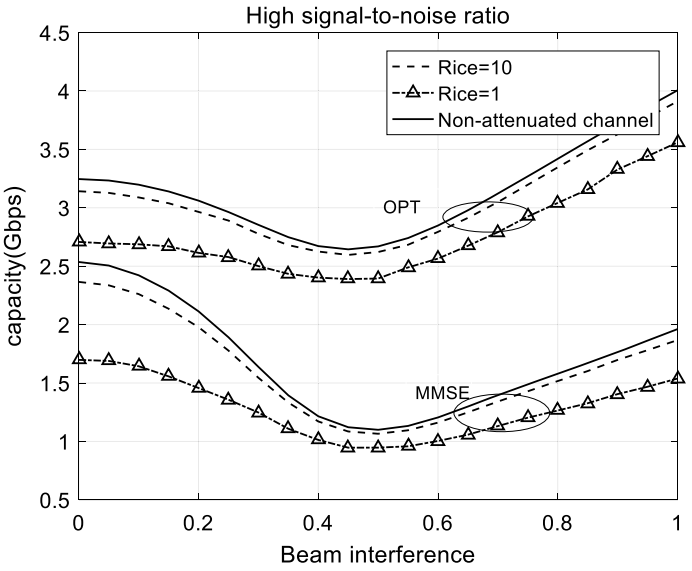


Fig. 6 Performance comparison of two precoding algorithms for Wiener 2D model with high signal-to-noise ratio

References

1. Letzepis N, Grant AJ (2008) Capacity of the Multiple spot beam satellite channel with Rician fading. IEEE Trans Inf Theory 54(11):5210–5222
2. Moher M, Erup L, Lyons RG (1996) Interference statistics for multibeam satellites. In: Proceedings of the second European workshop on mobile/personal Satcoms. Mobile and personal satellite communications, vol 2, pp 366–384
3. Taricco G (2014) Linear precoding methods for multi-beam broadband satellite systems. In: European wireless 2014, European wireless conference, Proceedings of VDE, pp 1–6
4. Caire G, Shamai S (2003) On the achievable throughput of a multiantenna Gaussian broadcast channel. IEEE Trans Inf Theory 49(7):1691–1706
5. Arnau J, Christopoulos D, Chatzinotas S et al (2014) Performance of the multibeam satellite return link with correlated rain attenuation. IEEE Trans Wireless Commun 13(11):6286–6299
6. Christopoulos D, Arnau J, Chatzinotas S et al (2013) MMSE performance analysis of generalized multibeam satellite channels. IEEE Commun Lett 17(7):1332–1335
7. Wyner AD (1994) Shannon-theoretic approach to a Gaussian cellular multiple-access channel. IEEE Trans Inform Theory 40(6):1713–1727
8. Rinaldo R, Gaudenzi RD (2004) Capacity analysis and system optimization for the reverse link of multi-beam satellite broadband systems exploiting adaptive coding and modulation. Int J Satell Commun Network 22(4):425–448
9. Vazquez MA, Perez-Neira A, Christopoulos D et al (2015) Precoding in multibeam satellite communications: present and future challenges. IEEE Wirel Commun 23(6):88–95
10. Devillers B, Perez-Neira A, Mosquera C (2012) Joint linear precoding and beamforming for the forward link of multi-beam broadband satellite systems. In: Global telecommunications conference. IEEE, pp 1–6

Main Lobe Interference Suppression Method Based on Fractional Fourier Transform

Jian Yang, Yuwei Tu, Jian Lu, Fengtao Xue, and Zhi Geng

Abstract The Fractional Fourier transform has a good energy aggregation effect for linear frequency modulation (LFM) signals commonly used in radar systems. Therefore, this paper proposes a mainlobe interference suppression method based on Fractional Fourier transform (FRFT). Firstly, the mixed radar echo signal containing main lobe interference is processed by FRFT transform with specific LFM signal characteristics, then the interference and most noise energy are removed by filtering in the FRFT domain. Finally, FRFT inverse transformation recovers the target signal. Simulation verifies the effectiveness of the algorithm.

Keywords Anti-mainlobe interference · Fractional Fourier transform · FRFT domain filtering · Pulse compression

1 Introduction

The anti-jamming performance of radar is increasingly important in modern electronic warfare. Mainlobe interference reduces the output SINR of radar, increases false alarm rate, and makes it difficult for previous interference suppression algorithms against sidelobe interference to suppress mainlobe interference, which may cause mainlobe distortion and suppression of desired signals as interference. As one of the important means of electronic jamming, mainlobe interference can easily

J. Yang (✉) · Y. Tu
Xi'an Research Institute of Hi-Tech, Hongqing Town, Xi'an 710025, China
e-mail: yangjian@nudt.edu.cn

J. Lu
96921 Troops of the Chinese People's Liberation Army, Beijing 100001, China

F. Xue
Beijing Remote Sensing Equipment Research Institute, Beijing 100854, China

Z. Geng
Logistics Information Center, Former Joint Logistics Department of Chengdu Command, Chengdu 610015, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_8

deceive target distance, velocity, angle and other parameters, seriously affecting the normal operation and detection performance of radar system. In paper [1], a mainlobe interference suppression method based on low sidelobe constraint blocking matrix preprocessing (BMP) is proposed. The Angle of mainlobe interference is estimated using the singular value decomposition (SVD) method, and a blocking matrix is constructed to suppress mainlobe interference. In paper [2], the main lobe interference suppression technique of wideband noise based on space–time adaptive processing (STAP) is studied. A mainlobe interference suppression method based on oblique projection of eigenvalues is proposed in [3]. The eigenvalue oblique projection preprocessing matrix is constructed to filter out the main lobe interference, and then the influence of covariance mismatch matrix is reduced by diagonal loading. In [4], a four-channel monopulse technique for suppressing mainlobe interference was proposed. In Ref. [5], a method for anti-deception interference of an array radar was proposed. The idea based on FRFT transformation mentioned in the paper was applied to suppress the main lobe interference proposed in this paper. We propose a FRFT-based radar's primary lobe jamming suppression technique that uses FRFT transformation's special feature for linear frequency modulation signal. The algorithm can effectively suppress jamming and improve peak signal-to-noise ratio after pulse compression according to theoretical analysis. The performance of the proposed method is verified through comprehensive performance analyses and simulation results.

1.1 Signal Model

Using a linear frequency modulation signal as the transmission signal for a radar system, its mathematical expression in the time domain is (Fig. 1):

$$s(t) = \mathrm{rect}(t \Big/ T_p)\mathrm{e}^{\mathrm{j}2\pi f_c + \mathrm{j}\pi\kappa t^2} \tag{1}$$

In the formula, $s(t)$ is the radar transmission signal, $\mathrm{rect}(t \Big/ T_p)$ represents the rectangular pulse function, T_p represents the pulse width, f_c represents the signal carrier frequency, and κ represents the modulation frequency.

The mathematical model of the target in the radar echo needs to consider the time delay and frequency shift caused by the target distance and movement, then the mathematical model of the target echo signal is:

$$s_1(t) = s_0(t - \tau_1)\mathrm{e}^{-\mathrm{j}2\pi f_c \frac{2v_{r1}}{c}t} \tag{2}$$

In the formula, $s_1(t)$ is the echo signal of the target, τ_1 is the time delay of the echo signal. Because there is a distance between the radar and the target, there will be a time delay between the echo signal and the radar transmitted signal. The relationship between the time delay τ_1 of the target and the distance r_1 of the target is $\tau_1 = \frac{2r_1}{c}$.

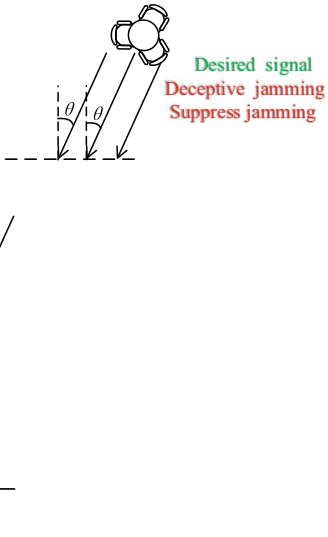


Fig. 1 Radar antenna signal reception schematic diagram

1.2 Principles Related to FRFT

In 1980, Namias first proposed the definition of Fractional Fourier transform from a purely mathematical perspective in terms of eigenvalues and eigenfunctions [6]. Subsequently, McBride provided a more rigorous mathematical definition for Fractional Fourier transform in integral form [7]. Below we give the basic definition of Fractional Fourier transform from the perspective of integral transformation.

The p-th order Fractional Fourier transform of a time-domain signal x(t) defined in the t-domain is a linear integral operation:

$$F^p(u) = \int_{-\infty}^{+\infty} K_p(u,t)x(t)dt \tag{3}$$

In the formula, $K_p(u,t)$ is the kernel function of FRFT, and p represents the transformation order, which can be specifically written as

$$K_p(t,u) = \begin{cases} \sqrt{(1-j\cot\alpha)/(2\pi)}\exp\left(j\frac{u^2+t^2}{2}\cot\alpha - jut\csc\alpha\right), & \alpha \neq n\pi \\ \delta(t-u), & \alpha = 2n\pi \\ \delta(t+u), & \alpha = (2n+1)\pi \end{cases} \tag{4}$$

After variable substitution $u = u/\sqrt{2\pi}$ and $t = t/\sqrt{2\pi}$, Formula (3) can be further written as

$$F^p(u) = \{ F^p[x(t)](u)\} = \int_{-\infty}^{+\infty} K_p(u,t)x(t)dt, \quad 0 < |p| < 2, 0 < |\alpha| < 2$$

$$= \begin{cases} {}_\alpha \exp(j\frac{u^2+t^2}{2}\cot\alpha - jut\csc\alpha)x(t)dt, & \alpha \neq n\pi \\ x(t), & \alpha = 2n\pi \\ x(-t), & \alpha = (2n+1)\pi \end{cases} \tag{5}$$

where $A_\alpha = \sqrt{(1 - j\cot\alpha)/(2\pi)}$, the definition of FRFT given by formula (5) is linear but not shift-invariant (except for $p = 4n$), because the kernel function is not only a function of (u, t) but also a function of p. It is worth noting that F^{4n} and F^{4n+2} are equivalent to the identity operator I and parity operator P respectively. For $p = 1, \alpha = \pi/2, A_\alpha = 1$, and:

$$F^1(u) = \int_{-\infty}^{+\infty} \exp(j2\pi ut)x(t)dt \tag{6}$$

It can be seen that $F^1(u)$ is the ordinary Fourier transform of $x(t)$, and similarly, $F^{-1}(u)$ is the ordinary inverse Fourier transform of $x(t)$.

Linear frequency modulation signals are typical non-stationary signals with large time-bandwidth products, widely used in radar, communication, sonar detection and other fields, especially in radar systems. As a linear transform that decomposes signals into chirp bases without cross-term interference, Fractional Fourier transform is particularly suitable for processing chirp-like signals. Analogous to how any vector on a two-dimensional plane can be decomposed into two basis vectors along x and y axes, any LFM signal can also be decomposed into two chirp bases. By choosing an appropriate order p, FRFT can concentrate the energy of the chirp signal on the basis vector along this direction, forming an impulse pulse, whereas conventional Fourier transform usually does not align with this "optimal direction", resulting in more dispersed energy with a certain bandwidth. Therefore, FRFT has good energy aggregation characteristics for LFM signals in some fractional transform domains [8].

2 Mainlobe Interference Suppression Algorithm

Assume that the time-domain representation of an arbitrary LFM signal is expressed as $\phi(t) = \exp(\mathrm{j}2\pi f_0 t + \mathrm{j}\pi\kappa t^2)$. Then, the p th order FRFT transform of LFM signal can be written as [9]:

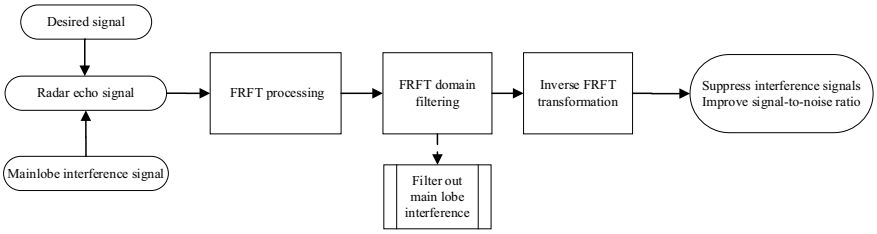


Fig. 2 Flowchart of main lobe interference suppression algorithm based on FRFT

$$F^p[x_{LFM}(t)](u) = F^p[\phi(t)](u)$$
$$= \begin{cases} \sqrt{\frac{1+\mathrm{j}\tan\alpha}{1+\kappa\tan\alpha}}\exp\left[\mathrm{j}\pi\frac{(\kappa-\tan\alpha)u^2+2uf_c\sec\alpha-f_c^2\tan\alpha}{1+\kappa\tan\alpha}\right], & \alpha-\arctan\kappa\neq(2i+1)\pi/2 \\ \sqrt{\frac{1}{1-\mathrm{j}\kappa}}\exp[\mathrm{j}\pi\cos\alpha(-f_c^2\sin\alpha+2uf_c)]\delta(u-f_c\sin\alpha), & \alpha-\arctan\kappa=(2i+1)\pi/2 \end{cases} \tag{7}$$

Utilizing the characteristic of linear frequency modulation signals in Fractional Fourier transform, we propose a radar mainlobe interference suppression technique based on FRFT. The received mainlobe interference mixed echo signal is first processed by FRFT, followed by filtering in the FRFT domain to remove most of the suppressed interference and noise energy. The target signal is then recovered by inverse FRFT. Theoretical analysis suggests that this algorithm can effectively suppress interference and significantly improve the peak signal-to-noise ratio after pulse compression. By analyzing the phase relationship between peak values of continuous pulse echoes and improving the signal-to-noise ratio of fractional domain echo signals through coherent accumulation, it is possible to extract peak values of LFM signals under low signal-to-noise ratio conditions. The design method process is shown in the Fig. 2.

Since both deceptive interference signals and target echoes are wideband LFM signals with the same modulation rate κ, for ease of analysis, we first analyze the phase relationship in the fractional domain of the target echo signals received by radar antenna after p_0 th order FRFT processing. It can be derived that by representing the target echo component in the m-th echo received by the radar antenna as a phase-shifted and frequency-shifted form of an LFM signal.

$$s_0(t_m,\tilde{t}) = \chi_m\gamma(m)\exp(-\mathrm{j}2\pi f_m\tilde{t}+\mathrm{j}\pi\kappa\tilde{t}^2) \tag{8}$$

where,

$$\chi_m = \mathrm{rect}\left[\frac{\tilde{t}-\tau(t_m)}{T_p}\right], \quad f_m = \kappa\tau(t_m)$$
$$\gamma(m) = \exp\{\mathrm{j}2\pi f_c[-\tau(t_m)]+\mathrm{j}\pi\kappa[\tau(t_m)]^2\} \tag{9}$$

The above equation can be regarded as the result of an LFM signal with a carrier frequency of f_m and a modulation rate of κ undergoing data truncation and linear transformation. In the actual digital sampling process, the radar antenna uses the same distance gate for digitizing samples of the same echo.

According to the aforementioned analysis, FRFT has many excellent properties such as reversibility, linearity, exponential additivity, commutativity, associativity, time-shift property and frequency-shift property [7–10]. Utilizing these properties, we analyze the phase relationship in fractional domain of target echo signals received by radar antenna. Since $\tan\alpha_0 = \kappa \gg 1$, then $\kappa\cos\alpha_0 = \sin\alpha_0 \approx 1$. According to Eqs. (7) and (8), the fractional domain form $F^{p_0}[\breve{s}_0(t_m, \tilde{t})](u)$ can be further expressed as:

$$\begin{aligned}
&F^{p_0}[\breve{s}_0(t_m, \tilde{t})](u) \\
&= A_\kappa \gamma(m) \exp\left\{-\mathrm{j}\pi \frac{\kappa^3}{\kappa^2+1}[\tau(t_m)]^2\right\} \cdot \exp(-\mathrm{j}2\pi f_m \cos\alpha_0 u) \\
&\approx A_\kappa \gamma(m) \exp\left\{-\mathrm{j}\pi \frac{\kappa^3}{\kappa^2+1}[\tau(t_m)]^2\right\} \cdot \exp\{-\mathrm{j}2\pi[\tau(t_m)]u\} \\
&\approx A_\kappa \gamma(m) \exp\{-\mathrm{j}\pi\kappa[\tau(t_m)]^2\} \cdot \exp\{-\mathrm{j}2\pi[\tau(t_m)]u\} \\
&= A_\kappa \exp\{-\mathrm{j}2\pi f_c[\tau(t_m)\} \cdot \exp\{-\mathrm{j}2\pi[\tau(t_m)]u\}
\end{aligned} \tag{10}$$

wherein, the transformation order p_0 is a specific transformation order that converts an LFM signal into a complex single-frequency point signal in a fractional Fourier domain. That is, after performing FRFT with an order of p_0, the LFM signal becomes a sine wave with a frequency of $f = f_0 \cos\alpha_0$ [9]. In digital processing, its specific transformation order is expressed as: $p_0 = \frac{2}{\pi}\arctan\left(\frac{N_g\kappa}{f_s^2}\right)$.

Since discrete FRFT operations also require dimension normalization processing, this paper adopts discrete scaling proposed in literature to perform scale transformation[11]. By decomposing the above equation according to the radar antenna received signal model, we obtain the fractional domain data $\mathbf{S}_m(u)$ obtained by processing the target signal of sampling mth echo with p_0-order FRFT as follows:

$$\mathbf{s}_m(u) = A_\kappa \exp[-\mathrm{j}2\pi\tau(t_m)(f_c+u)] = A_\kappa \exp[-\mathrm{j}2\pi f_c\tau(t_m) - \mathrm{j}2\pi\tau(t_m)u)] \tag{11}$$

From Eq. (11), it can be seen that the incident signal of the m-th echo received by the radar is equivalent to the fractional domain signal $\mathbf{s}_m(u)$, and the fractional domain frequency of signal $\mathbf{s}_m(u)$ is determined by time delay $\tau(t_m)$. According to the dimension normalization method in literature [11], the fractional domain frequency corresponding to signal $\mathbf{s}_m(u)$ can be expressed as:

$$f_{nu} = \left| \frac{\frac{\tau_{\max}+\tau_{\min}}{2} - \tau(t_m)}{\sqrt{\frac{\tau_{\max}-\tau_{\min}}{f_s}}} \right| \tag{12}$$

In the equation, $|\cdot|$ represents taking absolute value, $\tau_{\min}$ and $\tau_{\max}$ respectively represent the time delay corresponding to the minimum value $R_{\min}$ and maximum value $R_{\max}$ of distance gate. Since the frequency of target echo signal in fractional domain is determined by target's time delay relative to center of gate, prior information about target position is usually known before transmitting broadband signal. For example, it is known that target distance interval is $[r_{\min}, r_{\max}] \subset [R_{\min}, R_{\max}]$, and assuming that two distance intervals have a common center. Accordingly, the highest frequency of target echo in fractional domain is:

$$f_{mu} = \frac{\frac{\tau_{\max}+\tau_{\min}}{2} - \frac{2r_{min}}{c}}{\sqrt{\frac{\tau_{\max}-\tau_{\min}}{f_s}}} \tag{13}$$

Therefore, if the echo signal of the interference target is outside the distance interval $[r_{\min}, r_{\max}]$, the frequency of the interference signal in the fractional domain will be greater than f_{mu}, and the interference signal can be filtered out by a low-pass filter in the fractional domain.

3 Simulation Results

A desired signal and a noise suppression interference signal were set up along with two deception interference signals. All the signals are coming from the same direction. The velocities of interference signals 1 and 2 were set to − 10 m/s and − 15 m/s respectively, while the velocity of the desired signal was set to 15 m/s. The signal-to-noise ratio was set as 0 dB, and the interference noise ratio of two deceptive interferences and one suppressing interference was set as 18 dB,18 dB and 35 dB respectively (Fig. 3).

From the time–frequency diagrams before and after algorithm processing, it can be seen that the algorithm has clearly filtered out most of the interference in the frequency band after processing, especially suppressing interference.

By processing the echo signal and performing pulse compression and coherent processing, an R-D plane is obtained for moving target detection and parameter estimation. The algorithm simulation results are shown in the Fig. 4.

According to the simulation results, the proposed algorithm can significantly suppress main lobe interference from two signals in the same direction as the desired signal, thereby making the desired signal more prominent and significantly improving the signal-to-noise ratio.

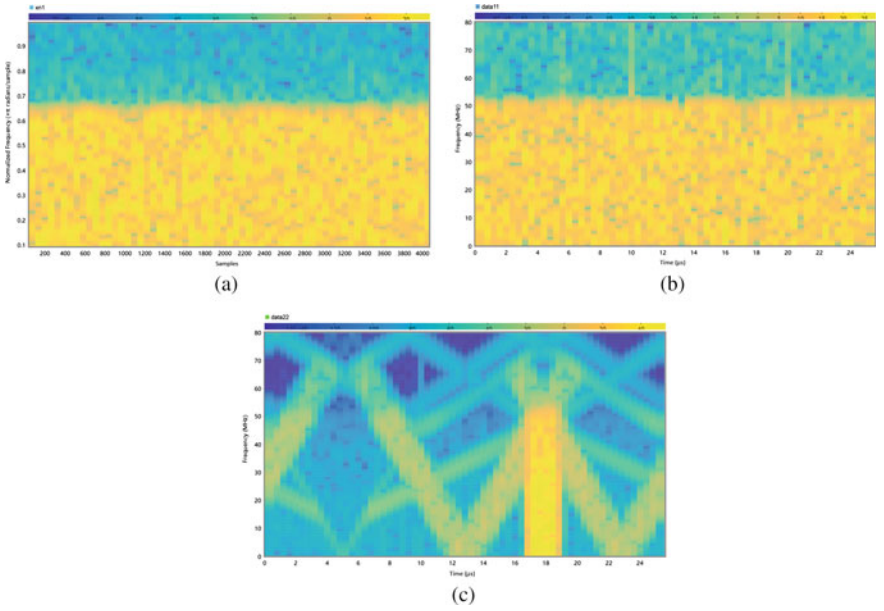


Fig. 3 The time–frequency diagram of the signal: **a** noise suppression interference signal time–frequency distribution diagram, **b** time–frequency distribution of transmitted signal, **c** time–frequency distribution of the signal after algorithm processing

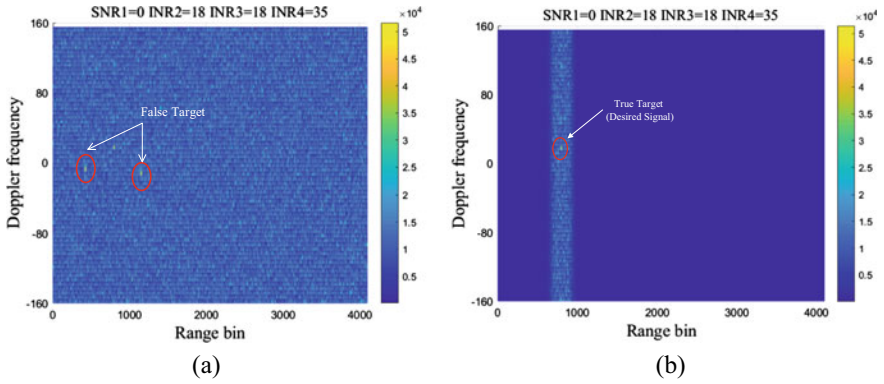


Fig. 4 **a** R-D distribution of simulated echo received signal **b** R-D distribution of signal after algorithm processing

4 Conclusions

In this paper, we propose an algorithm that involves FRFT processing of the received mixed echo signal with primary lobe jamming. Most of the energy of suppressed jamming and noise in FRFT domain is filtered out. Simulation experiments show that the algorithm can effectively improve the estimation accuracy under the condition of low signal-to-noise ratio with the presence of the main lobe interference. The algorithm can improve the peak signal-to-noise ratio after pulse compression and improve the detection performance of the pulse compression radar.

Acknowledgements This research was funded by National Natural Science Foundation of China, grant numbers 62071481 and 61501471.

References

1. Haoyu M, Xiaodong Q, Xingyu Z, Wolin L, Zhengyan Z, Xiaopeng Y (2022) Mainlobe interference suppression method based on blocking matrix preprocessing with low sidelobe constraint. In: 2022 Asia-Pacific signal and information processing association annual summit and conference (APSIPA ASC), Chiang Mai, Thailand, pp 2065–2070. https://doi.org/10.23919/APSIPAASC55919.2022.9980066
2. Liao Y, Wu X, Wang Y (2018) The main lobe interference suppression algorithm for wideband interference. In: 2018 14th IEEE International conference on signal processing (ICSP), Beijing, China, pp 148–152. https://doi.org/10.1109/ICSP.2018.8652409
3. Meng H, Yang X, Gao S, Yu Z (2022) Main lobe interference suppression method based on eigenvalue oblique projection. J Signal Process 38(2):439–444
4. Xu A, Zhao C, Xue J, Li L, Meng F (2019) Four-channel monopulse technique for main-lobe interference suppression. In: 2019 IEEE 8th Joint international information technology and artificial intelligence conference (ITAIC), Chongqing, China, pp 1354–1357. https://doi.org/10.1109/ITAIC.2019.8785875
5. Yang J, Lu J, Tu Y et al (2022) Spatial deception suppression for wideband linear frequency modulation signals based on fractional Fourier transform with robust adaptive beamforming. Digital Signal Process 126:103485
6. Namias V (1980) The fractional Fourier transform and its application in quantum mechanics. J Inst Math Appl 25(3):241–265
7. McBride AC, Kerr FH (1987) On Namias' fractional Fourier transforms. IMA J Appl Math 39(2):159–175
8. Cowell DM, Freear S (2010) Separation of overlapping linear frequency modulated (LFM) signals using the fractional Fourier transform. IEEE Trans Ultrason Ferroelectr Freq Control 57(10):2324–2333. https://doi.org/10.1109/TUFFC.2010.1693. PMID: 20889420
9. Ma Y, Wang R, Du J (2013) Amplitude characteristics of linear frequency modulation signal in FRFT domain. In: 2013 IEEE international conference on signal and image processing applications, pp 431–434. https://doi.org/10.1109/ICSIPA.2013.6708045
10. Ozaktas HM et al (1996) Digital computation of the fractional Fourier transform. IEEE Trans Signal Process 44(9):2141–2141
11. Zhao XH, Deng B, Tao R (2005) Dimensional normalization in the digital computation of the Fractional Fourier transform. Trans Beijing Inst Technol 25(4):360–364

Mangrove Species Classification in Qi'ao Island Based on Gaofen-2 Image and UAV LiDAR

Yuchao Sun, Zheng Wei, Yang Gao, Hongkai Ren, Qidong Chen, Di Dong, and Ping Hu

Abstract Mangrove species classification is of great significance to the study of mangrove community structure and biodiversity. Most researches use foreign high-resolution remote sensing images or UAV images for mangrove species classification. In order to improve classification accuracy, LiDAR and hyper-spectral data are often used to assist classification. In this paper, based on the Gaofen-2 image and the CHM data obtained by the UAV Lidar, the mangroves in Qi'ao Island, Zhuhai are classified among species by using the random forest classification method. The classified species include 5 types of true mangroves, 3 types of semi mangroves, *Phragmites australis* and non-vegetation. The results show that the use of Gaofen-2 image can only effectively distinguish the *Sonneratia apetala*, *Acrostichum aureum* and non-vegetation, the accuracy of distinguishing other mangrove species is not ideal; After Gaofen-2 image fusion of CHM data, the classification accuracy of each mangrove species has been significantly improved, with the overall classification accuracy reaching 91.44%, which verifies the effectiveness of Gaofen-2 image fusion of external data in mangrove species classification research.

Keywords Mangroves · Qi'ao Island · Species classification · Gaofen-2 Image · CHM

Y. Sun · Z. Wei (✉) · Y. Gao · Q. Chen · D. Dong · P. Hu
South China Sea Institute of Planning and Environmental Research, State Oceanic Administration, Guangzhou 510300, China
e-mail: weizheng0628@foxmail.com

Key Laboratory of Marine Environmental Survey Technology and Application, Ministry of Natural Resources, Guangzhou 510300, China

H. Ren
Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519082, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_9

1 Introduction

Mangrove is a woody plant community that grows in intertidal zone of tropical and subtropical coast and is periodically submerged by seawater. It is an important coastal wetland ecosystem, and plays an important role in purifying seawater quality, protecting seawall, maintaining ecological balance and biodiversity [1, 2]. Due to the special habitat in the intertidal zone, field survey of mangrove is quite difficult. In recent years, remote sensing technology combined with field survey is commonly used to extract the mangrove distribution range [3], species classification [4] and biomass estimation [5, 6]. Mangrove species classification is of great significance to the study of mangrove community structure and biodiversity, but it also puts forward higher requirements for spectral resolution and spatial resolution of remote sensing images [7].

Mangrove species classification is mainly based on supervised classification methods, which carried out by combining the image spectral features from the band index and texture features [7]. While the spectral features were usually extracted from the band index and the texture features were from the gray level co-occurrence matrix. To improve the classification accuracy, many studies used hyperspectral images such as Gaofen-5 [8] and radar images such as RADARSAT-2 [9] or combined multiple types of remote sensing image data to carry out mangrove species classification [10]. With the wide application of unmanned aerial vehicle (UAV), more and more UAV remote sensing data are also applied to the mangrove species classification. Using the light detection and ranging (LiDAR) to assist the UAV orthophoto image can effectively improve the accuracy of classification [11, 12]. The combination of mangrove canopy height model (CHM) extracted from UAV Lidar and UAV hyperspectral images can also improve the classification accuracy [13].

In the past decade, most of the data sources used for mangrove species classification are mainly foreign high-resolution images, which with high cost. In recent years, domestic high-resolution images such as Gaofen series and Ziyuan series have been gradually applied to various applications due to their high resolution and low cost, and have also been applied to the mangrove species classification [14]. However, due to the small number of spectral bands and other reasons, it is difficult to obtain ideal classification accuracy. To improve the classification accuracy of domestic high-resolution images, this paper takes the mangrove reserve of Qi'ao Island in Zhuhai as an example, fuses the Gaofen-2 image with the UAV LiDAR data and discusses the adaptability of the domestic high-resolution image in mangrove species classification.

2 Study Area and Data Sets

2.1 Study Area

Qi'ao Island is in the northeast of Zhuhai city and on the west bank of the Pearl River Estuary, covering an area of about 24 km^2. It has a subtropical monsoon climate, with an annual average temperature of 24.5 °C, the lowest temperature in January and the highest temperature in July. The average annual sunshine hours can reach 1907.4 h, the average annual precipitation is about 1964.4 mm, mainly concentrated in April to September, and the average annual relative humidity of the air is about 79%. The tide in this area is irregular semidiurnal tide, and the annual average value of seawater salinity is 18.4 ‰, belonging to the coastal saline meadow marsh soil.

According to the field survey, the mangroves of Qi'ao Island are mainly distributed in the northwest and west of Qi'ao Island (Fig. 1). There are 10 families, 13 genera and 15 species of true mangroves in Qi'ao Island, mainly composed of *Sonneratia apetala* (SA), *Acrostichum aureum* (AA), *Acanthus ilicifolius* (AI), *Kandelia candel* (KC) and *Aegiceras corniculatum* (AC), among which SA community is the absolute dominant community. There are 7 families, 9 genera and 9 species of semi-mangroves, mainly including *Heritiera littoralis* (HL), *Bruguiera gymnorrhiza* (BG, mainly refer to seedling of BG) and *Hibiscus tiliaceus* (HT). In addition, there are many reeds (RE) associated with mangrove growing areas [15, 16]. According to the field survey and measurement, there are obvious differences in the height of mangrove species (Table 1), and the distribution of mangroves is characterized by clustering in species [17].

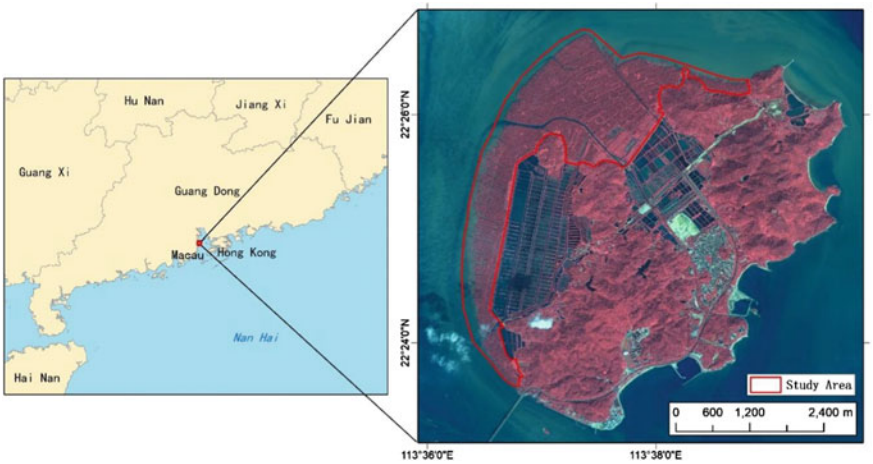


Fig. 1 Location of Qi'ao island and region of study area

Table 1 Height of different mangrove species

Species	SA	AA	AI	KC	AC	HL	BG	HT	RE
Height (m)	10–20	1–2	1–2	3–5	2–3	6–10	2–3	6–8	1–2

2.2 Data Sets

2.2.1 Gaofen-2 Image

Gaofen-2 satellite is the first civil optical remote sensing satellite with a spatial resolution better than 1 m independently developed by China, which was launched in August 2014. Gaofen-2 carries two high-resolution spectral sensors: 1 m panchromatic and 4 m multispectral cameras, has the characteristics of high radiation accuracy, high positioning accuracy and rapid attitude mobility. Gaofen-2 images are widely used in natural resource survey, agricultural crop yield estimation, water conservancy flood facilities and disaster monitoring [18, 19].

We obtained the Level-1A product of Gaofen-2 image on November 1, 2021, which covers the whole Qi'ao Island. We used ENVI 5.3 software for radiometric calibration, and the "FLAASH Atmospheric Correction" tool for atmospheric correction, the 30 m resolution DEM data released by the National Aeronautics and Space Administration (NASA) in February 2020 was used for orthophoto correction, and the "Gram-Schmidt Pan Sharpening" tool was used for image fusion to obtain an image with a resolution of 1 m and four multispectral bands.

2.2.2 UAV LiDAR Data

The UAV Lidar data is obtained by Shenzhen FEIMA V10 UAV system with DV-LiDAR20 airborne laser. The FEIMA V10 UAV has a length of about 175 cm, a wingspan of 415 cm, a takeoff weight of 25 kg, a maximum load of 6 kg, a cruise speed of 20 m/s, a endurance time of about 160 min (6 kg load), a wind resistance level of 6, a standard configuration of network RTK/PKK and its integrated solution service (Fig. 2a). The FEIMA DV-LiDAR20 laser system weighs about 3.5 kg and has a size of 216 mm × 384 mm × 166 mm, the maximum range is 1350 m, the wavelength is 1550 nm, and the transmission frequency is 50–550 kHz, which integrates a 42 megapixel full-frame camera with a focal length of 18 mm (Fig. 2b).

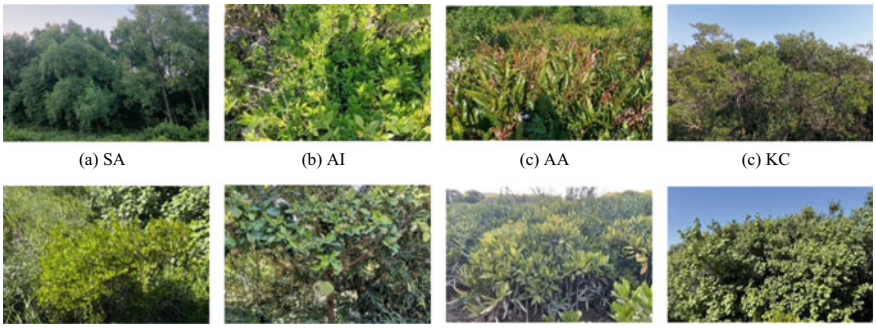
The Lidar data was obtained in November 17, 2021, the flight altitude was 300 m, and the elevation accuracy was about 7 cm. We used FEIMA "UAV Manager" and LiDAR360 Software for pre-processing, used "denoising" tool for point cloud denoising, used "classification" tool for ground point and vegetation point classification, exported digital elevation model (DEM) with a resolution of 1 m based on ground point, and digital surface model (DSM) with the same resolution based on vegetation point, canopy height model (CHM) was calculated by difference between DSM and DEM.

(a) FEIMA V10 UAV (b) FEIMA DV-LiDAR20

Fig. 2 FEIMA UAV LiDAR system

2.2.3 Mangrove Species Data

To understand the distribution characteristics of mangrove species in Qi'ao island and collect the sample data, a field survey was carried out in December 10, 2021. We mainly used camera and GPS to take photos of different mangrove species and record their coordinates. We established position and mangrove species association based on GPS data and remote sensing image (Fig. 3).



(a) SA (b) AI (c) AA (c) KC

(e) AC (f) HL (g) BG (h) HT

Fig. 3 Photos of different mangrove species in Qi'ao island

3 Methods

3.1 Mangrove Species Samples Selection

According to the field photos and GPS data, we established remote sensing interpretation marks of different mangrove species based on the GaoFen-2 image, and selected sample points data for different mangrove species on the GaoFen-2 image (Table 2). Non-vegetation (NV) refers to water, bare ground and artificial buildings, half of the samples data are used as training data and half as validation data.

3.2 Random Forest Classification

There are many algorithms applied to mangrove species classification, including support vector machine (Ting Wang et al., 2015), random forest [8, 9] and other supervised classification methods [20] based on pixel or object-oriented. Among them, the RF algorithm proposed by Breiman [21], has been proved to have high prediction accuracy, good tolerance for outliers and noise, and is not prone to overfitting [22], which is widely used in mangrove and other vegetation classification [23]. The RF algorithm is an ensemble algorithm for supervised classification based on classification and regression trees (CART). By combining the characteristics of CART, bootstrap aggregating, and random feature selection, independent predictions can be established and therefore improve accuracy.

Before RF classification, we used object-oriented method to segment and merge GaoFen-2 image in ENVI 5.3 software. Through experiments, we got the best object-oriented segmentation effect when the segmentation threshold is 60 and the merge threshold is 20. RF classification was also carried out in ENVI 5.3 software after segmentation, we selected the training data in Sect. 3.1 as sample data, the number of trees was set to 100, the number of features was set to square root, the impurity function was set to Gini coefficient.

Table 2 Mangrove species samples data

Species	SA	AA	AI	KC	AC	HL	BG	HT	RE	NV	Total
Training data	55	20	20	20	15	20	20	20	30	40	260
Validation data	55	20	20	20	15	20	20	20	30	40	260

3.3 Mangrove Species Classification Based on GaoFen-2 Image

We conducted the object-oriented RF classification for GaoFen-2 image using the method in Sect. 3.2. Considering the rich texture information in GaoFen-2 image and large texture difference among mangrove species [24], we extracted the texture features by calculating Grey Level Co-occurrence Matrix (GLCM) for blue band (Band3) of GaoFen-2 image [25]. The GLCM parameter included mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment and correlation. By combining the GLCM bands with the spectral bands of GaoFen-2 image, the object-oriented RF method was used for classification.

3.4 Mangrove Species Classification Based on GaoFen-2 and CHM

Due to the high spatial resolution and less spectral band, GaoFen-2 image combined with hyperspectral images, synthetic aperture radar (SAR) images or LiDAR should be adopted to achieve more accurate classification results [26]. Considering the obvious differences of tree height among mangrove species in Qi'ao Island, which can be reflected in the CHM data of LiDAR, the combination of GaoFen-2 image (including GLCM) and CHM data should effectively improve the classification accuracy. The object-oriented RF classification was also applied to combined data.

3.5 Accuracy Assessment

Confusion matrix is used to evaluate the classification accuracy. The evaluation indicators of confusion matrix include Overall Accuracy (OA), Kappa coefficient (Kappa), Producer Accuracy (PA) and User Accuracy (UA). OA represents the proportion of the number of correctly classified samples to the total number of samples; Kappa reflects the consistency between classification results and reference samples; PA represents the proportion of the number of samples correctly classified to the total number of reference samples; UA represents the proportion of the number of samples correctly classified to the total number of true samples.

4 Results and Discussion

4.1 Results

The object-oriented RF method was adopted to classify the GaoFen-2 image, GaoFen-2 + GLCM data and GaoFen-2 + GLCM + CHM data refer to Sects. 3.2–3.5. The validation data in Sect. 3.1 was used to calculate the classification accuracy (Table 3).

4.2 Discussion

Texture information and CHM data can effectively improve the classification accuracy of GaoFen-2 image according Table 3, By fusing texture information and CHM data, the OA of GaoFen-2 image reached 84.88% from 74.50%. From Table 4, using the spectral band of GaoFen-2 image can distinguish SA, BG and NV (whose PA and UA are both above 70%) due to the spectral difference of mangrove species. By comparing Tables 4 and 5, the PA and UA of AA, KC and PA are obviously improved after fusing texture information due to the textural difference of mangrove species. By comparing Tables 5 and 6, the PA and UA of AI, HL and HT have greatly

Table 3 Classification accuracy of different GaoFen-2 image

Input data	OA (%)	Kappa	Confusion matrix
GaoFen-2 image	74.50	0.7103	Table 4
GaoFen-2 + GLCM	80.92	0.783	Table 5
GaoFen-2 + GLCM + CHM	84.88	0.8278	Table 6

Table 4 Confusion matrix of GaoFen-2 image

%	SA	AA	AI	KC	AC	HL	BG	HT	RE	NV	PA	UA
SA	90.85	1.13	0	4.68	0	5.56	0	0.56	6.3	0	90.85	92.16
AA	0	63.28	1.69	0	0	2.78	1.67	0	5.93	0	63.28	80.58
AI	0	0	61.02	0	0	7.22	5.56	26.11	5.93	0	61.02	55.67
KC	5.49	0	0	91.81	42.22	10	1.67	3.33	0	0	91.81	58.58
AC	1.42	28.81	0	1.75	52.59	16.11	0	0.56	5.19	0	52.59	40.34
HL	0	0	0	0.58	0.74	41.67	5	0.56	11.85	0	41.67	63.03
BG	0	0	20.34	0	0	0	81.11	8.89	0	0	81.11	73.74
HT	0	0	3.95	0	1.48	5.56	4.44	45	0.74	0	45	73.64
RE	2.24	6.78	12.99	1.17	2.96	11.11	0.56	15	64.07	0	64.07	63.37
NV	0	0	0	0	0	0	0	0	0	100	100	100

Table 5 Confusion matrix of GaoFen-2 + GLCM data

%	SA	AA	AI	KC	AC	HL	BG	HT	RE	NV	PA	UA
SA	95.93	0	0	8.77	0	4.44	0	0	3.33	0	95.93	93.65
AA	0	71.75	1.69	0	0	0	0	0	1.48	0	71.75	94.78
AI	0	0	69.49	0	0	0	14.44	13.89	10	0	69.49	61.19
KC	1.02	2.82	0	71.93	21.48	0	1.67	3.33	0	0	71.93	71.93
AC	0	7.91	0	14.04	66.67	12.78	0	0.56	2.22	0	66.67	56.96
HL	3.05	7.34	0	4.68	6.67	66.67	0	12.22	3.33	0	66.67	61.22
BG	0	0	12.43	0	0	4.44	78.33	0	0	0	78.33	82.46
HT	0	0	12.43	0	0	11.11	0	70	6.67	0	70	67.74
RE	0	10.17	3.95	0.58	5.19	0.56	5.56	0	72.96	0	72.96	81.74
NV	0	0	0	0	0	0	0	0	0	100	100	100

Table 6 Confusion matrix of GaoFen-2 + GLCM + CHM data

%	SA	AA	AI	KC	AC	HL	BG	HT	RE	NV	PA	UA
SA	97.97	0	0	10.53	0	2.78	0	0	3.33	0	97.97	93.77
AA	0	71.75	0	0	0	0	0	0	0.37	0	71.75	99.22
AI	0	0	75.71	0	0	0	11.11	7.78	10.37	0	75.71	68.37
KC	0.41	1.69	0	74.85	13.33	1.11	0	1.67	0	0	74.85	82.05
AC	0	15.25	0	9.94	71.11	0.56	1.67	0	3.33	0	71.11	62.75
HL	1.22	1.13	0	4.09	6.67	82.78	0	13.33	1.48	0	82.78	74.13
BG	0	0	14.69	0	0	4.44	81.67	0	0	0	81.67	81.22
HT	0	0	2.82	0	0	8.33	0	77.22	3.7	0	77.22	82.25
RE	0	10.17	6.78	0.58	8.89	0	5.56	0	77.41	0	77.41	79.77
NV	0.41	0	0	0	0	0	0	0	0	100	100	99.45

improved after fusing CHM data due to the height difference of mangrove species. Take HL as an example, parts of HL are confused with SA, AA and RE by using the spectral and texture information of GaoFen-2 image, the discrimination is improved after fusing CHM data.

Although fusing texture information or CHM data can improve the overall accuracy and PA/UA of most mangrove species, there is also a reduction in the accuracy of individual mangrove species. Take KC in Tables 4 and 5 as an example, the PA reduces from 91.81 to 71.93% after fusing texture information, this may be caused by the lack of universality of KC validation sample selection. Similar situation appears in AC of Tables 5 and 6, the proportion of AA wrongly classified into AC increased from 7.91 to 15.25%, which decreased from 28.81 to 15.25% in Tables 4 and 5 due to the large difference in texture, while the similar height of AA and AC increases the difficulty of distinguishing between the two types of mangrove species, which also increases the proportion of wrongly classification after fusing CHM data.

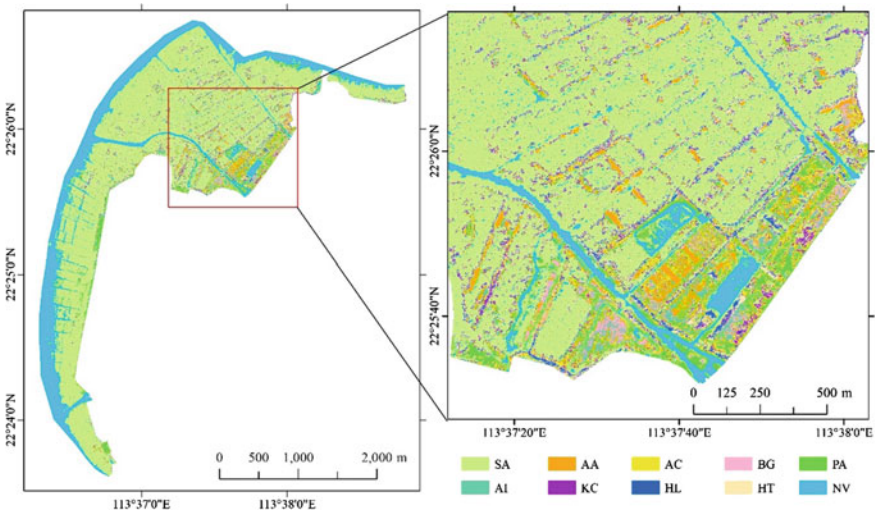


Fig. 4 Map of mangrove species in Qi'ao Island

4.3 Mapping

We use the best classification result for mapping, the classification result of GaoFen-2 + GLCM + CHM was converted into vector and mapped by ArcGIS (Fig. 4). From Fig. 4, SA has the widest distribution area among the mangrove species in Qi'ao Island, which distributed throughout the study area. AI, AA and RE are distributed between SA, while KC, AC, HL, BG, and HT are mainly distributed in the southeast of the study area and is relatively concentrated.

5 Conclusions

In this paper, the mangrove reserve of Qi'ao Island in Zhuhai is taken as the study area, and the UAV CHM data is fused with the GaoFen-2 image to classify the mangrove species. The research results show that due to the lack of spectral bands, Gaofen-2 image cannot better identify mangrove species other than AC and BG. After fusing textural information from GaoFen-2 image and CHM data from UAV LiDAR, the classification accuracy of mangrove species is significantly improved. The main achievements of this study are as follows:

(1) The spatial resolution of GaoFen-2 image up to 1 m provides the possibility of mangrove species classification. Due to its small number of bands, it cannot reflect the spectral differences between different mangrove species well, nor

can it achieve high classification accuracy. However, the abundant texture information contained in the GaoFen-2 image plays a good auxiliary role in distinguishing different mangrove species and can greatly improve the classification accuracy.

(2) Although the fusion of texture information can effectively improve the classification accuracy of the GaoFen-2 image, the accuracy for mangrove species classification is still unsatisfactory. The CHM data obtained by UAV LiDAR can well reflect the difference between mangrove tree heights, and can further improve the accuracy of mangrove classification after being fused with the GaoFen-2 image. The experiment shows that the Gaofen-2 image combined with external CHM data can meet the requirements of mangrove species classification, and can achieve reasonable classification accuracy.

Acknowledgements This research was supported by Independently setting up projects of Key Laboratory of Marine Environmental Survey Technology and Application, Ministry of Natural Resources (MESTA-2021-C005; MESTA-2022-C002); Key Program of Marine Economy Development Special Foundation of Department of Natural Resources of Guangdong Province (GDNRC [2022]19).

References

1. Giri C, Ochieng E et al (2011) Status and distribution of mangrove forests of the world using earth observation satellite data. Global Ecol Biogeogr
2. Katherine E, Robert T, Jin O (1998) Different kinds of mangrove forests provide different goods and services. Glob Ecol Biogeogr Lett 7(1):83–94
3. Deng G (2002) Application of remote sensing technology in mangrove resources investigation. Central South Forest Invent Plan 01:27–28
4. Liu K, Li X, Wang S, Zhong K, Qian J (2005) Monitoring of the changes of mangrove wetland around the Zhujiang Estuary in the past two decades by remote sensing. Trop Geogr 2005(02):111–116
5. Wang Y (2018) Estimation of mangrove biomass in Shenzhen Bay based on multi-source remote sensing data. Southwest University
6. Xu F (2020) mangrove extraction and carbon storage estimation by using Sentinel-2 images. Lanzhou Jiaotong University
7. Li S (2012) Mangroves spatial distribution extraction and species discrimination based on remote sensing data in Beibu Gulf. Nanjing University
8. Wan L, Lin Y, Zhang H, Wang F, Liu M, Lin H (2020) GF-5 Hyperspectral data for species mapping of mangrove in Mai Po. Hong Kong Remote Sens 12:656
9. Zhang H, Wang T, Liu M, Jia M, Lin H, Chu LM, Devlin AT (2018) Potential of combining optical and dual polarimetric SAR data for improving mangrove species discrimination using rotation forest. Remote Sens 10(3):467
10. Ferrentino E, Nunziata F, Zhang H et al (2020) On the ability of PolSAR measurements to discriminate among mangrove species [J]. IEEE J Sel Top Appl Earth Observ Remote Sens 13:2729–2737
11. Liu K, Gong H, Cao J, Zhu Y (2019) Comparison of mangrove remote sensing classification based on multi-type UAV data. Tropical Geography 39(4):492–501

12. Liu X, Bo Y (2015) Object-based crop species classification based on the combination of airborne hyperspectral images and LiDAR data. Remote Sensing 7(1):922–950
13. Cao J, Liu K, Zhuo L, Liu L, Zhu Y, Peng L (20121) Combining UAV-based hyperspectral and LiDAR data for mangrove species classification using the rotation forest algorithm. Int J Appl Earth Observ Geoinform 102(1):1569–8432
14. Li X, Liu K, Zhu Y et al (2018) Study on mangrove species classification based on ZY-3 image. Remote Sens Technol Appl 33(2):360–369
15. Qiu N, Xu S, Qiu P et al (2019) Community distribution and landscape pattern of the mangrove on the Qi'ao Island, Zhuhai. Scientia Silvae Sinicae 55(1):1–10
16. Zhen W, Weijun C, Wei G et al (2017) Study on Zhuhai Qi'ao island main mangrove community characteristics. J Central South Univ Forest Technol 37(4):86–91
17. Hu Y, Zhu N, Liao B et al (2019) Carbon density and carbon fixation rate of mangroves of different restoration types in Qi'ao island. J Central South Univ Forest Technol 39(12):101–107
18. Cao W, Wang W, Wang X et al (2021) Research on crop classification based on GF-2 satellite. Geomatics Spat Inform Technol 44(04):158–161
19. Wu D, Yu W, Xie T (2020) Application of GF-2 satellite data for monitoring organic pollution delivered to water bodies in the Guangdong-Hong Kong-Macao Greater Bay Area. Trop Geogr 40(4):675–683
20. Dezhi W, Bo W, Penghua Q, Yanjun S, Qinghua G, Xincai W (2018) Artificial mangrove species mapping using Pléiades-1: an evaluation of pixel-based and object-based classifications with selected machine learning algorithms. Remote Sens 10:294
21. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
22. Zhao LJ, Tang P (2016) Scalability analysis of typical remote sensing data classification methods: a case of remote sensing image scene. J Remote Sens 20(2):157–171
23. Wang W, Dong Z, Fu D et al (2020) Classification of mangrove in Leizhou Bay based on ZY-3. Hydrogr Survey Chart 40(01):35–39
24. Huang X, Zhang L, Wang L (2009) Evaluation of morphological texture features for mangrove forest mapping and species discrimination using multispectral IKONOS imagery. IEEE Geosci Remote Sens Lett 6:393–397
25. Rao CN, Sastry SS, Mallika K, Tiong HS, Mahalakshmi KB (2013) Co-occurrence matrix and its statistical features as an approach for identification of phase transitions of mesogens. Int J Innov Res Sci Eng Technol 2:4531–4538
26. Peng L, Liu K, Cao J, Zhu Y, Li F, Liu L (2020) Combining GF-2 and RapidEye satellite data for mapping mangrove species using ensemble machine-learning methods. Int J Remote Sens 41(3):813–838

Mobile Traffic Prediction Based on AR-GARCH-LightGBM Hybrid Model

Linxiao Che, Li Wang, Feng Li, and Jiancheng Ge

Abstract Accurate prediction of mobile network traffic is the basis for public network planning, mobile base station management and service quality enhancement. Traditional predictive modeling methods, such as hidden Markov model (HMM) and support vector machine (SVM), mainly focus on the linear time characteristics of traffic data, lacking the capability to address the situation of dense complex networks. In this paper, we propose a hybrid model containing autoregressive moving average (ARIMA) linear prediction and Autoregressive conditional heteroskedasticity mode (GARCH) Residual test, and introduce Light Gradient Boosting Machine (LightGBM) to conduct the network traffic prediction. We decompose the data set into linear and residual parts using DWT. ARIMA model is adopted to deal with the linear component, and the residual parts are predicted by LightGBM. In this process, GARCH is used to test and optimize the prediction of the linear part of ARIMA to improve the prediction accuracy of ARIMA for the linear part. Finally, wavelet reconstruction is used to combine the predicted results. The simulation results show that the prediction accuracy of the hybrid model is improved by 7% compared to the LSTM model, and the time is shortened by 36%.

Keywords LightGBM · ARIMA · GARCH · Mobile traffic forecast

1 Introduction

With the ongoing technological innovation, the development of smartphones in the past decade has accelerated the era of big data [1]. Global mobile data traffic has risen from 2017 to 2022 with an annual growth rate of 42%, especially with the rapid development of voice, short video, and live stream in recent years [2]. The

L. Che · L. Wang (✉) · J. Ge
College of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, China
e-mail: liwang2002@dlmu.edu.cn

F. Li
School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_10

unprecedented number of interconnected devices increases mobile traffic and the difficulty of mobile networks in processing such a large amount of blooming data [3, 4].

This article proposes a hybrid prediction model (DARGL) DWT-ARIMA-GARCH-LightGBM to address the challenges faced by traffic prediction. This model includes a DWT algorithm with decomposed linear and residual parts, an ARIMA part for linear autoregressive prediction, a GARCH part for testing residual characteristics, and a LightGBM ensemble learning model for integrating nonlinear relationships in residual networks. The general flow chart of the system is shown in Fig. 1.

- This article proposes the construction of a DWT preprocessing model in the traditional concept of time series traffic prediction. The time series is decomposed into linear and nonlinear parts through discrete wavelet transform using the DWT model. After processing these two parts, the prediction results are fused using wavelet reconstruction.

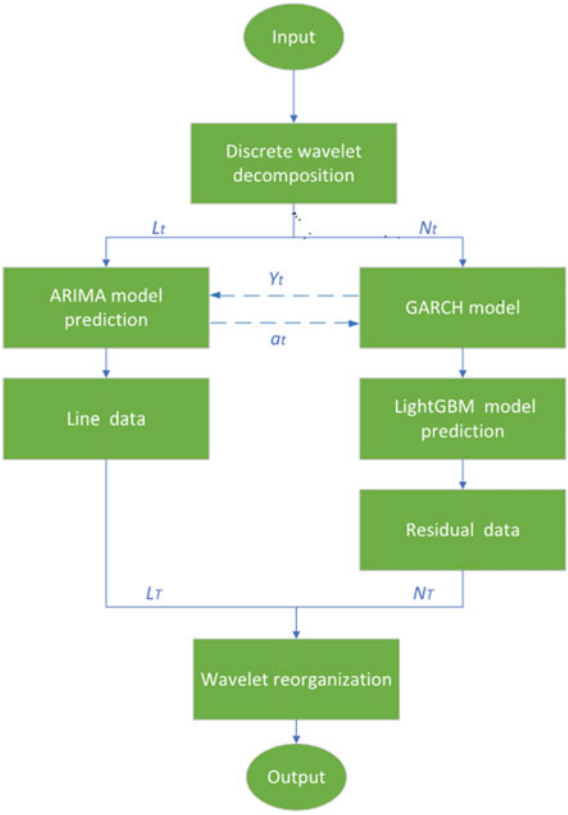


Fig. 1 Model flowchart

- Perform linear autoregressive prediction on mobile traffic data through ARIMA. Mining hidden cycle rules for target area traffic. Simultaneously analyze the stationarity of the residual part of the data and construct a residual network.
- After the ARIMA model fits the original sequence, ARCH effect test is carried out on the residual sequence. If ARCH effect exists, GARCH model is used to fit it, extract residual information from the residual sequence as far as possible, and use its prediction result to correct the ARIMA prediction result.
- When LightGBM predicts the nonlinearity of community traffic, it uses CART trees to process the nonlinear relationships and interdependence between variables in the model, capturing deep features. Regularization is implemented through a Histogram algorithm to reduce model variance and prevent model overfitting. Effectively improving prediction accuracy and convergence speed.

2 System Structure

The heterogeneous network geared to 5G [5] is used in this study, and its network architecture is depicted in Fig. 2. It primarily consists of a central controller module and an edge server module. To maintain the smooth operation of the whole network, the central controller manages the functioning condition of each edge server, resource block allocation, and other information. Perception, data gathering, processing, and storage are all functions of each edge server module. The suggested model still has some flaws in both spatial and temporal perspectives. For starters, it seldom examines the distribution of many external elements. When time and space aspects are included, and cross-domain data are merged, obtaining greater prediction accuracy and quicker convergence rate of wireless cellular service is also a tough task. In this case, we propose utilizing a hybrid deep learning method to forecast the unexpected spike in data.

3 Model

3.1 DWT Model

WT [6] is an ideal tool for signal time-frequency analysis and processing. It overcomes the disadvantage that the window size does not vary with frequency and can provide a "time-frequency" window that varies with frequency. The signal (function) is gradually multi-scale refined by extending and shifting operations. The wavelet transform formula is shown in (1), as follows

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{a}}\varphi\left(\frac{t-b}{a}\right) \tag{1}$$

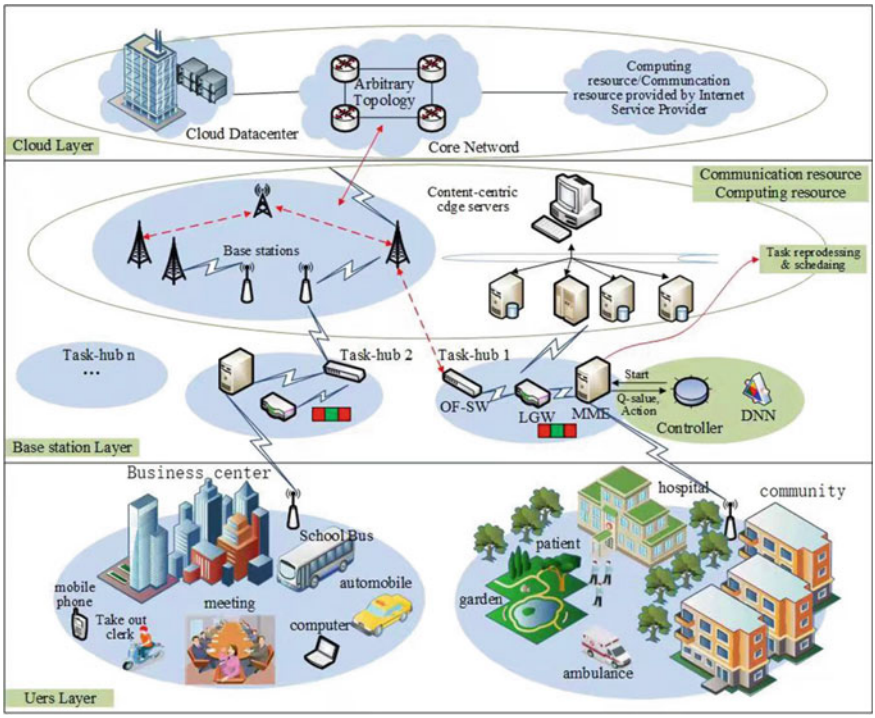


Fig. 2 Mobile flow frame diagram

where a represents the offset factor, and b represents the displacement coefficient. a and b are constants, and $a > 0$ is the basis function; $\varphi(t)$ is initially shifted and then scaled. If a and b constantly change, then a function cluster $\varphi_{m,n}(t)$ can be obtained. Given the square integrable signal $f(t)$, namely, $f(t) \in L^2(R)$, the corresponding DWT (2) is given by the following

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{a}}\varphi\left(\frac{t-b}{a}\right) W_f(m,n) = \left\langle f, \varphi_{m,n}(t)\right\rangle = \int_{-\infty}^{+\infty} f(t)\varphi^*_{m,n}(t)dt \qquad (2)$$

where $\langle *, * \rangle$ denotes the inner product. $*$ denotes a complex conjugate. m and n are the results of the discretization processing of a (scaling factor) and b (displacement factor).

3.2 ARIMA-GARCH Model

The autoregressive comprehensive moving average (ARIMA) method is one of the traditional random time series models invented by BOX and JENKINS, also known as the B-J method [7].

$$Y_t - \varphi_0 - \varphi_1 Y_{t-1} - \cdots - \varphi_p Y_{t-p} = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \tag{3}$$

where, $a_t, a_{t-1}, \ldots, a_{t-q}$ is a stationary white noise with mean of 0 and variance of ξ^2, p and q are the orders of the autoregressive model and the moving average model, respectively. The model is named as the autoregressive moving average series with model autoregressive order p and model moving average order q, which is called ARMA (p, q) series for short. When $q = 0$, it becomes an AR (p) sequence; When $p = 0$, it becomes the MA (q) sequence $\varphi_1, \varphi_2, \ldots, \varphi_p$ with autoregressive coefficients $\theta_1, \theta_2, \ldots, \theta_q$ with moving average coefficients, which are all parameters to be estimated. The ARIMA model can only handle time series of stationary processes. To analyze non-stationary time series, it is necessary to make them stationary.

On the basis of the ARCH model, BOLLERSLEV [8] introduced the lag stage into the conditional variance and obtained the generalized ARCH model, namely the GARCH (Generalized Autoregressive Conditional Heteroscedasticity) model.

$$\sigma_t^2 = \omega + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \cdots + \alpha_q a_{t-r}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \cdots + \beta_p \sigma_{t-s}^2 \tag{4}$$

GARCH (r, s) is a model with additional lag terms. This model is usually useful when using a long data span.

Combining (3) and (4), consider that Y_t follows the ARIMA (p, d, q) - GARCH (r, s) process as follows

$$\begin{aligned} Y_t &= c + \sum_{i=1}^{p} \varphi_i Y_{t-i} + \cdots + \sum_{j=1}^{q} \theta_j a_{t-j} + a_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^{r} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \\ a_t &= \sigma_t \varepsilon_t, \varepsilon_t \sim N(0, 1). \end{aligned} \tag{5}$$

3.3 LightGBM

LightGBM can implement a lightweight gradient lifting algorithm, which is an improved framework based on the GBDT model [9]. GBDT is a classic model in machine learning, and the algorithm idea is to use a weak classifier, that is, a decision tree iterative training, to obtain the optimal model [10]. This model can have a better

training effect while avoiding overfitting, Chen et al. [11] proposed an improved model XGBoost based on GBDT with good generalization ability. Although this improved model has powerful functions, it still has some drawbacks, such as the precise greedy algorithm with huge computational load and the Level wise growth method with low efficiency. To address these shortcomings, LightGBM uses the histogram algorithm The Leaf wise algorithm with depth constraints and parallel optimization have been further optimized.

The objective function of LightGBM can be expressed as

$$\widehat{y} = \sum_{k=1}^{k} f_x(\mathbf{x}), \quad f_k \in F \tag{6}$$

where K is the number of decision trees; F is the set corresponding to all decision trees. f_k is the k-th decision tree generated by the k-th iteration, which is a function in the function space F.

The loss function of LightGBM can be expressed as

$$L^t(\varphi) = \sum_{i=1}^{N} l\left(y_i, \hat{y}_i\right) + \sum_{j=1}^{t} \Omega\left(f_i\right) = \sum_{i=1}^{N} \left(g_i f_t\left(x_i\right) + \frac{1}{2} h_i f_t^2\left(x_i\right)\right) + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{t} w_j^2 \tag{7}$$

The loss function is deformed to obtain calculate the output fraction of leaf node w_j that minimizes the objective function in the t round, directly derive w_j, make the derivative 0, get $w_j = -G_j / \left(H_j + \lambda\right)$ that minimizes the loss function, and substitute it into the loss function

$$L^t(\varphi) = \sum_{j=1}^{\gamma} \left(-\frac{1}{2} \frac{G_j^2}{H_j + \lambda}\right) + \gamma T \tag{8}$$

where, $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$ needs to optimize the objective function, that is, calculate the output fraction of leaf node w_j that minimizes the objective function in the t round, directly derive w_j, make the derivative 0, get $w_j = -G_j/(H_j + \lambda)$ that minimizes the loss function, and substitute it into the loss function.

$$L^t(\varphi) = \sum_{j=1}^{T} \left(-\frac{1}{2} \frac{G_j^2}{H_j + \lambda}\right) + \gamma T \tag{9}$$

4 Community Traffic Prediction with Hybrid Model

This article analyzes the mobile data traffic of a certain mobile network operator. The data collecting period runs from 00:00:00 on July 1, 2022 to 23:00:00 on September 30, 2022, with a one-hour time resolution. This dataset contains community uplink and downlink traffic as well as neighboring cell uplink and downlink traffic with timestamps, temperature and climate, holiday attribute RRC connection indicators, PRB resource use, and total PDCP layer traffic.

We initially took 2040 units of downlink traffic from the community's first 85 days as a test set to analyze the traffic distribution pattern, and then made predictions using the hybrid model suggested in this work. The next 7 days will be the test set, and the predicted results will be tested.

4.1 ARIMA Linear Prediction

According to simulation calculations, $P < 0.005$ indicates the existence of autocorrelation between the data, which can be used for time series analysis. Meanwhile, from the ADF test, it can be seen that the original time series is unstable, and we need to perform a first-order difference on the original sequence. Determine $p = 1$ and q = 0 in the ARIMA model through ACF and PACF tests.

For the ARIMA (1, 1, 0) model, the residual density is shown in Fig. 3. As shown in the figure, perform ARCH test on its standardized residual, p-value $< 2.2 \times 10^{-16}$, indicating that the model has heteroscedasticity. This article chooses GARCH (1, 1) to establish a residual model. The predicted results of the ARIMIA-GARCH model are shown in Fig. 4.

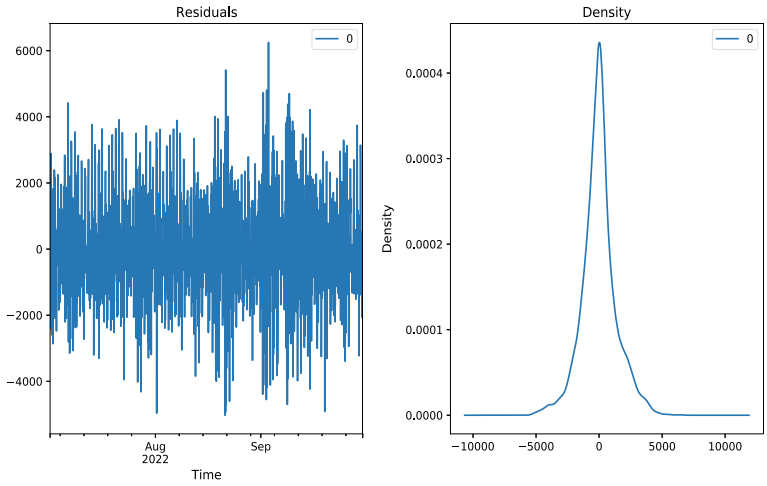


Fig. 3 Residual density

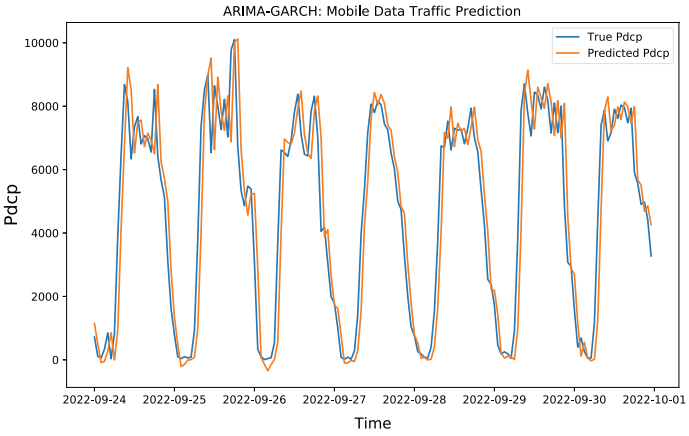


Fig. 4 ARIMA-GARCH model prediction chart

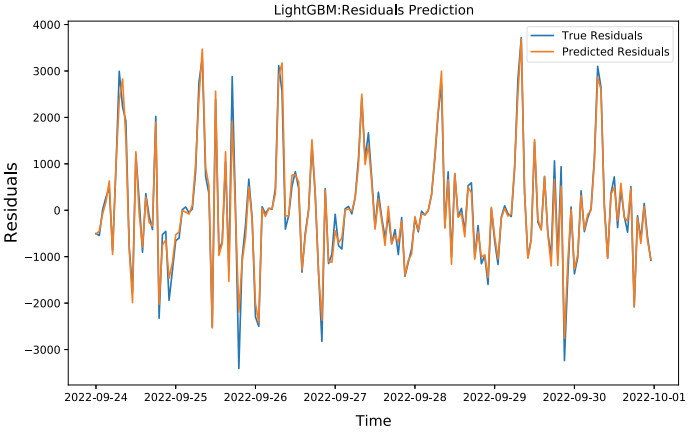


Fig. 5 Prediction diagram of the LightGBM approximation part

4.2 LightGBM

The residual part is introduced into the LightGBM model according to the results of extracting the characteristic variables, as shown in Fig. 5.

As shown in Fig. 6, the mixed model has good approximation and generalization ability, and the prediction accuracy can be greatly improved through the separate prediction and fusion of linear and nonlinear components.

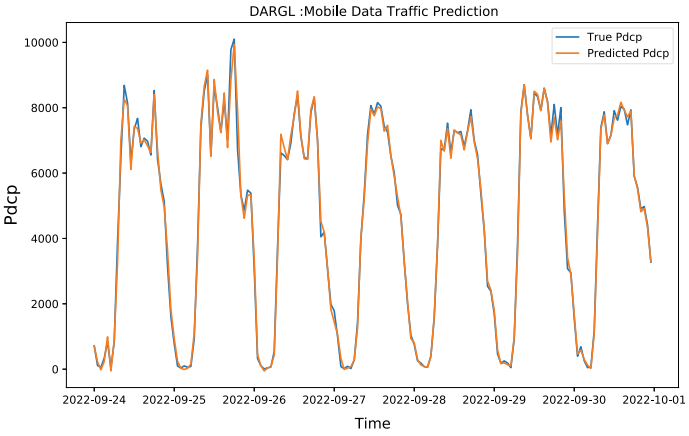


Fig. 6 Prediction of the DARGL model

4.3 Model Evaluation

In this study, the performance of the hybrid model was tested through the evaluation indexes of the mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), R^2, and the time used by the model. The time performance of the mixed model was compared by 1 in unit of the time used by the mixed model. The indicators are defined as follows

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|\hat{Y}_t - Y_t\right| \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\hat{Y}_t - Y_t\right)^2} \tag{11}$$

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{Y}_t - Y_t}{Y_t}\right| \tag{12}$$

$$R^2 = \frac{\sum_{t=1}^{n}\left\|\hat{Y}_t - \overline{Y_t}\right\|^2}{\sum_{t=1}^{n}\left\|Y_t - \bar{Y}\right\|_t^2} \tag{13}$$

where $\hat{Y}_t$ is the predicted value, Y_t is the true value, and $\bar{Y}$ is the mean value. The performance indexes of the proposed hybrid model and the conventional model are compared, as shown in Table 1.

Table 1 Index test diagram of the mixed model

Model	RMSE	MAE	MAPE	R^2	Time (%)
HMM	716.15684	492.8642	0.1937	0.8472	77
ARIMA	648.2872	477.2474	0.1905	0.8719	62
LSTM	476.15692	318.4682	0.0915	0.9104	136
ARIMA-LSTM	347.16845	269.36	0.0714	0.9472	138
ARIMA-Xgboost	293.16742	237.3349	0.0621	0.9462	127
AR-GARCH-Xgboost	127.12342	92.7743	0.0417	0.9761	124
DARGL	136.99512	86.12492	0.0466	0.9731	100

5 Conclusion

This article proposes a hybrid model of DARGL. Simulation shows that the model has good approximation and generalization ability in mobile network traffic prediction, greatly improving the prediction accuracy and enabling accurate and rapid prediction of the business volume within each subnet. Reasonable and effective use of mobile communication network resources can provide users with more stable and efficient services and improve the performance of the entire network. In our future work, we will consider some basic issues of security [12] and privacy when aggregating data in Cloud-Edge Collaboration networks, and will study more accurate and secure models for mobile traffic prediction in the field of wireless communication.

References

1. Naeem M, Jamal T, Diaz-Martinez J, Butt SA, Montesano N, Tariq MI, De-La-Hoz-Valdiris E (2022) Trends and future perspective challenges in big data. In: Advances in intelligent data analysis and applications. Proceeding of the sixth Euro-China conference on intelligent data analysis and applications. Springer Singapore, Dhoby Ghaut, pp 309–325
2. Zhang C, Zhang H, Qiao J, Yuan D, Zhang M (2019) Deep transfer learning for intelligent cellular traffic prediction based on cross-domain big data. IEEE J Sel Areas Commun 37:1389–1401
3. Wang J, Xu C, Zhang J, Zhong R (2022) Big data analytics for intelligent manufacturing systems: a review. J Manuf Syst 62:738–752
4. Zhang Z, Tian L, Shi J, Yuan J, Zhou Y, Cui X, Wang L, Sun Q (2019) Statistical multiplexing gain analysis of processing resources in centralized radio access networks. IEEE Access 7:23343–23353
5. Pateromichelakis E, Moggio F, Mannweiler C, Arnold P, Shariat M, Einhaus M, Wei Q, Bulakci O, De Domenico A (2019) End-to-end data analytics framework for 5G architecture. IEEE Access 7:40295–40312
6. Sinha S, Routh PS, Anno PD, Castagna JP (2005) Spectral decomposition of seismic data with continuous-wavelet transform. Geophysics 70:P19–P25

7. Pankratz A (2009) Forecasting with univariate Box-Jenkins models: concepts and cases. Wiley, London
8. Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. J Econom 31:307–327
9. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Liu TY (2017) LightGBM: a highly efficient gradient boosting decision tree. In: 2017 Advances in neural information processing systems. Curran Associates Inc., RedHook, NY, pp 3149–3157
10. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29:1189–1232
11. Chen TQ, Guestrin C (2016) XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. Association for Computing Machinery, New York, pp 785–794
12. Nan J, Ai M, Liu A, Duan X (2022) Regional-union based federated learning for wireless traffic prediction in 5G-advanced/6G network. In: 2022 IEEE/CIC international conference on communications in China (ICCC workshops). IEEE, London, pp 423–427

Named Entity Recognition of PCI Surgery Information Based on BERT+BiLSTM+CRF

Yuhang Zheng, Li Wang, Feng Li, Hongzeng Xu, and Jiancheng Ge

Abstract Percutaneous coronary intervention (PCI) is a vital treatment method for coronary artery disease, but the unstructured nature of its clinical data makes it challenging to utilize directly. The data for this study was obtained from the Cardiovascular Treatment Center of the People's Hospital of Liaoning Province, China. A representative dataset of 5.8% of PCI patients' surgical records was selected for labeling, and a language model-based PCI surgical information entity recognition model was developed. First, Encoder Representations from Transformers (BERT) was employed to express the semantic relationship between characters accurately. Then, BiLSTM was used as a feature extractor to extract contextual relations, and finally, conditional random field (CRF) was applied to optimize the prediction results. Experimental results demonstrated that the F1 score in the PCI surgical information entity recognition model reached 85.49%, which is 25.66% higher than the traditional HMM and 0.94% higher than BiLSTM in deep learning.

Keywords Percutaneous coronary intervention · Natural language processing · Clinical named entity recognition

Y. Zheng · L. Wang (✉) · J. Ge
College of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, China
e-mail: liwang2002@dlmu.edu.cn

F. Li
School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China

H. Xu
The People's Hospital of China Medical University, The People's Hospital of Liaoning Province, No. 33 Wenyi Road, Shenhe District, Shenyang 110011, China

Key Laboratory of Cardiovascular Imaging and Research of Liaoning Province, Shenyang 110016, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_11

1 Introduction

Percutaneous Coronary Intervention (PCI) is a cardiology procedure that involves balloon dilation or stent implantation to alleviate symptoms of coronary artery stenosis or occlusion. Although PCI operation information records the complete PCI process, clinical texts related to PCI are unstructured, which makes it difficult for clinicians to effectively utilize the information. Therefore, the use of artificial intelligence to process PCI information in a structured manner is necessary to uncover the full potential value of PCI operation information.

In 1996, the term Named Entity Recognition (NER) was introduced as a fundamental task of Natural Language Processing (NLP) at MUC-6 [1]. NER has demonstrated excellent performance in medical data mining. Early approaches relied on rule-based and dictionary-based methods [2] that used templates based on contextual semantic structures. However, these methods could not effectively summarize difficult-to-extract information, and were relatively expensive. To address these limitations, scholars have applied machine learning methods [3–5] such as Hidden Markov Models (HMM), which outperformed traditional methods but could not effectively use contextual semantics for named entity recognition. Currently, deep learning-based NER methods [6, 7] have become more prevalent than the previous two methods. The most popular method is the Bidirectional Long Short-Term Memory (BiLSTM) method. However, BiLSTM cannot constrain the relationship between predicted labels. The Conditional Random Field (CRF) can better constrain the relationship between labels through the emission probability matrix and transition probability matrix. For example, Li et al. [8] proposed an LSTM-CRF-based named entity recognition method.

The intersection of computer science and medicine has led to the emergence of Clinical Named Entity Recognition (CNER) as an important research field. CNER has been successfully utilized to identify body parts, diseases, drugs, and more in various medical fields [9]. However, its specific application in PCI clinical texts remains largely unexplored. The structure of PCI clinical information is complex, involving a combination of Chinese, English, and symbols. Clinical records in this field are relatively incomplete and disorganized, with a high degree of grammatical errors and context ambiguity. These factors pose significant challenges for the identification of named entities related to PCI clinical and surgical information within China's electronic medical records.

2 Methods

The BERT-BiLSTM-CRF model's structure diagram is presented in Fig. 1. This model consists of an input layer, an LSTM layer, a Linear layer, and a CRF layer. The input layer is responsible for converting the input corpus into character vector embedding matrices to facilitate subsequent global feature extraction using Bi-LSTM. The

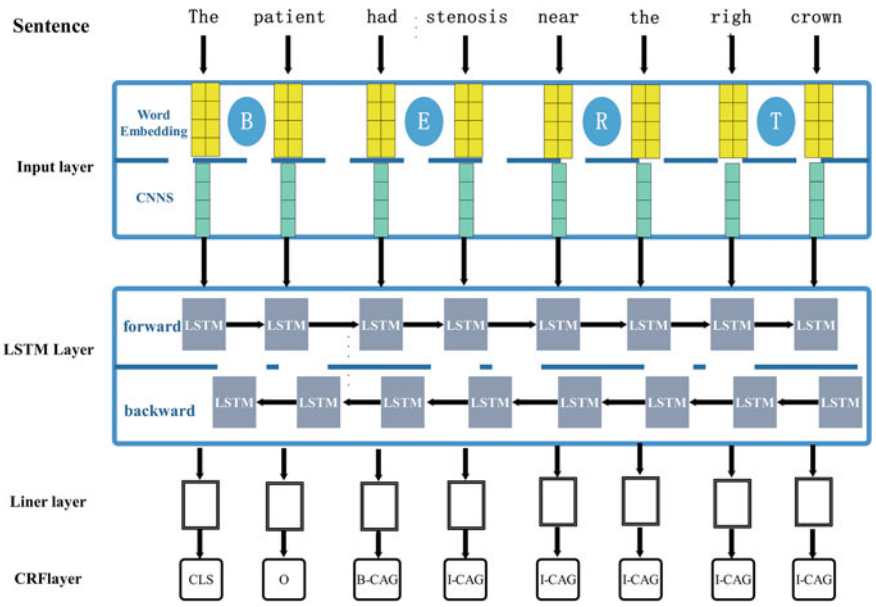


Fig. 1 BERT-BiLSTM-CRF model structure diagram

LSTM layer is tasked with extracting the global features and contextual semantic relationships of the time series. The Linear layer functions as a classifier to assign each entity a probability matrix based on the number of entities. The CRF layer then employs the probability matrix to constrain the relationship between the labels and determine the most probable label sequence.

2.1 BERT

In 2018, Devlin et al. [10] introduced BERT, a pre-trained language model based on the Transformer architecture. Unlike traditional language models, BERT employs a 12-layer Transformer Encoder for learning, with each Encoder consisting of a multi-head attention mechanism and a feedforward network. The multi-head attention mechanism calculates the relationship between words using Query, Key, and Value, and adjusts the weight to extract essential features from the text. Compared to previous pre-training models, BERT captures contextual information more accurately and learns the relationship between consecutive text fragments.

Traditional embedding methods, such as Word2vec [11], Glove [12], and FastText [13], represent all possible word-level meanings in vector form. However, the resulting embeddings are often limited in their ability to express the semantic and distance relationships between words. Additionally, in PCI operation information, there is

no clear boundary between Chinese, English, and symbols, which can lead to poor recognition performance. To address these issues, this paper proposes a model that utilizes BERT as an embedding method through transfer learning. Unlike traditional methods, BERT is capable of capturing contextual information and understanding relationships between consecutive text fragments, which can improve performance in recognizing PCI operation information.

2.2 LSTM

The LSTM layer's individual unit receives the output from the Input layer and extracts the sequence information of the text to learn the contextual features of the corpus. It can consider the semantic relationship between each character before and after the sentence simultaneously and combine them to create a more comprehensive representation. The LSTM layer takes the WordEmbedding of the corpus as an input sequence and produces another output sequence vector $h = (h_1, h_2, \ldots, h_n)$ that represents the sequence at each time step in the input sequence. LSTM addresses the challenges of long sentence dependencies and gradient explosion in RNNs by incorporating three gate control units that regulate the retention and forgetfulness of specific information. The computation formula for the LSTM hidden layer output representation at a given input time X_t is expressed as Formula (1)–(5).

$$i_t = \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}C_{t-1} + b_i\right) \tag{1}$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}C_{t-1} + b_f\right) \tag{2}$$

$$C_t = f_t C_{t-1} + i_t \tanh\left(W_{xC}x_t + W_{hC}h_{t-1} + b_C\right) \tag{3}$$

$$o_t = \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}C_{t-1} + b_o\right) \tag{4}$$

$$h_t = o_t \tanh\left(C_t\right) \tag{5}$$

In the equation, W represents the weight matrix and b is the offset vector. C represents the state of the memory unit, σ is the sigmoid activation function, and tanh is another activation function. The input gate, forget gate, and output gate are represented by i_t, f_t, and o_t, respectively. The hidden state output, H_t, includes the text information in the clinical information. The gate's frequency threshold mechanism is capable of effectively filtering out irrelevant information and retaining important information that needs to be preserved.

However, unidirectional LSTM can only capture information in a forward direction. As sentences become longer, local features at the beginning of the sentence make up a smaller proportion. Therefore, it is necessary to use a Bidirectional LSTM to capture information both forward and backward, and concatenate the outputs. This enhances the information captured from both the beginning and end of the sentence and improves the ability to capture semantic dependencies within the context.

2.3 CRF

While BiLSTM is effective at extracting semantic features from the entire text, it does not impose constraints on the label relationships. During the entity output prediction stage, the softmax function is typically used as a classifier to address multi-classification problems, resulting in incorrect outcomes and affecting model performance. To address this issue, this paper employs the CRF model in decoding.

The Conditional Random Field (CRF) is a type of undirected probabilistic graphical model that can constrain the relationship between labels and improve the accuracy of entity predictions. By defining a starting probability matrix and a transition probability matrix, the constraints between tags and the prediction of entities can be enforced. Specifically, because the entity labels starting with "B" must be followed by labels of the same type starting with "I", the launch probability matrix and transition probability matrix effectively model these constraints and reduce errors. Given an input sequence $H = h_1, h_2, h_3, \ldots, h_n$ and the corresponding output from the LSTM layer, the CRF model calculates the conditional probability distribution of the output sequence $Y = y_1, y_2, y_3, \ldots, y_n$ and assigns a score to each label. The label with the highest score is then selected as the final prediction label using the calculation formula (6):

$$s(h, y) = \sum_{i=1}^{N} P_{i,y_i} + \sum_{i=1}^{N} T_{y_{i-1},y_i} \tag{6}$$

When decoding, the Viterbi algorithm [14] is used to obtain the optimal output sequence y_R. This algorithm computes the score of each label in the sequence based on the transition matrix T and the score vector P. The optimal path is determined by selecting the label with the highest score for each position in the sequence, taking into account the constraints between labels. The calculation formula for the Viterbi algorithm is given by Eq. (7):

$$y_R = \arg\max_{y \in Y} s(h, y) \tag{7}$$

3 Experiments and Results

3.1 Dataset

The clinical data utilized in this research was sourced from the People's Hospital of Liaoning Province (Shenyang, China), a prominent and comprehensive third-class hospital in China. The clinical data relates to PCI (percutaneous coronary intervention) procedures involving coronary angiography, balloon dilatation, and stent implantation. Specifically, the clinical texts pertaining to three distinct categories of

Table 1 Distribution of eight entities in the dataset

Entity	Training set	Validating set	Testing set
Guding	1030	194	264
GuideBrand	1501	307	396
BalloonModel	2136	524	628
SupportModel	1558	273	398
SupportBrand	1578	273	389
Anticoagulants	1012	166	223
ContrastMedium	936	163	225
CAGResult	2900	476	635
All	12,651	2376	3158

fine-grained PCI surgical information were extracted. By conducting a comprehensive analysis of various clinical and surgical records, this study chose to use data from 1340 inpatients who were admitted to the Cardiology Department as the corpus for analysis. This study was finally approved by the Ethics Committee of the People's Hospital of Liaoning Province (Ethics number: (2023)K021).

For the experimental study, a corpus of 1340 cases of PCI surgery information was utilized. From this, 930 cases were selected as the training set, while 180 cases were assigned as the validation set. Following model training, the remaining 230 cases were utilized as the test set. Table 1 displays the distribution of the eight types of entities in the training, validation, and test sets.

3.2 Result Analysis

Table 2 shows that traditional HMM achieved excellent results due to the simpler format of anticoagulants than the normative structure, but it performed poorly in identifying entities that require contextual semantics. Bi-LSTM showed the best performance in Chinese and English entity recognition, but it did not perform well in identifying complex entities such as contrast results. However, after adding CRF, the relationship between labels was effectively constrained, resulting in significant improvements in angiography results and entities with complex stent signal structures. This greatly reduced false recognition. Additionally, with the addition of BERT, sensitivity to entities such as the combination of numbers and symbols, such as the bracket model, was increased, and the accuracy of other entities was improved.

The overall performance of a model can be evaluated by its ability to recognize all types of entities correctly. The F1 score for each model is calculated based on the proportion of each entity in the PCI operation information. The results are presented in Table 3.

Table 2 F1 values of eight entities in each model

Entity	HMM	BiLSTM	BiLSTM-CRF	BERT-BiLSTM-CRF
Guding	25.7	**89.2**	85.9	86.8
GuideBrand	52.0	89.2	89.9	**91.2**
BalloonModel	73.9	91.2	91.6	**93.1**
SupportModel	50.6	75.9	76.7	**77.2**
SupportBrand	56.7	90.9	93.1	**93.2**
Anticoagulants	**91.9**	86.4	90.1	90.5
ContrastMedium	79.5	94.2	**94.2**	89.7
CAGResult	15.5	68.5	70.5	**71.6**

Bold indicates the best performing of all models

Table 3 Performance between models

Model	Strict			Relaxed		
	Precision	Recall	F1	Precision	Recall	F1
HMM	52.08	72.89	59.83	88.98	81.48	82.79
BiLSTM	84.25	84.91	84.55	97.43	97.36	97.35
BiLSTM-CRF	84.71	85.1	84.9	97.43	97.36	97.35
BERT-BiLSTM-CRF	85.27	85.49	**85.49**	97.43	97.37	**97.37**

Bold indicates the best performing of all models

To ensure the rigor of the experiment, the overall performance of each model is evaluated under both strict and loose standards. As shown in the table, the traditional HMM performed poorly, with an F1 score of only 59.83 under the strict standard, due to its limited ability to extract features and combine contextual semantics, resulting in a large number of recognition errors and limitations in identifying entities. Although BiLSTM achieved higher performance than HMM, its performance on complex imaging entities was not satisfactory, with incorrect label order leading to a reduced F1 score. The addition of CRF effectively constrained label relationships and improved accuracy for some complex entities, such as those with mixed numbers and symbols, resulting in an overall F1 score improvement of 0.35. In comparison, BERT-BiLSTM-CRF performed better, as the WordEmbedding of the former two was randomly initialized and could not accurately reflect the relative distance of the word vector space in character information representation. This improved accuracy in the recognition of most entities.

4 Discussion and Conclusion

This study focuses on the extraction of clinical information from PCI surgery information based on Chinese EMRs. Determine and identify eight entities including angiography result, catheter type, guidewire brand, stent model, stent brand, balloon

model, contrast agent, and anticoagulant from three clinical surgical records. The performance of the BERT-BiLSTM-CRF model for extracting clinical PCI operation information has been further improved, and the accuracy rate basically meets the needs of clinical applications. The results demonstrate that deep learning methods can be used to automatically extract PCI surgical information from EMRs for clinical named entity recognition.

References

1. Grishman R, Sundheim BM (1996) Message understanding conference-6: a brief history. In: COLING 1996: the 16th international conference on computational linguistics, vol 1
2. Song M, Yu H, Han WS (2015) Developing a hybrid dictionary-based bio-entity recognition technique. BMC Med Inf Decis Making 15(1):1–8
3. McCallum A, Li W (2003) Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons
4. Chieu HL, Ng HT (2002) Named entity recognition: a maximum entropy approach using global information. In: COLING 2002: the 19th international conference on computational linguistics
5. Bender O, Och FJ, Ney H (2003) Maximum entropy models for named entity recognition. In: Proceedings of the seventh conference on natural language learning at HLT-NAACL 2003, pp 148–151
6. Wang X, Zhang Y, Ren X, Zhang Y, Zitnik M, Shang J, Langlotz C, Han J (2019) Cross-type biomedical named entity recognition with deep multi-task learning. Bioinformatics 35(10):1745–1752
7. Cai X, Dong S, Hu J (2019) A deep learning model incorporating part of speech and self-matching attention for named entity recognition of chinese electronic medical records. BMC Med Inf Decis Making 19(2):101–109
8. Li L, Jiang Y (2018) Integrating language model and reading control gate in BLSTM-CRF for biomedical named entity recognition. IEEE/ACM Trans Comput Biol Bioinform 17(3):841–846
9. Chen X, Ouyang C, Liu Y, Bu Y (2020) Improving the named entity recognition of Chinese electronic medical records by combining domain dictionary and rules. Int J Environ Res Publ Health 17(8):2687
10. Devlin J, Chang MW, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805
11. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J (2013) Distributed representations of words and phrases and their compositionality. Adv Neural Inf Process Syst 26
12. Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
13. Bojanowski P, Grave E, Joulin A, Mikolov T (2017) Enriching word vectors with subword information. Trans Assoc Comput Linguist 5:135–146
14. Viterbi AJ (2006) A personal history of the Viterbi algorithm. IEEE Sig Process Mag 23(4):120–142

Stacking Ensemble Models Predict Mortality from Acute Myocardial Infarction

Yu Zhang, Li Wang, Feng Li, Hongzeng Xu, and Songrui Pei

Abstract In this study, a stacking ensemble model is proposed to predict in-hospital mortality for NSTEMI. The stacking ensemble model uses a two-tier structure to improve the performance of the algorithm. The first layer of the model uses decision trees (DT), support vector machines (SVM), random forests (RF), gradient boosting decision trees (GBDT) as the base model, and logistic regression (LR) is selected as the metamodel in the second layer. By comparing with four machine learning models, the results of experiments show that the AUC of the stacking model is (0. 958), higher than DT (0.934), SVM (0.942), RF (0.945), GBDT (0.948). In terms of AUC, Accuracy, Precision, Recall, and F1, the indicators improved by 1%, 1%, 0.7%, 0.4% and 0.9%, respectively, compared with the highest of the four single models.

Keywords Non-ST-segment elevation myocardial infarction · In-hospital mortality · Stacking model

1 Introduction

Cardiovascular disease (CVD) kills 17.9 million people and is the leading cause of death globally, accounting for 31% of all deaths globally each year. Of these, 85% were caused by stroke and acute coronary syndrome (ACS), which underscores that

Y. Zhang · L. Wang (✉) · S. Pei
College of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, China
e-mail: liwang2002@dlmu.edu.cn

F. Li
School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China

H. Xu
The People's Hospital of China Medical University, The People's Hospital of Liaoning Province, No. 33 Wenyi Road, Shenhe District, Shenyang 110011, China

Key Laboratory of Cardiovascular Imaging and Research of Liaoning Province, Shenyang 110016, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_12

coronary artery disease (CAD) remains a challenging global health problem until now [1]. The differences in clinical features and risk of different ACS make the risk assessment of death greatly affect the level of medical care and the choice of management strategy [2]. Among them, NSTEMI accounts for about 70% of ACS patients [3].

It can accurately predict the in-hospital mortality of NSTEMI patients and play an important role in the rational allocation of resources and timely intervention [4]. Several effective risk scores already exist to predict in-hospital mortality with ACS, such as the Global Acute Coronary Event Registry (GRACE) score and the Thrombolysis in Myocardial Infarction (TIMI) risk score [5, 6]. Due to the small number of Asians included in these scores, they are not fully applicable to Asians. The emergence of machine learning provides a new way of medical data analytics. For example, a combination of machine learning and deep learning is used to predict heart disease [7] and improve the predictive power of COVID-19 [8]. For the prediction of in-hospital deaths in ACS, such as [9] using logistic regression to analyze the correlation between variables, using five machine learning methods to predict the prognosis of patients with STEMI.

Because NSTEMI has a high risk of death, the study of in-hospital mortality is clinically important. Therefore, this paper uses the Stacking [10] ensemble method to integrate multiple MLs to predict in-hospital mortality in patients with NSTEMI.

2 Material and Methods

2.1 Data Collection and Data Preprocessing

Overview of the Research Framework. The structure of this study consists of three parts as shown in Fig. 1. Firstly, the collected data were preprocessed and re-sampled using SMOTE oversampling method. After that, the Recursive Feature Elimination (REF) [11] was used to select features based on the AUC of the model. Finally, a Stacking model based on multiple models is established for dynamic prediction.

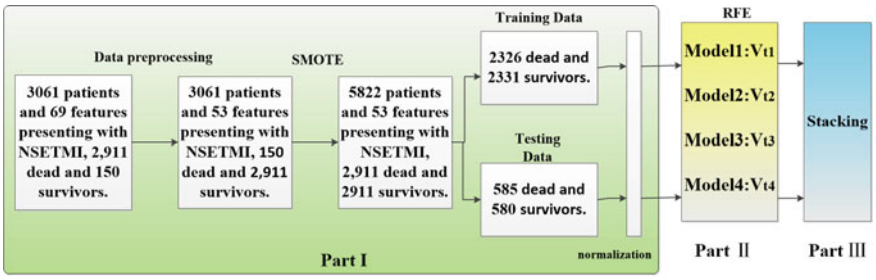


Fig. 1 Structural framework

Data Source. The clinical data used in this study are from patients diagnosed with NSTEMI inpatients in Liaoning Provincial People's Hospital (Shenyang, China) during the 5-year period from 2017 to 2022. The data excluded data under 18 years of age without any laboratory information and drug information, and finally included data for a total of 3061 patients, of which 2911 were in the surviving group and 150 in the death group. This study was finally approved by the Ethics Committee of the People's Hospital of Liaoning Province (Ethics number: (2023)K021). The study was retrospective, with patient data being de-identified and processed.

Data Preprocessing. To remove the features with a deletion rate of more than 30%, this paper uses multiple imputation (MI) to fill in the features with a deletion rate of less than 30%, which can make the filled data better maintain the original distribution of the data and the relationship between variables. Due to the serious imbalance of positive and negative samples of the dataset, the oversampling technique SMOTE was used to achieve the balance of the dataset by synthesizing data with a low sample size, and the total number of balanced datasets was 5822. The balanced dataset is divided into training and testing sets in an 8:2 ratio. Then standardize the data, such as Eq. (1) to eliminate numerical differences between variables. x is the input, mean is the average and σ is the standard deviation, and x^* is the normalized output value.

$$x^* = \frac{x - mean}{\sigma} \tag{1}$$

2.2 Feature Selection

In this study, RFE was used for feature extraction, which integrates the process of feature selection and training, although it takes a little time, but it is only done once in preprocessing, and subsequent prediction does not consume additional time.

2.3 Model Building

This paper uses the stacking ensemble model as the training model, which consists of two layers of structure, the first layer is the base model, and the second layer is the meta-model. The base model selects four learners, DT, SVM, RF, GBDT [12–15], and predicts by 5-fold cross-detection stacking, and the meta-model uses logistic regression. The output of the first layer is used as the input of the second layer to make the final prediction. This paper uses Stacking's ensemble model framework as shown in Fig. 2a, Model 1 to Model 4 represent the four base models, V_1 to V_n represent the characteristics of the data, where V_{t1}, V_{t2}, V_{t3}, V_{t4} represent the feature subset of the 4 models, V-Model1, V-Model2, V-Model3, V-Model4 represent the optimal feature subset of the 4 models after RFE screening, and the final The detailed steps of the Stacking integration model are described below.

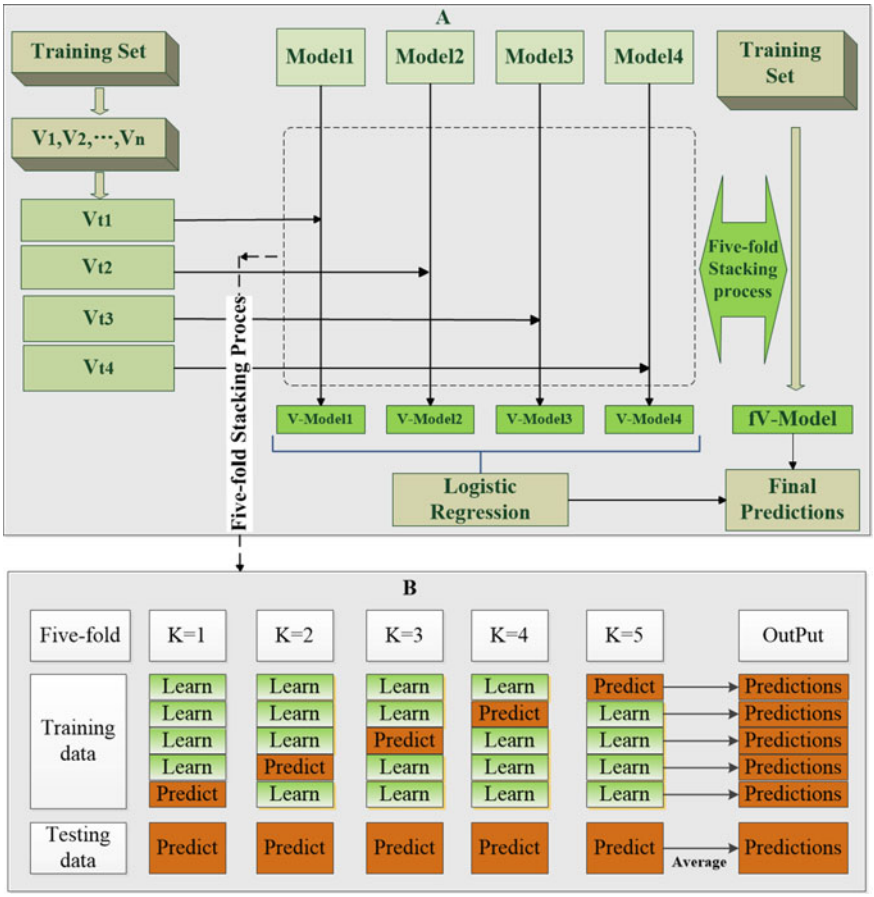


Fig. 2 This article uses the Stacking ensemble learning model frame

In the first layer, for the base model1 to Model4, the corresponding feature subsets V_{t1}, V_{t2}, V_{t3}, and V_{t4} are used as inputs, and the prediction V-Model1, V-Model2, V-Model3, and V-Model4 of the base classifier are generated by a 5-fold cross-test. As shown in Fig. 2b, we divide the training set into 5 folds for cross-validation. In each iteration, 4 folds are used to train the classifier, and the remaining 1 fold is used for prediction. At the same time, at each iteration, the trained classifier predicts the test set. After 5 iterations, the prediction results of the training set can be obtained. The mean of the predictions of the classifier in the test set is used as the prediction of the base model.

In the second layer, since the features of this layer are extracted based on complex nonlinear transformations, there is no need to select complex classifiers. LR [16] is a good candidate classifier due to its simple structure and the fact that L2 regularization can further prevent overfitting [17]. Therefore, we use LR as the prediction model

of the second layer, train the training set predictions generated by the first layer, and make the final predictions based on the test set predictions generated by the first layer.

2.4 Evaluation Metrics

The main evaluation index of this study is the area AUC below ROC, and AUC is selected as the standard on the performance standard based on parameter adjustment and feature selection. Accuracy, Precision, Recall and F1 are used as model evaluation indicators. As shown in Eqs. (2)–(5), where TP = true positive, FP = false positive, TN = true negative, FN = false negative.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$F1 = \frac{2 * Precision * Rrcall}{Precision + Recall} \tag{5}$$

2.5 Statistical Analysis

Categorical data are expressed as percentages, all continuous data for normal distributions are expressed as mean (standard deviation), and all continuous data for nonnormal distributions are expressed as median (IQR). The t-test, nonparametric test, and chi-square test were used for comparison. p-value < 0.05 was considered statistically significant.

3 Result

3.1 Baseline Characteristics

A total of 3061 hospitalized patients diagnosed with NSTEMI were included in this study, and their in-hospital mortality rate was 4.9% (2911 in the survival group and 150 in the death group). As can be seen in Table C, the P values of the 10 characteristics of HDL_C, MCH, TBIL, IBIL, ZDB, RVOT, AOtop, AOmid, AObottom, and AV are > 0.05, which is not statistically significant, so these characteristics are excluded. The final number of features included in subsequent model training is 53.

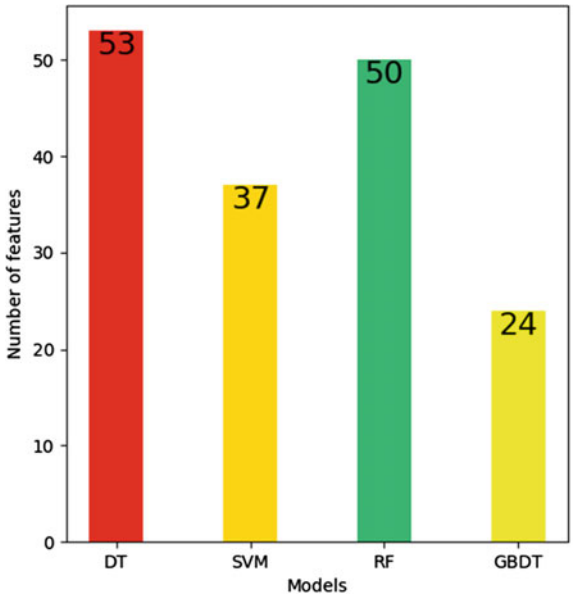


Fig. 3 Number of model features after RFE selection

3.2 Results of Feature Selection

After using machine learning methods for feature selection, you can further reduce the number of features required in the model to some extent Fig. 3 shows the results of RFE feature selection using four different methods. It can be seen from the figure that the number of features of DT, SVM, RF and GBDT is 53, 37, 50 and 24. Each model removes a large number of redundant features, especially for SVMs and GBDTs, which are as few as 37 and 24, respectively.

3.3 Results of Model Comparison

In order to evaluate the Stacking model used in this paper has higher predictive performance than the 4 single models, this paper compares with these 4 models, the ROC curve comparison chart of the model is shown in Fig. 4, and the comparison results of other indicators are shown in Table 1, the Stacking model proposed in this paper has higher performance than the 4 single models in all evaluation indicators, among which the AUC, Accuracy, Precision, Recall and F1 are improved by 1%, 1%, 0.7%, respectively. 0.4% and 0.9%.

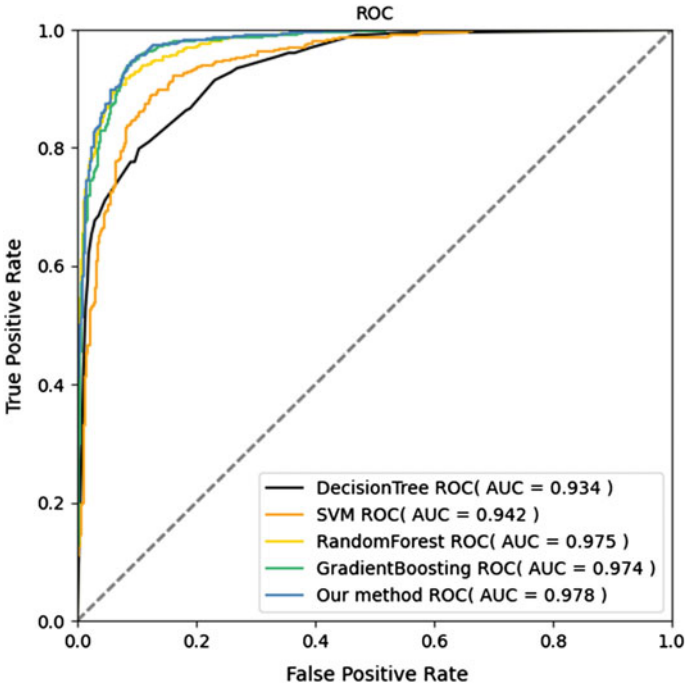


Fig. 4 ROC curve of the model

Table 1 Comparison of performance metrics of different models

Models	AUC	Accuracy	Recall	Precision	F1
DT	0.934	0.847	0.798	0.886	0.84
SVM	0.942	0.881	0.896	0.87	0.883
RF	0.945	0.866	0.894	0.848	0.87
GBDT	0.948	0.886	0.897	0.878	0.888
Our method	0.958	0.896	0.901	0.893	0.897

4 Discussion

This study used a stacking ensemble learning model to predict in-hospital mortality in patients with NSTEMI. The model uses 53 features, all of which can be obtained from the hospital's electronic information system. The addition of these data, which is not available in previous scores such as the Grace score and the TIMI score, will make our model more convincing. With the development of science and technology, the practical research of machine learning applications in the medical field is becoming more and more extensive, but there are fewer studies in the risk prediction of hospitalization of patients with NSTEMI patients. The submissions of this study.

Moreover, the Stacking ensemble model has a great improvement in performance compared with a single ML algorithm, can make predictions more accurately, and remind doctors to provide medical help to patients in a more timely manner. In this study, GBDT performed best among the four candidate models, with AUC, Accuracy, Precision, and F1 improvements of 1%, 1%, 0.7%, and 0.9%, respectively, for the stacking model compared to GBDT. It shows that the proposed model can further improve the overall prediction performance compared with the best individual model. The criteria for adjusting model parameters and selecting base classifiers are AUC indicators, not F1, so as to achieve the optimal classification effect as the standard. The research in this paper has some limitations, such as the time cost of RFE feature selection, and it is necessary to further explore effective methods to save time and cost, so that the model can obtain the optimal subset of features faster and further improve the accuracy of the model.

5 Conclusion

This study used a stacking ensemble model to predict NSETMI in-hospital mortality risk. Compared with a single ML model, it has outstanding performance in Accuracy, auc, Precision, Recall, and F1 evaluation indicators, indicating that the Stacking ensemble model in this study can integrate the advantages of multiple models to achieve better prediction performance. Effective death risk prediction models can provide doctors with valuable insights into a patient's condition and target high-level patients with early clinical intervention to minimize patient death.

References

1. Cardiovascular disease. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
2. Ryan TJ, Anderson JL, Antman EM, Braniff BA, Brooks NH, Califf RM, Hillis LD, Hiratzka LF, Rapaport E, Riegel BJ, Russell RO (1996) ACC/AHA guidelines for the management of patients with acute myocardial infarction: a report of the American College of Cardiology/American Heart Association task force on practice guidelines (committee on management of acute myocardial infarction). J Am Coll Cardiol 28:1328–1419
3. O'gara PT, Kushner FG, Ascheim DD, Casey DE, Chung MK, De Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA et al (2013) 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines. J Am Coll Cardiol 61:e78–e140
4. Kumar D, Saghir T, Zahid M, Ashok A, Karim M (2021) Validity of TIMI score for predicting 14-day mortality of non-ST elevation myocardial infarction patients. Cureus
5. Roffi M, Pa-trono C, Collet JP, Mueller C, Windecker S (2015) ESC guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: task force for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation of t. G Ital Cardiol 13(2011):171–228

6. Worner F, Cequier A, Bardají A, Bodí V, Pan M (2012) ESC guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. Eur Heart J 33:2569–619
7. Bharti R, Khamparia A, Shabaz M, Dhiman G, Pande S, Singh P (2021) Prediction of heart disease using a combination of machine learning and deep learning. Comput Intell Neurosci
8. Kwekha-Rashid AS, Abduljabbar HN, Alhayani B (2021) Coronavirus disease (COVID-19)cases analysis using machine-learning applications. Appl Nanosci 1–13
9. Xiao C, Guo Y, Zhao K, Liu S, He N, He Y, Guo S, Chen Z (2022) Prognostic value of machine learning in patients with acute myocardial infarction. J Cardiovasc Dev Dis 9:56
10. Wolpert DH (1992) Stacked generalization. Neural Netw 5:241–259
11. Wu C, Liang J, Wei W, Li C (2017) Random forest algorithm based on recursive feature elimination. Stat Decis
12. Singh M, Singh P, Singh H (2006) Decision tree classifier for human protein function prediction. In: 2006 international conference on advanced computing and communications. IEEE, pp 564–568
13. Danenas P, Garsva G (2012) Credit risk evaluation modeling using evolutionary linear SVM classifiers and sliding window approach. Proc Comput Sci 9:1324–1333
14. Waljee AK, Liu B, Sauder K, Zhu J, Govani SM, Stidham RW, Higgins PD (2018) Predicting corticosteroid-free endoscopic remission with vedolizumab in ulcerative colitis. Aliment Pharmacol Ther 47:763–772
15. Niu F, Recht B, Re C, Wright SJ (2011) Hogwild!: a lock-free approach to parallelizing stochastic gradient descent. Adv Neural Inf Process Syst 24:693–701
16. Krishnapuram B, Carin L, Figueiredo M, Hartemink AJ (2005) Sparse multinomial logistic regression: fast algorithms and generalization bounds. IEEE Trans Pattern Anal Mach Intell 27:957–68
17. Tang J, Liang J, Han C, Li Z, Huang H (2019) Crash injury severity analysis using a two-layer stacking framework. Accid Anal Prev 122:226–238

LF-GANet: Local Frame-Level Global Dynamic Attention Network for Speech Emotion Recognition

Shuwei Dou, Tingting Han, Ruqian Liu, Wei Xia, and Hongmei Zhong

Abstract Speech emotion recognition (SER) is an important field of human–computer interaction. Although humans have various ways of expressing emotions, speech is one of the most direct ways. Therefore, it is an important technical challenge to extract the emotional information from the speech signal as much as possible. To address this issue, we proposed a local frame-level global dynamic attention network (LF-GANet) to extract emotional information from speech signals. This network mainly consists of two parts, a local frame-level module (LFM) and a global dynamic attention module (GAM). To extract rich frame-level emotional information from speech signals, the LFM was designed to extract features from forward and reverse time series separately; the GAM real-time extracted the global correlations from speech signals. We conducted experiments on the EMODB and SAVEE datasets. The results showed that our method outperformes the existing SOTA model in UAR on both datasets, verifying the effectiveness of the model.

Keywords Speech emotion recognition · Frame-level features · Bidirectional time information · Multi-attention

1 Introduction

Human–computer interaction has been reflected in all aspects of our daily life, such as mobile phone application software, intelligent customer service, intelligent home system, mental health care and other fields. The biggest issue that exists today is the lack of natural human–machine interaction, where machines couldn't understand emotional information conveyed by humans very well and respond in a human-like way, resulting in poor user experience. Therefore, the most important goal of Speech

S. Dou · T. Han (✉) · R. Liu · W. Xia · H. Zhong
Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
e-mail: hantt@tjnu.edu.cn

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

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_13

Emotion Recognition (SER) is to extract the emotions that the speaker wants to convey from the speech signals.

Traditional SER methods usually were based on machine learning methods using artificial feature extraction, such as GMM (Gaussian Mixture Model), SVM (Support Vector Machine), random forest, etc. Yildirim et al. [1] indicated that SVM could serve as an effective approach for implementing speech emotion recognition. Assuncao et al. [2] used the Logistic Model Tree algorithm to train a classifier model, and experimental results showed that the model had good performance in handling emotion recognition problems. However, these methods required hand-crafted feature design, and the performance of these models was often limited by the feature representation. With the development of deep learning, there have been many new breakthroughs in speech emotion recognition, such as convolutional neural network (CNN) [3, 4], recurrent neural network (RNN) [5], long short-term memory (LSTM), Mustaqeem and Kwon [6] used a bidirectional long short-term memory network (Bi-LSTM) to directly map speech signals to emotional outcomes. In order to improve the performance of the model, Ye et al. [7] proposed Gated Multi-scale Temporal Convolutional Network (GM-TCNet), which considered the causality of speech emotion and greatly improved the performance of emotion classification. However, these researchers have not yet delved into the global correlation of speech signals. In 2017, Google proposed the Transformer model [8], which includes self-attention mechanism (self-attention) and multi-head self-attention mechanism (multi-head self-attention), and achieved excellent results in machine translation tasks. Among them, multi-head attention enables the model to simultaneously focus on information from different positions, thereby better capturing potential features from the input data. Therefore, the multi-head attention effectively considers the global correlation of the signal, which can improve the representation ability of the model and thus improve the performance. Inspired by this, we proposed a local bi-directional dynamic global attention network, which aimed at extracting emotional information contained in the speech signals as comprehensively as possible. The local frame-level module was designed to extract frame-level emotional information from speech signals. The global dynamic attention module integrated information from speech signals in real time. The introduction of multi-head attention focused on the information from different positions of the input features and considered the global correlation, so as to better capture potential emotional information in the input features. Our main contributions were:

(1) A novel local bidirectional dynamic global attention network (LF-GANet) was proposed, which not only considered the frame-level features of speech signals, but also deeply considered their global correlation, which can effectively extract multi-scale emotional information from speech signals.
(2) The evaluation results of LF-GANet on EMODB and SAVEE datasets exceeded most SOTA methods. The model achieved a UAR of 91.77% on the EMODB dataset and 76.04% on the SAVEE dataset.

2 Methodology

As shown in Fig. 1, the LF-GANet mainly consists of three modules, including the input module, the local frame-level module (LFM) and the global dynamic attention module (GAM). The LFM is to obtain rich frame-level emotional information, and it extracts speech information from forward and backward time sequences respectively. The GAM is to obtain the global correlation of speech. It integrates the information from forward and backward time sequences and enhances the correlation by inputting them into a multi-head attention layer. Finally, the output of GAM directly enters the fully connected layer for emotional classification.

2.1 Feature Extraction

MFCCs (Mel Frequency Cepstral Coefficients) [9] are the most commonly used features in speech emotion. It is not affected by factors such as environmental noise and differences in recording equipment, and its excellent anti-interference ability can extract features related to a specific speaker. Even if the speaker's pronunciation, speech speed, volume and other conditions are different, MFCCs can still perform relatively stable and have good robustness. Although MFCCs themselves can well describe the frequency domain information of speech signals, they are not capable

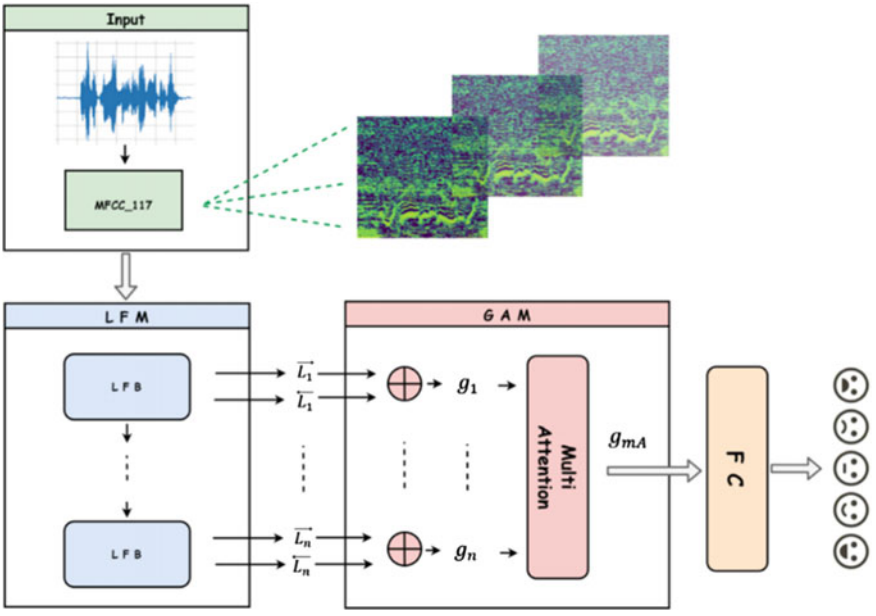


Fig. 1 The framework of LF-GANet

of modeling dynamic features in the time domain. However, its first and second derivative can better capture the dynamic changes of the speech signal, which can improve the ability to model audio signals. Therefore, we fused the 39-dimensional MFCCs and their first and second derivative into 117-dimensional MFCCs features as the input of our model. We set the sampling rate as the original sampling rate of each dataset, with a frame length of 50 ms, a frame shift of 12.5 ms, and a 2048-point FFT per frame.

2.2 *Local Frame—Level Module (LFM)*

As shown in Fig. 2, the core of LFM is a well-designed local frame-level block (LFB) used to extract frame-level features of speech signals. The core of LFM is to extract frame-level features from forward time series and reverse time series, respectively, by using well-designed LFB. Among them, LFB includes Conv1D, BN, ReLU, and SD layers. Conv1D is used to extract advanced emotional features, BN is for data normalization, and ReLU is for nonlinear activation. In order to overfit the way, we added the SpatialDropout1D layer for random deactivation. The four layers are stacked sequentially, and the input features are activated as frame-level features using Sigmoid. The activation method is given by Eqs. (1) and (2):

$$\overrightarrow{L_i} = \overrightarrow{L_{i-1}} \odot \delta\left(\overrightarrow{L_{i-1}}\right) \tag{1}$$

$$\overleftarrow{L_i} = \overleftarrow{L_{i-1}} \odot \delta\left(\overleftarrow{L_{i-1}}\right) \tag{2}$$

Among them, $\overrightarrow{L_i}$ represents the vector of the i-th forward time series, and $\overleftarrow{L_i}$ represents the vector of the i-th reverse time series.

2.3 *Global Dynamic Attention Module (GAM)*

The core of the GAM is to real-time fuse the frame-level features from the two directions and send them to the multi-head attention layer to dynamically obtain global information. The specific calculation process is given by Eqs. (3) and (4):

$$g_i = G\left(\overrightarrow{L_i}\right) \oplus G\left(\overleftarrow{L_i}\right) \tag{3}$$

$$g_{mA} = m\_A\left(\sum_i^n w_i g_i\right) \tag{4}$$

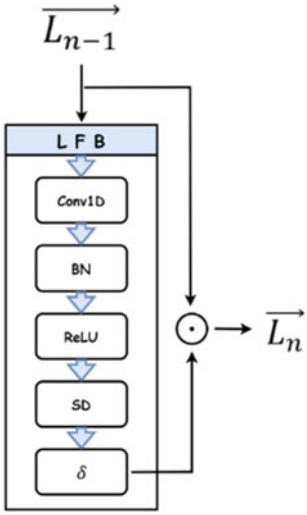


Fig. 2 The framework of LFB

where G represents the global time pooling operation, w_i represents a learnable weight matrix, m_A represents an m-head multi-head attention operation.

The features output by the multi-head attention layer not only contain locally grasped speech emotion information, but also consider the global correlation of speech information.

Finally, we used the cross-entropy loss function to input the features into a fully-connected layer for speech emotion classification. Due to the unbalanced sample size of the existing speech emotion datasets, we chose Unweighted Average Recall (UAR) as the evaluation index of the model. UAR represented the average recall rate under each emotion category, which could objectively evaluate our model. The calculation process was given by Eq. (5):

$$UAR = \frac{1}{K}\sum_{\alpha=1}^{K}\frac{\sum_{\beta=1}^{M} TP_{\alpha}^{\beta}}{\sum_{\beta=1}^{M}\left(TP_{\alpha}^{\beta}+FN_{\alpha}^{\beta}\right)} \tag{5}$$

Here, K, M and N represent the number of sentiment classes, the number of speech signals (β) of class α and the number of all speech signals, respectively. TP_{α}^{β}, TN_{α}^{β}, FP_{α}^{β} and FN_{α}^{β} represent the true positive, true negative, false positive, and false negative values of class α for speech signal β respectively.

Table 1 The detailed information of speech emotion datasets

Dataset	Language	Speakers	Numbers	Emotion	Frequency (kHz)
EMODB	German	5 males, 5 females	535	Angry, Boredom, Disgust, Fear, Happy, Neutral, Sad	16.0
SAVEE	English	4 males	480	Angry, Disgust, Fear, Happy, Neutral, Sad, Surprise	44.1

Table 2 The details of data distributions in two datasets

Dataset	Angry	Boredom	Disguise	Fear	Happy	Neutral	Sad	Surprise
EMODB	127	81	46	69	71	79	62	–
SAVEE	60	–	60	60	60	120	60	60

3 Experiments

3.1 Datasets

We used two publicly available sentiment datasets in the experiments. EMODB was an emotional database created by the Institute of Communication Sciences of the Technical University of Berlin [10]. The sampling rate of the audio files in the EMODB database was 16 kHz. The SAVEE dataset was a cross-modal speech emotion recognition database that included speech and facial expressions [11]. The sampling rate of the audio files in the SAVEE database was 44.1 kHz. Each audio file contained a speech record of a speaker in different emotional states. Tables 1 and 2 gave the details of these datasets.

3.2 Experimental Setup

During the experiment, we extracted a 39-dimensional MFCC from the Librosa toolbox [12], computed its first and second derivatives, and fused them into a 117-dimensional MFCC. The model was optimized using the Adam algorithm with an initial learning rate of 0.001 and a batch size of 64. To avoid overfitting during the training phase, we implemented the 0.1 factor label smoothing as a form of regularization and set the spatial dropout rate to 0.1. The attention head was set to 8, the dimension of the mapped vector was 128, and the dropout was 0.2. Due to the limited sample size of the speech dataset, we chose ten-fold cross-validation to fully utilize the data, improve the generalization ability of the model, and obtain more stable and reliable results.

Table 3 The performances of different approaches on the EMODB dataset

Study	Method	Year	UAR (%)
Yildirim et al. [1]	SVM	2021	78.89
Mustaqeem et al. [6]	Bi-LSTM	2020	85.57
Ye et al. [7]	GM-TCNet	2022	90.45
Our proposed	LF-GANet	2023	**91.77**

Note Bold indicates the result with the highest indicator

3.3 Results and Analysis

Table 3 shows the performance of our method on the EMODB dataset, and the UAR reached 91.77%. It is worth noting that although the Bi-LSTM model [6] considers the interdependence between bidirectional time series information, it fails to extract the global correlation from speech information. Our global dynamic attention module (GAM) effectively solves this problem, resulting in a 6.2% increase in UAR compared to the Bi-LSTM model. Although the GM-TCNet model [7] considers the causality in emotions and has strong advantages over previous methods, it still fails to learn the differences between global information well, and the model complexity is relatively high. Our method not only has a refreshing structure, but also the perfect combination of local bidirectional module and global dynamic attention module captures the emotional information in speech signals very well, resulting in a 1.31% improvement in UAR compared to the GM-TCNet model.

Table 4 shows the performance of our method (LF-GANet) on the SAVEE dataset, and the UAR reached 76.04%. The performance of our method on the SAVEE dataset is not as good as that on the EMODB dataset. Our analysis may have two reasons: Firstly, there were differences in the two datasets themselves. Secondly, we did not consider the causality of emotion in speech time series. Although there were few related studies on the SAVEE dataset for speech emotion recognition, our method still had good performance advantages compared to other SOTA methods [2, 13, 14]. However, this also reflects that in the field of speech emotion recognition, the scarcity of datasets hinders research work, and the narrow way to improve the generalization ability of the model is also a problem that needs to be solved urgently.

Table 4 The performances of different approaches on the SAVEE dataset

Study	Method	Year	UAR (%)
Ibrahim et al. [13]	Echo state networks	2022	64.05
Assuncao et al. [2]	Logistic model tree	2020	68.00
Kanwal et al. [14]	SVM	2021	69.80
Our proposed	LF-GANet	2023	**76.04**

Note Bold indicates the result with the highest indicator

4 Conclusions

In this paper, we proposed a local frame-level dynamic global attention network (LF-GANet), which mainly consists of two modules: a local bidirectional module and a global dynamic attention module. The local bidirectional module was used to extract the frame-level features from speech signals, and the global dynamic attention module was used to integrate the global correlation in the speech signal in real time. The combination of the two modules can not only have a clear model structure, but also extract rich emotional information from speech signals at multiple scales. It had significant performance advantages on both EMODB and SAVEE datasets. The performance on the SAVEE dataset was not perfect due to the sparsity limitation of the existing public datasets for speech emotion recognition. Therefore, the generalization ability of the model was still an important problem that needs to be solved urgently. In the future, we will seek better ways to extract the emotional causality in speech signals, and also want to generate a rich speech emotion dataset that can well solve the validation of model generalization ability.

Acknowledgements This work was supported by the Tianjin Science and Technology Planning Project under Grant No. 20JCYBJC00300 and the National Natural Science Foundation of China under Grant No. 62001328.

References

1. Yildirim S, Kaya Y, Kılıc F (2021) A modified feature selection method based on metaheuristic algorithms for speech emotion recognition. Appl Acoustics 173:107721
2. Assuncao G, Menezes P, Perdig˜ao F (2020) Speaker awareness for speech emotion recognition. Int J Online Biomed Eng 16(4):15–22
3. Ozer I (2021) Pseudo-colored rate map representation for speech emotion recognition. Biomed Signal Process Control 66:102502
4. Muppidi A, Radfar M (2021) Speech emotion recognition using quaternion convolutional neural networks. In: ICASSP 2021, Toronto, ON, Canada, June 6–11. IEEE, pp 6309–6313
5. Rajamani ST, Rajamani KT, Mallol-Ragolta A et al (2021) A novel attention-based gated recurrent unit and its efficacy in speech emotion recognition. In: ICASSP 2021, Toronto, ON, Canada, June 6–11. IEEE, pp 6294–6298
6. Mustaqeem M, Kwon S (2020) Clustering-based speech emotion recognition by incorporating learned features and deep bilstm, IEEE. Access 8:79861–79875
7. Ye JX, Wen XC, Wang XZ, Xu Y, Luo Y, Wu CL, Chen LY, Liu KH (2022) GM-TCNet: gated multi-scale temporal convolutional network using emotion causality for speech emotion recognition. Speech Commun 145:21–35. ISSN 0167-6393. https://doi.org/10.1016/j.specom.2022.07.005
8. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. Adv Neural Inf Process Syst 30
9. Peng Z, Lu Y, Pan S et al. (2021) Efficient speech emotion recognition using multi-scale CNN and attention. In: ICASSP 2021, Toronto, ON, Canada, June 6–11. IEEE, pp 3020–3024
10. Burkhardt F, Paeschke A, Rolfes M et al (2005) A database of German emotional speech. In: INTERSPEECH 2005, Lisbon, Portugal, September 4–8, vol 5, pp 1517–1520

11. Philip Jackson and SJUoSG Haq (2014) Surrey audio-visual expressed emotion (savee) database. University of Surrey, Guildford, UK
12. McFee B, Raffel C, Liang D, Ellis DP, McVicar M, Battenberg E, Nieto O librosa (2015) Audio and music signal analysis in python. In: Proceedings of the 14th python in science conference, vol 8, pp 18–25
13. Ibrahim H, Loo CK, Alnajjar F (2021) Grouped echo state network with late fusion for speech emotion recognition. In: Neural information processing—28th international conference, ICONIP 2021, Sanur, Bali, Indonesia, December 8–12, 2021, Proceedings, Part III. (Lecture Notes in Computer Science), vol 13110. Springer, pp 431–442
14. Kanwal S, Asghar S (2021) Speech emotion recognition using clustering based ga-optimized feature set. IEEE Access 9:125830–125842

Performance Analysis of DF-Based Satellite and UAV Relay Networks

Mingyi Ji, Bo Xu, Liwen Zhu, and Shupei Huang

Abstract In this paper, we investigate the performance of downlink transmission in a satellite and unmanned aerial vehicle (UAV) relay network. Specifically, given that the UAV equipped with multiple antennas adopts decoded-and-forward (DF) protocol, we propose three transmission schemes to satisfy the quality-of-service requirement of a terrestrial user, including direct transmission, cooperative transmission and joint transmission. Then, assuming that satellite-UAV link follows shadowed-Rician fading while the UAV-user link obeys Nakagami-*m* distribution, we derive the closed-form expression for outage probability (OP) of the considered system with different schemes. Based on the derived OP expressions, we further analyze the energy efficiency (EE) of the system. Finally, the simulations results verify the correctness of the theoretical analysis, and the influence of UAV antenna number on the performance in terms of OP and EE is revealed.

Keywords Satellite and unmanned aerial vehicle relay network · Decoded-and-forward · Outage probability · Energy efficiency

1 Introduction

Compared with terrestrial cellular communication, satellite communication has attracted a great deal of attention due to its advantages of wide coverage, long communication distance, and not being restricted by geographical conditions, etc. Since satellite network can not only provide reliable communication services for remote areas as a supplement to the terrestrial network, but also show broad application prospects in the next generation mobile communication, as an important means of constructing the integration of air and earth communication [1–3]. However, several problems such as the long delay and large path loss between the satellite and terrestrial users, as well as blocked line-of-sight transmission caused by shadow effect,

M. Ji (✉) · B. Xu · L. Zhu · S. Huang
College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China
e-mail: Jimy_email@163.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_14

can seriously affect the quality-of service (QoS) of terrestrial users. In order to address these issues, the concept of cooperative transmission techniques has been proposed in [4]. For example, the authors of [5] investigated BER (Bit Error Ratio) and outage performance for a satellite-terrestrial cooperative network under a single-user scenario, where a terrestrial relay uses amplify and forward (AF) protocol. In [6], the authors derived a theoretical expression for the outage probability (OP) of the system based on opportunistic user selection scheme. It should be noted that the terrestrial relay in remote areas, will suffer from high construction costs and deployment difficulties, so it is urgent to explore the other options to compensate for the shortcomings of terrestrial facilities.

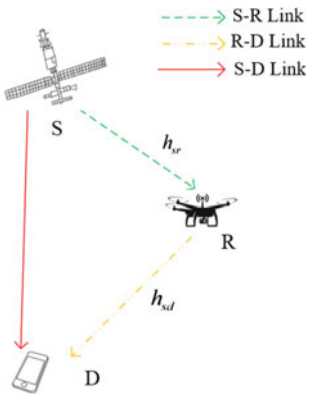
In recent years, unmanned aerial vehicle (UAV) communication has gained importance in the field of wireless communication due to its unique advantages such as low cost, easy control and flexible deployment. Compared with the terrestrial relay, UAV can be used as an aerial relay to quickly set up a direct communication link for different tasks and enhance the received signal power, and hence increase the system capacity and communication coverage. With this regard, the UAV-assisted cooperative network has gradually become a preferred communication network architecture for most researchers. For example, in a satellite-terrestrial cooperative network with an amplify-and-forward (AF) based UAV relays, the author of [7] proposed a multi-user scheduling scheme and analyze the performance. In [8], by utilizing either AF protocol or decode-and forward (DF) protocol, terrestrial relay nodes are designed to forward the signal received from satellite to users. Although the aforementioned works have in-depth studies among UAV-relay assisted satellite-terrestrial collaborative networks, and demonstrates that relay collaboration techniques can significantly improve the performance of satellite communication systems in the presence of occlusion effects [9]. It should be pointed out that they only focus on the case of relay nodes configured with a single antenna, while the multi-antennas technique is an effective means to improve power efficiency and information transmission rate in practical wireless communication.

Motivated by the aforementioned observations, we investigate the performance of downlink transmission in a satellite and UAV relay network, and then propose a transmission scheme based on the qualities of the channels between the satellite and the user. After analyzing the energy efficiency of the system according to the derived OP expression, we also carry out Monte Carlo simulations to verify the accuracy of the proposed theoretical analysis and reveal the influence of UAV antenna number on the system performance.

2 System and Channel Model

We consider a cooperative transmission network as shown in Fig. 1, which consists of a satellite source (S) equipped with a directional antenna, a terrestrial relay (R) equipped with N antennas, and a single-antenna user (D). It is assumed that the

Fig. 1 System model



direct link between S and D and the S-R link experience shadowed Rician (SR) fading, whereas the R-D link undergoes Nakagami-m fading [10].

2.1 Channel Model

As illustrated in Fig. 1, we assume that R and D are within the coverage of satellite. By considering the effects of fading characteristics of the wireless channel, the propagation loss of the radio wave on the channel, we model the S-D link channel as [11].

$$h_{sx} = l_{s,x} g_{sx} \tag{1}$$

where $h_{s0}(x = 0)$ and $h_{sx}(x = 1, 2, ..., N)$ denote the channel response of the S-D link and S-R link, respectively. g_{sx} is the small-scale fading, and l_{sx} denotes the path loss. According to the Friis transmission formula, l_{sx} can be expressed as [12]

$$l_{sx} = \frac{c\sqrt{G_s G_x}}{4\pi f D_x} \tag{2}$$

where c is the speed of light and f denotes the operating frequency. $D_0(x = 0)$ and $D_x(x = 1, 2, ..., N)$ are the distances between satellite S to the terrestrial user D and the UAV respectively. G_s and G_x denote the antenna gain of the satellite and the corresponding antenna gain respectively. G_s can be modeled as [13].

$$G_s = G_{\max}\left(\frac{J_1(\nu)}{2\nu} + 36\frac{J_3(\nu)}{\nu^3}\right)^2 \tag{3}$$

where $G_{\max}$ is the maximum satellite antenna gain, $J_1(v)$ and $J_3(v)$ represent the first kind Bessel function of order 1 and 3 respectively, and $v = 2.07123 \sin \vartheta_1 / \sin \vartheta_{3dB}$, with ϑ_1 being the angle between the terrestrial and the beam line of sight, and ϑ_{3dB} is the 3 dB angle.

Similar to the satellite model, the R-D link can be written as

$$\boldsymbol{h}_{rd} = l_{rd} \boldsymbol{g}_{rd} \tag{4}$$

where $l_{rd} = -\frac{1}{2}(10\alpha_D \lg D_3 + 28 + 20 \lg f)[dB]$ denotes the free-space path loss from the UAV to the terrestrial user [14]. α_D is the path loss factor, D_3 is the distance between the UAV relay and the user, $\boldsymbol{g}_{rd} = [g_{rd.1}, g_{rd.2} \cdots g_{rd.N}]$ is the small-scale fading, whose elements obey a Nakagami-m distribution with a fading parameter of μ and an average power of Ω.

2.2 Signal Model

In this paper, we propose three transmission schemes to satisfy the quality-of service (QoS) requirement of a terrestrial user, including direct transmission (DT), cooperative transmission (CT) and joint transmission (JT).

Direct transmission signal model

For the direct transmission scheme, we consider the case where a direct link exists between the satellite and the terrestrial user, the satellite can send the signal $x_s(t)$, satisfying $E\left[|x_s(t)|^2\right] = 1$, directly to the terrestrial user D, then the received signal of the terrestrial user can be expressed as

$$y_{sd}(t) = \sqrt{P_s} h_{sd} x_s(t) + n_0(t) \tag{5}$$

where P_s is the transmit power at S, h_{sd} is the channel coefficient of S-D link, $n_0(t)$ is the Additive White Gaussian Noise (AWGN) with zero mean and variance $\sigma_0^2 = \kappa B T$, where κ is the Boltzmann constant, B is the noise bandwidth and T is the noise temperature. Then, the output signal-to-noise ratio can be expressed as

$$\gamma_{sd} = \frac{P_s}{\sigma_0^2} |h_{sd}|^2 \triangleq \overline{\gamma}_{sd} |g_{sd}|^2 \tag{6}$$

where $\overline{\gamma_{sd}} = P_s l_{sd}^2 / \sigma_0^2$ denotes the average signal-to-noise ratio of the S-D link.

Collaborative transmission signal model

In a collaborative transmission scheme, the entire communication occurs in two phases. In the first stage, the satellite sends the signal $x_s(t)$ to the UAV, then the signal received by UAV can be written as

$$y_{sr}(t) = \sqrt{P_s}\boldsymbol{w}_1^H \boldsymbol{h}_{sr} x_s(t) + \boldsymbol{w}_1^H \boldsymbol{n}_1(t) \tag{7}$$

where $n_1(t)$ is a noise vector of size $N \times 1$ satisfying $n_1(t) \sim N_c(0, \sigma_2^2 I_N)$, $\sigma_1^2 = \kappa BT$. $h_{sr} \in \mathbb{C}^{N\times 1}$ is the channel response vector of the S-R link, and $w_1 \in \mathbb{C}^{N\times 1}$ denotes the received beamforming power vector at the UAV.

According to Eqs. (1) and (6), the received signal-to-noise ratio at R can be given by:

$$\gamma_{sr} = \frac{P_s}{\sigma_1^2}\|h_{sr}\|_F^2 \triangleq \overline{\gamma}_{sr}\|g_{sr}\|_F^2 \tag{8}$$

where $\overline{\gamma_{sr}} = P_s l_{sr}^2/\sigma_1^2$ denotes the average signal-to-noise ratio of the S-R link.

In the second stage, the UAV relay R decodes the received signal according to the DF protocol, and then sends the re-encoded signal to the terrestrial user D after processing it through the transmit beamforming power vector $w_2 \in \mathbb{C}^{N\times 1}$, then the signal received by the terrestrial user D from the relay R can be expressed as:

$$y_{rd}(t) = \sqrt{P_r}\boldsymbol{h}_{rd}^H \boldsymbol{w}_2 x_r(t) + n_2(t) \tag{9}$$

where P_r is the transmit power of the UAV relay R, $h_{rd} \in \mathbb{C}^{N\times 1}$ is the channel response vector of the R-D link, and $n_2(t)$ is the AWGN with zero mean and variance $\sigma_2^2 = \kappa BT$.

According to Eqs. (4) and (9), the signal-to-noise ratio at D can be written as

$$\gamma_{rd} = \frac{P_r}{\sigma_2^2}\|h_{rd}\|_F^2 \triangleq \overline{\gamma}_{rd}\|g_{rd}\|_F^2 \tag{10}$$

where $\overline{\gamma_{rd}} = \frac{P_r l_{rd}^2}{\sigma_2^2}$ indicates the average SNR of the R-D link.

Since the UAV relay R uses the DF transmission protocol, based on (9), the output SNR at D can is given by:

$$\gamma = \min\{\gamma_{sr}, \gamma_{rd}\} \tag{11}$$

Joint transmission signal model

The joint transmission scheme takes both of the schemes mentioned before into account. In the first stage, the satellite signal sends a signal $x_s(t)$ to R and D. In the second stage, the UAV relay forwards the received signal to the terrestrial user D via the DF protocol. To enhance the strength of the received signal, the terrestrial user uses maximal ratio combining (MRC) scheme. Thus, the output SNR of terrestrial user D can be written as

$$\gamma = \gamma_{sd} + \min\{\gamma_{sr}, \gamma_{rd}\} \tag{12}$$

Integrated model

In summary, the system output SNR of three satellite communication system transmission schemes can be written respectively as

$$\gamma = \begin{cases} \gamma_{sd}, & DT \\ \min\{\gamma_{sr}, \gamma_{rd}\}, & CT \\ \gamma_{sd} + \min\{\gamma_{sr}, \gamma_{rd}\}, & JT \end{cases} \tag{13}$$

2.3 UAV Beamforming Scheme

After receiving the signal, an optimization problem is established to maximize the SNR of the UAV, and then the optimization objective can be expressed as $\max\{\gamma_{sr}\}$.

To improve system performance, the UAV uses the MRC scheme, so the received beamforming weight vector satisfies $w_1 = h_{sr}/\|h_{sr}\|_F^2$.

Similarly, for UAV transmissions, the beamforming weight vector was derived after the optimization problem was established to maximize the signal-to-noise ratio received by users, i.e. $w_2 = h_{rd}/\|h_{rd}\|_F^2$.

3 Performance Analysis

OP is an important indicator to measure the performance of a wireless communication system. The OP is usually defined as the probability that the SNR falls below a certain threshold γ_{th}, namely

$$P_{out} = \Pr(\gamma \le \gamma_{th}) = F_{\gamma}(\gamma_{th}) \tag{14}$$

where $F_{\gamma}(x)$ denotes the Cumulative Distribution Function (CDF) of γ.

Assuming that the S-D link experience SR fading [15], the Probability Density Function (PDF) of γ_{sd} can be expressed as

$$f_{\gamma_{sd}}(x) = \frac{\chi_{s1}}{\overline{\gamma}_{sd}} \sum_{i=0}^{\tau_{s1}-1} \frac{(-1)^i (1-\tau_{s1})_i \eta_{s1}^i}{(i!)^2} \left(\frac{x}{\overline{\gamma}_{sd}}\right)^i \times \exp\left(-\frac{(\zeta_{s1}-\eta_{s1})x}{\overline{\gamma}_{sd}}\right) \tag{15}$$

where $\chi_{s1} = \frac{1}{2q_{s1}}\left(\frac{2q_{s1}\tau_{s1}}{2q_{s1}\tau_{s1}+E_{s1}}\right)^{\tau_{s1}}$, $\eta_{s1} = \frac{E_{s1}}{2q_{s1}(2q_{s1}\tau_{s1}+E_{s1})}$, $\zeta_{s1} = \frac{1}{2q_{s1}}$, E_{s1} denotes the average power of the direct path component. $2q_{s1}$ denotes the average power of the multipath component [14]. τ_{s1} denotes the channel fading parameter corresponding to the direct path component.

Similarly, the PDF of γ_{sr} can be expressed as:

$$f_{\gamma_{sr}}(x) = \frac{\chi_{s2}^{N}}{\overline{\gamma}_{sr}\Gamma(N)} \sum_{i=0}^{\tau_{s2}-1} \frac{(-1)^i (N - N\tau_{s2})_i \eta_{s2}^i}{(N)_i (i!)^2} \times \left(\frac{x}{\overline{\gamma}_{sr}}\right)^{i+N-1} \exp\left(-\frac{(\zeta_{s2} - \eta_{s2})x}{\overline{\gamma}_{sr}}\right) \tag{16}$$

where $\chi_{s2} = \frac{1}{2q_{s2}} \left(\frac{2q_{s2}\tau_{s2}}{2q_{s2}\tau_{s2}+E_{s2}}\right)^{\tau_{s2}}$; $\eta_{s2} = \frac{E_{s2}}{2q_{s2}(2q_{s2}\tau_{s2}+E_{s2})}$; $\zeta_{s2} = \frac{1}{2q_{s2}}$; E_{s2} denotes the average power of the direct path component; $2q_{s2}$ denotes the average power of the multipath component; τ_{s2} denotes the channel fading parameter corresponding to the direct path component.

Assuming that the S-D link undergoes Nakagami-m fading, the PDF of γ_{rd} can be written as

$$f_{\gamma_{rd}}(x) = \frac{\tau_r^{\tau_r N} E^{-\tau_r N}}{\Gamma(\tau_r N)\overline{\gamma}_{rd}^{\tau_r N}} x^{\tau_r N - 1} \exp(-\frac{\tau_r x}{E\overline{\gamma}_{rd}}) \tag{17}$$

In order to solve the expression for the system OP more conveniently, by integrating Eqs. (16), (17), and (18), the CDFs of γ_{sd}, γ_{sr}, and γ_{rd} can be, respectively, written as following

$$F_{\gamma_{sd}}(x) = \sum_{i=0}^{\tau_{s1}-1} \chi_{s1} \frac{(-1)^i (1 - \tau_{s1})_i \eta_{s1}^i}{(\zeta_{s1} - \eta_{s1})^{i+1} i!} \times \left(1 - \exp\left(-\frac{(\zeta_{s1} - \eta_{s1})x}{\overline{\gamma}_{sd}}\right) \sum_{j=0}^{i} \frac{(\zeta_{s1} - \eta_{s1})^j}{j!} \left(\frac{x}{\overline{\gamma}_{sd}}\right)^j\right) \tag{18}$$

$$F_{\gamma_{sr}}(x) = \sum_{i=0}^{N(\tau_{s2}-1)} \chi_{s2}^{N} \frac{(-1)^i (N - N\tau_{s2})_i \eta_{s2}^i}{(\zeta_{s2} - \eta_{s2})^{i+N} \Gamma(N)} \frac{(i + N - 1)!}{(N)_i i!} \times \left(1 - \exp\left(-\frac{(\zeta_{s2} - \eta_{s2})x}{\overline{\gamma}_{sr}}\right) \sum_{j=0}^{i+N-1} \frac{(\zeta_{s2} - \eta_{s2})^j}{j!} \left(\frac{x}{\overline{\gamma}_{sr}}\right)^j\right) \tag{19}$$

$$F_{\gamma_{rd}}(x) = 1 - \exp\left(-\frac{\tau_r x}{\overline{\gamma}_{rd} E}\right) \sum_{i=0}^{N\tau_r - 1} \frac{1}{i!} \left(\frac{\tau_r x}{\overline{\gamma}_{rd} E}\right)^i \tag{20}$$

According to Eqs. (13) and (15), the OP of three options can be written as

$$P_{out} = \begin{cases} F_{\gamma_{sd}}(\gamma_{th}), & DT \\ 1 - (1 - F_{\gamma_{sr}}(\gamma_{th}))(1 - F_{\gamma_{rd}}(\gamma_{th})), & CT \\ \int_0^{\gamma_{th}} \left[1 - (1 - F_{\gamma_{sr}}(\gamma_{th} - t))(1 - F_{\gamma_{rd}}(\gamma_{th} - t))\right] f_{\gamma_{sd}}(t) dt, & JT \end{cases} \tag{21}$$

Finally, the OP theoretical expressions for the three transmission schemes can be obtained by substituting Eqs. (19), (20) and (21) into Eq. (22) respectively.

Under the direct transmission scheme of the theoretical, the system OP can be expressed as

$$P_{out1} = \sum_{i=0}^{\tau_{s1}-1} \chi_{s1} \frac{(-1)^i (1-\tau_{s1})_i \eta_{s1}^i}{(\zeta_{s1}-\eta_{s1})^{i+1} i!} \times \left(1 - \exp\left(-\frac{(\zeta_{s1}-\eta_{s1})\gamma_{th}}{\overline{\gamma}_{sd}}\right) \sum_{j=0}^{i} \frac{(\zeta_{s1}-\eta_{s1})^j}{j!} \left(\frac{\gamma_{th}}{\overline{\gamma}_{sd}}\right)^j\right) \tag{22}$$

Under the collaborative transmission scheme of the theoretical, the system OP can be expressed as

$$P_{out2} = 1 - \sum_{i=0}^{N(\tau_{s2}-1)} \varphi_{s2} \exp\left[-\left(\frac{\zeta_{s2}-\eta_{s2}}{\overline{\gamma}_{sr}} + \frac{\tau_r}{\overline{\gamma}_{sr}E}\right)\gamma_{th}\right] \times \sum_{j=0}^{i+N-1} \sum_{k=0}^{\tau_r N-1} \frac{1}{j!k!} \left(\frac{\zeta_{s2}-\eta_{s2}}{\overline{\gamma}_{sr}}\right)^j \left(\frac{\tau_r}{\overline{\gamma}_{sd}E}\right)^k \gamma_{th}^{j+k} \tag{23}$$

of which $\varphi_{s2} = \chi_{s2}^N \frac{(-1)^i (N-N\tau_{s2})_i \eta_{s2}^i}{(\zeta_{s2}-\eta_{s2})^{i+N} i!}$. Under the joint transmission scheme of the theoretical, the system OP we can be derived in several steps. Firstly, we simplify the OP expression for JT in Eq. (23), i.e.

$$\begin{aligned} P_{out3} &= \int_0^{\gamma_{th}} \left[f_{\gamma_{sd}}(x) - (1 - F_{\gamma_{sr}}(\gamma_{th}-x))(1 - F_{\gamma_{rd}}(\gamma_{th}-x)) f_{\gamma_{sd}}(x)\right] dx \\ &= \int_0^{\gamma_{th}} f_{\gamma_{sd}}(x) dx - \int_0^{\gamma_{th}} (1 - F_{\gamma_{sr}}(\gamma_{th}-x))(1 - F_{\gamma_{rd}}(\gamma_{th}-x)) f_{\gamma_{sd}}(x) dx \end{aligned} \tag{24}$$

Then, according to the characteristics of the CDF function and the nature of the integral, P_{out3} can be further given by

$$\begin{aligned} P_{out3} = F_{\gamma_{sd}}(\gamma_{th}) &- \int_0^{\gamma_{th}} (1 - F_{\gamma_{rd}}(\gamma_{th}-x)) f_{\gamma_{sd}}(x) dx \\ &+ \int_0^{\gamma_{th}} F_{\gamma_{sr}}(\gamma_{th}-x)(1 - F_{\gamma_{rd}}(\gamma_{th}-x)) f_{\gamma_{sd}}(x) dx \end{aligned} \tag{25}$$

Finally, Eqs. (17), (21) and (22) are brought in according to the table of integral functions in the literature, resulting in the theoretical expression for OP as

$$
\begin{aligned}
P_{out3} =& F_{\gamma_{sd}}(\gamma_{th}) - \sum_{p_1=0}^{N\tau_r-1}\sum_{l_1=0}^{p_1}\sum_{n=0}^{\tau_{s1}-1}\phi_{p_1}\phi_{l_1} f_{s1} u_1^{-(l_1+n)-1}\gamma(l_1+n+1, u_1\gamma_{th}) \\
&+ \sum_{k=0}^{N(\tau_{s2}-1)}\sum_{n=0}^{\tau_{s1}-1}\sum_{p_1=0}^{N\tau_r-1}\sum_{l_1=0}^{p_1} f_{s2} f_{s1}\phi_{p_1}\phi_{l_1} u_1^{-(l_1+n)-1}\gamma(l_1+n+1, u_1\gamma_{th}) \\
&- \sum_{k=0}^{N(\tau_{s2}-1)}\sum_{n_1=0}^{k+N-1}\sum_{l_2}^{n_1}\sum_{n=0}^{\tau_{s1}-1}\sum_{p_1=0}^{N\tau_r-1}\sum_{l_1=0}^{p_1} f_{s2}\varphi_1\phi_{l_2} f_{s1}\phi_{p_1}\phi_{l_1} u_2^{-(l_1+n+l_2)-1} \\
&\times \gamma(l_1+n+l_2+1, u_2\gamma_{th}) \qquad (26)
\end{aligned}
$$

of which $\phi_{p_1} = \exp(-\frac{\tau_r\gamma_{th}}{\overline{\gamma}_{rd}E})\frac{1}{p_1!}(\frac{\tau_r}{\overline{\gamma}_{rd}E})^{p_1}$, $\phi_{l_1} = C_{p_1}^{l_1}\gamma_{th}^{p_1-l_1}(-1)^{l_1}$, $\phi_{l_2} = C_{n_1}^{l_2}\gamma_{th}^{n_1-l_2}(-1)^{l_2}$, $f_{s1} = \chi_{s1}\frac{(-1)^n(1-\tau_{s1})_n\eta_{s1}^n}{\overline{\gamma}_{sd}^{n+1}(n!)^2}$,
$\varphi_1 = \exp(-\frac{\zeta_{s2}-\eta_{s2}}{\overline{\gamma}_{sr}}\gamma_{th})\frac{(\zeta_{s2}-\eta_{s2})^{n_1}}{n_1!\overline{\gamma}_{sr}^{n1}}$, $u1 = \frac{\zeta_{s1}-\eta_{s1}}{\overline{\gamma}_{sd}} - \frac{\tau_r}{\overline{\gamma}_{rd}E}$, $f_{s2} = \chi_{s2}^N\frac{(-1)^k(N-N\tau_{s2})_k\eta_{s2}^k}{(\zeta_{s2}-\eta_{s2})^{k+N}\Gamma(N)}\frac{(k+N-1)!}{(N)_k k!}$, $u_2 = \frac{\zeta_{s1}-\eta_{s1}}{\overline{\gamma}_{sd}} - \frac{\zeta_{s2}-\eta_{s2}}{\overline{\gamma}_{sr}} - \frac{\tau_r}{\overline{\gamma}_{rd}E}$, and $\gamma(\cdot,\cdot)$ is the incomplete Gamma function.

In order to reflect the relationship between OP and different transmission schemes intuitively, SNR and energy efficiency, this paper will further analyze the energy efficiency of the system. The energy efficiency of a communications system can usually be defined as the ratio of system throughput to total transmit power, namely

$$\eta = \frac{P_{TPS}}{P_{\text{tot}}} = \frac{(1 - P_{\text{out}})R_0}{P_u + P_{\text{c}}} \qquad (27)$$

where P_{TPS} is the system throughput, R_0 is the system target rate, P_{tot} is the total system transmit power, P_u is the UAV transmit power, $P_c = NP_l + P_0$ is the fixed system power, P_l is the circuit power consumption per antenna of the UAV, and P_0 is the fixed power consumption of the UAV.

4 Analysis of Simulation Results

In this section, Monte Carlo simulations are performed to verify the analytical results. In our simulation, we assume that the number of the fixed antennas is 2, and the SR fading S-D or S-R link experiences either heavy shadowing (HS) or average shadowing (AS). So, their SR fading parameters are $\{m, b, \Omega\} = \{2, 0.063, 0.0005\}$ and $\{m, b, \Omega\} = \{6, 0.265, 0.275\}$ respectively. Other simulation parameters are shown in the following Table 1.

Assuming the same average SNR for each path, i.e. $\overline{\gamma} = \overline{\gamma_{sd}} = \overline{\gamma_{sr}} = \overline{\gamma_{rd}}$. The OP of the different transmission schemes is given as shown in Fig. 2. It can observed that the analytical results match well with the Monte Carlo simulations, showing that

Table 1 Simulation parameter settings

Parameters	Value	Parameters	Value
Satellite orbit	GEO	Bandwidth	5 MHz
Carrier frequency	2 GHZ	Noise temperature	300 K
Maximum satellite antenna gain	49 dB	3 dB angle	0.4°
Noise bandwidth	5 MHz	Interruption threshold	3 dB

the exact performance analysis is accurate. It can also be found that the performance of DT mode is better than CT, and the JT performance is superior to the DT and CT.

Figure 3 shows the OP of the CT and DT schemes with different numbers of antennas. We can obverse that the Outage performance of the system can be effectively improved as the number of antennas increases. This is because the increase of antenna can obtain the higher coding gain in the cooperative relay transmission, and allow the system to obtain more diversity and array gain, which can compete interference effectively, demonstrating the superiority of multi-antenna UAV relay transmission.

The satellite transmit power P_s is assumed to be 20 dBm, the circuit power consumption per antenna of the UAV P_c is 10 dBm and the fixed power consumption of the UAV P_0 is 30 dBm. Simulations of the system energy efficiency of CT and JT schemes are performed by controlling the transmit power of the UAV P_R, the

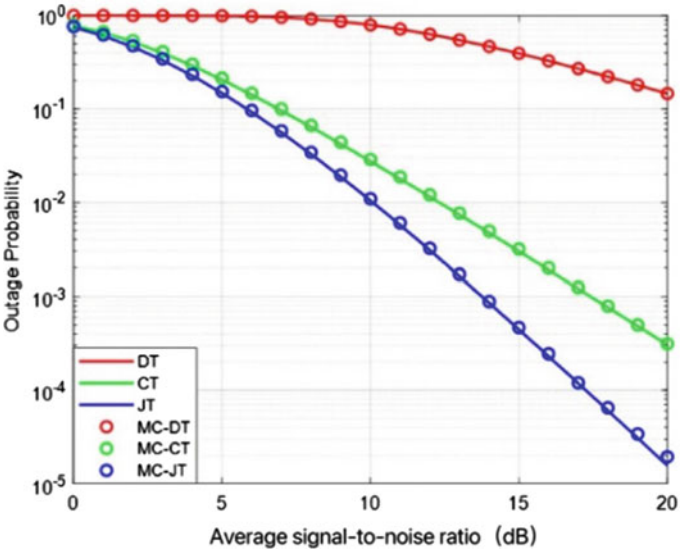


Fig. 2 Impact of satellite channel parameterson on the OP of the considered system

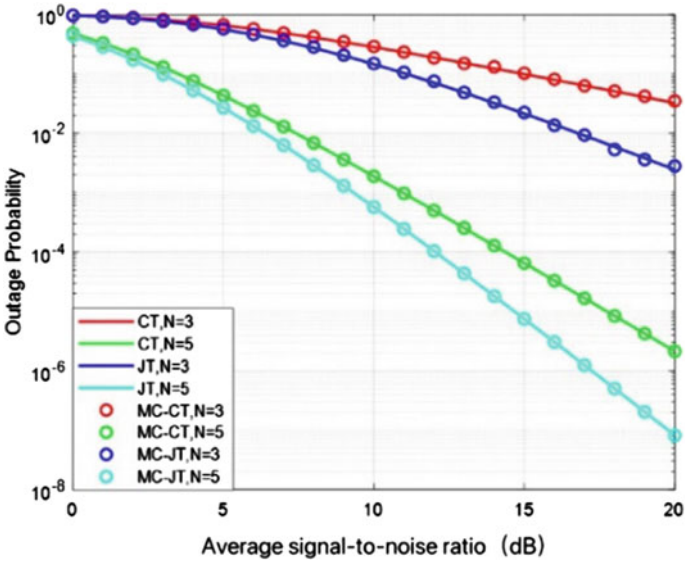


Fig. 3 Effective of antenna count on the performance of a relay collaboration

results are shown in Fig. 4. It can be seen that before reach maximum energy efficiency, the increase of transmit power can significant improve the system energy efficiency, for the main constraint is the limited transmit power. However, after reaching the maximum energy efficiency, increasing transmit power will decrease the system performance, this is because power increase does not significantly increase the system rate. Besides, with the same number of antennas, the energy efficiency of JT is always better than that of the CT, confirming the superiority of JT.

5 Concluding Remarks

In this paper, we have proposed three transmission schemes based on the qualities of the channels between the satellite and user and derived the closed-form expression for the OP of the considered system. Analytical and simulation results have shown that the system performance can be improved by using the joint transmission. Moreover, by changing the number of fixed antennas, we can meet the users' requirement accordingly. This presented work can be regarded as a starting point that the system relies on the use of UAV relay network to improve the wireless transmission performance.

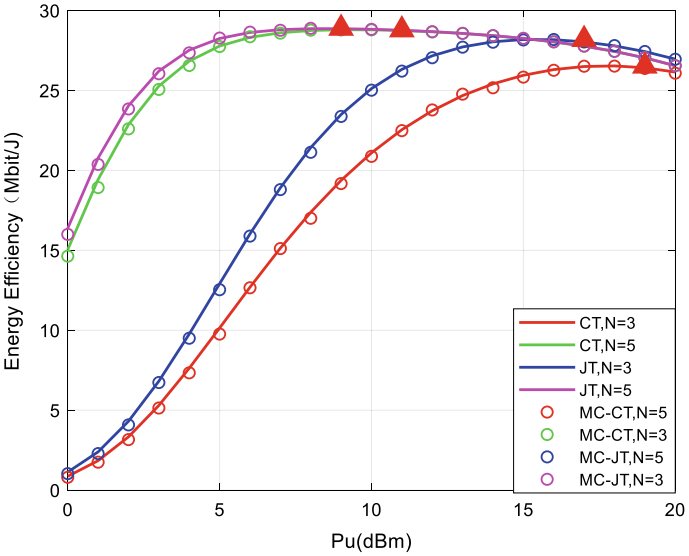


Fig. 4 System energy efficiency analysis diagram

Acknowledgements This work was supported by The Provincial College Student Innovation and Entrepreneurship Training Program in 2022 under Grands 202210293075Y. *(Mingyi Ji and Bo Xu contributed equally to this work.)*

References

1. Yao Y, Peng Z, Xiao B (2018) Parallel hyper-heuristic algorithm for multi-objective route planning in a smart city. IEEE Trans Veh Technol 67(11):10307–10318
2. Lin Z, Lin M, Ouyang J, Zhu W-P, Panagopoulos AD, Alouini M-S (2019) Robust secure beamforming for multibeam satellite communication systems. IEEE Trans Veh Technol 68(6):6202–6206
3. Lin Z, Lin M, Wang JB, Huang YM, Zhu W-P (2018) Robust secure beamforming for 5G cellular networks coexisting with satellite networks. IEEE J Sel Areas Commun 36(4):932–945
4. Ruan YH, Li YZ, Wang CX, Zhang R (2018) Energy efficient adaptive transmissions in integrated satellite-terrestrial networks with ser constraints. IEEE Trans Wireless Commun 17(1):210–222
5. Kong H, Lin M, Zhang J, Ouyang J, Wang J-B, Upadhyay PK (2022) Ergodic sum rate for uplink NOMA transmission in satellite-aerial-terrestrial integrated networks. Chin J Aeronaut 35(9):58–70
6. Wang Z, Lin M, Ouyang J, Wang J-B, Zhu W-P, Al-Dhahir N Multi-objective robust beamforming for integrated satellite and aerial networks supporting heterogeneous services. IEEE Trans Wirel Commun. https://doi.org/10.1109/TWC.2023.3246460
7. Liu XY, Zhao B, Lin M, Ouyang J, Wang JB, Wang J (2023) IRS-aided uplink transmission scheme in integrated satellite-terrestrial networks. IEEE Trans Vehic Tech 72(2):1847–1861
8. Huang Q, Lin M, Wang J-B, Tsiftsis TA, Wang J (2020) Energy efficient beamforming schemes for satellite-aerial-terrestrial networks. IEEE Trans Commun 68(6):3863–3875

9. Zeng S, Zhang H, Bian K, Song L (2018) UAV relaying: power allocation and trajectory optimization using decode-and-forward protocol. In: Proceedings 2018 IEEE international conference on communications workshops (ICC Workshops), pp 1–6
10. Arti MK, Jain V (2017) Relay selection-based hybrid satellite-terrestrial communication systems. IET Commun 11(17):2566–2574
11. Lin M, Lin Z, Zhu W-P, Wang J-B (2018) Joint beamforming for secure communication in cognitive satellite terrestrial networks. IEEE J Sel Areas Commun 36(5):1017–1029
12. Yan S, Malaney R (2016) Location-based beamforming for enhancing secrecy in Rician wiretap channels. IEEE Trans Wireless Commun 15(4):2780–2791
13. Bao J, Liu X, Xiang Z, Wei G (2020) Multi-objective optimization algorithm and preference multi-objective decision-making based on artificial intelligence biological immune system. IEEE Access 8:160221–160230
14. Bankey V, Upadhyay PK, Da Costa DB, Bithas PS, Kanatas AG, Dias US (2018) Serformance analysis of multi-antenna multiuser hybrid satellite-terrestrial relay systems for mobile services delivery. IEEE Access 6:24729–24745
15. Zhang JX, Evans B, Imran MA, Zhang X, Wang WB (2016) Green hybrid satellite terrestrial networks: fundamental trade-off analysis. In: IEEE Vehicular Technology Conference. IEEE

Composition Analysis and Identification of Ancient Glass Products Based on Random Forest Algorithm

Tao Wang and Cheng Wang

Abstract This work investigates the machine learning-based identification method of ancient glass products. Exposure to burial environments can cause significant chemical exchange between the glass and environmental elements, leading to alterations in the composition and appearance of the artifacts and potentially influencing their correct classification. To address this issue, we have developed a statistical model to predict the original composition content of weathered glass artifacts based on composition and appearance data from known glass artifacts. In addition, we have employed the Random Forest algorithm to classify unknown glass samples. Our results demonstrate that the Random Forest algorithm is highly effective in distinguishing between different classes of glass artifacts. Through the analysis of composition content, we have observed substantial differences in the composition of glass artifacts of different types. Overall, our study contributes valuable insights into the identification of ancient glass products and presents a promising approach for predicting their original composition and classification.

Keywords Ancient glass product · Weathering effect · K-sample Anderson–Darling test · Random forest algorithm

T. Wang · C. Wang (✉)
Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
e-mail: cwang@tjnu.edu.cn

T. Wang
e-mail: wangtgr@163.com

T. Wang
Department of Communication Engineering, College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China

C. Wang
Department of Intelligence Science and Technology, College of Artificial Intelligence, Tianjin Normal University, Tianjin 300387, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_15

1 Introduction

Glassware is valuable evidence of the early trade between China and the West. The utilization of glass as an ornamental material can be traced back to its production in West Asia and Egypt, and then introduced to China. The Chinese people then assimilated this foreign technology and began producing glassware utilizing local materials. Despite bearing similarities in appearance to imported glass, the chemical composition of Chinese glassware was different. Specifically, the primary constituent of glass, i.e., quartz sand (SiO_2), remains consistent, but the fluxes utilized in the firing process were different, which lead to variations in the final product's chemical composition. For instance, the glass produced using lead ore as a flux, i.e., lead-barium (Pb-Ba) glass contains elevated levels of lead and barium, whereas the utilization of plant ash with high potassium content as a flux resulted in high-K glass [1].

Ancient glass products are susceptible to burial environments which lead to the weathering effect. The weathering effect causes a large amount of element exchange between the glass and the environment, which changes the composition and appearance of glass products, and thus decreases the ratio of classification accuracy. Therefore, accurately predicting the original composition content of weathered glass products and classifying category-unknown glass samples are significantly important to scientific research on ancient glass products. Based on the dataset of a group of category-known glass samples' composition, this paper aims to resolve two problems as follows.

(a) Predict the original composition of the weathered glass products using a statistic model;
(b) Classify category-unknown glass samples via a machine learning method.

2 Dataset

The dataset used in this work comes from the C-question of the 2022 China Undergraduate Mathematical Contest in Modelling [2] (CUMCM), which includes three forms. Form 1 provides some external appearance features of the category-known samples. Form 2 provides the composition content and weathering conditions of category-known glass instances. Form 3 gives the data of category-unknown instances.

Form 2 will be used as the main dataset in two problems. In Sect. 3, the original composition content of all valid weathered instances will be predicted based on Form 2. In Sect. 4, all valid instances in Form 2 and corresponding external appearance features existed in Form 1, will be used as train and test sets, and instances in Form 3 are to be predicted. Not all the instances in Form 2 are valid, so the first step of prediction is data preprocessing.

3 Predict Original Composition Content

3.1 Data Preprocessing

1. Elimination of Invalid Instances

The instances taking the sum of all composition content out of the range between 85 and 105% were not used in this work.

2. Correction of Weathering Instance Types

Since the weathering effect is a continuous process, unweathered areas may exist in a weathered sample. Table 1 illustrates the classification detail of the final type of instances.

3.2 Model

$\mathcal{D} = X_{iw}^{c}$ $(c = 0\,or\,1, i = 0, 1, \ldots 13, w = 0\,or\,1)$ denotes the dataset where $c = 0, 1$ corresponds to high-K glass and Pb-Ba glass, $i = 0, 1, \ldots 13$ corresponds to chemical composition SiO_2, Na_2O, …SO_2, and $w = 0, 1$ corresponds to unweathered glass and weathered glass. The symbol X_{iw}^{c} denotes the content set of corresponding glass type, chemical composition, and weathering condition. For instance, X_{10}^{0} represents the Na_2O content set of all unweathered high-K glass instances.

Weathering causes a large amount of exchange between the elements in the glass and environmental elements. However, since the difference in the chemical composition's existence forms and element contents, for every chemical composition the effect caused by weathering is not equal, namely weathering effect with compositional differences. This paper establishes a model from the composition view rather than the instance view.

We assume that the discrepancy in chemical composition content between weathered and unweathered glass samples is mainly derived from weathering. Based on this assumption, there are two possible conditions relating to the difference in some

Table 1 The final types of samples

Sample type	Sampling area type	Final type
Weathered	Unweathered	Unweathered
Weathered	Others	Weathered
Unweathered	Weathered	Weathered
Unweathered	Others	Unweathered

composition content of two types of glass. If the difference is not significant, considering that the weathering effect hardly impacts the composition, therefore, it's reasonable to directly view the content of a weathered instance as the content of the corresponding unweathered instance. If the difference is statistically significant, the difference needs to be calculated and should be used to predict the original composition content. In this paper, a nonparametric test method, the K-sample Anderson–Darling test [3, 4], was used to estimate whether a weathered set and an unweathered set for some type of glass are sampled from an identical population. The details of the model are as follows.

For a given glass type of c and a given composition, do the K-sample Anderson–Darling test ($\alpha = 0.05$).

H_0: S_{i0}^c and S_{i1}^c subject to identical distribution. H_1: S_{i0}^c and S_{i1}^c not subject to identical distribution.

Under the given significance level, if the null hypothesis cannot be rejected, the composition content before weathering $\hat{S}_{i0}^c = S_{i0}^c$. Otherwise considering that $\hat{S}_{i0}^c$ satisfies the following equation

$$S_{i0}^c = \hat{S}_{i0}^c + \Delta_w \tag{1}$$

where Δ_w denotes the variation of composition content caused by weathering. According to Eq. (1), $\Delta_w = S_{i0}^c - \hat{S}_{i0}^c$. The variation can be estimated by $\Delta_{\hat{w}} = \overline{S} - \overline{S}$. Since there exist some outliers in the composition content set, using the median of the samples to represent the mean of samples, therefore the variation is finally estimated by the following equation

$$\Delta_{\hat{w}} = median\left(S_{i0}^c\right) - median\left(S_{i1}^c\right) \tag{2}$$

where $median(\cdot)$ returns the median. Based on Eq. (2), $\hat{S}_{i0}^c = S_{i0}^c - \Delta_{\hat{w}}$.

In conclusion, the algorithm pseudocode for predicting chemical composition content is as follows.

Algorithm 1: Algorithm for predicting original composition content

Input: Dataset, $\mathcal{D} = X_{iw}^{c}(c = 0, 1, i = 0, 1, \cdots 13, w = 0, 1)$

for $c = 0, 1$ do:

  for $i = 0, 1, \cdots, 13$ do:

    K-sample AD_test(S_{i0}^{c}, S_{i1}^{c}), where $\alpha = 0.05$

    if can not reject the null hypothesis

      $\hat{S}_{i0}^{c} = S_{i0}^{c}$

    else

      $\Delta_{\hat{w}} = median\left(S_{i0}^{c}\right) - median\left(S_{i1}^{c}\right)$

      $\hat{S}_{i0}^{c} = S_{i0}^{c} - \Delta_{\hat{w}}$

    end

  end

end

output: $\mathcal{D}_{out} = \left\{\hat{S}_{i0}^{c}\right\}$

3.3 Results and Discussion

We can get the original composition content of the weathered instances, through the algorithm in Sect. 3.2. Calculating whether the sum of a weathered instance's all original composition content is between 85 and 105%, the result is that 92.86% of instances satisfy the criteria. Using K-samples Anderson–Darling test ($\alpha = 0.05$) to test the dataset again whose weathered instances' contents are replaced by corresponding original content, approximately 96.88% of the instances not can reject the null hypothesis, which means that 96.88% of instances are regarded as sampled form identical population. Table 2 shows the test methods and pass ratio.

This model is effective in predicting the original composition content of the weathered instances. However, there exist some disadvantages. One of them is that this model omits the relationship between chemical compositions during model establishing, which might mainly contribute to the pass ratio of the sum of composition content just 92.86% lower than 100%.

Table 2 The test of prediction and result

Test method	Result (pass ratio) (%)
The sum of composition content	92.86
K-samples Anderson–darling test	96.88

4 Classification of Category-Unknown Glass Instances

4.1 Classification and Results

This question belongs to the small sample classification. Since the data set includes both numeric and category features, this paper established a Random Forest classifier [5], (Tree num is 100, the criterion is Gini index [6]), with AUC of ROC curve as the criteria to finish this task. After training on the train set, the AUC of the test set is 1.0, which means that all of the test instances are classified correctly (Fig. 1).

Using the trained model to predict the category-unknown glass instances, Table 3 shows the result. All instances are classified correctly.

4.2 Analysis and Discussion

The AUC score equal to 1.0 means that the classifier knows which chemical compositions determine the types of glass instances precisely. Figure 2 illustrates the contribution score to the final prediction of each composition, and it's evident that the contribution score of the top five compositions is higher than other compositions. Therefore, the top five compositions will be focused on in future analysis.

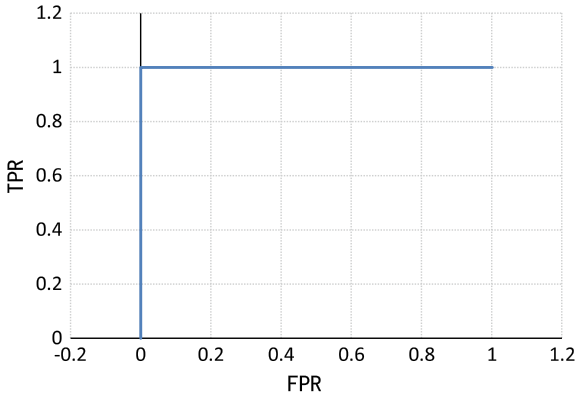


Fig. 1 ROC curve on the test set

Table 3 The types of category-unknown instances

Glass samples	Category	Glass samples	Category
A1	High-K	A5	Pb-Ba
A2	Pb-Ba	A6	High-K
A3	Pb-Ba	A7	High-K
A4	Pb-Ba	A8	Pb-Ba

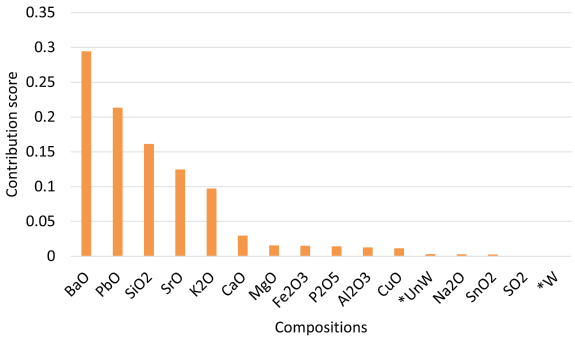


Fig. 2 Contribution score of compositions (starting with * means the weathering conditions)

Adding PbO content and BaO content together as a new composition, PbO-BaO, content, thus making the Pearson correlation analysis finding that PbO-BaO and SiO_2 are highly related in content. Just keeping, therefore, three compositions, PbO-BaO, K_2O, and SrO in future analysis.

Figures 3, 4 and 5 illustrate the pairwise analysis abort the above three compositions. It's evident that the distributions of PbO-BaO, K_2O, and SrO of different types of glass are very different, almost distributing in two orthogonal directions. In other words, categories of glass have great distinguishability in some special chemical compositions, which also accounts for why the AUC score of the test is 1.0. This distinguishability, however, is derived from the occasionality of a small dataset or is the intrinsic structure of the population embodied in the dataset, more work is needed in the future.

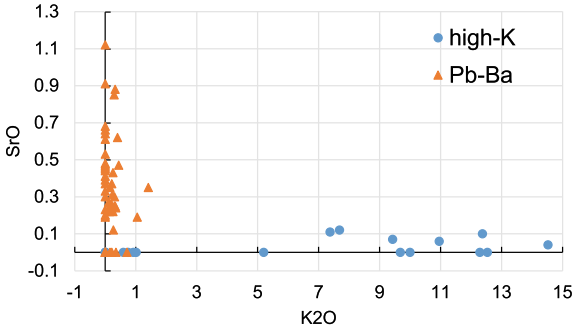


Fig. 3 The relation of SrO and K2O of different glass types

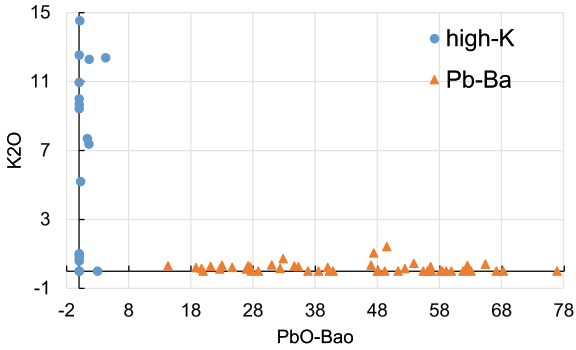


Fig. 4 The relation of PbO-BaO and K2O of different glass types

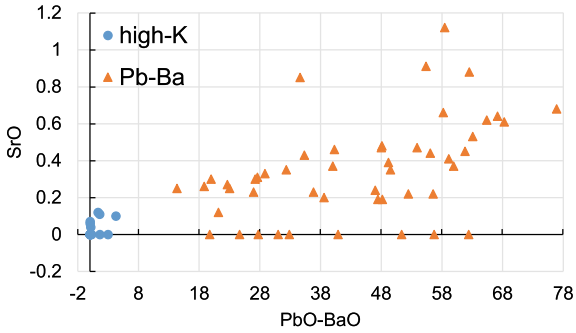


Fig. 5 The relation of SrO and PbO-BaO of different glass types

5 Conclusions

In this paper, we built a statistic model, based on the *K*-sample Anderson–Darling test, to predict the composition content before weathering of the weathered instances and a Random Forest learner to classify type-unknown glass instances. From the point of result, both models are effective to handle the corresponding task.

References

1. Wen R, Zhou ZQ, Ma J, Wang JX (2016) Chemical analysis of ancient glass beads from Shirenzigou sites in Balikun county. Xinjiang Spect Spect Anal 36(09):2961–2965
2. http://www.mcm.edu.cn/
3. Scholz FW, Stephens MA (1987) K-sample Anderson-darling tests. J Am Stat Assoc 82(399):918–924
4. Scholz F, Stephens M (1986) K-sample anderson-darling tests of fit, for continuous and discrete cases. University of Washington, Technical Report (81)
5. Breiman L (2001) Random forests. Machine learning 45:5–32
6. Breiman L, Friedman J, Olshen R, Stone C (1984) Cart. Classification and regression trees

CNN-Based Sea-Surface Target Detection Using Continuous Wavelet Transform

Jingchen Ni, Haoru Li, Lilin Xu, and Jing Liang

Abstract Building a high-performance target detector under sea clutter background has always been a difficult but vital problem leading to various detection methods. In this work, we propose a ResNet-CWT detector. ResNet is adopted to gain a higher target detection probability (P_d) and solve the network degradation problem as the network is deepening, while CWT is applied to extract time-frequency feature effectively. The result of tests shows that the P_d of ResNet-CWT detector is higher than other CNN detectors including LeNet-5-STFT, AlexNet-STFT and ResNet-STFT. Also, when increasing the observation window length, the P_d of target rises and the false alarm rate decreases.

Keywords CWT · ResNet · Target detector · Sea clutter

1 Introduction

Detection of small targets on the sea surface based on real radar data is a hard but essential topic in many fields [1]. However, it's hard to build an effective detectors as marine radar echoes are buried in the sea clutter in the process of target detection. Traditional statistical methods tried to fit the amplitude distribution of the sea clutter and deploy constant false alarm detection (CFAR) to detect the target [2], but they failed to achieve high performance as the sea state varies. Thus, it's essential to establish new detection methods.

Along with the fast development of artificial intelligence, deep learning methods have been widely studied and employed in the radar field [3]. Convolutional neural network (CNN), one of the deep learning methods, has already been applied to detect

Jingchen Ni, Haoru Li and Lilin Xu are co-first authors.

J. Ni · H. Li · L. Xu · J. Liang (✉)
University of Electronic Science and Technology of China, Chengdu, China
e-mail: liangjing@uestc.edu.cn

J. Ni
e-mail: 2020020908023@std.uestc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_16

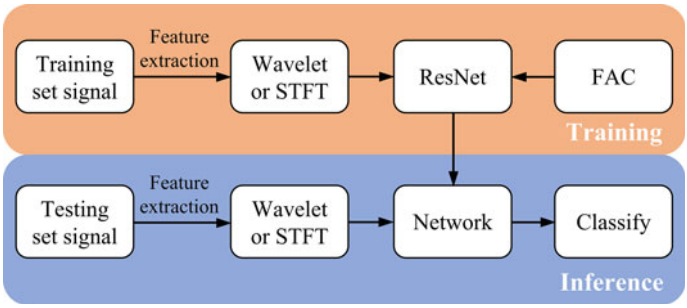


Fig. 1 Working flow of the proposed ResNet-CWT detector

marine target from real radar data [4] and suppress sea clutter component in radar signals [5] due to its excellent ability in processing two-dimensional data [6]. As the computing capability of chips has been greatly improved recently, more complex CNNs are applied in this field. In 2019, Liu's team used fractional Fourier transform (FRFT) and residual network (ResNet) [7] to distinguish target [8].

For detection via CNN, it is necessary to transform sequential echo data into two-dimensional data via time-frequency transform. Variants of Fourier transform, including short-time Fourier transform (STFT) [4, 9] and FRFT [8], are widely used to extract two-dimensional time-frequency features as the input of CNN.

However, most Fourier transforms encounter a common problem that infinite-length trigonometric functions are not capable of capturing sudden variations in sea clutter background, which is critical for target detection. Instead, wavelet transform with finite length is preferred for this task. Furthermore, wavelet transform is widely employed in data fusion and data processing for remote sensing.

In this work, we propose a ResNet-CWT detector with false alarm control (FAC) method [10]. In this detector, CWT is innovatively used in radar sea-surface target detection based on CNN. ResNet shows a higher target detection probability and a lower false alarm under the same conditions than simpler networks containing AlexNet [11] and LeNet-5 [12]. Experiments are also conducted to evaluate the effect of increasing the observation window size on the performance of the detector. Our results demonstrate that enlarging the feature representation can significantly enhance the detection target probability as the network captures more details. The working flow of the proposed detector is shown in Fig. 1.

The structure of the paper is as follows. Section 2 introduces evaluation criteria of detectors. Section 3 shows the feature extraction methods used and the data preprocessing. Section 4 describes the ResNet detector. Section 5 provides the test performance and analysis. Section 6 draws the conclusion.

Table 1 Four situations of classification

		Prediction	
		Positive	Negative
Sample	Target	TP	FN
	Clutter	FP	TN

2 Evaluation Criteria of Detectors

The marine target detection can be expressed as a binary classification problem, where each sample is classified into two types: clutter (negative) or target (positive). There are four situations of classification shown in Table 1.

(1) True positive (TP) means a target sample is predicted to be positive.
(2) True negative (TN) means a clutter sample is predicted to be negative.
(3) False positive (FP) means a clutter sample is predicted to be positive.
(4) False negative (FN) means a target sample is predicted to be negative.

Hence, target detection probability (P_d) ,false alarm probability (P_{fa}) and detection accuracy (Acc) can be defined in the following form.

$$\begin{aligned} P_d &= \frac{TP}{TP + FP} \\ P_{fa} &= \frac{FP}{TN + FP} \\ Acc &= \frac{TP + TN}{TP + TN + FP + FN} \end{aligned} \tag{1}$$

3 Feature Extraction

3.1 STFT

The short-time Fourier transform (STFT) is a widely used signal processing technique that provides a time-frequency analysis of non-stationary signals.

The STFT is defined by the following formula:

$$X(\tau, \omega) = \int_{-\infty}^{\infty} x(t) w(t - \tau) e^{-j\omega t} dt \tag{2}$$

where $x(t)$ is the signal, $w(t)$ is Blackman window function, τ is the time shift, ω is the angular frequency, and $X(\tau, \omega)$ is the STFT of the signal.

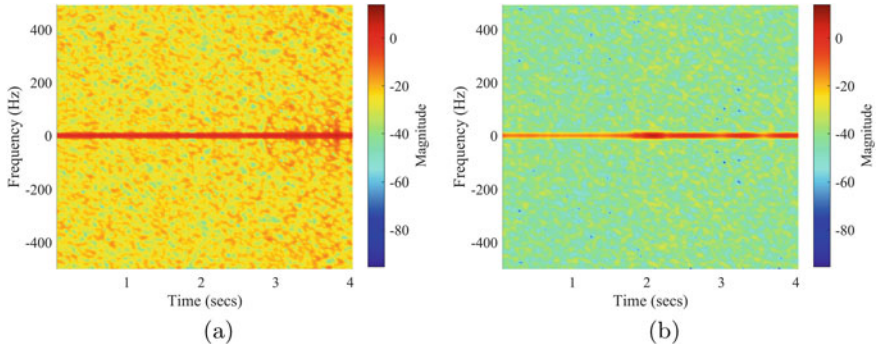


Fig. 2 STFT of **a** clutter sample and **b** target sample in 19980204_220046_ANTSTEP of VH polarization

For actual discrete signal, the short time Fourier transform is defined as:

$$X(m,n) = \sum_{k=-\infty}^{\infty} x(k)w(k-m)e^{-j\frac{2\pi}{N}kn} \tag{3}$$

where $X(m, n)$ is the STFT coefficient at time m and frequency n, $x(k)$ is the discrete-time signal, $w(k)$ is a window function, N is the length of the window (Fig. 2).

3.2 Continuous Wavelet Transform

Compared with STFT, continuous wavelet transform (CWT) is generally considered to be more effective in processing non-stationary time series, because STFT is limited by the fixed window function size [13]. This transform is particularly useful for signals that have components with varying frequencies over time.

The continuous wavelet transform of a signal $x(t)$ can be defined as:

$$C(a,b) = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{a}}\psi^*\left(\frac{t-b}{a}\right)dt \tag{4}$$

where ψ^* is the complex conjugate of the mother wavelet function $\psi(t)$, and a and b are the scale and time parameters, respectively. The scale parameter controls the frequency resolution of the wavelet and the time parameter controls the location of the wavelet in time.

In this work, Morse wavelet is chosen as the wavelet function ψ, which is defined in the frequency domain as

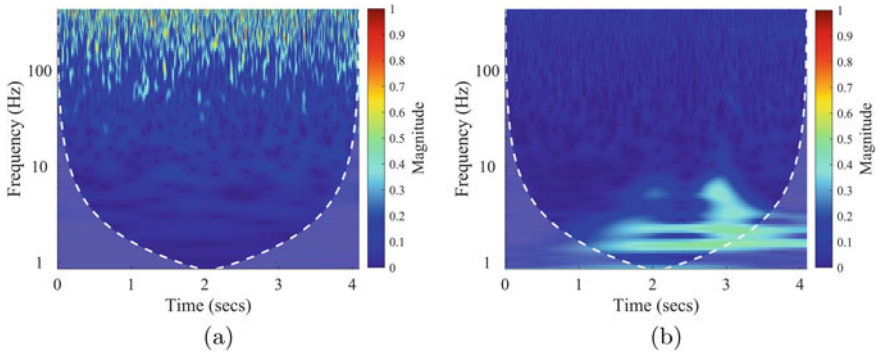


Fig. 3 CWT of **a** clutter sample and **b** target sample in 19980204_220046_ANTSTEP of VH polarization

$$\psi_{P,\gamma}(\omega) = U(\omega) a_{P,\gamma} \omega^{\frac{P^2}{\gamma}} e^{-\omega^{\gamma}} \tag{5}$$

where $U(\omega)$ is the unit step, $a_{P,\gamma}$ is a normalizing constant, P^2 is the time-bandwidth product, and γ characterizes the symmetry of the Morse wavelet. $\gamma = 3$ and $P^2 = 60$ is selected to obtain a symmetric, nearly Gaussian, and time-frequency concentrated member in Morse wavelet family [14].

Two CWT scalogram, whose frequency domain is plotted on a logarithmic scale are shown in Fig. 3. Regions with significant edge effects are delineated in the gray regions outside the dashed white line. CWT provides a highly recognizable feature that can be learned by neural network.

3.3 Data Sampling and Preprocessing

This study employs 10 IPIX datasets acquired from Lake Ontario's shore in varying sea clutter environments. Table 2 displays the fundamental details of these 10 datasets.

Each data set, except the first, contains 28 distance units, while the first data set has 27 distance units. Each distance unit comprises approximately 60 s of 60,000 time series, which were sampled at a frequency of 1000 Hz. In this work, we sample the each unit into sequence vectors for detection, using the same observation window length N, with a step size of M:

$$\boldsymbol{x}_i = [x(M \cdot (i-1)+1), \ldots, x(M \cdot (i-1)+N)] \quad i = 1, 2, \ldots \tag{6}$$

Samples that correspond to targets are assigned a label of 1, while those corresponding to clutter are assigned a label of 0. After randomly shuffling the order of these samples, 70% of the samples constitute the training dataset, 20% of the samples

Table 2 Primitive data selected from 1998 IPIX dataset

Data set	Target bin
19980204_220046_ANTSTEP	13–17
19980204_220325_ANTSTEP	18–20
19980204_220546_ANTSTEP	11–13
19980204_220849_ANTSTEP	7, 8
19980204_221104_ANTSTEP	4, 5
19980205_170935_ANTSTEP	7, 8
19980205_171203_ANTSTEP	7, 8
19980205_171437_ANTSTEP	7, 8
19980205_171654_ANTSTEP	7, 8
19980205_180558_ANTSTEP	7, 8

constitute the cross-validation dataset, and the rest 10% of the samples constitute test dataset.

The $\boldsymbol{x}_i$ will be processed by CWT to get the input of ResNet. The raw output of CWT is too large to serve as an input for ResNet, which considerably reduces the detection speed. Therefore, we evenly sample the CWT data to ensure that the length of its longest dimension does not exceed 128. For the same reason, for tests using STFT, we set the length of window function to 128.

4 ResNet Detector

4.1 Residual Network

Generally, increasing the network depth can improve the network performance. However, when the network depth reaches a certain level, the network performance will decline, which is due to gradient explosion and disappearance. To solve this problem, Microsoft Research proposed ResNet with residual block structure in 2015 [7]. The structure of the residual block is shown in Fig. 4.

With the residual block, the convolutional layer can combine the feature outputs from multiple layers, which can obtain a more accurate feature calculation effect. Also, adding extra branch can make more gradient flow paths in network training.

In our ResNet detector, we choose ResNet50 as our backbone as it is deep enough for binary classification task of marine time-frequency data.

The structure of ResNet50 is shown in Fig. 5.

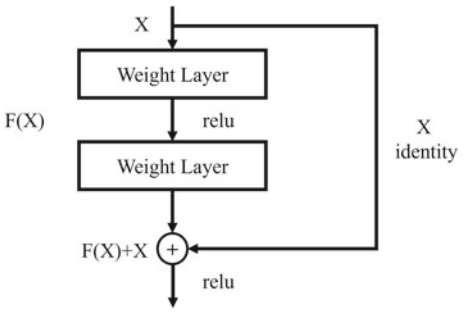


Fig. 4 Residual block

$$7 \times 7, 64 \text{ stride } 2 \xrightarrow{3 \times 3, \text{ max pool, stride } 2} \begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3 \rightarrow \begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$$
$$\rightarrow \begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6 \rightarrow \begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3 \rightarrow \text{average pool, fully connected layer, softmax}$$

Fig. 5 Structure of ResNet50

4.2 Network Training

The parameters of the network are set as follows. The batch size is set at 128. The max epoch number is 100 and one epoch includes 171 iterations and the parameter optimizer is Adam. In order to make the model converge better, we use exponential learning rate decay to automatically decay the learning rate in each epoch, the learning rate is initially set to 0.01 and drop rate is set to 0.9.

The loss function used in this work is cross-entropy loss function which can be expressed as:

$$Loss = -\frac{1}{n}\sum_{i=1}^{n} t_i \log(P_i) \tag{7}$$

where n is the total number of classes, t_i is a one-hot coded vector that indicates the label of each sample, and P_i is the output of softmax layer representing the probability of belonging to each given class.

4.3 False Alarm Rate Control

To restrict the false alarm probability to desired value, a false alarm rate control method proposed in a previous study is applied to this work. The output of the network contains two items, where the first one represents the probability that the input sample is a clutter sample. When the first item of all clutter samples is sorted in ascending order, the threshold η which decides the type of sample is located at the $P_{fa} \times lenth\_of(queue)$-th place of the probability queue. The process above can be expressed as

$$\begin{aligned} &Q_{clutter}\text{: item of all clutter samples in ascending order} \\ &\eta = Q_{clutter}[P_{fa} \times length\_of(Q_{clutter})] \end{aligned} \tag{8}$$

However, it should be noticed that we use cross-validation dataset to set the value of η, and the test is implemented on test set. The method cannot guarantee exact equality between the actual P_{fa} and desired P_{fa} in test set. In our experiment, the relative error between actual P_{fa} and desired P_{fa} is less or equal to 30% in most situation.

5 Results and Comparison

Firstly, we set the length of the observation window at 512 and test the performance in different polarization channels, convolutional networks and feature extraction methods. In Table 3 we can see that when setting the controlled P_{fa} at 0.01, all tests get an actual P_{fa} close to 0.01 which demonstrates the effectiveness of false alarm control approach. Also, ResNet, a complex network, gets a higher P_d under controlled P_{fa} 0.01 compared to AlexNet and LeNet-5. For four polarization channels, VH channel gets a higher P_d in all network tests than other channels. In the test which takes CWT as the input, the same conclusion can be obtained. Therefore, VH channel can better express the features of the presence and absence of the target under sea clutter background, and we choose VH channel to finish the follow-up experiments. Moreover, for feature extraction, we can find that P_d of CWT method is slightly lower than that of STFT method. This is possibly caused by points selection of the result of CWT, leading to 75% information dropout in the CWT. As the observation window length is set to 512, the size of STFT input is 128 in frequency domain and 13 in time domain, while the size of CWT input is 61 in frequency domain and 128 in time domain. STFT provides more abundant information in frequency domain, which is more valuable for detecting targets (Fig. 6).

To further study the effect of feature extraction and obtain a detector that is qualified to be applied to the real situation, we set the target P_{fa} to 0.001 and implement more tests for larger Ns, whose results are shown in Table 4.

Table 3 Performance of 3 networks in different polarization channels

Feature	Channel	Network	Actual P_{fa}	P_d	Acc	Feature	Channel	Network	Actual P_{fa}	P_d	Acc
STFT	**VH**	**ResNet50**	**0.0085**	**0.8997**	**0.9820**	**CWT**	**VH**	**ResNet50**	**0.0134**	**0.8784**	**0.9753**
STFT	VH	AlexNet	0.0081	0.8085	0.9728	CWT	VH	AlexNet	0.0060	0.7964	0.9734
STFT	VH	LeNet-5	0.0092	0.8207	0.9731	CWT	VH	LeNet-5	0.0117	0.7964	0.9683
STFT	VV	ResNet50	0.0152	0.5957	0.9443	CWT	VV	ResNet50	0.0155	0.5957	0.9440
STFT	VV	AlexNet	0.0148	0.3982	0.9240	CWT	VV	AlexNet	0.0134	0.3131	0.9164
STFT	VV	LeNet-5	0.0134	0.4438	0.9300	CWT	VV	LeNet-5	0.0134	0.2523	0.9101
STFT	HH	ResNet50	0.0145	0.5046	0.9354	CWT	HH	ResNet50	0.0111	0.5127	0.9475
STFT	HH	AlexNet	0.0184	0.3374	0.9145	CWT	HH	AlexNet	0.0120	0.2766	0.9139
STFT	HH	LeNet-5	0.0120	0.3283	0.9193	CWT	HH	LeNet-5	0.0106	0.2128	0.9085
STFT	HV	ResNet50	0.0110	0.7812	0.9674	CWT	HV	ResNet50	0.0090	0.7818	0.9728
STFT	HV	AlexNet	0.0110	0.7173	0.9607	CWT	HV	AlexNet	0.0099	0.7173	0.9617
STFT	HV	LeNet-5	0.0078	0.7112	0.9630	CWT	HV	LeNet-5	0.0113	0.6474	0.9531

Bold text shows two best results in two features

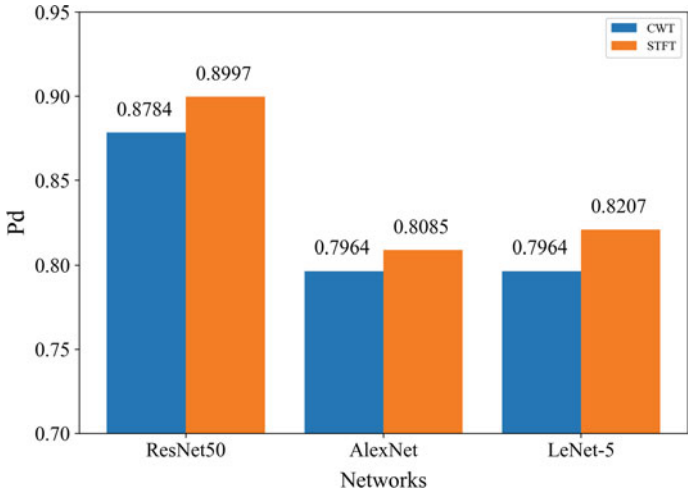


Fig. 6 P_d of different networks in VH channel when observation window length is 512 and controlled P_{fa} is 0.01

Table 4 Performance of VH channel CWT and STFT based on ResNet

Window length	Feature	P_d	P_{fa}	Acc
1024	STFT	0.7069	0.0014	0.9716
1024	**CWT**	**0.7406**	**0.0011**	**0.9758**
2048	STFT	0.8963	0.0011	0.9899
2048	**CWT**	**0.9173**	**0.0007**	**0.9919**
4096	STFT	0.9937	0.0018	0.9977
4096	**CWT**	**0.9964**	**0.0007**	**0.9990**

Bold text shows the best results when Window length is consistent

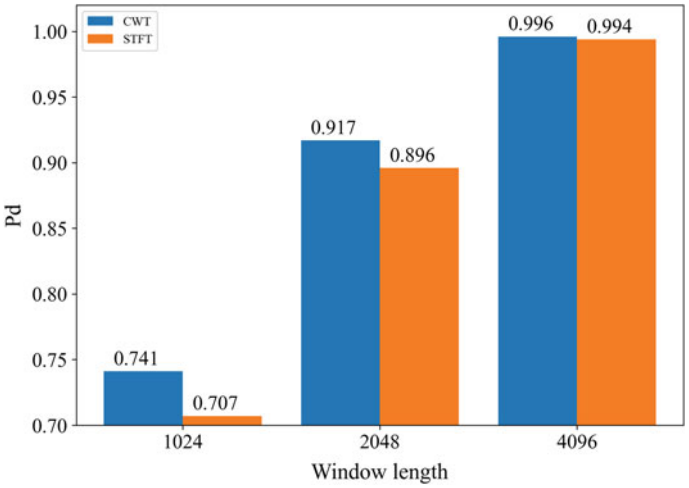


Fig. 7 Comparison of P_d between STFT and CWT using ResNet changing observation window when P_{fa} is 0.001

Since the larger observation window length includes more information and features, we try to increase the length of observation window from 512 to 4096. The result of test is shown in Table 4. When larger observation window length is set to 1024 and 2048, the size of frequency domain of CWT input increases to 71 and 81. As the result, P_d of CWT exceeds the P_d of STFT with a lower P_{fa}. If the observation window length is 4096, both feature extraction methods can achieve a near 1.0 P_d, while the size of input of CWT is 37.2% less than that of STFT and the P_{fa} of CWT is 40.97% of that of STFT. It is worth noting that while the size of frequency domain of CWT input rises, the size of the time domain of CWT input is kept at 128 due to the points selection, which means in the case of 4096 observation window length, the size of CWT input will be shrunk from 4096 to 128 and nearly 97% original CWT result is abandoned. It proves that the growth of observation time provides more original information for CWT analysis and improves its performance effectively, even though a large portion of CWT result is not used by the detector. It also demonstrates that CWT can extract different resolutional features, capture the local similarity of signals and clearly detect the outliers in the signal. In conclusion, compared with STFT which is limited by observation window size, CWT achieves better results under the same network conditions (Fig. 7).

6 Conclusion

In order to detect small targets on the sea surface, this work proposes a ResNet-CWT detector that achieves a high level of target detection probability while maintaining a low false alarm rate. After training via a real database, the detector is tested under dif-

ferent controlled false alarm rates. When the false alarm rate is restricted to 10^{-3}, the P_d of our ResNet-CWT detector is 0.917 using a 2048-point observation window, as well as 0.996 by a 4096-point observation window, which exceeds the performances of LeNet-5, AlexNet and ResNet using STFT feature. Besides, CWT performs better than STFT as the observation window length increases. when the observation window length is set to 4096, the input extracted from CWT is 37.2% smaller than that of STFT, while the false alarm rate is reduced to half. In this work, we select the points from CWT to establish the input of networks, where more complex methods might be taken into consideration in the future.

References

1. Chen X, Guan J, Huang Y, Xue Y, Liu N (2019) Radar signal processing for low-observable marine target-challenges and solutions. In: 2019 IEEE international conference on signal, information and data processing (ICSIDP), pp 1–6
2. Zhao W, Jin M, Liu W (2018) A modified matrix CFAR detector based on maximum eigenvalue for target detection in the sea clutter. In: 2018 IEEE radar conference (RadarConf18), pp 0896–0901
3. Wan H, Tian X, Liang J, Shen X (2022) Sequence-feature detection of small targets in sea clutter based on Bi-LSTM. IEEE Trans Geosci Remote Sens 60:1–11
4. Lopez-Risueno G, Grajal J, Diaz-Oliver R (2003) Target detection in sea clutter using convolutional neural networks. In: Proceedings of the 2003 IEEE radar conference (Cat. No. 03CH37474), pp 321–328
5. Wen L, Zhong C, Huang X, Ding J (2019) Sea clutter suppression based on selective reconstruction of features. In: 2019 6th Asia-Pacific conference on synthetic aperture radar (APSAR), pp 1–6
6. Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In: International conference on engineering and technology (ICET), pp 1–6
7. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778
8. Liu N, Xu Y, Ding H, Xue Y, Guan J (2019) High-dimensional feature extraction of sea clutter and target signal for intelligent maritime monitoring network. Comput Commun 147:76–84
9. Qu Q, Wang Y-L, Liu W, Li B (2022) A false alarm controllable detection method based on CNN for sea-surface small targets. IEEE Geosci Remote Sens Lett 19:1–5
10. Chen X, Su N, Huang Y, Guan J (2021) False-alarm-controllable radar detection for marine target based on multi features fusion via CNNs. IEEE Sens J 21(7):9099–9111
11. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
12. Lecun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
13. Rhif M, Ben Abbes A, Farah IR, Martínez B, Sang Y (2019) Wavelet transform application for/in non-stationary time-series analysis: a review. Appl Sci 9(7):1345
14. Lilly JM, Olhede SC (2012) Generalized Morse wavelets as a superfamily of analytic wavelets. IEEE Trans Signal Process 60(11):6036–6041

Interference Exploitation Beamforming for Cognitive Radio Network with Energy Transfer

Yongjian Xiao, Jiaqing Song, Tong Liu, and Xuewen Liao

Abstract Instead of fully mitigating the interference, we design a data-aided beamforming scheme that exploits the multiuser interference to further improve the quality of the received signal at each secondary user, termed constructive interference (CI) scheme. A power minimization problem is formulated by jointly optimizing the beamforming design and the power splitting (PS) ratio, subject to the constraints of the received signal-to-interference-plus-noise ratio (SINR), the harvested energy requirement of the secondary users and interference power of primary users. To tackle this non-convex problem, we provide an approximate second-order cone programming (SOCP) relaxation method for the optimization problem, and derive the robust beamformer for CI exploitation in the presence of imperfect CSI. Simulation results show that the proposed CI scheme can reduce the transmission power significantly compared with the conventional interference suppression transmit beamforming designs.

Keywords Power minimization problem · Data-aided beamforming · Constructive interference · Approximate second-order cone programming

1 Introduction

The increasing traffic demand and number of user equipments (UEs) in the wireless environment have led to a dramatic increase in the power consumption of wireless communication networks [1]. Towards this direction, energy harvesting (EH) techniques have been proposed to prolong the battery life of UEs. Since RF signals can carry both information and energy, simultaneous wireless information and power transfer (SWIPT) technology has been developed as a promising technique to meet service requirements between EH and information transmission [2]. There are two

Y. Xiao (✉) · J. Song
China Electronics Technology Group 54th Research Institute, Shijiazhuang 050011, China
e-mail: haierdy@163.com

J. Song · T. Liu · X. Liao
School of Electronics and Information Engineering, Xian Jiaotong University, Xian 710049, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_17

receiving strategies commonly used in SWIPT system, i.e., time switching (TS) and power splitting (PS), where it is shown in [3] that the PS structure can achieve a better performance than TS in terms of rate-energy tradeoff, and has therefore received extensive research attention.

Recently, a new idea that exploits the interference power instead of minimizing it has been proposed in the downlink transmission [4]. This concept termed constructive interference (CI) observes interference from a per symbol basis, and tries to manipulate the instantaneous interference such that the interference align or overlap constructively at each user side, which leads to an improved detection performance. Compared with conventional downlink beamforming designs, where multiuser interference is suppressed or eliminated at the receiver, exploiting CI can effectively reduce the transmission power [5, 6]. The existing works on CI mainly focus on the beamforming design in the downlink transmission of a multiuser MISO system, where the closed-form precoding scheme of maximizing the distance between the CI optimization region and the detection thresholds for PSK modulations is proposed in [7]. CI-based beamforming design is also combined with the SWIPT technique to further reduce the required transmit power and thus improve the energy efficiency, as studied in [8, 9]. Moreover, the authors in [10] have considered the beamforming design that exploits CI in the secondary link of a MISO underlay CR network to improve the QoS of the secondary users, under transmission power and interference temperature constraints of the primary network.

The remainder of this paper is organized as follows. In Sect. 2, we introduce our system model. Section 3 consider a robust beamforming scheme for imperfect CSI. In Sect. 4, the corresponding simulation results are provided. Section 5 concludes the paper.

2 System Model

Consider a MU-MISO downlink CR network in Fig. 1, where primary links and secondary links operate over the same frequency band. There is one secondary base station with N transmit antennas, K single-antenna SUs with PS SWIPT architecture, denoted as $\mathrm{SU}_1, \ldots, \mathrm{SU}_K$, and L single-antenna PUs, denoted as $\mathrm{PU}_1, \ldots, \mathrm{PU}_L$. We focus on the transmission of SBS to SUs, which will cause interference to the PUs. In the following analysis, $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$ and $\mathbf{g}_l \in \mathbb{C}^{N\times 1}$ denote the channel vectors from the SBS to SU_k and PU_l. In this paper, we investigate the beamforming designs when both perfect CSI and only imperfect CSI are known at the SBS.

Accordingly, the received signal at SU_k from the SBS can be expressed as

$$y_k = \mathbf{h}_k^T \sum_{i=1}^{K} \mathbf{v}_i s_i + n_k, \tag{1}$$

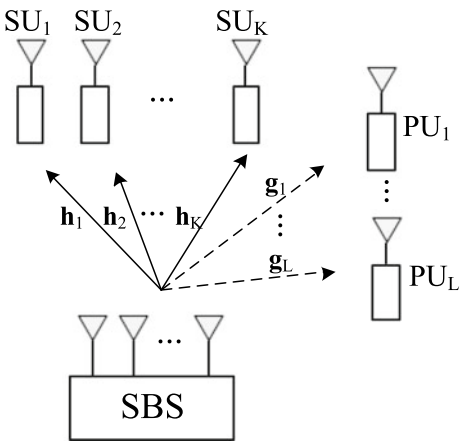


Fig. 1 System model

where s_i and $\mathbf{v}_i \in \mathbb{C}^{N\times 1}$ are the data symbol and the beamforming vector for SU_k. $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ denotes the complex Gaussian noise at the SU_k, where we assume $\sigma_k^2 = N_0, \forall k$.

Since each SU is equipped with a PS receiver, the signal received at SU_k is split into two part. The information decoding part is given by

$$y_k^{ID} = \sqrt{\rho_k} y_k + n_{wk} = \sqrt{\rho_k}\mathbf{h}_k^T \sum_{i=1}^{K} \mathbf{v}_i s_i + \sqrt{\rho_k} n_k + n_{wk}, \tag{2}$$

where $\rho_k \in (0, 1)$ is the PS ratio, $n_{wk} \sim \mathcal{CN}(0, \sigma_{wk}^2)$ is the additional processing noise generated by the SU_k. It is assumed that n_{wk} is independent of n_k and σ_{wk}^2 is identical at each SU_k. Besides, the residual signal at SU_k for EH is

$$y_k^{EH} = \sqrt{1-\rho_k} y_k = \sqrt{1-\rho_k} \left(\mathbf{h}_k^T \sum_{i=1}^{K} \mathbf{v}_i s_i + n_k \right), \tag{3}$$

and the harvested energy of SU_k can be written as

$$P_k = \eta(1-\rho_k) \left(\left| \sum_{i=1}^{K} \mathbf{h}_k^T \mathbf{v}_i \right|^2 + \sigma_k^2 \right), \tag{4}$$

where $\eta \in (0, 1)$ is the energy conversion efficiency. Based on the information decoding part in (2), the received SINR of SU_k is written as

$$\mathrm{SINR}_k = \frac{\left|\mathbf{h}_k^T \mathbf{v}_k\right|^2}{\sum_{j \neq k} \left|\mathbf{h}_k^T \mathbf{v}_j\right|^2 + \sigma_k^2 + \sigma_{wk}^2/\rho_k}. \tag{5}$$

Moreover, the interference power received at PU_l from the SBS can be expressed as

$$I_l = \left| \sum_{i=1}^{K} \mathbf{g}_l^T \mathbf{v}_i \right|^2. \tag{6}$$

3 Robust Beamforming with Constructive Interference

In this section, we consider the transmission power minimization problem when the CSI is imperfectly known by the SBS. We assume that PS ratios ρ of each SU_k are fixed for simplicity. For robust beamforming design based on CI, the channel between SBS and SUs, PUs are modeled as

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \mathbf{e}_k, \ \forall k; \mathbf{g}_l = \hat{\mathbf{g}}_l + \mathbf{e}_l, \ \forall l, \tag{7}$$

where $\hat{\mathbf{h}}_k$ and $\hat{\mathbf{g}}_l$ denote the estimated CSI known at the SBS, respectively. $\mathbf{e}_k$ and $\mathbf{e}_l$ are the CSI uncertainty with $\|\mathbf{e}_k\| \leq \varepsilon_k$ and $\|\mathbf{e}_l\| \leq \varepsilon_l$, respectively.

We separate the real and imaginary parts of each complex vector, defining $\bar{\mathbf{h}}_k = [\hat{\mathbf{h}}_{R,k}; \hat{\mathbf{h}}_{I,k}]$ and $\bar{\mathbf{e}}_k = [\mathbf{e}_{R,k}; \mathbf{e}_{I,k}]$. As such, $\mathrm{Im}\left(\tilde{\mathbf{h}}_k^T \mathbf{w}\right)$ and $\mathrm{Re}\left(\tilde{\mathbf{h}}_k^T \mathbf{w}\right)$ can be expressed as

$$\begin{aligned} \mathrm{Im}\left(\tilde{\mathbf{h}}_k^T \mathbf{w}\right) &= \hat{\mathbf{h}}_{R,k}^T \mathbf{w}_I + \hat{\mathbf{h}}_{I,k}^T \mathbf{w}_R + \mathbf{e}_{R,k}^T \mathbf{w}_I + \mathbf{e}_{I,k}^T \mathbf{w}_R \\ &= \bar{\mathbf{h}}_k^T \mathbf{w}_1 + \bar{\mathbf{e}}_k^T \mathbf{w}_1, \end{aligned} \tag{8}$$

$$\mathrm{Re}\left(\tilde{\mathbf{h}}_k^T \mathbf{w}\right) = \bar{\mathbf{h}}_k^T \mathbf{w}_2 + \bar{\mathbf{e}}_k^T \mathbf{w}_2, \tag{9}$$

where $\mathbf{w}_1 = [\mathbf{w}_I; \mathbf{w}_R]$ and $\mathbf{w}_2 = [\mathbf{w}_R; -\mathbf{w}_I]$.

Thus, the robust formulations of SINR constraint (16b) can be rewritten as

$$\bar{\mathbf{h}}_k^T \mathbf{w}_1 - \bar{\mathbf{h}}_k^T \mathbf{w}_2 \tan\theta + \varepsilon_k \left\|\mathbf{w}_1 - \mathbf{w}_2 \tan\theta\right\| + \tilde{\gamma}_k \tan\theta \leq 0, \tag{10}$$

$$-\bar{\mathbf{h}}_k^T \mathbf{w}_1 - \bar{\mathbf{h}}_k^T \mathbf{w}_2 \tan\theta + \varepsilon_k \left\|\mathbf{w}_1 + \mathbf{w}_2 \tan\theta\right\| + \tilde{\gamma}_k \tan\theta \leq 0. \tag{11}$$

For the EH constraint (16c), the robust expression can be given by considering $\left\|\tilde{\mathbf{h}}_k^T \mathbf{w}\right\|^2 = \mathrm{Re}\left(\tilde{\mathbf{h}}_k^T \mathbf{w}\right)^2 + \mathrm{Im}\left(\tilde{\mathbf{h}}_k^T \mathbf{w}\right)^2$. Then, we use Schur complement theorem and S-procedure [11] to transform it into

$$\begin{bmatrix} \lambda_k \mathbf{I} + \mathbf{W}_1 + \mathbf{W}_2 (\mathbf{W}_1 + \mathbf{W}_2)\, \bar{\mathbf{h}}_k \\ \bar{\mathbf{h}}_k^T\, (\mathbf{W}_1 + \mathbf{W}_2) \qquad u_k \end{bmatrix} \succeq 0, \tag{12}$$

where $\mathbf{W}_1 = \mathbf{w}_1 \mathbf{w}_1^T$ and $\mathbf{W}_2 = \mathbf{w}_2 \mathbf{w}_2^T$. λ_k is the slack variables and $u_k = \bar{\mathbf{h}}_k^T\, (\mathbf{W}_1 + \mathbf{W}_2)\, \bar{\mathbf{h}}_k - E_k/\eta\, (1 - \rho_k) - \lambda_k \varepsilon_k^2$.

Following a similar approach, we further transform the interference power constraint (16d) into linear matrix inequality (LMI). Accordingly, the robust beamforming with CI transmission power minimization problem can be formulated as following

$$\min_{\mathbf{w}_1, \mathbf{w}_2, \mathbf{W}_1, \mathbf{W}_2, \lambda_k, \mu_l} \|\mathbf{w}_1\|^2 \tag{13a}$$

$$s.t.\ (30), (31), (32), \forall k, \tag{13b}$$

$$\begin{bmatrix} \mu_l \mathbf{I} + \mathbf{W}_1 + \mathbf{W}_2\ (\mathbf{W}_1 + \mathbf{W}_2)\, \bar{\mathbf{g}}_l \\ \bar{\mathbf{g}}_l^T\, (\mathbf{W}_1 + \mathbf{W}_2) \qquad \nu_l \end{bmatrix} \preceq 0, \forall l, \tag{13c}$$

$$\begin{bmatrix} \mathbf{W}_1 & \mathbf{w}_1 \\ \mathbf{w}_1^T & 1 \end{bmatrix} \succeq 0, \begin{bmatrix} \mathbf{W}_2 & \mathbf{w}_2 \\ \mathbf{w}_2^T & 1 \end{bmatrix} \succeq 0, \tag{13d}$$

$$\text{rank}\,(\mathbf{W}_1) = 1, \text{rank}\,(\mathbf{W}_2) = 1. \tag{13e}$$

where $\bar{\mathbf{g}}_l = [\hat{\mathbf{g}}_{R,l}; \hat{\mathbf{g}}_{I,l}]$ and μ_l are the slack variables, $\nu_l = \bar{\mathbf{g}}_l^T\, (\mathbf{W}_1 + \mathbf{W}_2)\, \bar{\mathbf{g}}_l - \beta_l - \mu_l \varepsilon_l^2$. The problem (33) can be readily solved by the SDR approach.

4 Simulation Results

In this section, we present the simulation results to illustrate the effect of SINR threshold and EH threshold on the total transmission power of SBS, respectively. The CI-based approach is compared with the traditional downlink beamforming design through Monte-Carlo simulations, where $N = K = 3$, $L = 2$, and QPSK modulation is adopted. In addition, the effects of different constraint thresholds on the transmission power are simulated. It is assumed that all SUs an PUs have the same constraint threshold parameters, $\gamma_k = \gamma$, $E_k = E$, $\beta_l = \beta$, $\sigma_k^2 = N_0 = -55$ dBm and $\sigma_{wk}^2 = N_w = -65$ dBm in all simulations. In order to compare with the traditional scheme under the same condition, considering the distance between the users and the base station is 5 m, the path loss coefficient is 2.7. The energy harvesting efficiency of all SUs is $\eta = 0.7$.

In Fig. 2, we compare the performance of algorithms SOCP UB, SOCP LB, and conventional beamforming based on (7), (23) and (26), and study the transmit power of SBS versus different SINR threshold γ. EH and the interference power thresholds are fixed at $E = -30$ dBm and $\beta = -15$ dBm, respectively. Compared with the conventional beamforming design, the CI design needs less transmission power. When the SINR threshold $\gamma = 15$ dB, beamforming based on CI can reduce the transmission power by about 8 dB. Importantly, we observe that the upper bound

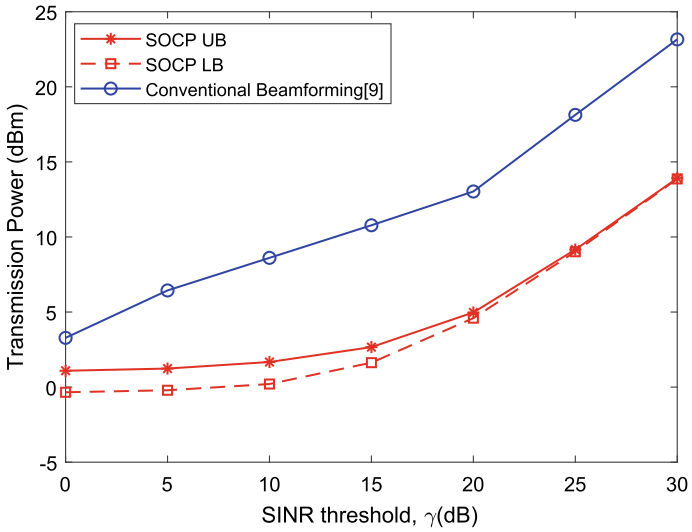


Fig. 2 Transmission power versus SINR threshold, γ

and the lower bound of the transmit power converge with the increase of the SINR threshold.

Figure 3 shows the transmit power achieved with varying EH threshold E under the fix SINR threshold $\gamma = 10\,\text{dB}$. It can be observed that the total transmission power increases with the increasing of the EH threshold. Compared with conventional beamforming scheme, the CI beamforming design can reduce the transmission power by 6 dB approximately.

Figure 4 shows the effect of the threshold of different constraint parameters on the transmission power. Simulation results show that under certain constraint threshold, the SINR or interference power constraint has little influence on the transmission power, which numerically validates the effectiveness of our proposed approximation method. To be specific, when the EH threshold E is small, the transmission power is greatly affected by SINR threshold γ. When $E \geq -10\,\text{dBm}$, the influence of the γ on the transmission power is weakened. Hence, the SINR constraint can be ignored, and we only need to consider EH and interference power constraints. Moreover, when γ varies between $0 \sim 15\,\text{dB}$ and E is around $-30\,\text{dBm}$, interference power threshold β has less effect on the transmission power, so the interference power constraint can be ignored. While when γ and E increase, the interference power constraint cannot be ignored since the transmission power is increasing.

Figure 5 shows the change of transmission power with the increasing channel estimation error ε_k and ε_l is fixed at 10^{-4}. The thresholds of constraints are fixed at $\gamma = 30\,\text{dB}$, $E = -30\text{dBm}$ and $\beta = -15\,\text{dBm}$. Compared with the CI scheme with perfect CSI, the transmission power of robust beamforming with imperfect CSI is increasing obviously while ε_k increasing. And when $\rho = 0.1$, the robust

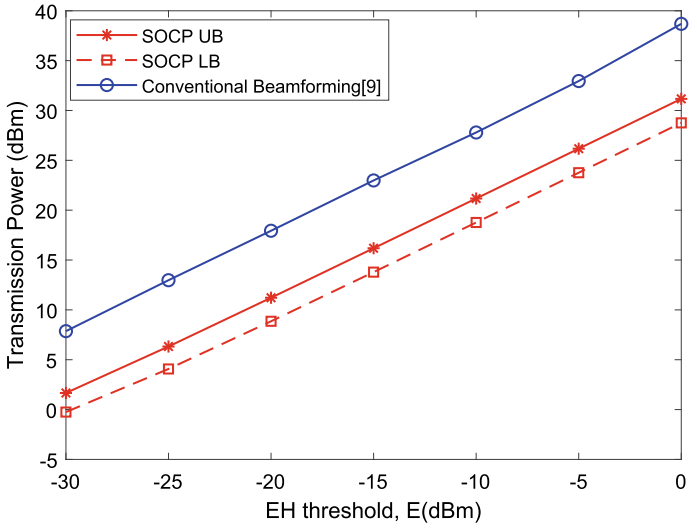


Fig. 3 Transmission power versus EH threshold, E

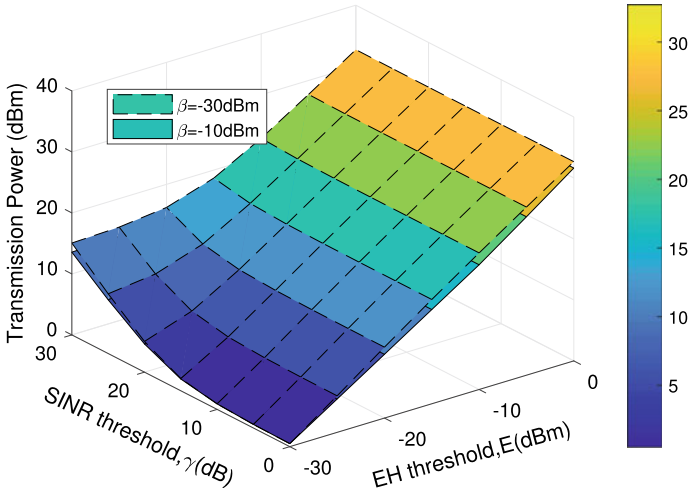


Fig. 4 Transmission power versus constraint thresholds

beamforming scheme needs more transmission power to resist greater channel estimation uncertainty and ensure the qualified QoS. Furthermore, the lower transmission power can be achieved in the optimized ρ compared to the fixed PS ratio.

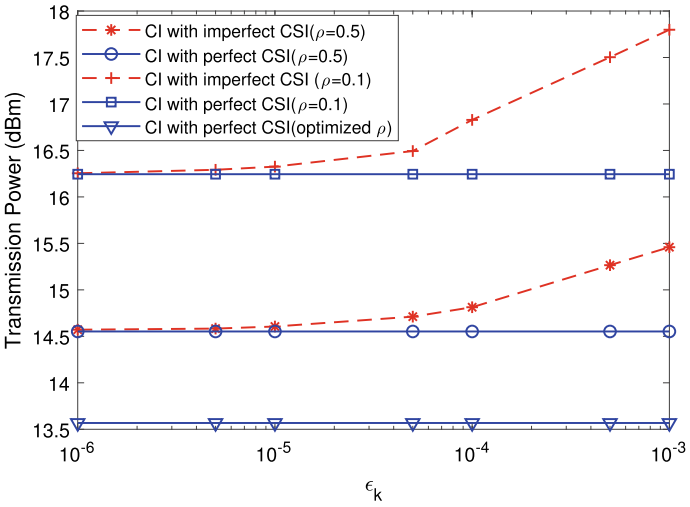


Fig. 5 Transmission power versus channel estimation error bound

5 Conclusion

This paper investigated the joint beamforming design based on CI and PS ratios optimization for a MU-MISO CR network with SWIPT. The formulated optimization problem aimed to minimize the transmit power at the SBS subject to the constraints of the interference power of each PU, SINR and energy harvesting of SUs. The problem is shown to be non-convex, for which we managed to derive its upper bound and lower bound by transforming the problem into SOCP. The robust CI beamforming scheme was proposed in the case of imperfect CSI. Simulation results showed that compared with the conventional beamforming, the proposed method can effectively reduce the power consumption at SBS.

References

1. Ku M, Li W, Chen Y, Liu KJR (2016) Advances in energy harvesting communications: past, present, and future challenges. IEEE Commun Surv Tutor 18(2):1384–1412
2. Ponnimbaduge Perera TD, Jayakody DNK, Sharma SK, Chatzinotas S, Li J (2018) Simultaneous wireless information and power transfer (SWIPT): recent advances and future challenges. IEEE Commun Surv Tutor 20(1):264–302
3. Zhou X, Zhang R, Ho CK (2013) Wireless information and power transfer: architecture design and rate-energy tradeoff. IEEE Trans Commun 61(11):4754–4767
4. Mohjazi L, Ahmed I, Muhaidat S, Dianati M, Al-Qutayri M (2017) Downlink beamforming for SWIPT multi-user MISO underlay cognitive radio networks. IEEE Commun Lett 21(2):434–437

5. Alodeh M, Chatzinotas S, Ottersten B (2015) Constructive multiuser interference in symbol level precoding for the MISO downlink channel. IEEE Trans Signal Process 63(9):2239–2252
6. Masouros C, Zheng G (2015) Exploiting known interference as green signal power for downlink beamforming optimization. IEEE Trans Signal Process 63(14):3628–3640
7. Li A, Masouros C (2018) Interference exploitation precoding made practical: optimal closed-form solutions for PSK modulations. IEEE Trans Wireless Commun 17(11):7661–7676
8. Li A, Masouros C (2018) Energy-efficient SWIPT: from fully digital to hybrid analog–digital beamforming. IEEE Trans Veh Technol 67(4):3390–3405
9. Timotheou S, Zheng G, Masouros C, Krikidis I (2016) Exploiting constructive interference for simultaneous wireless information and power transfer in multiuser downlink systems. IEEE J Sel Areas Commun 34(5):1772–1784
10. Law KL, Masouros C, Pesavento M (2017) Transmit precoding for interference exploitation in the underlay cognitive radio Z-channel. IEEE Trans Signal Process 65(14):3617–3631
11. Boyd S, Vandenberghe L (2004) Convex optimization. Cambridge University Press, Cambridge

Kinematic Analysis and Programming Implementation of Patrol Robots Based on Ackermann's Principle

Renjun Wang, Lu Li, Xiaojiang Li, and Qunfeng Zhao

Abstract The kinematic analysis of the patrol robot is based on Ackermann theory to derive the relationship between front wheel deflection angle, rear wheel speed and robot target speed. The fitted curve is obtained by plotting a scatter plot of the relationship between the deflection angle of the robot's rudder and the deflection angle of the front wheel and fitting the curve. Ultimately the whole robot is programmed for motion control. This paper provides a basis for theoretical analysis for further research on the Ackermann principle patrol robot.

Keywords Ackermann's Theory · Patrol robot · Programming Control

1 Introduction

The Ackermann steering robot is in fact the steering structure of a modern car, the Ackermann type is similar to the two-wheel differential [1–3]. It also relies on the differential speed of the drive wheels to make the turn. However, it is also necessary to control the angle of the front wheels to match the turn. Otherwise the friction between the front wheels and the ground would be very high and would seriously affect the steering movement of the robot and wear out the wheels [4–6]. As shown in Fig. 1, the Ackermann robot steering state requires the perpendiculars of the 4 wheels to intersect at a point, i.e. the 4 wheels rotate around the same centre.

2 Kinematic Analysis

We first discuss the kinematic inverse solution as shown in Fig. 2, i.e. the target velocities of the drive wheels are derived from the target velocities of the V_X, V_Y and V_Z axes. The driving wheels are the two rear wheels, whose velocities are denoted by

R. Wang · L. Li (✉) · X. Li · Q. Zhao
Zhonghuan Information College Tianjin University of Technology, Tianjin, China
e-mail: lilu569@126.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_18

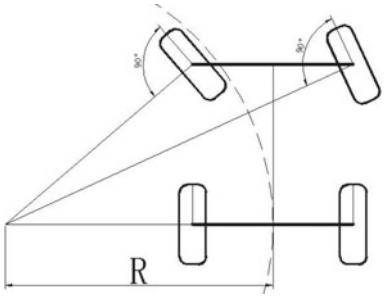


Fig. 1 Typical Ackermann steering state

V_L and V_R respectively, with forward being positive and backward being negative [7, 8]. The Ackermann robot also requires the deflection angles of the two front wheels, $Angle_L$ for the left front wheel and $Angle_R$ for the right front wheel, with the left deflection being positive and the right deflection being negative [9]. The Ackermann robot cannot move directly laterally, so the Y-axis target velocity is not discussed.

For Ackermann-type robots, the kinematic inverse solution is required and the relationship between V_X, V_Z and the turning radius R must first be determined (here the centre of the rear axis is the centre of rotation of the robot) [10, 11]. Assume that the robot moves for t seconds at velocities V_X, V_Z. The length of arc $O_1O_2$is the integral of velocity V_X over time t: $O_1O_2 = V_X * t$. The robot's attitude is rotated by θ, which is the integral of the velocity V_Z over time t: $\theta = V_Z * t$. It is easy to see from the properties of a circle that the radius of the turn it is shown in Eq. (1).

$$R = \frac{O_1O_2}{\theta} = \frac{V_X * t}{V_Z * t} = \frac{V_X}{V_Z} \tag{1}$$

Now that V_X, V_Z and the turning radius $R = \frac{V_X}{V_Z}$. Next, find the velocities V_L and V_R of the left and right rear wheels and the left and right front wheel deflection angles $Angle_L$ and $Angle_R$. The other parameters that need to be known are the wheelbase W and the wheelbase H, as shown in Fig. 3.

Knowing that the robot as a whole is rotating around point O during steering, it is easy to know the velocities V_L, V_R of the left and right rear wheels. It is is shown in Eqs. (2) and (3).

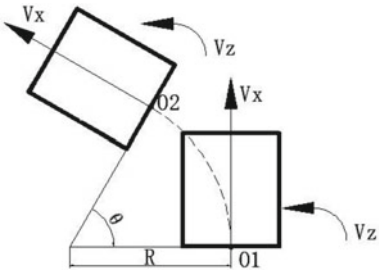


Fig. 2 Target speed and turning radius

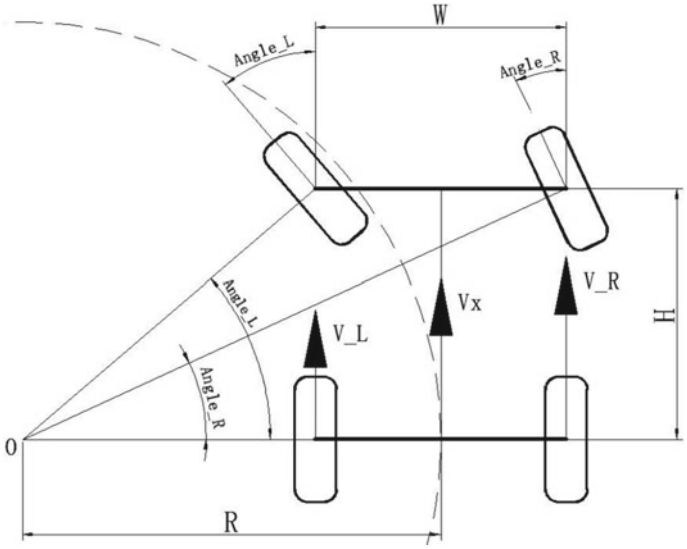


Fig. 3 Target speed versus wheel speed, deflection angle

$$V_L = V_X \frac{R-0.5W}{R} \tag{2}$$

$$V_R = V_X \frac{R+0.5W}{R} \tag{3}$$

From the geometric relationship, the left and right front wheel deflection angles AngleL, AngleR are shown in Eqs. (4) and (5).

$$Angle_L = \tan^{-1}\left(\frac{H}{R-0.5W}\right) \tag{4}$$

$$Angle_R = \tan^{-1}\left(\frac{H}{R+0.5W}\right) \tag{5}$$

Then Eqs. (1), (4) and (5) together form the kinematic inverse solution. Combining Eqs. (3) and (4) yields Vx, as shown in Eq. (6).

$$V_X = \frac{V_L+V_R}{2} \tag{6}$$

Combining Eqs. (1), (3), (4) and (6) yields Vz, as shown in Eq. (7).

$$V_Z = \frac{V_R-V_L}{2} \tag{7}$$

At this point we have the kinematic positive solution.

3 Controlling the Steering Angle of the Front Wheels

The Ackermann steering structure based on a crank rocker mechanism is shown in Fig. 4. The steering wheel is driven by the drive servo, which in turn controls the steering of the two front wheels by mechanical transmission [12]. The difference in deflection angle between the two front wheels is determined by the mechanical design, which requires that the plumb line of the two front wheels intersects the rear wheel axis at one point [13].

From Eqs. (1), (6) and (7) we know that for each set of determined target speeds V_X and V_Z, there is a set of determined $Angle_L$ and $Angle_R$. So we only need to know the relationship between the deflection angle of the rudder disc and the corresponding deflection angle of one of the front wheels. This is because the deflection angle of the rudder disc corresponds to the deflection angle of the front wheel. We use a curve fitting model to approximate the relationship between rudder deflection and front wheel deflection. Figure 5 shows a scatter plot of the relationship between rudder deflection angle and front wheel deflection angle for the Ackermann patrol robot.

It can be seen that the curve can be fitted with either a primary curve or a secondary curve. For simplicity, MATLAB's Curve Fitting Toolbox can be invoked to view the fit of both curves [14], and the results of the primary and secondary curve fits are shown in Figs. 6 and 7.

Where SSE: Sum of the Squared Error, the smaller the sum of the squares, the better the curve fits the observations. R2: Goodness of Fit, the closer the Goodness of Fit is to 1 the better the curve fits the observations. RMSE: Root Mean Squared Error, the smaller the RMSE, the better the curve fits the observations. Therefore, the quadratic curve fits the observed values slightly better than the primary curve, and the quadratic curve is chosen as the fitted curve here [15].

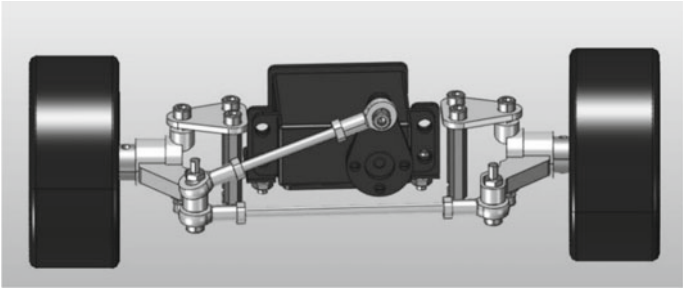


Fig. 4 Ackermann steering structure based on a crank rocker mechanism

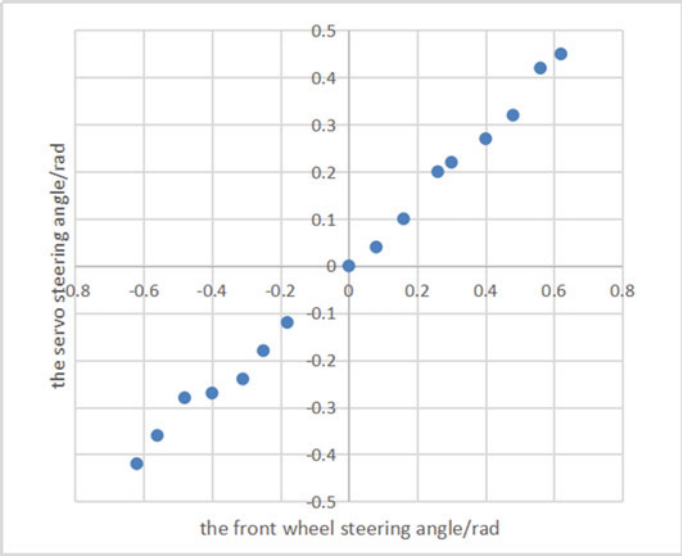


Fig. 5 Scatter plot of rudder deflection versus front wheel deflection

Fig. 6 Fitting results for primary curves

Linear model Polyl:
f(x)=p1*x+p2
Coefficients(with 95% confidence bounds):
p1= 1.432(1.374,1.49)
p2= -0.005732(-0.02132,0.009859)
Goodness of fit:
SSE:0.01364
R-square:0.9946
Adjusted R-square:0.9942
RMSE:0.03016

4 The Programming of the Control

The following is the control of the entire robot movement in C. There are two main steps. The first step is to find the Ackermann turning radius from the X-axis target velocity and the Z-axis target velocity. It is shown in Fig. 8.

The second step is to find the target velocity of the left and right motors and to control the servo rotation from the target velocity of the X-axis and the target angle of rotation of the right front wheel. The function input parameter V_Z represents the target angle of rotation of the right front wheel. It is shown in Fig. 9.

Linear model Poly2:
f(x)=p1*x^2+p2*x+p3
Coefficients(with95%confidence bounds):
p1= -0.2137(-0.4413,0.01388)
p2= 1.439(1.385,1.493)
p3= 0.009599(-0.0121,0.0313)
Goodness of fit:
SSE:0.01058
R-square:0.9958
Adjusted R-square:0.9952
RMSE:0.02749

Fig. 7 Fitting results for secondary curves

```
float Vz_to_Akm_Angle(float Vx, float Vz)
{    float R, AngleR, Min_Turn_Radius;
     Min_Turn_Radius=MINI_AKM_MIN_TURN_RADIUS;
     if(Vz!=0 && Vx!=0)
     {     //If it is less than the minimum turning radius
           if(float_abs(Vx/Vz)<=Min_Turn_Radius)
           {
                //Reducing the angular velocity of
                the target
                     Vz= Vx/(Min_Turn_Radius);
                else
                     Vz=-Vx/(Min_Turn_Radius);
           }
           R=Vx/Vz;
           //AngleL=atan(Axle_spacing/(R-0.5*Wheel_spacing));
           AngleR=atan(Axle_spacing/(R+0.5*Wheel_spacing));
     }
     else
     { AngleR=0;}
     return AngleR;
}
```

Fig. 8 First step code

```
float Vz_to_Akm_Angle(float Vx, float Vz)
{    float R, AngleR, Min_Turn_Radius;
     Min_Turn_Radius=MINI_AKM_MIN_TURN_RADIUS;
     if(Vz!=0 && Vx!=0)
     {     //If it is less than the minimum turning radius
           if(float_abs(Vx/Vz)<=Min_Turn_Radius)
           {
                //Reducing the angular velocity of
                the target
                     Vz= Vx/(Min_Turn_Radius);
                else
                     Vz=-Vx/(Min_Turn_Radius);
           }
           R=Vx/Vz;
           //AngleL=atan(Axle_spacing/(R-0.5*Wheel_spacing));
           AngleR=atan(Axle_spacing/(R+0.5*Wheel_spacing));
     }
     else
     { AngleR=0;}
     return AngleR;
}
```

Fig. 9 Second code step

5 Conclusion

This paper presents a kinematic analysis of the Ackermann principle for patrol robots. The inverse and positive solutions of the kinematic relationship are derived by analysing the relationship between the front wheel deflection angle, the rear wheel velocity and the robot target velocity. The control of the front wheel deflection angle is investigated by plotting a scatter plot of the relationship between the robot's rudder deflection angle and the front wheel deflection angle and fitting the curve. The final control of the entire robot's motion is achieved by programming.

Acknowledgements This work was supported by the Scientific research project of the Tianjin Education Commission: No. 2022KJ039.

References

1. Liao L, Wang J, Xie M et al (2010) Optimization of automotive steering trapezoidal mechanism under objective functions. Mech Electr 02:3–6

2. Qiang L, Yang C, Zhang W et al (2023) Research on 4WID/4WIS vehicle tracking control based on Ackermann's principle. J Dyn Control 21(01):36–44
3. Ma K (2016) A preliminary investigation of the steering angle relationship of wheeled vehicle deflector. Mech Prod Develop Innov 29(01):37–39
4. Xu L (2017) Analysis of steering trapezoid design based on Ackermann theory. Mech Eng 2017(05):69–70
5. Lan J, Gao X (2018) Design and simulation of a container loading robot based on Ackermann's principle. Mechatronics 24(Z1):32–38
6. Li C, Fu Y, Zhang J, Wan F et al (2018) Study of modified Ackermann's principle turning angle when considering lateral deflection characteristics. Missile Space Launch Tech 06:96–99
7. Zhou R, Yang Q, Wang Y et al (2017) Research and application of cooperative control algorithm of rudder and differential motor based on Ackermann's principle. J Chifeng College (Nat Sci Edn) 33(14):14–15
8. Li Q (2019) Teaching and extending the Ackerman steering principle. Curriculum Educ Res 2019(32):189
9. Zhang Z (2016) Intelligent wheelchair dynamics analysis and steering trapezoid design based on Ackermann's principle. Manuf Autom 38(12):78–80+84
10. Zhang X (2016) Design of steering mechanism of 16-wheel crane based on Ackermann principle. Port Handling (01):10–13
11. Lu H, Zhang C, Guo Y et al (2020) (2023) Optimal control of all-wheel steering for three-axis vehicles based on Ackermann's principle. J Guangxi Univ (Nat Sci Edn) 45(03):558–568
12. Gao X (2015) A case of analyzing and verifying steering trapezoid by using MATLAB software. Hebei Agric Mach 2015(11):60
13. Han Y, Shi Z, Chen L et al (2022) Study on tire bias grinding of dual front axle vehicle based on steering system arrangement optimization. Heavy Duty Vehicle (06):15–16+19
14. Ren X, Cai Z (2009) Kinematic modeling of a vehicle-based mobile robot based on Ackerman's principle. J Intell Syst 4(06):534–537
15. Cui Y, Liu X, Ma C et al (2008) Design analysis of steering mechanism for multi-wheeled heavy vehicles. Mech Sci Technol 08:1052–1055

Task Allocation Algorithm for Multiple UAVs in IoT Networks

Liang Ye, Yu Yang, Rangang Zhu, and Xiaoshuai Li

Abstract As UAV and sensor techniques develop, UAVs are becoming more and more popular in both military and civilian fields. In post-earthquake evaluation scenes, UAVs and sensors can form an IoT network and collect environmental information. By allocating tasks reasonably and planning drone flight paths, data collection tasks within the entire earthquake area can be efficiently and stably completed. This paper proposes an improved task allocation algorithm based on the immune algorithm. The proposed algorithm can allocate tasks reasonably, and thus shorten the flight distance while avoiding meteorological threat areas.

Keywords Improved immune algorithm · Task allocation · UAV · IoT network

1 Introduction

Unmanned aerial vehicles (UAVs) [1] and other unmanned flight systems have extremely strong manoeuvrability. When performing tasks in hazardous areas, they do not rely on onboard personnel. They are not only convenient and intelligent to use, but also reduce pilot training costs. Therefore, they are widely used in various application scenarios in various fields such as reconnaissance, strike, surveillance, search and rescue. With the increasing popularity of drone applications [2], research issues such as autonomous algorithms, intelligent control algorithms [3], and collaborative planning algorithms for drones are becoming increasingly prominent. Especially the ability of drones to perform tasks in dry, harsh or dangerous environments for humans has attracted more and more scientific researchers to conduct research on drone flight formation algorithms, task allocation algorithms, trajectory planning algorithms, and efficiency evaluation methods. Drones not only play an increasingly

L. Ye (✉) · Y. Yang
Communication Research Center, Harbin Institute of Technology, Harbin 150001, China
e-mail: coldwound@163.com

R. Zhu · X. Li
School of Electronic Countermeasures, National University of Defense Technology, Hefei 230031, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_19

important role in the military field, but also have a high demand in civilian field such as IoT (Internet of Things) networks. Especially in today's increasingly frequent occurrence of various disasters, drones are increasingly being used to monitor disaster scenes, such as mountain fire monitoring, maritime emergency rescue, and medical material transportation.

The main scenario considered in this paper is that after a severe earthquake, the early warning measurement center needs to collect soil, pressure, water quality and other information collected by wireless sensor nodes pre-set at certain key points in the earthquake area, such as plate boundaries, in order to analyze the disaster situation, whether a barrier lake will form, and the arrival time of aftershocks. Due to the threat of aftershocks and the impact of landforms being destroyed, it is not feasible to manually collect information using land transportation facilities. Utilizing the high mobility of drones, each early warning measurement center sends a drone cluster to collect data from a task target point. Due to the cluster distribution of wireless sensor nodes, when earthquakes occur on a large scale, it is difficult to network between each group of sensor nodes. At the same time, there is no support from communication facilities such as base stations, which makes it impossible to gather all information from various locations. That is, there are far more locations in a large range that require data collection tasks than drones. On a small scale, due to the possibility of node damage during earthquakes and limited energy, maintaining a data aggregation state in scenarios where batteries cannot be replaced during earthquakes can quickly deplete battery energy. Therefore, each group of nodes does not choose a separate data aggregation node, but instead uses drones as the aggregation nodes, maximizing energy efficiency by planning the acquisition path of unmanned aerial vehicles. By allocating tasks reasonably and planning drone flight paths, data collection tasks within the entire earthquake area can be efficiently and stably completed.

2 Task Allocation Model

The issue of multi-drone task allocation [4] can generally be described as: assigning a set of target point sequences to all drones under the condition of known drone formation groups, target point information, and partial environmental information, in order to complete all tasks with the minimum cost and achieve the optimal efficiency.

The drone clusters in this project meet the following conditions:

1. Isomorphism: All the drones have the same structure and function.
2. Communication guarantee: All the drones have established communication links through flight Ad hoc networks.
3. Functionality: All drones are equipped with relevant equipment for data collection.

Assume that the number of drones is N, and the set of drones is $U = \{U_1, U_2, \ldots, U_N\}$. The number of target points to traverse is M, and the set of tasks is $T = \{T_1, T_2, \ldots, T_M\}$. The task allocation problem meets the following principles:

1. Each drone has the same structure, which satisfies the same resource and dynamic constraints;
2. Tasks at any location can only be assigned to a single drone cluster, meaning that a single target point can only be accessed once;
3. The drones visit all target points at the shortest possible distance and return to the base, while ensuring the shortest total length of the drone formation flight route;
4. The task allocation of all drone clusters participating in the task should be as fair as possible;
5. Each drone cluster should try to avoid meteorological threat areas as much as possible.

The task allocation problem is a combinatorial optimization problem with multiple constraints, which needs to meet the following constraints:

1. Maximum flight distance constraint: The maximum flight distance of a single drone cluster is limited by the amount of fuel carried. Assume that the maximum flight distance is D_{max}, and the route of the ith drone is L_i. Then the maximum flight distance constraint is expressed as,

$$D(L_i) \leq D_{\max}, \forall i = 1, ..., N$$

2. Target traversal constraint: All the target points should be assigned only once, and this constraint can be expressed as,

$$\sum_{i=1}^{N} \sum_{j=1}^{M_i} T_i^j = T \&\& \sum_{j=1}^{M_i} T_i^j \cap \sum_{j=1}^{M_k} T_k^j = \emptyset, \forall i, j \in N, i \neq j$$

where T_i^j is the jth target point that the ith drone will visit.

In dynamic scenarios, task allocation should consider the total length of the drone cluster's execution path, the flight load of each drone cluster, and the penalty for passing through the threat zone. Considering the above factors, the fitness function is established as,

$$Fitness = \alpha \sum_{i=1}^{N} D(L_i) + \beta \max(D(L_i)) + \gamma \sum_{i=1}^{N} D_o(L_i)$$

where $\sum_{i=1}^{N} D(L_i)$ is the total length of all the drones, $\max(D(L_i))$ is the maximum flight distance, and $D_o(L_i)$ is the penalty path for drones flying through threat zones. α, β, γ are three weighting coefficients and,

$$\alpha + \beta + \gamma = 1$$

By minimizing the fitness function, an optimal solution can be obtained for the comprehensive evaluation of the total distance, the balanced degree of task allocation, and the flight safety degree. The final optimization problem can be expressed as,

$$\begin{aligned} &\min Fitness \\ &s.t.\ D(L_i) \le D_{\max}, \forall i = 1, ..., N \\ &\sum_{i=1}^{N}\sum_{j=1}^{M_i} T_i^j = T \&\& \sum_{j=1}^{M_i} T_i^j \cap \sum_{j=1}^{M_k} T_k^j = \emptyset, \forall i,\ j \in N, i \ne j \end{aligned}$$

3 Task Allocation Based on Improved Immune Algorithm

Based on the concept of artificial immune algorithm [5] and multi-agent system, an improved immune multi-agent algorithm is proposed to solve the task allocation problem. By limiting some operations within the neighborhood, the ability of traditional immune algorithms to jump out of local optima is improved, while also maintaining the fast convergence speed of traditional immune algorithms in the initial stage. It is also suitable for dynamic and static task allocation problems.

3.1 Memory Population Initialization

Immunological memory refers to the concept of memory population in Clonal selection algorithm, which is used to save the optimal part of agents in each evolution process. Due to the limited size of the memory agent group, it is necessary to update and eliminate it in real-time during each iteration. The specific steps are as follows,

1. Assuming that the size of the initial memory agent group is N, select k antibody agents with the best fitness in the current antibody agent grid, and try to add them to the memory agent group. If the memory agent group has not reached the pre-set size, this addition is allowed, and the k antibody agents are directly added to the agent group.
2. If the memory agent group is full, or the number of antibody agents that can be accommodated is less than k, then the elimination operation is performed, all antibody agents in the current memory agent group are compared with the new k antibody agents for fitness, and the antibody agents with the lowest fitness that exceed the size of the memory agent group are removed from the memory agent group.

The overall process can be expressed as,

$$N_{best} = SelectN(N_{best}, L)$$

where N_{best} is the memory agent group, $SelectN(\cdot)$ means to select the first N agent functions with the best fitness.

3.2 Neighborhood Cloning Operation

Neighborhood cloning refers to one of the artificial immune algorithms, the antibody cloning operation in the clonal selection algorithm. It is worth noting that the antibody cloning operation in the clonal selection algorithm only multiplies the best few antibodies in the population and then mutates, which is unfavorable for maintaining the diversity of the population. This paper will keep the cloning process in the neighborhood. For each antibody agent, select the agent with the best fitness in its neighborhood and its own several agents, multiply and mutate it to a certain extent, and then select the antibody agent with the best fitness to replace the original antibody agent. Neighborhood cloning operations can be represented as,

$$A_{i,j} = Select\big(Mutation\big(Clone\big(Select\big(\{A_{i,j}, Neigh.A_{i,j},\}\big)\big)\big)\big)$$

where $Select(\bullet)$ represents selecting the best agent, $Mutation(\bullet)$ represents the mutation operation, and $Clone(\bullet)$ represents the cloning operation.

3.3 Neighborhood Suppression Operation

The neighborhood suppression operator draws inspiration from the suppression concept in artificial immune algorithms, which suppresses high affinity antibodies to maintain population diversity. Because the cloning operation of this algorithm is carried out in the neighborhood, if an antibody agent is the local optimal agent, its excellent fitness will make it the parent of all antibody agent clones in its neighborhood, that is, the antibodies in a small range all preserve the characteristics of the local optimal agent, and the affinity is naturally high, which needs to be suppressed. So within each neighborhood, neighborhood suppression operations can be represented as,

$$A_{i,j} = \begin{cases} B_{i,j}, if\ A_{i,j} \in Neigh.\{A_{i,j}\} \\ A_{i,j}, otherwise \end{cases}$$

where B_{ij} is a randomly generated new antibody agent. It can be noted that due to the small neighborhood range, for each agent, if it is found that there are identical agents in the neighborhood, they will be directly replaced with randomly generated antibody agents to maintain the diversity of the agent group.

3.4 Population Crossover Operation

The objects selected for the crossover operation in this algorithm are different. For each antibody agent, a roulette wheel based selection method is used to select another parent agent in the entire grid for cross operation, and then replace the original agent with the better offspring after cross operation. This operation not only provides an opportunity for every antibody agent, regardless of its excellence, to intersect, improve the diversity of the population, conduct global search, but also has mutual motivation between excellent individuals, ensuring the direction of evolution. If the crossover operator uses a partially matched crossover operator, the population crossover operation can be represented as,

$$A_{i,j} = \begin{cases} PMX\big(A_{i,j}, Roulette(L)\big), if\ unifrnd(0,1) < (times - i)/times \\ A_{i,j}, otherwise \end{cases}$$

where *Roulette*(*L*) represents the roulette choice function for the entire grid *L*, *unifrnd* (0, 1) represents a random number that follows a uniform distribution between 0 and 1, *times* represents the overall number of iterations, and *i* indicates the current number of iterations. It can be noted that the probability of population crossover operations gradually decreases as the number of iterations increases, with the aim of accelerating population convergence and preventing the loss of too much excellent information due to excessive crossover.

3.5 Self Learning Operation

The self-learning operation refers to the self-learning ability of agent in multi-agent system. In this paper, multi-agent is the antibody agent group, and its most representative individual is the individual with the best fitness. Through cloning and mutation of the individual, selecting the optimal individual after mutation to replace the original optimal agent, and completing local search in a small range, the antibody agent group will further evolve. Self learning operations can be represented as,

$$A_{best} = Select(Mutation(Clone(A_{best})))$$

3.6 Partial 2-opt Operator

The 2-opt operator was first used in traveling salesman problems, where the goal is to find a Hamiltonian circuit with the minimum sum of weights. The traveling

salesman problem is one of the most widely studied problems in the field of combinatorial optimization. This problem is also NP hard, and without adding any additional assumptions, there is no constant approximation of polynomial time for the traveling salesman problem. Researchers are committed to designing a better approximation algorithm for the traveling salesman problem in metric space. The 2-opt optimization algorithm is one of them, which is the core of the classic Lin Kernighan algorithm. The 2-opt optimization algorithm starts from an arbitrary feasible solution and repeatedly searches. If a crossing edge is found, the encoding in the middle of this pair of crossing edges is reversed. Perform this substitution operation until the local optimal is reached, and finally use the local optimal solution as the algorithm's solution. The algorithm execution process is as follows.

Algorithm 2-opt local optimization algorithm
Input: feasible solution, distance matrix D between nodes
Output: 2-opt optimized feasible solution
For a = 1:length(Solution)-2
For b = a + 2: length(Solution)-1% Traverse the entire feasible solution
If D(a,a + 1) + D(b,b + 1) > D(a,b) + D(a + 1,b + 1) % If cross edges are found
Solution(a + 1:b) = Solution(b:− 1:a + 1)
End
End
End

In the task allocation problem in this paper, although it is not possible to directly perform 2-opt operations on the overall antibody, for the target points under each drone, partial 2-opt operations can be used to accelerate the convergence of the algorithm. The specific operations are as follows,

$$\text{Solution} = \left\{ \begin{array}{l} 2-opt(\text{Solution}\left[1:\left\lfloor \frac{N}{M} \right\rfloor\right]), \ldots, \\ \ldots, 2-opt(\text{Solution}\left[M*\left\lfloor \frac{N}{M} \right\rfloor : N\right]) \end{array} \right\}$$

where $2-opt(\bullet)$ represents performing 2-opt optimization, N is the number of target points, M is the number of drones, and $\{\bullet\}$ represents the splicing of various parts.

4 Simulation Results

The simulation parameters are given as follows: the grid parameter is 6, the size of the agent group is 36, the memory population size is 36, the iterations are 5000, and the clone multiplier is 10. The simulation scene is 50 km × 50 km × 400 m, the

Table 1 Simulation results

	Theoretical optimal solution	Simulation optimal solution	Simulation average solution
Fitness	75.0640	75.0640	76.8237

number of drones is 4, the number of targets is 30, and the truncation vector is [7, 14, 21].

Assume that the fuel of the drones is limited, so the distances that the drones fly should be as short as possible. On the other hand, in the post-earthquake areas, there are often meteorological threats such as strong wind, so the route plan of the drones should take the meteorological threats into consideration, and let the drones avoid such areas. In the simulation scene, we set a random meteorological threat area with a radius of 3 km. When there is a meteorological threat, the task allocation problem can be expressed as follows: while ensuring the shortest distance and load balance, try to stay as far away from the meteorological threat as possible. Due to the dominant factor in the threat zone, the penalty path weighting coefficient $\gamma = 1/3$. α should be larger than β, so we set $\alpha = 7/15$, and $\beta = 1/5$. The simulation results are shown in Table 1.

It can be seen that our proposed algorithm can reach the theoretical optimal solutions. Besides, the total flight distance of the clonal selection algorithm is 349.6837 km, whereas that of our proposed algorithm is 285.1336 km.

Acknowledgements This work was supported by China Chile ICT "the Belt and Road" Joint Laboratory under grant number SQ2020YFE020708 and National Natural Science Foundation of China under grant number 61971161.

References

1. Liu K, Zheng J (2022) UAV trajectory optimization for time-constrained data collection in UAV-enabled environmental monitoring systems. IEEE Internet Things J 9(23):24300–24314
2. Seno DP, Yohana SKD, Muhammad IH et al (2022) A drone technology implementation approach to conventional paddy fields application. IEEE Access 10:120650–120658
3. Jiang M (2021) Binocular stereo vision intelligent control framework for the landing position of quadrotor UAV based on fuzzy control. In: 2021 6th international conference on communication and electronics systems (ICCES), pp 1348–1352
4. Hasan SY, Hüseyin G, Hakan Ç et al (2020) Optimal task allocation for multiple UAVs. In: 2020 28th signal processing and communications applications conference (SIU)
5. Zhang Y, Liu C (2021) An immune algorithm for network data security detection. In: 2021 17th International conference on computational intelligence and security (CIS), pp 247–251

Research on Distributed Dynamic Spectrum Access Based on Deep Reinforcement Learning

Yanchao Liu, Xiaohua Zhang, and Shubin Wang

Abstract Dynamic Spectrum Access (DSA) is a critical technology for Cognitive Wireless Sensor Network (CWSN). The main challenge of DSA is how Secondary Users (SUs) can quickly and accurately identify vacant spectrum, while ensuring that the service of the Primary User (PU) is not interrupted. The current DSA solutions do not satisfy the requirements of high throughput, low interference and fast convergence simultaneously for large scale multiple users and multiple channels access scenarios. In this paper, we propose a distributed DSA algorithm based on Deep Reinforcement Learning (DRL). First, we construct a Cognitive Wireless Sensor Network (CWSN) environment with multiple users and multiple channels. Next, based on the spectrum sensing results, each SU provides channel observations to our proposed Deep Q-Network (DQN) model for training in order to learn the optimal spectrum access policy. Finally, using the output of the DQN model, each SU intelligently accesses the appropriate channel. In order to improve the training accuracy and address the performance degradation problem caused by the network depth in deep neural networks, we added the Residual Network (ResNet) structure to the DQN. Simulation results show that the proposed algorithm achieves faster convergence speed, completely avoids collisions between SUs, greatly reduces the interference of SUs to PU, and significantly improves the success rate of channel access.

Keywords Dynamic spectrum access · Deep reinforcement learning · Spectrum allocation · Cognitive wireless sensor network · Deep Q-Network

Y. Liu · S. Wang (✉)
College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China
e-mail: wangshubin@imu.edu.cn

X. Zhang
Department of Foreign Languages, Guizhou University of Commerce, Guiyang, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_20

1 Introduction

CWSN combines cognitive radio technology with Wireless Sensor Network (WSN) to address the problem of scarce spectrum resources by allowing a large number of sensor nodes as SUs to access the authorized spectrum. DSA is one of the key technologies in CWSN, and its task is to make a decision based on spectrum sensing data from cognitive sensor nodes to access a vacant spectrum licensed to a PU. However, when using this technique, the issues that need to be addressed are: how to minimise the interference to the PU while accessing and using the authorised spectrum, and how to avoid conflicts between SUs when multiple SUs try to access the same spectrum [1, 2].

Traditional optimization algorithms such as Game Theory, Particle Swarm Optimization and Genetic Algorithm have been used to address the DSA problem [3, 4]. Although these methods achieve spectrum reuse, their model design is complex, easily get trapped in local optima and less flexible and adaptive. In contrast, Reinforcement Learning (RL) can adaptively learn optimal strategies without a priori information in uncertain and dynamic complex environments. Therefore, in recent years, RL has been applied to DSA. In literature [5], a Q-learning based spectrum access algorithm is proposed to improve the transmission performance through intelligent utilisation of spectrum resources. Document [6] proposes a decentralised multi-intelligence reinforcement learning-based resource allocation scheme to address resource allocation problem without complete channel state information. The Q-learning used in the literatures [5, 6] performs well on small-scale models. However, it shows significant performance degradation when the state or action space is large. Deep Neural Network (DNN) is used in DRL to overcome this limitation. In literature [7], a centralised dynamic multichannel access framework based on DQN is proposed to minimise conflicts and optimise multi-user channel allocation through a centralised allocation policy. However, the centralised approach to spectrum access can lead to high communication overheads and may be difficult to implement in practice. In addition, the algorithm's performance may be limited as it doesn't account for imperfect spectrum sense that occur in real-world environments. Literature [8, 9] proposed using multi-intelligent deep reinforcement learning at medium access control layer for channel access. In this approach, users make transmission decisions through centralised training and decentralised execution to maximise the long-term average rates or to improve the performance of the network in terms of throughput, delay and jitter. However, this centralized training approach has single point of failure and necessitates high communication and computational resources, and decentralized execution requires transmission and synchronisation of parameters. In literature [10], a new DSA method is proposed for multichannel wireless networks that can find near-optimal policies in fewer iterations and can be applied to a wide range of communication environments. However, this method is limited as it targets at only one DSA user and does not consider the collision problem between SUs and PUs. The authors of [11, 12] employ reservoir computing or echo state networks, a type of Recurrent Neural Network (RNN), in DRL to enable distributed dynamic

spectrum access for multiple users. They mitigate the effects of spectrum sensing errors by taking advantage of the temporal correlation of RNNs, thereby reducing conflicts among users. Nonetheless, the Q-networks used are complicated and the convergence speed of the algorithm needs to be improved.

2 System Model and Problem Formulation

We consider a multi-user, multi-channel CWSN environment with N PUs and M SUs. Figure 1 depicts the intricate association of desired links and interfering links when PU_1, SU_1, and SU_2 operate on the same channel. We calculate the received signal of SU_i on channel m:

$$y_i^m = x_i^m \cdot h_{ii}^m + x_m^m \cdot h_{mi}^m + \sum_{j \in \Phi_m, j \neq i} x_j^m \cdot h_{ji}^m + z_i^m \tag{1}$$

where x_i^m represents the desired signal from SU_i on channel m, while x_m^m and x_j^m represent interfering signals from PU_m and SU_j, respectively. Similarly, the variables h_{ii}^m, h_{mi}^m, and h_{ji}^m represent the channel gain from the transmitter to SU_i at SU_i, PU_m, and SU_j, respectively. Additionally, z_i^m represents additive white Gaussian noise (AWGN). The corresponding signal to interference plus noise ratio (SINR) is:

$$SINR_i^m = \frac{p_i^m \cdot \left|h_{ii}^m\right|^2}{p_m^m \cdot \left|h_{mi}^m\right|^2 + \sum_{j \in \Phi_m, j \neq i} p_j^m \cdot \left|h_{ji}^m\right|^2 + B \cdot N_0} \tag{2}$$

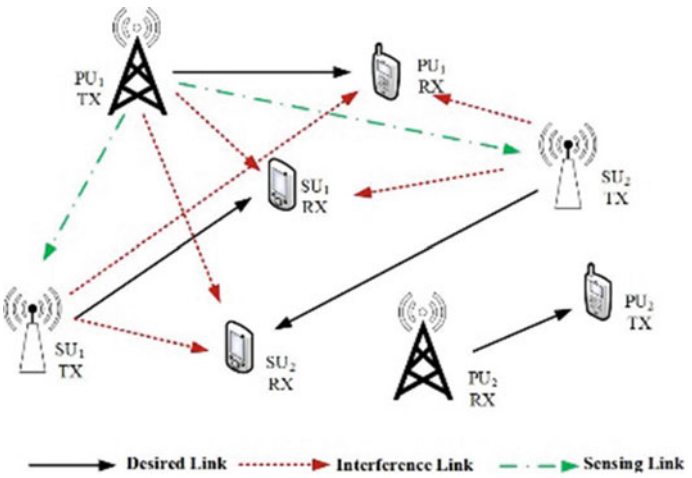


Fig. 1 System model

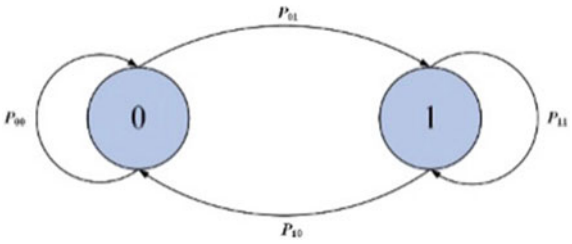


Fig. 2 Two-state Markov chain

where p_i^m, p_m^m and p_j^m denote the transmit power of users i, m and j on channel m. B and N_0 are the channel bandwidth and noise spectral density, respectively. The transmission rate C_i received by the receiver of SU_i is:

$$C_i = \log_2(1 + SINR_i) \tag{3}$$

Equations (2) and (3) show that optimal for only one SU to transmit on an inactive channel.

We divide the spectrum hole of the authorised channel into multiple time slots. The channel occupancy as a two-state Markov chain, as shown in Fig. 2, where 0 represents an occupied channel and 1 represents a vacant channel. The transition probability of the two-state Markov chain on the ith channel is:

$$p_i = \begin{bmatrix} p_{00}^i & p_{01}^i \\ p_{10}^i & p_{11}^i \end{bmatrix} \tag{4}$$

where $p_{xy} = \{\text{the next state is } x | \text{the current state is } y\}$, $(x, y \in \{0, 1\})$.

2.1 State

At the beginning of each time slot, SU_i conducts spectrum sensing on N channels to obtain information about the state of the channel. The state of the channel in the t-th time slot is expressed as follows:

$$s_i = \left[s_i^1, s_i^2, \cdots, s_i^N\right] \tag{5}$$

where $s_i^n = 1$ or $s_i^n = 0$. Since the spectrum detector is not perfect, the results of sensing the channel state may contain errors. We define the probability of sensing error for SU_i on channel n as P_i^n. Therefore, the probability of observing the true state o_i of the channel is given by:

$$\Pr(o_i) = s_i \cdot \left(1 - P_i^n\right) + (1 - s_i) \cdot P_i^n \tag{6}$$

The SU does not konw whether a spectrum sensing error will occur. Consequently, the observed results are mainly used as historical channel state data in this paper. The perception outcomes acquired by the SU in the presence of possible spectrum sensing errors are denoted as:

$$o_i = \left[o_i^1, o_i^2, \cdots, o_i^N\right] \tag{7}$$

2.2 Action

After spectrum sensing, the SU determines whether to access a channel based on the sensing result. The action of SU_i is denoted by $a_i \in \{0, 1, \cdots, N\}$, where $a_i = n(n > 0)$ indicates that at time slot t, SU_i chooses to transmit on the nth channel, while $a_i = 0$ indicates that SU_i chooses not to transmit. The action of each SU is denoted as:

$$A = \{a_1, a_2, \cdots, a_N\} \tag{8}$$

2.3 Reward

SUs receive rewards based on the actions they take. Principles for SU access to a channel include minimizing collisions with other SUs and avoiding interference with the PU to maximize their own transmission rate. The reward function is defined as:

$$r_i = \begin{cases} -\mathrm{C} & \text{, collision with PU} \\ 0 & \text{, no channel access} \\ \log_2(1 + SINR_i) & \text{, successful access} \end{cases} \tag{9}$$

Specifically, the reward is set to $-\ C\ (C > 0)$ when the SU collides with the PU, and 0 when the SU does not transmit data. Otherwise, the SU's reward is the transmission rate of its receiver.

2.4 Policy

SUs don't konw the probability of the channel state transmission and the sensing errors, so they use these rewards to form an access policy that maximizes their cumulative discounted returns, which can be expressed as:

$$R_i = \sum_{t=1}^{\infty} \gamma_{t-1} r_i(t+1) \tag{10}$$

where $\gamma \in [0, 1]$ is a discounted factor.

In summary, the ultimate goal of DSA is to maximise the reward as given in Eq. (10). The optimal Q value is calculated using the following equation to find the optimal policy π^*.

$$\pi^* = \underset{a_i \in A}{\operatorname{argmax}}\, Q_{\pi^*}(o_n, a_i) \tag{11}$$

3 Proposed DRL Algorithm

Since the efficiency of Q-learning deteriorates as the state and action space increases, we address the inefficiency of Q-learning by incorporating DNNs. The DQN architecture we use is shown in Fig. 3.

In the training phase of the DQN, as intelligent agent, each SU uses its observations at each time slot as input to the DQN evaluation network. The evaluation network selects actions using the ε-greedy strategy. After the SU takes action a_i, it receives a reward r_i from the environment and inputs channel observations $o_i\prime$ into the target network at the next time slot to obtain the next time slot action $a_i\prime$ and the target Q value $max_{a_i\prime} Q\left(o_i\prime, a_i\prime; \theta'\right)$. $(o_i, a_i, r_i, o_i\prime)$ represents an experience that is collected and stored in the experience pool by the ε-greedy strategy before training starts. The accumulated experiences in the experience pool are used to calculate the loss value during the DQN training:

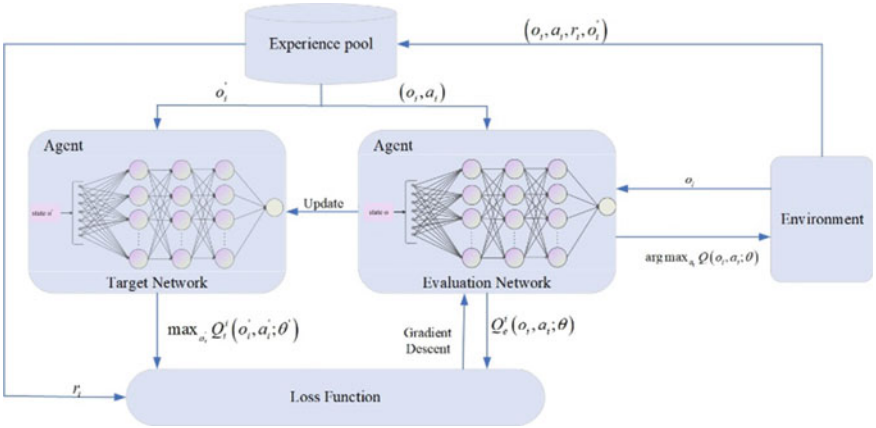


Fig. 3 The framework of DQN

$$loss = \left[r_i + \gamma \max_{a_i} Q_t^i\left(o_i\prime, a_i; \theta'\right) - Q_e^i(o_i, a_i; \theta) \right]^2 \quad (12)$$

The parameters θ of the evaluation network are updated using the calculated loss values through back propagation, and the parameters of the evaluation network are periodically copied to the target network to update its parameters θ'.

4 Simulation Results

We conducted simulation experiments in an environment where 2 SUs coexist with 6 PUs, and their positions were randomly set within a 150 m × 150 m area. The SUs were placed within range of 20–40 m from each other. We used the WINNER II and Rician models to calculate the path loss and channel model, respectively. We randomly selected p_{11} from the uniform distribution [0.7, 1] and p_{00} from [0, 0.3]. We then calculated $p_{10} = 1 - p_{11}$ and $p_{01} = 1 - p_{00}$. The parameters of the system model are shown in Table 1.

To improve the training accuracy and address the performance degradation of deep neural networks due to network depth, we designed the DNN structure in our DQN as a ResNet structure with four hidden layers, as shown in Fig. 4. Each hidden layer contains 64 neurons with Rectified Linear Unit (ReLU) as the activation function. In order to avoid sub-optimal decision strategies before gaining sufficient learning experience, we used the decaying ε-greedy algorithm with an initial value of ε set to 1. At each time slot, ε was decayed according to $\varepsilon \leftarrow \max\{0.995*\varepsilon, 0.005\}$. The hyperparameters are provided in Table 2.

We conducted simulations using Python and TensorFlow to evaluate the performance of our proposed algorithm DQN + MLP4 + ResNet against several other algorithms: myopic algorithm [13], DQN + RC [11], Q-learning, and DQN with only four fully connected layers (DQN + MLP4). We compared the algorithms based on their cumulative rewards, success rate, and conflicts with PUs and other SUs.

Our proposed algorithm has demonstrated superior performance compared to other algorithms, as shown in Figs. 5, 6, 7 and 8. Figure 5 shows that our algorithm achieved the highest average reward compared to other algorithms, while Fig. 6 shows

Table 1 Parameters of system model

Parameters	Value
Number of PUs N	6
Number of SUs M	2
Noise spectral density N_0	− 174 dBm/Hz
Transmission power of PU	40 mW
Transmission power of SU	20 mW

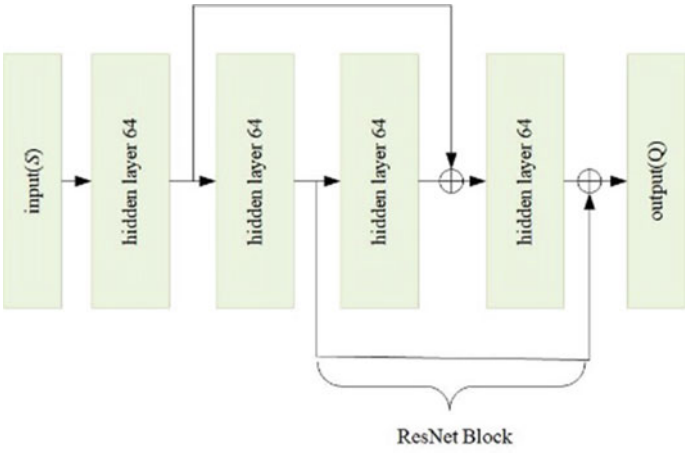


Fig. 4 The structure of deep neural networks for DQN algorithm

Table 2 Hyperparameters of DQN algorithm

Hyperparameters	Value
ε in ε-greedy policy	$1.0 \rightarrow 0.005$
Learning rate α	0.01
Discount rate γ	0.9
Activation function	ReLu
Memory size	2000
Optimzer	Adam
Target network update frequency	300

that our algorithm achieved a much higher access channel success rate, reaching approximately 95%. Figure 7 shows that all learning-based algorithms, except for the myopic policy, eventually reach a zero conflict rate with other SUs, indicating that they learn the access policies of other SUs by interacting with the environment. However, the myopic policy only accesses the channel that brings the maximum expected reward based on the known system channel information, and cannot learn the access policies of other SUs. To prevent conflicts with PUs, we set the reward to -2, and as depicted in Fig. 8, our proposed algorithm achieves the lowest collision rate with PUs, even lower than the myopic policy.

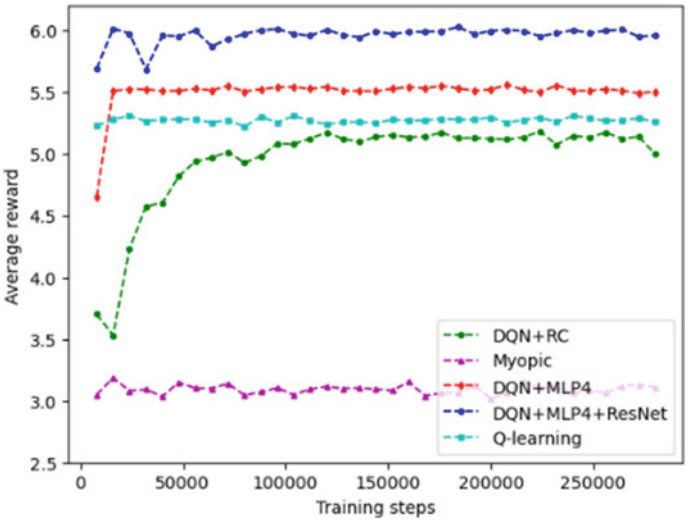


Fig. 5 The average reward

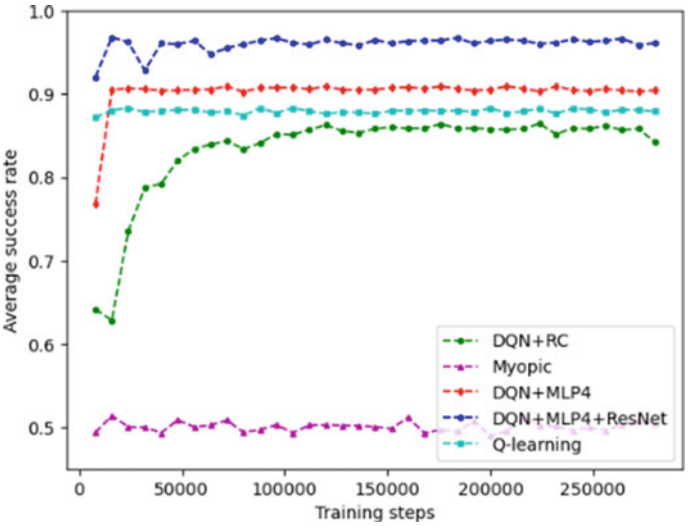


Fig. 6 The average success rate

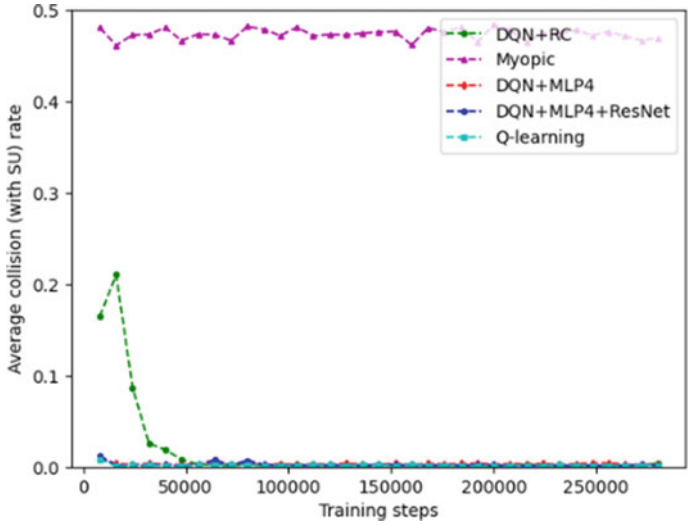


Fig. 7 The average collision with SU

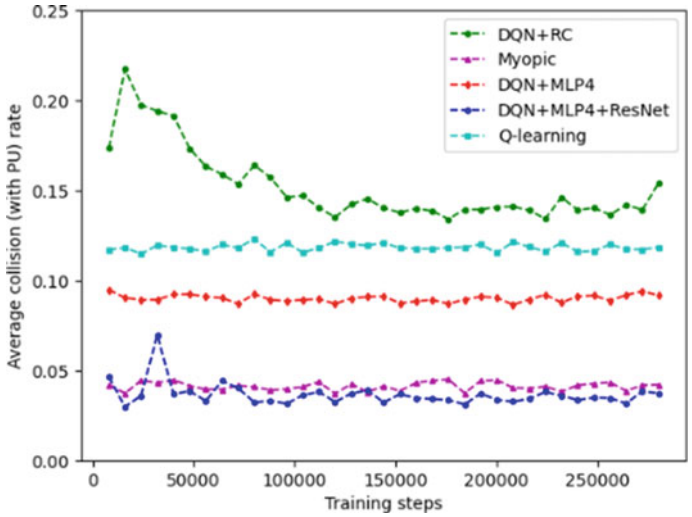


Fig. 8 The average collision with PU

5 Conclusion

This study addresses the spectrum access problem in distributed DSA networks with spectrum sensing errors, and proposes a DSA algorithm that combines DQN with ResNet. Simulation results demonstrate that the proposed DQN + MLP4 +

ResNet algorithm facilitates SUs to learn the optimal channel access policy more efficiently, improves spectrum access opportunities, and effectively reduces inter-user collisions when SUs have incomplete knowledge of the environment and face certain perception errors. In future work, we plan to consider more practical spectrum sharing scenarios and further improve the performance of the algorithm.

Acknowledgements Shubin Wang (wangshubin@imu.edu.cn) is the correspondent author and this work was supported by the National Natural Science Foundation of China (61761034).

References

1. Carie A, Li M, Marapelli B et al (2019) Cognitive radio assisted WSN with interference aware AODV routing protocol. J Ambient Intell Humaniz Comput 10:4033–4042
2. Song H, Liu L, Ashdown J et al (2021) A deep reinforcement learning framework for spectrum management in dynamic spectrum access. IEEE Internet Things J 8(14):11208–11218
3. Cai P, Zhang Y (2020) Intelligent cognitive spectrum collaboration: Convergence of spectrum sensing, spectrum access, and coding technology. Intelligent and Converged Networks 1(1):79–98
4. Qian B, Zhou H, Ma T et al (2020) Leveraging dynamic stackelberg pricing game for multi-mode spectrum sharing in 5G-VANET. IEEE Trans Veh Technol 69(6):6374–6387
5. Liu X, Sun C, Yu W et al (2021) Reinforcement-Learning-based dynamic spectrum access for software-defined cognitive industrial internet of things. IEEE Trans Industr Inf 18(6):4244–4253
6. Kaur A, Kumar K (2020) Imperfect CSI based intelligent dynamic spectrum management using cooperative reinforcement learning framework in cognitive radio networks. IEEE Trans Mob Comput 21(5):1672–1683
7. Cong Q, Lang W (2021) Double deep recurrent reinforcement learning for centralized dynamic multichannel access. Wirel Commun Mob Comput 2021:1–10
8. Doshi A, Yerramalli S, Ferrari L et al (2021) A deep reinforcement learning framework for contention-based spectrum sharing. IEEE J Sel Areas Commun 39(8):2526–2540
9. Guo Z, Chen Z, Liu P et al (2022) Multi-agent reinforcement learning-based distributed channel access for next generation wireless networks[J]. IEEE J Sel Areas Commun 40(5):1587–1599
10. Cong Q, Lang W (2021) Deep multi-user reinforcement learning for centralized dynamic multichannel access/2021. In: 6th international conference on intelligent computing and signal processing (ICSP). IEEE, pp 824–827
11. Chang HH, Song H, Yi Y et al (2019) Distributive dynamic spectrum access through deep reinforcement learning: A reservoir computing-based approach. IEEE Internet Things J 6(2):1938–1948
12. Chang HH, Liu L, Yi Y (2020) Deep echo state Q-network (DEQN) and its application in dynamic spectrum sharing for 5G and beyond. IEEE Trans Neural Netw Learn Syst 33(3):929–939
13. Li Y, Jayaweera SK, Bkassiny M et al (2012) Optimal myopic sensing and dynamic spectrum access in cognitive radio networks with low-complexity implementations. IEEE Trans Wireless Commun 11(7):2412–2423

Study on the Inversion of Dissolved Oxygen in Water Based on HJ-1A Hyperspectral Data

Lei Yang, Xiaoyan Tang, and Haotian Hu

Abstract The content of dissolved oxygen (DO) is a necessary condition for the survival of various marine organisms and an important water quality parameter for monitoring water pollution. Hyperspectral data is gradually applied to the inversion of water quality parameters due to its advantages such as high spectral resolution and wide spectral range. In this study, HJ-1A hyperspectral images were used to invert the dissolved oxygen concentration in the waters near Hong Kong. Firstly, the 2nd-level image product taken by the HJ-1A satellite's HSI sensor was converted into the appropriate format, and then atmospheric correction using FLAASH and geometric calibration were performed. Next, a correlation analysis was conducted on the measured dissolved oxygen concentration and spectral reflectance to select sensitive bands. Finally, the bands corresponding to reflection peaks and absorption troughs were selected for band combination, and a semi-empirical linear regression method was used for modeling. The model determination coefficient R^2 was 0.639, indicating a relatively high accuracy of inversion.

Keywords Hyperspectral remote sensing · HJ-1A · Dissolved oxygen · Semi-empirical linear regression

1 Introduction

With the development of industry and technology, both urban domestic sewage and industrial wastewater are discharged into the ocean, causing pollution to marine organisms. A key issue in protecting the marine ecosystem is monitoring specific water quality parameters in the marine environment. The content of dissolved oxygen [1] (DO) in water quality parameters is a necessary condition for the survival of

L. Yang (✉)
Henan Polytechic Institute, Nangyang 473009, China
e-mail: 547717964@qq.com

L. Yang · X. Tang · H. Hu
Nanyang Institute of Technology, Nanyang 473004, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_21

various marine organisms, and it is an important indicator for monitoring the water quality status and eutrophication degree of seawater.

Traditional water quality monitoring methods are constrained by multiple factors, with long and limited sampling that can only monitor water quality conditions of a section, unable to achieve large-scale rapid monitoring. Remote sensing technology has advantages such as fast monitoring speed, a large monitored range, and periodic monitoring, and is widely used in fields such as land resource surveys, urban environmental monitoring, agriculture, forestry, atmosphere, hydrology, etc. [2]. Sharaf et al. [3] proposed a backpropagation neural network model based on Landsat 8 for the estimation of optical and non-optical parameters of water bodies. Cui et al. [4] used the combination of the third and fourth bands of Landsat 8 OLI remote sensing images to construct a model for the spatial distribution of dissolved oxygen in the coastal waters of the Pearl River Estuary. Zhuang et al. [5] indirectly inverted dissolved oxygen using multispectral Landsat 8 remote sensing with hydrocarbon concentration as an intermediate variable, after analyzing field measured data in Dalian Port. Du et al. [6] used the unique water quality spectral information of salt lake areas to perform a correlation analysis between water surface reflectance combined with hyperspectral bands and dissolved oxygen, and developed a dissolved oxygen inversion model using band ratio method.

In general, there have been more satellite remote sensing studies on the retrieval of water quality parameters such as chlorophyll-a and nutrients, and less research on the remote sensing retrieval of dissolved oxygen in seawater. Existing studies have mainly used Landsat series multispectral image data, which has low spectral resolution and generally yields moderate model accuracy. The hyperspectral imaging spectrometer carried by the Environmental Satellite 1 has a high spectral resolution. Therefore, in this study, HJ-1A satellite hyperspectral imaging (HSI) data were used to invert dissolved oxygen (DO) concentrations in the Hong Kong offshore area. A DO inversion model based on hyperspectral data was established to provide a basis for remote sensing monitoring of DO in the offshore area.

2 Data Acquisition and Preprocessing

2.1 Acquisition of Measured Data

The land area of the HKSAR is 1106 km^2 and the sea area is 1649 km^2. The study area is located at 22°26′ north latitude and 114°12′ east longitude. The measured DO data used in this paper came from 76 water quality monitoring stations in Hong Kong Special Administrative Region [7], and the sampling points of the monitoring stations are shown in Fig. 1.

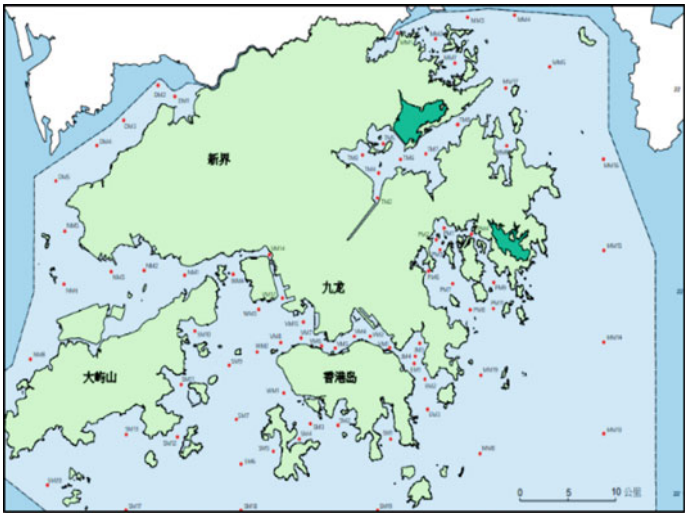


Fig. 1 Location of sampling points of monitoring stations

2.2 Acquisition and Preprocessing of Remote Sensing Image Data

The HJ-1A hyperspectral satellite image used in this paper can be downloaded from the website of China Resources Satellite Application Center. According to the basic requirements of remote sensing image required: high resolution, less striping noise, less cloud coverage than 5%, and keeping the date of satellite image shooting consistent with the date of water source data as far as possible to ensure the quasi-synchronization of ground monitoring station data and satellite hyperspectral data. The downloaded level 2 product of HJ-1A satellite HSI sensor [8] is the product data after spectral repair, radiometric correction and systematic geometric correction, and the corrected image is mapped to the map projection coordinates of the satellite image shooting location through ArcGIS.

The preprocessing process of hyperspectral image includes: spectral curve radiation brightness value conversion, hyperspectral data storage format conversion, low signal-to-noise ratio band removal. Since the sensor carried by the HJ-1A hyperspectral satellite is inevitably absorbed and scattered by the atmosphere, there are some differences between the surface reflection data and the actual measured values. In order to minimize the influence of the atmosphere on the ground reflection spectral data, the hyperspectral image is used to invert the reflection spectral information of ground objects, and atmospheric correction is needed.

After atmospheric correction, in order to fully match the latitude and longitude of sampling points with hyperspectral satellite images and improve the accuracy and accuracy of subsequent modeling, geometric precision correction was carried out on the images. In this paper, satellite images of Landsat-8 in the same period and region

downloaded from geospatial data cloud website are adopted. As Landsat-8 image products are of higher grade and have undergone geometric precision correction, the longitude and latitude of hyperspectral images taken by HJ-1A can be calibrated by Landsat-8.

3 Semi-Empirical Linear Regression Modeling

3.1 Pearson Correlation Analysis

Semi-empirical method [9] combines hyperspectral remote sensing data with ground measured water quality data, uses statistical analysis to find out the best correlation bands of water quality parameters, and then uses linear regression, nonlinear regression, principal component analysis and other methods to build water quality inversion models. When the inversion model of water quality parameters is established, spectral characteristics of water quality parameters are first found from spectral curves of hyperspectral images, and Pearson correlation analysis is carried out between the concentration of water quality parameters and the reflectance of water bodies. Pearson correlation coefficient [10] is used to measure the strength of the linear relationship between X and Y of two groups of data. Moreover, when the absolute value of Pearson correlation is in different intervals, it represents different correlations, so the absolute value of correlation should be as close to 1 as possible. The band with the largest absolute correlation value is found and the corresponding inversion model is established.

In this paper, SPSS 25 software is used to analyze the correlation between the single-band reflectance of HJ-1A remote sensing satellite and the DO concentration. Correlation between measured data and remote sensing reflectance established in SPSS, Pearson correlation coefficient figure is shown in Fig. 2.

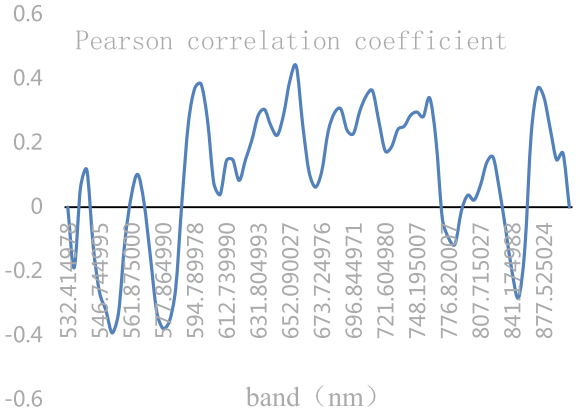


Fig. 2 Pearson correlation coefficient of single band

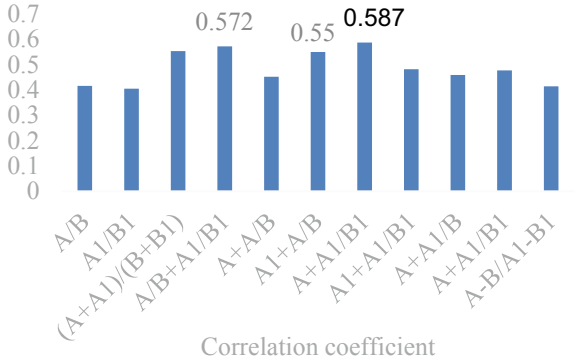


Fig. 3 Pearson correlation coefficient of band combination

Through the analysis of the measured dissolved oxygen data after the envelope removal and the Pearson correlation coefficient image established by the spectral curve of HJ-1A hyperspectral image, it is found that there are obvious reflection peaks at 598 and 656 nm, and obvious absorption valleys at 609 and 669 nm. Through multiple combination tests on the reflection peak band and absorption valley band in these two places, the band combination with high correlation coefficient can be found. Therefore, the phase addition and subtraction, phase multiplication and division of reflection peaks and absorption valleys of 609 nm, 669 nm, 598 nm and 656 nm were respectively conducted, and the following correlation relationships were finally obtained. The Pearson correlation coefficient of band combination is shown in Fig. 3.

3.2 Establishment of DO Inversion Model

In this paper, A + A1/B1, which has the highest correlation coefficient, is selected, and a linear regression equation is established between the reflectance and dissolved oxygen after the combination of this band. In order to ensure the accuracy of the model and reduce random errors as much as possible, 20 monitoring points are randomly selected as training sets and the remaining 9 as prediction sets, and the average results are taken. The final inversion model is $y = 39.68x4 - 290.83x3 + 787.52x2 - 931.03x + 409.6$, and the R^2 obtained by calculation is 0.639, which has a good effect. The dissolved oxygen concentration of the remaining 9 monitoring stations was predicted by the model, and the accuracy of the dissolved oxygen model was verified. The fitting curve between the measured value and the predicted value can well show the applicability of the model. The fitting curve between the measured value and the predicted value of DO concentration is shown in Fig. 4. According to the linear regression curve analysis of measured value and predicted value, it can be seen that the degree of fitting is good, and the R^2 index is up to 0.6698, indicating that the mathematical model is effective in establishing the relationship between dissolved oxygen content and remote sensing image.

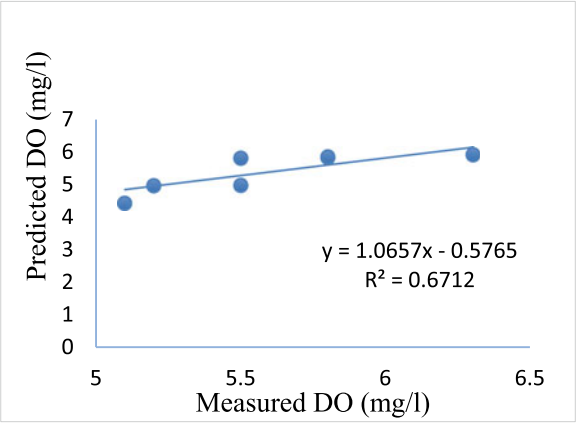


Fig. 4 Fitting curve of measured value and predicted value of DO

4 Conclusion

The traditional water quality monitoring methods are laborious and limited by various factors, so it is difficult to meet the requirements of large water area, real-time spatial dynamics, convenient and rapid monitoring. The research of water quality monitoring technology based on remote sensing becomes more and more important with the serious water pollution. In this paper, the advantages of hyperspectral remote sensing were given full play, and the offshore region of Hong Kong was taken as the research object. The field monitoring data and the hyperspectral data image of environmental and disaster monitoring and prediction satellite HJ-1A were used to establish a linear regression equation to monitor the concentration of dissolved oxygen in the offshore region of Hong Kong. The model determination coefficient R^2 was 0.639, indicating a high inversion accuracy. It provides powerful technical support for environmental monitoring in Hong Kong offshore area.

References

1. Mallin MA, Williams KE, Esham EC et al (2000) Effect of human development on bacteriological water quality in coastal watersheds. Ecol Appl 10(4):1047–1056
2. Chen LF, Yan J, Fan WJ et al (2016) Twentieth anniversary of the journal of remote sensing. J Remote Sens 20(5):794–806
3. El Din ES, Zhang Y, Suliman A (2017) Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework. Int J Remote Sens 38(3–4):1023–1042
4. Cui WJ, Xia LH, Xie XT et al. (2017) A model of dissolved oxygen in the Pearl River estuary based on measured spectrum. J Guangzhou Univ (Nat Sci Edn) 16(6):84–92
5. Zhuang Y (2020) Study om remote sensing estimation of dissolved oxygen in petroleum-contaminated water in Dalian Port. Guangdong Ocean University, Zhan Jiang
6. Du C, Li DL, Li GJ (2021) Application and exploration of dissolved oxygen inversion of plateau salt lakes based on spectral characteristics. Remote Sens Nat Resour 33(03):246–252

7. Sun HL, He HC, Fu BL et al (2020) Quantitative inversion and analysis of spatio-temporal changes of chlorophyll-a concentration in Hong Kong's coastal waters. China Environ Sci 40(5):2222–2229
8. Qiu Y, Yu Y, Guo A, Feng ZF, Song ZL (2015) Technology process of pre-processing for HJ-1A HSI data. Forest Eng 31(04):83–86
9. Xing QG (2007) Retrieval of water quality in the pearl river estuary using hyperspetral technique. In: Guang Z (eds) Graduate University of Chinese Academy of Sciences (South China Sea Institute of Oceanology)
10. Gao XH, Li XQ (2022) Comparison of dimensionless methods in multiple linear regression models. Stat Decis 38(06):5–9

Research on the Impact of 5G Terminals on Electromagnetic Radiation of 5G Base Stations

Ye Tian and Hui Kang

Abstract The Ministry of Ecology and Environment released the "5G mobile communication base station electromagnetic radiation environmental monitoring methods (for trial implementation)" (HJ1151-2020) standard in 2020, which specifies the content, methods and other technical requirements for electromagnetic radiation environmental monitoring of 5G mobile communication base stations. In this standard, it is clearly stated that when monitoring, the monitored mobile communication base station should be in normal working condition, and the 5G terminal equipment should be connected to the monitored 5G mobile communication base station and be in at least one typical application scenario. It is also stated that the 5G terminal equipment and the monitoring equipment probe should be kept within the range of 1–3 m during monitoring. This paper uses frequency-selective electromagnetic radiation field meter (EMF Meter) and 5G NR spectrum analyzer to test different application scenarios of 5G terminals and different relative positions with the EMF probe for different effects on the electromagnetic radiation monitoring values of 5G base stations from several aspects.

Keywords 5G NR · EMF · Application scenarios

Y. Tian
Zhonghuan Information College Tianjin University of Technology, Tianjin, China

H. Kang (✉)
Deviser Instruments Co., Ltd, Tianjin, China
e-mail: kanghui@deviser.com.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_22

1 5G NR Mobile Communication System

1.1 Introduction of 3GPP Release 15, 16, 17 for the History of 5G

The 3GPP (Third Generation Mobile Partnership Project) announces a version freeze every once in a while, which means that no new technical features will be added to that version, and newly generated technological innovations will be reflected in the next version. 5G is a technology that continues to evolve over a decade, and currently 5G has gone through three versions, R15, R16 and R17.

The 5G R15 version was frozen in June 2018, with R15 focusing on eMBB (enhanced mobile broadband), laying the foundation for 5G.

The 5G R16 version freezes in July 2020, adding NR-U, eURLLC, NR V2X, 5G broadcast, etc., expanding industry applications, especially in vertical areas.

The 5G R17 release freeze in June 2022 further extends the 5G technology base in terms of network coverage, mobility, power consumption and reliability, and broadens 5G to new use cases, deployment methods and network topologies [1, 1].

1.2 Channel and Beam Profile for 5G NR

In 3GPP TS38.104, the bandwidth requirements for the use of the 5G FR1 band range from a minimum of 5 MHz to a maximum of 100 MHz.5G NR defines two downlink synchronization signals: Primary Synchronization Signal (PSS) and Secondary Synchronization Signal (SSS). PSS, SSS and the physical broadcast channel PBCH together form a structure called SS/PBCH Block, or SSB for short, which occupies a limited and fixed physical bandwidth [3].

The SSB is sent in the 5G NR downlink frame with a default period of 20 ms for terminal synchronization [4]. The number of SSBs per 20 ms is the number of beams. The number of beams varies from 1 to 8 and is transmitted in different directions. This is shown in Fig. 1.

In the downlink transmission between base station and terminal, SSB is only responsible for synchronization and broadcasting information, while the service information of each terminal is carried in the physical downlink shared channel with the maximum bandwidth up to 100 MHz.

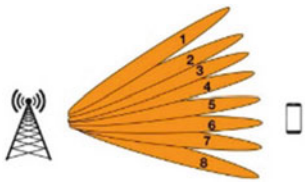


Fig. 1 5G NR downlink beam transmitting schematic

2 Electromagnetic Radiation from 5G Base Stations in Different 5G Terminal Application Scenarios

2.1 Application Scenarios Specified in HJ1151-2020

According to the "5G mobile communication base station electromagnetic radiation environmental monitoring methods (for trial implementation)" (HJ1151-2020), Sect. 2.4 application scenario requirements, 5G mobile communication application scenarios include: enhanced mobile broadband (eMBB), ultra-high reliability and low latency communication (uRLLC), large-scale machine-like communication (mMTC), such as data transmission, video interaction, gaming and entertainment, virtual shopping, smart medical, industrial applications [5].

In order to get several application scenarios, Sect. 2.3 in HJ1151-2020 requires that: when monitoring, the monitored mobile communication base station should be in normal working condition, and the 5G terminal equipment should be connected to the monitored 5G mobile communication base station and be in at least one typical application scenario.

When monitoring, the monitoring instrument probe is placed on the monitoring instrument stand, and the distance between the tip of the probe and the operator's torso is not less than 0.5 m, and the 5G terminal equipment is kept within 1–3 m; avoid or minimize the interference of other electromagnetic radiation sources incidental to the surrounding area and the impact of leakage current from the monitoring instrument stand [6].

2.2 Problem Formulation

We can't help but ask the following questions: In the HJ1151-2020 standard, why does the application scenario for 5G terminal equipment and the relative distance between the 5G terminal setup and the probe make a description of the requirements, namely:

1. What impact will the different application scenarios of 5G terminals have on the measured electromagnetic radiation values of 5G base stations?
2. What effect will the different relative positions of the 5G terminal and the probe of the monitoring instrument have on the electromagnetic radiation measurement value of the 5G base station?

In order to solve the above two questions, we use the base station electromagnetic radiation function of the EMF meter to measure a 5G base station, and use the 5G NR spectrum analyzer as an auxiliary instrument to verify the cause of the corresponding phenomenon.

2.3 Practical Testing

To verify question 1, we use the 5G terminal working in non-data download state (other mode) as well as data transmission in two application scenarios, respectively, at the same measurement location and when the 5G terminal is close to the probe (1–3 m range).

To verify question 2, we move the 5G terminal away from the field intensity meter probe, while the 5G terminal and the electromagnetic probe are in the same beam area and different beam areas of the 5G base station, respectively, and the 5G terminal performs data transmission application scenarios.

During the experiment, a EMF meter and a 5G NR spectrum analyzer are placed close to a 5G base station at a distance of about 100 m, and the probe of the EMF meter is at the same height as the antenna of the spectrum analyzer, with a position difference of 20 cm. The tester places the 5G terminal at different ranges from the EMF probe and works in different application scenarios to observe the measured radiation values of the base station and the spectrum tested by the spectrum analyzer.

Use the frequency selective measurement function of the EMF meter; use the waterfall spectrum function of the spectrum analyzer to observe the 5G base station signal emission and electromagnetic radiation in different application scenarios of the 5G terminal, respectively.

For this test, we selected China Mobile's 5G NR n41 band. The China Mobile n41 band uses TDD time division multiplexing mode with a bandwidth of 100 MHz.

- Group 1 and Group 2 tests

Relative distance from the 5G terminal to the electromagnetic probe

The distance between the 5G NR terminal and the probe of the EMF meter is 1 m, and the 5G NR terminal and the probe are located in the same beam radiation direction of the base station. As shown in Fig. 2.

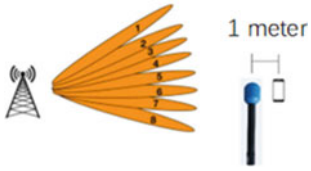


Fig. 2 The first group and the second group test layout diagram

Table 1 Group 1 and group 2 test application scenarios

Test group	Group 1	Group 2
Application scenarios	5G NR terminal working in non-data download state	5G NR terminal working in data transmission state

Fig. 3 The first group of test measurement results

The first group of test results

From the electromagnetic field intensity meter's frequency selection measurement mode we measured that the average value of electromagnetic radiation within 6 min for the China Mobile n41 5G band was 0.003 uW/cm^2. Use the waterfall spectrum function of the spectrum analyzer. As shown in Fig. 3.

The second group of test results

From the selected frequency measurement mode of the EMF strength meter we measured that the average value of electromagnetic radiation in 6 min for the China Mobile n41 5G band is 8.249 uW/cm^2. Use the waterfall spectrum function of the spectrum analyzer, as shown in Fig. 4. This value is much higher than the value of the first group of tests.

We observe that the power values are strong for the vast majority of frequencies in the 100 MHz spectrum. Also on the waterfall plot, we observe that the vast majority is red, indicating a strong signal. It also contrasts with the first set of tests. In this scenario we see that the 5G base station service channel has to use the majority of the service channel transmit power within the 100 MHz frequency to transmit data as it has to carry large data volumes of data transmission information.

As seen from the test findings in Table 2, the radiation field strength values of the 5G base station are low when the 5G terminal is not performing data interaction

Fig. 4 The second group of test measurement results

services. The radiation field intensity value of the 5G base station is high when the 5G terminal is carrying out data transmission application scenarios.

The low radiated field strength of the 5G base station in the first test group can also be verified from the spectrum analyzer measurements because most of the RBs in the 100 MHz bandwidth are not loaded and do not send strong power signals. The reason for the high radiated field strength of the 5G base stations in the second group of tests is that most of the RBs in the 100 MHz bandwidth are loaded and send strong power signals in order to achieve high throughput data compliance.

- Group 3 test

Compared with the first and second group of tests, in the third group of tests, the position of the electromagnetic probe relative to the base station remains the same, while the relative distance from the 5G terminal to the electromagnetic probe is changed. The distance between the 5G terminal and the EMF meter probe is 20 m and the 5G NR terminal and the electromagnetic probe are located in the same beam radiation direction of the 5G base station. As shown in Fig. 5.

The application scenarios for the third group of tests: 5G NR terminal working in data transmission state.

Results of Group 3 tests

From the electromagnetic field intensity meter's frequency selection measurement mode we measured that the average value of electromagnetic radiation within 6 min for the China Mobile n41 5G band was 4.932 uW/cm^2, and waterfall spectrum as shown in Fig. 6.

During the test, it can be observed that the real-time measurement value of the electromagnetic radiation field strength meter immediately becomes larger while the

Table 2 Conclusions of group 1 and group 2 tests

Test group	The first group	The second group
5G terminal and probe relative position	Relative distance 1 m in the same beam radiation direction	
Application scenarios	5G NR terminal operating in non-data download state	5G NR terminal working in data transmission state
Electromagnetic radiation value	0.003 uW/cm^2	8.249 uW/cm^2
Waterfall spectrum	Low spectral amplitude on the time axis	High spectral amplitude on the time axis

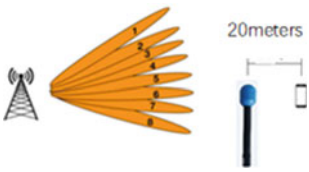


Fig. 5 The third group of test layout diagram

Fig. 6 The third group of test measurement results

spectrum amplitude of the spectrum analyzer waterfall graph becomes larger, but the color of the spectrum analyzer waterfall graph display is not darker than that of the second group test, indicating that the received power at the spectrum analyzer of the third group test is not as large as the received power of the second group test, which also verifies the test value of the electromagnetic field strength meter.

As seen in Table 3, when the 5G terminal is 20 m away from the probe of the EMF meter, and the 5G terminal and the probe are located in the same beam radiation direction of the base station, the tested downlink electromagnetic radiation value of the 5G base station is smaller than the value when the 5G terminal is closer to the electromagnetic probe, but it can still capture the downlink signal during data transmission.

- Group 4 test

Compared with the first and second group of tests, the position of the electromagnetic probe relative to the base station remains unchanged in the fourth group of tests, while the relative distance from the 5G terminal to the electromagnetic probe is changed. The distance between the 5G NR terminal and the probe of the EMF meter is 20 m, but the 5G NR terminal and the probe are located in different beams radiation direction of the base station. As shown in Fig. 7.

The application scenarios of the fourth group of tests: 5G NR terminal working in data transmission state.

Table 3 Third group of test findings

Test group	The third group
5G terminal and probe relative position	Relative distance of 20 m in the same beam radiation direction
Application scenarios	5G NR terminal working in data transmission state
Electromagnetic radiation value	4.932 uW/cm^2
Waterfall spectrum	Medium spectral amplitude on the time axis

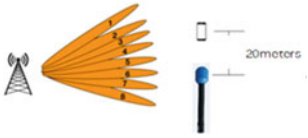


Fig. 7 The fourth group of test layout diagram

Results of Group 4 test

From the electromagnetic field intensity meter's frequency selection measurement mode we measured that the average value of electromagnetic radiation within 6 min for the China Mobile n41 5G band was 0.389 uW/cm^2, and waterfall spectrum as shown in Fig. 8.

During the test, it was observed that the real-time measurements of the electromagnetic radiation field strength meter did not become significantly stronger when the 5G terminal went from the resting state to the moment of starting data transmission, and the spectrum amplitude of the spectrum analyzer waterfall plot could not be observed to become significantly larger.

As seen in Table 4 in the 5G NR terminal and EMF meter probe distance of 20 m, and 5G terminal and the probe is located in different beam radiation direction of the base station, the tested 5G base station downlink electromagnetic radiation value than the 5G terminal distance electromagnetic 20 m but in the same base station beam direction is much smaller, indicating that the 5G base station according to the beam to 5G terminal to transmit downlink signals, in different beam coverage of the electromagnetic field intensity meter can test the radiation value is very low.

Fig. 8 The fourth group of test frequency selection electromagnetic measurement results

Table 4 Group 4 test findings

Test group	Group 4
5G terminal and probe relative position	Relative distance 20 m different beam radiation direction
Application scenarios	5G NR terminal working in data transmission state
Electromagnetic radiation	0.389 uW/cm^2
Waterfall spectrum	Low spectral amplitude on the time axis

Four groups of test conclusions:

- When the 5G terminal is in two different application scenarios of non-data transmission and data transmission, the electromagnetic radiation values of 5G base stations are different. It shows that the test report of HJ1151-2020 standard indicates that the application scenario is critical.
- When the 5G terminal is far away from the electromagnetic probe, but still in the same beam coverage of the 5G base station as the electromagnetic probe, the tested electromagnetic radiation value is lower than that when the 5G terminal is close to the electromagnetic probe.
- When the 5G terminal is far away from the EM probe, and the EM probe is in a different beam coverage of the 5G base station, the tested EM radiation value will be much lower. This means that the 5G base station's beam is aligned with the 5G terminal, and the EMF radiation value is low because the EM probe is not in the same downlink beam signal emission direction.

References

1. Adda S et al (2020) A theoretical and experimental Investigation on the measurement of the electromagnetic field level radiated by 5G base stations. IEEE Access 8:101448–101463
2. Pawlak R, Krawiec P, Zurek J (2019) On measuring electromagnetic fields in 5G technology. IEEE Access 7:29826–29835
3. Werner R, Knipe P, Iskra S (2019) A comparison between measured and computed assessments of the RF exposure compliance boundary of an in-situ radio base station massive MIMO antenna. IEEE Access 7:170682–170689
4. Keller H (2019) On the assessment of human exposure to electromagnetic fields transmitted by 5G NR base stations. Health Phys 5:483–524
5. Riederer M (2003) EMF exposure due to GSM base stations: measurements and limits. IEEE Int Sympos Electr Compat 1:402–405
6. Olivier C, Martens L (2007) Optimal settings for frequency-selective measurements used for the exposure assessment around UMTS base stations. IEEE Trans Instrum Meas 56:1901–1909

Lightweight Infrared and Visible Image Fusion Based on Attention Mechanism and Receptive Field Enhancement

Ting Liu, Yuxin Zhang, Yunsheng Fan, Peiqi Luo, and Guofeng Wang

Abstract To balance the accuracy and speed of image fusion, a lightweight infrared and visible image fusion algorithm is proposed in this paper. Firstly, to improve the feature extraction capability, ECA attention is introduced to enhance information exchange among channels. Secondly, to improve the feature reconstruction ability and lightweight network, RFB is used to enlarge the receptive field and minimize parameters. Finally, experimental results show that the proposed method outperforms other algorithms in quality and speed. The quantitative results demonstrate that our method effectively preserves infrared features and visible details, with minimal color distortion and uniform brightness. Among them, params decreased by 59.5%, and time shortened by 4.38%. Therefore, the proposed method has high engineering application potential.

Keywords Image fusion · Attention mechanism · Receptive field enhancement · Lightweight network

1 Introduction

Infrared and visible light are the most common visual perception methods. Specifically, infrared images can perceive targets under low light or occlusion conditions by capturing thermal radiation information of the scene, but lack texture and environmental brightness details. On the contrary, visible images contain rich texture information [1, 2]. The fusion of infrared and visible light images can fully utilize the complementarity of visual information between the two types of images [3], improving the accuracy of target recognition and tracking under challenging visual conditions such as low light, smoke, fog and occlusion [4].

Traditional image fusion methods [3] are mainly divided into sparse representation based methods, multi-scale transformation based methods, subspace based methods, and saliency based methods. However, extensive manual design are involved,

T. Liu (✉) · Y. Zhang · Y. Fan · P. Luo · G. Wang
Marine Electrical Engineering College, Dalian Maritime University, 116026 Dalian, China
e-mail: liuting0910@dlmu.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_23

resulting in poor robustness and time-consuming. Methods based on deep learning utilize a large amount of data for training and testing, and are effective with accuracy and processing speed. It can be divided into CNN based, GAN based, and AE based methods. In 2017, Liu et al. [5] introduced CNN to multifocal image fusion, but not for infrared and visible image fusion. In 2018, Li et al. [6] improved image fusion quality by utilizing AE and introducing denseblock in DenseFuse. In 2020, Li et al. [7] further enhanced network performance with NestFuse by introducing nest connection and attention fusion strategies. In 2019, Ma et al. [8] applied GAN to achieve infrared and visible image fusion in FusionGAN, and generated fusion images via adversarial games between the generator and discriminator. Overall, complex network structures are adopted for most deep learning based methods to improve image fusion quality. Thus, the algorithms are unable to be deployed on most practical embedded platforms for the large number of parameters and computational complexity.

Therefore, to achieve a balance between image fusion accuracy and speed, we proposed a lightweight infrared and visible image fusion algorithm based on DenseFuse [6] and improved from three aspects: attention channel interaction, receptive field enhancement and lightweight network. The main innovation points of this article are as follows:

- In response to the problem of insufficient information exchange between channels, we introduce Efficient Channel Attention (ECA) into denseblock to fully extract image features.
- To solve the problem of limited decoding part of receptive field and lack of multi-scale reconstruction, we use Receptive Field Block (RFB) to replace the three-layer ordinary convolution to increase the network receptive field. Besides, the use of RFB also enables lightweight networking.
- Numerous experiments indicate that our method can quickly and effectively fuse infrared features and visible details compared to other methods.

The remaining chapters of this article are composed as follows: The proposed image fusion algorithm is introduced in detail in Sect. 2. Section 3 analyzes the superiority of our method over other methods through a large number of experiments. Section 4 provides the conclusion.

2 Method

The proposed infrared and visible image fusion algorithm is detailed in this section. As the fusion algorithm in this article is designed based on DenseFuse [6], the structure of DenseFuse and all improvements are explained in this section.

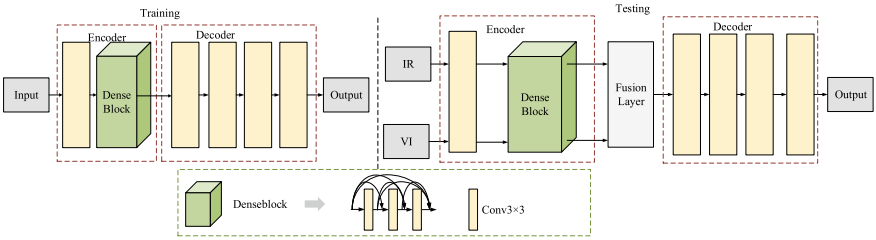


Fig. 1 The structure of DenseFuse

2.1 *DenseFuse: A Fusion Method to Infrared and Visible Images*

The network structure of DenseFuse is shown in Fig. 1, which consists of encoding, decoding, and fusion layer. During training, fusion layer is not used. 3×3 convolution is used for encoding module to extract image features while denseblock is used to preserve detailed information. Stacked 3×3 convolutions are employed to reconstructs the fused image for decoder. During testing, a fusion layer is added to combine infrared and visible image features using either addition or l1 norm fusion. However, DenseFuse still has room for improvement in network design. Firstly, the importance of different channels is not taken into account in denseblock, resulting in inaccurate feature extraction of key parts of the image. Secondly, the stack of 3×3 convolution has limited range of receptive field, which affects the performance of image reconstruction. Therefore, we improved DenseFuse in this paper.

2.2 *Lightweight Infrared and Visible Image Fusion Based on Attention Mechanism and Receptive Field Enhancement*

To solve the problems of DenseFuse in design, we propose a lightweight infrared and visible image fusion algorithm. The structure is shown in Fig. 2. Firstly, ECA is introduced to achieve cross channel interaction and improve feature extraction capabilities. Secondly, RFB is used to replace the three-layer convolution to enhance the receptive field and achieve multi-scale feature reconstruction.

Attention denseblock. To improve the feature extraction capability of the network, the Efficient Channel Attention (ECA) [9] was added to denseblock. ECA is an improved version of SE, and excels at channel attention without reducing dimensionality during cross-channel communication. Firstly, a $1 \times 1 \times C$ vector is obtained through global average pooling. Then, one-dimensional convolution is used to achieve cross channel information exchange. Finally, adaptive convolution kernel size is applied in one-dimensional convolution. The ECA improves computational efficiency with minimal parameters and complexity, resulting in significant gains in

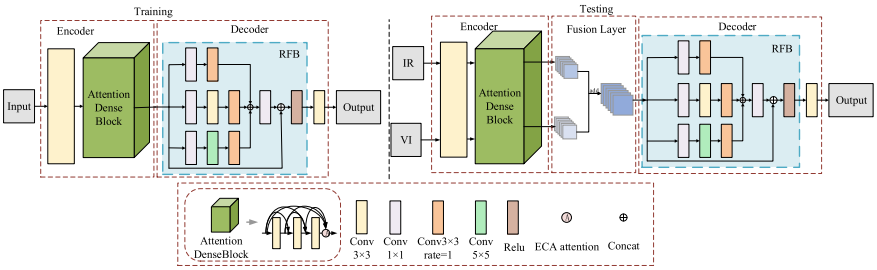


Fig. 2 The structure of our method

feature extraction. Thus, ECA attention was incorporated after dense links in dense block to achieve efficient cross-channel communication.

Receptive field block. In order to make full use of the image features extracted from encoding and retain more details of the image, RFB [10] is used to replace the three-layer convolution of the DenseFuse decoding, so as to enhance the receptive field and realize multi-scale feature reconstruction. The RFB module is a multi-branch convolutional module, as shown in the blue box in Fig. 2. Firstly, 1×1 convolution is used to reduce the number of channels before each branch is convolved. Secondly, convolution of different sizes is used for different branches, and the normal convolution of each branch is followed by a 3×3 empty convolution. Compared with ordinary convolution, the convolution kernel of empty convolution will multiply with pixels at fixed interval points. Under the premise of constant parameters, the sensitivity field is effectively increased, thus filtering the redundant information between adjacent pixels in the image. Therefore, using RFB improves both receptive field and network lightness.

Fusion layer. We use the simpler addition fusion strategy proposed by DenseFuse to ensure a lightweight network. The addition fusion strategy is shown in the Fusion Layer box in Fig. 1, which meets:

$$f_m(x, y) = \sum_{i=1}^{k} \emptyset_i^m (x, y) \tag{1}$$

Among them, m represents the number of feature channels, and k represents the number of input feature maps. (x, y) represents the position of feature map and the fused feature map. $\emptyset_k^m$ represents the mth channel of the kth type of data, f_m is the fusion result.

Loss function. The loss function combines pixel loss and ssim loss to reconstruct the input image more accurately, and is defined as follows:

$$L = L_{pixel} + \lambda L_{ssim} \tag{2}$$

Among them, λ is a parameter set to balance the weight of pixel loss and ssim loss. In the text λ is set to 1. L_{pixel} is the pixel loss, defined as:

$$L_{pixel} = ||O - I||_2 \tag{3}$$

L_{ssim} is the ssim loss, defined as:

$$L_{ssim} = 1 - SSIM(O, I) \tag{4}$$

3 Experiment and Result Analysis

To verify the improvement of our method, detailed experiments and result analysis will be conducted in this section.

3.1 Experimental Details and Parameter Settings

All experiments in this article were completed on a laboratory workstation with a dual core Intel (R) Xeon (R) Silver 4210R processor and NVIDIA GeForce RTX 2080 Ti graphics card. EN [11], SD, Qabf [12], SCD [13], SSIM [14], MS-SSIM [15], params and image processing time were used to evaluate method in this paper. LLVIP [16] dataset was used to train and test the network, and the training set and test set were randomly selected at a ratio of 0.05.

3.2 Comparative Experiment

In order to comprehensively evaluate our method, six different image fusion methods were used to evaluate the fusion effect of our method, including CBF [17], GTF [18], CSR [19], Dualbranch [20], ICAFusion [21] and DenseFuse [6].

The qualitative comparison results are shown in Fig. 3. It was observed that CBF and CSR fusion generated noise, resulting in blurry visible light details and less infrared features. GTF fusion resulted in severe distortion of visible details. Dual-branch produced high-quality images, but when compared to our method, its overall brightness is lower. ICAFusion had an uneven illumination distribution, causing some visible light details to disappear. DenseFuse showed a relatively superior fusion effect among the selected methods with obvious infrared features, clear visible light details. However, our method has even clearer images with uniform overall brightness, which is beneficial for low-light object detection and tracking tasks.

The quantitative results are shown in Table 1, and the optimal values are shown in bold. It can be seen that four evaluation indexes of our method are the first, indicating

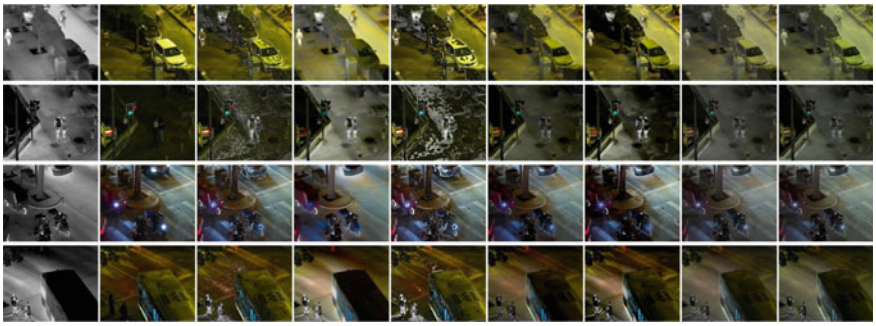


Fig. 3 Qualitative comparison results of different fusion methods

Table 1 Quantitative comparison results of different fusion methods

	EN	SD	Qabf	SCD	SSIM	MS-SSIM	Params	Time
CBF	7.242	8.985	0.622	0.784	0.796	0.733	–	85.320
GTF	**7.450**	**10.022**	**0.640**	1.132	0.916	0.870	–	33.446
CSR	7.322	9.224	0.491	0.743	0.873	0.796	–	40.317
Dualbranch	6.857	9.536	0.383	1.386	0.908	0.905	1.08M	0.411
ICAFusion	6.791	9.233	0.393	**1.493**	0.909	0.874	10.68M	0.931
DenseFuse	6.902	9.480	0.349	1.233	0.911	0.896	0.299M	0.297
Our	6.943	9.576	0.460	1.299	**0.912**	**0.906**	**0.121M**	**0.277**

that our method has the best performance in image brightness, contrast, structural similarity, parameters and processing speed. For EN, SD and Qabf, GTF achieved the best. It shows that this method has certain advantages in the processing of image information richness. For deep learning-based methods, the above three indicators are lower than traditional methods. But our method still takes the top spot among deep learning-based methods. It shows that deep learning based-methods are poor in processing image information richness, which should be paid more attention to in the subsequent research. In addition, our method SCD ranks third, indicating that there is still room for improvement in the correlation between the fusion image and the source image. In summary, our method has certain advantages in image brightness, contrast and image processing speed, and is more suitable for practical applications.

3.3 Ablation Experiment

In order to verify the effectiveness of all improvement points in this article, ablation experiments is conducted on various modules in this section.

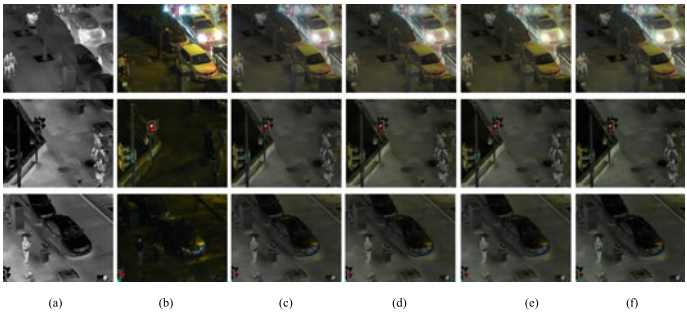


Fig. 4 Qualitative comparison of ablation experiment results

Table 2 Quantitative comparison of ablation experiment results

ECA	RFB	EN	SD	Qabf	SCD	SSIM	MS-SSIM	Params	Time
		6.902	9.480	0.349	1.233	0.911	0.895	0.299M	0.297
✓		6.905	9.496	0.343	1.235	0.912	0.896	0.299M	0.297
	✓	6.918	9.509	0.399	1.285	**0.920**	0.905	0.121M	0.277
✓	✓	**6.943**	**9.576**	**0.460**	**1.299**	0.912	**0.906**	**0.121M**	**0.277**

The qualitative comparison of the ablation experiment results is shown in Fig. 4a–f are infrared images, visible images, DenseFuse fusion results, ECA fusion results, RFB fusion results, and the fusion results of this article. It shows that when ECA was added, the fused image performed better in detail processing. After adding RFB, the infrared features become more prominent. When all improvements are added, the infrared features and visible light details of the image are improved significantly.

Table 2 shows that adding ECA improved all image quality indicators except Qabf, while adding RFB improved all indicators and significantly reduced params and time, and the optimal values are shown in bold. Adding both ECA and RFB improved EN by 0.59%, SD by 1.01%, Qabf by 31.81%, SCD by 53.5%, SSIM by 0.11%, MS-SSIM by 1.23%, reduced params by 59.5%, time by 4.38%. Overall, these improvements had a positive impact on the final image fusion performance.

4 Conclusion

A lightweight infrared and visible image fusion algorithm based on attention mechanism and receptive field enhancement has been proposed in this paper. It is based on DenseFuse. ECA has been added to improve cross channel communication, while RFB is introduced to enhance the reception field and reduce network complexity. The results indicate that our method outperforms other methods in both quantitative

and qualitative analysis. However, our network has limitations in evaluating image richness and future research should aim to improve this aspect and consider color constraints. Combining traditional and deep-learning-based algorithms may also be beneficial.

Acknowledgements This work was supported by China Postdoctoral Science Foundation (Grant No. 2019M661076).

References

1. Ma W, Wang K, Li J, Yang SX, Li J, Song L, Li Q (2023) Infrared and visible image fusion technology and application: a review. Sensors 23(2):599
2. Jin X, Jiang Q, Yao S, Zhou D, Nie R, Hai J, He K (2017) A survey of infrared and visual image fusion methods. Infrared Phys Technol 85:478–501
3. Karim S, Tong G, Li J, Qadir A, Farooq U, Yu Y (2022) Current advances and future perspectives of image fusion: a comprehensive review. Inf Fusion
4. Singh S, Mittal N, Singh H (2021) Review of various image fusion algorithms and image fusion performance metric. Arch Comput Methods Eng 28:3645–3659
5. Liu Y, Chen X, Peng H, Wang Z (2017) Multi-focus image fusion with a deep convolutional neural network. Inf Fusion 36:191–207
6. Li H, Wu XJ (2019) DenseFuse: a fusion approach to infrared and visible images. IEEE Trans Image Process 28(5):2614–2623
7. Li H, Wu XJ, Durrani T (2020) NestFuse: an infrared and visible image fusion architecture based on nest connection and spatial/channel attention models. IEEE Trans Instrum Meas 69(12):9645–9656
8. Ma J, Yu W, Liang P, Li C, Jiang J (2019) FusionGAN: a generative adversarial network for infrared and visible image fusion. Inf Fusion 48:11–26
9. Wang Q, Wu B, Zhu P, Li P, Zuo W, Hu Q (2019) ECA-Net: efficient channel attention for deep convolutional neural networks
10. Liu S, Huang D et al (2018) Receptive field block net for accurate and fast object detection. In: Proceedings of the European conference on computer vision (ECCV), pp 385–400
11. Roberts W, van Aardt J, Ahmed F (2008) Assessment of image fusion procedures using entropy, image quality, and multispectral classification. J Appl Remote Sens 2:1–28
12. Bondzulic B, Petrovic V (2008) Objective image fusion performance measures. Vojnoteh glas 56:181–193
13. Aslantas V, Bendes E (2015) A new image quality metric for image fusion: the sum of the correlations of differences. AEU-Int J Electron Commun 69(12):1890–1896
14. Wang Z, Bovik A, Sheikh H, Simoncelli E (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
15. Ma K, Zeng K, Wang Z (2015) Perceptual quality assessment for multi-exposure image fusion. IEEE Trans Image Process 24(11):3345–3356
16. Jia X, Zhu C, Li M, Tang W, Zhou W (2021) LLVIP: a visible-infrared paired dataset for low-light vision. In: 2021 IEEE/CVF international conference on computer vision workshops (ICCVW), pp 3489–3497
17. Shreyamsha Kumar BK (2015) Image fusion based on pixel significance using cross bilateral filter. Signal Image Video Process 9:1193–1204
18. Ma J, Chen C, Li C, Huang J (2016) Infrared and visible image fusion via gradient transfer and total variation minimization. Inf Fusion 31:100–109
19. Liu Y, Chen X, Ward RK, Wang ZJ (2016) Image fusion with convolutional sparse representation. IEEE Signal Process Lett 23(12):1882–1886

20. Fu Y, Wu XJ (2021) A dual-branch network for infrared and visible image fusion. In: 2020 25th international conference on pattern recognition (ICPR), pp 10675–10680
21. Wang Z, Shao W, Chen Y, Xu J, Zhang X (2022) Infrared and visible image fusion via interactive compensatory attention adversarial learning. IEEE Trans Multimed 1–13

Lightweight Feature Fusion for Single Shot Multibox Floater Detection

Ting Liu, Peiqi Luo, and Yuxin Zhang

Abstract A lightweight feature fusion SSD for USV (unmanned surface vessel) platforms with limited memory is proposed in this paper. Firstly, MobileNetV2 is used to replace VGG to reduce parameter redundancy and improve speed without compromising accuracy. Secondly, a lightweight feature fusion module is utilized to improve accuracy for small targets. On the Flow dataset the average accuracy of the network proposed in this paper is improved by nearly 10%, parameters and GFLOPs is reduced by nearly 50% and 90%, respectively. What's more, the FPS meets the real-time detection requirements. This approach is considered suitable for real-time target detection on USV with limited memory resources.

Keywords Floater · USV · Lightweight backbone networks · Lightweight feature fusion

1 Introduction

USV are intelligent platforms on water that can perform a variety of tasks by carrying different devices. When equipped with optoelectronic systems and target detection algorithms, USV become capable of cleaning and maintaining inner lake water surfaces in a cost-effective and continuous manner, outperforming traditional manual salvage operations. However, USV typically rely on embedded control platforms that demand strict constraints on the number of object detection algorithm parameters and the calculation volume size.

Existing object detection algorithms like Faster R-CNN [1] have complex network structures and low speeds, which are not suitable for real-time tasks. Lightweight networks like MobileNet [2] and ShuffleNet [3] are proposed for general-purpose object detection, which have poor accuracy for small objects and are not suitable for water surface scenarios with many floaters. In contrast, SSD (single shot multibox detector) [4] can detect objects at different scales well by selecting multiple fea-

T. Liu (✉) · P. Luo · Y. Zhang
Marine Electrical Engineering College, Dalian Maritime University, Dalian 116026, China
e-mail: liuting0910@dlmu.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_24

ture maps. Additionally, it has a simple structure with fewer parameters than other algorithms.

Although SSD [4] has shown success in detecting small objects using shallow feature maps, its limitations arise from the absence of semantic information, resulting in poor performance for small objects. Furthermore, the computational demands of SSD [4] necessitate more GPU power, limiting its utility for pure CPU or embedded devices. In this paper, we improve on the SSD [4] by considering the characteristics of floaters and combining them with the requirements of embedded real-time object detection. The main contributions are summarised as follows:

- Lightweighting of the backbone network. A lightweight network MobileNetV2 [2] was used to replace VGG16 [5] as the backbone network. The number of parameters was reduced from 23.745M to 5.610M after the backbone network was lightened and the detection accuracy was not significantly affected.
- Adding a lightweight feature fusion network. In this paper, some shallow feature layers are selected for fusion with deep feature layers to improve the detection accuracy from 51.92 to 62.52%.

The remaining sections of this paper are organised as follows: Related work on object detection lightweighting and SSD networks is introduced in Sect. 2. The lightweight feature fusion SSD network proposed in this paper is detailed in Sect. 3. Experimental demonstration that the network proposed in this paper is more suitable for embedded platforms will be provided in Sect. 4, and conclusions will be drawn in Sect. 5.

2 Related Work

2.1 Lightweight Network

Most existing target detection networks achieve accuracy at the expense of speed by complicating the model and increasing its depth, making it difficult to achieve real-time object detection on embedded platforms. Therefore, the design and development of more efficient object detection networks for real-time embedded tasks is becoming an increasingly popular area. Existing approaches to network lightweighting typically include the following two.

(1) Lightweight network structure. Quan [6] added dilated convolution, which increases feature map resolution and receptive field size. Kanimozhi [7] replaced the backbone network of SSD with a MobileNetDet structure, resulting in faster processing times by using depth-separable convolutions instead of downsampling convolutions. Amudhan [8] improved upon the Shuffle block by adding a channel splitting operation to split the channel into two branches, improving detection accuracy compared to simply replacing the backbone network.

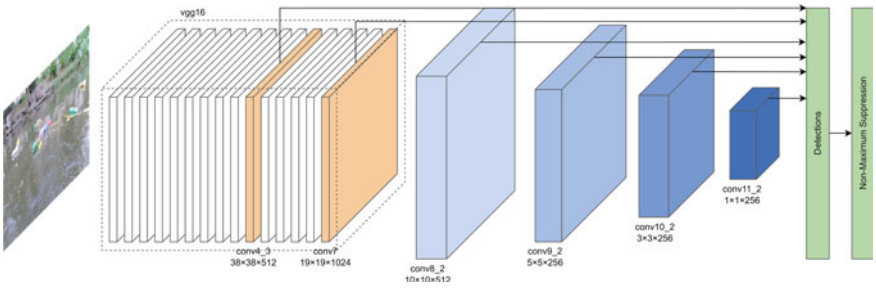


Fig. 1 SSD network structure

(2) Model compression. Zhang [9] arranged the channels according to their importance and pruned a certain percentage of channels according to a threshold value to obtain a more compact model. Chu [10] proposed a group channel pruning method that improved pruning accuracy by obtaining thresholds based on groups with similar feature layers. Chen [11] applied knowledge distillation to train compact models without accuracy loss. Li [12] used Deep Mutual Learning (DML) technique to improve the accuracy of both models before and after pruning.

Although both lightweight network structure and model compression have significant effects in reducing model parameters and computational effort. The lightweight network structure is more effective in reducing the number of parameters and computational effort, and has less impact on model accuracy.

2.2 SSD Network

The structure of SSD [4] is shown in Fig. 1. The basic idea is to extract the image features through the backbone network, then get the adjustment parameters by convolution on the anchor boxes on each pixel point of the feature map, and finally get the results. Specifically, the features will get the position and class prediction of all anchor after passing through the detection network, and then the result will be obtained by non-maximum suppression. The model is trained using a weighted sum of local loss and confidence loss on the predicted anchors, so that the predicted values keep approximating the true values.

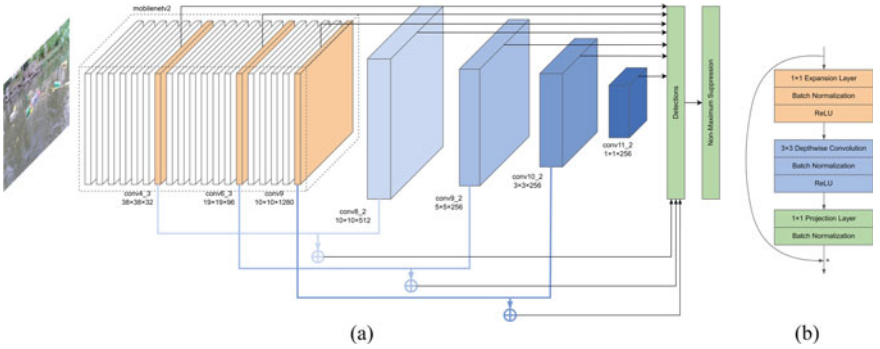


Fig. 2 Lightweight feature fusion SSD structure. **a** Framework. **b** Inverted residual

3 SSD Networks Based on Lightweight Feature Fusion

3.1 Lightweight Backbone Network

The structure of lightweight feature fusion SSD proposed in this paper is shown in Fig. 2a. Considering that the original backbone alone contains 20.48M parameters, making up 86% of the overall number of parameters. To effectively reduce parameter and improve parameter efficiency, this paper uses the more lightweight MobileNetV2 [2] for replacement.

The structure of the Inverted Residual, the fundamental building block of MobileNetV2 [2], is presented in Fig. 2b. The Inverted Residual module considers the input data as compressed real data and performs decompression using the expansion layer to extract high-dimensional features. These high-dimensional features are then passed through the depthwise layer for filtering before compression with the projection layer to obtain low-dimensional features. As non-linear activation results in the partial loss of low-dimensional features, the linear bottleneck does not involve a non-linear activation layer. This structural design strategy provides a commendable trade-off between model size and detection accuracy for small target detection tasks.

3.2 Lightweight Feature Fusion Module

Feature fusion is a method to organically combine shallow and deep feature maps, which helps the network to learn both detailed and semantic information at the same time. This paper borrow the deconvolutional module from DSSD [13], which uses three times of deconvolution to enlarge the deep feature map to the size of the shallow feature map before fusion. The module structure is shown in Fig. 3a, and the SandGlass that used to replace the normal convolution is shown in Fig. 3b.

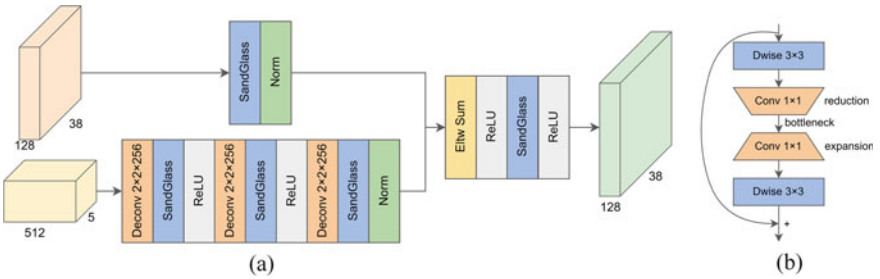


Fig. 3 Lightweight feature fusion module. **a** Framework. **b** SandGlass

To speed up gradient back propagation and improve feature retention, the SandGlass module [14] uses wider, shorter connections to pass more information. This module encodes inter-channel information by placing a bottleneck structure consisting of a 1×1 convolution between two deep convolutions, and maximises the retention of shallow feature information. The improved feature fusion module both enhances feature representation and reduces the computational burden.

4 Experiments and Analysis of Results

The experiment utilized a self-contained workstation featuring an NVIDIA Ge Force RTX 2080 Ti graphics card to train and test the Flow dataset [15]. To randomly divide the dataset into training and test sets, a 9:1 ratio was employed. To address overfitting when training the model on small sample datasets, a data random augmentation strategy was used to expand the training set. The evaluation metrics used for assessing target detection results included mAP (mean average precision), FPS (frames per second), Parameters and GFLOPs (Giga floating-point operations).

In this research paper, the backbone network weights that were trained on the VOC2007 dataset [16] are utilized as pre-training weights. During the training process, a batch size of 16 is adopted, and the training iterations are set to 200 rounds. The initial learning rate is established with a value of 0.01, and a cosine annealing learning rate decay strategy is employed. Moreover, to train the images effectively, adjustments are made based on the input sizes specified in Table 1. Furthermore, as part of the training process, the default number of boxes is set to 4, and their sizes are defined as [15, 21, 45, 99, 153, 207, 261, 315].

Table 1 Performance comparison of different detection models on the test set

Model	Input size	Backbone	mAP (%)	FPS (frame/s)	Params (M)	GFLOPs (G)
Faster RCNN [1]	600 × 600	VGG16	52.92	20.9	136.710	369.743
SSD [4]	300 × 300	VGG16	51.92	46.7	23.745	60.855
RFBNet [17]	300 × 300	–	53.36	43.8	39.693	72.930
RetinaNet [18]	600 × 600	ResNet50	32.87	23	37.969	170.091
M2Det [19]	320 × 320	VGG16	50.51	33.4	62.227	146.890
EfficientDet [20]	512 × 512	–	42.27	22.6	3.829	4.743
Our	300 × 300	MobileNetV2	56.97	50.6	11.051	5.958

4.1 Comparative Analysis of Different Algorithms

In order to compare the advantages and disadvantages between the models more clearly, the network proposed in this paper is compared with the classical two-stage network Faster RCNN [1], RFBNet [17] and RetinaNet [18] which are based on an improved SSD network, M2Det [19] which uses fusion of feature layers, and the light-weighted EfficientDet [20].

The results of each model's qualitative analysis is shown in Fig. 4. It includes target location, classification results and confidence level. It can be seen from the figure that apart from Faster RCNN [1] and the model presented in this paper, which have good detection abilities for small targets, other models fail to detect objects in most cases. Compares with Faster RCNN [1] which has the problem of false detection and overlapping detection frames in (b), the model proposed in this paper can detect most of the objects with accurate positions. Although some individual objects remain undetected as shown in (b).

The quantitative analysis of each model is shown in Table 1. It can be seen that the mAP of the model proposed outperforms Faster RCNN [1], SSD [4], RFBNet [17], RetinaNet [18], M2Det [19] and EfficientDet [20] by 4.05, 5.05, 3.61, 24.1, 6.46, and 14.7%. Although the improved model is not as good as EfficientDet [20] in terms of number of parameters and computation, according to the current market conditions where the frame rate of optoelectronic devices is between 25 and 30 Hz, EfficientDet is not good enough for real-time detection of the captured frames. Therefore, the proposed model provides higher accuracy and real-time performance on the embedded platform.

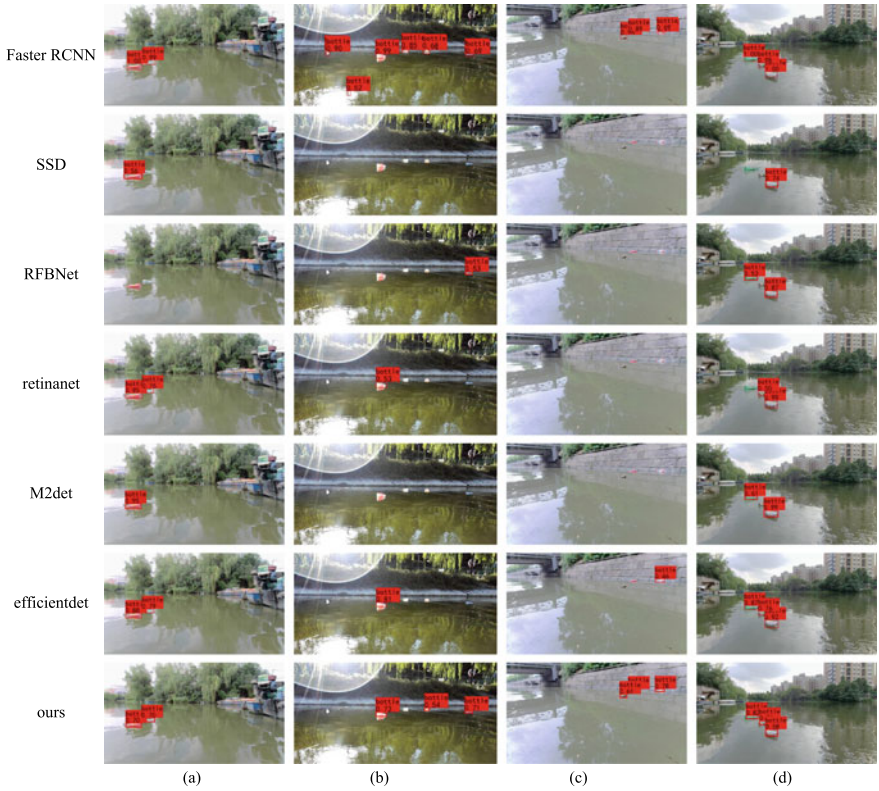


Fig. 4 Visualisation of test results for each model

4.2 Ablation Experiment

To further illustrate the effectiveness of the method in this paper, ablation experiments were conducted. The experimental results, as shown in Table 2, indicate that the introduction of the feature fusion network model improved detection accuracy by 10.6% at the expense of some detection time. Lightening the backbone network led to a 15.9 increase in FPS but decreased accuracy. Combining lightening and feature fusion of the SSD achieved a balanced trade-off between detection accuracy and speed, improving accuracy by 5.05% and FPS by 3.9. Moreover, this approach significantly reduced the number of model parameters and computational effort, which is beneficial for resource-limited devices.

Table 2 Ablation experiment

Backbone	Fusion	mAP (%)	FPS (frame/s)	Params (M)	GFLOPs (G)
		51.92	46.7	23.745	60.855
	✓	62.52	43.9	29.313	78.410
✓		49.47	62.6	5.610	1.477
✓	✓	56.97	50.6	11.051	5.958

5 Conclusion

This paper proposed a floater detection model for USV. Compared with the original SSD model, the improved SSD used a lightweight feature fusion strategy to improve the mAP by nearly 10% and the FPS by 4.1. In addition, the number of parameters and the computational effort of the backbone network are reduced by 53.5% and 90.2% respectively after the backbone lightweighting, which greatly reduces the parameter redundancy and improves the parameter efficiency. Although the proposed network has achieved a good balance between lightness and accuracy, the confidence level for objects with backgrounds such as water reflections and strong light reflections is still low, which is a direction for further improvement.

Acknowledgements This work was supported by China Postdoctoral Science Foundation (Grant No. 2019M661076).

References

1. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst 28
2. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) MobileNetV2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4510–4520
3. Zhang X, Zhou X, Lin M, Sun J (2018) ShuffleNet: an extremely efficient convolutional neural network for mobile devices. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 6848–6856
4. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) SSD: single shot multibox detector. In: Computer vision—ECCV 2016: 14th European conference, proceedings, part I 14, Amsterdam, The Netherlands, 11–14 Oct 2016. Springer, pp 21–37
5. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
6. Quan Y, Li Z, Zhang F, Zhang C (2019) D_dNet-65 R-CNN: object detection model fusing deep dilated convolutions and light-weight networks. In: PRICAI 2019: trends in artificial intelligence: 16th Pacific Rim international conference on artificial intelligence, proceedings, part III 16, Cuvu, Yanuca Island, Fiji, 26–30 Aug 2019. Springer, pp 16–28
7. Kanimozhi S, Gayathri G, Mala T (2019) Multiple real-time object identification using single shot multi-box detection. In: 2019 international conference on computational intelligence in data science (ICCIDS). IEEE, pp 1–5

8. Amudhan A, Sudheer A (2022) Lightweight and computationally faster hypermetropic convolutional neural network for small size object detection. Image Vis Comput 119:104396
9. Zhang J, Wang P, Zhao Z, Su F (2021) Pruned-YOLO: learning efficient object detector using model pruning. In: Artificial neural networks and machine learning—ICANN 2021: 30th international conference on artificial neural networks, proceedings, part IV, Bratislava, Slovakia, 14–17 Sept 2021. Springer, pp 34–45
10. Chu Y, Li P, Bai Y, Hu Z, Chen Y, Lu J (2022) Group channel pruning and spatial attention distilling for object detection. Appl Intell 1–19
11. Chen G, Choi W, Yu X, et al. (2017) Learning efficient object detection models with knowledge distillation. Adv Neural Inf Process Syst 30
12. Li Z, Xu X, Xie L, Su H (2019) Learning slimming SSD through pruning and knowledge distillation. In: 2019 Chinese automation congress (CAC). IEEE, pp 2701–2705
13. Fu CY, Liu W, Ranga A, Tyagi A, Berg AC (2017) DSSD: deconvolutional single shot detector. arXiv preprint arXiv:1701.06659
14. Zhou D, Hou Q, Chen Y, Feng J, Yan S (2020) Rethinking bottleneck structure for efficient mobile network design. In: Computer vision—ECCV 2020: 16th European conference, proceedings, part III 16, Glasgow, UK, 23–28 Aug 2020. Springer, pp 680–697
15. Cheng Y, Zhu J, Jiang M, Fu J, Pang C, Wang P, Sankaran K, Onabola O, Liu Y, Liu D et al (2021) Flow: a dataset and benchmark for floating waste detection in inland waters. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 10953–10962
16. Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A (2010) The PASCAL visual object classes (VOC) challenge. Int J Comput Vis 88:303–338
17. Deng L, Yang M, Li T, He Y, Wang C (2019) RFBNet: deep multimodal networks with residual fusion blocks for RGB-D semantic segmentation. arXiv preprint arXiv:1907.00135
18. Lin TY, Goyal P, Girshick R, He K, Dollár P (2017) Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision, pp 2980–2988
19. Zhao Q, Sheng T, Wang Y, Tang Z, Chen Y, Cai L, Ling H (2019) M2Det: a single-shot object detector based on multi-level feature pyramid network. In: Proceedings of the AAAI conference on artificial intelligence, vol 33, pp 9259–9266
20. Tan M, Pang R, Le QV (2020) EfficientDet: scalable and efficient object detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10781–10790

LOS/NLOS Detection in UWB Localization System with Anchors Deployment Along a Narrow Path

Wan-Ning He, Xin-Lin Huang, and Zhong-Wei Xu

Abstract Recently, the ultra-wideband (UWB) localization technology has been widely used in outdoor scenarios with high accuracy, fast refresh rate and non-sensitive to multipath fading. However, the deployment of UWB anchors along a narrow path has a large condition number of the observation matrix formed by the coordinates of UWB anchors and thus quite sensitive to the non-line-of-sight (NLOS) ranging error. In this paper, a LOS/NLOS detection algorithm is proposed to identify the NLOS from LOS propagation with UWB anchors deployment along a narrow path. Firstly, a ranging error correction model is constructed to estimate the NLOS ranging error under the blocking of single path between one anchor and the tag. Secondly, an iterative algorithm for ranging error estimation is proposed in the case of the multi-path blocked. Finally, a detection method based on the tetrahedron volume is combined to improve the detection accuracy. The simulation results show that the proposed algorithm outperforms the maximum likelihood method. When the average NLOS error is 20 cm, the probability of detection achieves an improvement of about 12% with 5 anchors, compared with other algorithms. When the NLOS error increases to 50 cm, the improvement is about 3% and the probability of detection reaches 90%.

Keywords UWB localization · LOS/NLOS detection · Ranging error estimation

1 Introduction

With the rapid development of automatic driving technologies in metro and high-speed train, the location information of the train plays an important role in safety driving [1]. Up to now, how to promote the accuracy of outdoor localization for safety driving is still an open hotspot.

UWB is a promising localization technology in outdoor scenarios and has a significant advantage in localization accuracy, such as 10–30 cm errors in some cases [2].

W.-N. He · X.-L. Huang (✉) · Z.-W. Xu
Department of Information and Communication Engineering, Tongji University, Shanghai, China
e-mail: xlhuang@tongji.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_25

As we all know, the typical methods of UWB localization mainly include angle-of-arrival (AOA), time-of-arrival (TOA), time-difference-of-arrival (TDOA) and their hybrids [3]. Since an antenna array is required to installing on each node for AOA algorithm, TOA and TDOA have low-price advantages in practice. However, the ranging results in TOA and TDOA can be affected by the non-line-of-sight propagation (NLOS) when the transmission path is blocked by metal or large obstacles and a slight ranging error will cause a large position error. In this paper, we focus on the LOS/NLOS detection with the TOA localization.

Regarding NLOS detection, its methods can be classified into two types: determining with the statistical characteristics of UWB pulses, utilizing ranging information and training the classification model via machine learning (ML). Among them, the energy statistics of the direct-path signal and its statistics of delay were analyzed in [4] to identify the channel. Then, Lin Liu utilized the ranging information sufficiently and incorporated the geometric constraints within maximum likelihood to detect NLOS [5]. Besides, in order to improve the detection accuracy, convolutional neural networks (CNN), k-nearest neighbors (KNN), support vector machine (SVM) and other ML methods were introduced to distinguish LOS from NLOS channels [6, 7]. Based on ML algorithm, the statistical distribution of Gaussian noise and generalized Gaussian noise were analyzed to improve the accuracy of NLOS channel identification [8].

Compared with detection by the pulse and ranging information, the classification model has a great advantage on detection accuracy at the cost of higher complexity and longer training time [9, 10]. Hence, this paper will propose an ML algorithm for NLOS detection with UWB anchors deployment along a narrow path, which makes full use of UWB ranging information and determines LOS/NLOS by a threshold. Based on the ranging information, we firstly construct an equation to estimate the NLOS ranging error for the single-path blocking, and in the case of the multi-path blocked, an iterative algorithm for multiple-ranging errors estimation is proposed. Then, an error-tetrahedron method is introduced and the volume of the tetrahedron is calculated. Finally, based on statistical method, the appropriate thresholds are set to determine whether the NLOS transmission is present or not. Through the proposed algorithm, the LOS/NLOS can be identified with high accuracy and low computational complexity.

2 Problem Statement

In the train localization system, anchors are deployed along a narrow path to realize wireless localization and its architecture is illustrated in Fig. 1. As one can see, the anchors alternatively communicate with the tag and their distances can be calculated by time of flight (TOF). Then, each anchor sends their ranging results to the primary anchor connected with a computer. The following part of this Section will take a brief review on the TOA algorithm and NLOS-related scenarios.

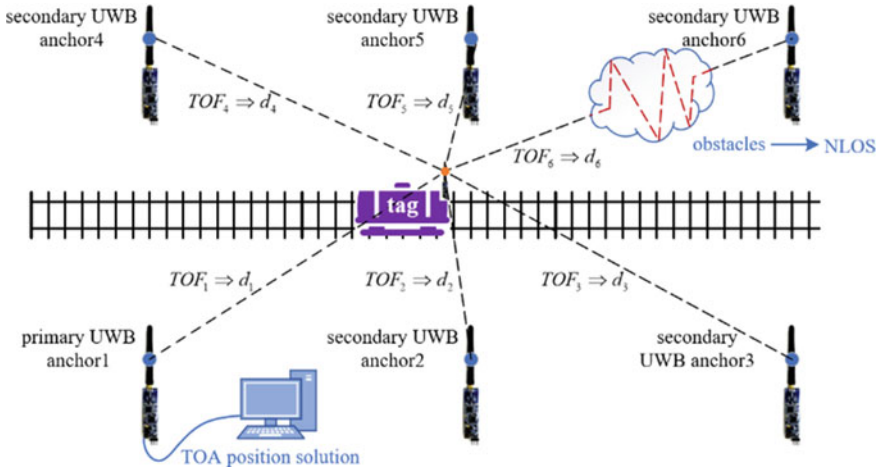


Fig. 1 The architecture of UWB localizing system

Assuming that the transmission rate of electromagnetic wave is c and the propagation time between the i-th anchor and the tag is t_i, the corresponding distance is

$$d_i = ct_i. \tag{1}$$

Considering both measurement noise and the NLOS range error, the ranging model in LOS/NLOS case can be represented as:

$$\tilde{r}_i = \begin{cases} r_i, & LOS \\ r_i + l_i, & NLOS \end{cases} \tag{2}$$

where

$$r_i = d_i + n_i. \tag{3}$$

Here, the errors caused by measurement noise and NLOS are represented by n_i and l_i, respectively. We assume $n_i \sim N(0, \sigma^2)$ and $l_i \sim N(\mu_i, \sigma_i^2)$ with the positive mean μ_i of the NLOS error.

Let $S_i = (x_i, y_i, z_i)^T$ be the coordinate of the i-th anchor and the coordinate of the tag is donated as $X = (x, y, z)^T$. In the Cartesian coordinate system, the distance between the tag and n anchors can be represented as:

$$\begin{cases} (x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2 = r_1^2 \\ (x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2 = r_2^2 \\ \vdots \\ (x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2 = r_n^2 \end{cases}. \tag{4}$$

Each equation minus the i-th row in Eq. (4), and we can get

$$A_i X = B_i \tag{5}$$

where the observation matrix and the constant vector are

$$A_1 = 2\begin{pmatrix} x_{12} & y_{12} & z_{12} \\ x_{13} & y_{13} & z_{13} \\ \vdots & \vdots & \vdots \\ x_{1n} & y_{1n} & z_{1n} \end{pmatrix} or A_i = 2\begin{pmatrix} x_{i1} & y_{i1} & z_{i1} \\ x_{i2} & y_{i2} & z_{i2} \\ \vdots & \vdots & \vdots \\ x_{in} & y_{in} & z_{in} \end{pmatrix} (i \neq 1) \tag{6}$$

$$B_1 = \begin{pmatrix} r_2^2 - r_1^2 - k_2 + k_1 \\ r_3^2 - r_1^2 - k_3 + k_1 \\ \vdots \\ r_n^2 - r_1^2 - k_n + k_1 \end{pmatrix} or B_i = \begin{pmatrix} r_1^2 - r_i^2 - k_1 + k_i \\ r_2^2 - r_i^2 - k_2 + k_i \\ \vdots \\ r_n^2 - r_i^2 - k_n + k_i \end{pmatrix} (i \neq 1). \tag{7}$$

Here, we donate $x_{mn} = x_m - x_n$, $y_{mn} = y_m - y_n$, $z_{mn} = z_m - z_n$, $(m \neq n)$ and $k_m = x_m^2 + y_m^2 + z_m^2$. Then, Eq. (5) can be solved by least square (LS) method as:

$$X = A_i^+ B_i \tag{8}$$

where donate A_i^+ as the pseudo inverse of A_i, i.e., $A_i^+ = \left(A_i^T A_i\right)^{-1} A_i^T$. When the anchors are deployed along a narrow path, the condition number of matrix A_i will be large and a slight ranging error will cause a significant interference on localization precision. Hence, in this paper, we focus on LOS/NLOS detection in the case of deploying anchors along a narrow path.

3 LOS/NLOS Detection Algorithm

In this Section, we will propose a LOS/NLOS detection algorithm. In the case of NLOS blocking of single path between one anchor and the tag, the ranging equation with NLOS error will be established to solve error estimations. When NLOS exists between multiple anchors and the tag, an iterative method is introduced to estimate each ranging error alternatively. At last, in the absence of a proper solution for the

ranging equation [see Eq. (4)], the error-tetrahedron method is used to identify LOS/NLOS.

3.1 Single Path LOS/NLOS Detection

As we can see from Eq. (7), the ranging results (i.e., r_i) in the equation appears in the form of square, so the error in vector B_i caused by the NLOS ranging error can be represented as:

$$\varepsilon_i = \tilde{r}_i^2 - r_i^2 = 2l_i r_i + l_i^2. \tag{9}$$

which is determined by the NLOS error l_i and the LOS distance r_i.

Assume that the NLOS occlusion occurs in the i-th path. The localization equation of Eq. (5) can be rewritten as:

$$A_i \tilde{X} = \begin{pmatrix} r_1^2 - r_i^2 - k_1 + k_i - \varepsilon_i \\ r_2^2 - r_i^2 - k_2 + k_i - \varepsilon_i \\ \vdots \\ r_n^2 - r_i^2 - k_n + k_i - \varepsilon_i \end{pmatrix} = B_i - \varepsilon_i \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{10}$$

where the estimated coordinate of the tag under the NLOS scenario is denoted as $\tilde{X}$ and the LS solution is solved as:

$$\tilde{X} = X - \varepsilon_i e_i. \tag{11}$$

Here, we denote the vector $e_i = A_i^+(1, 1, \ldots, 1)^T$ to simplify the solution. Then, the range distance between the estimated results in Eq. (11) and the location of the i-th anchor can be calculated as:

$$r_{i\_cal}^2 = \left(\tilde{X} - S_i\right)^T \left(\tilde{X} - S_i\right). \tag{12}$$

Substituting Eq. (11) into Eq. (12), we can obtain

$$r_{i\_cal}^2 = \tilde{r}_i^2 - \varepsilon_i + \varepsilon_i^2 e_i^T e_i - 2\varepsilon_i e_i^T \left(\tilde{X} - S_i + \varepsilon_i e_i\right). \tag{13}$$

Hence, the error in vector B_i caused by the NLOS ranging error can be calculated as:

$$\varepsilon_i = \frac{-\left[1 + 2e_i^T\left(\tilde{X} - S_i\right)\right] \pm \sqrt{\left[1 + 2e_i^T\left(\tilde{X} - S_i\right)\right]^2 - 4e_i^T e_i\left(r_{i\_cal}^2 - \tilde{r}_i^2\right)}}{2e_i^T e_i}. \quad (14)$$

Then, the NLOS ranging error can be estimated as:

$$\hat{l}_i = \tilde{r}_i - \sqrt{\tilde{r}_i^2 - \varepsilon_i}. \quad (15)$$

During the process of NLOS detection, each path is assumed as blocked alternatively and the parameter ε_i and the corresponding $\hat{l}_i$ are calculated. Comparing with a threshold Th (i.e., 5 cm), the NLOS path can be determined if

$$\max\{\hat{l}_i\} \geq Th. \quad (16)$$

3.2 Multipath LOS/NLOS Detection

For multiple NLOS paths, we propose an iterative method to continuously estimate the error elements in vector B_i. The localization equation in Eq. (10) can be rewritten as:

$$A_i\tilde{X} = \begin{pmatrix} r_1^2 - r_i^2 - k_1 + k_i - \varepsilon_i + \varepsilon_1 \\ r_2^2 - r_i^2 - k_2 + k_i - \varepsilon_i + \varepsilon_2 \\ \vdots \\ r_n^2 - r_i^2 - k_n + k_i - \varepsilon_i + \varepsilon_n \end{pmatrix} = B_i + \varphi_i \quad (17)$$

where the form of the iterative vector corresponds to matrix A_i as:

$$\varphi_1 = \begin{pmatrix} \varepsilon_2 - \varepsilon_1 \\ \varepsilon_3 - \varepsilon_1 \\ \vdots \\ \varepsilon_n - \varepsilon_1 \end{pmatrix} or\ \varphi_i = \begin{pmatrix} \varepsilon_1 - \varepsilon_i \\ \varepsilon_2 - \varepsilon_i \\ \vdots \\ \varepsilon_n - \varepsilon_i \end{pmatrix}, (i \neq 1). \quad (18)$$

Then we can update the error elements in vector B_i caused by the NLOS ranging error as:

$$\varepsilon_i = \tilde{r}_i^2 - r_{i\_cal}^2 - \left(A_i^+\varphi_i\right)^T\left(A_i^+\varphi_i\right) + 2\left(A_i^+\varphi_i\right)^T\left(\tilde{X} - S_i\right) \quad (19)$$

until their variations are less than a threshold $tol > 0$. Besides, an upper bound of times is also adopted to terminate the iteration. After the iterative method stopped, the parameter ε_i and the corresponding $\hat{l}_i$ are calculated finally.

3.3 LOS/NLOS Detection Based on Tetrahedron Volume

Considering the case of no solution available from Eq. (13), we propose an error-tetrahedron volume method to determine whether the wireless channel is blocked or not. We assume four anchors available as an example, and select three anchors of them alternatively to solve out the coordinates of the tag [see Eq. (8)], expressed as $X_{Si} = (x_{Si}, y_{Si}, z_{Si})^T$. We connect the four solutions in Cartesian space and obtain a tetrahedron $X_{S1} - X_{S2}X_{S3}X_{S4}$ as shown in Fig. 2.

Generally, in the LOS environment, the four coordinates in Fig. 2, are closer while they are far away from each other in the NLOS environment. As a result, the volume of the tetrahedron can be set as a key parameter. With vertex coordinates, the tetrahedron volume can be calculated by

$$v_{X_{S1}X_{S2}X_{S3}X_{S4}} = \frac{1}{6}\begin{vmatrix} x_{S1S2} & y_{S1S2} & z_{S1S2} \\ x_{S1S3} & y_{S1S3} & z_{S1S3} \\ x_{S1S4} & y_{S1S4} & z_{S1S4} \end{vmatrix} \tag{20}$$

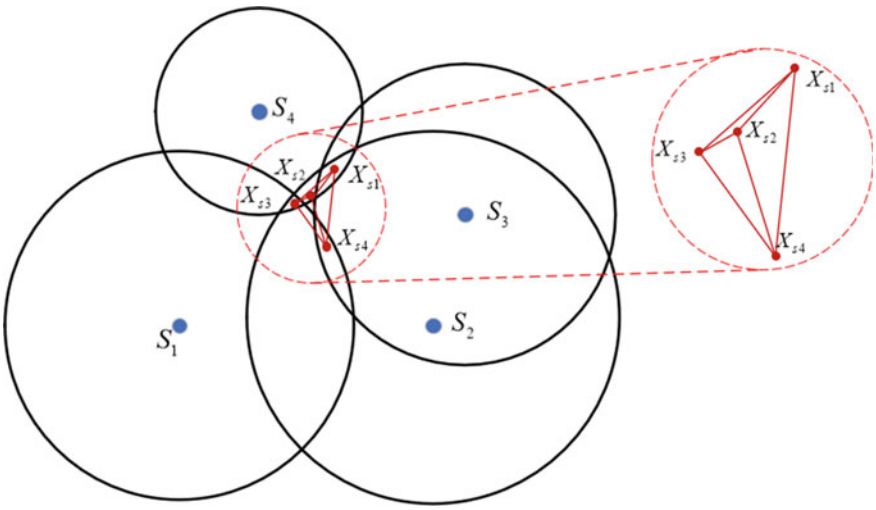


Fig. 2 The diagram of the error tetrahedron

where denote $x_{SmSn} = x_{Sm} - x_{Sn}$, $y_{SmSn} = y_{Sm} - y_{Sn}$, $z_{SmSn} = z_{Sm} - z_{Sn}$, $(m \neq n)$. As a result, when there is no solution available in Eq. (13), LOS/NLOS can be detected by comparing Eq. (20) with a threshold.

4 Simulation Results

In this Section, we evaluate the performance of LOS/NLOS detection in term of the probability of detection and compare the proposed algorithm with the maximum likelihood detection method [5]. Assume the anchors and the tag are placed in a narrow path region with size 5 m × 800 m × 1 m to simulate the train localization along the track. Generally, the probability of detection is influenced by the mean value of the NLOS error and the number of anchors. Therefore, in this Section, we analyze such two factors' impact on the performance of the proposed algorithm. Assume the average NLOS error ranges from 0.1 to 1 m with an interval of 0.1 m and the variances of ranging error and NLOS error are both $4cm^2$, the simulation results of the probability of detection under 5–8 anchors localization are shown in the Fig. 3.

As we can see, the proposed algorithm always outperforms the maximum likelihood detection method. When the mean ranging error caused by NLOS is 0.1 m, compared with the maximum likelihood method, the proposed algorithm achieves an improvement of 18.17% with 5 anchors. When the average NLOS error is 0.7 m, the improvement is 1.83%. Besides, as the number of anchors increases, the probability of detection of the proposed algorithm is also improved. With the average ranging error as 0.5m, the probabilities of detection are 89.14%, 89.58%, 92.38% and 92.81% under localization with 5 to 8 anchors, respectively, and the corresponding improvements are 2.99%, 2.55%, 2.49% and 2.09%. Overall, the improvement of the proposed algorithm is more significant with fewer anchors and smaller NLOS error.

5 Conclusion

In this paper, an algorithm of LOS/NLOS detection has been proposed to apply to anchors deployment along a narrow path, such as lanes, tracks and so on. During the LOS/NLOS detection, we have firstly introduced a ranging equation with NLOS errors and identified NLOS under a single path blocked. Then, an iterative method has been introduced to update each ranging error when NLOS exists between multiple anchors and the tag. At last, when the ranging error equation has no proper solution, an error-tetrahedron method has been utilized and the LOS/NLOS propagation has been identified based on a volume threshold. Through the proposed algorithm, the range information and ML method have been well combined and the NLOS could be detected with high accuracy and low complexity. The simulation results have shown

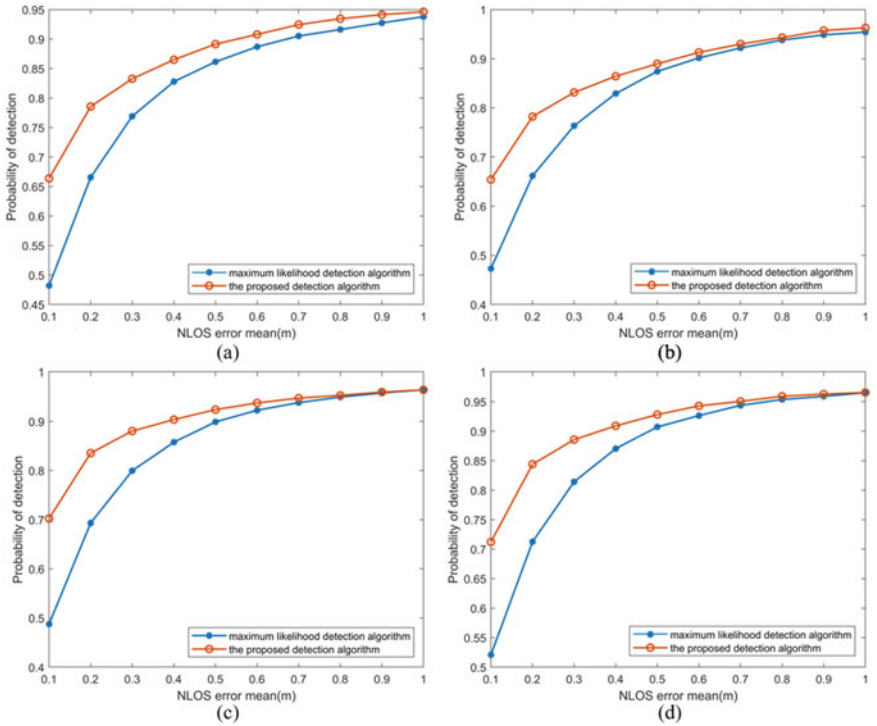


Fig. 3 The probability of detection under various NLOS error mean and the number of anchors. **a** 5 anchors, **b** 6 anchors, **c** 7 anchors, **d** 8 anchors

that the proposed algorithm could achieve a better performance than the maximum likelihood detection method, especially with fewer anchors and smaller NLOS error.

Acknowledgements This work was supported by the National Key R&D Program of China under Grant 2022YFB4300504.

References

1. Akai N, Akagi Y, Hirayama T, Morikawa T, Murase H (2022) Detection of localization failures using Markov random fields with fully connected latent variables for safe LiDAR-based automated driving. IEEE Trans Intell Transp Syst 23(10):17130–17142
2. Güler S, Abdelkader M, Shamma JS (2021) Peer-to-peer relative localization of aerial robots with ultrawideband sensors. IEEE Trans Control Syst Technol 29(5):1981–1996
3. Su Z, Shao G, Liu H (2018) Semidefinite programming for NLOS error mitigation in TDOA localization. IEEE Commun Lett 22(7):1430–1433
4. Maali A, Mimoun H, Baudoin G, Ouldali A (2009) A new low complexity NLOS identification approach based on UWB energy detection. In: 2009 IEEE radio and wireless symposium. IEEE Press, San Diego, pp 675–678

5. Liu L, Fan P (2010) An efficient geometry-constrained NLOS mitigation algorithm based on ML-detection. In: IET 3rd international conference on wireless, mobile and multimedia networks (ICWMNN 2010). IET, Beijing, pp 348–352
6. Wang J, Yu K, Bu J, Lin Y, Han S (2022) Multi-classification of UWB signal propagation channels based on one-dimensional wavelet packet analysis and CNN. IEEE Trans Veh Technol 71(8):8534–8547
7. Zandian R, Witkowski U (2018) Differential NLOS error detection in UWB-based localization systems using logistic regression. In: 2018 15th workshop on positioning, navigation and communications (WPNC). IEEE Press, Bremen, pp 1–6
8. Che F et al (2022) Feature-based generalized Gaussian distribution method for NLOS detection in ultra-wideband (UWB) indoor positioning system. IEEE Sens J 22(19):18726–18739
9. Stahlke M, Kram S, Mutschler C, Mahr T (2020) NLOS detection using UWB channel impulse responses and convolutional neural networks. In: 2020 international conference on localization and GNSS (ICL-GNSS). IEEE Press, Tampere, pp 1–6
10. Zhao Z, Huang W, Liang Y, Zhang Y (2020) A NLOS detection method based on machine learning in UWB indoor location system. In: 2020 international conference on virtual reality and intelligent systems (ICVRIS). IEEE Press, Zhangjiajie, pp 214–217

Stackelberg Game Based Adaptive Value Evaluation Strategy

Chen-Ye Huang, Yi-Zhan Chen, and Xin-Lin Huang

Abstract In the process of recycling and reusing household appliances, implementing extended producer responsibility (EPR) has become increasingly important. Therefore it is necessary to design a value evaluation model to solve optimal pricing problems in closed-loop supply chain to drive producers to fulfil EPR. Based on Stackelberg Game and various recycle chain members, this paper provides two kinds of value evaluation models with the same structure but different parameters. Compared to three basic models, these two models are more concrete and practical. Besides, the profit of recycle chain in the first model is relatively high but the recycle rate is fixed at 12.5%. On the contrary, the profit of recycle chain in the second model is lower but the recycle rate can vary freely with the parameters and its total profit is a little higher. So these two models can be applied according to different situations.

Keywords Value evaluation · Extended producer responsibility (EPR) · Stackelberg game · Waste home appliance recycling

1 Introduction

In recent years, with the continuous development of global society and economy, the life standard of residents is constantly improving. Correspondingly, the electrical and electronic products industry is also expanding at a high speed and the products are constantly updated and iterated, which results in the greatly shortened service life of electronic products. These products have become waste electrical and electronic equipment prematurely and the quantity is huge, causing serious pollution and waste to the environment and resources. Furthermore, there are still problems existing in the recycling industry such as poor operation of recycling system and low reusing rate of products [1]. Therefore, some measures need to be executed to curb this situation.

Extended producer responsibility (EPR) is an effective means to recycle waste electronic products. It emphasizes that producers should not only consider the impact

C.-Y. Huang · Y.-Z. Chen · X.-L. Huang (✉)
Department of Information and Communication Engineering, Tongji University, Shanghai, China
e-mail: xlhuang@tongji.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_26

on the environment caused by the production process in the forward supply chain, but also assume the corresponding responsibility for the recycle process in the reverse supply chain [2]. The implementation of EPR also helps to improve the profits of supply chain members, enhance the corporate image and promote sales volume [3]. Moreover, EPR also encourages producers to incorporate environmental considerations into the design of their products [4].

Based on EPR, this paper will study the optimal sales and recycle price of each member in the closed-loop supply chain and solve the conflict among each member in the system, which will optimize the benefit and profit distribution of the closed-loop supply chain and minimize the extra cost of producers.

2 Research Method

2.1 Stackelberg Game

The pricing model of this paper is based on Stackelberg Game, a dynamic game model with complete information which reflects the asymmetric competition between enterprises [5]. In this model, there is a decision order among the members and these members make continuous decisions. The party that makes the decision first is called the "leader" and the party that makes the decision later is called the "follower". In the decision-making process, the leader will revise its decision according to the follower's possible decision and the follower will also consider its decision in the same way. These two sides will repeatedly consider their final decisions according to this logic until they reach Nash equilibrium.

2.2 Backward Induction

Stackelberg game model is usually solved by backward induction: In dynamic games like Stackelberg game, the leader should consider the decision chosen by the followers when making its decision. Only the last follower will not be affected by subsequent decisions and can directly make the optimal decision. Therefore, when the decision of last follower is determined, the game is solved. That is to say, it is assumed that all the possible decisions of each member in the game are known to all the others. Then the follower who makes the final decision starts to analyse all the possible decisions of its leaders and obtains the optimal expected decision based on these decisions. Next, according to the above logic, all the optimal decisions can be determined, so that the Nash equilibrium of the whole game can be obtained. When solving a game, we can also draw a relation tree to understand game relation and decision order more intuitively. It is important to note that in the whole game, all the members should have symmetric and complete information and they must fully understand the inter-

Fig. 1 Flow chart of backward induction

est function of each member and the order structure of the game. In addition, all members should be rational and they must know that other members are also rational, which is the prerequisite and ideal condition to solve the game. The process of backward induction is shown in Fig. 1.

3 Model Building

3.1 Assumptions and Parameters

This model sets three basic subjects in the closed-loop supply chain, namely producers, sellers and consumers and it also refines the recycle chain. Recycle chain in the general model is broken down into individual and enterprise recyclers, disintegrators and reproducers, among which the enterprise recyclers need to sell the recycled waste products to the disintegrators for further dismantling into usable materials. And the disintegrators will sell them to the reproducers. Individual recyclers, on the other hand, combine recycling and dismantling in one process, selling materials directly to the reproducers.

This model eliminates the cost of miscellaneous affairs such as production, operation and transportation of whole process, which is incorporated into the purchase cost and selling price, because these costs are not affected by the price optimization and adjustment strategy. In addition, it is necessary to assume that the market's supply and demand are completely balanced, which means that the quantity of products sold in the closed-loop supply chain is consistent with the actual demand of the home appliance market. The parameters are set as shown in Table 1, in which β_{rp} means sensitivity coefficient of reproducer to price and β_r means sensitivity coefficient of consumer to recycle price. These two parameters will divide the model into two different cases for calculation and discussion.

3.2 Model Building

In the supply chain of this paper, the producer is the leader of Stackelberg game, while the seller and recycler are the followers. The order of the game is: the producer decides the wholesale price and recycle price, then the seller decides the retail price. The model legend is shown in Fig. 2.

Table 1 Symbols and meanings

Symbols	Meaning	Symbols	Meaning
x	Cost of new means of production	Q	Quantity of waste products
w	Producer's price to seller	λ	Recycle rate of waste products
p	Seller's price to consumers	β_{rp}	Sensitivity of reproducer to price
y	Individual recycler's price to consumers	β_r	Sensitivity of consumer to recycle price
b	Enterprise recycler's price to consumers	π_m	Producer's profit
z	Disintegrator's price to reproducer	π_{er}	Enterprise recycler's profit
r	Enterprise recycler's price to disintegrator	π_{ir}	Individual recycler's profit
e	Individual recycler's price to reproducer	π_s	Seller's profit
c	Reproducer's price to producer	π_d	Disintegrator's profit
a	Demand for home appliances	π_r	Reproducer's profit
β_s	Consumer sensitivity to sales price	π	Total profit

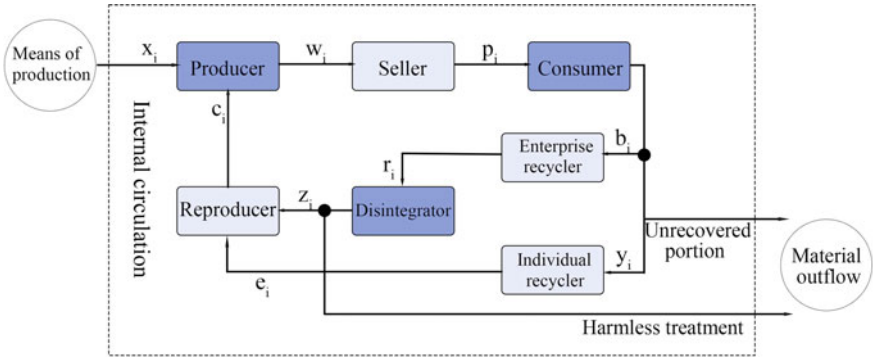


Fig. 2 Material circulation in the manufacturing and recycling process of household appliances

In the first model involving β_{rp}, both the enterprise and individual recyclers decide the recycle price to the consumers and the price to the disintegrator and reproducer based on market condition. The disintegrator will also determine the price to the reproducer. In fact, this model can be roughly regarded as two Stackelberg games with a similar structure. One is the game in supply chain and the other is the game in recycle chain. This separation is due to the fact that there is no clear direct quantitative relationship between the quantity of waste products and the sales status of products in the current market. Therefore, there is no game relationship between the consumer and recycler. The intersection point of the two chains is between the reproducer and the producer because there are transactions between them. But the price is determined

by the producer, so the game relationship still does not exist. Therefore, this model needs to analyse and solve two simple Stackelberg games.

In the second model involving β_r, each member of recycle chain decides its purchase price rather than its selling price, so there is no game existing in the recycle chain. Therefore, the optimal price can be calculated in turn according to the decision order. While the game in supply chain can be solved according to the method of the first model.

In order to make the model conform to the actual situation, the result of each subtraction operation in the profit function of the two models must be greater than zero.

4 Calculation

4.1 The First Model

The profit functions of members in the first model are shown in the following formulas. All of the profit functions are greater than 0 in order to be realistic. The reproducer's pricing should be less than the cost of new means of production to enable the producer to fulfill EPR.

$$\begin{aligned}
&\pi_m = \left\{a - \beta_s p - \left[Q - \beta_{rp}(z+e)\right]\right\}(w-x) + \left[Q - \beta_{rp}(z+e)\right](w-c)\\
&\pi_{er} = \left[Q - \beta_{rp}(z+e)\right]\frac{rb-b^2}{b+y},\ \pi_{ir} = \left[Q - \beta_{rp}(z+e)\right]\frac{ey-y^2}{b+y}\\
&\pi_s = (a-\beta_s p)(p-w),\ \pi_d = \left[Q - \beta_{rp}(z+e)\right]\frac{b(z-r)}{b+y}\\
&\pi_r = \left[Q - \beta_{rp}(z+e)\right]\left(c - \frac{bz}{b+y} - \frac{ye}{b+y}\right),\ \lambda = \frac{Q-\beta_{rp}(z+e)}{Q}\\
&z < c \le x
\end{aligned} \tag{1}$$

Solving by backward induction can get the result as follows:

$$y = b = \frac{Q}{8\beta_{rp}}, e = r = \frac{3Q}{8\beta_{rp}}, z = \frac{Q}{2\beta_{rp}}, w = \frac{a}{2\beta_s} + \frac{x}{2}, p = \frac{3a}{4\beta_s} + \frac{x}{4} \tag{2}$$

$$\begin{aligned}
&\pi_m = \frac{a^2}{8\beta_s} - \frac{ax}{4} + \frac{\beta_s x^2 + Qx - Qc}{8},\ \pi_{er} = \pi_{ir} = \frac{Q^2}{64\beta_{rp}}\\
&\pi_s = \frac{(a-\beta_s x)^2}{16\beta_s},\ \pi_d = \frac{Q^2}{128\beta_{rp}}\\
&\pi_r = \frac{Q}{8}c - \frac{7Q^2}{128\beta_{rp}} > \frac{Q^2}{128\beta_{rp}},\ \lambda = 12.5\%
\end{aligned} \tag{3}$$

Fig. 3 Seller recycling mode block diagram

4.2 The Second Model

The profit functions of members in the second model are shown in the following formulas. Unlike the first model, there is an optimal solution for the pricing of all members.

$$\begin{aligned}
&\pi_m = [a - \beta_s p - \beta_r (b + y)]\,(w - x) + \beta_r (b + y)(w - c) \\
&\pi_{er} = \beta_r b(r - b),\ \pi_{ir} = \beta_r y(e - y) \\
&\pi_s = (a - \beta_s p)\,(p - w),\ \pi_d = \beta_r b(z - r) \\
&\pi_r = \beta_r (b + y)c - \beta_r bz - \beta_r ye,\ \lambda = \frac{\beta_r (b + y)}{Q}
\end{aligned} \tag{4}$$

Solving by backward induction can get the result as follows:

$$b = \frac{x}{16}, y = r = \frac{x}{8}, z = e = \frac{x}{4}, c = \frac{x}{2}, p = \frac{3a + \beta_s x}{4\beta_s}, w = \frac{a + \beta_s x}{2\beta_s} \tag{5}$$

$$\begin{aligned}
&\pi_m = \frac{(a - \beta_s x)^2}{8\beta_s} + \frac{3\beta_r x^2}{32},\ \pi_{er} = \frac{\beta_r x^2}{256},\ \pi_{ir} = \frac{\beta_r x^2}{64} \\
&\pi_s = \frac{(a - \beta_s x)^2}{16\beta_s},\ \pi_d = \frac{\beta_r x^2}{128},\ \pi_r = \frac{3\beta_r x^2}{64},\ \lambda = \frac{3\beta_r x}{16Q}
\end{aligned} \tag{6}$$

5 Example Analysis

5.1 Three Basic Recycle Models

The block diagram of the seller recycling model and the optimal pricing and profit functions of its members are shown in Fig. 3. This is a model where the seller participates in recycling.

$$w = \frac{a + \beta_s x}{2\beta_s}, c = \frac{x}{2}, b = \frac{x}{4}, p = \frac{3a + \beta_s x}{4\beta_s} \tag{7}$$

$$\pi_m = \frac{(a + \beta_s x)^2}{8\beta_s} + \frac{\beta_r x^2}{8} - \frac{ax}{2},\ \pi_s = \frac{(a - \beta_s x)^2}{16\beta_s} + \frac{\beta_r x^2}{16} \tag{8}$$

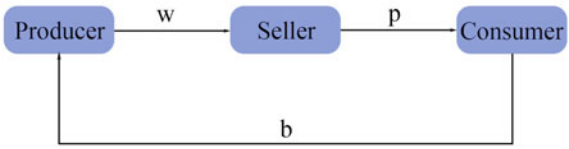


Fig. 4 Producer recycling mode block diagram

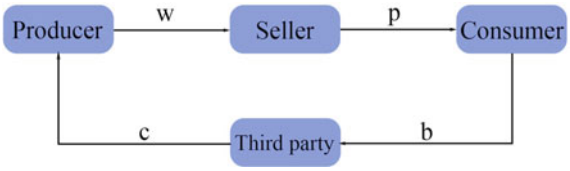


Fig. 5 Third party recycling model

The block diagram of the producer recycling model and the optimal pricing and profit functions of its members are shown in Fig. 4. In this model, only the producer is responsible for recycling.

$$w = \frac{a + \beta_s x}{2\beta_s}, b = \frac{x}{2}, p = \frac{3a + \beta_s x}{4\beta_s} \tag{9}$$

$$\pi_m = \frac{(a + \beta_s x)^2}{8\beta_s} + \frac{\beta_r x^2}{4} - \frac{ax}{2}, \ \pi_s = \frac{(a - \beta_s x)^2}{16\beta_s} \tag{10}$$

The block diagram of the third party recycling model and the optimal pricing and profit functions of its members are shown in Fig. 5. There is a third party responsible for recycling in this model.

$$w = \frac{a + \beta_s x}{2\beta_s}, c = \frac{x}{2}, b = \frac{x}{4}, p = \frac{3a + \beta_s x}{4\beta_s} \tag{11}$$

$$\pi_m = \frac{(a - \beta_s x)^2 + \beta_s \beta_r x^2}{8\beta_s}, \ \pi_s = \frac{(a - \beta_s x)^2}{16\beta_s}, \ \pi_r = \frac{\beta_r x^2}{16} \tag{12}$$

5.2 Examples

Example of the models of this paper.
Parameter setting: $x = 10, a = 300, \beta_s = 15, \beta_{rp} = 20 (\beta_r = \frac{50}{3}), c = 8, Q = 250$.

It can be seen from Table 2 that when the recycle rate is the same, total profit of the second model is slightly higher than that of the first model and the profit of the producer is much higher. While the profit of all members in the recycle chain is lower. Therefore, it can be inferred that the game of the recycle chain increases its members' profit, but decreases the total profit of whole closed-loop supply chain.

Table 2 Example of two models in this paper

	π_m	π_s	π_{er}	π_{ir}	π_d	π_r	π	λ (%)
Model 1	250	93.75	48.83	48.83	24.41	79.10	544.92	12.5
Model 2	343.75	93.75	6.51	26.04	13.02	78.13	561.20	12.5

Table 3 Comparison of profit and recycle rate of all parties

	Seller recycling model	Producer recycling model	Third party recycling model	Model 2 in this paper
π_m	437.5	687.5	437.5	375
π_s	218.75	93.75	93.75	93.75
π_r	–	–	125	93.75
π	656.25	781.25	656.25	617.19
λ (%)	20	40	20	15

Example of three basic models and the second model in this paper.
Parameter setting: $x = 10$, $a = 300$, $\beta_s = 15$, $\beta_r = 20$, $Q = 250$.

As can be seen from the results in Table 3, when the parameters of models are the same, the total profit and recycle rate of the optimal pricing strategy will decrease with the increase of the number of members of the recycle chain. This is because with the increase of members, the price of recycling waste products from consumers will gradually decrease in order to ensure the profits of each member. Moreover, in all models, producers account for the largest proportion of total profit, which is their advantage as the game leader and also improves their motivation to reproduce and fulfill EPR.

6 Conclusion

1. The first model of the paper is applicable to the situation where there is two games existing in both the recycle chain and the supply chain. The second model is applicable to the case that there is no game existing in the recycle chain.
2. The total profit of the second model is slightly higher than that of the first. But the profit of members in the recycle chain is quite low, which may lead to the decline of recycle efficiency. This shows that game can make the recycle chain gain more profits.
3. In order to maximize the profit of closed-loop supply chain and the producer, the number of members in the recycling chain should be reduced as much as possible to ensure that the waste products can be recycled from consumers at the highest possible price.

Acknowledgements This work was supported by the National Key R&D Program of China under Grant 2022YFB3305801.

References

1. Han H, Fan X, Zhang Q, Du Y (2022) Situation and development trend for recycling of electrical and electronic equipment. Comput Integr Manuf Syst 28(7)
2. Jia J (2020) Study on recycling mode of waste electrical appliances and electronic products under extended producer responsibility system. Taiyuan University of Technology
3. Zhou J, Tao X (2016) Research on revenue-sharing contract of sales-recycling closed-loop supply chain with EPR. Sci Decis Mak 02:39–57
4. Subramanian R, Gupta S, Talbot B (2009) Product design and supply chain coordination under extended producer responsibility. Prod Oper Manag 18(3):259–277
5. Zhou L, Ma Y, Ki F (2021) Research on dual-model recycling pricing of waste household appliances based on Stackelberg game. In: 2021 international conference on E-commerce and E-management (ICECEM), pp 170–173

Surface Defect Detection Method of Strip Steel Based on Improved YOLOv5

Bin Wang, Ronaldo Juanatas, Jasmin Niguidula, and Hainan Luo

Abstract A strip steel surface defect detection method called YOLOv5-ABS is developed in order to address the issues of low detection accuracy, insufficient feature extraction ability, and insufficient feature fusion of YOLOv5. Firstly, in order to enhance the backbone network's capacity for feature extraction, the C3 module is swapped out for the SeC3 module with an attention mechanism. Secondly, in order to improve the network feature fusion capability, the bidirectional weighted feature pyramid network BiFPN is added in the Neck section. Finally, by introducing the SPPFCSPC spatial pyramid pooling structure, speed and accuracy are improved while keeping the receptive field unchanged. According to the experimental findings, the revised YOLOv5-ABS algorithm's mAP on the NEU-DET dataset is 78.6%, 3.8% larger compared to the initial YOLOv5s algorithm, and the detection speed is 142.8 FPS, enabling the quick and precise identification of strip steel surface defects.

Keywords Detected defects · Attention · YOLOv5 · BiFPN · SPPFCSPC

1 Introduction

Strip steel is an important industrial raw material in the field of mechanical manufacturing. Due to the influence of external factors such as the production process, raw material quality, and processing equipment, the surface of the product may produce various types of defects, such as rolled-in scale, scratches, and patches, during the strip steel production process. These flaws will not only degrade the product's look but will also reduce its steel strength, corrosion resistance, and wear resistance, and

B. Wang (✉) · R. Juanatas · J. Niguidula
Technological University of the Philippines, Manila, Philippines
e-mail: wangbin@ahcme.edu.cn

B. Wang
Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241002, Anhui, China

H. Luo
WuHu Hit Robot Technology Research Institute, Wuhu 241002, Anhui, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_27

may even result in a safety mishap. As a result, one of the hottest topics in the world of mechanical manufacturing is how to quickly and precisely accomplish the identification of strip surface flaws.

Surface defect detection methods based on machine vision have received a lot of attention as machine vision technology has advanced. A typical machine vision defect detection system is made up of two parts: hardware and software. The hardware system mainly refers to the image acquisition device composed of an industrial camera, an industrial lens, a light source, and other equipment, which mainly completes the image acquisition work; the software system first performs noise reduction, segmentation, and morphological processing on the collected images, and then completes recognition and classification according to the extracted image feature information. Based on the shape template matching algorithm, some researchers have proposed an automatic detection method for shape defects of stamped parts, but this method is greatly affected by the light environment and is sensitive to noise [1]. Liu uses bilateral filtering and the Hilditch algorithm to improve the traditional Sobel operator and proposes a new edge detection algorithm for strip steel surface defect images. However, this method is not accurate for edge positioning, and the extracted edge lines are thicker [2]. A defect identification approach for solar cells based on SVM has been presented by some academics to address a number of common solar cell flaws, but it is challenging to train this system on huge amounts of data [3].

In the area of machine vision, deep learning has advanced quickly in recent years. Deep learning-based surface defect detection techniques have recently attracted a lot of attention in the target detection field. The deep learning-based approach may effectively address the complexity and unpredictability of manual feature extraction in conventional machine vision by automatically learning and extracting the input data's features [4]. To increase the effectiveness and accuracy of defect identification, academics domestically and internationally have applied deep learning to the detection of surface defects in a variety of products. A surface defect detection algorithm for a metal workpiece based on improved Faster RCNN is proposed by introducing multi-level ROI pooling layer structure and a bilinear interpolation method. This algorithm aims to address the issues of low accuracy and low speed of surface defect detection of metal materials [5]. Based on YOLOv3, Kou uses Anchor-Free to improve the speed of the model and designs dense blocks to extract richer feature information. The accuracy and robustness of the model have been improved [6]. Aiming at the problem of low accuracy of metal surface scratch detection, a method of metal surface scratch detection based on YOLOv4 is designed by introducing the small target detection idea, data enhancement, and adjusting anchor frame size [7]. Aiming at the problem of low detection efficiency of different scale defects on metal surfaces, an improved YOLOv5 detection network is designed. An adaptive anchor frame method is proposed, and a feature layer is added to the main components to enhance the useful feature information. In the prediction part, an effective loss function is used to solve the problem of data imbalance caused by small targets [8]. At present, there are many methods applied to the detection of strip steel surface defects, but these methods have problems such as low detection accuracy, low detection speed, and large model calculations.

In order to further improve the accuracy of strip steel surface defect detection, this study improves the YOLOv5s model and proposes the YOLOv5-ABS, which integrates an attention mechanism and a bidirectionally weighted feature pyramid structure. Firstly, a channel attention mechanism called SENet (squeeze and excitation network) is introduced into the backbone network to improve the network's feature extraction ability. Secondly, the BiFPN network is introduced to replace the feature fusion network PANET to enhance the fusion of shallow feature information and deep feature information. Finally, the spatial pyramid pooling structure SPPFCSPC is introduced to reduce the amount of calculation and improve the detection speed. Our experiments demonstrate that the updated YOLOv5-ABS network structure can achieve a maximum mAP of 78.6% on the NEU-DET dataset, which is 3.8% higher than the mAP of the previous YOLOv5s network. A better balance between detection speed and accuracy is achieved with a detection speed of 122 FPS.

2 Related Works

The YOLO series algorithm [9] is a one-stage target detection model based on deep learning and convolutional neural networks. Compared with previous versions, the YOLOv5 model has higher detection accuracy, faster detection speed, and a smaller model volume. According to the difference between network width and depth, the YOLOv5 model has five versions: YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. The detection accuracy of the five versions continues to increase, but as the model size and model parameters continue to increase, the detection speed decreases significantly. Considering the requirements of strip steel surface defect detection in terms of detection accuracy, detection speed, and real-time performance, YOLOv5s is selected for improved design in this study.

3 Methodology

In this section, we provide an improved model. Figure 1 represents the structure of the upgraded YOLOv5-ABS model, which is primarily made up of the Backbone, Neck, and Head.

Backbone: This component's primary job is to extract information about image features. It consists of the Conv module, the SeC3 module with channel attention mechanisms, and SPPFCSPC (Cross-Stage Partial Fast Spatial Pyramid Pooling). Conv performs convolution on the input features, BN prevents overfitting and facilitates model training, and the SiLU activation function is used to improve the expression ability of the model. The SeC3 module first uses the channel attention mechanism to enhance the feature extraction ability of the model, and then uses the C3 module to reduce the amount of calculation and memory. SPPFCSPC is used to fuse the feature

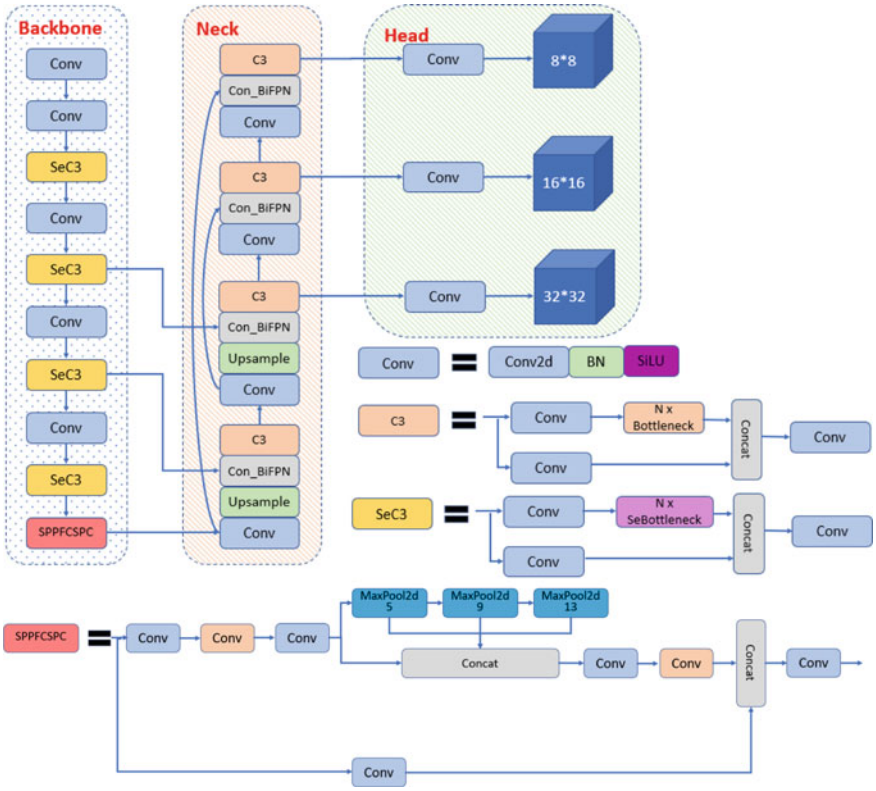


Fig. 1 YOLOv5-ABS model structure

maps of different receptive fields, enrich the feature expression ability, and improve the running speed.

Neck: Using BiFPN structure, the image feature information processed by the channel attention mechanism is integrated into the BiFPN network, which enhances the network's feature fusion ability and solves the problem of multi-scale feature fusion.

Head: the output of the model, which is mainly used to generate bounding boxes and predict defect categories. The loss function of the Bounding box is GIOU Loss, and the target box is filtered by NMS (Non-Maximum Suppression) to eliminate redundant boxes.

3.1 SeC3 Module Integrated with SENet

SENet (Squeeze-and-excitation networks) [10] is a channel attention mechanism network. The network mainly enhances the important features and suppresses the

general features by obtaining the importance of each feature channel. Consequently, the model's capacity to extract target features and detection precision are enhanced.

The workflow of SE mainly includes three steps: Sequeeze, Excitation and Scale, as shown in Fig. 2. First, Sequeeze average pooling or maximum pooling the feature maps obtained by convolution to obtain a feature map of 1 × 1 × C. The Sequeeze formula is:

$$Z = Fsq(Xc) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} Xc(i, j) \tag{1}$$

In the formula (1), $H \times W$ is the size of channel space, X_c is the input feature map; (i, j) is the point with abscissa i and ordinate j on the feature map; $F_{sq}(X_c)$ represents the compression operation of the feature map; Z is the weight obtained by the compression channel.

Then, the relationship between the feature channels is grasped by the Excitation operation to generate weights for each feature channel. The Excitation formula is:

$$\mathrm{Sc} = Fex(Z, W) = Sigmoid[W2 \times \mathrm{Re}LU(W1, Z)] \tag{2}$$

In the formula (2), W is the dimension, S_c is the attention weight generated after F_{ex} operation.

Finally, the weight of the Excitation output is multiplied by the input feature to complete the recalibration of the original feature, so that the model can distinguish the features of each channel. The Scale formula is:

$$\widetilde{X} = F\mathrm{scale}(Xc, Sc) \otimes Sc \tag{3}$$

In the formula (3), $\otimes$ is the multiplication of elements, F_{scale} is a weight resetting operation, $\widetilde{X}$ is the result of SE channel attention.

The SeC3 module in this study was created by combining SE with the C3 unit. The SeC3 module is frequently utilized in the YOLOv5-ABS backbone, making it easier for the model to extract image features and increase detection accuracy.

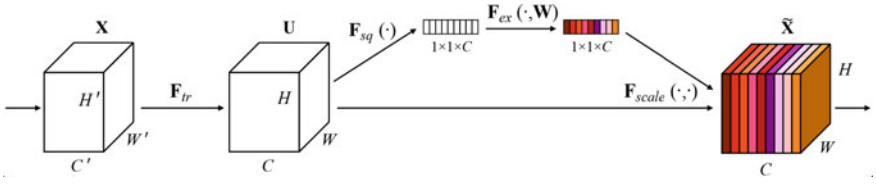


Fig. 2 SENet network structure

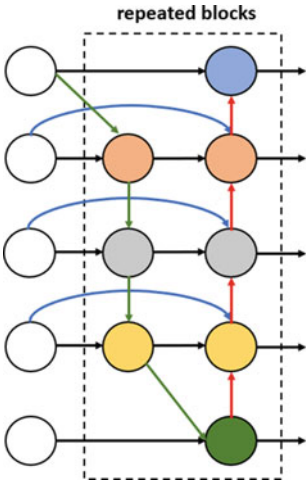


Fig. 3 BiFPN structure

3.2 Multi-scale Features Fusion Network BiFPN

Multi-scale feature fusion is frequently employed in target detection networks to enhance the model's detection performance. At present, the commonly used features of fusion networks are FPN [11], PANet [12], and BiFPN [13]. FPN establishes a top-down feature pyramid structure to achieve effective fusion of image semantic level and feature level, but it is limited by one-way feature information transmission. In order to make it simpler for the underlying information to be transferred to the top of the high-level, PANet integrates a bottom-up channel on top of FPN; however, this also makes the model more complex and decreases the efficiency of information transmission. BiFPN is a Bidirectional Feature Pyramid Network proposed by the Google team, as shown in Fig. 3. The network first deletes nodes with just one input edge; then, a new channel is established between the previous input point and the output point, allowing additional feature information to be fused without incurring excessive costs. In this study, BiFPN is added to the neck network to improve the network's capacity for feature fusion.

3.3 SPPFCSPC Module

According to the design scheme of SPPCSPC, this study combines the SPPF pyramid pooling structure and CSPNet to obtain the SPPFCSPC structure [14], which achieves speed improvement while keeping the receptive field unchanged. SPPF obtains different receptive fields through maximum pooling, which makes the model adapt to different resolution images. SPPF effectively solves the problem of repeated feature extraction, improves the generation speed of candidate boxes, and saves computing

costs. CSPNet divides the features into two parts, only one of which is processed routinely. Finally, the two parts are combined, the calculation amount is reduced by half, and the speed and accuracy are improved.

4 Experimental Results and Analysis

4.1 Experimental Dataset

In this research, the strip steel surface defect dataset NEU-DET produced by Northeastern University was used, which contains six common types of surface defects: Rolled-in Scale (RS), Crazing (Cr), Inclusion (In), Patches (Pa), Scratches (Sc), Pitted Surface (PS). The dataset has a total of 1800 images, and each type of defect has 300 images. The grayscale image is 200 × 200 pixels in size in its original form. A total of 1440 images were chosen at random to make up the training-set; 72 images served as the validation-set; and 288 images served as the test-set. In order to increase the amount of training data and enrich the training scene, this study uses mosaic data enhancement technology to randomly crop four images and splice them into one image as training data. Considering the characteristics of the model and the original resolution, the input image's measurement is set to 256 × 256.

4.2 Experimental Environment and Evaluation Metrics

The Windows 11 operating system serves as the foundation for the study's experimental environment. The CPU is an Intel Core i7-12700H, the memory is 32 GB, and the GPU is a NVIDIA GeForce RTX 3070 Ti Laptop 8 GB. Using PyTorch 1.12.1 as the deep learning framework, the Python version is 3.8 and the CUDA version is 11.3. During the experiment, the epoch is set to 100 and the Batchsize is set to 128. The momentum is set to 0.937, the initial learning rate lr0 is set to 0.01, the cosine annealing algorithm is used to dynamically adjust the learning rate, and the weight _ decay is set to 0.0005.

This study uses Average Precision (*AP*) and Mean Average Precision (*mAP*) to evaluate the improved model. *AP* (formula 6) represents the mean value of Precision (*P*, formula 4) under different Recall (*R*, formula 5), that is, the *PR* curve obtained with *P* as the ordinate and *R* as the abscissa.

$$P = \frac{TP}{TP + FP} \times 100\% \tag{4}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{5}$$

$$AP = \int_0^1 p(r)dr \tag{6}$$

where *TP* means True Positives, *FP* means False Positives, and *FN* means False Negative.

The mean *mAP* (formula 7) represents the average accuracy of all target detection categories.

$$mAP = \frac{\sum_{i=0}^{n} AP(i)}{n} \tag{7}$$

where *n* represents the number of defect categories.

4.3 Ablation Experiment

The performance of the proposed method in strip surface defect detection is verified by ablation experiments. Based on the YOLOv5s network, the SeC3 module is used to replace the C3 module in the backbone named YOLOv5s-A; the introduction of the BiFPN module based on YOLOv5s-A is named YOLOv5s-AB; and based on YOLOv5s-AB, the SPPFCSPC module is used to replace the SPPF module named YOLOv5s-ABS. Table 1 displays the experimental outcomes.

According to the experimental results, after introducing the SE channel attention mechanism into the C3 module, the mAP of YOLOv5s-A is 75.8%, which is 1% higher than the original algorithm. This demonstrates that the SE attention mechanism effectively increases the model's ability to extract the image's global feature information and boosts detection accuracy. After introducing the BiFPN structure, the parameters and GFLOPs of YOLOv5s-AB are slightly improved compared with the original algorithm, but its mAP is further increased to 77.9%, which indicates that the use of BiFPN can effectively improve the detection accuracy of the network model. On this basis, the SPPFCSPC module is introduced. It can be found that the parameters and GFLOPs of YOLOv5s-ABS increase greatly, but the mAP is further improved to 78.6% and the detection speed is basically the same as the original algorithm, so it can still meet the real-time detection requirements.

Table 1 Results of ablation test

Model	Parameters	GFLOPs	FPS	mAP
YOLOv5s	6.29×10^6	15.0	147.4	74.8
YOLOv5s-A	6.06×10^6	13.8	145.1	75.8
YOLOv5s-AB	7.12×10^6	15.1	144.7	77.9
YOLOv5s-ABS	13.55×10^6	20.3	142.8	78.6

Figure 4 depicts the AP of each type of defect in YOLOv5s-ABS and YOLOv5s. The figure shows that the AP of YOLOv5s-ABS is enhanced as compared to the initial algorithm; the AP of RS, Cr, Pa, Sc, and Ps is increased by 1.6%, 19%, 1.4%, 0.7%, and 1.1%, respectively, except that the AP of In is reduced by 1.1%. The detection accuracy is substantially better than the previous algorithm, especially for Cr. It has been discovered that the application of the BiFPN multi-scale features fusion network, the SE channel attention mechanism, and the SPPFCSPC structure may significantly increase the detection accuracy of strip steel defects.

Figure 5 shows the test results for six kinds of defect samples. Different defects use different colors of Bounding-Box, and the upper right of each Bounding-Box has the defect name and confidence. The illustration demonstrates that for the same image, the enhanced algorithm has a better detection effect on all kinds of defects and can accurately detect the missing parts of the original algorithm.

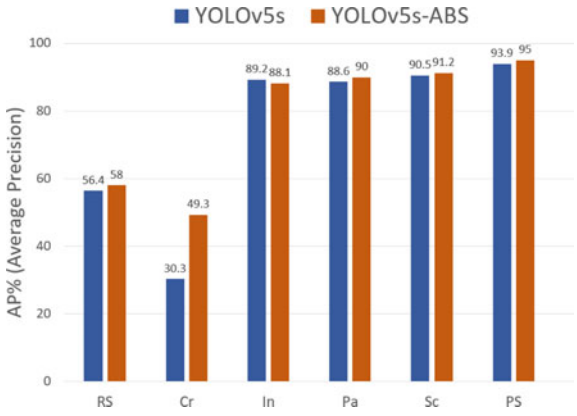


Fig. 4 Comparison of various types of AP between the two algorithms

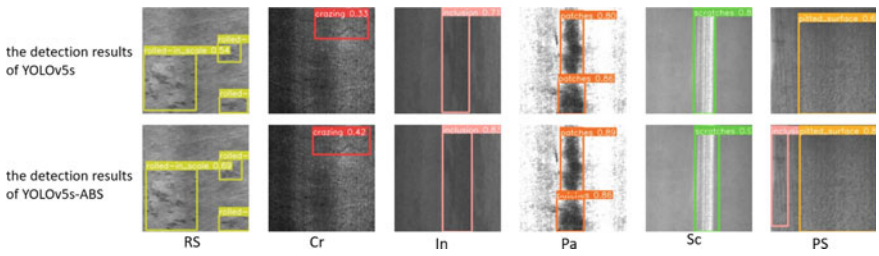


Fig. 5 Detection results of six kinds of defect samples

Table 2 Comparison of 5 different models on NEU-DET

Model	AP%						Weight/MB	mAP@0.5%
	RS	Cr	In	Pa	Sc	PS		
YOLOv3	53.1	36.7	62.8	77.4	72.2	68.5	235	61.8
YOLOv4	45.6	31.7	75.3	81.8	78.2	72.4	244	64.2
SSD	54	32.1	62.3	82.5	71.9	71.6	93.9	62.4
Faster-RCNN	57	37.3	83.6	89.7	90.6	90	108	74.7
YOLOv5s-ABS	58	49.3	88.1	90	91.2	95	26.7	78.6

4.4 Comparative Experiment

To further confirm the effectiveness of the improved YOLOv5s-ABS algorithm, this study selects four mainstream target detection algorithms YOLOv3, YOLOv4, SSD, and Faster-RCNN to compare with the improved YOLOv5s-ABS model on the NEU-DET dataset. Table 2 presents the outcomes.

According to Table 2, the mAP of the improved model YOLOv5s-ABS in this study reached 78.6%, which has a higher detection accuracy than the other four mainstream target detection models. Among them, the mAP of Faster-RCNN is 74.7%, which is close to the improved model. However, Faster-RCNN is a two-stage detection model with a complex training process, a large amount of calculation, and slow speed. The mAP of SSD is low, and it performs poorly in Cr defects but better in Pa defects. This is because Pa defects are mostly large targets, and Cr defects are mostly small targets, which reflects the shortcomings of SSD in small target defect detection. The detection accuracy of YOLOv3 and YOLOv4 models is low, the model volume is large, and the industrial deployment cost is high. In general, YOLOv5s-ABS has higher accuracy on the NEU-DET dataset, which can maintain a good balance between accuracy and speed, and a smaller model volume can be more convenient for industrial deployment.

5 Conclusions

This paper offers an improved YOLOv5s-ABS algorithm to solve the problem of strip steel surface defect detection and increase the accuracy and speed of strip steel surface defect detection. The algorithm first inserts the SENet channel attention mechanism into Backbone, which improves the feature extraction ability of the network. Secondly, the bidirectional weighted feature pyramid network BiFPN is introduced into Neck to strengthen the network features fusion ability. Finally, the SPPFCSPC structure is introduced to further improve speed and accuracy while keeping the receptive field unchanged. The algorithm's usefulness is demonstrated by experimental findings. Since the improved model has a low recognition rate for

the two types of defects of Crazing and Rolled-in Scale, the next step will be to study the detection of these two types of defects and further optimize the detection accuracy of the model.

Acknowledgements This work was also supported by The Key project of natural science research in colleges and universities of Anhui Province (2022AH052365, KJ2020A1102) and Wuhu Engineering Technology Research Center (KJCXPT202204) and The Open Research Fund of Anhui Province Key Laboratory of Machine Vision Inspection.

References

1. Li YJ, Zhu PY, Sun XP (2017) Shape defect detection algorithm of stamping parts based on shape template matching. J Guangzhou Univ (Natural Science Edition) 16(5):62–66
2. Liu Y, Xia CL (2021) An image edge detection algorithm for strip steel surface defects based on Sobel opera-tor. Electronic Measurement Technol 44(3):138–143
3. Liu L, Wang C, Zhao SW (2018) Research on solar cells defect detection technology based on machine vision. J Electronic Measurement Instrument 32(10):47–52
4. Zhang T, Liu YT, Yang YN (2020) Review of surface defect detection based on machine vision. Sci Technol Eng 20(35):14366–14376
5. Dai XH, Chen HJ, Zhu CP (2020) Surface defect detection and realization of metal workpiece based on improved faster RCNN. Surface Technol 49(10):362–371
6. Kou XP, Liu SJ, Cheng KQ (2021) Development of a YOLO-V3-based model for detecting defects on steel strip surface. Measurement 182(2):109–119
7. Zhang BY, Leng YB (2022) Detection of metal surface scratch based on YOLOv4 network model. Acta Armamentarii 43(1):214–221
8. Wang K, Teng ZX, Zou TY (2022) Metal defect detection based on Yolov5. J Phys: Conf Ser 2218(1):120–128
9. Shao Y, Zhang D, Chu H (2022) A review of YOLO object detection based on deep learning. J Electron Inf Technol 44(2):1–12
10. Hu J, Shen L, Albanie S (2020) Squeeze-and-excitation networks. IEEE Trans Pattern Anal Mach Intell 42(8):2011–2023
11. Lin TY, Dollar P, Girshick R (2017) Feature pyramid networks for object detection. In: Computer vision and pattern recognition
12. Liu S, Qi L, Qin H (2018) Path aggregation network for instance segmentation
13. Tan M, Pang R, Le QV (2020) EfficientDet: scalable and efficient object detection. In: IEEE conference on computer vision and pattern recognition
14. He TJ, Li H (2023) The pavement damage detection model based on improved YOLOv5. Chin Civil Eng J 17(2):1–12

Optimization of Ultrasonic Atomization by Electrolytic Polishing

Wei-Qing Huang, Jia-Li Liang, Jian-Hui Zhang, and Fan Zhang

Abstract Because of structural simplicity, low production costs and small atomizing droplets, ultrasonic atomization used in aerospace, medical, energy saving and other fields. With the increasing need to increase improve the atomization efficiency, Researchers have made a lot of efforts. However, the micro-tapered hole group on atomizer has not been deeply studied. In this study, the electrolytic polishing processing is proposed to improve the quality of the micro-tapered holes. First, a mathematical relation of atomization amount and polishing time was established. Second, the relevant parameter were measured using a microscope. Lastly, the mathematical relation was determined and revised through an atomization amount experiment. The experimental results show that when the polishing time is 4 min, the atomization amount reaches the maximum of 4.53 ml/min, which is much higher than that of the untreated atomizer. This study demonstrates that polishing can effectively remove burr in the process of hole machining, which will further improve the atomization efficiency of atomizers.

Keywords Ultrasonic atomization · Piezoelectric · Electrolytic polishing · Laser processing

1 Introduction

Ultrasonic atomization is the use of electronic high-frequency oscillation, through the ceramic atomization plate high-frequency resonance, the molecular bonds between liquid water molecules to break up and produce natural floating water mist. The micro conical hole on the metal plate of atomizer is the core position to produce the moving cone Angle effect and realize atomization [1–3]. Maehara pioneered the atomizer structure consisting of multiple pinhole sheets and show that the maximum atomization rate is approximately proportional to the number of micro-tapered holes [4]. Toda proposed a plate atomizer with small holes to obtain the smallest piezoelectric

W.-Q. Huang (✉) · J.-L. Liang · J.-H. Zhang · F. Zhang
School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou, China
e-mail: meehuangweiqing@gzhu.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_28

atomizer, which was not only convenient to carry but also improved the atomization efficiency [5, 6]. Than Lu and Shen studied an atomizer with a cymbal-type high-power driver [7, 8].

At the same time, researchers have optimized vibration to improve the atomization amount. Maehara further studied the existing atomizer to improve its atomization efficiency [9]. Whereafter, Lam found that a PMN-PT single crystal with ultrahigh voltage electric coefficient could generate higher axial displacement [10]. In recent years, Yan and Pallavi has conducted extensive research on the influence of dynamic cone Angle, frequency and other parameters on the atomization rate of ultrasonic atomizer [11–14].

At present, micro-tapered holes of metal sheets of ultrasonic atomizer are mostly made by laser thermal processing technology, so the molten metal fragments will gather near micro-tapered holes to form uneven burrs, which will affect the atomization efficiency of the atomizer sheets [15]. In order to solve the influence of burr on atomization effect caused by this process, electrolytic polishing process is proposed to treat metal sheet in this study.

2 Mathematical Relation of Atomization Amount and Polishing Time

2.1 Theoretical Model of Atomization Amount and Polishing Time

Given $D^2 + d^2 + Dd = D_1$, where D is the diameter of the large end of a micro-tapered hole, d is the diameter of the small end of a micro-tapered hole. The actual metal sheet mass is the mass of the entire metal sheet without laser processing minus the mass of micro-tapered hole

$$m = \frac{\pi}{12}\rho\xi\left(12R^2 - N_1 D_1\right) \tag{1}$$

where ρ is the density of the stainless steel, R is the radius of the metal sheet, and ξ is the thickness of the metal sheet, and N_1 is the number of micro-tapered holes.

Figure 1 shows a structural drawing of the atomization sheet. Given that the radius R of the metal sheet changes minimally in electrolytic polishing, it can be ignored. Hence, According to Faraday's law of electrolysis and the law of conservation of mass

$$\frac{\pi}{12}\rho\xi N_1\left(D_1' - D_1\right) = \frac{MIt}{N_2 F} \tag{2}$$

This paper adds a correction coefficient μ when calculating the burr. Given $D + d = D_2$, the following equation is obtained:

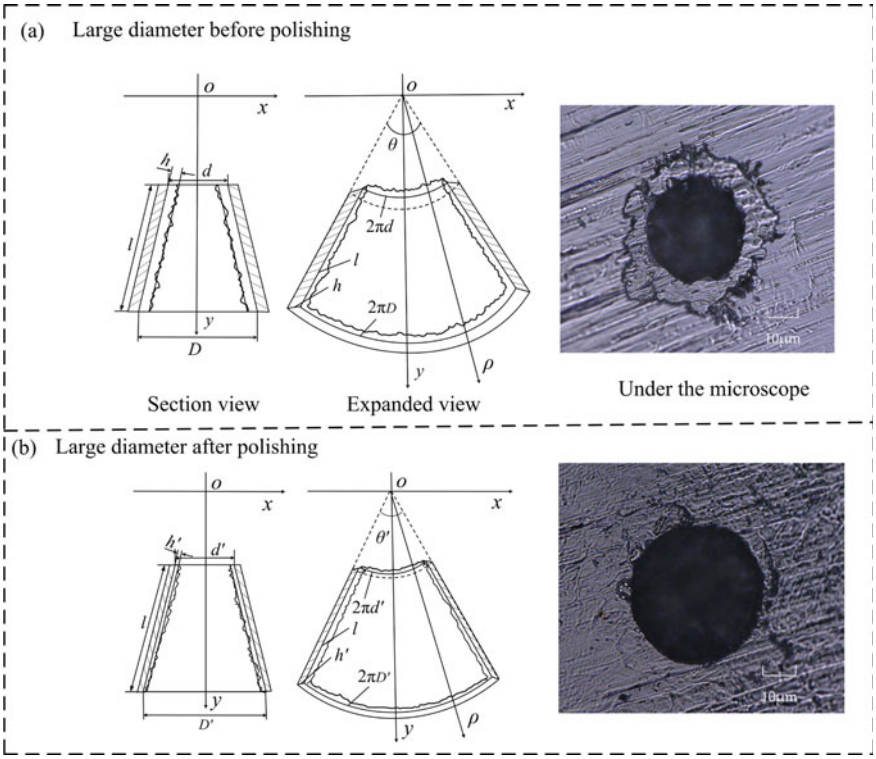


Fig. 1 Comparison of the state of the metal sheet before and after polishing: **a** large end of a micro-tapered hole before polishing; **b** large end of a micro-tapered hole after polishing

$$V = \frac{\pi}{12}\xi D_1 - \frac{\pi}{2}\mu h l D_2 \tag{3}$$

where μ represents the correction coefficient, h is the burr height, and l is the busbar length. The difference between the volume of the cavity where the atomization sheet reaches the maximum deformation when vibrating after electrolytic polishing and the maximum deformation when vibrating before electrolytic polishing is

$$\Delta V = \frac{\pi}{12}\xi\left(D_1' - D_1\right) + \frac{\pi}{2}\mu\left(h l D_2 - h' l' D_2'\right) \tag{4}$$

For the micro-tapered hole, the atomization volume delivered is as follows [16]:

$$Q = \Delta V f \frac{(\xi_d - \xi_n)}{(2 + \xi_d + \xi_n)} \tag{5}$$

where Q is the amount of atomization produced by the atomization sheet, f is the vibration frequency, ξ_d is the conical diffuser loss coefficient, and ξ_n is the nozzle

diffuser loss coefficient. Let $hlD_2 - h'l'D_2' = P$, $\frac{(\xi_d - \xi_n)}{(2+\xi_d+\xi_n)} = K$. In summary, the expression of the atomization quantity of the atomization sheet can be written as

$$Q = \begin{cases} f\left(\frac{\pi}{12}\xi D_1 - \frac{\pi}{2}\mu hlD_2\right)K, & (t=0) \\ f\left(\frac{MIt}{N_1N_2F\rho} + \frac{\pi}{2}\mu P\right)K, & (t>0) \end{cases} \tag{6}$$

Given that f, M, I, N_1, N_2, F, ρ, and μ are constants or measurable quantities and the change in K is minimal in a short electrolytic time, the change in K is temporarily ignored. The relationship between electrolytic time t and K is then discussed in the following experiment. Moreover, the value of atomization quantity Q can be calculated by integrating the parameters D, d, and h measured in the experiment into Eq. (6).

3 Electrolytic Polishing Experiment

3.1 The Composition of the Electrolyte

The main factors altering the effect of electrolytic polishing are time, voltage, temperature, and the distance between anode and cathode plates [17, 18]. The traditional electrolytic polishing with strong acid composite solution is easy to lead to excessive corrosion. Since this experiment only requires the removal of burrs at a micron height, the chromic anhydride component is removed, while the amount of phosphoric acid and sulfuric acid is greatly reduced, and polyethylene glycol is added as a thickening agent to replace the traditional process containing chromium anhydride [19, 20]. The electrolyte components used in this experiment are shown in Table 1.

Table 1 Electrolyte composition table

Serial number	Component	Mass fraction
1	Phosphoric acid	10–15%
2	Sulfuric acid	5–10%
3	The cationic surfactant	1.1–2.3%
4	Polyethylene glycol	1.5–3%
5	Glycerin	1–3%
6	Brightener	1.2–3.1%
7	Water	65–75%

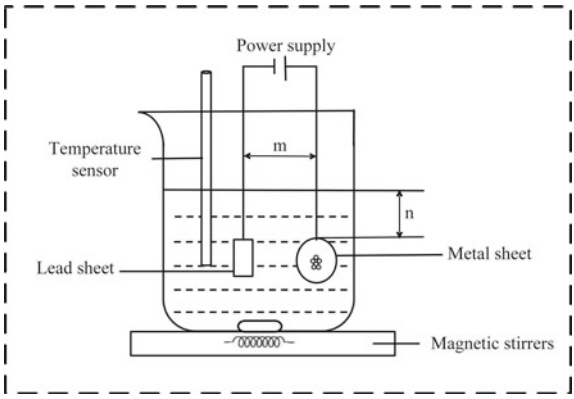


Fig. 2 Experimental diagram of electrolytic polishing

3.2 Experimental Equipment System

In this experiment, electrolytic polishing was used to remove the uneven burr on the surface of microporous metal sheet. In the test, the atomizing plate is the anode and the lead plate is the cathode. Figure 2 is the experimental device and schematic diagram of electrolytic polishing, respectively. Where, n a is the distance from the anode plate to the liquid surface, and the experimental value is 40 mm; m is the distance between the anode and cathode plates, and the experimental value is 45 mm. According to the above analysis, the experimental control of polishing voltage is 6 V, and electrolyte temperature is 35 °C. The polishing time range is 1–6 min, and the interval time is 1 min in the experiment to conduct the group experiment.

4 Measurement Experiment of Metal Sheet

After the electrolytic polishing experiment, the large hole diameter D, small hole diameter d and burr height h of the metal sheet before and after electrolytic polishing were measured by the ultra-depth of field three-dimensional microscope system, so as to explore the influence of electrolytic polishing on the diameter of the large and small ends of micro-tapered holes and the surface quality of a laser-machined metal sheet.

4.1 Measurement of Related Parameters of Metal Sheet

According to the different electrolytic time, the experiment was divided into seven groups. The first group was atomized sheet without treatment, and the polishing time was 1 min, 2 min to 6 min, and the second to the seventh group successively.

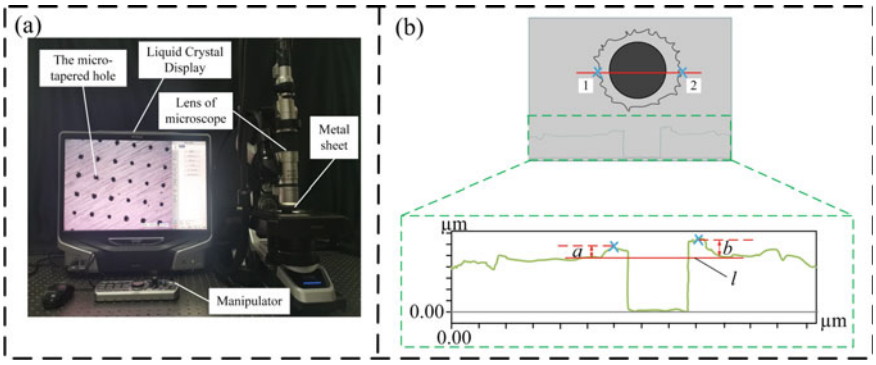


Fig. 3 Picture of ultra-depth-of-field 3D microscopic measurement: **a** Microscope measurement object diagram; **b** schematic of enlarging the burr height of the metal sheet

When measuring the diameter of large holes of metal sheets, the diameter of 100 holes of each group of metal sheets was measured and counted, and 700 large hole diameters were obtained in 7 groups. Similarly, a total of 700 small hole diameters were measured by measuring the small hole diameters of the above 7 groups of metal sheets. Figure 3a is the measurement diagram of the ultra-depth-of-field 3D microscope. Figure 3b is a schematic of enlarging the burr height of the metal sheet.

4.2 Data Analysis

According to the statistical data of the previous experiments, a box plot based on the statistical data of the previous experiments is drawn in Fig. 4. Figure 4a presents the relationship between the electrolytic polishing time and the diameter of micro-tapered holes of the metal sheet, and Fig. 4b presents the relationship between the electrolytic polishing time and the burr height of micro-tapered holes of the metal sheet.

It can be found that with the increase of polishing time, the diameter of the metal plate increases, but the burr height decreases. Burr height dispersion is also reduced.

5 Experimental Analysis of Atomization Amount

In order to further verify the influence of electrolytic polishing on atomization amount. Figure 5 is the atomization test device. In this experiment, the atomization volume of 7 groups of nebulizing tablets with the operating voltage of 40 V and the polishing time of 1–6 min were measured.

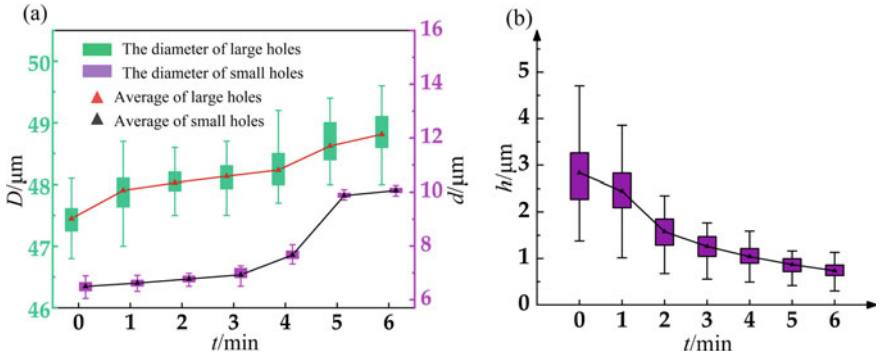


Fig. 4 The relationship between electrolytic polishing time and metal sheet: **a** Relationship between electrolytic polishing time and the diameter of micro-tapered holes; **b** relationship between electrolytic polishing time and burr height

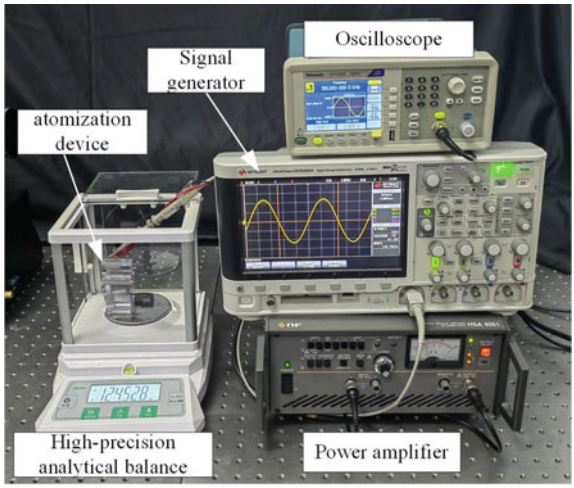


Fig. 5 Atomization experiment device

From the atomization experiment, the atomization amount of different polishing time is basically consistent with the theory analyzed above. However, with the increase in electrolytic time, the influence of K on Q can no longer be ignored, that is, when $t > 4$, K cannot be regarded as a constant. The reason is that with the increase in electrolytic polishing time, K will affect the cone angle θ along with the change in t. The conical and nozzle diffuser loss coefficients are then affected gradually, thus altering K. According to the experiment, when $t > 4$, the correction coefficient should be added to make the correction.

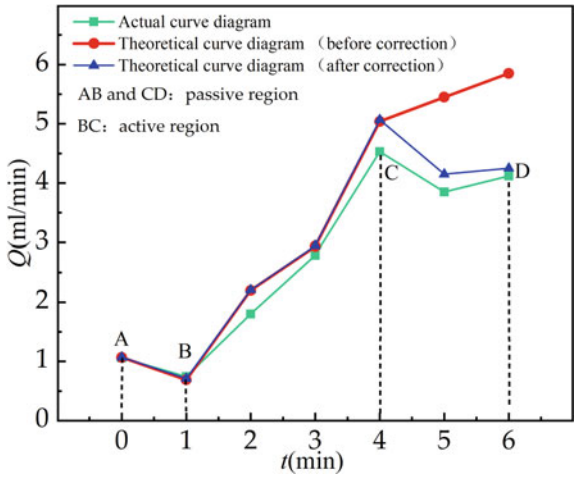


Fig. 6 Atomization amount under different polishing time

$$Q=\begin{cases} f\left(\dfrac{\pi}{12}\xi D_1-\dfrac{\pi}{2}\mu hlD_2\right)\mathrm{K}, & (t=0) \\ f\left(\dfrac{MIt}{N_1N_2F\rho}+\dfrac{\pi}{2}\mu P\right)\mathrm{K}, & (1\le t\le 4) \\ f\left(\dfrac{MIt}{N_1N_2F\rho}+\dfrac{\pi}{2}\mu P\right)\gamma\mathrm{K}, & (t>4) \end{cases} \tag{7}$$

From Eq. (7), the change in atomization amount with polishing time can be obtained, as shown in Fig. 6. The atomization amount basically increases with the increase in electrolytic polishing time, indicating that the theoretical analysis is consistent with the experimental results. When the electrolysis time is 4 min, the atomization amount reaches the maximum value of 4.53 ml/min.

As can be seen from Fig. 6, In the AB region, the presence of an initial passivation film on the surface of the sheet metal prevents the metal from dissolving, resulting in atomized volume. Later, in BC zone, the atomization volume increases with the increase of polishing time. When the polishing time is 4 min, the atomization volume reaches the maximum, reaching 4.53 ml/min. Then the metal sheet enters the CD passivation zone and a new passivation film is formed on the anode surface. Due to the obstruction of the passivation film, the metal dissolution rate is very slow, resulting in the atomization volume beginning to decrease. Therefore, in order to achieve a good polishing effect, the electrolytic time should be controlled within a reasonable range of the over-passivation zone.

6 Conclusions

- With the increase in polishing time, the diameter of the metal sheet increases, but its burr height decreases. The dispersion of the burr height also decreases. Polishing can dissolve the uneven burr on the surface of the atomization sheet and improve the atomization effect of atomizers. Specifically, when the polishing time is 4 min, the maximum atomization amount reaches 4.53 ml/min.
- When the polishing time is extremely long, the atomization amount will decrease. When the polishing time t is excessively long, K can no longer be treated as a constant. According to the atomization experiment, K should be corrected when $t > 4$.

References

1. Guerra-Bravo E, Lee H-J, Baltazar A, Loh KJ (2021) Vibration analysis of a piezoelectric ultrasonic atomizer to control atomization rate. Appl Sci 11:8350
2. Jianhui Z, Qiufeng Y, Wanting S (2020) Advances in piezoelectric atomizers. Trans Nanjing Univ Aeronautics Astronautics
3. Pritchard JN, Hatley RH, Denyer J, Hollen DV (2018) Mesh nebulizers have become the first choice for new nebulized pharmaceutical drug developments. Therapeutic Deliv 9:121–136
4. Maehara N, Ueha S, Mori E (1986) Influence of the vibrating system of a multipinhole-plate ultrasonic nebulizer on its performance. Rev Sci Instrum 57:2870–2876
5. Toda K, Akimura Y (1994) An ultrasonic atomizing device using coupled-mode vibration. Rev Sci Instrum 65:3276–3278
6. Toda K, Ishii JIJ (1995) Operation performance of self-oscillation ultrasonic vibrating device for liquid atomization. Jpn J Appl Phys 34:5332
7. Lu C, Fu C, Yang J, Chen C, Cheng C (2006) 4A-3 traveling wave driven micro-dispenser for CPU cooling application. In: Proceedings of the 2006 IEEE ultrasonics symposium, pp 54–57
8. Shen S-C, Wang Y-J, Chen Y-Y (2008) Design and fabrication of medical micro-nebulizer. Sens Actuators A 144:135–143
9. Maehara N, Ueha S, Mori E (1987) Optimum design procedure for multi-pinhole-plate ultrasonic atomizer. Jpn J Appl Phys 26:215
10. Lam KH, Chan H, Luo H, Yin Q, Yin Z (2005) Piezoelectrically actuated ejector using PMN–PT single crystal. Sens Actuators A 121:197–202
11. Dupuis E, Momen AM, Patel VK, Shahab S (2018) Multiphysics modeling of mesh piezoelectric atomizers. In: Proceedings of the active and passive smart structures and integrated systems XII, pp 429–437
12. Sar G, Aydn YN, Bilekli MG, Gl MB (2021) Dynamics of liquid drop on a vibrating microperforated plate. Zeitschrift fur Naturforschung A 76
13. Sharma P, Jackson N (2022) Vibration analysis of MEMS vibrating mesh atomizer. J Micromechanics Microengineering 32
14. Yan Q, Sun W, Zhang J (2020) Study on the influencing factors of the atomization rate in a piezoceramic vibrating mesh atomizer. Appl Sci 10:2422
15. Zhang L, Tong H, Li Y (2015) Precision machining of micro tool electrodes in micro EDM for drilling array micro holes. Precis Eng 39:100–106
16. Akiyoshi O (2004)Flow direction of piezoelectric pump with nozzle/diffuser-elements. Chin J Mech Eng 107–109

17. Han W, Fang F (2019) Fundamental aspects and recent developments in electropolishing. Int J Mach Tools Manuf 139:1–23
18. Huang Y, Wang C, Ding F, Yang Y, Li N (2021) Principle, process, and application of metal plasma electrolytic polishing: a review. Int J Adv Manuf Technol
19. Han W, Fang F (2019) Electropolishing of 316L Stainless Steel using sulfuric acid-free electrolyte. J Manuf Sci Eng 141
20. Zhou C, Su H, Qian N, Zhang Z, Xu J (2022) Characteristics and function of vapour gaseous envelope fluctuation in plasma electrolytic polishing. Int J Adv Manuf Technol 119:7815–7825

Research on Automatic Modulation Recognition Method Based on Deep Learning

Sen Yan, Xiaohua Zhang, and Shubin Wang

Abstract In non-cooperative communication systems, Automatic Modulation Recognition (AMR) is a key technology for spectrum sensing, spectrum monitoring and spectrum utilization. Most traditional AMR methods ignore the diversity and intrinsic connections of features, and also face the challenges of low recognition rate and weak generalization ability. In this paper, we propose a parallel neural network combining skip-connected Convolutional Neural Network (CNN) with Gated Recurrent Unit (GRU) to extract spatiotemporal features in parallel from both the In-phase Quadrature (IQ) and Amplitude Phase (AP) components of the signal. The proposed network uses skip-connection structures to effectively reduce the problems of gradient vanishing and network degradation. GRU reduces the computational parameter count while retaining the advantages of Long-Short Term Memory (LSTM) networks. In addition, based on the interdependence between different feature channels, we introduce a lightweight and efficient channel attention network to re-weight all features, further improving the signal recognition rate. The results show that the proposed network achieves a recognition rate of 90% in environments where the Signal to Noise Ratio (SNR) is above 4 dB, and can effectively improve the recognition rate in low SNR conditions. Compared with two other parallel neural networks, the dual-stream CNN and dual-stream CNN-LSTM, our network can achieve higher recognition accuracy with lower complexity.

Keywords Automatic modulation recognition · Neural network · Attention · Signal to noise ratio

S. Yan · S. Wang (✉)
College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China
e-mail: wangshubin@imu.edu.cn

X. Zhang
Department of Foreign Languages, Guizhou University of Commerce, Guiyang, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_29

1 Introduction

AMR is a necessary step between signal detection and demodulation. In non-cooperative communication systems, it can help to manage spectrum resources, achieve dynamic allocation of spectrum and improve spectrum utilization [1]. Traditional AMR methods can be broadly classified into Likelihood Based (LB) methods and Feature Based (FB) methods [2]. Compared to LB methods, FB methods have demonstrated stronger performance and lower computational cost, but traditional machine learning algorithms have limited generalization capability and cannot handle complex modulation recognition tasks. To address these challenges in AMR, academic research has recently focused on Deep Learning (DL) methods, and various DL-based AMR methods have been developed in recent years.

O'Shea et al. first used CNN to extract features from the IQ of the original signal and then identified the modulation types with a classifier, which outperformed traditional machine learning methods [3]. Considering that the sampled modulated signal is essentially a time series, Rajendran et al. input the amplitude and phase features of the signal into a 2-layer LSTM for modulation recognition [4]. West et al. designed a Convolutional Long short-term Deep Neural Network (CLDNN) that combines CNN and LSTM [5]. The above studies have made certain progress in the field of signal modulation recognition, but they did not consider the diversity and inherent connections of signal features, which resulted in reduced recognition performance. Zhang et al. designed a CNN-LSTM based dual-stream structure, which takes both IQ features and AP features of signals as inputs [6]. Although multiple features of the signal are utilized to improve the recognition performance and learning ability in a high SNR environments, the impact of low SNR on modulation recognition performance has not been taken into account. Huang et al. input the amplitude and phase features of the received signal into a gated recurrent residual neural network, which effectively improved the robustness of the model [7]. However, they did not fully utilize the superior performance of IQ features in resisting low SNR environments, as well as the interdependent relationship between feature channels. Qi et al. proposed a multi-modal fusion model for waveform and spectrum, which uses multiple residual networks to compute related features [8]. Despite high accuracy, this model requires significant computational resources for extracting multimodal features. Kumar et al. proposed a modulation classification algorithm based on Constellation Density Matrix (CDM) [9]. Although this method ingeniously converts the AMR problem into a well-studied image recognition problem, it introduces complex image processing steps. In low SNR environments, large amounts of noise can mask the differences between different modulation schemes. As a result, most models face challenges such as low recognition rates, weak noise immunity, and difficulty in balancing model complexity and recognition accuracy.

In order to improve the AMR recognition rate and enhance the noise immunity of the model, this paper proposes a parallel neural network combining skip-connected CNN with GRU, which inputs the IQ and AP features of the signal in parallel to obtain deeper connections between the features and enhance the specificity of each

modulated signal. We also perform feature scaling on AP data to improve the model generalization ability. The proposed network adopts skip-connection structures to alleviate the problems of gradient vanishing and network degradation, and GRU is used to reduce the computational complexity while ensuring recognition accuracy. In addition, a lightweight and Efficient Channel Attention Network (ECANet) is used [10] to assign weights to the learned features to reduce data redundancy for the purpose of network optimization. Finally, the effectiveness of the network model is verified by simulations under different SNRs. The experimental results show that the parallel neural network proposed in this paper outperforms the existing advanced AMR methods.

2 Related Work

DL-based AMR typically involves two consecutive stages. In the first stage, the received signals are preprocessed for subsequent use, and in the second stage, we utilize Deep Neural Networks (DNNs) to process different features of the signal and determine the modulation scheme [11].

The input data for AMR tasks can be represented in three formats: IQ format, AP format, and constellation format [12]. Generally, the AP format is generated by transforming the IQ format from the Cartesian coordinate system to the polar coordinate system, which can be derived from a simple sequence transformation. The constellation format is a graphical representation of the modulation scheme used for baseband signals in digital communication [9]. However, the transformed constellation format can easily lose temporal information, so the IQ and AP features are widely used in AMR. From previous research on DL-based AMR, it can be observed that the IQ feature and AP feature can complement each other at high and low SNR. Specifically, the DL-based IQ models outperform the AP model in low SNR environments, while the opposite result is observed in high SNR environments. This suggests that IQ models exhibit good noise resistance in low SNR environments. Furthermore, combining the strengths of CNN and RNN in the model can enhance network performance. Therefore, it is of great significance to form a neural network that can effectively integrate multiple different features of signals in AMR.

2.1 Signal Model

In wireless communication systems, AMR is regarded as a multi-classification problem, which determines the modulation scheme based on the characteristics of the received signal sequence. Generally, the received baseband signal can be expressed as:

$$r(n) = Ae^{j(2\pi f_0 nT+\theta_n)} \sum_{-\infty}^{\infty} s(k)h(nT - kT + \varepsilon T) + g(n), \tag{1}$$

where A is amplitude factor, f_0 represents frequency offset, T represents the symbol spacing, θ_n represents phase jitter, $s(k)$ is the transmitting symbol sequence, $h(\cdot)$ is the channel response function, ε is the time error, $g(n)$ represents noise.

SNR is a crucial factor that affects the quality and reliability of signals in AMR. It is generally expressed as:

$$SNR = 10lg\left(\frac{P_S}{P_N}\right), \tag{2}$$

where P_S is the signal power, P_N is the noise power, and lg denotes the logarithm with a base of 10.

2.2 Signal Pre-processing

The IQ data vector of a signal is composed of the in-phase component $S_I(n)$ and the quadrature component $S_Q(n)$, which are represented as:

$$S_I(n) = real(r(n)), S_Q(n) = imag(r(n)). \tag{3}$$

Secondly, the AP data vector consists of the amplitude component $S_A(n)$ and phase component $S_P(n)$, represented as:

$$S_A(n) = \sqrt{(S_I(n))^2 + (S_Q(n))^2}, S_P(n) = \arctan\left(\frac{S_Q(n)}{S_I(n)}\right). \tag{4}$$

The IQ data of a signal can be obtained directly from the raw signal, but normalization is required while obtaining the AP data, including instantaneous amplitude normalized by L2 and instantaneous phase normalized between -1 and 1. Such normalization helps to avoid dimension influence between different indicators and improve the model's generalization ability.

3 Proposed AMR System Model

In this paper, a parallel neural network combining skip-connected CNN with GRU is proposed to effectively enhance signal recognition rate, noise resistance, and generalization ability, as well as to reduce computational complexity and dependence on expert knowledge. The overall network structure is shown in Fig. 1.

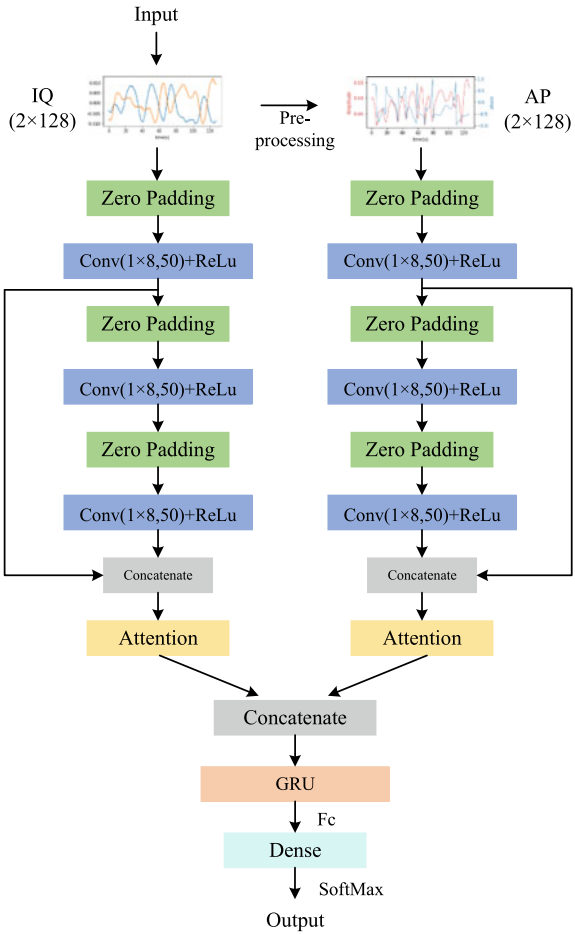


Fig. 1 Parallel neural network combining skip-connected CNN and GRU

In each path of feature extraction, there are first three convolutional layers with a kernel size of 50 and a size of 1×8. The output of the first convolutional layer is then connected to the output of the third convolutional layer through a skip connection to enhance the dependency between the preceding and succeeding convolutional layers. In addition, to preserve edge information features, zero-padding is applied before each convolutional layer for adjustment.

The convolutional layer is mainly responsible for receiving sample data from each channel and learning feature information from it, while the attention layer assigns corresponding attention weights to the feature set learned in the convolutional layer, helping the deep learning model focus on the key features of the signal data set. In this paper, we introduce a lightweight ECANet, as shown in Fig. 2. The ECANet module aggregates convolutional features through non-dimensional reduction Global Average Pooling (GAP), and learns channel attention by using the Sigmoid function

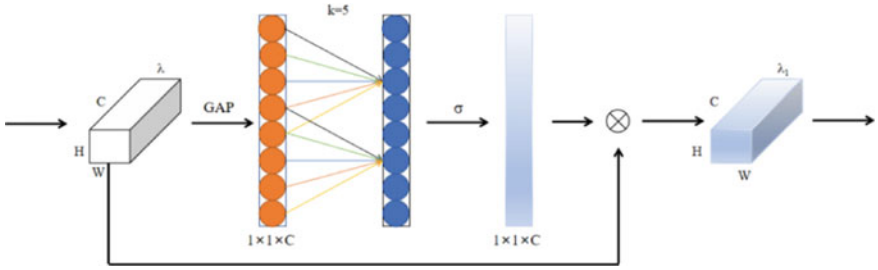


Fig. 2 Structure of efficient channel attention network

to determine the weight of each channel. Finally, the weights are multiplied by the corresponding elements of the original input features to obtain the final output features.

After the experience attention layer, in order to fully utilize the differential characteristics between the signal IQ and AP, a feature fusion is performed by using the sequence concatenation method. In DL networks, GRU has a simpler structure and fewer number of parameters than LSTM, reducing the space required for the model and increasing the operation speed while ensuring accuracy. Therefore, in this network, the fused features are input into the GRU layer, allowing the network to perform better in recognizing modulation with long-term dependencies.

Finally, we pass the obtained feature vector to the fully connected layer, normalize it through SoftMax classifier, and output the classification probability. The proposed network uses ReLu activation function for the convolutional layer and tanh activation function for the GRU layer.

4 Results and Discussion

The experiments in this paper use the benchmark modulation recognition dataset Radio ML 2016.10a [3], which contains 11 types of modulated signals, including 8 digital modulations: BPSK, CPFSK, GFSK, PAM4, QAM16, QAM64, QPSK, and 8PSK, as well as 3 analog modulations: WBFM, AM-DSB, and AM-SSB. During the experiments, the data are divided into training set (67%), validation set (13%) and test set (20%), the epoch of training is set to 100, and the batch size is set to 1024. To prevent overfitting, dropout layers are added to the network with a dropout rate of 0.5.

Figure 3a gives the comparison of recognition accuracy with the DL single-input based AMR models under different SNR environments of -20 to 18 dB. The compared models include CNN_IQ, CNN_AP, Inception_IQ, Inception_AP, LSTM_IQ, LSTM_AP, CLDNN_IQ, CLDNN_AP, CGDNN_IQ, CGDNN_AP. We can observe that the recognition accuracy of our network model and the comparison models gradually improves as the signal-to-noise ratio increases. In low SNR environments, the

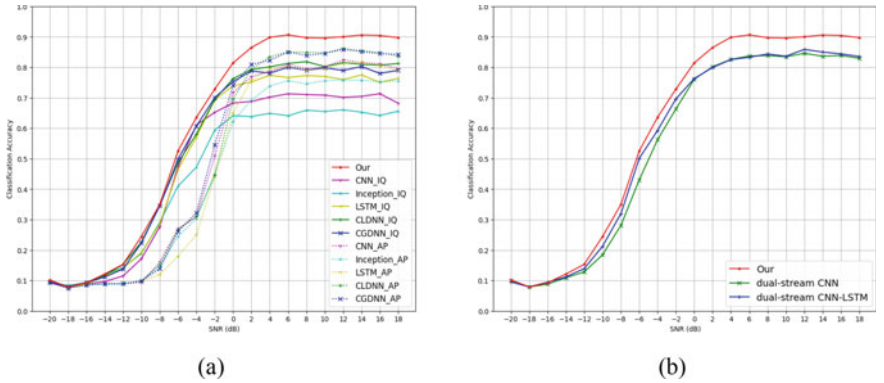


Fig. 3 Signal recognition rates. **a** Signal recognition rates at different SNRs; **b** signal recognition rates of different parallel neural networks

IQ model based on the same deep learning architecture has a higher recognition rate than the AP model. However, in high SNR environments, the AP model has a higher recognition rate than the IQ model, indicating that IQ provides better noise resistance for the model in low SNR environments, while AP has more prominent feature representation in high SNR environments. It can also be found that the CLDNN and CGDNN models, which combine CNN and RNN, perform better than CNN, Inception, and LSTM. Our network model combines these advantages and makes some improvements, achieving a recognition rate of over 80% in high SNR environments above 0 dB, and all of which are higher than the compared AP model. In low SNR environments of − 10 to 0 dB, the recognition rate is respectively 7.15, 11.69, 5.08, 3.32, and 2.84% higher than CNN_IQ, Inception_IQ, LSTM_IQ, CLDNN_IQ, and CGDNN_IQ.

Figure 3b compares the recognition accuracy of two popular parallel neural networks, dual-stream CNN and dual-stream CNN-LSTM. It can be seen that multimodal features of parallel extraction can enable the network to learn more useful information and have good recognition ability in both high and low SNRs. Due to the combination of CNN's excellent performance in extracting spatial features and LSTM's outstanding ability in processing sequential data, dual-stream CNN-LSTM has better performance than dual-stream CNN. The parallel neural network proposed in this paper perfectly integrates the attention mechanism to effectively enhance and suppress features. Therefore, our network is superior to the dual-stream CNN and the dual-stream CNN-LSTM in both recognition rate and noise resistance.

Due to the use of skip connections in our network, which helps with gradient backpropagation and speeds up training. And the use of GRU for extracting temporal features reduces the number of computational parameters. In addition, the lightweight and efficient channel attention network is invoked to re-weight the features and reduce the data redundancy. As shown in Table 1, our parallel neural network outperforms the other two parallel neural networks in terms of computational parameters, training time, and recognition accuracy.

Table 1 Comparison of required parameter count, training time, and average recognition rate among different parallel neural networks

Model	Parameters	Time (ms)	Average accuracy (%)
Dual-stream CNN	3,657,899	322	54.40
Dual-stream CNN-LSTM	1,347,979	360	55.69
Our	265,221	137	59.62

(a) (b) (c)

Fig. 4 Confusion matrices for the proposed network at different SNRs. **a** SNR range of − 20 to 18 dB; **b** SNR range of − 10 to 0 dB; **c** SNR range of 1–18 dB

The confusion matrix is a widely used method for assessing accuracy and performance, where the predicted categories are listed as columns and the actual categories are listed as rows. The squares where rows and columns intersect are darker in color when the accuracy is higher. Figure 4 gives the confusion matrix of the proposed network in all SNR of − 20 to 18 dB, low SNR of − 10 to 0 dB, and high SNR of 1–18 dB environments. Significant differences can be observed in the recognition rates of different modulation schemes. In high SNR environments, most modulation schemes have a recognition rate of over 90%, and some even approach 100%. However, QAM16 may be confused with QAM64, and WBFM can be easily misclassified as AM-DSB. Because their time-frequency spectra are quite similar, neural network models find it difficult to learn effective features that distinguish between them, resulting in some misjudgments in recognition results.

5 Conclusion and Future Work

In this paper, to address the AMR problems such as low modulation recognition rate and weak noise immunity performance, we propose a parallel neural network combining skip-connected CNN with GRU. The proposed network takes into account the IQ and AP difference characteristics of the signal and introduces an attention mechanism to obtain the key information in the signal. The results show that com-

pared with a variety of existing DL-based AMR methods, our network has significant advantages in recognition accuracy and noise immunity performance. Since multiple feature extraction modules often result in higher complexity, we will continue to reduce computational parameters and training time while maintaining high recognition accuracy in future work. We also aim to further improve the AMR performance between certain easily-confused modulation schemes.

Acknowledgements This work was supported by the National Natural Science Foundation of China (61761034).

References

1. Jiao X, Wei X-L, Xue Y, Wang C, Duan Q (2022) Automatic modulation recognition based on deep learning. Comput Sci 49(5):266–278
2. Zheng J, Lv Y (2018) Likelihood-based automatic modulation classification in OFDM with index modulation. IEEE Trans Veh Technol 67(9):8192–8204
3. O'Shea TJ, Corgan J, Clancy TC (2016) Convolutional radio modulation recognition networks. In: Engineering applications of neural networks, pp 213–226
4. Rajendran S, Meert W, Giustiniano D et al (2018) Deep learning models for wireless signal classification with distributed low-cost spectrum sensors. IEEE Trans Cogn Commun Netw 4(3):433–445
5. West NE, O'Shea T (2017) Deep architectures for modulation recognition. In: 2017 IEEE international symposium on dynamic spectrum access networks (DySPAN). IEEE, pp 1–6
6. Zhang Z, Luo H, Wang C et al (2020) Automatic modulation classification using CNN-LSTM based dual-stream structure. IEEE Trans Veh Technol 69(11):13521–13531
7. Huang S, Dai R, Huang J et al (2020) Automatic modulation classification using gated recurrent residual network. IEEE Internet Things J 7(8):7795–7807
8. Qi P, Zhou X, Zheng S et al (2020) Automatic modulation classification based on deep residual networks with multimodal information. IEEE Trans Cogn Commun Netw 7(1):21–33
9. Kumar Y, Sheoran M, Jajoo G et al (2020) Automatic modulation classification based on constellation density using deep learning. IEEE Commun Lett 24(6):1275–1278
10. Wang Q, Wu B, Zhu P et al (2020) ECA-Net: efficient channel attention for deep convolutional neural networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 11534–11542
11. Peng S, Sun S, Yao YD (2021) A survey of modulation classification using deep learning: signal representation and data preprocessing. IEEE Trans Neural Netw Learn Syst 33(12):7020–7038
12. Chang S, Huang S, Zhang R et al (2021) Multitask-learning-based deep neural network for automatic modulation classification. IEEE Internet Things J 9(3):2192–2206

Design of a Radar Moving Target and Clutter Environment Simulator

Zhonglin Wei, Suochang Yang, Yuan Li, and Zhikai Zhao

Abstract This paper introduces the design of a laboratory radar moving target and clutter environment simulator, which can be used for radar function test and performance verification. The simulator uses DDWS (direct digital waveform synthesis) and DDS (direct digital Synthesis) technology, which can generate corresponding simulated target echo RF signals according to the target echo characteristic parameters, and generate interference signals, noise signals and clutter signals in specific frequency bands according to the experimental needs, so as to construct the electromagnetic signal environment for radar testing. The variable difference method model and the least square fitting method model are used to calculate the targets' simulated routes, and the real-time position of the radar feed is controlled by the feed motion controller to realize echo simulation of high-speed moving targets. The BOX-MULLER transform is used to generate Gaussian random sequences, and a high-speed Gaussian noise synthesis method based on FPGA is used to generate high-precision clutter and interference simulation.

Keywords Target simulator · Radar echo · Interference simulation · Moving target

1 Introduction

Radar target and clutter environment simulator is an indispensable tool for debugging radar and testing radar performance [1, 2]. In order to test the function and verify the performance of the radar seeker in the laboratory environment, a radar moving target and clutter environment simulator was designed. By simulating the target echo characteristics and electromagnetic environment characteristics of radar, this

Z. Wei · S. Yang · Y. Li
Army Engineering University of PLA, Shijiazhuang Campus, No. 97, Heping West Road, Shijiazhuang 050003, Hebei, China

Z. Zhao (✉)
The 31668 Army of PLA, 59 Bayi East Road, Xining 810007, Qinghai, China
e-mail: 582389350@qq.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_30

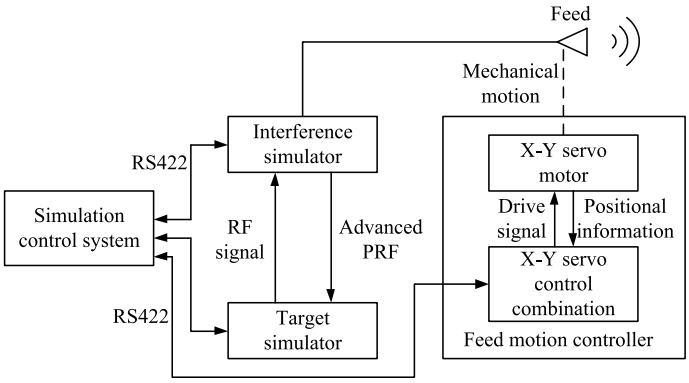


Fig. 1 The composition diagram of simulator

simulator can radiate specific RF signals for conducting radar target search and tracking experiments.

The simulator uses DDWS and DDS technology to simulate target echo signals, and generates interference signals by FPGA and DAC, and simulates high-speed moving targets by controlling the real-time position of the radar feed. It has mature technology, low development difficulty, and low cost.

2 System Consists

The simulator consists of target simulator, interference simulator, feed, feed motion controller, simulation control system and so on. As shown in Fig. 1.

The simulation control system is the control center, which is used to establish and calculate the mathematical model of target echo, clutter and multiple interferences, and then sends control information to target simulator, interference simulator and feed motion controller through RS422 interface. The target simulator generates a simulated target echo radio frequency signal, which is sent to the interference simulator. The RF signal is synthesized with a variety of interference signals (suppression, deception, frequency sweep, etc.), clutter signals and noise signals generated by the interference simulator and sent to the feed, and finally radiated to space.

3 Target Simulator

The target simulator is used to generate radar transmitting signal, target echo signal, synchronization signal and timing signal.

3.1 Radar Echo Signal Model

In a radar timing pulse width, the target motion distance is often relatively small. Therefore, the following assumptions are made that the target Doppler frequency is constant, and the gain of the scanning antenna in the target direction is constant and is only a function of the angle between the target and the beam axis, and the complex reflection coefficient of the target is constant. Radar echo signal can be described as formula (1).

$$X_r(t) = X_T[t - \tau(t)]e^{j2\pi(f_0+f_d)t}\left[\lambda^2/((4\pi)^3 R^4)\right]^{1/2} G(\beta, \varepsilon)\gamma \tag{1}$$

In the formula, $X_r(t)$ is the pulse signal emitted by radar, and $\tau(t)$ is round-trip delay time of the target, and f_0 is the radar microwave frequency, and f_d is the Doppler frequency of the target, and λ is the wavelength of radar, and R is the distance of target, and β, ε is the difference of azimuth angle and pitch angle between target and beam axis, and $G(\beta, \varepsilon)$ is the antenna power gain in the target direction, and γ is the complex reflection coefficient of the target which is used to characterize the fluctuation characteristics of the target echo.

It can be seen from the formula (1) that the radar echo signal is related to many factors such as the waveform of the transmitted signal, the target distance, the frequency (or wavelength), the antenna gain, the fluctuation characteristics and the relative motion characteristics of the target. Therefore, the target simulator must have the precise control ability of signal amplitude, frequency, phase and delay.

3.2 Hardware Design of the Target Simulator

The target simulator adopts DDWS (direct digital waveform synthesis) and DDS (direct digital Synthesis) technology [3, 4], which is mainly composed of reference frequency source, frequency standard components, DDWS source and mixing components, as shown in Fig. 2.

The reference frequency source uses a 120 MHz anti-vibration constant temperature crystal oscillator to generate five 120 MHz reference frequency signals.

The frequency standard component is used to generate 18 frequency points of radar local oscillator and auxiliary local oscillator signals, and adopts a direct, 3×6 frequency synthesis scheme. The 27 / 28 / 29 / 30 / 31 / 32 harmonics of the 120 MHz comb generator are processed by the switching filter component to provide 6 channels of medium frequency standard signals. The 29 / 30 / 31 harmonics of the 360 MHz comb generator are processed by the switching filter component to provide 3 channels of local oscillator frequency standard signals. The intermediate frequency standard signal is used as the intermediate frequency signal of the mixer after the frequency division, and the power is divided into two channels after the down conversion filter amplification of the local oscillator frequency standard signal. One is used as the local

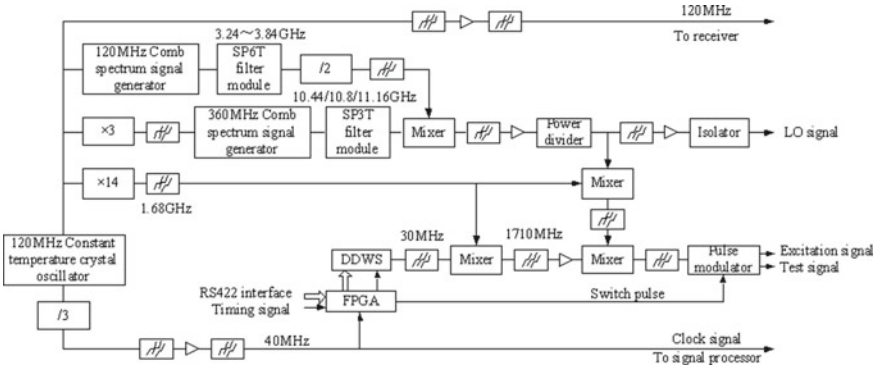


Fig. 2 Principle diagram of target simulator

oscillator signal of the radar, and the other is used as the auxiliary local oscillator of the analog signal source.

The DDWS source is used to complete the intermediate frequency simulation of the echo signal, including radar waveform, distance, Doppler frequency, signal strength, etc. The DDWS source uses MAX5891 as the main chip to receive the target information of the simulation control system, generates an intermediate frequency analog modulation waveform with a frequency of 30 MHz, and then performs up-conversion filtering with the 1680 MHz signal to generate a 1710 MHz analog signal.

The radar waveform adopts digital storage playback technology. The radar waveform data is stored in the FPGA internal ROM, and the data playback process is started under the control of the external trigger signal. In order to simulate the distance characteristics of the echo signal, the advanced PRF is used as the reference. According to the distance information of the track, the advanced PRF signal is accurately delayed, and the signal is used as the trigger signal of the modulation waveform. Doppler frequency is generated by 40 MHz, 28-bit DDS technology. The echo intensity control of the target simulator is realized by the secondary control attenuator. The target simulator adjusts the attenuation of the attenuator according to the difference of the received medium amplitude information to achieve the purpose of simulating the change of echo signal intensity.

4 Interference Simulator

The interference simulator is used to generate interference signals, noise signals and clutter signals of specific frequency bands for suppressing, deceiving and interfering targets. It is mainly composed of local oscillator module, intermediate frequency module and up-conversion module, as shown in Fig. 3.

The local oscillator module uses a high-performance crystal oscillator to generate three 100 MHz coherent clock signals, one of which is sent to the FPGA to do the

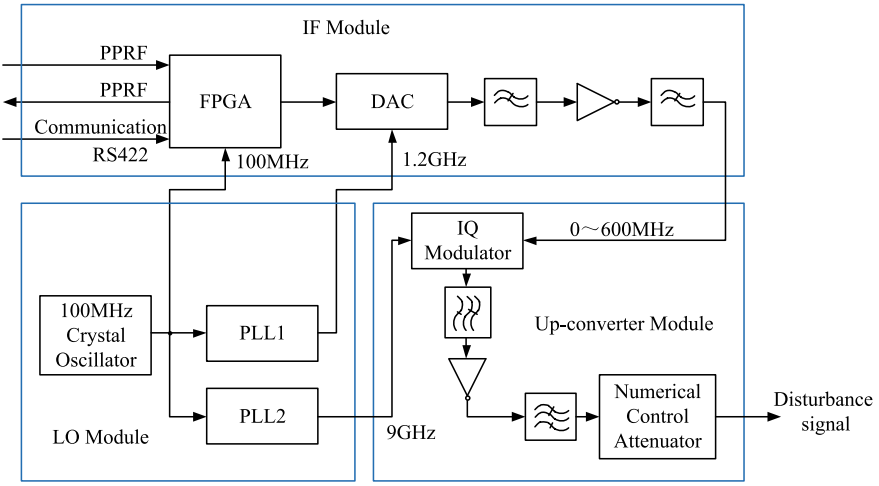


Fig. 3 Principle diagram of reference simulator

clock signal, and the other two are used as the phase discrimination input of the two phase-locked loop circuits. One phase-locked loop circuit generates the clock signal of the DAC chip, and the other phase-locked loop circuit generates the high-stability local oscillator signal of the up-conversion module mixer.

The intermediate frequency module uses FPGA and DAC to generate a baseband intermediate frequency signal of 0–600 MHz. After amplification and filtering, it is sent to the up-conversion module for mixing. The IF module receives the timing signal from the signal processor, and forwards it to the target simulator. The RS422 interface is used for communication control, and the control signals required by other modules in the system are generated.

The up-conversion module mainly amplifies and filters the intermediate frequency signal sent by the intermediate frequency source and the local oscillator signal sent by the local oscillator source, and adds a 31.5 dB digital attenuator at the last stage to adjust the power of the final output interference signal.

The suppression jamming bandwidth ($\geq$ 1 GHz) is wide, and it is designed to directly simulate the IQ modulation implementation, which can reduce the difficulty of baseband implementation. The jamming includes spoofing jamming and suppression jamming. The spoofing jamming implementation approach is similar to the target simulation, including distance deception, speed deception, and signal strength deception. The jamming simulator is only used as the actuator of the deception jamming strategy, and the jamming strategy and model are calculated by the simulation computer. Suppression jamming is divided into two categories, one is the suppression jamming system based on frequency scanning, and the other is the broadband noise suppression system based on high-speed PN code.

5 Feed Motion Controller

The feed motion controller can move the feed to the specified position according to the control command, and simulate the spatial position of the target. It is composed of an X–Y linear module and a servo control combination, as shown in Fig. 4.

The X–Y linear module is composed of an X-axis linear module, a Y-axis linear module, two servo motors, two linear light rulers and a fixed bracket. The X-axis linear module is used to control the feed to move left and right, the Y-axis linear module is used to control the feed to move up and down, the servo motors are the actuator of the motion, and the linear light rulers are used to read the position of the axis motion.

The servo control combination uses ARM chip as the core controller to collect the linear position of the X-axis and Y-axis linear grating ruler, responds to the command information sent by the simulation control system through the RS422 serial communication interface, controls the movement of the X-axis and Y-axis of the linear module, and simulates the movement of the target.

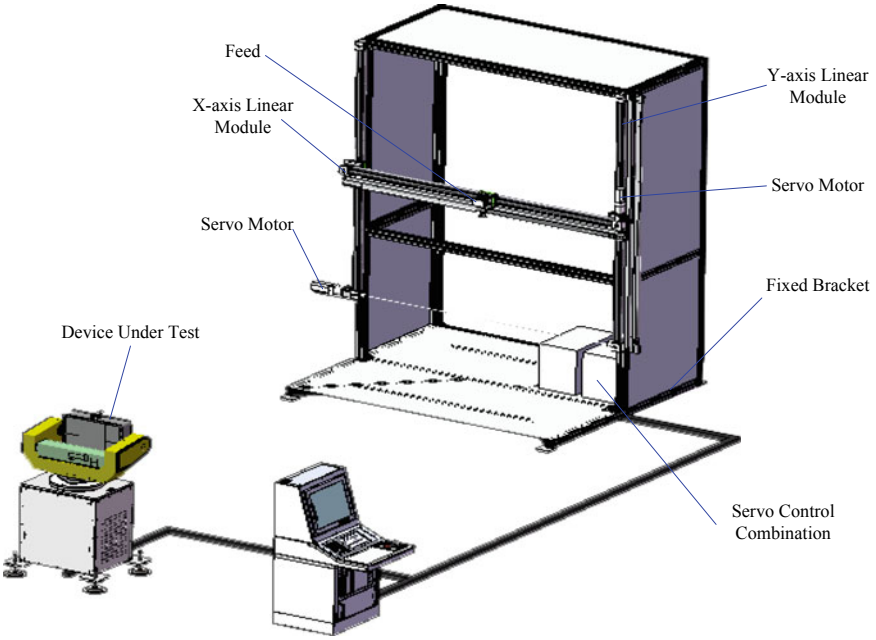


Fig. 4 Structure diagram of feed motion controller

6 Simulation Control System

The simulation control system is responsible for the establishment and control of mathematical models for simulating targets, clutter and multiple interferences, and provides a system control human–computer interaction interface. The simulation control system consists of control parameter setting module, analog route generation module, clutter data generation module, interference data generation module and network communication module.

In order to improve the accuracy, compatibility and flexibility of the clutter signal model, the simulation control system uses the BOX-MULER transform method to realize the Gaussian random sequence. The BOX-MULLER transform method is a commonly used change method for generating high-precision Gaussian random sequences [5, 6]. The specific principles is shown in formula (3).

$$\begin{aligned} V_1 &= -2\ln U_1 \cos(2\pi U_2) \\ V_2 &= -2\ln U_1 \sin(2\pi U_2) \end{aligned} \tag{3}$$

In the formula, U_1 and U_2 are uniform random distribution sequences which obey the interval [0, 1] respectively, V_1 and V_2 are two simple deformed sine and cosine Gaussian random sequences, respectively. Its logic generation principle is shown in Fig. 5.

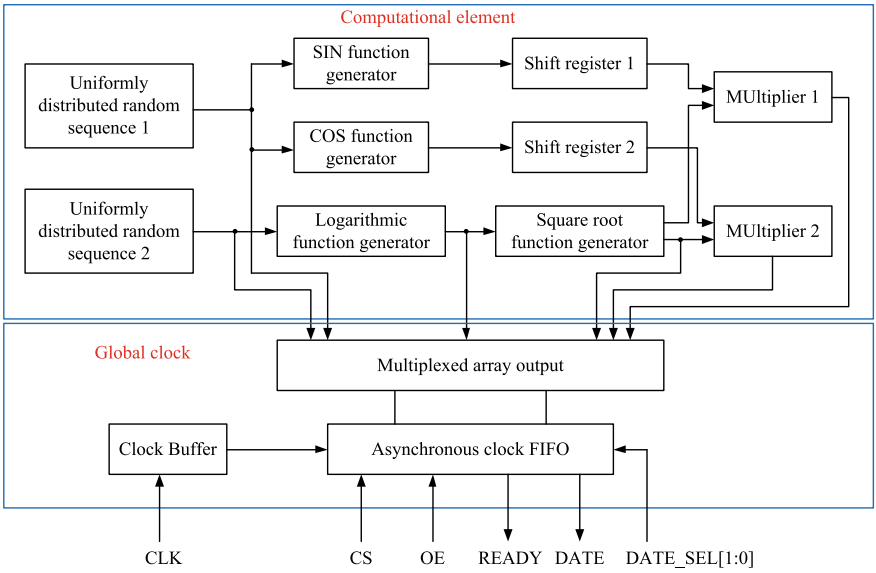


Fig. 5 Schematic diagram of Gaussian random sequence generation

7 Conclusions

The design adopts DDWS and DDS technology, which can flexibly changes the target echo and clutter interference signals according to experimental needs. Through route simulation and feed motion control, it can provide echo signals similar to real moving targets, and conduct in-depth testing on the search and tracking performance of the seeker. This design has been applied to a radar guidance simulation experimental system, and has demonstrated excellent experimental performance.

References

1. Liu Z, Shen FM, Zhang SH (1997) A radar signal simulator of targets and clutter background. J Xidian Univ 24(2):290–295
2. Li BH (2018) A study on design of radar target simulator. Modern Radar 40(2):85–90
3. Zhao ZY, Chang W, Li XY, Tu RB (2012) Pre-distortion for DDWS system. In: 19th international radar symposium. https://doi.org/10.1109/IRS.2012.6233344
4. Tian Y, Shao GP, Xie W (2012) Implementation of DDWS-based complicated waveform generators. J Inf Eng Univ 13(1):54–59
5. Scott DW (2011) Box-Muller transformation. In: Wiley interdisciplinary reviews: computational statistics. https://doi.org/10.1002/wics.148
6. Wang YT, Bie ZS (2014) A novel hardware Gaussian noise generator using Box-Muller and CORDIC. In: Wireless communications and signal processing. https://doi.org/10.1109/WCSP.2014.6992107

Radar Maritime Target Detection Method Based on Decision Fusion and Attention Mechanism

Jurong Hu, Yongruo Sun, Mohammed Mutahar Abduljalil Shujaa Aldeen, and Ning Cao

Abstract This paper presents a decision fusion model based on two-channel convolutional neural network (DF-TCNN) as a way to solve the problems of insufficient feature representation, low performance detection, and poor fault tolerance in radar target detection based on deep learning. In the pro-posed model, a mean strategy is incorporated on each branch's predictions, and a decision fusion algorithm is applied to refine the classification results. Moreover, a dual-channel network structure with an attention mechanism is embedded for feature enhancement and learning adaptation. Verification of the radar data shows that the method offers a high fault tolerance rate and strong anti-interference ability, which can significantly improve radar target detection in the background of complex sea clutter.

Keywords Radar target detection · Decision fusion · Attention mechanisms · Convolutional neural network

1 Introduction

Target detection technology under the background of sea clutter is widely used in military and civilian settings, and it is a radar research hotspot [1]. Sea clutter has non-Gaussian, nonlinear, and non-stationary characteristics [2–4], which lead to mismatches of the statistical model which reduce the target detection performance of the constant false alarm rate (CFAR) method [5]. Contrary to CFAR detections, convolutional neural networks (CNN) [6–9] are data-driven and construct models using deep perceptual networks, which overcom1es the limitation of relying solely on statistical characteristics to simulate clutter distributions [10].

In Ref. [11], radar target detection is defined as a binary classification problem between targets and clutter, which is realized by using the Doppler domain information of the echo signal. To classify maritime targets, clutter, and coastline, CNN is used to classify the segmented maritime radar image samples in [12]. Since the

J. Hu (✉) · Y. Sun · M. M. A. S. Aldeen · N. Cao
Hohai University, Nanjing, Jiangsu Province, China
e-mail: hujurong@sina.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_31

target has a variety of motion characteristics, the Doppler velocity of the clutter unit tends to be large in low sea conditions, and the Doppler spectrum of the target and the clutter overlaps partially, detecting it with only a single feature is unreliable.

In this paper time frequency and amplitude characteristics of sea clutter were extracted using two-channel convolutional neural networks. Based on the viewpoints of feature fusion and feature extraction, the feature vector layer fusion model and the decision layer fusion model are then developed. Similarly, an attention mechanism is simultaneously implemented in the detection of sea clutter objects to improve the model's ability to extract features. For the proposed strategy, simulation analysis is used to determine the most suitable model, taking into account detection performance, parameter amounts, calculation complexity, and detection time.

2 Proposed Method Descriptions

In this section, a two-channel feature extraction network structure was developed, along with a convolutional attention module. The predicted results from each branch were fused at the decision layer, and the target detection problem was transformed into a binary classification problem using the Softmax classifier. The decision threshold was adjusted to effectively control the system's false alarm rate.

2.1 Fusion Attention Mechanism for Feature Extraction Networks

Convolutional attention module (CBAM) [13, 14] is integrated into the feature extraction network VGG16, and end-to-end training is performed together with VGG16 to form an improved feature extraction network. Figure 1 illustrates the network structure. There are two types of CBAM: Channel Attention Modules (CAM) and Spatial Attention Modules (SAM). In order to adapt the features, the two modules infer the attention weights in turn along the channel and space dimensions.

As shown in Fig. 2, an input feature map F is compressed to reduce its spatial dimension, and information about its spatial location is consolidated using average pooling and maximum pooling to create a $1 \times 1 \times C$ feature map: $F^c_{\max}$ and F^c_{avg}. As a result, the multilayer perceptron (MLP) of the hidden layer is applied to the two feature maps, and the output feature vectors are combined element-by-element. As a final step, the Sigmoid activation operation is used to produce a detailed feature map of channel attention A_c.

Formulas (1)–(3) show the calculation process:

$$A_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \tag{1}$$

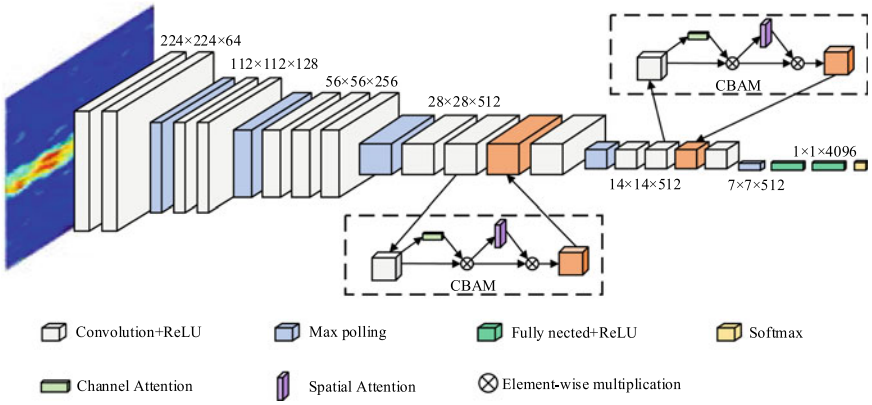


Fig. 1 Network of VGG16 feature extraction with CBAM

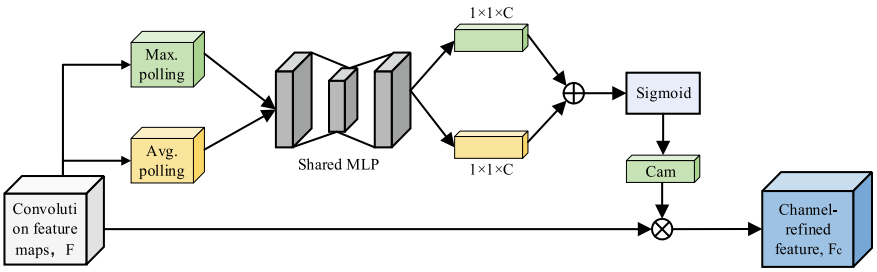


Fig. 2 Process of CAM operation

$$A_c(F) = \sigma\left(W_1\left(W_0\left(F^c_{\max}\right)\right) + W_1\left(W_0\left(F^c_{avg}\right)\right)\right) \tag{2}$$

$$F_c = A_c(F) \otimes F \tag{3}$$

where $Maxpool$ and $Avgpool$ represent average pooling and maximum pooling, W_0 and W_1 represent the two weights of MLP, σ represents sigmoid function, and $\otimes$ represents element-wise multiplication.

Figure 3 shows the operation process, which is the most informative part of the SAM. As a first step, the average pooling and maximum pooling operations will be performed in the channel dimension to obtain two feature maps of $H \times W \times 1$. As a result, two features are spliced together into a feature map of $H \times W \times 2$, and a 7×7 convolution kernel is used to reduce the dimension again into a feature map of $H \times W \times 1$. As a final step, the spatial attention map A_s is generated via the Sigmoid activation method and the final salient feature map F_s is obtained through element-wise multiplication.

Formulas (4)–(6) show the calculation process:

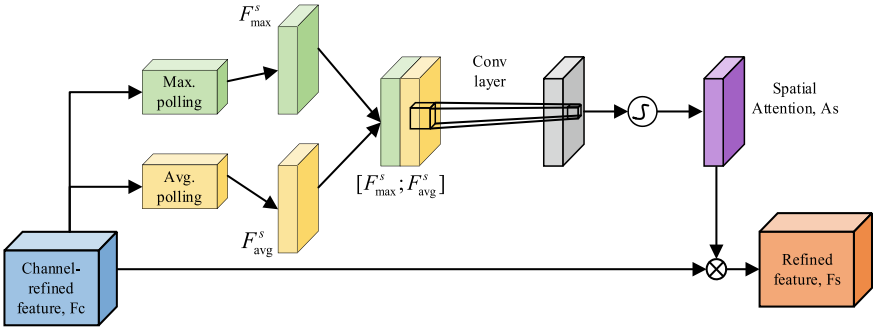


Fig. 3 Process of SAM operation

$$A_s(F_c) = \sigma(Conv_{7\times7}([MaxPool(F_c); AvgPool(A_c)])) \tag{4}$$

$$A_s(F_c) = \sigma\left(Conv_{7\times7}\left(\left[F^s_{\max}; F^s_{avg}\right]\right)\right) \tag{5}$$

$$F_R = A_s(A_c) \otimes F_c \tag{6}$$

2.2 *Algorithm for Decision Fusion*

In this framework, decision-level feature fusion is implemented through two modules: classification and decision fusion. LeNet-5 and VGG16 are used as feature extraction channels in the classification module; the mean fusion algorithm is used in the decision fusion module (Fig. 4).

A VGG16 channel contains two fully connected layers with an output vector size of 4096×1, while a LeNet-5 channel has an output vector size of 120×1. In the last fully connected layer of each network branch, vector probability is calculated using

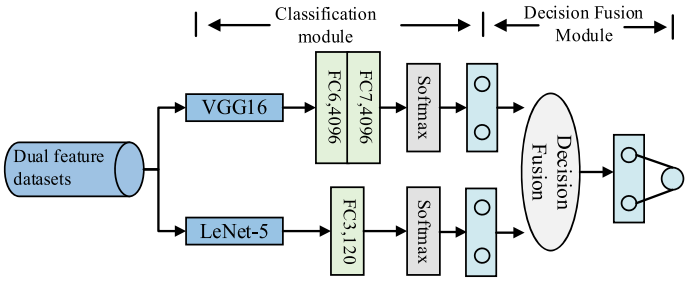


Fig. 4 Model based on fusion of decision layers

the Softmax function. During feature extraction, the weights and parameters of each layer of the trained network are loaded using the transfer learning method.

Each branch of the classification module's prediction results is fused at the decision level in the fusion module, denotes the prediction result of the i-th branch, then there are two branches, so $i \in \{1, 2\}$ is obtained. By using different fusion rules, $P_f = \left(p_{f,1}, p_{f,2}\right)$ can be predicted for the final fusion module. In this paper, the inter-element mean strategy is used for fusion. The calculation rules for the fusion of the j-th element $P_{f,j}$ through the element mean strategy are as follows:

$$P_{f,j} = \frac{1}{2} \sum_{i=1}^{2} p_{i,j} \tag{7}$$

3 Experimentals and Analysis

The proposed method is comprised of three parts: data preprocessing, dataset construction, and model training. During the forward propagation of the dataset in the network model, preliminary prediction results are obtained. Subsequently, in the backpropagation process, the model weights are adjusted by computing the error between the predicted and expected values. This adjustment enables obtaining the optimal network parameter model, facilitating binary classification of targets and clutter.

3.1 Data Set Description and Settings

In this paper, a dual-feature dataset that combines sea clutter amplitude and Doppler velocity is produced for the parallel dual-channel feature network structure described in Sect. 2. IPIX sea clutter data is used for training and testing.

A sea clutter dual-feature image sample is made by splicing and packaging the data obtained from the same signal sequence after different preprocessing methods to ensure that the time–frequency map and amplitude map are identical. Figure 5 displays a double-featured image of sea clutter. The time–frequency features are represented by lines 1–224, while the compressed amplitude features are represented by lines 225–229.

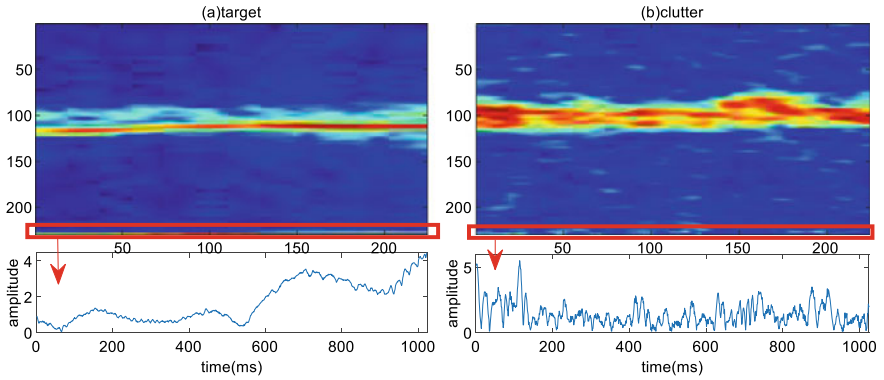


Fig. 5 Sea clutter dual feature dataset

3.2 Evaluations of the Proposed Method

As shown in Table 1, compared with models 1–4, VGG16 achieves higher clutter classification accuracy, while LeNet shows higher target classification accuracy. Moreover, compared with single-channel models 5–7, dual-channel models significantly improve target sample classification accuracy.

In order to verify the detection performance of Model 6 based on the decision fusion algorithm, we designed Model 5, which uses different feature fusion strategies but the same feature extraction channel model. The experimental results show that Model 6 has a target accuracy of 91.27% and a clutter accuracy of 98.33%. Target classification performance improved by 4.45%. Afterwards, the detection probability increased by 1.54% and 0.86% after the CBAM module was introduced in VGG16.

As shown in Fig. 6, Model 7 outperforms Models 5 and 6 if a variable threshold Softmax classifier is used. With false alarm probabilities greater than 10^{-2}, models 5

Table 1 Extraction model performance

Model	Preprocessing	Feature extraction channel	Fusion method	Accuracy (Target)/%	Accuracy (Clutter)/%
1	AMP	LeNet-5	/	78.89	91.12
2	STFT	LeNet-5	/	85.12	94.34
3	AMP	VGG	/	76.52	92.86
4	STFT	VGG	/	84.23	97.67
5	Ch1:AMP Ch2:STFT	LeNet-5 VGG	Feature vector	86.82	98.69
6	Ch1:AMP Ch2:STFT	LeNet-5 VGG	Decision-making layer	91.27	98.33
7	Ch1:AMP Ch2:STFT	LeNet VGG (CABM)	Decision-making layer	92.81	99.19

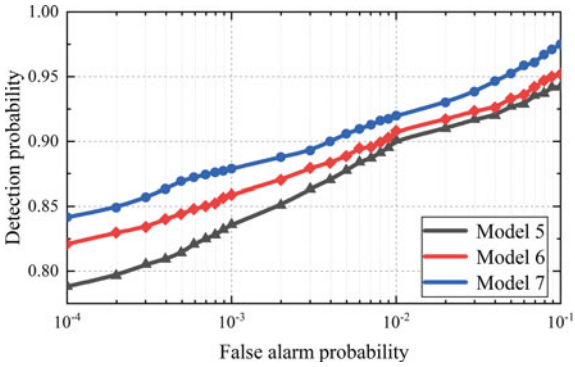


Fig. 6 ROC curves for model 5–7

and 6 display similar detection performance, but both perform worse than model 7. When the false alarm rate is less than 10^{-3}, model 7 can achieve a higher detection probability, and when the false alarm rate is 10^{-4}, model 7 can still achieve a detection accuracy of 84.16%.

3.3 Analysis of Influencing Factors

In practical detection tasks, environmental factors such as wind speed, temperature, and weather can contribute to a complex and variable radar operating environment. Moreover, high-speed acquisition of sea surface information is necessary during detection, requiring control of radar dwell time. The shorter the dwell time, the shorter the observation time. Therefore, this section conducts experimental tests to verify the detection performance of the proposed detector under different observation durations and sea conditions.

Figure 7a presents a comparison of the detector's performance for different observation periods under the HH polarization. The results indicate that an increase in observation time significantly improves the detection probability. Notably, under high false alarm rates, increasing the observation time from 2048 to 4096 ms slightly improves the detector's detection performance. Moreover, when the observation time is only 256 ms and the false alarm probability is 10^{-3}, the detector's detection accuracy can exceed 80%, which can meet the requirements of low observation time, low false alarm probability, and high detection accuracy in practical applications.

Figure 7b shows detection of Class 2, Class 3 and Class 4 sea states with 1024 ms observation time and a false alarm probability of 10^{-3}. Due to the difference between the sea clutter and the Doppler spectrum of the target unit in the third sea state, the network is better able to extract and learn image features based on the differences between those two factors; Nevertheless, in level 2 sea states, the target Doppler spectrum overlaps with the clutter Doppler spectrum, which limits the performance of target detection; Similarly, when the sea state is level 4, backward electromagnetic

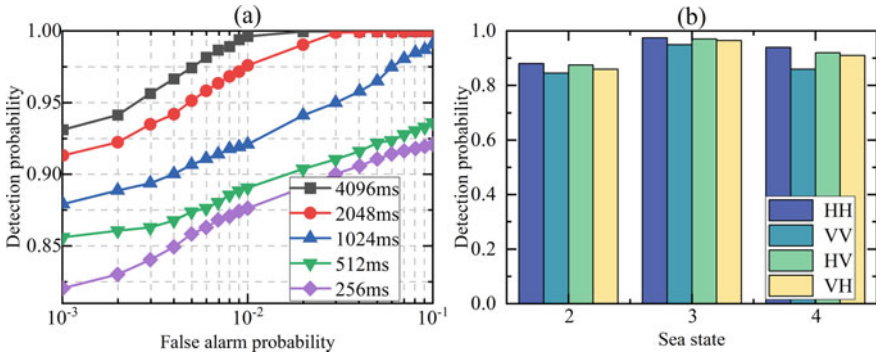


Fig. 7 Model detection figures under different observation durations and sea states

scattering characteristics of sea clutter are strong, and sea peaks due to waves and swells on the sea surface are similar to the target echo and have a high amplitude, making it easier to cover up the target. At this time, the detection probability decreases slightly.

4 Conclusion

As a part of this paper, radar signal target detection is converted into a binary classification problem, and the measured sea clutter and target radar signal data are used to test the performance of different feature extraction models, as well as an attention mechanism-based method.

We process radar signals by short-time Fourier transform and modulo method, VGG16 and LeNet networks are used for feature extraction, and CBAM is fused into the VGG16 network in the decision layer. The detection probability is 87.88% under the conditions of 10^{-3} false alarm rate, and 84.16% under the conditions of 10^{-4} false alarm rate.

References

1. Liu X, Xu S, Tang S (2020) CFAR strategy formulation and evaluation based on Fox's H-function in positive Alpha-Stable Sea clutter. Remote Sens 12(8)
2. Fiche A, Angelliaume S, Rosenberg L, Khenchaf L (2015) Analysis of X-band SAR Sea-clutter distributions at different grazing angles. IEEE Trans Geosci Remote Sens 53(8)
3. Zhihui et al (2017) Analysis of distribution using graphical goodness of fit for airborne SAR Sea-clutter data. IEEE Trans Geosci Remote Sens
4. Conte E, Maio AD, Galdi C (2004) Analysis of distribution using graphical goodness of fit for airborne SAR Sea-clutter data. IEEE Trans Aerosp Electron Syst 40(3):903–918

5. Yuan J, Ma X, Zhang C (2001) Real data based analysis of sea clutter characteristics. J Airforce Radar Acad
6. Oquab M, Bottou L, Laptev I (2014) Learning and transferring mid-level image representations using convolutional neural networks. Comput Vision Pattern Recogn IEEE 1717–1724
7. Chen Y, Jiang H, Li C et al (2016) Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Trans Geosci Remote Sens 54(10):6232–6251
8. Xu C, Hong X, Yao Y et al (2020) Multi-scale region-based fully convolutional networks. In: Intelligent computing and systems (ICPICS), pp 500–505
9. Carrera EV, Lara F, Ortiz M et al (2020) Target detection using radar processors based on machine learning. In: The technical and scientific conference of the Andean Council of the IEEE
10. Oquab M, Bottou L, Laptev I et al (2014) Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1717–1724
11. Wang L, Tang J, Liao Q (2019) A study on radar target detection based on deep neural networks. IEEE Sens Lett 3(3):1–4
12. Hu C (2018) Target recognition for marine radar using deep learning methods. Informatization Res 44(2):63–67
13. Woo S, Park J, Lee JY et al (2018) Cbam: convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV), pp 3–19
14. Yan Y, Kai J, Zhiyuan G et al (2021) Attention-based deep learning system for automated diagnoses of age-related macular degeneration in optical coherence tomography images. Med Phys 48(9):4926–4934

Analyze the Development of ChatGPT Based on Technical Perspective

Yi Sun, Baoju Zhang, Bo Zhang, and Cuiping Zhang

Abstract ChatGPT is a large-scale language model developed by OpenAI that dates back to 2018. Initially, OpenAI developed the GPT-1 model, which was trained using a large-scale corpus and showed excellent performance in natural language processing tasks. The future development of ChatGPT will also continue to promote the technological progress of the dialogue system, to provide people with more convenient, efficient and intelligent services.

Keywords ChatGPT · OpenAI · Natural language processing

1 Introduction

Generative Artificial Intelligence [1] aims to use artificial intelligence technology to automatically generate multimodal data such as text, images, video, audio, and so on, which has attracted wide attention in the field of education. Among them, ChatGPT system shows high application potential in many fields because of its good natural language understanding and generation ability.

In recent years, with the improvement of artificial intelligence technology, computing power and the amount of available data, generative artificial intelligence technology can achieve better content generation effect [2] by relying on language, image and multi foundation model [3] . And in the media, retail, legal, medical, financial and other fields gradually began to provide professional and personalized content generation services.

This paper will take ChatGPT as the main research object and analyze the opportunities and challenges brought by ChatGPT based on its core competence in technical dimension.

Y. Sun · B. Zhang (✉) · B. Zhang · C. Zhang
College of Electronic and Communication Engineering, Tianjin Normal University, Xiqing, China
e-mail: wdxyzbj@163.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_32

2 Overview of ChatGPT

Chat generative pre-training transformer is an artificial intelligence dialogue system developed by OpenAI. Its ability to give clear and detailed answers in many areas of knowledge, and even to write articles close to those written by real people, has gained rapid attention since its launch. What ChatGPT is can be grasped from the following five aspects.

2.1 The External Representation of ChatGPT

The external representation is a chatbot. It can communicate with people by learning and understanding human language [4], and has the ability to answer questions based on the context of the conversation, just like people chat with each other.

2.2 The Actual Nature of ChatGPT

The actual nature is artificial intelligence generation technology. It is a specific application of Artificial Intelligence Generate Content (AIGC) technology. Based on learning human language and related domain knowledge, it has the ability of intelligent content creation and can automatically generate specific content [5].

2.3 The Key Foundation of ChatGPT

The key foundation is a generative large-scale language model. Generative Pre-trained Transformer [6], which is based on generative self-supervised learning to learn implicit language laws and patterns from terabytes of training data and train large-scale language models with hundreds of billions of parameters.

2.4 The Core Technology of ChatGPT

The core technology is instructGPT. It uses Reinforcement Learning with Human Feedback (RLHF) [7] to align the output of AI models with human common sense, cognition, demands, and values.

2.5 The Main Feature of ChatGPT

To a certain extent, it solves the defects of traditional language models in the knowledge utilization, deductive reasoning, deceptive reaction and other aspects of complex and multiple fields, making the answers more useful and authentic, and has the important characteristics of greatly reducing the fabricated facts, generating imitative lies and less wrong content.

3 Analysis of ChatGPT from the Technical Aspect

To achieve high quality generated content, the GPT series system, represented by ChatGPT, involves five key technologies and architectures. They are Transformer model, basic Transformer-based architecture, Reinforcement Learning from Human Feedback (RLHF), Instruction Tuning (IT) technology and Chain of Thought (CT) technology. The scientific and rational use of the above mentioned technologies and basic principles has enabled ChatGPT and other systems to show four outstanding core capabilities in natural language understanding and content generation, inspiring content generation, dialogue context understanding, sequential task execution and program language parsing.

The GPT series system is based on the Transformer model which constitutes its basic system architecture. Since the specific technical structure information of the ChatGPT system has not been fully disclosed, we take its predecessor GPT-3 as an example [8]. As shown in Fig. 1, GPT-3 is mainly composed of 96-layer Transformer decoder, where each layer of the decoder contains a masked multi-headed attention mechanism sublayer and a fully connected feedforward neural network sublayer, word embedding dimension and context window length are extended, and a sparse attention model is used to improve the operation efficiency. The model training process is based on the autoregressive idea [9] that given above content predicts below words or given below content predicts above words. In addition, GPT-3 converts text corpus in different formats for different natural language processing tasks for model training.

ChatGPT in the technical field can help people better understand and apply various technical knowledge [10] . Through natural language generation, it explains and expresses various technical concepts and knowledge, such as programming languages, data science, artificial intelligence, and so on. Moreover, ChatGPT can also help people solve various technical problems by means of intelligent question and answer. However, there are some challenges in the application of ChatGPT in the technical field. Firstly, ChatGPT may not be able to follow and update the relevant technical knowledge in time due to the constant updates and changes of technical knowledge. Second, ChatGPT may have understanding bias or misunderstand-

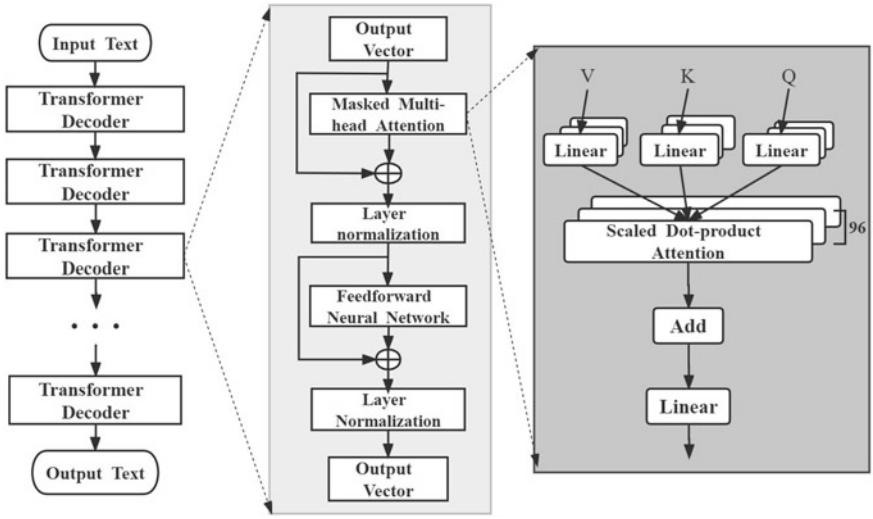


Fig. 1 The basic architecture of the GPT-3 model

ing when explaining and expressing certain technical concepts, which requires the involvement of human technologists to correct. Finally, ChatGPT may require further analysis and processing by human technologists in solving technical problems, especially in some complex and severe technical problems.

4 Conclusion

Artificial intelligence technology represented by ChatGPT has made rapid breakthroughs in recent years, and the related achievements are widely used in different fields, which have a huge impact on various industries in society. The ability to quickly and efficiently learn the knowledge hidden in various data resources is the essence of the rapid progress of AI. The rapid development of AI technology represented by ChatGPT originates from the dramatic improvement of knowledge learning ability, and it has the following implications for improving knowledge learning.

4.1 Change in Solving Problem Approach

Machine learning has changed the model of computer problem solving. Human knowledge characterized by corpus is the key to the rapid breakthrough of machine learning [11]. Machine learning has become an important means of acquiring the knowledge needed to solve problems.

4.2 The Performance Improvement of Deep Learning

All kinds of deep learning models are the foundation, large amounts of computable data resources are the premise, and large-scale computing power is the catalyst. The performance improvement of it is attributed to corpus and computing power in addition to model breakthroughs.

4.3 Rewriting of the Related Techniques

Natural language processing techniques have been rewritten, based on the two-stage learning method of Pre-Training and Fine-Tuning, the natural language processing (NLP) mode is rewritten, and unsupervised pre-training [12] is valuable for knowledge acquisition.

4.4 The Breakthrough of ChatGPT in Learning Ability

Looking back at the history, from the initial GPT-1 model with 117 million parameters, 5 GB corpus, and 12 layers of Transformer to the current ChatGPT model with 175 billion parameters, 45 TB corpus, 96 layers of Transformer, and reinforcement learning using human feedback [13]. Every small progress in the ability of artificial intelligence knowledge learning is valuable, and the continuous progress over time has finally achieved the conversion from quantitative to qualitative change.

Overall, the application of ChatGPT in education, translation, law, technology, and other fields brings unprecedented opportunities and challenges for humanity. While general AI has gradually approached human society, education, as the cornerstone of human civilization progress, we need to be vigilant to avoid blindly pursuing technology and to be able to handle challenges with ease and confidence.

References

1. Askell A, Bai Y, Chen A, Andy J et al (2021) A general language assistant as a laboratory for alignment. Comput Lang 3:8–61
2. Brown TB, Mann B, Ryder N, Prafulla D et al (2020) Language models are few-shot learners. Adv Neural Inform Process Syst 33:1877–1901
3. Blodgett S, Madaio M (2021) Risks of AI foundation models in education. Comput Soc 110:24–32
4. Qingtian X (2020) Unsupervised neural machine translation for multilingualism. J Xiamen Univ 2:192–197
5. Jiayi LY, Zhao LW (2018) A multi-structure and text fusion approach to network characterization. Comput Sci 7:45–48

6. Vaswani A, Shazeer N, Parmar N, Nike P et al (2017) Attention is all you need. Adv Neural Inform Process Syst 30:2702–2712
7. Ouyang L, Wu J, Jiang X, Diogo A et al (2022) Training language models to follow instructions with human feedback. Comput Lang 203:21–55
8. Floridi L, Chiriatti M (2021) GPT-3: its nature, scope, limits, and consequences. Soc Sci Electr Publ 30:681–694
9. Gao F, Yila S, Niu X, Zhao Y (2020) Research on Mongolian-Chinese neural machine translation based on transformer. Comput Appl Softw 225:141–146
10. Xiuquan L (2017) Characteristics and prospects of the new round of artificial intelligence development. High Technol Ind 6:18–19
11. Yu Z, Yu J, Zhu J et al (2022) Multimodal transformer scene text visual quiz with fused knowledge representation. Chin J Graph 27:9–14
12. Baoju Z (2023) Ji yu ji shu jiao du tan tan ChatGPT dai lai de ji yu yu fa zhan [Talk about the opportunities and development brought by ChatGPT from a technical perspective]. J Tianjin Norm Univ
13. Radford A, Wu J, Child R, Luan D, Amodei D (2019) Language models are unsupervised multitask learners. Comput Sci 212:177–192

Design of Fiber Grating Demodulation System Based on Tunable F-P Filter

Tao Li, Tianmin Zhang, Ying Geng, Dingding Zhao, Xidong Ni, and Yinguo Huang

Abstract Aiming at dynamic torque measurement system, fiber Bragg grating sensing principle is used to measure rotating shaft torque, and a fiber Bragg grating demodulation system based on tunable F-P filter is designed. Based on the influence of hysteresis and creep of piezoelectric ceramics, a tunable F-P filter is calibrated with a standard to locate the central wavelength reflected by fiber Bragg grating. In this paper, a photoelectric conditioning circuit for fiber Bragg grating demodulation is designed. The experimental results show that this method can accurately demodulate fiber Bragg grating wavelength, which is a preparation for the subsequent research.

Keywords Fiber Bragg grating · Tunable F-P filter · Etalon

1 Introduction

Fiber Bragg grating (FBG) has the advantages of small size, high sensitivity, strong anti-interference ability and mature production technology, and has been widely used in the field of optical fiber communication and sensing. Since the resonant wavelength of fiber Bragg grating is sensitive to the changes of external environment such as temperature, strain and concentration, fiber Bragg grating sensors have been developed rapidly in recent years and have been widely used in civil engineering, environmental monitoring and transportation. With the rapid development of fiber Bragg grating sensing technology, various demodulation modes of fiber Bragg grating sensing signals have emerged, including intensity demodulation, phase demodulation, frequency demodulation, polarization demodulation and wavelength demodulation. Among them, wavelength demodulation technology is to obtain sensing information through the modulation of external physical parameters

T. Li · D. Zhao · X. Ni
Shanghai Marine Equipment Research Institute, Shanghai, China

T. Zhang · Y. Geng · Y. Huang (✉)
Tianjin University, Tianjin, China
e-mail: hyg@tju.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_33

of fiber Bragg wavelength. Therefore, it is an absolute measurement based on wavelength coding, which has the advantages of high feasibility, strong anti-interference ability and stable performance. Thus it has been widely used in fiber Bragg grating demodulation technology.

Because the optical signal received is weak due to the rotation of the rotating shaft during dynamic torque measurement, weak signal demodulation is needed. In this paper, the photoelectric conditioning circuit is designed, the standard equipment is used to calibrate, and the wavelength of fiber Bragg grating can be demodulated accurately.

2 Demodulation System

2.1 Principle of Tunable F-P Filter

Tunable F-P filter demodulation is one of the most widely used demodulation methods in fiber Bragg grating demodulation at present. It is mainly composed of two symmetrical convex lenses, two precise parallel mirrors and piezoelectric ceramic actuators [1]. The structure is shown in Fig. 1.

L1 and L2 represent two lenses, M1 and M2 represent two mirrors parallel to each other, and the distance d represents the length of the F-P cavity. When the piezoelectric ceramic is driven by an external signal, the piezoelectric ceramic will vibrate with the driving signal, making the M1 mirror move left and right, while the M2 mirror does not move, thus changing the cavity length of the F-P cavity. When the light source is incident, the incident light is reflected by the resonant cavity, and then the lens M2 converges and transmits the light. It is proved by theoretical studies that the transmission wavelength of the tunable F-P cavity is as follows [2]:

$$\lambda = \frac{2nd}{K} \quad (K = 1, 2, 3 \ldots) \tag{1}$$

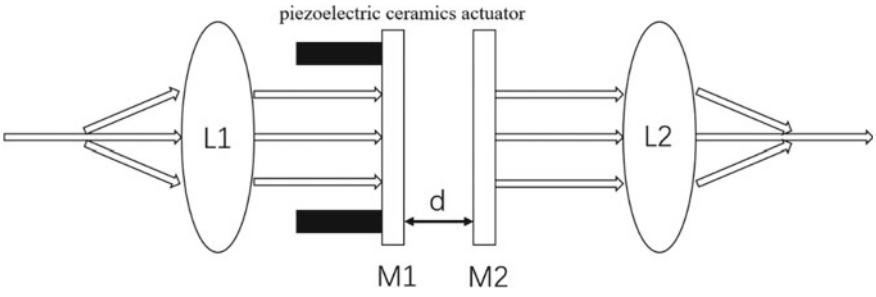


Fig. 1 Schematic diagram of tunable F-P filter

It can be obtained from the above formula that the transmission wavelength of the filter is proportional to the length of the cavity d, so the transmission wavelength of the filter can be changed by changing the length of the cavity.

2.2 Design of Demodulation System

When the filter is connected with the fiber Bragg grating, in order to maximize the output light intensity of the filter, only the center wavelength of the fiber Bragg grating and the transmission wavelength of the filter need to be consistent, so we can use this matching relationship to demodulate the wavelength of the fiber Bragg grating. When the central wavelength of fiber Bragg grating changes due to strain, the piezoelectric ceramics can be controlled by sawtooth wave or triangle wave to adjust the cavity length. When the transmission wavelength is consistent with the central wavelength, the maximum optical power can be obtained, and the driving voltage at this time can reflect the central wavelength of fiber Bragg grating. The overall demodulation system diagram is as follows (Fig. 2).

The transmission wavelength of the F-P filter is controlled by the voltage of the driving circuit. The system uses periodic triangular wave for scanning. Two peaks will be generated within one cycle, and the voltage corresponding to the two peaks corresponds to the corresponding transmission wavelength, thus corresponding to the central wavelength of the sensor grating. Because the traditional F-P filter demodulation method is affected by hysteresis and creep in the moving process of

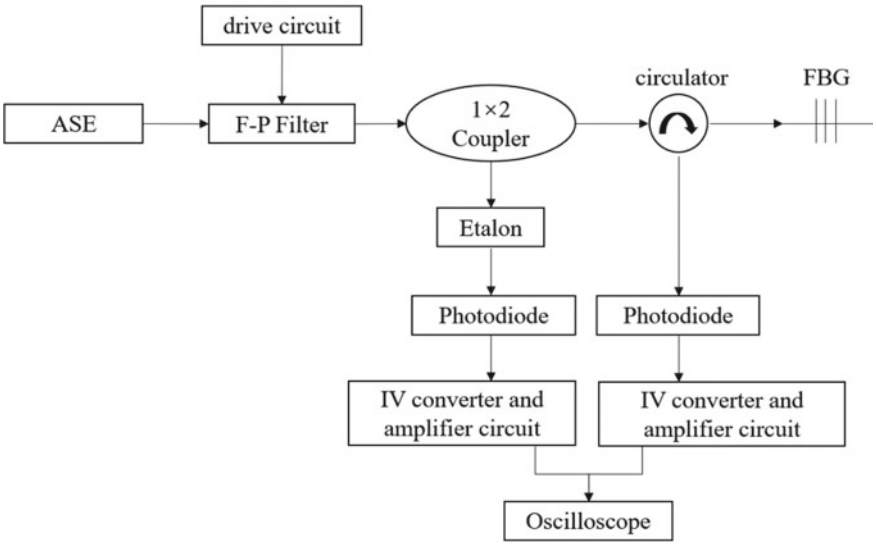


Fig. 2 Schematic diagram of demodulation system of tunable F-P filter

piezoelectric ceramics, its linearity is poor, so this system introduces optical fiber standard to calibrate F-P filter [3–7]. The output light of the broadband light source passes through the filter and becomes a narrow-band light. The wavelength of narrow-band light is related to the voltage of the driving circuit. The optical path is divided into two ways through the 1 × 2 coupler [8]. By calibrating the wavelength-voltage relationship of the filter with the transmission peak of the standard, the relationship between the central wavelength of the fiber Bragg grating and the driving voltage of the filter can be obtained, and the wavelength demodulation can be realized [5].

2.3 Hardware Circuit Design

The main hardware circuits include the F-P filter [9–12] driver circuit and the photoelectric conversion circuit. Since the optical fiber standard is used for calibration in this paper, the driving voltage of the filter is not high [13]. The signal generator is directly used to generate the triangular wave with a frequency of 10 Hz, and the amplifying circuit is used in the same direction. Triangular waves before and after amplification are shown in Fig. 3.

In order to realize the conversion of photoelectric signal, this paper adopts the InGaAs photodiode [8] of Beijing Min Light Company for measurement, the selected model is LSIPD-A75, the wavelength range is 800–1700 nm [14], when the wavelength [10] is 1550 nm, the sensitivity is 0.90 mA/mW, when the reverse voltage is 5 V dark current is 18 pA. Photodiode converts optical signals into electrical signals

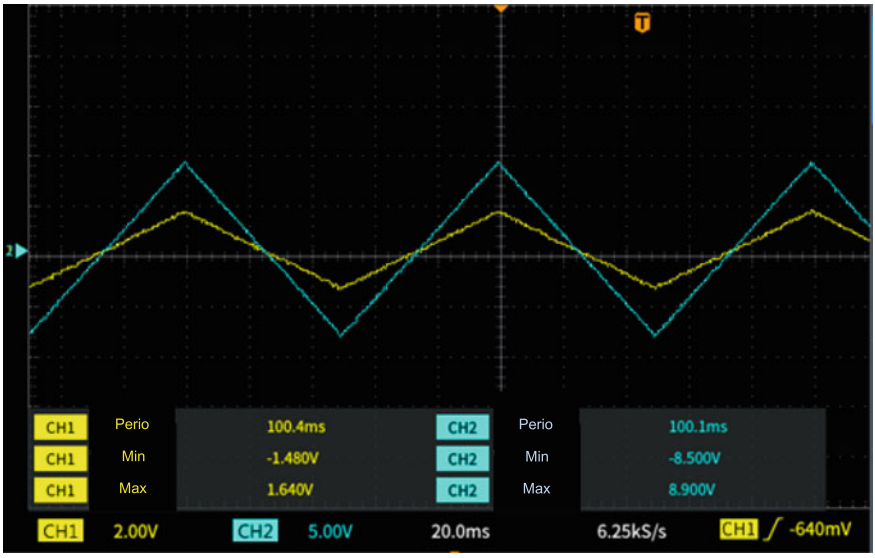


Fig. 3 Filter drive voltage

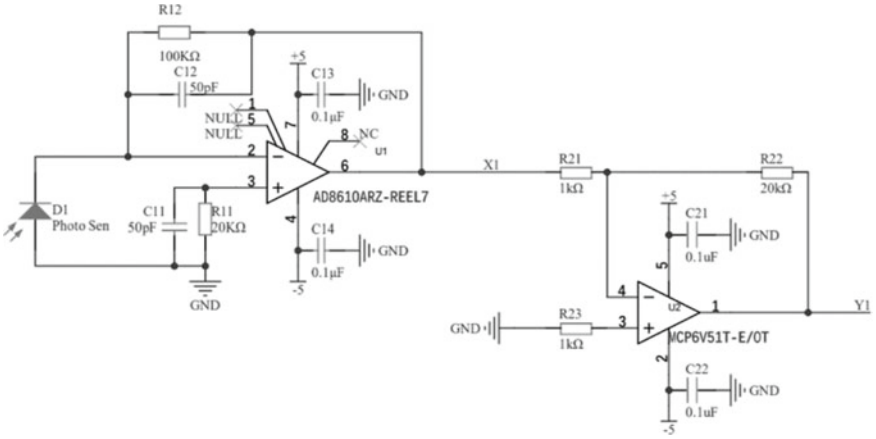


Fig. 4 Photoelectric conversion circuit

and outputs weak current signals. In this paper, IV conversion circuit and secondary amplifier circuit are used to process current signals, with a total amplification factor of 2 M. The design circuit diagram is shown in Fig. 4.

3 Result

The optical path structure of the whole system is shown in Fig. 5.

The output results of the etalon and the circulator measured by oscilloscope are shown in Figs. 6 and 7 respectively.

The triangular wave in the figure is the driving voltage of the F-P filter, the comb wave peak is the transmission peak of the standard, and the two narrow-band peaks are the reflection peak of the fiber Bragg grating. It can be seen from the figure that there are two marker points within a period of the standard, both of which are positions with wavelength of 1550 nm. The relationship between wavelength and voltage can be calibrated with the standard, and then the position voltage of the reflection peak of the fiber Bragg grating can be substituted into the relationship to obtain the wavelength of the fiber Bragg grating, thus achieving wavelength demodulation.

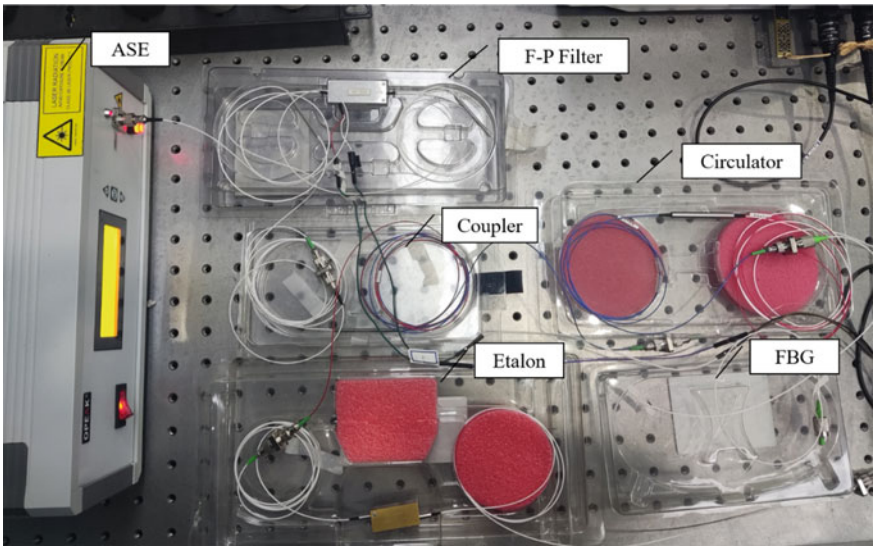


Fig. 5 Actual picture of the experimental system

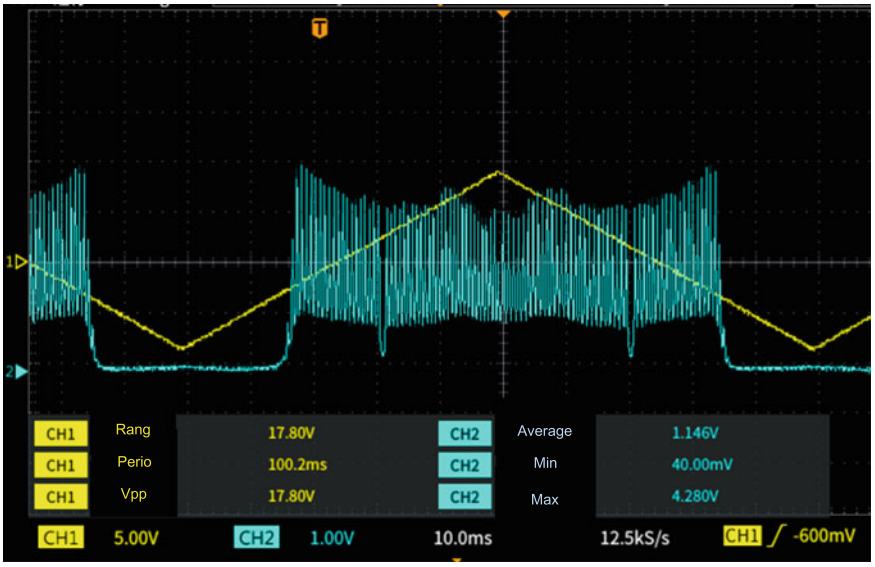


Fig. 6 Etalon output result

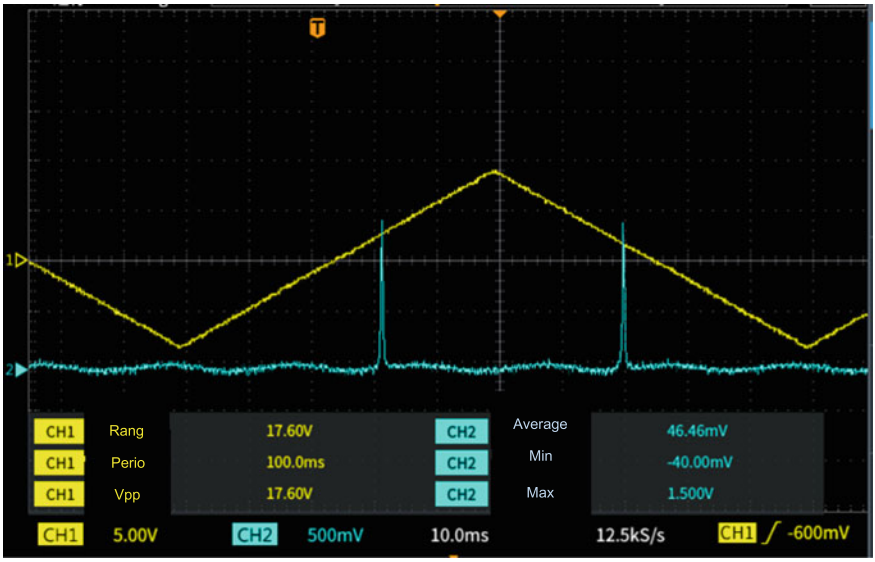


Fig. 7 Circulator output result

4 Conclusion

In this paper, the demodulation system of fiber Bragg grating is designed by using the demodulation method of tunable F-P filter, and the effect of hysteresis and creep of piezoelectric ceramics is reduced by adding a standard for calibration. The photoelectric conditioning circuit is designed to realize the demodulation of wavelength signals. Based on dynamic torque measurement, this paper provides the basis for the realization of understanding modulation. Subsequently, it is necessary to increase the filter scanning frequency and coordinate with the rotating shaft speed to realize dynamic demodulation.

References

1. Fu H, Ping K, Qiao X et al (2013) Research on linear drive of piezoelectric ceramics based on F-P cavity. Piezoelectricity AcoustoOptics 35(2):226–228
2. Wang Y, Huang J (2015) Optical fiber sensing technology. Xidian University Press, pp 78–90
3. Wen F, Zhang Y, Jia X (2022) Design of fiber Bragg grating demodulation system based on F-P tunable filter. Electronic Measurement Techn 45(09):38–43
4. Liu C, Sun Z (2021) Design and fabrication of a metallic irregular F-P filter array for a miniature spectrometer. Appl Opt 60(16):4948–4953
5. Han X, Wang P, Cai SW et al (2014) Design of tunable FP filter driver based on FPGA[C]. Appl Mech Mater 602:2680–2683 (Trans Tech Publications Ltd)
6. Gao H, Yuan S, Bo L et al (2008) InGaAs spectrometer and FP filter combined FBG sensing multiplexing technique. J Lightwave Technol 26(14):2282–2285

7. Tunable F-P filter with dual transmission peaks suitable for demultiplexing in mm-wave radio-on-fibre dense WDM systems
8. Han M, Xu A, Tang S et al (1996) Wavelength control precision using FP filter in a WDM or FDM system. Fiber Optic Components Optical Commun SPIE 2893:60–63
9. Cui Y, Xiao F, Ma X et al (2021) All-optical clock extraction and wavelength conversion based on FP filter and cross gain modulation for WDM-PON. In: International conference on laser, optics and optoelectronic technology (LOPET 2021). SPIE, vol 11885, pp 245–248
10. Chen M, Liu X, Li H (2001) Tunable FP optical filter based on microcavity structure. Adv Microelectronic Device Technol SPIE 4600:170–174
11. Yang G, Guo J, Xu G et al (2014) A novel fiber Bragg grating wavelength demodulation system based on FP etalon. Optoelectronic Devices Integr V SPIE 9270:142–148
12. Feng H, Deng ZL, Luo XH (2002) The simulation on the microelectromechanical tunable FP optical filter. Int J Nonlinear Sci Numer Simul 3(3–4):585–588
13. Fang L, Cheng X, Peng X et al (2017) Resolution improved monochromator based on tunable Fabry-Perot filter and grating hybrid modulator. In: Selected papers of the Chinese Society for optical engineering conferences held October and November 2016. SPIE, vol 10255, pp 1217–1222
14. Atherton PD, Reay NK, Ring J et al (1981) Tunable fabry-perot filters. Opt Eng 20(6):806–814

Remote Monitoring and Early Warning System of Subway Construction Based on Wireless Sensor Network Technology

Liang Zhao, Qu Wang, Meixia Fu, Jing Gao, and Hanming Wei

Abstract To enhance the intelligence level of monitoring risk sources in the process of subway construction, this system deploys a variety of sensors in the target area using wireless sensor network technology. Relevant data from the target area is collected through these sensors, which serve as data reference for monitoring risks during the subway construction. The collected data is transmitted to the management terminal in real-time through wireless network and is analyzed and processed for decision-making assistance. This remote monitoring system enables administrators to keep track of the construction project at all times and provide timely feedback to adjust ongoing efforts. Ultimately, the analysis and processing of the collected data provide powerful support and assistance for decision-making. In summary, this subway construction monitoring system has a high intelligence level, which greatly improves the safety and efficiency of subway construction and allows administrators to be always informed of the target area's conditions for timely decision-making and adjustments, guaranteeing the smooth progress of subway construction.

Keywords Wireless sensor network (WSN) · Subway · Remote monitoring early warning · Sensor

1 Introduction

As urbanization accelerates in different countries, infrastructure construction like subways is experiencing rapid development. The mileage of subway construction is increasing in every country, but the number of accidents occurring during the subway construction period is also increasing, and the harm caused by these accidents is becoming increasingly significant. For comprehensive and systematic monitoring of

L. Zhao (✉) · J. Gao · H. Wei
Shanxi Information Industry Technology Research Institute Co., Ltd, Taiyuan 030012, China
e-mail: zhaoliang0817yu@163.com

Q. Wang · M. Fu
School of Automation and Electrical Engineering, Institute of Industrial Internet, University of Science and Technology Beijing, Beijing 100083, People's Republic of China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_34

the support structure in the foundation pit and the surrounding soil during subway construction. This is extremely important to adjust construction methods in a timely manner to ensure the safe and smooth progress of the project [1, 2]. When risks are identified during the construction process, management personnel can be alerted promptly, and emergency measures can be quickly activated to minimize losses caused by risks [3]. By collecting data on the construction process of subway tunnels and deep foundation pits and analyzing the change patterns, we can not only ensure the safety of subway construction, but also provide reference for subsequent subway construction. Therefore, it's become an important research direction for subway construction to use intelligent methods to improve the level of safety risk control and strengthen safety risk management [4, 5].

With the rapid development of Internet of Things technology, wireless sensor networks have high applicability and potential in subway construction. It is a self-organizing network system formed by many sensor nodes deployed in the monitoring area to communicate with each other in a multi-hop manner. It can obtain target information in any location, time, and environmental conditions. We apply it to the monitoring of safety risks in subway construction. By deploying multiple types of sensors in the target area to collect target information, and using wireless communication networks to transmit the data collected by the sensor nodes to the management terminal, managers can real-time understand the construction progress of the subway project on their PCs and mobile devices [6, 7]. Managers can simultaneously see different types of data collected from different locations by different sensors, which facilitates comprehensive judgment on the construction progress of the subway project. The application of wireless sensor network technology can not only reduce the human cost involved in subway construction but also enable more comprehensive monitoring of subway construction. It also promotes the intelligence level of subway construction and ensures the safety of the subway construction process [8, 9].

2 System Structure Design

The wireless sensor network system consists of sensor nodes, aggregation information nodes, management terminal nodes, and a wireless communication network, as shown in Fig. 1 [10, 11]. Multiple types of sensors can be connected to the sensor nodes, which possess the functions of data acquisition, storage, wireless communication, and more. The aggregation nodes collect information from multiple sensor nodes and forward it. When the management terminal node receives the information from the aggregation node, administrators can remotely understand the construction situation of the subway construction area on their PC or mobile device [12].

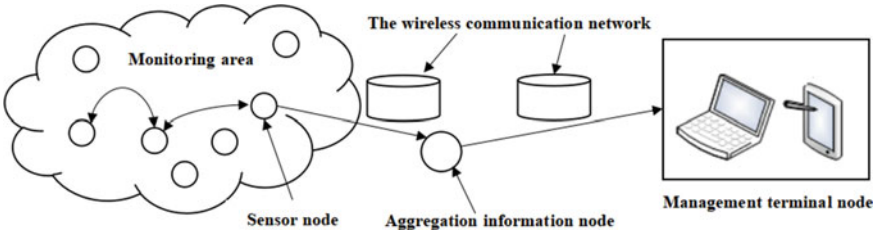


Fig. 1 Schematic diagram of wireless sensor network system structure

3 Hardware Design

3.1 Sensor Node Design

The sensor node is composed of a data acquisition module, an analog to digital conversion module, a storage module, a wireless communication module, a control module, and a power supply module [13, 14]. The structure of the sensor node is shown in Fig. 2. The data acquisition module is composed of various sensors that collect information from the monitoring area of the subway construction. The analog to digital conversion module converts the analog signal collected by the data acquisition module into digital signal. The control module manages the control of each module. The storage module stores the digital signals. The wireless communication module sends digital signals wirelessly. The power supply module supplies all the modules with power.

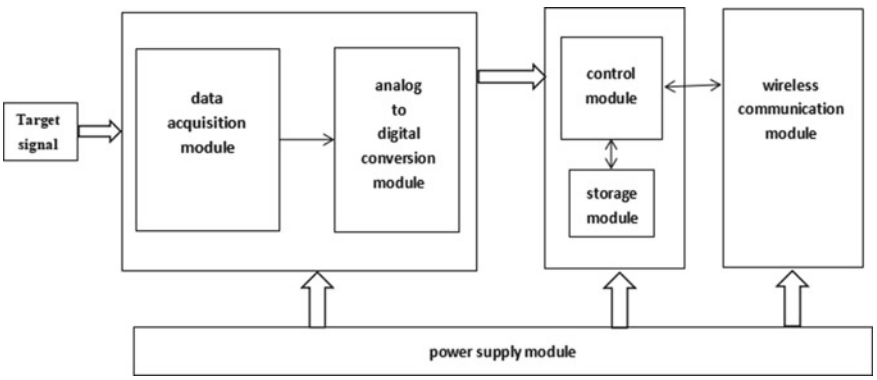


Fig. 2 Structure of the sensor node

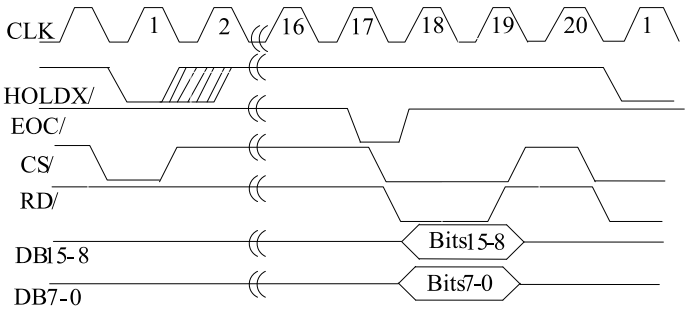


Fig. 3 The timing diagram of the analog to digital conversion

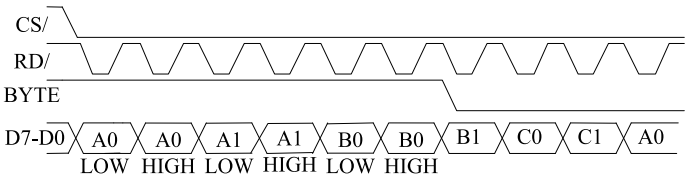


Fig. 4 The timing diagram of the data reading

3.2 Acquisition Module Design

Firstly, the data acquisition module collects target information through various types of sensors. Then, the signal collected is adjusted by a second-order low-pass filter circuit. Finally, the processed signal enters the analog to digital conversion module to convert the analog signal into a digital signal. The A/D selected in this design is a 6-channel high-speed analog-to-digital converter with a 16-bit sampling accuracy. The timing diagram of the analog to digital conversion and the data reading are shown in Figs. 3 and 4, respectively.

3.3 The Use of the Sensor

The system mainly uses vibrating wire piezometers to monitor underground water level, steel bar meters and axial force meters to monitor support axial force, inclinometers to monitor the displacement of the pile body, and total stations to monitor the horizontal displacement of the pile top [15].

The sensor for measuring the water level, a vibrating wire piezometer, mainly consists of permeable stones, pressure-bearing film, steel strings, fastening clamps, and a coil [16]. Under certain conditions of chord length and force, the natural frequency of the steel strings is fixed. When the chord length is constant, the square of the natural frequency of the steel strings is proportional to the tension of the

string. When the external water pressure acts on the pressure-bearing film of the vibrating wire piezometer, causing it to undergo a slight deformation, the tension of the steel strings connected to the pressure-bearing film changes, and the natural frequency of the strings also changes accordingly. The square of the natural frequency of the steel strings is proportional to the pressure of the water on the membrane. By measuring the change in the steel string frequency, the size of the measured pore water pressure can be obtained. The vibrating wire piezometer can obtain the water pressure value through the change in steel string frequency. The function for monitoring underground water level using a vibrating wire piezometer is as follows:

$$P = K\left(f_i^2 - f_0^2\right)$$

where P represents pore water pressure. K is the calibration coefficient of the pore water pressure gauge being measured. fi is the average frequency value of the pore water pressure gauge during measurement. f0 is the initial frequency average value of the pore water pressure gauge before installation.

4 Software Design

Based on previous research, the following software process is designed. Firstly, the risk knowledge base is established, based on the risk sources that have caused accidents during the subway construction process. Figure 5 shows the software design process diagram of the system. Then, experts are organized to assess the possible risk sources at specific locations and create corresponding risk mitigation plans. Subsequently, if the collected values during the sensor collection process exceed the preset warning requirements of the system, a risk warning is triggered. If the risk source that triggers the warning exists in the risk knowledge base, the relevant personnel will be informed to eliminate the risk according to the corresponding risk mitigation plan. If the risk source has no corresponding risk mitigation plan in the knowledge base, the relevant personnel will be informed to organize the experts and develop a corresponding risk mitigation plan. Finally, if the risk is eliminated with no accident, the warning is removed; if an accident occurs due to the risk, a risk accident report will be generated.

In the subway construction project, based on the abnormal changes in the monitoring data of the foundation pit, tunnel, and surrounding environment, the warning is divided into three levels: yellow warning, orange warning, and red warning.

(1) Yellow monitoring warning: when the change in monitoring data and the rate of change exceed 70% of the control value, or when the change in monitoring data or the rate of change exceeds 85% of the control value.
(2) Orange monitoring warning: when the change in monitoring data and the rate of change exceed 85% of the control value, or when the change in monitoring data or the rate of change exceeds the control value.

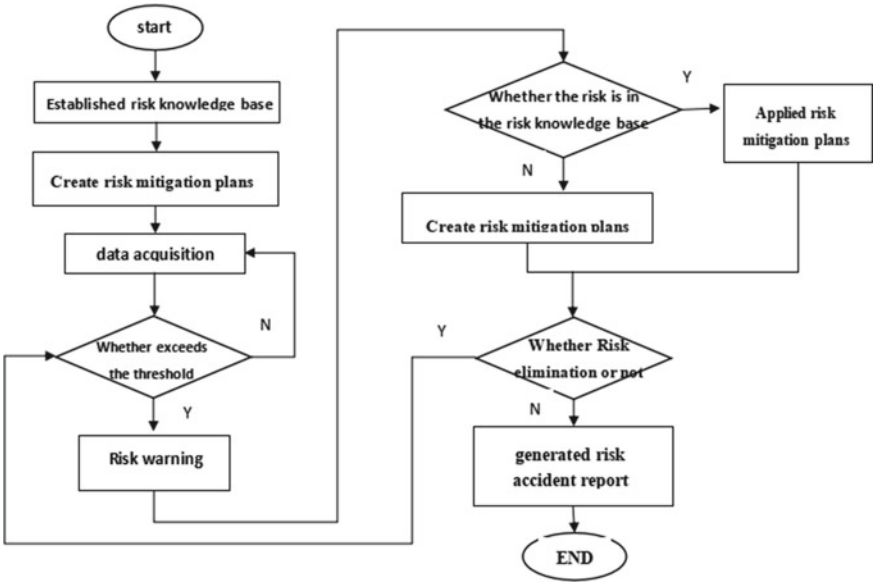


Fig. 5 System software design flow chart

(3) Red monitoring warning: when both the change in monitoring data and the rate of change exceed the control value.

5 Experimental Results

Deploying sensor nodes in the subway construction monitoring area to collect target information and transmitting relevant information to the management terminal for display. Figure 6 shows the water level change trend chart of Four-hole well, which presents a multi-variable display of the water level change trend of Four-hole well within the same period of time. All the data collected by these sensors are collected from within the same monitoring area. The combination of single-variable display and multi-variable display of the collected data solves the problem of untimely and inaccurate judgment of a particular area by a single sensor or a single type of sensor, and the difficulty in the timely discovery and control of risk sources.

6 Conclusions

The system utilizes wireless sensor network technology, deploying various types of sensor nodes within the subway construction monitoring area to collect target information, which enables managers to understand the real-time data information

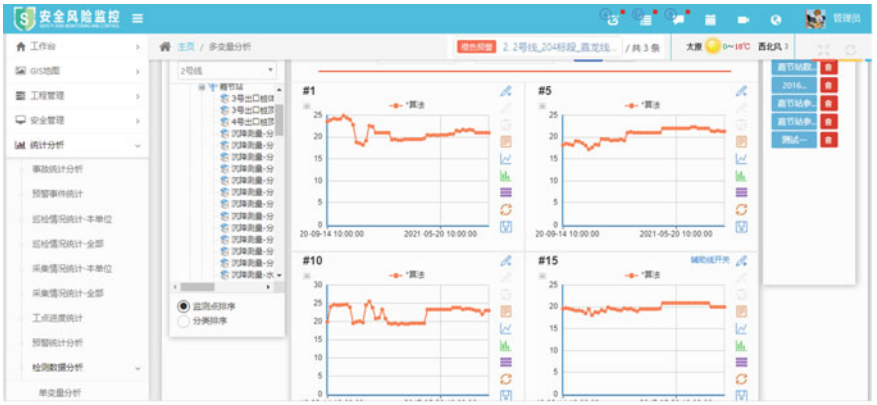


Fig. 6 Multi-variable display of the water level change chart

of various monitoring points in the subway construction process from a distance. The system is capable of monitoring, data transmission, calculation, analysis, and warning of target data, allowing managers to assess and forecast the subway construction realistically, provide corresponding instructions, and remotely monitor the subway construction process. The system provides a powerful guarantee for safety in subway construction.

Acknowledgements This study was funded by the Shanxi Provincial Key Research and Development Project (Project Number 202101020101020), the Shanxi Provincial Basic Research Program (Project Number 20210302124622), and the Shanxi Provincial Enterprise Technology Center Construction Project.

References

1. Zhang X, Liu J, Chen J (2021) A review of wireless sensor network-based intelligent monitoring systems for subway tunnels. Sensors 21(8):2839
2. Wang J, Gao Z, Li J, Ding X (2019) Real-time wireless sensor network monitoring system for railway tunnel deformation. Sensors 19(16):3595
3. Zeng S, Xie J, Lei Y (2020) Application of wireless sensor network in early-warning system of tunnel fire. In: 2020 international conference on artificial intelligence, information processing and cloud computing (AIIPCC). IEEE, pp 159–163
4. Jiang Y, Li J (2018) Subway tunnel structural health monitoring based on wireless sensor networks. Adv Struct Eng 21(5):732–744
5. Yang L, Lai S, Chen W (2021) A novel method for subway tunnel structural health monitoring based on fiber optic sensing technology. In: Proceedings of the 2021 international conference on artificial intelligence and smart systems, pp 212–217
6. Wu H, Guo D, Zhang X, Huang Y (2019) A wireless sensor network-based structure monitoring system for urban subway tunnels. IET Wireless Sens Syst 9(6):298–305
7. Zhang X, Liu J, Chen J, Wang T (2018) Fault diagnosis and early warning system for subway mechanical equipment based on wireless sensor networks. J Phys: Conf Ser 1063:032037

8. Cui X, Li Z, Shi L (2019) Performance evaluation of wireless sensor networks for rail track deflection monitoring. J Ambient Intell Humaniz Comput 10(5):1961–1970
9. Wu H, Guo D, Chang X, Song J (2019) A remote monitoring system for subway structures using wireless sensor networks. J Sens
10. Liu M, Wen C, Bai J, Lu H (2020) Structural health monitoring of subway based on wireless sensor network and machine learning. IEEE Sens J 20(24):15076–15085
11. Zhang C, Liu J, Xie L et al (2021) Age-optimal data gathering and energy recharging of UAV in wireless sensor networks. In: The international society for applied computing.proceedings of 2021 3rd international conference on advanced information science and system (AISS 2021). ACM, vol 6. https://doi.org/10.26914/c.cnkihy.2021.052260
12. Chueh YL, Hsiao YT, Hsieh SH (2019) An early-warning system for tunnel fire hazard based on a heterogeneous wireless sensor network. IEEE Sens J 19(23):10962–10973
13. Su L (2017) Research on a dynamic sampling scheme of wireless sensor node. Information Engineering Research Institute, USA, Singapore Management and Sports Science Institute, Singapore. In: Proceedings of 2017 7th ICEA international conference on information science, library science and social sciences (ILS 2017). Information Engineering Research Institute Press, vol 6
14. Liu Y, Zhou H (2013) Optimize on data correlation of sensor nodes with adaptive fault-tolerant algorithm. Information Engineering Research Institute, USA. In: Proceedings of 2013 2nd AASRI conference on computational intelligence and bioinformatics (CIB 2013 V6). Elsevier, vol 7
15. Chen H, Zhang Y, Li W (2019) Monitoring of subway tunnel settlement using an inclinometer. Geotech Test J 42(1):20180040
16. Mourlas NJ, Gilchrist KH, Giovangrandi L, Maluf NI, Kovacs GTA (2002) An in-line osmometer for application to a cell-based biosensor system. Sens Actuators: B Chem

Advancing Internet of Things Through Statistical Pruning of Neural Networks

Chengchen Mao, Qilian Liang, Chenyun Pan, and Ioannis Schizas

Abstract Pruning is showing huge potential for compressing and accelerating deep neural networks by eliminating redundant parameters. Along with more terminal chips integrated with AI accelerators for Internet of Things (IoT) devices, structured pruning is gaining popularity with the edge computing research area. Different from filter pruning and group-wise pruning, stripe-wise pruning (SWP) conducts pruning at the level of stripes in each filter. By introducing filter skeleton (FS) to each stripe, the existing SWP method sets an absolute threshold for the values in FS and removes the stripes whose corresponding values in FS could not meet the threshold. Starting with investigation into the process of stripe wise convolution, we use the statistical properties of the weights located on each stripe to learn the importance between those stripes in a filter and remove stripes with low importance. Our pruned VGG-16 achieves the existing results by a 4-fold reduction in parameter with only 0.4% decrease of accuracy. Results from comprehensive experiments on IoT devices are also presented.

Keywords Prune stripe-wise · Edge device · Normal distribution · Internet of things

1 Introduction

In the internet of Things (IoT) realm, sensors and actuators seamlessly integrate with the environment [1], enabling cross-platform information flow for environmental metrics, while numerous connected devices generate massive data, offering convenience but also high latency. However, applications such as vehicle-to-vehicle communication which enhances the traffic safety by automobile collaboration, require low latency and high security. Edge computing is a promising technology that has the potential to improve the performance and security of IoT applications [2].

C. Mao (✉) · Q. Liang · C. Pan · I. Schizas
Department of Electrical Engineering, The University of Texas at Arlington, 701 S. Nedderman Drive, Arlington, TX 76019, USA
e-mail: chengchen.mao@mavs.uta.edu

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_35

Even though chip giants are integrating more and more AI accelerators into their design for the IoT devices, the massive number of parameters and the huge amount of computation would bring horrible experience to the consumers when Deep Neural Networks (DNNs) are employed in their devices [3]. To alleviate such kind of problems, researchers have made efforts in many directions, which could be mainly categorized into two types: unstructured ones and structured ones.

Unstructured pruning methods prune individual weights based on the importance of themselves. For example, by using the second-order derivatives of the error function, Optimal Brain Damage and Optimal Brain Surgery proposed to remove unimportance weights from a trained network [4]. Deep Compression compressed neural networks by pruning the unimportant connections, quantizing the network, and applying Huffman coding [5]. With Taylor expansion that approximates the change in the cost function [6], pruned convolutional kernels to enable efficient inference and could handle the transfer learning tasks effectively. A major downside of the unstructured methods is the sparse matrix and the relative indices after pruning, which leads to the complexity and inefficiency on hardware [7].

Structured methods prune weights in a predictable way. Li et al. [8] pruned unimportant filters with L_1 norm. Luo et al. [9] pruned filters based on statistics information computed from its next layer, not the current layer. He et al. [10] pruned channels by LASSO regression. By using scaling factors from batch normalization layers, [11] removed unimportant channels. Lebedev and Lempitsky [12] revisited the idea of brain damage and extended it to group wise, obtaining the sparsities in new neural network. To the best of our knowledge, one recent study [13] proposed a stripe-wise pruning based methods by introducing filter skeleton to learn the shape of filters and then performed pruning on the stripes according to the corresponding values of the filter skeleton.

However setting an absolute threshold sometimes could not express the relative importance of each stripe in a filter. To resolve this problem, in this work, we put forward a new method, using the statistical properties of the weights located on each stripe, to learn the importance between those stripes in a filter. The intuition of this method is triggered by the process during stripe wise convolution and the properties of normal distributions.

Our principal contributions in this paper could be summarized as follows: (1) The research proposes a new method for determining which weights in a neural network can be pruned without sacrificing accuracy. Our pruned VGG16 achieves results comparable to the existing model, with a fourfold reduction in parameters and only a 0.4% decrease in accuracy. (2) The proposed method is based on sound theoretical principles, making it more trustworthy and easier to understand and apply. (3) The effectiveness of the proposed approach is tested on different neural network architectures (VGG16 and ResNet56) and evaluated on edge devices with limited computational resources.

The paper is arranged as follows: In Sect. 2, we present our method as well as the theoretical framework behind it. In Sect. 3, we explain the experimental details and demonstrate comparisons between our method and the original method. Additionally,

we showcase the performance of our method deployed on edge devices. Finally, concluding remarks are provided in Sect. 4.

2 The Proposed Method

2.1 Stripe Wise Convolution

In l-th convolution layer, suppose the weight 4-D matrix W is of size $\mathbb{R}^{N\times C\times K\times K}$, where N, C and K are the numbers of filters, the channel dimension and the kernel size, respectively.

Let $x^l_{c,h,w}$ be one point of feature map in the l-th layer and $x^{l+1}_{n,h,w}$ be the convolution result in the $l+1$-th layer. We modify the calculation order of standard convolution in the stripe-wise way (1) as illustrating in Fig. 1a.

$$x^{l+1}_{n,h,w} = \sum_{i}^{K}\sum_{j}^{K}\left(\sum_{c}^{C} w^l_{n,c,i,j} \times x^l_{c,h+i-\frac{K+1}{2},w+j-\frac{K+1}{2}}\right) \tag{1}$$

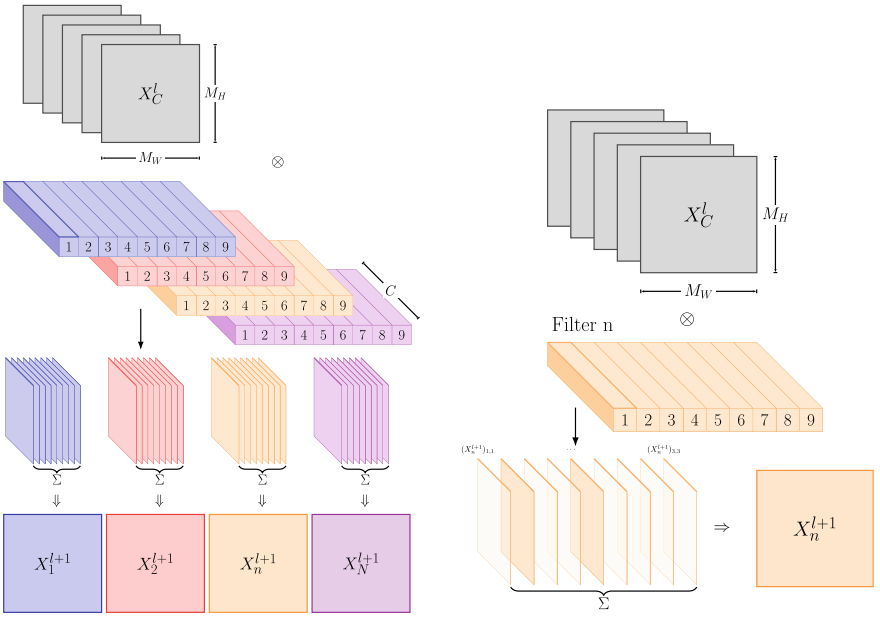
$x^l_{c,p,q} = 0$, when $p < 1$ or $p > M_H$ or $q < 1$ or $q > M_W$. M_H is the height of the feature map, while M_W represents the width.

From Fig. 1a, we could find that in stripe wise convolution, the convolution result of individual filter is the summation of the convolution result of the stripes which belongs to this filter. One intuition is that if the convolution result of the stripe 1 is much smaller than the convolution result of the stripe 2, Stripe 1 could be pruned and Stripe 2 could be kept as shown in Fig. 1b. The following part will prove it in a theoretical manner.

2.2 Theoretical Analysis

Batch normalization (BN) is widely used in a neural network. This method could make DNN faster and more stable [14]. In one filter, suppose B is a mini-batch of size m, i.e., $B = \{a_1, \ldots a_m\}$. BN layer processes these following transformation steps: $\mu_B = \frac{1}{m}\sum_{i=1}^{m} a_i$, $\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m}(a_i - \mu_B)^2$, $\hat{a}_i = \frac{a_i-\mu_B}{\sqrt{\sigma_B^2+\epsilon}}$, $x_i = \gamma\hat{a}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(a_i)$, where μ_B and σ_B are the empirical mean and standard deviation of B. To resume the representation ability of the network, scale γ and shift β are learned during the whole process.

After transformation in the BN layer, in c-th channel of l-th layer, the input feature map could be $X^l_c \sim \mathcal{N}(\beta^l_c, (\gamma^l_c)^2)$. When M_H is large, $(X^l_c)_{i,j} \sim \mathcal{N}(\beta^l_c, (\gamma^l_c)^2)$. From (1), we could get $X^{l+1}_n = \sum_i^K \sum_j^K (\sum_c^C w^l_{n,c,i,j} \times (X^l_c)_{i,j})$.



(a) The filter's kernel size is 3 (b) Single filter case

Fig. 1 Stripe wise convolution

Assuming all data is independently identically distribution, with the properties of normal distribution [15], we have $X_n^{l+1} \sim \mathcal{N}(\mu_n^{l+1}, (\sigma_n^{l+1})^2)$, where $\mu_n^{l+1} = \sum_i^K \sum_j^K (\sum_c^C w_{n,c,i,j}^l \beta_c^l)$ and $(\sigma_n^{l+1})^2 = \sum_i^K \sum_j^K (\sum_c^C (w_{n,c,i,j}^l)^2 (\gamma_c^l)^2)$

To reduce the number of parameters $w_{n,c,i,j}^l$ and avoid the value of μ_n^{l+1} change, we introducing an importance indicator $Q_{n,i,j}^l$ to the output of convolution of each stripe and have the following loss function.

$$L_n = \text{loss}\left(\mu_{\text{n}}^{\text{l+1}}, \sum_{\text{i}}^{\text{K}} \sum_{\text{j}}^{\text{K}} \text{Q}_{\text{n,i,j}}^{\text{l}} \left(\sum_{\text{c}}^{\text{C}} \text{w}_{\text{n,c,i,j}}^{\text{l}} \beta_{\text{c}}^{\text{l}}\right)\right) + \alpha \text{g}_{\text{n}}(\text{Q}) \tag{2}$$

where $g_n(Q) = \sum_i^K \sum_j^K \left| Q_{n,i,j}^l \right|$, $Q_{n,i,j}^l = 1$ or 0.

Let $s_{n,i,j}^l \triangleq \sum_c^C w_{n,c,i,j}^l$. If we assume $\beta_1^l = \beta_2^l \cdots = \beta_c^l = \beta^l$, (2) could be written as $L_n = \text{loss}(\beta^{\text{l}} \sum_{\text{a}}^{\text{K}} \sum_{\text{b}}^{\text{K}} \text{s}_{\text{n,a,b}}^{\text{l}}, \beta^{\text{l}} \sum_{\text{i}}^{\text{K}} \sum_{\text{j}}^{\text{K}} \text{Q}_{\text{n,i,j}}^{\text{l}} \text{s}_{\text{n,i,j}}^{\text{l}}) + \alpha \text{g}_{\text{n}}(\text{Q})$ and it can be further written as

$$L_n = \text{loss}\left(1, \sum_{\text{i}}^{\text{K}} \sum_{\text{j}}^{\text{K}} \text{Q}_{\text{n,i,j}}^{\text{l}} \text{T}_{\text{n,i,j}}^{\text{l}}\right) + \alpha' \text{g}_{\text{n}}(\text{Q}) \tag{3}$$

where

$$T^l_{n,i,j} = \frac{s^l_{n,i,j}}{\sum_a^K \sum_b^K s^l_{n,a,b}} \tag{4}$$

Obviously, $\sum_i^K \sum_j^K T^l_{n,i,j} = 1, 0 \leq T^l_{n,i,j} < 1$ To minimize (3), we could set $Q^l_{n,i,j} = 0$ to those $T^l_{n,i,j}$ close to 0, which means the corresponding stripes will be pruned. $T^l_{n,i,j}$ could be used to describe the relative importance of $\text{stripe}_{i,j}$ in filter_n. When $T^l_{n,i,j} \to 1$, $\text{stripe}_{i,j}$ contributes more than other stripes. When $T^l_{n,i,j} \to 0$, $\text{stripe}_{i,j}$ contributes less than other stripes and could be pruned.

Before setting a threshold for $T^l_{n,i,j}$ to prune stripes, we need to impose regularization on the whole neural network to achieve sparsity. This method could avoid so-called "Train, Prune, Fine-tune" pipeline. The regularization on the FS will be

$$L = \sum_{(x,y)} \text{loss}(\text{f}(\text{x}, \text{W}), \text{y}) + \alpha \text{g}(\text{W}) \tag{5}$$

where α adjusts the degree of regularization. $g(W)$ is L_1 norm penalty on W and could be written as $g(W) = \sum_{l=1}^{L}(\sum_{n=1}^{N} \sum_{c=1}^{C} \sum_{i=1}^{K} \sum_{j=1}^{K} \left| W^l_{n,c,i,j} \right|)$.

To avoid using sub-gradient at non-smooth point, instead of the L_1 penalty, we deploy the smooth-L_1 penalty [16].

3 Experiments

In order to assess the performance of the proposed model and confirm its effectiveness, we carry out experiments using the CIFAR-10 dataset. Our method is implemented using the publicly available Torch. **Dataset and Model:** CIFAR-10 [17] is one of the most popular image collection data sets. This dataset contains 60K color images from 10 different classes. 50K and 10K images are included in the training and testing sets respectively. By adopting CIFAR-10, we evaluated the proposed method mainly on VGG [18] and ResNet56 [19]. The inference time refers to the total amount of time needed to classify 3270 image patches with a size of 224×224. **Baseline Setting:** We train the model using mini-batch size of 64 for 100 epochs. The initial learning rate is set to 0.1, and is divided by 10 at the epoch 50. Random crop and random horizontal flip are used as data augmentation for training images. Image is scaled to 256×256. Then a 224×224 part is randomly cropped from the scaled image for training. The testing is the center crop with 224×224. **Experiment environment:** NVidia 1080-TI and Intel Core i5-8500B are selected as two different computing platforms representatives of the server and the edge device, respectively. The first is a GPU which has high computation ability, however needs communication with sensors and actuators. The second is a CPU to represent the restricted computer power of an edge device.

3.1 Comparing with the Original SWP

To compare our method with the original SWP, we revisit the concept of filter skeleton (FS) from [13]. Each value in FS corresponds to a stripe in the filter. During training, the filters' weights are multiplied with FS. With I representing the FS, the stripe wise convolution could be written as

$$x_{n,h,w}^{l+1} = \sum_{i}^{K} \sum_{j}^{K} I_{n,i,j}^{l} \left(\sum_{c}^{C} w_{n,c,i,j}^{l} \times x_{c,h+i-\frac{K+1}{2},w+j-\frac{K+1}{2}}^{l} \right) \tag{6}$$

where $I_{n,i,j}^{l}$ is initialized with 1.

The regularization on the FS will be

$$L = \sum_{(x,y)} \text{loss}(\text{f}(\text{x}, \text{W} \odot \text{I}), \text{y}) + \alpha \text{g}(\text{I}) \tag{7}$$

where $\odot$ denotes dot product and α adjusts the degree of regularization. $g(I)$ is written as: $g(I) = \sum_{l=1}^{L} (\sum_{n=1}^{N} \sum_{i=1}^{K} \sum_{j=1}^{K} \left| I_{n,i,j}^{l} \right|)$.

For convenience, in Table 1 for the comparison on CIFAR-10, both the original method and our method use FS to train and prune the whole neural network. Both of them use the coefficient α of regularization, which is set to $1e-5$ and $5e-5$. The difference is that for the original method, pruning is based on the value in FS which corresponds to a stripe and for our method, pruning is based on $T_{n,i,j}^{l}$ which combines the weights located in a stripe. Regarding the choice of T, we used the value corresponding to the highest accuracy.

From the table, we could find both methods could reduce the number of parameters and the amount of computation (FLOPs) in a considerable volume without losing network performance. For the backbone is VGG16 situation, when $\alpha = 1e-5$, the number of parameters and the amount of computation of our method are larger than the original approach. This is because our method will keep at least one stripe in a filter, while the original approach might prune a whole filter. However, when $\alpha = 5e-5$, the original approach could not converge and our method could reach a high compression rate both in the number of parameters and the amount of computation. Our pruned VGG16 could achieve 95% reduction in memory demands.

For the backbone is Resnet56 situation, we present our result of $\alpha = 5e-5$. To compare with the original approach's result of $\alpha = 1e-5$, our method could see a large reduction in the number of parameters and the amount of computation while sacrificing a bit of accuracy. Our pruned Resnet56 could achieve 75% reduction in memory demands.

In our method, there are two decisive hyper-parameters in the neural network, the coefficient α of regularization in (7) and the weight combination threshold T in (4). As the outcomes of the experiment demonstrated in Table 2, we display the effects of the hyper-parameters in pruning consequences. It could be noticed that $\alpha = 5e-5$ and $T = 0.005$ holds an acceptable pruning ratio as well as test accuracy.

Table 1 Comparison with the original SWP on CIFAR-10

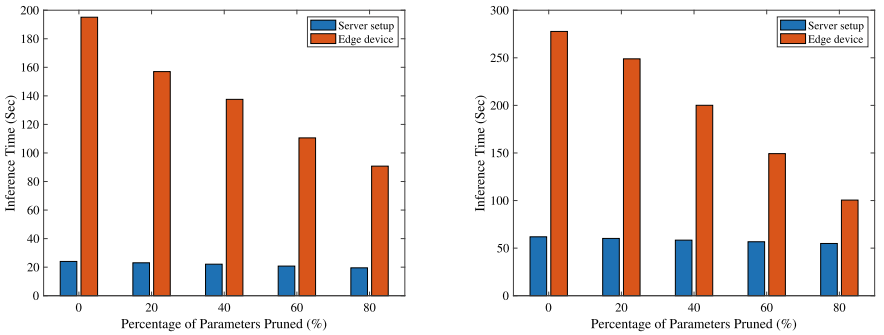
Backbone	Metrics	Params	FLOPS	Accuracy
VGG16	Baseline	14.76 M	627.37 M	93.76 %
	Original($\alpha = 1e-5$)	3.62M	350.28M	93.46 %
	Original($\alpha = 5e-5$)	Could not converge		
	Ours ($\alpha = 1e-5$, $T = 0.0001$)	4.63M	385.49M	93.43 %
	Ours ($\alpha = 5e-5$, $T = 0.005$)	0.84M	126.03M	93.06 %
ResNet56	Baseline	0.87 M	251.50 M	93.11 %
	Original($\alpha = 1e-5$)	0.60M	150.63M	93.41 %
	Ours ($\alpha = 5e-5$, $T = 0.001$)	0.23M	60.76M	92.96 %

Table 2 Different coefficient α and weight combination threshold

α	1e − 5			5e − 5			
T	0.0001	0.001	0.01	0.0001	0.0005	0.001	0.005
Params (M)	4.63	4.17	2.89	0.78	0.80	0.79	0.84
FLOPS (M)	385.49	327.17	200.09	130.96	135.56	134.68	126.03
Accuracy (%)	93.43	93.28	92.99	92.79	92.86	92.96	93.06

3.2 Edge Device Performance

We further verify our approach in an edge device. Pruning is executed on the server as training consumes computing resources on learning the importance between the stripes and serval complete passes of the training dataset through the whole neural network. The pruned networks are then deployed on these two computing platforms to test results and get the inference time. The comparison is shown in Fig. 2a, b. It should be noted that stripe wise convolution is not yet optimized in CUDA. Along with the increase in percentage of parameters pruned, the decline in inference time in servers is not quite clear. However, the inference time in edge device drops by half when 75–95% of parameters are pruned.



(a) Pruned VGG16 (b) Pruned ResNet56

Fig. 2 Required inference time for pruned models

4 Conclusion

In this work, we avoid using an absolute threshold in existing stripe-wise pruning by combining the weights located on each stripe. This allows us to learn the importance between stripes in a filter and remove those with low importance. Our pruned method effectively reduces the parameters and inference time of our VGG16 model without significantly impacting accuracy. In future work, we will explore the introduction of regularizers to prune filters with single stripes, which may further compress deep neural networks and improve performance.

Acknowledgements This work was supported by U.S. National Science Foundation (NSF) under Grant CCF-2219753.

References

1. Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of things (IoT): a vision, architectural elements, and future directions. Fut Gener Comput Syst 29(7):1645–1660
2. Yu W, Liang F, He X, Hatcher WG, Lu C, Lin J, Yang X (2018) A survey on the edge computing for the internet of things. IEEE Access 6:6900–6919
3. Li H, Ota K, Dong M (2018) Learning IoT in edge: deep learning for the internet of things with edge computing. IEEE Netw 32(1):96–101
4. Bishop CM et al (1995) Neural networks for pattern recognition. Oxford University Press, Oxford
5. Han S, Mao H, Dally WJ (2015) Deep compression: compressing deep neural networks with pruning, trained quantization and huffman coding. arXiv preprint arXiv:1510.00149
6. Molchanov P, Tyree S, Karras T, Aila T, Kautz J (2016) Pruning convolutional neural networks for resource efficient inference. arXiv preprint arXiv:1611.06440
7. Han S, Liu X, Mao H, Pu J, Pedram A, Horowitz MA, Dally WJ (2016) EIE: efficient inference engine on compressed deep neural network. ACM SIGARCH Comput Architect News 44(3):243–254

8. Li H, Kadav A, Durdanovic I, Samet H, Graf HP (2016) Pruning filters for efficient convnets. arXiv preprint arXiv:1608.08710
9. Luo JH, Wu J, Lin W (2017) Thinet: a filter level pruning method for deep neural network compression. In: Proceedings of the IEEE international conference on computer vision, pp 5058–5066
10. He Y, Zhang X, Sun J (2017) Channel pruning for accelerating very deep neural networks. In: Proceedings of the IEEE international conference on computer vision, pp 1389–1397
11. Liu Z, Li J, Shen Z, Huang G, Yan S, Zhang C (2017) Learning efficient convolutional networks through network slimming. In: Proceedings of the IEEE international conference on computer vision, pp 2736–2744
12. Lebedev V, Lempitsky V (2016) Fast convnets using group-wise brain damage. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2554–2564
13. Meng F, Cheng H, Li K, Luo H, Guo X, Lu G, Sun X (2020) Pruning filter in filter. Adv Neural Inform Process Syst 33:17629–17640
14. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning, pp 448–456
15. Feller W (2008) An introduction to probability theory and its applications, vol 2. Wiley, Amsterdam
16. Schmidt M, Fung G, Rosales R (2007) Fast optimization methods for l1 regularization: a comparative study and two new approaches. European conference on machine learning. Springer, New York, pp 286–297
17. Krizhevsky A, Hinton G (2009) Learning multiple layers of features from tiny images. University of Toronto, Toronto
18. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
19. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778

Design of a Joint Radar-Communication System for Target Recognition

Qilong Miao, Jing Liang, and Ge Zhang

Abstract Joint radar-communication system (JRCS) that shares the same infrastructure and employs dual-function signal, has advantages of low cost, compact size and less power consumption. At present, most of previous investigations in JRCS are focused on detection, but system design and analysis for target recognition are scarce due to its complexity. In this work, we establish a JRCS aiming at both target recognition and communication using a single signal. Compared with detection oriented system that provides position of the targets, the proposed JRCS further identifies the categories of the targets. Frequency modulated continuous wave (FMCW) and orthogonal frequency division multiplexing (OFDM) signals are adopted as dual-function fundamental waveforms. Also, the optimal spectrum distribution of recognition and communication is theoretically analyzed. Simulation results show that the proposed JRCS performs well for both target recognition and communication.

Keywords HRRP · Joint radar-communication system (JRCS) · MKL · MRA · Target recognition

1 Introduction

JRCS that meets the demands of both detection and communication, has seen a significant amount of research effort in recent years [1, 2]. Unlike the traditional independent radar and communication systems, the JRCS combines both of them in a unitary hardware platform and adopts dual-functional signal [3]. Shared structure helps to reduce the system size, cost, and energy. Because of the efficient resource

Please note that the LNCS Editorial assumes that all authors have used the western naming convention, with given names preceding surnames. This determines the structure of the names in the running heads and the author index.

Q. Miao · J. Liang (✉) · G. Zhang
University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave, Chengdu 611731, China
e-mail: 201911012021@std.uestc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_36

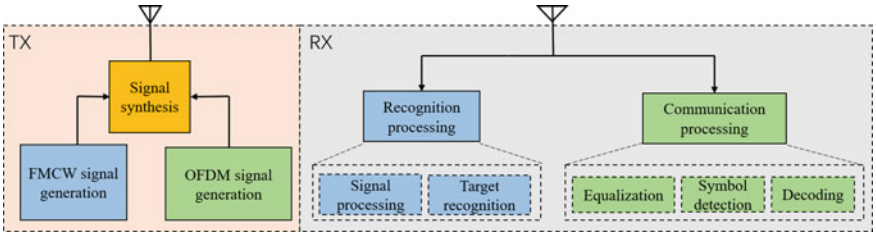


Fig. 1 System structure of JRCS

utilization, the JRCS is adapted to resource constrained platforms, such as unmanned aerial vehicle (UAV) airborne sensor.

Compared with the detection that estimates the location of the target, recognition is indispensable to identify what the category the detected object is. Unlike the detection, recognition generally treats target as an extended object other than a simple point [4]. However, at present, a lot of research about JRCS focus on the target detection, and there are few investigations into target recognition for JRCS due to the complexity.

High-resolution range profile (HRRP), a one-dimensional real vector contains rich target information including size, structure and reflection intensity, has received intensive attention from the radar automatic target recognition (RATR) community [5–7]. Compared with the recognition methods based on electromagnetic scattering characteristics and synthetic aperture radar (SAR) images [8], the recognition methodology of HRRP offers the advantages of obvious features and less computation. Thus, there have been a lot of recognition methods on a basis of HRRP, such as pattern recognition [5], deep learning [6], information theory [9] and fuzzy logic [7]. However, there are still few investigations into the JRCS in view of HRRP for target recognition.

In this work, we establish a JRCS that has ability of both target recognition and communications between JRCS based on one dual-functional signal. Figure 1 illustrates the bistatic structure of the proposed system. For the joint waveform, FMCW is selected due to its simplicity and the low-cost receiver [1], while OFDM waveform is applied due to its stable performance in multipath fading and relatively simple synchronization. FDM is applied as the multiplexing solution of FMCW and OFDM waveforms. The optimal bandwidth allocation strategy of JRCS is theoretically analyzed in this work.

2 Transmitting Model of JRCS

The proposed JRCS adopts FMCW + OFDM signal as the joint signal, and the FMCW signal and OFDM signal are divided in frequency domain. The transmitting signal $x(t)$ of JRCS is

$$x(t) = [x_r(t) + x_c(t)\exp(j2\pi f_0 t)]\exp(j2\pi f_c t), \tag{1}$$

where $x_r(t)$ is the FMCW signal of bandwidth B_r, $x_c(t)$ is the OFDM signal of bandwidth B_c. $f_0 = (B_r + B_c)/2 + B_g$ is the shift frequency of $x_c(t)$, B_g is the guard bandwidth between $x_r(t)$ and $x_c(t)$. f_c is the carrier frequency. In (1), the FMCW waveform $x_r(t)$ is

$$x_r(t) = \text{rect}\left(\frac{t}{T_p}\right)\exp\left(j\pi St^2\right), \tag{2}$$

where T_p is the ramp duration, S is the frequency modulation slop with bandwidth $B_r = ST_p$, $\text{rect}\left(\frac{t}{T}\right)$ is rectangular pulse.

In (1), the OFDM signal $x_c(t, q)$ is

$$x_c(t) = \sum_{m=0}^{M-1} g_m \exp\left(j2\pi m\Delta f t\right)\text{rect}\left(\frac{t}{T_p}\right), \tag{3}$$

where g_m is the sent code of mth subcarrier, M is the number of subcarriers, Δf is the subcarriers spacing.

In JRCS, the system design is mainly affected by B_r and B_c. Assume JRCS has system bandwidth of B_S, then $B_S = B_r + B_c + B_g$. We define $\mu \in [0, 1]$ to adjust the bandwidth ratio of FMCW and OFDM signals. That is, $B_r = \mu(B_S - B_g)$ and $B_c = (1 - \mu)(B_S - B_g)$.

2.1 Optimal Selection of μ

To recognize the target, testing HRRP frames (HRRPs obtained by JRCS's signal of certain μ) are compared to training HRRP frames (HRRPs obtained by FMCW signal of bandwidth B_t). Before recognition, interpolation to test HRRP frames is needed for B_r is very likely different from B_t. Therefore, interpolation error, which deteriorates recognition performance, will be introduced. We model the problem of selecting μ as an optimization problem whose objective is minimizing the interpolation error $|\varepsilon|$ while the communication symbol rate R_s is limited. Note that $R_s = B_c$, the problem can be modeled as follow.

$$\begin{aligned} &\underset{\mu}{\text{argmin}} \quad |\varepsilon| \\ &\text{s.t.} \quad R_c \geq R_0. \end{aligned} \tag{4}$$

where R_0 is the required symbol rate.

According to the interpolation theorem, assume that the function $f(x)$ has a second derivative on $[x_1, x_2]$, for any interpolation point $x \in [x_1, x_2]$, the error between the linear interpolation function $u(x)$ and $f(x)$ is

$$\varepsilon = f(x) - u(x) = \frac{f^{(2)}(\varphi)}{2}(x - x_1)(x - x_2), \tag{5}$$

where $\varphi \in [x_1, x_2]$ depends on x. It is challenging to optimize $|\varepsilon|$ directly, so we turn to optimizing the upper bound of $|\varepsilon|$. We choose a positive A such that $f^{(2)}(\varphi)/2 \leq A$. Then (5) becomes

$$|\varepsilon| \leq A(x - x_1)(x_2 - x) = -A(x_2 - x)x_1 + Ax(x_2 - x). \tag{6}$$

In (6), if x and x_2 are fixed, The upper bound of $|\varepsilon|$ becomes a linear function of $x_1 \in (-\infty, x]$. Thus, $|\varepsilon|$ increases as the error between x_1 and x increases, and $|\varepsilon|$ achieves the minimum when $x_1 = x$. The conclusion for x_2 is similar.

Suppose R_1 and R_2 are two adjacent range bins of test HRRP, and $R \in [R_1, R_2]$ is the range bin that needs to be interpolated. If $B_r \leq B_t$, R_1, R_2, R have the following relationship

$$\begin{aligned} R_1 &= a\Delta R_m, \\ R_2 &= (a+1)\Delta R_m, \\ R &= b\Delta R_t, \end{aligned} \tag{7}$$

where a and b are some non-negative integers, ΔR_m is the range space between range adjacent bins of test HRRP, ΔR_t is the range space of training HRRP. Note that if $B_r = B_t$, then $R = R_1$ and $a = b$.

According to previous conclusions, the larger the error between R and R_1 or R_2, the bigger the interpolation error. The error η between R and R_1 is

$$\eta = R - R_1 = a\Delta R_m - b\Delta R_t. \tag{8}$$

According to relationship between B_r and S,

$$\begin{aligned} \Delta R_m &= \frac{T_p c^2}{4R_w B_r}, \\ \Delta R_t &= \frac{T_p c^2}{4R_w B_t}, \end{aligned} \tag{9}$$

where R_w is the maximum range of interest. Alter simplification, (8) becomes

$$\eta = \frac{T_p c^2}{4R_w B_t}\left(a\frac{B_t}{B_r} - b\right). \tag{10}$$

It is easy to find that η is a linear function of B_t/B_r, and $\eta = 0$ when $B_r = B_t$. The interpolation error is minimal when the radar signal bandwidth is equal to the signal bandwidth to generate training HRRPs, and when the error between B_r and B_t is larger, the interpolation error is bigger. The same conclusion can be drawn for R_2.

If $B_r > B_t$, the following relation holds

$$\lim_{B_r \to \infty} R = R_1 = R_2. \tag{11}$$

In this case, the interpolation error decreases as B_r increases. When B_r tends to approach infinity, the interpolation error is 0. To sum up, the solution of (4) is as follows

$$\mu = \begin{cases} \frac{B_t}{B-B_g}, & B - B_g - R_0 \geq B_t, \\ \frac{B-B_g-R_0}{B-B_g}, & B - B_g - R_0 < B_t. \end{cases} \tag{12}$$

3 Receiving Model of JRCS

Based on (1), the received FMCW and OFDM signals can be recovered by passing through constant phase low-pass filters (CPLPFs) $H_r(t)$ and $H_c(t)$, respectively.

3.1 Recognition Processing

The first step of recognition is to obtain HRRP of the target. The targets are modeled as a collection of main scatters if the range resolution of JRCS is much smaller than the target size. Assume the doppler shift of the target are perfectly estimated and compensated the target is static, the echo $y_r(t)$ is

$$y_r(t) = \sum_{k=1}^{K} \alpha_k x(t - \tau_k) + n(t), \tag{13}$$

where K is the number of the main scatters, for the kth main scatter, α_k is the attenuation coefficient caused by channel and reflection, $\tau_k = 2R_k/c$ is the time delay, R_k is the radial distance, c is the speed of light, $n(t)$ is the noise. The beat signal $\hat{y_r}(t)$ is obtained by removing the carrier frequency, being filtered by $H_r(t)$, and deramping, that is

$$\hat{y_r}(t) = \{H_r(t) * [y_r(t)\exp(-j\pi f_c t)]\}\exp\left(-j\pi S t^2\right), \tag{14}$$

After simplification,

$$\begin{aligned} \hat{y_r}(t) = &\sum_{k=1}^{K} \alpha_k \text{rect}\left(\frac{t-\tau_k}{T_p}\right)\exp\left(-j2\pi f_{R_k} t\right)\exp\left(-j2\pi f_c \tau_k\right) \\ &\times \exp\left(j\pi S\tau_k^2\right) + n(t). \end{aligned} \tag{15}$$

In (15), $\exp\left(-j2\pi f_{R_k} t\right)$ is related to range of the main scatters, $f_{R_k} = S\tau_k$. Once f_{R_k} is determined, R_k can be calculate by

$$R_k = -\frac{cf_{R_k}}{2S}. \tag{16}$$

HRRP of the target can be obtained by the following formula.

$$z = \|FT[\hat{y_r}(t)]\|_2, \tag{17}$$

where $FT[\cdot]$ is Fourier transform, $\|\cdot\|_2$ is l_2-norm.

The second step of recognition processing is to get the categories of obtained HRRPs. In this work, recognition based on minimum Kullback–Leibler (MKL) distance and minimum resistor-average (MRA) distance proposed in [9] has been adopted.

3.2 Communication Processing

In this work, we focus on problem of communication between two nodes. This framework can be easily extended to multiple nodes situation by frequency division multiple access (FDMA), time division multiple access (TDMA) and code division multiple access (CDMA). The received communication signal is

$$y_c(t) = \kappa x(t - \tau_l) + n(t), \tag{18}$$

where τ_l is the time delay between these two nodes, κ is the channel coefficient. The base band OFDM signal $\hat{y_c}(t)$ is obtained by removing the carrier frequency and being filtered by $H_c(t)$, that is

$$\hat{y_c}(t) = H_c(t) * [y_c(t)\exp(-j\pi f_c t)\exp(-j\pi f_0 t)], \tag{19}$$

After simplification,

$$\begin{aligned}\hat{y_c}(t) = \kappa \sum_{m=0}^{M-1} g_m \exp(j2\pi m\Delta f t)\exp(-j2\pi f_c\tau_l - j2\pi m\Delta f\tau_l)) \\ \times \text{rect}\left(\frac{t-\tau_l}{T_s}\right) + n(t).\end{aligned} \tag{20}$$

In OFDM processing, symbol detection of OFDM signals can be easily done by Fast Fourier transform (FFT). The detected symbols are then decoded to obtain the transmitted information.

Table 1 Simulation parameters settings

Simulation parameters	Symbol	Quantity
System bandwidth	B_S	610 MHz
Duration	T_p	35 μs
Guard bandwidth	B_g	10 MHz
Sampling frequency	f_s	700 MHz
No. of main scatters in target 1	K_1	8
No. of main scatters in target 2	K_2	10
No. of main scatters in target 3	K_3	12
No. of training samples	W_t	7200
No. of test samples	W_e	500
No. of template	N_t	1600
Signal Bandwidth to get training HRRPs	B_t	300 MHz
Angle of each training frame	a_t	0.2°
No. of HRRPs per training frame	N_0	5
No. of HRRPs per test frame	N_1	5
No. of target categories	N_c	3

4 Simulation and Analysis

In the simulation, we validate the relationship between μ and performance of recognition and communication, respectively. 16 quadrature amplitude modulation (16-QAM) is chosen as digital modulation method due to its high resource efficiency. Assume that channel side infomation (CSI) is known. Table 1 shows the simulation parameters. An example of transmitting signal spectrum of JRCS is shown in Fig. 2a.

4.1 Target Recognition Performance

Figure 2b shows a HRRP example of a targets with 8 main scatters. Figure 3a, b show the recognition performance of JRCS for MKL and MRA, respectively. The following conclusions can be drawn: (1) If μ makes $B_r = B_t$, correct recognition rate of both MKL criterion and MRA criterion are the best, especially when SNR $\leq 10\,d$B. (2) If $B_r < B_t$, system with μ making B_r closer to B_t obtains better recognition performance. (3) If $B_r > B_t$, system with μ making larger B_r obtains better recognition performance. This result confirms the conclusion (12) about optimal selection of μ.

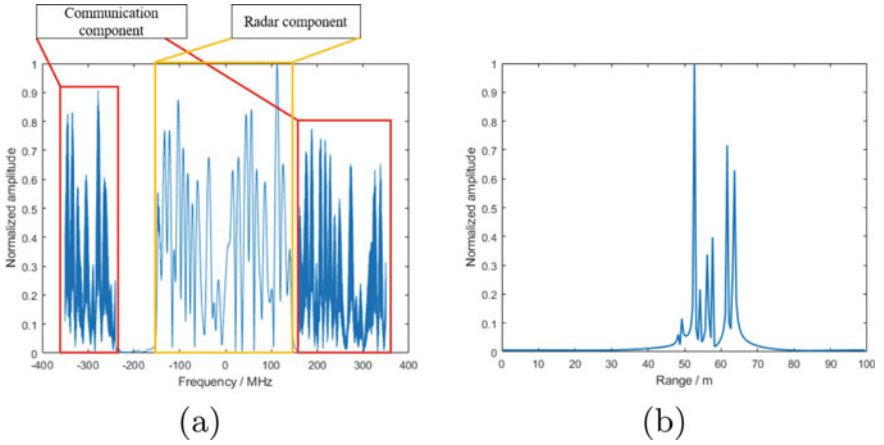


Fig. 2 HRRP and transmitting signal's spectrum of JRCS: **(a)** The spectrum of JRCS transmitting signal; **(b)** HRRP of target with 8 main scatters

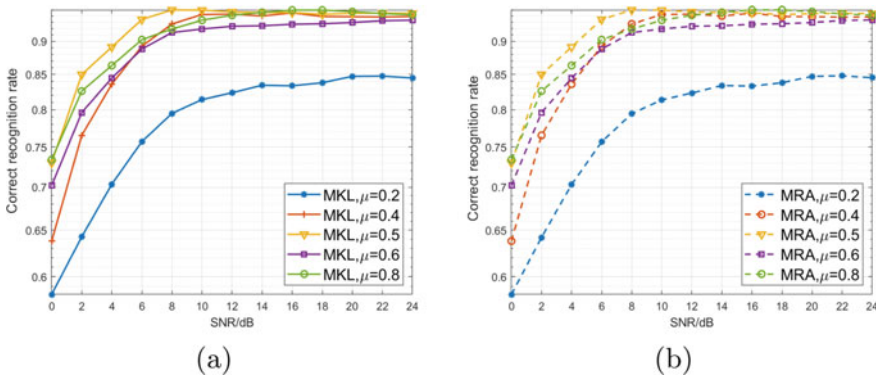


Fig. 3 Correct recognition rate under different μ for: **(a)** MKL and **(b)** MRA

4.2 Communication Performance

Communication performance of JRCS is measured via BER for different μ. Figure 4 shows that BER decreases with the increase of μ. The reason of this result is that larger B_c is equivalent to higher symbol rate, which means more symbols will be transmitted in one signal. For fixed transmitting power, more transmitting symbols will reduce SNR per symbol, which result in increment of BER.

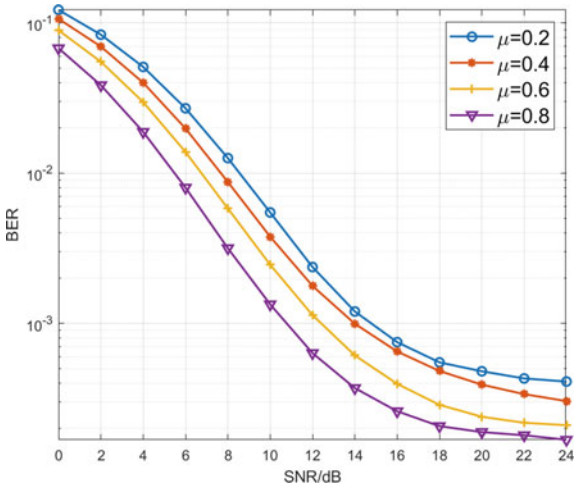


Fig. 4 BER of JRCS for different μ

5 Conclusion

A novel JRCS that performs both target recognition and communication via a single transmitting signal has been established in this work. The proposed framework has advantages of low cost, compatible size and less power consumption. Through simulation, the conclusions be drawn as follows.

- The target recognition performance of JRCS is impacted by the bandwidth distribution scheme. The bandwidth distribution scheme is optimal such that the bandwidth of signal to obtain test HRRPs and training HRRPs are equal.
- Communication performance of JRCS is affected by bandwidth allocation schemes as well. If more bandwidth is allocated to communication, the symbol rate and the BER become higher simultaneously.

This article shows how the JRCS works. In future work, a framework of multiple JRCSs that form a radar sensor network (RSN) will be designed to obtain synergistic effect of cooperation.

Acknowledgment 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. Mishra KV, Bhavani Shankar M, Koivunen V, Ottersten B, Vorobyov SA (2019) Toward millimeter-wave joint radar communications: a signal processing perspective. IEEE Signal Process Magaz 36(5):100–114
2. Dokhanchi SH, Mysore BS, Mishra KV, Ottersten B (2019) A mmwave automotive joint radar-communications system. IEEE Trans Aerosp Electr Syst 55(3):1241–1260
3. Paul B, Chiriyath AR, Bliss DW (2017) Survey of RF communications and sensing convergence research. IEEE Access 5:252–270
4. Du L, Liu H, Bao Z, Zhang J (2006) A two-distribution compounded statistical model for radar HRRP target recognition. IEEE Trans Signal Process 54(6):2226–2238
5. Du L, Liu H, Wang P, Feng B, Pan M, Bao Z (2012) Noise robust radar HRRP target recognition based on multitask factor analysis with small training data size. IEEE Trans Signal Process 60(7):3546–3559
6. Zhao C, He X, Liang J, Wang T, Huang C (2019) Radar HRRP target recognition via semi-supervised multi-task deep network. IEEE Access 7:114788–114794
7. Liang J, Zhu F (2014) Fuzzy logic classifier design for air targets recognition based on HRRP. Phys Commun 13:205–210
8. Guo D, Chen B, Zheng M, Liu H (2021) Sar automatic target recognition based on supervised deep variational autoencoding model. IEEE Trans Aerosp Electr Syst 57(6):4313–4328
9. Mao C, Liang J (2017) HRRP recognition in radar sensor network. Ad Hoc Netw 58:171–178

A Semantic Communication Based Wireless Image Transmission for Internet of Things Devices

Rangang Zhu, Shengxian Huang, Chenguang He, Shaojing Su, and Hao Chen

Abstract In recent years, wildfires occur frequently as global warming. People set a large number of sensors to monitor the wild land. However, the poor network in remote area can hardly afford the big data transmission while low delay. In this paper, we firstly propose a semantic communication framework, which consists of the GAN-based semantic extraction and LDPC code. To make it affordable for Internet of Things (IoT) devices, we then compression the pre-trained model by parameters pruning and clustering with the acceptable price of inference performance. Based on our analysis, the proposed semantic communication system can significantly reduce the volume of transmission data by extracting the semantic information of images and preform robustness in fading channel.

Keywords Semantic communication · Image compression · Internet of Things

1 Introduction

As global warming continues, extreme heat and dry weather occurs frequently, and the subsequent wildfire damage has got people's attention. People used to patrol the forest as guards, but now, a more common solution is that setting many sensors or edge devices to keep monitoring their surrounding in real time. Besides sending the warning signal in a faster and cheaper way, they can log more details and send them

R. Zhu
College of Electronic Engineering, National University of Defense Technology, Hunan, China

S. Huang · C. He (✉) · H. Chen
Communications Research Center, Harbin Institute of Technology, Harbin, China
e-mail: hechenguang@hit.edu.cn

S. Huang
e-mail: 21s105173@stu.hit.edu.cn

S. Su
College of Intelligence Science and Technology, National University of Defense Technology, Hunan, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_37

back to the center for further analysis. However, sensors are generally supposed to be energy-efficient and performance-limited. A sensor has a restricted service area, which means that it may need thousands of sensors to cover a hill. Although applying source compression, it is still a great burden for the transmission networks if lots of sensors send their data at the same time. Especially in some unusual cases like emergency rescue and damage assessment, live pictures or videos of interested area are always required.

To reduce the networks traffic and the probability of message collision, we need a more efficient image compression. However, conventional image compress like JPEG, BPG are regarded as high-efficiency engineering implementation to approximate the entropy of the image. Benefitting from the advancements of deep learning and end-to-end communication, semantic communication is promising to break the compression limit defined by Shannon information theory. Semantic communication system interprets received information at the semantic level rather than bit level, which is what we exactly do in convention communication system [1]. The difference between conventional communication and semantic communication application is illustrated in Fig. 1.

With the DL-based source coding and channel coding module, semantic features can be extracted and reconstructed quickly in high-performance computing environment. However, sensors can hardly afford the semantic interpreting processing of large semantic models. In this paper, we will focus on the lightweight semantic communication for edge devices.

There have been some initial works related to semantic communication and lite semantic communication for image transmission. [2, 3] proposed efficient joint source-channel coding methods for wireless image transmission based on the convolution neural network (CNN), respectively named DeepJSCC and DeepJSCC-f, where the latter firstly exploited the channel output feedback in training and surpass

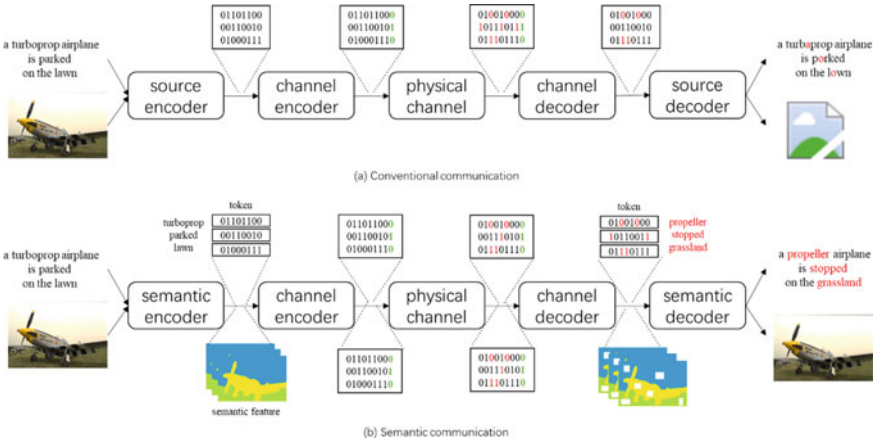


Fig. 1 Conventional communication and semantic communication

the traditional structured coding-based designs. However, the compression rate seems to be constrained by the bottleneck of CNN, lacking of the comparison with advanced image compression code. A lite semantic communication model for limited computing capability IoT devices is discussed in [4]. It proposed a low complexity text transmission model based on transformer and developed a channel state information (CSI) aided training processing to promise IoT devices to get the correct data and train the distributed model locally, but it is difficult to design a lite image transmission model which prefers convolution layers rather than dense layers.

2 System Model

Referencing to the block-based design in conventional communication system, we take the semantic communication system apart into two modules. One is for semantic feature extraction and the other is for semantic feature transmission. Observed that wireless channels in physical world change stochastically, and the random data makes the model training difficult to converge. There has a paradox that the data driven model usually requires the correct input for a more accurate interpretation which means DL-based channel coding module needs a huge number of parameters to learn statistical characteristics of channel. It significantly increases the model complexity and becomes unaffordable for IoT devices, otherwise model cannot provide helpful semantic features for convergence. The structure of semantic transmission system is shown in Fig. 2.

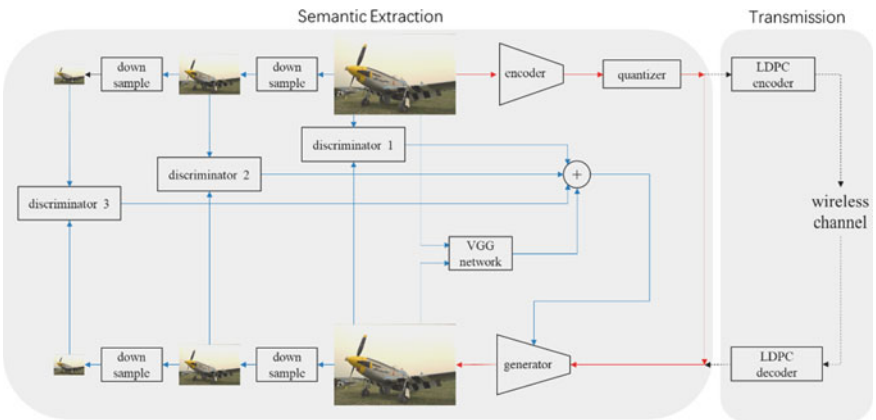


Fig. 2 System structure

2.1 Transmission Module

To simplify the semantic communication system, we adapted LDPC coding to keep the semantic information high fidelity while transmission. LDPC coding is a mature scheme and widely used in 5G. It provides a high efficiency and high reliability channel coding within low complexity, which can be implemented at low hardware cost.

2.2 Semantic Extraction Module

Once we solved the effect of fading channel with low consumption, more computing resource of IoT devices can be allocated into running a DL model. The semantic extraction module can concentrate on image extreme compression. The target of semantic extraction model can be modeled after the semantic rate-perception-distortion theory [5], which gives the limit of semantic information rate R within the conventional average distortion D and semantic perceptual distortion P, i.e.,

$$R(D, P) = \min_{p(\hat{s}|s)} I(s, \hat{s}) \quad \text{s.t.} \quad \mathbb{E}\left[\Delta(s, \hat{s})\right] \leq D, d[p_s, p_{\hat{s}}] \leq P \tag{1}$$

where $I(s, \hat{s})$ is the mutual semantic information between the transmitter s and the receiver $\hat{s}$. Since the divergence distance is considered as an effective index of semantic perceptual quality, we define the divergence distance between distributions as $d[\cdot, \cdot]$. GANs [6], which is firstly proposed by Goodfellow, has been demonstrating superior performance than CNNs on image application like generation, reconstruction and so on. The structure of alternately training the generator $G(\cdot)$ and the discriminator $D(\cdot)$ for a saddle point of min–max objective with the loss function, which is formulated as

$$L_{\text{GAN}}(G, D) = \max_{D} \mathbb{E}[f(D(s))] + \mathbb{E}[g(D(G(z)))] \tag{2}$$

is proven to be equivalent to measure the divergence distance between the probability distribution of the origin dataset and the generated dataset. The proposed training structure is based on the Least-Squares GAN [7], which corresponds to the Pearson χ^2 divergence with $f(x) = (x-1)^2$ and $g(x) = x^2$.

We adapt a multi-scale discriminator, which consisted of three independent and identical PatchGANs [8], to measure the semantic features loss of the origin image and its downsampled images. The quantizer q maps the encoder output from float-point number to integer. Moreover, we adopt VGG loss to navigate the model to generate low average distortion images. The optimization objective can be formulated as

$$\min_{E,G}\max_{d_k\in D}\left[\sum_{k=1}^{N_D} L_{\mathrm{LSGAN}}(G,d_k)+\lambda\mathbb{E}[\mathrm{VGGLoss}(s,G(q(E(s))))]\right] \tag{3}$$

where coefficient λ controls the rate of semantic perceptual distortion and conventional average distortion.

3 Model Compression and Acceleration

As revealed in the lottery ticket hypothesis [9], it is difficult to train a pruned model from start. A better solution is to prune while training. However, it is impracticable for performance-limited and power-limited IoT devices to retrain the pruned model within locally collected data. We tend to deploy the model one time and maintain the IoT devices if and only if necessary in the remote mountainous region with poor infrastructure. Although retraining with the backhaul data and redistributing the updated model is viable, the low bandwidth and high delay network might exhaust itself by such huge data transmission. Hence, a plug and play IoT device with a pre-compression is better suited for wildfire monitoring scenario. Considering that the IoT devices usually work in a preset position even the cameras face toward a fixed direction and we can deal with the trade-off between accuracy and convenience in a simpler way. In this section, we will give out some general model compression and acceleration method. These all operations directly act on a pretraining large model without fine-tuning.

3.1 Parameter Quantification

The parameters of model are set to floating-point numbers by default, which needs 32 bits to save a single parameter (FLOAT32). The operation and storage of a large number of FLOAT32 are tolerable for compute unified device architecture (CUDA) supported computer. However, IoT devices are equipped with limited performance CPUs and without GPUs. They are supposed to handle integer data and provide low floating-point operations per second (FLOPS). Hence, we can quantize the parameter from FLOAT32 to INT8, and a general uniform quantization is as

$$\begin{aligned} Q&=\frac{R}{S}+Z\\ R&=(Q-Z)\times S\\ S&=\frac{R_{\max}-R_{\min}}{Q_{\max}-Q_{\min}}\\ Z&=Q_{\max}-\frac{R_{\max}}{S} \end{aligned} \tag{4}$$

where Q, R represent the quantized value and real value, respectively. S is the scale of quantization and Z is the zero point. Z is usually equal to 0 exactly because it plays an important role in the whole model, which deserves a special treatment. Furthermore, to quantize the whole model, additional operation should be applicated. We firstly took several typical valves of dataset to translate the type of the input and output. Then we reinterpreted the model with the filtered dataset to help the activation layer calibrate.

3.2 Parameter Pruning and Clustering

To find the potential winning ticket, the over-parameterization models are widely accepted. The redundant parameters can help handle corner cases and improve robustness while training, however, they become a burden for inference instead in a simple scenarios like wildfire monitoring. A portion of parameters in the over-parameterization model have minor effect on inference accuracy in fact, which can be replaced by their statistical characteristics without a heavy cost. Parameter pruning can be controlled by a hyper-parameter named sparsity, which is ranging from 0 to 1. All the parameters are sorted and the parameters below the threshold are set to 0 brutally. Parameter clustering is similar but gentler. It build a set of statistical characteristics of parameters and represent the parameters with their index in the set. The details are shown in Fig. 3.

Both pruning and clustering do not modify the network structure directly and they just give out the sparse representation of the model for further compression. Because the calculation of most DL model is based on the computation graphs, which are divided into dynamic graph and static graph. The difference between both is that the latter is invariant once the network have initialized. We prefer the dynamic graph for clearer debugging procedure while the static graph for higher execution efficiency.

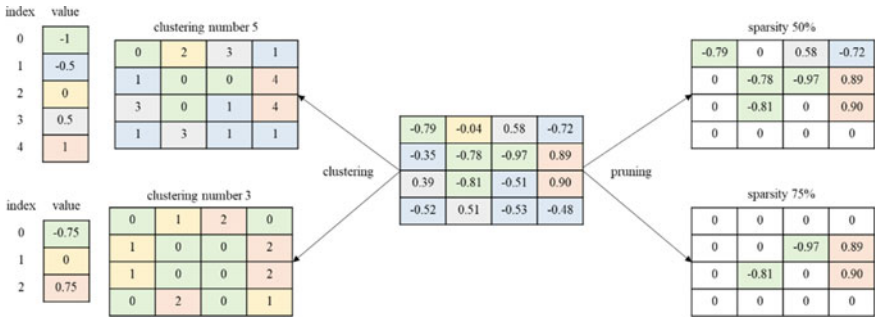


Fig. 3 Pruning and clustering

4 Numerical Results

In this section, we firstly compare the proposed semantic communication system with different conventional image compression. Then we compare the performance penalty of compressed model.

As shown in Fig. 4, the proposed semantic extraction module generated clearer images within less bit cost. Benefitting from the generator, which interprets and reconstructs with the received semantic information, the volume of data reduce significantly. Unlike the conventional image compression, they treat the every pixel of images equally and reconstruct blur results. The proposed model keeps a good visual perception separately in both the subject and the background of images but the border, which can be demonstrated from the difference of the flame shape. Figure 5 shows the compression performance of ours module and conventional compression. We can see that.

Figure 6 shows the transmission performance between the proposed system and DeepJSCC over the AWGN channel and Rayleigh channel. DeepJSCC performed better in low SNR environment with the additional training against channel fading. However, the fading effect is simulated by the complex Gaussion distribution and it is difficult to apply the theoretical results into practice. The semantic transmission system interpreted negative results from the incorrect semantic information but its

Fig. 4 Example of visual comparison

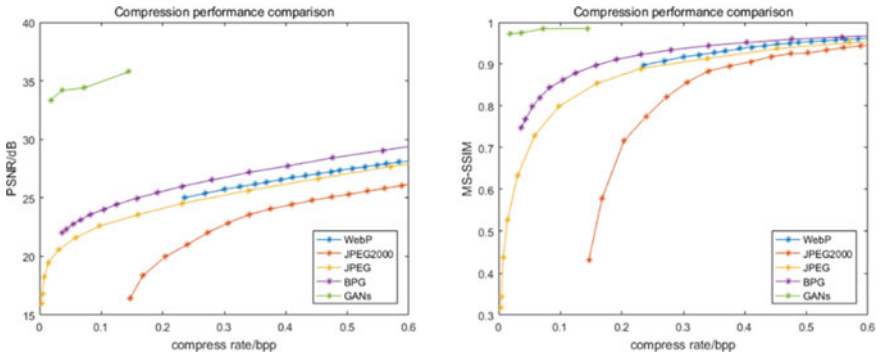


Fig. 5 Compression performance comparison

performance increases rapidly with as the SNR increases owing to the robustness of LDPC code.

Table 1 shows the result of the proposed model before and after the compression. Noticed that the number of parameters is constant because of the computation graph mentioned above. And the loss of the generated results are acceptable, where the

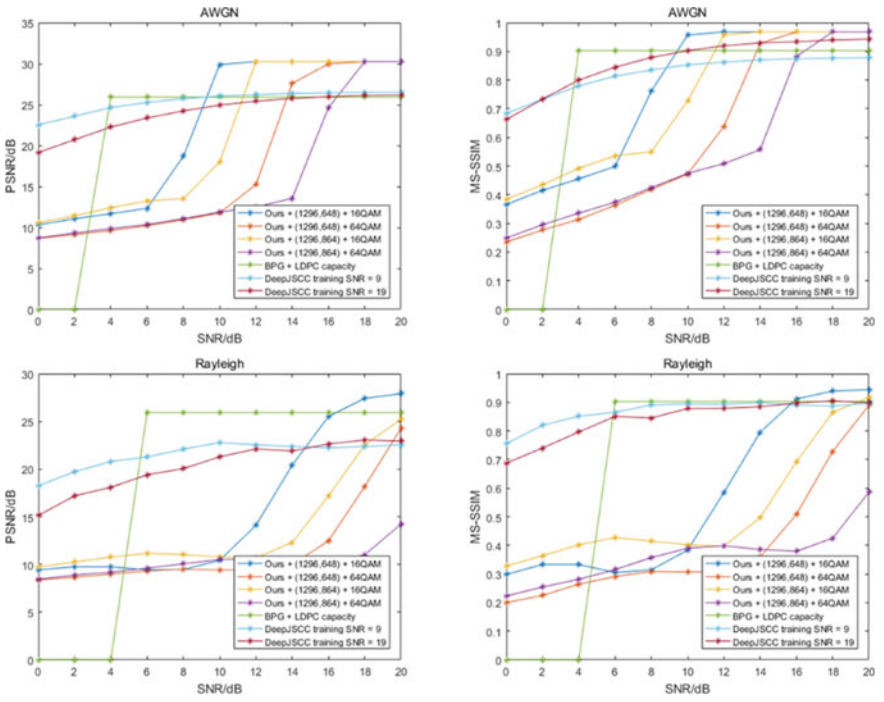


Fig. 6 Transmission comparison, where LDPC denotes (block length, information length)

Table 1 Model compress and acceleration comparison

Quantization	measurement	Semantic extraction module				Pruning	Measurement	Semantic extraction module			
		Encoder		Generator				Encoder		Generator	
Weight only (int8)		Before	After	Before	After	Sparsity 50%		Before	After	Before	After
	Number of parameter	5,591,544	5,591,544	154,942,563	154,942,563		Number of parameter	5,591,544	5,591,544	154,942,563	154,942,563
	Model size	21.8 MB	5.35 MB	592 MB	141.37 MB		Model size	21.8 MB	19.51 MB	592 MB	544.87 MB
		Before		After				Before		After	
	PSNR	30.3 dB		29.1 dB			PSNR	30.3 dB		25.7 dB	
	MS-SSIM	0.97		0.94			MS-SSIM	0.97		0.92	
Parameter all (int8)		Before	After	Before	After	Sparsity 75%		Before	After	Before	After
	Number of parameter	5,591,544	5,591,544	154,942,563	154,942,563		Number of parameter	5,591,544	5,591,544	154,942,563	154,942,563
	Model size	21.8 MB	5.34 MB	592 MB	131.24 MB		Model size	21.8 MB	19.48 MB	592 MB	520.43 MB
		Before		After				Before		After	
	PSNR	30.3 dB		19.8 dB			PSNR	30.3 dB		23.8 dB	
	MS-SSIM	0.97		0.65			MS-SSIM	0.97		0.9	

Clustering	Measurement	Semantic extraction module			
		Encoder		Generator	
Num = 8		Before	After	Before	After
	Number of parameter	5,591,544	5,591,544	154,942,563	154,942,563
	Model size	21.8 MB	2.89 MB	592 MB	80.01 MB
		Before		After	
	PSNR	30.3 dB		26.6 dB	
	MS-SSIM	0.97		0.93	

(continued)

Table 1 (continued)

Clustering	Measurement	Semantic extraction module			
		Encoder		Generator	
Num = 3		Before	After	Before	After
	Number of parameter	5,591,544	5,591,544	154,942,563	154,942,563
	Model size	21.8 MB	1.65 MB	592 MB	30.32 MB
		Before		After	
	PSNR	30.3 dB		23.4 dB	
	MS-SSIM	0.97		0.9	

Table 2 Runtime comparison

Model	Encode time/s	Decode time/s	Compress rate/bpp
JSCC	0.0235	1.027	8
Ours	1.2766	4.524	0.094
JPEG	0.0498	0.0173	1.416
JPEG2000	0.2717	0.1916	1.036
BPG	1.9982	0.3122	0.755

average of PSNR and MS-SSIM are above 25 dB and 0.9, respectively. The result of model size may be counterintuitive because the parameters consist of weights and bias and we found that the bias make a significant impact while interpreting, which can be proofed by the result of all parameters quantization. Thus we only operated on the weights of parameters. Table 2 shows the speed of our system and the others, where we took PSNR 30 dB as the benchmark with the Raspberry Pi 4B. We see that the proposed compressed model spends more time on semantic coding due to the limit parallel computing power. However, considering the poor network in the wilderness, we should pay more attention to reducing the transmission delay with the higher compression rate. In such scenario, a feasible structure is that the IoT devices encode and send the semantic information while the center completes the semantic reconstruction. Besides, we performed the simulation again in GPU by the computer with NVIDIA GeForce RTX 2080 Ti, and it took 0.7 s to complete the semantic reconstruction of one frame.

5 Conclusion

In this paper, we proposed an available semantic communication for IoT devices, which can work in a limit computing capabilities environment. Unlike the DeepJSCC, we considered the design of source semantic coding and channel coding separately. The former, which is based on the LSGAN, helped image extreme compression by extracting the semantic information while the latter provided an efficient and inexpensive method to keep the semantic information low distortion over the fading channel. To avoid the model retraining and the data backhaul, we compress the parameters of the pre-trained model directly by quantization, pruning and clustering. The simulation result demonstrated that the proposed semantic communication provided a higher compress rate and better reconstruction than other systems within comparable runtime.

References

1. Luo X, Chen H-H, Guo Q (2022) Semantic communications: overview, open issues, and future research directions. IEEE Wirel Commun 29(1):210–219. https://doi.org/10.1109/MWC.101.2100269
2. Bourtsoulatze E, Burth Kurka D, Gündüz D (2019) Deep joint source-channel coding for wireless image transmission. IEEE Trans Cognitive Commun Netw 5(3):567–579. https://doi.org/10.1109/TCCN.2019.2919300
3. Kurka DB, Gündüz D (2020) DeepJSCC-f: deep joint source-channel coding of images with feedback. IEEE J Selected Areas Inf Theor 1(1):178–193. https://doi.org/10.1109/JSAIT.2020.2987203
4. Xie H, Qin Z (2021) A lite distributed semantic communication system for Internet of Things. IEEE J Sel Areas Commun 39(1):142–153. https://doi.org/10.1109/JSAC.2020.3036968
5. Blau Y, Michaeli T (2018) The perception-distortion tradeoff. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, Salt Lake City, UT, USA, pp 6228–6237. https://doi.org/10.1109/CVPR.2018.00652
6. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: Advances in neural information processing systems, pp 2672–2680
7. Mao X, Li Q, Xie H, Lau RYK, Wang Z, Smolley SP (2017) Least squares generative adversarial networks. In: IEEE international conference on computer vision (ICCV). IEEE, pp 2813–2821
8. Isola P, Zhu J-Y, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), Honolulu, HI, USA, pp 5967–5976. https://doi.org/10.1109/CVPR.2017.632
9. Frankle J, Carbin M (2019) The lottery ticket hypothesis: finding sparse, trainable neural networks. In: 7th international conference on learning representations, ICLR 2019, New Orleans, LA, USA, 6–9 May 2019, pp 1–42

Image Super Resolution Reconstruction Algorithm Based on Multiple Prior Constraints

Ting Liu and Kun Wang

Abstract High spatial resolution is necessary for several applications such as visual inspection. However, the conflict between resolution and image distance limits the applications of image devices. In this paper, a super-resolution framework with multiple priors is proposed. Firstly, the directional generalized total variational and the non-local self-similar constraint are incorporated to enhance image texture details and smooth edge effects. Especially, an adaptive Gaussian kernel is used to better descript the non-local prior. Secondly, the proposed multi-constraint problem is solved by the alternate direction multiplier method. Generally, a large number of qualitative and quantitative results demonstrated the effectiveness and superiority of our method over traditional methods.

Keywords Image super-resolution reconstruction · Non-local self-similarity · Second-order total variation · Multi-constraint optimization problem

1 Introduction

Image super-resolution (SR) aims at full reconstructing the rich details to restore low-resolution (LR) images to high-resolution (HR) images. The loss of details can be due to various degrading factors such as blur, decimation, noise or hardware limitations. For example, low resolution optical and acoustic images will be acquired with a large field of view [1–3]. SR is an ill-posed problem because for each LR image patch the number of corresponding HR image patches can be very large.

SR includes single-image super-resolution (SISR) and multi-frame super-resolution. The focus of this paper is SISR. SISR reconstruction can be categorized into interpolation-based method, learning-based method, and regularization-based method. The image pixels are directly processed by interpolation-based method [4], but the similarity of the internal structure is not taken into account. Learning-based methods [5] are popular for SISR reconstruction, but large datasets are required.

T. Liu (✉) · K. Wang
College of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, China
e-mail: liuting0910@dlmu.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_38

Extensive studies have been conducted on exploring appropriate prior for regularization-based method [6]– [13]. Wu et al. [6] has shown success in obtaining more accurate restoration effects by directional generalized total variational (DTGV), but step effects are generated. However, the non-local self-similar regularization (NLM) can reduce edge ladder effects and the process of blur can enhance the self-similarities within images, then the NLM constraint is widely used in [9–12].

High resolution images cannot be effectively recovered with a single priority, DTGV and NLM are employed in this paper to capture the edge information and higher-order details of images with specific directivity at the same time. The main contributions can be summarized as follows.

- A new NLM prior is designed by introducing adaptive Gaussian kernel to better remove the noise artifacts and blur from the LR image.
- NLM and DTGV are used to form a multi-constraints optimization problem for SISR, while the proposed model is effectively solved by the ADMM algorithm.

The structure is described as follows. The proposed SR model is presented in Sect. 2 and the solving method is detailed demonstrated in Sect. 3. Experimental results have been analyzed in detail in Sect. 4. Conclusion is drawn in Sect. 5.

2 The Proposed SR Model

The proposed SR model is described in detail, including NLM and DTGV.

2.1 Second-Order Directional Total Generalized Variation Prior Model

The anisotropic version of the directional total variational formula will effectively avoid the ladder effect on the reconstructed image edge, and texture details also will be better protected. DGTV is proposed [8] to incorporate the directional information of the image gradient while maintaining the advantages of total variation, which is shown as follows

$$\begin{aligned} &A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, W_\alpha = \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix}, W_\alpha^* = W_\alpha, A_\theta^* = A_{-\theta}, \\ &||X||_{DTGV} = ||W_\alpha^* A_\theta^* \nabla X - p||_1 + ||\varepsilon(p)||_1 \end{aligned} \tag{1}$$

where A_θ is a rotation transformation matrix and A_θ^* is the conjugate operator of A_θ. W_α is a scale contraction matrix and α represents the scale of expansion. ∇ is a linear operator that combines the horizontal and vertical difference operators. θ is

the directional total variation principal direction angle. $\varepsilon(p)$ represents symmetric difference.

In fact, the value of p lies within the range of 0 and $W\alpha A\theta\nabla X$. This means that in the edge area of the image, the reconstruction effect of DTGV is similar to total variational. However, in the smooth area, DTGV is represented in a smoother way rather than a jagged way, resulting in a reduction of jagged artifacts.

2.2 Non-local Similarity Prior Model

Repetitive structure and redundant information exist in natural images. Non-local pixels or blocks within the image tend to display similarities in the content and structure.

To fully exploit the similarity between columns, the $L_{2,1}$ norm [10] is proposed to better obtain the sparsity relationship between adjacent whole columns. Then the NLM term is shown in Eq. (2).

$$R \in R^{N\times N-1} = \begin{bmatrix} -1 & & \\ 1 & \ddots & \\ & \ddots & -1 \\ & & 1 \end{bmatrix} \tag{2}$$

$$||X||_{NLM} = ||XR||_{2,1} = [X_2 - X_1, X_3 - X_2, ..., X_N - X_{N-1}]$$

$||XR||_{2,1}$ represents the sparsity of each column counted, and then calculates the sparsity of N columns. R represents the relationships between two adjacent columns, N represents the Nth image block similar to the reference image block, and X_N represents the matrix of similar block groups of the Nth reference block.

Non-local self-similarity has the problems of excessive smoothing of image, loss of texture edge information and lack of self-adaptability of the model. Thus, we utilize an adaptive Gaussian kernel in the iterative process of NLM. The size of the Gaussian kernel is determined by the degradation process of the image and the edge features surrounding each pixel.

In general, the image degradation process can be modeled as Eq. (3):

$$Y = \Gamma(H * X) + n = \Gamma\mathcal{F}^H(\mathcal{F}(H)\mathcal{F}(X)) + n \tag{3}$$

where Y is the LR image and H represents the blurry kernel. X is the HR image. n represents the noise. Γ is a down-sampling operator and * represents the convolution operator. The inverse Fourier transform is denoted by $\mathcal{F}^H$, and $\mathcal{F}$ is the Fourier transform.

The goal of SR reconstruction is to reconstruct the HR image from the LR image. To solve this ill-posed problem, the minimal problem is defined:

$$\hat{X} = \arg\min_{X} \frac{1}{2}||\Gamma\mathcal{F}^{H}(\mathcal{F}(H)\mathcal{F}(X)) - Y||_F^2 + \alpha R(X) \tag{4}$$

where the image prior is represented as $R(X)$ and α denotes the regularization coefficient. The performance of the algorithm is highly dependent on the design of $R(X)$.

A main direction for each block selected in the DTGV, leading to uneven blending between adjacent blocks. While the NLM constraint can smooth blending between adjacent blocks. Thus, we have combined the NLM and DTGV constraint to improve image quality. The problem of solving the minimum value of an estimated image $\hat{X}$ is shown in Eq. (5).

$$\begin{aligned}\hat{X} = \arg\min_{X} &\frac{\beta}{2}||\Gamma\mathcal{F}^{H}(\mathcal{F}(H)\mathcal{F}(X)) - Y||_F^2 + \alpha_1 \sum_{i=1}^{2} ||W_\alpha^* A_\theta^* \nabla X(i,j) - p_i||_1 \\ &+ \alpha_2 \sum_{j=1}^{4} ||(\varepsilon(p))_j||_1 + \alpha_3 ||XR||_{2,1}\end{aligned} \tag{5}$$

where α_1,α_2 and α_3 are regularization parameters and β is a coefficient parameter. The first term is fidelity term. The second and third terms describe the DTGV prior, which promotes spatial smoothness and preserves sharp edges in the reconstructed image. The fourth term is the NLM prior. Some auxiliary variables are set to represent the compound variables: $\nabla = (M_1, M_2)$, $\varepsilon(p) = \frac{(\nabla p + \nabla p^T)}{2}$, $p = (p_1, p_2)$, then

$$\begin{aligned}(\varepsilon(p))_1 &= M_1(X(i,j))p_1 \\ (\varepsilon(p))_2 &= M_2(X(i,j))p_2 \\ (\varepsilon(p))_3 &= \frac{M_2(X(i,j))p_1 + M_1(X(i,j))p_2}{2} \\ (\varepsilon(p))_4 &= \frac{M_2(X(i,j))p_1 + M_1(X(i,j))p_2}{2}\end{aligned} \tag{6}$$

3 ADMM Based Solve Method

The problem in Eq. (5) can be solved by ADMM. By splitting variables, it can be written as

$$\begin{aligned}\hat{X} = \arg\min_{X} &\frac{\beta}{2}||\Gamma\mathcal{F}^{H}(\mathcal{F}(H)\mathcal{F}(X)) - Y||_F^2 \\ &+ \alpha_1 \sum_{i=1}^{2} ||W_\alpha^* A_\theta^* \nabla X(i,j) - p_i||_1 + \alpha_2 \sum_{j=1}^{4} ||(\varepsilon(p))_j||_1 + \alpha_3 ||XR||_{2,1}\end{aligned}$$

$$s.t.S = XR, D_{\mathrm{j}} = (\varepsilon(p))_j, C_i = W_\alpha^* A_\theta^* \nabla X(i, j) - p_i = K_i(X) - p_i \quad (7)$$

To solve the model with multiple norms, an augmented Lagrange function is constructed

$$\begin{aligned} L(X, S, p_i, r_{\mathrm{i}}, q_j, \lambda_3) = {} & \frac{\beta}{2}||\Gamma \mathcal{F}^H(\mathcal{F}(H)\mathcal{F}(X)) - Y||_F^2 + \alpha_1 \sum_{i=1}^{2} ||C_i||_1 \\ & + \alpha_2 \sum_{j=1}^{4} ||D_j||_1 \\ & + \alpha_3 ||S||_{2.1} + < \lambda_3, S - XR > + < q_j, D_j - (\varepsilon(p))_j > \\ & + < r_{\mathrm{i}}, p_i + C_i - K_i(X) > + \mu_1 \sum_{\mathrm{i}=1}^{2} ||p_i + C_i - K_i(X)||_2^2 \\ & + \mu_3 ||S - XR||_2^2 + \mu_2 \sum_{j=1}^{4} ||D_j - (\varepsilon(p))_j||_2^2 \end{aligned} \quad (8)$$

where μ_1, μ_2, μ_3 are augmented Lagrange multipliers, and $\langle \cdot, \cdot \rangle$ denotes the matrix inner product. The objective function is decomposed into sub-problems by ADMM.

The solution of C_i subproblem is given by a soft thresholding operator in Eq. (9).

$$C_i^{(k+1)} = shrink(r_i + K_i(X) - p_i, \frac{1}{\mu_1}), i = 1, 2 \quad (9)$$

D_j is updated by using the iterative shrinkage operator, the problem can be easily solved by imposing soft-thresholding operation.

$$D_j^{(k+1)} = shrink(q_j + (\varepsilon(p))_j, \frac{1}{\mu_2}), j = 1, 2, 3, 4 \quad (10)$$

the closed solution of S can be obtained by the following formula (11)

$$S^{(k+1)} = \begin{cases} \dfrac{||XR - \lambda_3|| - \frac{1}{\mu_3}}{||XR - \lambda_3||}(XR - \lambda_3), if||XR - \lambda_3|| > \dfrac{1}{\mu_3} \\ 0, otherwise \end{cases} \quad (11)$$

For simplicity of calculation, $Z(X)$ represents a composite operator of blurring and down-sampling $\Gamma \mathcal{F}^H(\mathcal{F}(H)\mathcal{F}(X))$. Then the estimate of X, p_1 and p_2 can be solved respectively by conjugate gradient method in Eqs. (12), (13) and (14).

$$\begin{aligned}&(\alpha_3\mu_3 R^{-1}R^T + \alpha_1\mu_1 \sum_{i=1}^{2} K_i^T K_i + \beta \mathrm{Z}^T \mathrm{Z})X \\ &= \beta \mathrm{Z}^T ZY + \alpha_1\mu_1 \sum_{i=1}^{2} K_i^T (p_i + C_i - r_i) + \alpha_3\mu_3 R^T (S - \lambda_3)\end{aligned} \tag{12}$$

$$\begin{aligned}&(\alpha_1\mu_1 + \alpha_2\mu_2 M_1^T M_1 + \frac{\alpha_2\mu_2 M_2^T M_2}{2})p_1 = \alpha_1\mu_1(r_1 + K_1 X - C_1) \\ &+ \alpha_2\mu_2 M_1^T (D_1 - q_1) + \alpha_2\mu_2 M_2^T (-\frac{M_1 p_2}{2} + D_3 - q_3)\end{aligned} \tag{13}$$

$$\begin{aligned}&(\alpha_1\mu_1 + \alpha_2\mu_2 M_2^T M_2 + \frac{\alpha_2\mu_2 M_1^T M_1}{2})p_2 = \alpha_1\mu_1(r_2 + K_2 X - C_2) \\ &+ \alpha_2\mu_2 M_2^T (D_2 - q_2) + \alpha_2\mu_2 M_1^T (-\frac{M_2 p_1}{2} + D_3 - q_3)\end{aligned} \tag{14}$$

where λ_3, r_i and q_j are the scaled dual variables, updating in Eq. (15).

$$\begin{cases} \lambda_3^{(k+1)} = \rho_3(S^{(k+1)} - X^{(k+1)}R) + \lambda_3^{(k)} \\ q_{\mathrm{j}}^{(k+1)} = \rho_2((\varepsilon(p))_j^{(k+1)} - D_j^{(k+1)}) + q_j^{(k)}, j = 1, 2, 3, 4 \\ r_i^{(k+1)} = \rho_1(K_i(X^{(k+1)}) - \mathrm{p}_i^{(k+1)} - \mathrm{C}_i^{(k+1)}) + q_i^{(k)}, i = 1, 2 \end{cases} \tag{15}$$

where ρ_1, ρ_2 and ρ_3 are the step-sizes which are chosen to guarantee the convergence of the above iterations.

4 Experimental Results

To demonstrate the effectiveness of the proposed method, both qualitative and quantitative experimental analyses will be conducted in this section. The proposed algorithm is implemented by using MATLAB R2019a running on a computer with an Intel Core-i7 processor, 8 GB RAM, and 500 GB HDD. The LR images used in this paper is acquired from HR images. HR images are blurred by a 5×5 Gaussian kernel, and then random noise is added, finally the image is down-sampled. The regularization parameters α_1,α_2 and α_3 are set to 6, 2 and 0.05. Augmented Lagrange multipliers μ_1, μ_2 and μ_3 are set to 0.2, 0.9 and 0.07. The step sizes ρ_1,ρ_2 and ρ_3 used in algorithm are set to 0.1,0.1 and 0.5. The tolerance $tol = 10^{-3}$, coefficient parameter $\beta = 0.5$ and $iteration = 800$.

4.1 Comparative Experiments

Comparative experiment is conducted to verify the effectiveness of the proposed method from both qualitative and quantitative perspectives. Each model's quantitative results are shown in Table 1, the first value is the SSIM value and the second one is the PSNR value, and qualitative results are shown in Fig. 1. Figure 1 shows the proposed algorithm NLMDTGV can preserve the edge information of the image well in the field of super resolution reconstruction of optical and acoustic images.

Although the PSNR of NLM is higher than that of other methods, the literature [6] shows that the use of SSIM seems to be more reasonable for these texture images. The methods with Bicubic, DTGV [7], and CSR [13] all conduct poor data results in SSIM. It is mainly because the separate reconstruction of each patch without considering the structural correlations of neighboring patches. The proposed method shows more richer textures by incorporating the structural correlations into the variational frame.

Table 1 Comparisons of peak signal-to-noise ratio (PSNR) values and SSIM values for test images with a scale factor of 3, noise level $\sigma = 10$ in different super-resolution approaches

Methods	Test image			
	Image-c	Image-a	Image-b	Image-h
Bicubic	0.6326/15.2323	0.9278/23.9144	0.7368/18.0087	0.7151/19.5639
NLM	0.6385/*23.9121*	0.9458/*32.1537*	0.7459/*25.0665*	0.7234/*26.9833*
DTGV [7]	0.6496/23.6527	0.9678/30.3071	0.7635/24.6063	0.7402/26.3327
CSR [13]	0.6526/23.5642	0.9618/29.6411	0.7569/24.2773	0.7414/26.2173
NLMDTGV	*0.6511*/23.6651	*0.9712*/30.4478	*0.7672*/24.5620	*0.7426*/26.3501

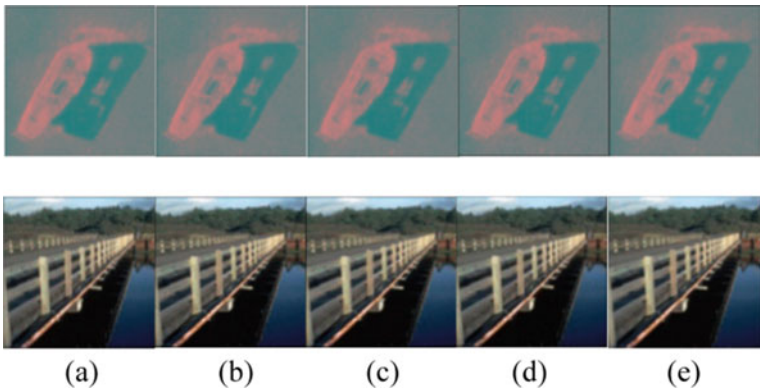


Fig. 1 Reconstructed results for various incomplete color image with a scale factor of 3, noise level $\sigma = 10$. The results from left to right: **a** Bicubic, **b** NLM, **c** DTGV, **d** CSR, **e** NLMDTGV

Table 2 Structural similarity (SSIM) results on test images with different degenerate parameters

Image	SR Methods	Noise level $\sigma = 5$		Noise level $\sigma = 10$		Noise level $\sigma = 30$	
		$\times 2$	$\times 3$	$\times 2$	$\times 3$	$\times 2$	$\times 3$
Image-g	Bicubic	0.8721	0.8259	0.8467	0.8067	0.6702	0.6599
	NLM	*0.8765*	0.8288	0.8486	0.8077	0.7080	0.6723
	DTGV	0.8751	0.8295	0.8627	0.8179	0.8036	0.7569
	NLMDTGV	0.8764	*0.8313*	*0.8642*	*0.8204*	*0.8156*	*0.7639*

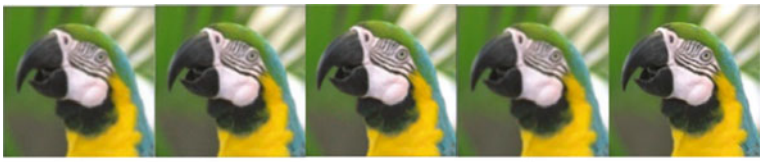


Fig. 2 Reconstructed results for image-c with a scale factor of 2, noise level $\sigma = 10$. The recovered results by **a** Bicubic, **b** NLM, **c** DTGV, **d** NLMDTGV, **e** the original image

4.2 Experiments with Different Levels of Random Noise

We present SR results of all comparison methods to test the robustness of these methods against noise. The proposed method is compared with Bicubic, NLM [10], and DTGV [7], the results are shown in Table 2 and Fig. 2. Compared with the traditional method, the improved method is more suitable to solve the problem of serious noise. Fewest noise and jaggy artifacts remained in Fig. 2 (g).

Compared with the traditional method, the improved method is more suitable to solve the problem of serious noise pollution because of adaptive Gaussian kernel. The best values are marked in incline in Table 2. Figure 2 shows that the noise in (g) is removed better than other methods, although fewer noise and jaggy artifacts remain.

5 Conclusions

The directional generalized total variational and the non-local self-similar constraint are incorporated and a super-resolution framework with multiple priors is proposed in this paper. Although results of proposed method have fewer noise and jaggy artifacts in the images, our method can avoid the stepped effect of the DTGV regular term and enhance image texture details.

We conduct several experiments on image denoising and super-resolution to evaluate the effectiveness of the proposed method. Experimental results show that the

proposed method achieves noticeably improved results. The proposed method can provide images whose numerical and visual qualities are higher than those obtained by traditional methods.

References

1. Chen WL, Yang Z, Zhang HQ, Zhu Y, Liu YT (2019) Super-Resolution Restoration for Sonar Images. Comprehensive Comparison, Communications in Computer and Information Science 1009:55–64
2. Chen LB, Zhan YC, Wang XC et al (2019) Super resolution reconstruction method of underwater image based on deep learning. Computer Applications 39(09):2738–2743
3. Desai, C., Tabib, R.A., et al.: Realistic underwater image generation towards restoration. In: CVF Conference on Computer Vision and Pattern Recognition Workshops. pp. 2181–2189 (2021)
4. Zhang Y, Fan Q, Bao F, Liu Y, Zhang C (2018) Single-image super-resolution based on rational fractal interpolation. IEEE Trans Image Process 27(8):3782–3797
5. Guo, Y., et al.: Dual regression networks for single image super-resolution. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5406–5415 (2020)
6. Wang, Q., Wu, Z., Sun, M., et al.: Single-image super-resolution using directional total variation regularization and alternating direction method of multiplier solver. Journal of Electronic Imaging **24**(2) (2015)
7. Wu ZH, Sun MJ, Gu ZS, Fan MY (2017) Image super resolution reconstruction method based on second order generalized directivity total variation. J Electron Imaging 45(11):2625–2632
8. Daneshmand PG, Mehridehnavi A, Rabbani H (2021) Reconstruction of optical coherence tomography images using mixed low rank approximation and second order tensor based total variation method. IEEE Trans Med Imaging 27(3):865–878
9. Zha Z, Yuan X, Wen B, Zhou J, Zhang J, Zhu C (2020) From rank estimation to rank approximation: rank residual constraint for image restoration. IEEE Trans Image Process 29:3254–3269
10. Lu, H.R.: Research on restoration method of cultural relic image based on rank minimization. Jiangxi University of Science and Technology (2016)
11. Shi, M.Z., Gong, X.W.: Parameters identification via cepstrum analysis for Mix blurred image. In: Liang, Q., Mu, J., Jia, M., Wang, W., Feng, X., Zhang, B. (eds) Communications, Signal Processing, and Systems. CSPS 2017. Lecture Notes in Electrical Engineering, vol 463, pp. 1515–1521. Springer, Singapore (2019)
12. Huang S, Sun J, Yang Y, Fang Y, Lin P, Que Y (2018) Robust single-image super-resolution based on adaptive edge-preserving smoothing regularization. IEEE Trans Image Process 27(6):2650–2663
13. Dong W, Zhang L, Shi G, Li X (2013) Nonlocally centralized sparse representation for image restoration. IEEE Trans Image Process 22(4):1620–1630

The Channel Selection in the Analysis of Binocular Disparity EEG Data Processing

Yuhang Shi, Tingting Zhang, Wei Zhou, Ling Xia, Yi Mao, and Xiaofeng Liu

Abstract Binocular disparity is a crucial cue for perceiving depth, and investigating its neural mechanisms using EEG can deepen our understanding of stereo vision. However, analyzing EEG data is often complex and time-consuming, which may hinder researchers' efficiency. To address this issue, we designed a channel selection method based on average power spectral density (PSD) and Iterative Deepening Search (IDS) to streamline data analysis while maintaining accuracy. We collected EEG signals from four subjects while they viewed random dot stereograms (RDS) with disparities of 0″ and 1000″. Our results demonstrate that by reducing the number of channels by half, the proposed channel selection method improved the classification accuracy of the two disparities by 20% compared to using all channels. This approach provides a promising tool for efficient and accurate EEG data analysis in the investigation of binocular disparity and stereo vision.

Keywords Binocular disparity · Random dot stereogram · EEG channel selection · Power spectral density

1 Introduction

In recent years, 3D materials have become increasingly common in our daily lives, which can bring more realistic experiences in the fields of entertainment, medical care, and education [1]. However, viewing stereoscopic images can cause discomforting symptoms such as eye strain, difficulty focusing, and headaches due to the process of fusing binocular disparity in the brain [2]. Consequently, more and more researchers are turning to EEG technology to explore the connection between stereo vision and brain activity.

While having a large number of electrodes can enhance the spatial–temporal resolution and signal-to-noise ratio of the signal, it can also result in information

Y. Shi · T. Zhang (✉) · W. Zhou · L. Xia · Y. Mao · X. Liu
Hohai University, Changzhou, China
e-mail: zhangtingting@hhu.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_39

redundancy and noise interference due to the volume conduction effect, which ultimately affects the signal quality and accuracy [3]. To address this issue, channel selection algorithms are used to effectively eliminate redundant information and reduce feature dimensions. Thus, finding the right balance between the number of channels and signal accuracy is a crucial consideration.

Channel selection can be achieved using various methods, including manual, statistical, signal processing, and machine learning approaches. The manual method, which is subjective and inefficient, has been replaced by automatic methods. Statistical methods leverage the statistical properties of the data to enhance signal discriminability effectively [4, 5]. Signal processing methods typically rely on time-domain or frequency-domain characteristics, which may have a low signal-to-noise ratio [6, 7]. The machine learning methods combine statistical and signal characteristics of the data, leading to a significant improvement in signal discriminability [8, 9].

Channel selection methods have been primarily applied to datasets related to motion imagination and emotion recognition [10], creating a theoretical gap in stereo vision research. To address this gap, we proposed a channel selection method based on average PSD and IDS for accurately classifying stimuli with varying levels of binocular disparity.

2 EEG Data Collection and Preprocessing

We recruited four participants for our study who were asked to watch square RDS stimuli with a disparity of $0''$ and $1000''$.The stimuli were presented on a screen with a resolution of 1920×1080 in a well-lit and quiet room, and the participants wore anaglyph glasses to view the stereoscopic images. The viewing distance was set at 1.6 m in front of the display to create a realistic 3D impression. As depicted in Fig. 1, each condition was presented 50 times for a duration of four seconds, with a one-second " + " symbol displayed between stimuli. All participants had normal or corrected-to-normal visual acuity and stereoscopic vision, ensuring that the data collected during the study was reliable and accurate.

EEG data were recorded throughout the experiment using a Neuroscan 64-channel EEG cap and NuAmps amplifier with a sampling frequency of 1000 Hz. The preprocessing process includes 0.5–30 Hz bandpass filtering, re-referencing with the mean value of M1 and M2, and ICA for artifact removal (such as eye movement, electromyographic noise, and linear noise).

3 Channel Selection Method

In order to select the best EEG channel combination, we proposed a channel selection method based on average PSD and IDS, which was called PSD-ID. The PSD-ID involves three main steps which are depicted in Fig. 2. First, the average PSD of each

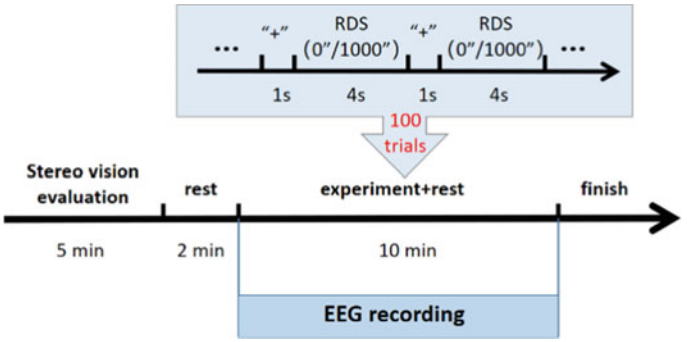


Fig. 1 The experimental procedure

channel was calculated and the channels were classified according to their weights. Second, the IDS method was employed to combine the channels. In the end, EEG data features were extracted under different channel combinations, and the classification performance was calculated to choose the best channel combination.

Channel Weight Sort: PSD provides a visualization of the energy distribution of EEG signals in different frequency bands, allowing the contribution of each channel in different frequency bands to be evaluated. The higher the PSD value of the channel, the higher the energy of the channel in the corresponding frequency band, indicating that the channel contributes more to the EEG signal in the frequency band. Thus, the total PSD value of each channel was divided by the sum of the total PSD of all channels to obtain the weight percentage of each channel in this study.

Channel Combination: IDS is an improved algorithm that deepens the search depth to find the target node or reach the maximum search depth. In this study, the

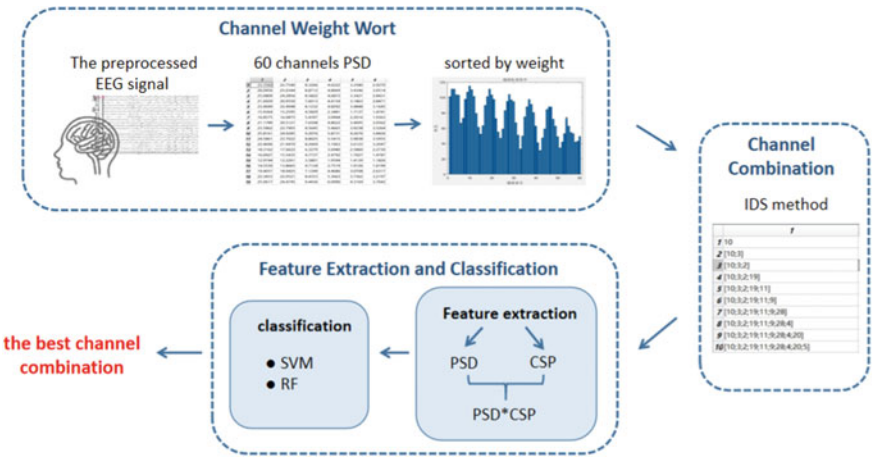


Fig. 2 The PSD-ID channel selection method

channel with the highest PSD weight was selected as the initial channel, and then other channels were gradually added until all channels are traversed. This method yielded a total of 60 channel combinations.

Feature Extraction and Classification: To comprehensively characterize the highly time-varying and spatially distributed EEG signals, feature extraction is needed from different perspectives. In this study, two features, PSD and CSP, were selected from the frequency domain and the spatial domain respectively. The PSD of EEG signals and the signal power of four frequency bands (Delta, Theta, Alpha, and Beta) under each channel were stacked together to form the PSD feature vector. Meanwhile, the CSP projection matrix was calculated, and the CSP filter corresponding to the highest and lowest eigenvalues was selected. The EEG signal was then projected onto the CSP filter to obtain the CSP feature. In addition, this study applied a convolution of these two features to generate a new feature, PSD * CSP.

This study used support vector machine (SVM) and Random Forest (RF) classifiers to compare the classification performance of two kinds of disparity under different channel combinations and features. The classification performance was evaluated in terms of accuracy, specificity, and sensitivity, which were calculated as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{3}$$

where TP and TN are the cases correctly predicted by the classifier in assigning to the positive or negative class, FP and FN are the cases incorrectly predicted by the classifier in assigning to the positive or negative class.

4 Results and Discussion

Channel weight percentage and corresponding brain topographic map are shown in Fig. 3. In Fig. 3b, the darker the color, the higher the weight of the region. The results show that the frontal lobe contributed the most to the EEG signals, with its weight gradually decreasing to the surrounding regions.

We selected three channel combinations and calculated the percentage contribution of each brain region, as shown in Table 1. The results indicate that the frontal and parietal lobes contributed more to classification accuracy, while the temporal lobe contributed less. These findings are generally consistent with the brain topography map of channel weight distribution. This may be attributed to differences in brain function. With the frontal lobe playing a key role in cognitive control and

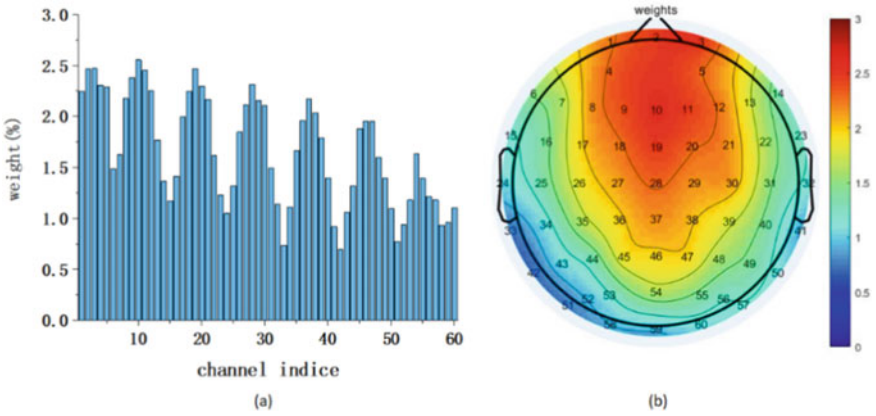


Fig. 3 Channel weight percentage and corresponding brain topographic map

decision-making and the parietal lobe being crucial for spatial perception and attention, neuronal activity in these regions is more intense during the stereoscopic task. In contrast, the temporal lobe, which is primarily associated with the recognition and memory of stimuli during visual processing, plays a minor role in stereoscopic tasks.

We further evaluated the classification performance of these channel combinations under different features and classifications. The results are presented in Fig. 4. Fig. 4a, b show the comparison of classification performance using different channel combinations and features with SVM and RF classifiers. In Fig. 4a, when using PSD as the feature and SVM as the classifier, the first channel combination yielded the highest accuracy of 66.25%, which was 20% higher than using the whole channel. When using PSD * CSP as the feature and SVM as the classifier, the second channel combination performed relatively well with an accuracy of 10% higher than using the whole channel. In Fig. 4b, when using CSP as the feature, the classification performance of all three channel combinations under the RF classifier was significantly better than under the SVM classifier, with accuracies ranging from 60 to 65%.

To sum up, PSD and PSD * CSP performed better with SVM, while CSP performed better with RF. It could be attributed to the dimension of the features and classification algorithm. SVM is better suited for high-dimensional spatial classification, so the high-dimensional features of PSD and PSD * CSP support SVM well. In contrast, RF is more effective in classifying multiple low-dimensional and differentiable features, whereas the low-dimensional features of CSP make it a better fit for RF. These findings provide a practical basis for selecting suitable features and classifiers in future studies. Both channel combinations proposed in this study significantly improved the classification performance of binocular disparity images, but from the perspective of the number of channels and the improvement of classification accuracy, the first channel combination was more effective. Compared with the full channel, the classification accuracy of 0″ and 1000″ disparity was improved by 20% based on reducing the number of channels by 1/2.

Table 1 Basic information and contribution rate of the channel combinations

	Channel combination 1		Channel combination 2		Full channel combination	
	Channels(31)	Percentage (%)	Channels(40)	Percentage (%)	Channels(60)	Percentage (%)
Frontal lobe	FZ FP2 FPZ FCZ F2 F1 FC2 F4 FP1 FC1 F3 FC4 FC3 F6 F5	51.07	FZ FP2 FPZ FCZ F2 F1 FC2 F4 FP1 FC1 F3 FC4 FC3 F6 F5 FC6 FC5 F8	48.12	FZ FP2 FPZ FCZ F2 F1 FC2 F4 FP1 FC1 F3 FC4 FC3 F6 F5 FC6 FC5 F8 FT8 FT7	40.36
Parietal lobe	CZ CPZ C2 C1 C4 CP2 CP1 C3 CP4 CP3	30.65	CZ CPZ C2 C1 C4 CP2 CP1 C3 CP4 CP3 P4 C6	29.45	CZ CPZ C2 C1 C4 CP2 CP1 C3 CP4 CP3 P4 C6 C5 P3 CP5 P5	28.04
Occipital lobe	P2 PZ P1 POZ	11.28	P2 PZ P1 POZ P7 P6 PO4	14.83	P2 PZ P1 POZ P7 P6 PO4 PO8 PO3 O2 P8 OZ PO5 O1 PO7 P7	20.55
Temporal lobe	AF3 AF4	7.00	AF3 AF4 CP6	7.60	AF3 AF4 CP6 PO6 T8 T7 TP8 TP7	11.05

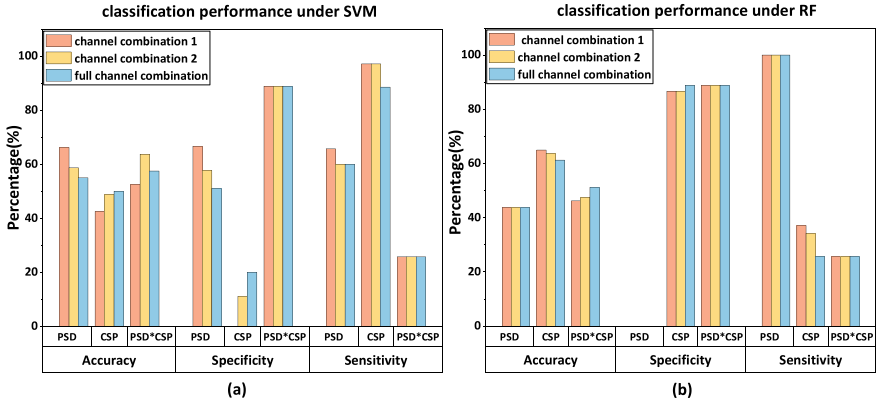


Fig. 4 Results of three channel combinations under different features and classifiers

5 Conclusion

In this study, we proposed the PSD-ID method for EEG channel selection in binocular disparity images which could choose the best channel combination. Results showed that the proposed channel combination could significantly improve the classification performance of EEG signals across images with different disparities. The number of channels was reduced by 1/2, and the classification accuracy was increased by 20%. Overall, the study highlighted the critical role of EEG channel selection in improving the accuracy of binocular disparity image classification and demonstrated the potential for EEG-based approaches to advance our understanding of the neural mechanisms underlying stereo vision. There are still some limitations in this study. First, the feature selection is relatively limited, and future research could explore more advanced EEG features to obtain better classification results. Second, the classification performance can be further improved by incorporating adaptive parameter adjustment algorithms or model fusion algorithms to enhance the accuracy and robustness of classification algorithms.

References

1. Minchev Z (2013) 2D vs 3D visualization and social networks entertainment games: a human factor response case study. In: International conference on evolutionary computation
2. Banks MS, Read JC, Allison RS, Watt SJ (2011) Stereoscopy and the human visual system. SMPTE Motion Imaging J 121(4):24–43
3. Peraza LR, Asghar AU, Green GG, Halliday DM (2012) Volume conduction effects in brain network inference from electroencephalographic recordings using phase lag index. J Neurosci Methods 207:189–199
4. An Y, Lam H, Ling S (2023) Multi-classification for EEG motor imagery signals using data evaluation-based auto-selected regularized FBCSP and convolutional neural network. Neural Comput Appl 35:12001–12027
5. Varsehi H, Firoozabadi SM (2020) An EEG channel selection method for motor imagery-based brain-computer interface and neurofeedback using Granger causality. Neural Netw: Official J Int Neural Netw Soc 133:193–206
6. Ma Y, Wu X, Zheng L, Lian P, Xiao Y, Yi Z (2022) Iterative outlier removal clustering based time-frequency-spatial feature selection for binary EEG motor imagery decoding. IEEE Trans Instrum Meas 71:1–14
7. Qi F, Wu W, Yu ZL, Gu Z, Wen Z, Yu T, Li Y (2020) Spatiotemporal-filtering-based channel selection for single-trial EEG classification. IEEE Trans Cybern 51:558–567
8. Xu M, Qi H, Ma L, Sun C, Zhang L, Wan B, Yin T, Ming D (2013) Channel selection based on phase measurement in P300-based brain-computer interface. PLoS ONE 8
9. Zhao Z, Anand R, Wang M (2019) Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform. In: 2019 IEEE international conference on data science and advanced analytics (DSAA), pp 442–452
10. Fauzi H, Azzam MA, Shapiai MI, Kyoso M, Khairuddin U, Komura T (2019) Energy extraction method for EEG channel selection. TELKOMNIKA (Telecommunication Computing Electronics and Control)

Semi-supervised Modulation Recognition Greatly Improved by Strong Data Augmentation

Weidong Wang, Cheng Luo, and Lu Gan

Abstract Modulation recognition plays an important role in modern wireless communications. Recent work has shown that modulation recognition based on deep learning significantly outperforms conventional approaches. However, this superiority relies largely on using plenty of labeled data for supervised learning, whereas training deep neural networks with limited data generally falls into overfitting, resulting in poor performance. In addition, it is also challenging to obtain plenty of labeled data in real-world communication activities, with expensive and time-consuming costs. To this end, we present a semi-supervised method for modulation recognition, which can take advantage of unlabeled samples that are more easily accessible in practice to enhance generalization and thus reduce such demand for labeled data. By introducing strong data augmentation, we improve supervised training and simultaneously perturb unlabeled data for consistency-based regularization, resulting in a remarkable generalization improvement. Experimental results on RadioML 2018.01A dataset demonstrate that our proposed method for semi-supervised modulation recognition is far superior to other competing ones and achieves almost fully supervised performance with a very small number of labeled samples.

Keywords Modulation recognition · Strong data augmentation · Deep semi-supervised learning · Consistency-based regularization

1 Introduction

In recent years, deep learning has been widely adopted for modulation recognition [1–3]. The powerful non-linear representation of deep neural networks makes it possible to learn high-level features directly from raw signal data, such as in-phase and quadrature (I/Q) waveform data. avoiding complicated and not always effective feature engineering. More importantly, it shows better performance than conven-

W. Wang · C. Luo · L. Gan (✉)
School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
e-mail: ganlu@uestc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_40

tional methods. However, it cannot be ignored that training deep neural networks requires a lot of data, especially labeled data. It is often challenging to obtain a large amount of labeled data in real-world communication activities, especially for blind recognition in some non-cooperative communication scenarios, where those received signals may lack relevant prior knowledge to be annotated. The signal acquisition and annotation process itself is also expensive and time-consuming, even with sufficient prior knowledge. Training a deep neural network on small datasets generally results in overfitting and degrades its generalization ability. In particular, deep neural networks are likely to be misled if some transmissions are mislabeled, resulting in worse performance. Hence, we expect to train deep neural networks for modulation recognition using relatively little supervised information, and deep semi-supervised learning is a good solution to address this issue.

Although plenty of studies have adopted deep learning for modulation recognition, most belong to supervised learning and only consider an ideal amount of labeled data. The efforts with semi-supervised learning are still very limited. Earlier, O'Shea et al. [4] employed a convolutional autoencoder to learn embeddings from massive signal data without labels. The encoder is then concatenated to a linear classifier, fine-tuned with a small number of labeled samples. However, such embeddings learned by minimizing a reconstruction error are not necessarily applicable to classification. The two-stage training pipeline is also complicated that easily accumulates more errors. The authors in [5, 6] employ generative adversarial nets (GAN) to synthesize more data for better generalization. But training such GANs itself also requires a certain amount of labeled data, which cannot provide satisfactory accuracy when only given limited labels. Recent work by Dong et al. [7] considers consistency-based regularization, a more modern semi-supervised paradigm that simultaneously involves supervised and unsupervised losses, capable of achieving excellent semi-supervised performance.

This paper presents a semi-supervised method for modulation recognition. In contrast to [7], which only considers consistency-based regularization, we introduce strong data augmentation both in supervised and unsupervised training. The strong data augmentation takes signal rotation and flipping as its primary transformation, followed by k-segmented stochastic permutation, and we employ it to increase labeled data while constructing perturbation over unlabeled samples to perform consistency-based regularization, leading to better generalization. Experimental results on RadioML 2018.01A [8] demonstrate that our proposed method for semi-supervised modulation recognition is far superior to other competing ones and achieves almost fully supervised performance with a very small number of labeled samples.

2 Methodology

Formally, we are provided with a radio dataset $\mathcal{D} = \mathcal{S} \cup \mathcal{U}$ collected from C modulation types, where signal samples in $\mathcal{S}$ are labeled, i.e., $\mathcal{S} = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{\mathrm{M}\times\mathrm{C}}$, and

those in $\mathcal{U}$ are not, i.e., $\mathcal{U} = \{\boldsymbol{x}_j\}_{j=1}^{\mathrm{N}\times\mathrm{C}}$, typically $\mathrm{M} \ll \mathrm{N}$. The core problem is how to utilize $\mathcal{U}$ to help a given deep model f_θ learning on $\mathcal{S}$ for modulation recognition with better generalization.

2.1 Strong Data Augmentation

Data augmentation has been proven to be an effective approach to improve generalization for deep neural networks. For example, one routinely uses rotation, translation, cropping, flipping, and random erasing to enforce visually plausible invariances in image classification [9]. Likewise, modulation recognition can also benefit from data augmentations customized for communication signals, including rotation, flipping, noise adding, and GAN-based sample generation [10, 11]. In our past work [12], we theoretically analyzed all these existing data augmentations with their effectiveness in RF fingerprinting. Among existing data augmentations for communication signals, we found that only rotation plays an effective role in RF fingerprinting, which can be explained by whether a constellation topology of signals is consistent with its original one after augmentation and whether additional noise is introduced.

Moreover, we proposed another powerful data augmentation technique, which divides a signal into k segments, then shuffles and reassembles them into a new signal, called k-segmented stochastic permutation, which mainly benefits from that each sampling point of radio signals follows an identical distribution, i.e., an underlying sampling distribution. Then, we combined signal rotation and k-segmented stochastic permutation to construct a composite data augmentation operation for RF fingerprinting. Similarly, we can employ this composite data augmentation operation to improve modulation recognition performance. Nevertheless, unlike RF fingerprinting, our analysis has shown that both rotation and flipping work in modulation recognition. As a result, this composite data augmentation operation needs to be further extended with flipping added. Specifically, given a signal sample, we randomly select any one of 6 available transformations (4 rotations and 2 flippings, see more details in [12]) to operate and then perform k-segmented stochastic permutation. For clarity, we denote this procedure as $\mathrm{g}(\cdot)$.

2.2 Training in Semi-supervised Fashion

Modern semi-supervised learning generally follows a form like

$$\mathcal{L} = \mathcal{L}_\mathrm{s} + \lambda \mathcal{L}_\mathrm{u} \tag{1}$$

where $\mathcal{L}_\mathrm{s}$ represents a supervised objective, $\mathcal{L}_\mathrm{u}$ is an unsupervised component worked as regularization, and λ is a penalty factor to combine these two terms. The supervised objective for modulation recognition is calculated by a standard cross-entropy loss:

$$\mathcal{L}_{\mathrm{s}} = \frac{1}{|\mathcal{S}|} \sum_{(\boldsymbol{x},\, y)\in\mathcal{S}} H(f_\theta(\mathrm{g}(\boldsymbol{x})),\ y) \tag{2}$$

Note that we employ strong data augmentation $\mathrm{g}(\cdot)$ to improve its learning effect since supervised training still plays a decisive role in semi-supervised learning. It is not difficult to find that modern semi-supervised learning largely depends on its unsupervised component—how to take full advantage of plenty of unlabeled samples easily available to improve generalization.

Recent work has shown that various consistency-based regularization approaches achieve SOTA performance [13]. Concretely, when a realistic perturbation is imposed on an input, its corresponding model prediction should not change significantly because these data points with distinct labels are separated by low-density regions, making it unlikely that one sample switches its class after perturbed [14]. The model $f_\theta(\cdot)$ is trained to have similar predictions for a sample $\boldsymbol{x}$ and its strongly augmented version $\mathrm{g}(\boldsymbol{x})$, and such similarity can be measured using Jensen-Shannon (JS) divergence, i.e.,

$$\mathcal{L}_{\mathrm{u}} = \frac{1}{|\mathcal{U}|} \sum_{\boldsymbol{x}\in\mathcal{U}} \mathrm{JS}(f_\theta(\mathrm{g}(\boldsymbol{x})) \,\|\, f_\theta(\boldsymbol{x})) \tag{3}$$

Note that reference [7] also uses a similar form for semi-supervised modulation recognition. In sharp contrast to [7], which perturbs unlabeled signal samples only by white Gaussian noise, we use strong data augmentation to build perturbation, which results in a better effect since our strong data augmentation does not change sample distribution and yields more diversity. In short, we exploit strong data augmentation to increase labeled samples and perturb unlabeled samples, resulting in a maximum generalization improvement for semi-supervised modulation recognition.

3 Experiments and Results

3.1 Data Preparation

Our experiments adopt RadioML 2018.01A [8], a public radio dataset includes both synthetic simulated channel effects and over-the-air (OTA) recordings of 24 digital and analog modulation types that has been heavily validated, available on DeepSig. The modulation types include OOK, 4ASK, 8ASK, BPSK, QPSK, 8PSK, 16PSK, 32PSK, 16APSK, 32APSK, 64APSK, 128APSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM, AM-SSB-WC, AM-SSB-SC, AM-DSB-WC, AM-DSB-SC, FM, GMSK, OQPSK, each of which involves different signal-to-noise ratios (SNR), varying from -20 to 30 dB with an interval of 2 dB. At each SNR, there are a total of 4096 complex-valued signal samples of length 1024 provided for each modulation type, among which we randomly select 2096 samples as a training set. The labeled and unlabeled data required in semi-supervised experiments are

randomly selected from this training set according to their specific amounts in each trial. The remaining 2000 signal samples are randomly split in half as a validation and test set.

3.2 Implementation and Training Details

Many previous studies for modulation recognition adopt convolutional neural networks (CNN). The network used in this work is a well-designed deep residual network (ResNet) [15], which has been widely used in our other works and can achieve excellent supervised performance for communication signal recognition [12]. To adapt for modulation recognition on RadioML 2018.01A, we slightly adjust it into a form with 3 convolution blocks stacked. The model is built with TensorFlow and then trained on a single NVIDIA RTX 2080S GPU utilizing an Adam optimizer for 180 epochs. The batch size could affect training stability and performance when training with a very limited number of examples. So we have conducted many experiments and given a relatively optimal configuration. The initial learning rate is set to 0.001. The recognition accuracy is used as a performance metric and calculated based on 20 trials.

3.3 Evaluation of Supervised Performance

Figure 1 shows supervised performance when training with different numbers of labeled samples. Three SNR conditions are considered, i.e., 0 dB (low), 10 dB (medium), and 20 dB (high). The supervised performance after using strong data augmentation is also indicated by square markers accordingly. As expected, we can see that supervised performance deteriorates when labeled data is insufficient. Nevertheless, we can also see that strong data augmentation can significantly improve recognition accuracy.

3.4 Evaluation of Semi-supervised Performance

The signal data with 10 dB provided in RadioML 2018.01A belongs to OTA data. So we employ this SNR condition to perform semi-supervised experiments. To begin with, we need to determine two hyper-parameters, i.e., λ and k. With k fixed at 2, and $M = 10$ and $N = 1000$, we investigate how semi-supervised performance varies with λ. The evaluation is mainly based on loss variation and average recognition accuracy, as presented in Fig. 2a–h. It can be seen that a smaller λ makes unsupervised training account for a small proportion so that consistency-based regularization is almost impossible to work, resulting in poor performance. However, when λ becomes larger, unsupervised training occupies more, resulting in another form of overfitting

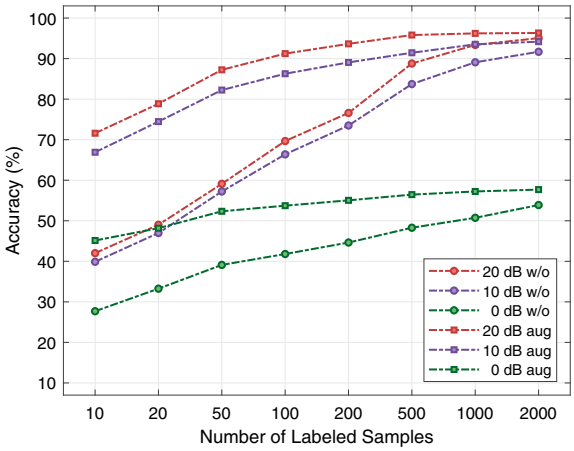
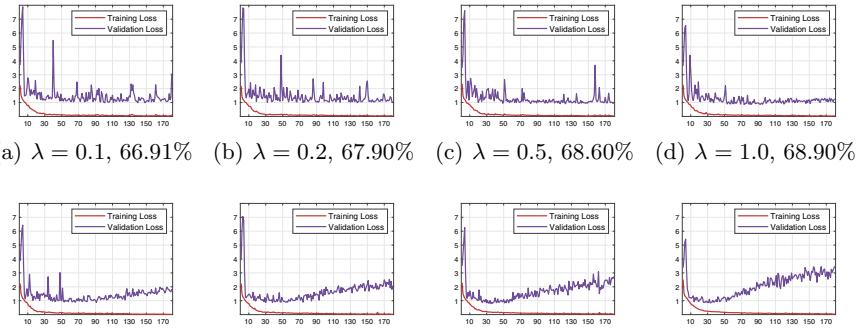


Fig. 1 Supervised performance vs different numbers of labeled samples



(a) $\lambda = 0.1$, 66.91% (b) $\lambda = 0.2$, 67.90% (c) $\lambda = 0.5$, 68.60% (d) $\lambda = 1.0$, 68.90%

(e) $\lambda = 2.0$, 68.77% (f) $\lambda = 5.0$, 67.67% (g) $\lambda = 10.0$, 67.41% (h) $\lambda = 20.0$, 66.85%

Fig. 2 Loss variations and average recognition accuracy under different λ

since only JS divergence cannot ensure the learning classification-related features. The best recognition accuracy is achieved at $\lambda = 1.0$ with a relatively smooth loss variation. The same ablation study was also performed with "20 + 1000" and "50 + 1000", and we have obtained a similar conclusion.

Then, with $\lambda = 1.0$, we investigate how semi-supervised performance varies with k, as shown in Fig. 3. Generally, as k increases, such enhanced data diversity could further promote our network to learn invariance for more robust features. We recommend that k should be less than an input signal's symbol length because only an average segmented length greater than one symbol is reasonable. It can be seen that semi-supervised performance keeps relatively stable when k varies from 48 to 80. The best recognition accuracy is achieved at $k = 64$.

The proposed method for semi-supervised modulation recognition is compared with other competing ones, including SSRCNN [7] and E3SGAN [6]. Table 1 gives

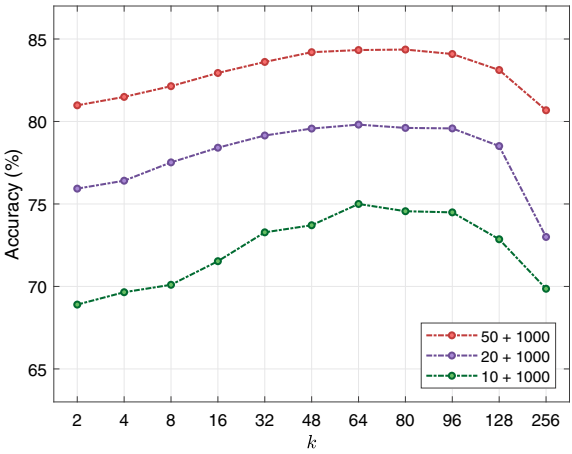


Fig. 3 Semi-supervised performance varies with k

Table 1 Comparsion with other semi-supervised modulation recognition algorithms

Method	10					20	50	180
	100	200	500	1000	2000	1000	1000	1800
Proposal	73.89 ± 0.86	74.38 ± 1.07	74.72 ± 0.84	75.00 ± 0.49	75.03 ± 0.49	79.81 ± 0.21	84.27 ± 0.31	90.03 ± 0.28
SSRCNN	42.38 ± 0.65	43.23 ± 0.79	43.06 ± 0.79	43.88 ± 0.68	44.00 ± 0.61	50.83 ± 0.53	56.82 ± 0.52	75.81 ± 0.66
E3SGAN	40.96 ± 0.64	40.87 ± 0.96	42.01 ± 1.06	41.65 ± 1.13	40.46 ± 1.73	47.65 ± 1.10	51.22 ± 1.62	63.20 ± 1.32

The table headers like "**180** and **1800**" means $M = 180$ and $N = 1800$

their performance under different data conditions. It can be seen that our proposed method is far superior to SSRCNN and E3SGAN. In particular, both ours and SSRCNN consider consistency-based regularization. However, we introduce strong data augmentation in supervised training and simultaneously use it to perturb unlabeled samples. It is not hard to conclude that data augmentation is critical to our performance improvement. The overall recognition accuracy obtained using all 2096 samples under supervised learning is 91.68%. The proposed method has a close performance of more than 90% at $M = 180$ and $N = 1800$. The demand for labeled data has been greatly reduced.

4 Conclusion

This paper has explored modulation recognition using deep semi-supervised learning. Compared to previous methods, we emphasize that data augmentation is necessary for improving generalization and then introduce strong data augmentation for semi-supervised modulation recognition. The experimental results demonstrate that our

proposed method for semi-supervised modulation recognition is far superior to other competing ones and achieves almost fully supervised performance with a very small number of signal samples. In future work, we will investigate using only a single labeled sample to achieve modulation recognition with satisfactory accuracy.

Acknowledgment This work is partially supported by Sichuan Science and Technology Program under Grant 2023YFSY0008 and 2023YFG0291, and partially supported by Yibin Science and Technology Program under Grant YBP-002.

References

1. O'Shea TJ, Corgan J, Clancy TC (2016) Convolutional radio modulation recognition networks. International conference on engineering applications of neural networks. Springer, New York, pp 213–226
2. Jdid B, Hassan K, Dayoub I, Lim WH, Mokayef M (2021) Machine learning based automatic modulation recognition for wireless communications: a comprehensive survey. IEEE Access 9:57851–57873
3. Peng S, Sun S, Yao YD (2021) A survey of modulation classification using deep learning: signal representation and data preprocessing. IEEE Trans Neural Netw Learn Syst
4. O'Shea TJ, West N, Vondal M, Clancy TC (2017) Semi-supervised radio signal identification. In: Proceedings of the 2017 19th international conference on advanced communication technology (ICACT). IEEE, pp. 33–38
5. Li M, Li O, Liu G, Zhang C (2018) Generative adversarial networks-based semi-supervised automatic modulation recognition for cognitive radio networks. Sensors 18(11):3913
6. Zhou H, Jiao L, Zheng S, Yang L, Shen W, Yang X (2020) Generative adversarial network-based electromagnetic signal classification: a semi-supervised learning framework. China Commun 17(10):157–169
7. Dong Y, Jiang X, Cheng L, Shi Q (2021) SSRCNN: a semi-supervised learning framework for signal recognition. IEEE Trans Cognit Commun Netw
8. O'Shea TJ, Roy T, Clancy TC (2018) Over-the-air deep learning based radio signal classification. IEEE J Select Top Signal Process 12(1):168–179
9. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(1):1–48
10. Huang L, Pan W, Zhang Y, Qian L, Gao N, Wu Y (2019) Data augmentation for deep learning-based radio modulation classification. IEEE Access 8:1498–1506
11. Tang B, Tu Y, Zhang Z, Lin Y (2018) Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks. IEEE Access 6:15713–15722
12. Wang W, Luo C, An J, Gan L, Liao H, Yuen C (2023) Semi-supervised RF fingerprinting with consistency-based regularization. arXiv preprint arXiv:2304.14795
13. Oliver A, Odena A, Raffel C, Cubuk ED, Goodfellow IJ (2018) Realistic evaluation of deep semi-supervised learning algorithms. arXiv preprint arXiv:1804.09170
14. Ouali Y, Hudelot C, Tami M (2020) An overview of deep semi-supervised learning. arXiv preprint arXiv:2006.05278
15. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778

Design and Research of the Control and Management System of Photovoltaic Cell

Youjie Zhou, Jinmao Chen, Chunhua Xiong, Xudong Wang, Liang Wen, Lianling Ren, Yongcheng Huang, Yaohui Wang, and Guang Hu

Abstract The idea of combining theoretical analysis and critical technologies is used in this article to design and study the MPPT algorithm, DC-DC control module, and output interface. It mainly relies on maximum power point tracking technology, while adopting a perturbance control strategy to maintain the maximum power output of photovoltaic cells. It provides a reference for achieving integrated system power supply management of photovoltaic cells.

Keywords Photovoltaic cells · Control and management system · MPPT

1 Introduction

The current shortage of fossil fuels is becoming increasingly prominent, and efficient utilization of renewable energy is a critical way to solve energy depletion. Compared with traditional forms of power generation, photovoltaic power generation technology with no fuel consumption requirement, no exhaust emissions, and no complex maintenance has received widespread attention as one of the mainstream technologies of renewable energy power generation.

In order to utilize solar energy effectively and convert it into usable electricity, it is necessary to allocate energy reasonably through a control and management system. The power management and control circuit is the brain of the entire power system, which manages the operation of the whole system. Therefore, it is significant for improving the performance of the photovoltaic cells to design a safe and efficient control and management system.

A control manager of a photovoltaic cell is the main research object in this paper, and the MPPT algorithm, DC/DC control module, and output interface are designed and studied relying on the maximum power point tracking (MPPT) technology, which

Y. Zhou · J. Chen · C. Xiong · X. Wang · L. Wen · L. Ren · Y. Huang · Y. Wang (✉) · G. Hu
System Engineering Institute, AMS, PLA, Beijing 100300, China
e-mail: energywyh@163.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_41

is in order to provide a technical reference for photovoltaic cells to realize integrated system power supply management in different environments and maintain the continuous output of electric energy to the load end at maximum power.

2 Design of Photovoltaic Cell Control Management

2.1 Overall Design

Due to the effect of external factors (such as sunlight intensity, angle, weather, etc.) on the output of photovoltaic power sources [1–7], it cannot guarantee stable power output for a long time. Therefore, photovoltaic cells often need to be used with control and management systems. The quality of the performance of the entire photovoltaic power supply system is usually determined by the control and management system of the photovoltaic cells.

In this paper, a block diagram of the photovoltaic cell power generation system (as shown in Fig. 1) is designed to realize the integrated power supply management function of the photovoltaic cells.

The voltage comparison circuits (charge and discharge) are introduced in this power supply system which also contains a bandgap reference voltage and a switching circuit that controls the charging and discharging of ceramic capacitors. In addition, limiting circuits of voltage and current are designed at the power output interface in order to guarantee the stability of three places: (1) charging of the energy storage battery; (2) discharging of the energy storage battery; (3) power supply to the load. Based on the overall design, the operation algorithm, voltage regulation control, and output interface function of the photovoltaic cell controller are studied as follows.

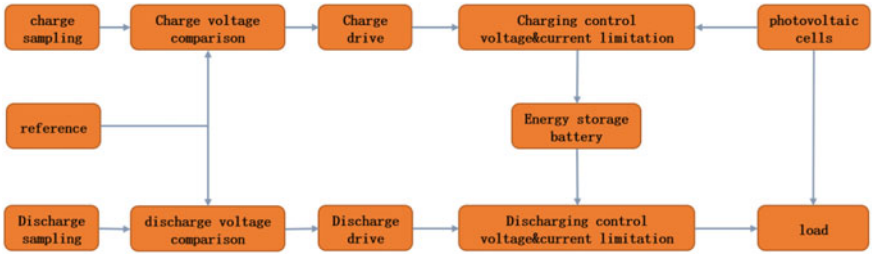


Fig. 1 Block diagram of photovoltaic cell power generation system

2.2 Submodule Design

2.2.1 MPPT Algorithm Module

The volt-ampere characteristics of the photovoltaic cells output are non-linear [8–13] (as shown in Fig. 2). At a specific temperature and light intensity, the voltage power curve (P–V) of the photovoltaic cells is an approximate parabolic shape with a maximum power point (MPP). And the MPP is generally obtained by introducing the MPPT algorithm module, which mainly consists of a current detection amplifier (voltage and current conversion circuit, current sampling amplifier), Analog multiplier, comparator, MPPT algorithm circuit, and logic control power supply.

MPPT algorithm circuit: Based on the compromise consideration of factors such as tracking accuracy, design cost, and load dependency, the perturb and observe (P&O) is adopted as the MPPT tracking algorithm [14, 15]. P&O is able to track and touch larger areas in a shorter time, scan more data, and meet the design requirements of photovoltaic cell control management in this article.

The mechanism of the P&O is continuously changing the working point of the photovoltaic cell and gradually adjusting the operating voltage of the solar cell to work near the maximum power point.

After the circuit starts, the output (voltage and current) of the photovoltaic cell is collected, the current power value is calculated, and the reference voltage is then perturbed. If the output power of the photovoltaic cell board increases after perturbance, the reference voltage will be perturbed in the same direction next time; If the output power of the photovoltaic panel decreases after perturbance, the reference voltage will be disturbed in the opposite direction next time. The algorithm logic judgment and relevant situation are shown in Fig. 3.

Current sense amplifier: The current sense amplifier consists of a current sampling amplifier and a voltage and current converter. Since the output current signal of the photovoltaic cell is difficult to enter the chip directly, it is necessary to enter the chip through the resistor sampled by current. Considering the power dissipation, the input chip voltage signal will be small due to the small resistor selection. Thus, an amplifier is required for subsequent processing.

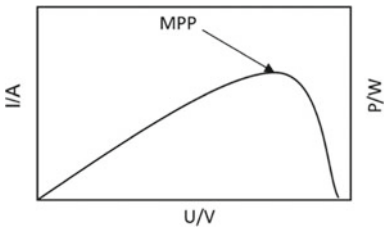


Fig. 2 Output characteristics of photovoltaic arrays under different light intensities

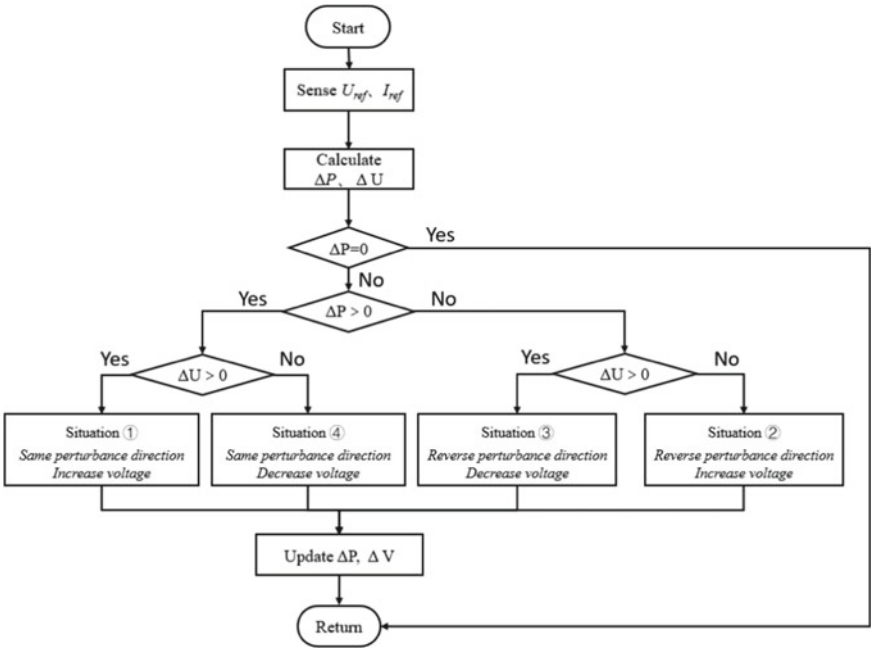


Fig. 3 Flow chart of disturbance observation method

Analog multiplier: The current multiplier is selected to calculate the output power of the photovoltaic cell after the output voltage of the photovoltaic cell is converted into a current signal inside the chip (Fig. 4).

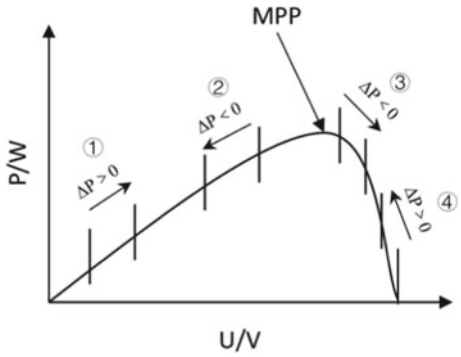


Fig. 4 Diagram of the relationship between the perturbance direction and step size

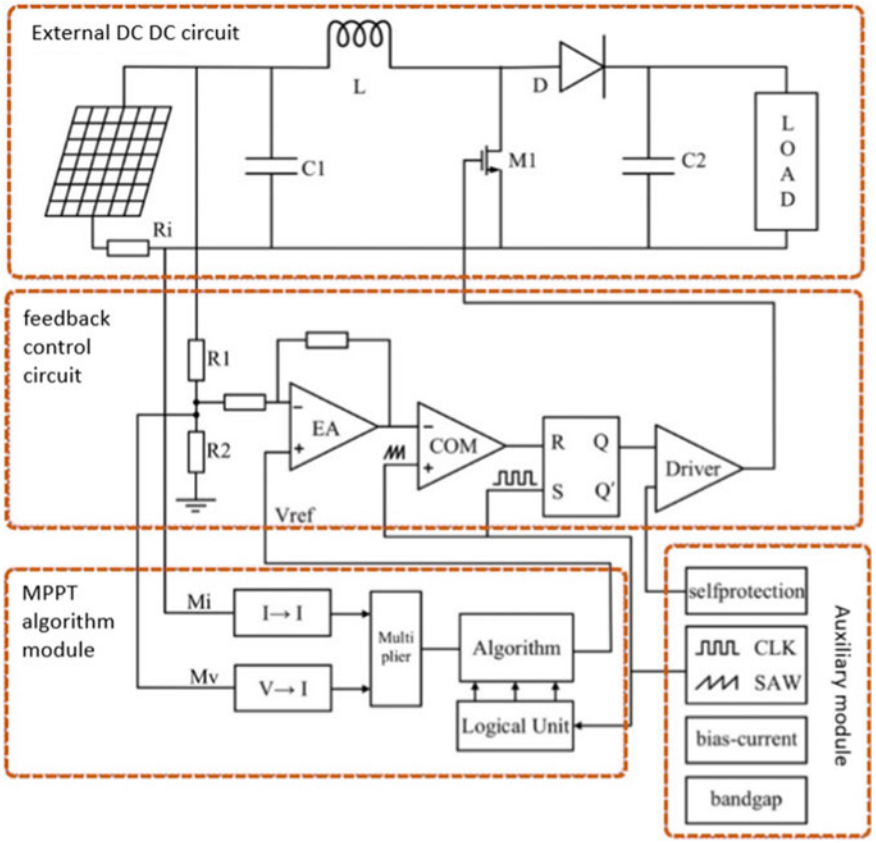


Fig. 5 The structure of the MPPT controller system

2.2.2 DC-DC Control Module

A Nonsynchronous rectified BOOST circuit is selected as the basic topology in the DC-DC module (Fig. 5). Although the implementation method of this circuit is simple, there is pulsating interference in the input and output currents. The research group will optimize the circuit design in the next step.

2.2.3 Interface Design

The photovoltaic cells are able to generate relatively weak DC voltage under light. Since charging energy storage batteries requires continuous stable voltage and current density, and driving load also requires the constant output of a specific voltage and current, the research of charging technology and specific interface circuit is one of the crucial issues and key technologies that must be overcome.

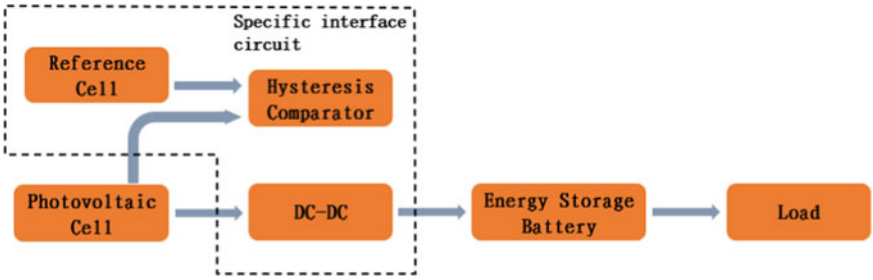


Fig. 6 Design of special interface circuit for photovoltaic cells

It is an efficient way to improve energy management and use efficiency by reasonably matching the impedance between the photovoltaic cell, energy storage battery, and loads. The open-circuit voltage proportional coefficient method is tried to adopt in this article which introduces the reference battery as well (as shown in Fig. 6). The proportional relationship between the open-circuit voltage of the reference battery (V_{ref}) and the photovoltaic cell (V_{OC}) is approximately unchanged ($V_{OC} = k_{ref} \cdot V_{ref}$).

By utilizing the linear proportional relationship between the MPP voltage (V_{MPP}) and the open-circuit voltage (V_{OC}) of the photovoltaic cell, the V_{MPP} can be obtained after the V_{ref} is divided by resistors. To connect the positive and negative voltage (V_{PV}) of the photovoltaic cell and the V_{MPP} obtained by resistance voltage division to the positive and negative input terminals of the hysteresis comparator LTC1440, and control the on–off of the DC-DC interface circuit. The circuit in the dashed block diagram in the figure achieves real-time tracking and control of the maximum output power of the photovoltaic cell.

The DC-DC conversion circuit is the interface circuit connecting power generation, energy storage components, and loads. It transforms the output DC voltage into a constant DC voltage of different values required by each module while matching impedance.

3 Conclusion

This article designs the photovoltaic power generation system and focuses on the MPPT algorithm module, DC-DC control module, and output interface. The purpose of the integrated power supply management of the photovoltaic cells is achieved by applying the P&O strategy combined with the asynchronous rectification BOOST topology and customized interface design.

References

1. Jiangqi G (2022) Research on maximum power point tracking control algorithm of photovoltaic power system under different illumination. Chang'an University, Shanxi
2. Quan L, Xuefei W (2019) Research on output characteristics of photovoltaic arrays with different structures under partial shading. Acta Energy Sol Sin 40(04):1003–1010
3. Huijing G (2021) Influence of incident angle of light on power of photovoltaic panels. Mech Eng Autom 05:186–188
4. Jinjun X, Wenlin F et al (2015) Output characteristic and application for solar panel. Infrared Laser Eng 1:176–181
5. Jinjun P, Yanbo S et al (2014) The influence of meteorological factors on the temperature of solar panels. J Appl Meteorol Sci 25(2):150–157
6. Yihua H, Hao C et al (2011) PV module characteristics effected by shadow problem. Trans China Electrotech Soc 26(01):123–128, 134
7. Yongqin C, Qi F et al (2018) Review on the impact of dust accumulation on photovoltaic power output. J Desert Res 38(2):270–277
8. Riza M, Minwon P et al (2003) A maximum power point tracking for photovoltaic-SPE system using a maximum current controller. Sol Energy Mater Sol Cells 75(3/4):697–706
9. Zaiteng Z, Jinlei D et al (2006) Influence of the electrical parameters on the serial solar panels mismatch losses. Proc CSEE 26(25):42–45
10. Mohalmad AS, Masourn HD, Ewald FF (2002) Theoretical and experimental analysis of photovoltaic system with voltage and current based maximum power point tracking. IEEE Trans Energy Convers 17(4):514–522
11. Chihchiang H, Jongrong L (2004) A modified tracking algorithm for maximum power tracking of solar array. Energy Convers Manage 45(6):911–925
12. Eugene VS, Shengyi L, Roger AD (2004) Power controller design for maximum power tracking in solar installations. IEEE Trans Power Electron 19(5):1295–1304
13. Kawamura T, Harada K, Ishihara Y et al (1997) Analysis of MPPT characteristics in photovoltaic power system. Sol Energy Mater Sol Cells 47(1/4):155–165
14. Yan C, Binghuang C et al (2006) Comparative studies on the MPPT control algorithms of solar energy photovoltaic system. Acta Energy Sol Sin 2006(06):535–539
15. Yan W (2007) Study on MPPT control method on PV generation system. North China Electric Power University, Beijing

A Cluster-Based Spectrum Allocation Method for Interference Mitigation of Multiple WBANs

Yuanyuan Li and Jiasong Mu

Abstract Wireless Body Area Network (WBAN) provides comfortable and reliable medical and non-medical applications for users by continuously sensing human data and information about the surrounding environment. WBANs may change the way people live everywhere in the foreseeable future which poses challenges in communication. In this paper, we proposed a clustering-based spectrum allocation method for large-scale body-domain networks aiming to mitigate co-channel interference when multiple users are involved, in which the algorithm uses a scheduling method that includes clustering and coloring algorithms to achieve optimal resource allocation when there is no infrastructure. The simulation results demonstrate that the proposed algorithm effectively promotes interference resistance of the network, minimizes the network delay and significantly improves the spectrum resource utilization.

Keywords Interference mitigation · WBAN · Spectrum allocation · Clustering · Coloring

1 Introduction

According to the latest medical research, many deadly malignant diseases have a high cure rate if they are detected and treated early. Similarly, the earlier some chronic diseases are detected and controlled by medical intervention, the less likely they are to deteriorate further [1]. On top of that, due to the aging of the population there is an urgent need for a technology that can solve these problems and shift the way in how people consider and manage their health and lives [2]. Wireless Body Area

Y. Li (✉)
School of Computer Science and Artificial Intelligence, Lanzhou Institute of Technology, Lanzhou, China
e-mail: lyytuffo@163.com

J. Mu
College of Electronics and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
e-mail: mujiasong@aliyun.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_42

Networks, an emerging technology, is regarded as a revolutionary technology that will help humanity solve the problems we face today.

Typically, a Wireless Body Area Networks consists of a collection of intelligent physiological sensor nodes which are capable of monitor the human body functions and characteristics from the surrounding environment, also with the ability to establishing a wireless communication link that operate in the proximity of a human body [3]. Figure 1 shows a typical WBAN. The communication architecture of WBANs can be defined as the following three layers, as shown in Fig. 2. Tier 1 described as the intra-WBAN communication, Tier 2 represents the inter-WBANs communication, Tier 3 denotes the extra-WBANs communication [1, 4, 5].

IEEE 802.15.6 is a communication standard optimized to provide an international standard for a short-range (i.e., about human body range), low power, and highly reliable wireless communication for use in close proximity to, or inside, a human body to satisfy an evolutionary set of entertainment, consumer electronics, healthcare services and other [6]. Figure 3 depicts the available frequency bands in IEEE 802.15.6.

According to the standard, a WBAN allows only one coordinator to exist, each WBAN needs to be able to support at most 256 nodes, and the physical layer needs to support up to 10 randomly distributed coexisting WBANs in a space of 6 m^3 [6]. Inevitably, there is some level of interference when several WBANs transmit data which use the same frequency. In this paper, we propose use clustering algorithm to achieve self-organizing clustering of WBANs in a multi-user environment as a way to solve the no-infrastructure problem of WBANs, and a coloring scheme is used to achieve spectrum allocation.

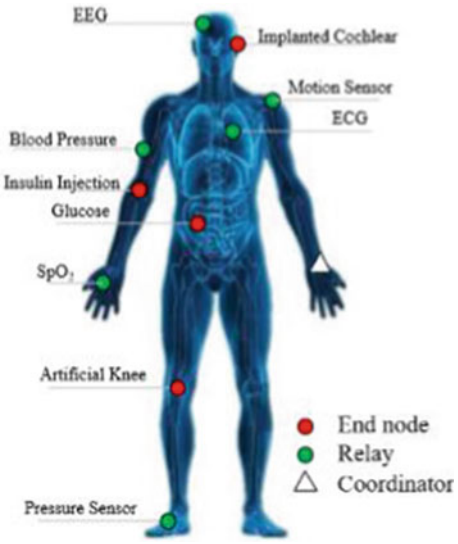


Fig. 1 An example of WBAN

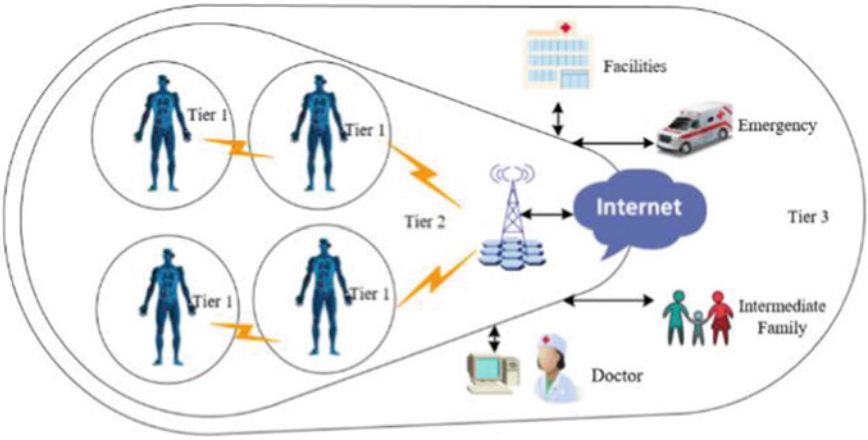


Fig. 2 Communication tiers for WBANs

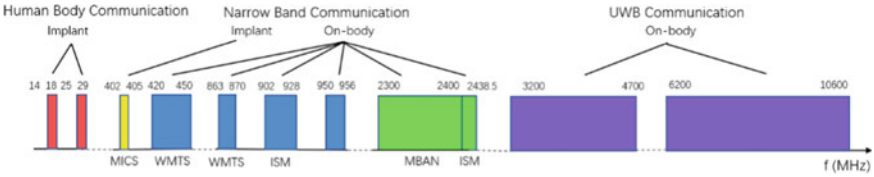


Fig. 3 Frequency bands for IEEE 802.15.6

The rest of the paper is organized as follows: Sect. 2 gives a brief review of existing related work. Section 3 presents the clustering algorithm and the vertex coloring scheme. The analysis of simulation results is shown in Sect. 4. Finally, Sect. 5 concludes the paper and presents ideas for further research.

2 Related Works

In [7], a communication channel resource allocation (PRA) algorithm based on a priority-based active superframe interleaving scheme is proposed. While it is concerned only with the allocation of communication resources in a single WBAN. Movassaghi et al. have done some research in the adaptive interference mitigation scheme for WBANs [8, 9]. Their proposed scheme that not treat each WBAN as a whole but analyzes it from the node-level, the Interference Region (IR) is determined by setting the receive power threshold.

Cheng and Huang [10] conduct a random incomplete coloring (RIC) research for WBANs to overcome interference, each WBAN in the algorithm is assigned a random value, based on which the exchange of coloring messages generates

significant network overhead. The clique-based WBAN scheduling was studied as shown in [11] which differs from RIC in that WBANs are not considered as a whole, where nodes with the same function in the algorithm work in the same time slot. In [12], the authors studied a Cooperative resource allocation method which are on extended vertex coloring (CEVC). Yet CEVC doesn't consider large-scale, high-density WBAN scenarios, which is not a global solution.

To mitigate co-channel interference between WBANs, and improve QoS, [13] borrows the concepts of cell and cluster in cellular networks. This is a global interference suppression solution for multiple-user high-density scenarios, while [13] accomplishes self-organizing clustering by relying on the received signal strength which is a locally optimal result the effect needs to be improved and the scalability is a serious restriction for the algorithm.

3 Cluster-Based Spectrum Allocation Method

In this paper, a spectrum allocation method based on clustering algorithm and a graph coloring scheme is proposed. The clustering algorithm is used to perform cell partitioning of multiple WBANs, thus solving the problem of no fixed infrastructure in WBANs, and then the interference mitigation is described as a graph coloring problem to complete the spectrum allocation.

3.1 Cluster-Based DBSCAN

DBSCAN (Density-based spatial clustering of applications with noise) [14] is a density-based clustering non-parametric algorithm, in this paper Euclidean distance is adopted. Conventionally, clustering of oddly-shaped can be accomplished with DBSCAN by setting only two parameters, MinPts and ε.

The improved DBSCAN for multi-WBANs environments, which adds a threshold setting for an upper limit on the number of WBANs within a cluster during clustering, when the maximum number of members in a cluster is reached, it stops joining and automatically splits into two clusters. When clustering, the WBANs location information is taken as a known input and MinPts, ε and threshold are set at the beginning. Multiple WBANs clustering processing can be abstracted into the following steps: Initially, marking all the WBANs in the service area as unvisited and starting clustering at an unvisited point of arbitrary selection. The ε neighborhood of P is checked afterwards. If the number of objects within the neighborhood is less than MinPts, P is temporarily tagged as a noise point. When the number of users contained within the cluster is greater than MinPts and below the threshold there cluster C is established, then iteratively processes all the connection components within the cluster which have not been labeled as visited, classifying them all as cluster C to complete the

expansion of the cluster. New cluster C + 1 is created when the number of users is excessive.

3.2 Coloring of Clustered Topologies

The multi-WBAN network scenario after clustering can be modeled as an undirected graph G = (V, E), where the set of vertices V(G) denotes clusters and the set of edges E(G) indicates two connected neighboring vertices (aka., cells) that would generate interference if they used the same frequency band, respectively. It is proposed that we use the backtracking algorithm for coloring, converting the undirected connected graph into an adjacency matrix, and generating a depth-first tree through depth-first search, coloring the nodes in succession. According to the four-color theorem, the number of colors k cannot be less than 4, the flow of the back-coloring method is as shown in Fig. 4.

In this context, the FDMA is instinctively supported. Figure 5 illustrates the channel assignment based on the clustering algorithm. The frequency resource available to each cell is Δf, and the resource accessible to each WBAN user is Δf/m. In this paper, two methods, FDMA and TDMA, are used to improve the spectrum allocation.

4 Simulation

In this section, the performance of CBSA for spectrum allocation is evaluated. The simulations were carried out in two network environments with different user distribution densities, 120 and 240 users. The experimental area for the simulations was 36 m * 36 m and users were randomly distributed within the area. The simulations assume no interference and other noise from adjacent channels, the communication distance is limited to a few meters due to the strict requirements on transmit power as a radio device close to or inside the body [15], the co-channel interference distance is set to 3 m, the data retransmission time is 2 ms and the maximum number of members that can be accommodated in each cluster is 20.

Compared the resources allocated after clustering case with the non-clustered, the SIR of the different methods in the two scenarios were shown in Fig. 6. The results shown that optimizing the spectrum allocation contributes to improving the interference resistance of the network which is due to the reduction in the number of users using the same frequency through the algorithm and the distance of users working with the same sub-channel is far enough, but as the data rate increases the SIR both decrease this is due to the increase in the amount of information.

Figure 7 illustrated the network delay for different conditions, network latency increases as the data rate increases. On the one hand, the increase in the amount of data in the network itself leads to an additional waiting time, and on the other hand,

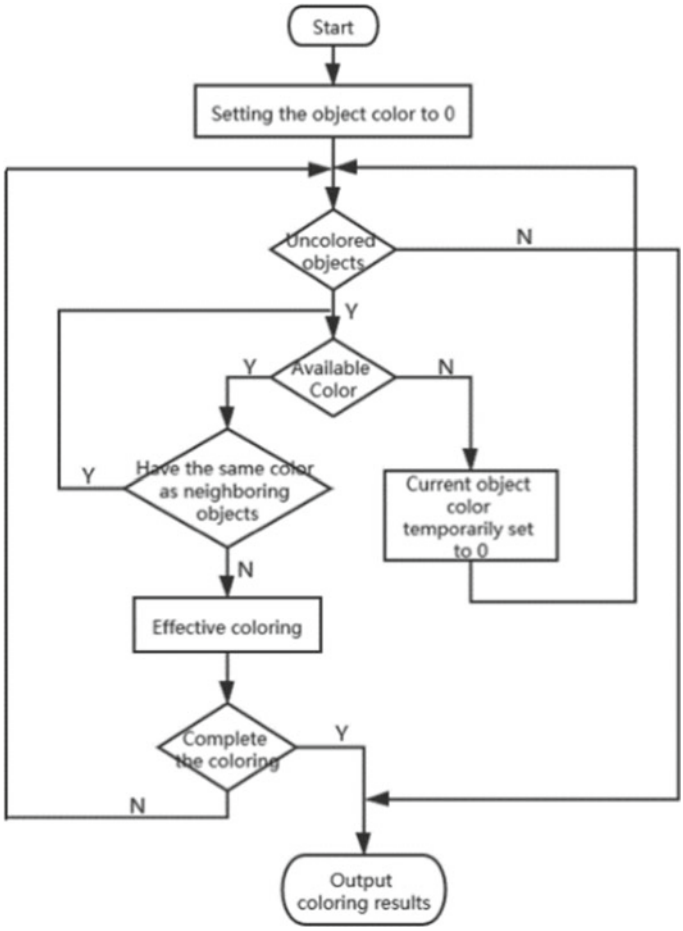


Fig. 4 Flow chart of back tracking coloring method

the increasement in data rate leads to a reduction in the SIR, which means the error rate goes up.

The spectrum reuse ratio can be expressed in terms of the average number of vertices for each color coloring Vpc (Vpc = number of coloring vertices/number of available colors). It can be seen from Fig. 8. This is because the clustering algorithm partitions the WBAN users into clusters, the users within these clusters would be able to simultaneously transmit data using the same section of spectrum resources. While the un-clustered case treats each WBAN as a coloring object, which greatly increases the time it takes to complete the coloring and the coloring cannot be completed when the number of coloring colors is three.

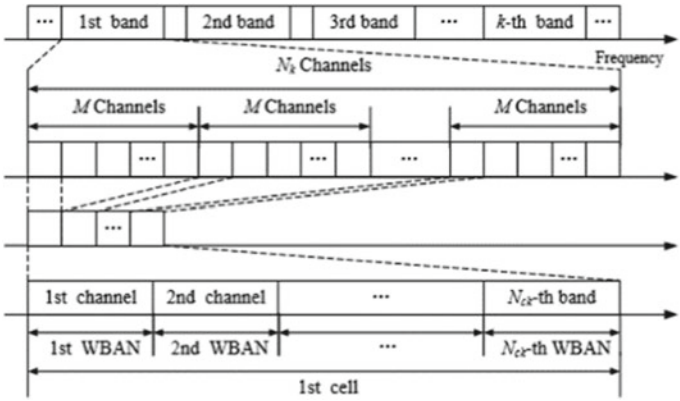


Fig. 5 Channel assignment based on the clustering algorithm

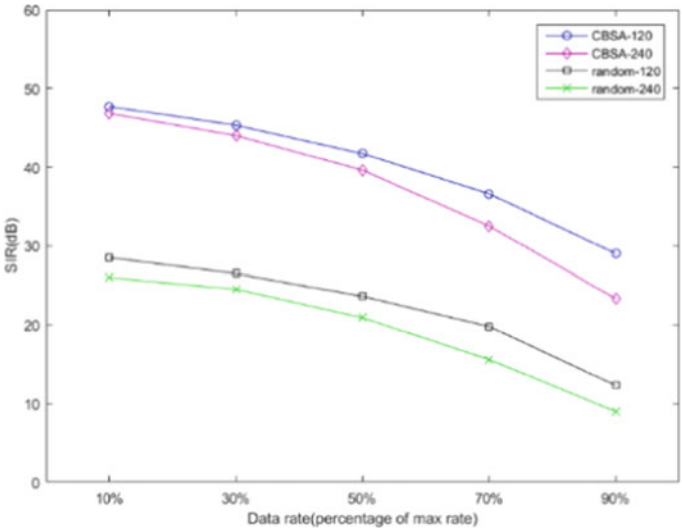


Fig. 6 SIR in different scenarios

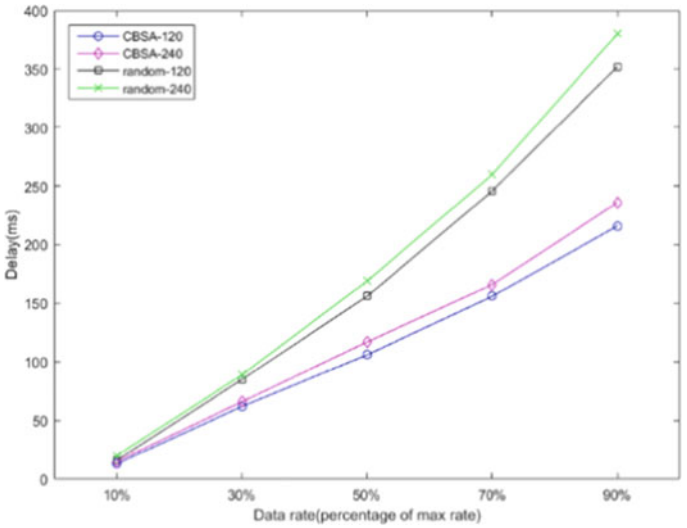


Fig. 7 Delay in different scenarios

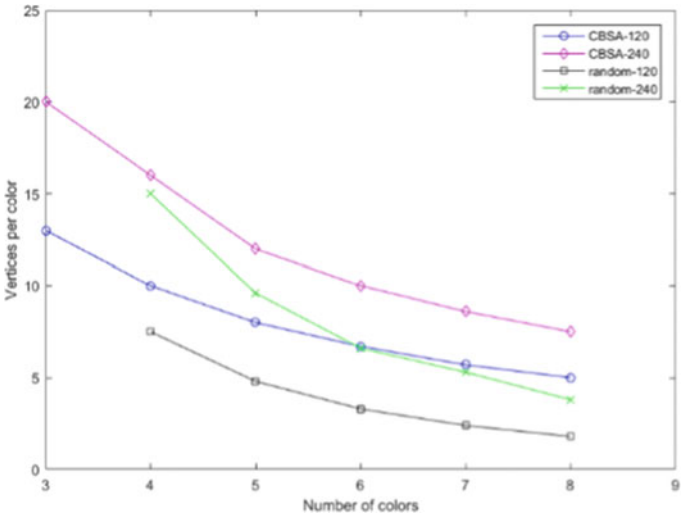


Fig. 8 Vpc in different scenarios

5 Conclusion

Wireless body area networks as an emerging technology different from mobile communication networks, an efficient management architecture needs to be redesigned for high-density deployments of WBANs. In this paper, we proposed

a clustering-based spectrum optimization method to mitigate co-channel interference and improve network performance in multi-WBANs environment. The experimental results show that the clustering-based spectrum optimization and allocation method effectively improves the signal-to-interference ratio, ensures the reliability of the network, and to some extent improves the reuse of spectrum resources. Moving forward, we will be concentrating on achieving further improvements in spectrum reuse, ensuring network stability during topology changes, and cloud-based solutions.

References

1. Movassaghi S, Abolhasan M, Lipman J et al (2014) Wireless body area networks: a survey. IEEE Commun Surv Tutor 16(3):1658–1686
2. Milenkovic A, Otto C, Jovanov E (2006) Wireless sensor networks for personal health monitoring: issues and an implementation. Comput Commun 29:2521–2533
3. Taleb H, Nasser A, Andrieux G et al (2021) Wireless technologies, medical applications and future challenges in WBAN: a survey. Wireless Netw 27:5271–5295. https://doi.org/10.1007/s11276-021-02780-2
4. Awan K, Qureshi KN, Mehwish M (2016) Wireless body area networks routing protocols: a review. Indones J Electr Eng Comput Sci 4(3):594–604
5. Ullah S, Higgins H, Braem B, Latre B, Blondia C, Moerman I, Saleem S, Rahman Z, Kwak K (2010) A comprehensive survey of wireless body area networks. J Med Syst 1–30
6. IEEE standard for local and metropolitan area networks—part 15.6: wireless body area networks, 2012
7. Kim S, Song BK (2017) A prioritized resource allocation algorithm for multiple wireless body area networks
8. Movassaghi S, Abolhasan M, Smith D (2014) Smart spectrum allocation for interference mitigation in wireless body area networks
9. Movassaghi S, Abolhasan M, Smith D, Jamalipour A (2014) AIM: adaptive internetwork interference mitigation amongst co-existing wireless body area networks. In: 2014 IEEE global communications conference, Austin, TX, pp 2460–2465
10. Cheng SH, Huang CY (2013) Coloring-based inter-WBAN scheduling for mobile wireless body area networks. IEEE Trans Parallel Distrib Syst 24(2):250–259
11. Xie Z, Huang G, He J et al (2014) A clique-based WBAN scheduling for mobile wireless body area networks. Procedia Comput Sci 31:1092–1101
12. Miri M, Mohamedpour K, Darmani Y et al (2019) An efficient resource allocation algorithm based on vertex coloring to mitigate interference among coexisting WBANs. Comput Netw 151:132–146
13. Mu J, Stewart R, Han L et al (2018) A self-organized dynamic clustering method and its multiple access mechanism for multiple WBANs. IEEE Internet Things J 1
14. Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the second international conference on knowledge discovery and data mining (KDD'96). AAAI Press, pp 226–231
15. C95.1-2005—IEEE standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz, 1999

Thermo-Acoustic Imaging for Tissue Mimicking (TM) Materials and Breast Tissues in Microwave Induced Thermo-Acoustic Tomography System

Jiazhou Liu, Yalin Wang, Yuxin Song, Wenlin Cheng, and Zhiqin Zhao

Abstract Microwave induced thermo-acoustic tomography (MITAT) has been investigated for decades due to its great potential in early breast cancer detection. Based on a precise MITAT system, the imaging experiments for tissue mimicking (TM) materials and ex-vivo breast cancer tissue are performed. First, the effectiveness of the TM materials is evaluated. The thermo-acoustic (TA) signal and response of the TM material are compared with those of the real breast tumor specimen. Second, the 3D thermo-acoustic image for TM material is reconstructed to validate the tomographic imaging ability of the MITAT system. Finally, the thermo-acoustic imaging experiment for hybrid breast cancer tissue is performed to verify the distinguishing sensitivity of the system. The potential of MITAT for breast cancer detection is further validated by the results presented in this paper.

Keywords MITAT · Breast cancer · Thermo-acoustic image

1 Introduction

According to the statistics from the American Cancer Society, breast cancer is the most diagnosed cancer type among women and the early detection is crucial to the cure of the cancer [1]. Microwave induced thermo-acoustic tomography (MITAT) is

J. Liu · Y. Wang
China Academy of Electronic and Information Technology, Beijing 100041, China

J. Liu
Tsinghua University, Beijing 100084, China

Y. Song (✉)
West China Forth University Hospital, Chengdu 610041, China
e-mail: 635195698@qq.com

W. Cheng
China Shipbuilding Trading Co. Ltd., Beijing 100094, China

Z. Zhao
University of Electronic Science and Technology of China, Chengdu 611731, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_43

a potential detection modality for the early breast cancer detection due to it combines the advantages of microwave imaging and ultrasound tomography [2, 3]. In MITAT system, the microwave pulses are radiated to the biological tissues and some of microwave energy is absorbed by the tissues. The absorbed microwave energy causes thermo-elastic expansion inside the tissue and then generates acoustic signals termed thermo-acoustic (TA) wave [4, 5]. The TA waves can be received by the ultrasound transducers around the tissue. By using the received TA signals, an image which reflects the TA response of the tissue can be reconstructed according to some specific inversion method.

The methods about TA imaging and numerical models have been reported in many literatures [6–8]. However, the experimental setups in these literatures were often simply constructed and were not considered much for the clinical application. The imaging system of MITAT, which is close to clinical application, is a combination of mechanics, microwave engineering, acoustic engineering and digital signal processing. In this paper, a MITAT system prototype which has potential in clinical application is proposed. The purpose of the system is to build a platform that can let the patient prone on it and make tomography of her breast one by one.

In order to validate the imaging ability of the MITAT system, tissue mimicking (TM) material to mimic tumor tissue is made and adopted to perform the imaging experiments. The dielectric properties are first measured to illustrate the validity of the TM material. The thermo-acoustic signals generated by TM and breast cancer specimen are also compared to further verify the validity of TM material. Afterwards, TM material sample with different shapes are utilized to achieve the imaging experiments due to the convenience of molding. The imaging ability of the MITAT system is approved by the TM images.

Toward the initial clinical application, the ex-vivo breast cancer specimen is utilized to perform the imaging experiment. In order to keep the character of the tissue as same as the condition in vivo, the specimen is adopted to achieve the experiment within two hours after the surgery. The imaging result efficiently shows the ability of MITAT system for breast cancer detection.

Moreover, a powerful imaging inversion algorithm is a major technologic factor in MITAT. Time reversal mirror (TRM) is a combination technique based on wave propagation and array signal processing [9]. It has the ability to effectively suppress system noise due to the characteristics of its spatial–temporal matched filtering. Furthermore, the statistic stability due to the self-average of TRM gives it the capability to decrease the effects induced by random fluctuating of the media. Therefore, image inversion method based on TRM technique is adopted to reconstruct the thermo-acoustic image.

2 MITAT System and Materials

2.1 Overview of MITAT System

Figure 1 is a picture of the MITAT system with scanning mode. In the system, a modulator is designed to drive and modulate a 2.45 GHz magnetron to produce rectangular microwave pulses with 500 ns duration. Microwave pulses are radiated to the scanning container from the bottom and the samples immersed in the coupling medium generate thermo-acoustic signals due to the thermo-elastic expansion caused by microwave energy accumulation. The thermo-acoustic signals can be received by the ultrasound transducers which are uniformly fixed on the container wall. In order to better capture thermo-acoustic signal, a kind of ultrasound transducer with fairly large aperture (Olympus, V314_SU) is adopted. Due to the large aperture and limit superficial area of the container, the imaging sensor array is realized by rotating four transducers to form a synthesized array with a scanning platform.

In a scanning procedure, the system controller controls the motor to drive the scanning container to rotate a scanning step. Afterwards, the trigger pulses are sent from the system controller to trigger the microwave modulator and the data acquisition card. The irradiated microwave from the radiator illuminates on the tissue in the container and thermo-acoustic signals will be emitted. The thermo-acoustic signals are amplified by a 54 dB preamplifier (Olympus, 5662). Then the thermo-acoustic signals are sampled with 7.5 MHz sampling rate after 20 times averaging. After finish all the scanning procedures, digital thermo-acoustic signals are processed to reconstruct an image. All the scanning procedures with 1.8° step can be finished within two minutes. More details about the system can be found in [10, 11].

The imaging method adopted in MITAT system is based on the time reversal mirror (TRM) technique. The procedure of TRM can be expressed as

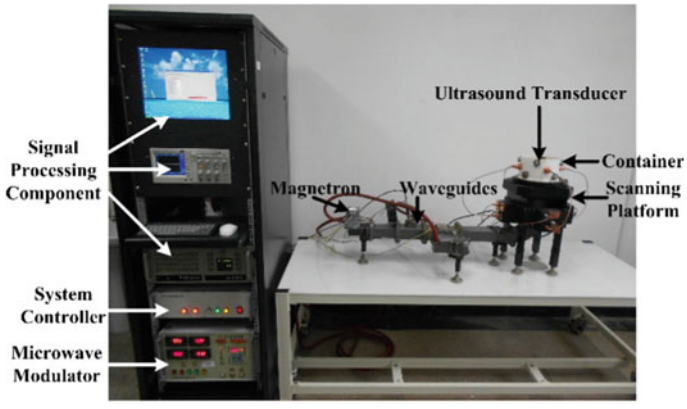


Fig. 1 Picture of MITAT system

$$I(\vec{r},t)=\sum_{s=1}^{M}\int_{-\infty}^{+\infty}\overline{p}_s(\omega)\sum_{k=1}^{K}G_c(\vec{r}_k,\vec{r},\omega)\overline{G}_s(\vec{r}_k,\vec{r}_s,\omega)e^{-j\omega t}d\omega \tag{1}$$

$$p_s(\omega)=\int_{-\infty}^{+\infty}p(t)e^{-j\omega t}dt \tag{2}$$

where $I(\vec{r},t)$ is the pixel value at location $\vec{r}$ in imaging zone at moment t, M is the totality of thermo-acoustic signals, K is the totality of sensors, G_c is a Green's function from the sensor to field location $\vec{r}$ and $\overline{G}_s$ is a conjugate Green's function from acoustic source to the sensor. By using TRM method, a tomographic image can be reconstructed within one minute.

2.2 *Tissue Mimicking (TM) Material*

The properties of ex-vivo breast tissues over a frequency range can be modeled by a one pole Cole–Cole relaxation equation

$$\varepsilon_{\text{rel}}(\omega)-j\frac{\sigma_{eff}(\omega)}{\omega\varepsilon_0}=\varepsilon_\infty+\frac{\Delta\varepsilon}{1+(j\omega\tau)^{1-\alpha}}-j\frac{\sigma_s}{\omega\varepsilon_0} \tag{3}$$

where ε_∞ is the high frequency permittivity, τ is the relaxation time, $\Delta\varepsilon$ is the pole amplitude and σ_s is the static ionic conductivity.

Tissue mimicking (TM) material in [12] is referred to make the breast tumor TM material. The tumor TM material is mainly focus on the dielectric property mimicking and expected to reflect the TA response of the real tumor tissue. The TM material presented in this paper is made of agar powder and distilled water etc. The dielectric property (relative permittivity and effective conductivity) is measured from 1 to 15 GHz and shown in Fig. 2.

As a reference, the Cole–Cole curves of the breast tumor [13] are also plotted in the figure. It is important to emphasize that the Cole–Cole curve matches well with the ex-vivo breast tumor. However, the in-vivo breast tumor has higher dielectric parameter than that of the ex-vivo one. Therefore, in considering of the increments of in-vivo tissue, the TM material is compounded with the higher dielectric parameter.

From Fig. 2, it is observed that the dielectric property of the TM material is dispersive along with the frequency. This consistency indicates that the TM material has the similar dielectric property with the in-vivo tumor tissue in an ultra-wideband.

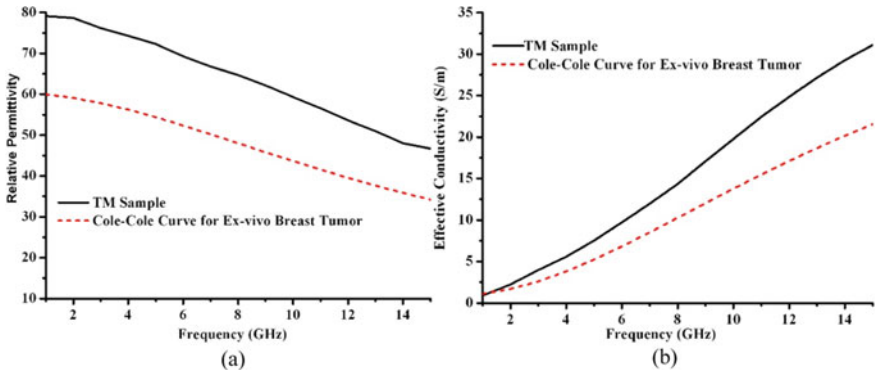


Fig. 2 Dielectric property of TM material. **a** Relative permittivity, **b** effective conductivity

2.3 Real Breast Cancer Tissue

Real breast tissue is utilized to verify the imaging potential of MITAT system in clinical application. For keeping the character of the tissue similar with the condition in vivo, the tissue is imaged within two hours after the surgery.

For evaluating the validity of TM material, real tumor tissue is compared with the TM material in TA signal generation and TA image. The TM material is processed to the same shape with the tumor and the picture is shown in Fig. 3a.

The breast tissue, which is diagnosed in advance to be a hybrid tissue including cancer tissue and normal tissue, is used to perform imaging experiment. With this hybrid tissue, the imaging result shows not only the contrast between tumor and normal tissue but also the ability of the MITAT system to distinguish the tumor from the surrounding normal tissue. The optical pictures of tissues are show in Fig. 3b. The cross section size of tissue is about 20 mm × 20 mm.

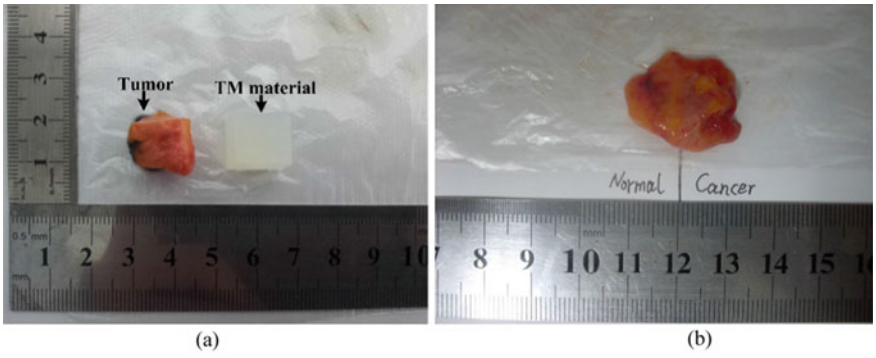


Fig. 3 Real breast tissue, **a** breast tumor and TM sample; **b** hybrid breast tissue

3 Results and Discussions

In this part, the comparison between breast tumor and TM material is shown and discussed.

3.1 TA Signal of TM Material and Breast Tumor

The breast tumor and TM sample shown in Fig. 3a are both placed in the scanning container of MITAT system to perform the experiment. Figure 4 shows the temporal TA waveforms as detected by the ultrasound transducer at three different receiving positions. The microwave pulses are radiated to the samples at t = 0 μs. Since the microwave pulses are delivered to the samples at the speed of light, the TA waves are induced at $t \approx 0$ μs. The approximately 37 μs time delay corresponds to the propagation time of the induced TA signal to reach the ultrasound transducer in the coupling medium.

There are two TA signals which are composed of mainly two lobes in each waveform shown in Fig. 4. The first TA signal is generated by breast tumor and the other one is excited from TM sample. In order to quantitatively analysis the similarity of the TA signals between tumor and TM sample, the peak-to-peak values as well as the temporal widths of TA signals are summarized in Table 1. The temporal width is the duration between the positive lobe and negative lobe.

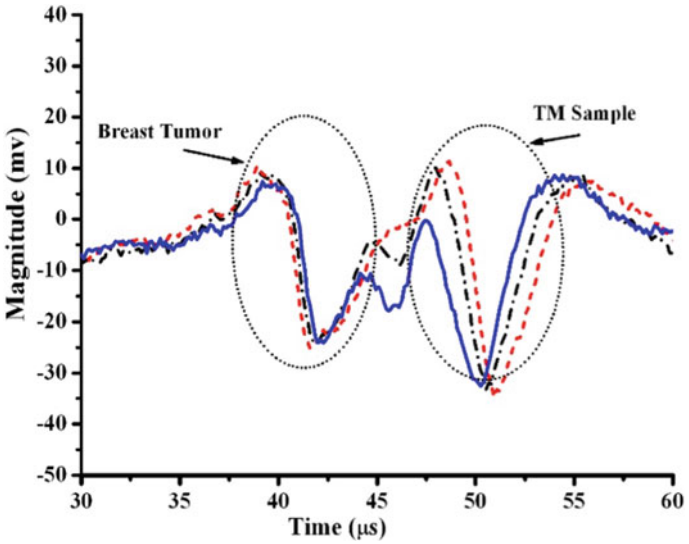


Fig. 4 TA signals of TM sample and breast tumor

Table 1 Summary of peak-to-peak values and temporal widths of TA signals

	Peak-to-peak values (mv)			Temporal widths (μs)		
Tumor	33.32	35.85	31.56	2.39	2.66	2.79
TM sample	43.64	46.18	32.74	2.66	2.26	2.79

Figure 4 and the data in Table 1 show that the TA peak-to-peak values of TM sample are higher than those of the tumors about 10 mv. This result is due to the higher dielectric parameters of the TM sample. The higher effective conductivity leads to the more microwave absorption and the magnitude of TA signal is proportional to the absorbed energy. The temporal fluctuations of TA signals have some differences. We attribute these variations in TA waveforms to small changes in the placements of the target relative to the ultrasound transducer.

3.2 TA Image for TM Material and Breast Tumor

After the comparison between the TA signals of the tumor and TM sample, the TA image for the ex-vivo breast tumor and TM sample is presented in this section. The tumor and TM samples shown in Fig. 3a are utilized to perform the imaging experiment by using the MITAT system described in Sect. 2.1. The TA image for the tumor and TM samples is shown in Fig. 5.

In the figure, it is observed that the shapes and locations of the TM and tumor samples are both reconstructed. The TA image is normalized by the maximum intensity value and the mean pixel values of the two samples are 0.44 and 0.38, respectively. This result indicates that the TM sample has higher TA response than the tumor tissue. Meanwhile, it is noted that the increment of the TM sample in pixel value is

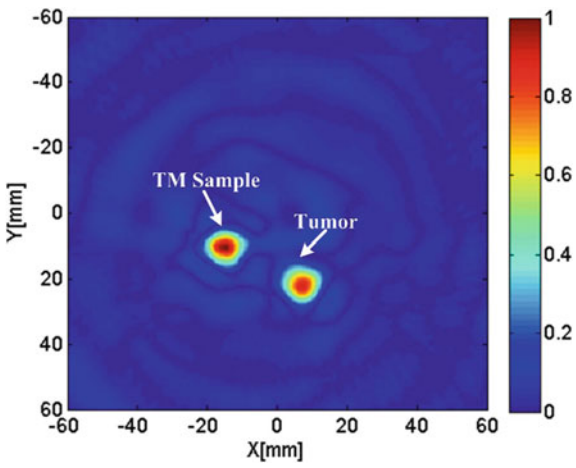


Fig. 5 TA image for TM sample and breast tumor

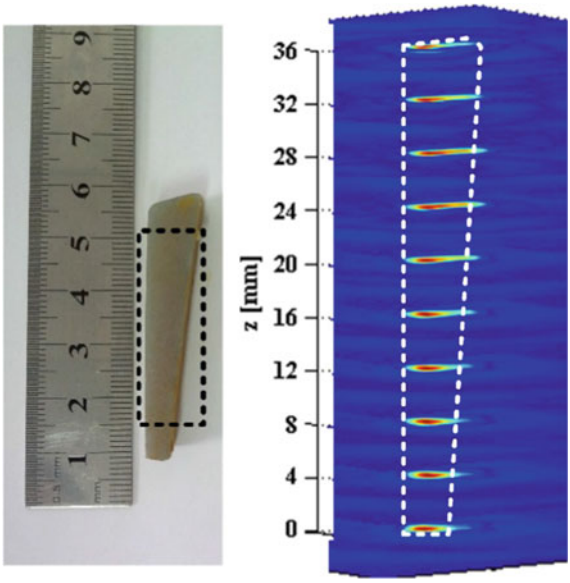


Fig. 6 Tomographic image for the trapezoid TM sample

not significant than the TA peak-to-peak value. This result is due to the refocusing loss in the image reconstruction.

According to the comparison between TA signals and image of TM and breast tumor samples, it can be concluded that the TM material can effectively mimic the breast tumor sample. Based on the substitution of TM material, the performances of established MITAT system can be initially evaluated without a real breast tumor sample. In order to evaluate the tomographic sensitivity of the MITAT system, a trapezoid TM sample is utilized to perform the imaging experiment. In the experiment, the TM sample was moved along the vertical direction with a 4 mm step to obtain the tomographic images of the target. The tomographic image for the TM sample is shown in Fig. 6.

The black dashed line marked in the picture (the left part of Fig. 6) is the imaging zone of the TM sample. The right part of the figure is the corresponding TA image to the TM sample. From the figure, it is observed that the TA tomographic image accurately reflect the variation trend of the TM sample edge and the shape of the TM sample is reconstructed. This result indicates that the MITAT system can obtain the tomographic image with micrometer level resolution in the vertical direction.

3.3 TA Image for Hybrid Breast Tissue

The contrast between separated cancer tissue and normal tissue has been investigated in [14, 15]. In order to verify the distinguishing ability from the complex tissue of the MITAT system, a hybrid breast tissue shown in Fig. 3b is adopted to perform

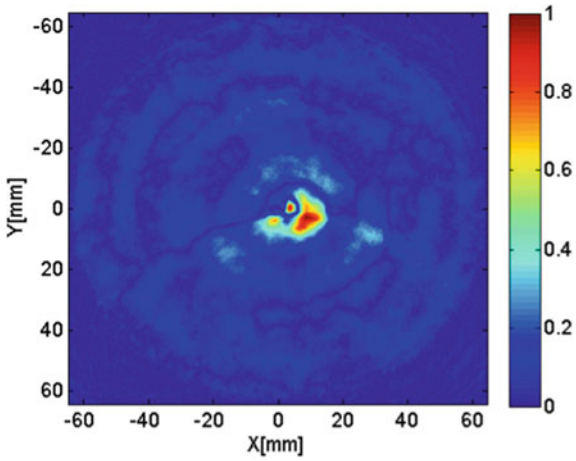


Fig. 7 TA image for hybrid breast tumor

the imaging experiment. In the experiment, the hybrid tissue is immersed with the coupling medium in the scanning container. After finishing the scanning procedure, the TA image of the breast tissue can be obtained. The TA image for the hybrid breast tissue is shown in Fig. 7.

In the figure, the cancer tissue part of the hybrid tissue is reconstructed while the normal tissue is not. This result indicates that the MITAT system with scanning mode not only has the significant contrast between the separated cancer and normal tissues but also can distinguish the cancer tissue from the hybrid tissue. The TA image for the hybrid tissue further suggests that MITAT has the potential to detect tumor from the normal tissues.

4 Conclusion

In this paper, thermo-acoustic (TA) imaging experiments for tissue mimicking (TM) material and real breast tissue are performed and discussed. The dielectric property of the TM material is measured and matches well with in-vivo breast tumor. Meanwhile, the TM material shows the similar TA response (TA signal and TA image intensity) with the real breast tissue. Therefore, it can be concluded that the TM material proposed in this paper can efficiently mimic the real breast tissue. The imaging and distinguishing ability of the MITAT system are verified by the TA image for the TM sample and real hybrid breast tissue. The imaging resolution along the vertical direction can achieve micrometer level. Based on these experiments, the potential of MITAT is further verified.

References

1. American Cancer Society (2023) Cancer facts and figures 2023. (Online). Available: http://www.cancer.org
2. Cui Y, Yuan C, Ji Z (2017) A review of microwave-induced thermoacoustic imaging: excitation source, data acquisition system and biomedical applications. J Innov Opt Health Sci 10(04):1730007
3. Lin JC (2021) Microwave thermoacoustic tomographic (MTT) imaging. Phys Med Biol 66(10):10TR02
4. Ye F et al (2015) Ultrashort microwave-pumped real-time thermoacoustic breast tumor imaging system. IEEE Trans Med Imaging 35(3):839–844
5. Bauer D, Wang X, Vollin J, Xin H, Witte R (2012) Spectroscopic thermoacoustic imaging of water and fat composition. Appl Phys Lett 101:033705
6. Wang X, Bauer DR, Witte R, Xin H (2012) Microwave-induced thermoacoustic imaging model for potential breast cancer detection. IEEE Trans Biomed Eng 59(10):2782–2791
7. Wang B et al (2019) Three-dimensional microwave-induced thermoacoustic imaging based on compressive sensing using an analytically constructed dictionary. IEEE Trans Microw Theory Techn 68(1):377–386
8. Witte RS, Tamimi EA (2022) Emerging photoacoustic and thermoacoustic imaging technologies for detecting primary and metastatic cancer and guiding therapy. Clin Exp Metastasis 39(1):213–217
9. Wang JG, Zhao ZQ, Song J, Zhu XZ, Nie ZP, Liu QH (2012) Reconstruction of microwave absorption properties in heterogeneous tissue for microwave induced thermo-acoustic tomography. Prog Electromagn Res PIER 130:225–240
10. Rahpeima R, Soltani M, Kashkooli FM (2020) Numerical study of microwave induced thermoacoustic imaging for initial detection of cancer of breast on anatomically realistic breast phantom. Comput Methods Programs Biomed 196:105606
11. Soltani M, Rahpeima R, Kashkooli FM (2019) Breast cancer diagnosis with a microwave thermoacoustic imaging technique—a numerical approach. Med Biol Eng Comput 57:1497–1513
12. Hacker L et al (2022) Criteria for the design of tissue-mimicking phantoms for the standardization of biophotonic instrumentation. Nat Biomed Eng 6(5):541–558
13. El Misilmani HM et al (2020) A survey on antenna designs for breast cancer detection using microwave imaging. IEEE Access 8:102570–102594
14. Wang B et al (2018) Efficient dictionary construction method for microwave induced thermoacoustic compressive sensing imaging. Appl Phys Lett 113(5):053701
15. Song J, Zhao ZQ, Wang JG, Zhu XZ, Wu JN, Nie ZP, Liu QH (2013) An integrated simulation approach and experimental research on microwave induced thermo-acoustic tomography system. Prog Electromagn Res PIER 140:385–400

IGWL: The Intersection Graph of Wireless Links Based Multi-target Localization in UHF RFID Scenarios

Bobo Wang, Yongtao Ma, Xianchao Zhang, and Kaiping Xue

Abstract Radio tomographic imaging (RTI) is one of the most favorable device-free localization methods with many advantages. However, since its model is approximate and ill-posed, it has low multi-target localization accuracy in indoor environments. To overcome the problem, we propose an IGWL scheme, the intersection graph of wireless links based multi-target localization, which directly utilizes target-affected links to localize targets. Firstly, the intersections of target-affected links are extracted from all intersections of all wireless links according to the first Fresnel zone. Then, analyze the distribution of extracted intersection graph and determine the intersections with high respect to targets. After, the gap statistic method is used to determine the number of targets. Finally, we modify the K-means algorithm to acquire the locations of targets. The simulation shows that the localization accuracy of the proposed IGWL scheme is significantly higher than the traditional RTI.

Keywords Device-free localization · Radio tomographic imaging · Passive UHF RFID · Multi-target localization · K-means algorithm

1 Introduction

Indoor localization serves as sensing and understanding of human behavior and will penetrate into ubiquitous things in different domains. Device-free localization (DFL) is one of the promising indoor localization technologies and can localize targets without any auxiliary device [1]. The mainstream techniques adopted to realize it

B. Wang · K. Xue
School of Cyber Science and Technology, University of Science and Technology of China, Hefei, Anhui, China

Y. Ma
School of Microelectronics, Tianjin University, Tianjin, China

B. Wang (✉) · X. Zhang
Key Laboratory of Medical Electronics and Digital Health of Zhejiang Province, Jiaxing, Zhejiang, China
e-mail: wangbobo@ustc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_44

include ultra-wideband (UWB) [2], WiFi [3], and RFID [4]. Compared with other techniques, RFID tags have many advantages of low maintenance, low cost, and easy deployment. Especially, passive UHF RFID tags are lighter, smaller, and require no battery [5]. Therefore, passive UHF RFID-based DFL deserves to study deeply.

The approaches of existing DFL techniques are often categorized into big data-based and model-based methods [6]. The big data-based method [7] must spend significant time and labor to build and update the fingerprint database. The model-based method builds a localization model to estimate targets' locations with low labor and fast speed. Radio tomographic imaging (RTI) is an auspicious model-based approach due to its low computational cost, good real-time, and high robustness. Depending on different characteristics of wireless signals, many approaches are proposed, including variance-based RTI [8], subspace variance-based RTI and least square variance-based RTI [9], and phase response shift-based RTI [10].

Both passive RFID systems and the RTI have obvious inherent superiority. Therefore, some scholars study the RTI in passive UHF RFID scenarios. The RFID-based RTI system uses the variations of target-affected signals to image targets' locations. Specifically, an adaptive bistatic weight model [11] and multi-layer perceptron estimation [12] utilize the characteristics of backscatter signals to localize people accurately. To eliminate artifacts and false targets in reconstruction images, a gray analysis and naive Bayes classifier method [4], and a novel scanning circle link analysis method [13] are proposed to determine the locations and the number of targets accurately.

Although some scholars have contributed substantially to improving the localization of RFID-based RTI, it is challenging to localize multiple targets indoors accurately [14]. The reason is that the RTI is an ill-posed and approximate model, which results in low localization performance. To enhance the performance, we propose a novel IGWL scheme. In the scheme, target-affected wireless links are determined according to their first Fresnel zone, and the intersections of affected links are extracted. Then, analyze the distribution of the extracted intersection graph and determine the intersections with high respect to targets. After, the gap statistic method is used to determine the number of targets. Finally, we modify the K-means algorithm to acquire the locations of targets. The main contributions of this paper are summarized as follows:

- We acquire target-affected wireless links according to the Fresnel zone and extract the intersections of affected links. Then, analyzing the extracted intersection graph, we determine the core intersections with high respect to targets.
- We combine the gap statistic method and K-means algorithm to count and cluster the core intersection graph for determining the number and locations of targets.
- We perform many simulations to compare the proposed IGWL scheme with the traditional RTI (TraRTI). The results show that the proposed IGWL scheme outperforms the TraRTI.

The rest of this paper is organized as follows. In Sect. 2, radio tomographic imaging is introduced. In Sect. 3, the IGWL scheme is presented. Many simulations are performed to validate the proposed IGWL scheme in Sect. 4. The conclusion is drawn in Sect. 5.

2 Radio Tomographic Imaging

2.1 Radio Tomographic Imaging Model for Passive UHF RFID

Radio tomographic imaging is a localization method for imaging physical targets' locations in a wireless network composed of wireless links. The locations in a localization area correspond to the pixels in a reconstruction image. When a target enters the localization area, wireless signals near the target are affected, and their strength changes significantly. The RTI localization model [4] can characterize the relationship between the target's location and the variations of target-affected signals:

$$\mathbf{y} = \mathbf{W} \cdot \mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is a vector recording received signal strength (RSS) variations of all wireless signals, $\mathbf{W}$ is a weight matrix, the symbol $\cdot$ represents the dot product, $\mathbf{x}$ is a vector recording the gray values in a reconstructed image, and $\mathbf{n}$ is a noise vector. Since the weight and knowledge of indoor environment are difficult to obtain, a weight model is required to approximate the knowledge of indoor environment. Currently, the elliptical model is widely used in RTI, that is, $w_{i,j}$ is the weight of pixel i for link j:

$$w_{i,j} = \frac{1}{\sqrt{d_j}} \begin{cases} 1, & d_{i,j}^{[1]} + d_{i,j}^{[2]} < d_j + \alpha \\ 0, & \text{other} \end{cases}, \tag{2}$$

where d_j is the length of link j, $d_{i,j}^{[1]}$ and $d_{i,j}^{[2]}$ are the distances between the pixel i and two elliptical foci, respectively, and α is the ellipse's width. The mathematical model [2] explains how the target affects a wireless signal:

$$\widehat{y} = a d_e^c \cos\left(2\pi \frac{d_e}{\lambda} + \phi\right), \tag{3}$$

where $\widehat{y}$ is RSS variation of an affected link, both a and c are tunable constants with respect to multipath, λ is the signal wavelength, ϕ is an initial phase, $d_e = d_{nl}^{[t]} + d_{nl}^{[r]} - d_l$, $d_{nl}^{[t]}$ and $d_{nl}^{[r]}$ are the distances from the target to the sender and the receiver under non-line-of-sight propagation, respectively, and d_l is the distance between the sender and the receiver under line-of-sight (LOS) propagation. The parameter $[\hat{y}_j]$ replaces $\mathbf{y}$ in Eq. (1) to determine the RTI model for passive UHF RFID.

2.2 Problem Statement for RTI Model

The RTI model characterizes the relationship between the target's location and the measurement and is solved to estimate the target's location through some algorithms. There are two problems in this process.

- The current RTI is an approximate model that cannot well characterize the relationship between the target's location and the measurement. Moreover, since the weight in it is difficult to obtain, only use approximate geometric models to represent this weight.
- The solution of RTI model is an ill-posed problem, that is, the number of measurements is significantly less than that of pixels in a RTI reconstruction image. Its solution accuracy is low. Moreover, small measurement errors may cause large localization errors.

In order to eliminate the impact of these two problems on localization performance, this paper proposes an IGWL scheme to directly use the measurement for localizing targets.

3 Multi-target Localization Based on the Intersection Graph of Wireless Links

The intersection graph of wireless links based multi-target localization scheme contains four steps. First, get target-affected wireless links and extract the intersections of these links. Then, core intersections are determined according to the intersection graph. After, the IGWL scheme uses the gap statistic method to analyze the distribution of core intersections to acquire the number of targets. Finally, the IGWL scheme utilizes the K-means algorithm to obtain estimated locations of targets. These four steps will be elaborated as follows.

3.1 The Intersections of Affected Link Extraction

Most of the power of wireless signals is concentrated in its first Fresnel zone. To simplify, we assume that only a target within the first Fresnel zone affects the corresponding wireless signal. The radius of the first Fresnel zone [15] is calculated as:

$$r = \sqrt{\frac{\lambda d_1 d_2}{d_1 + d_2}}, \tag{4}$$

where d_1 and d_2 are the distances between the point at the LOS path and two endpoints of LOS path, respectively. The parameters of the first Fresnel zone corresponding to two wireless links: $\{(x_1,y_1),a_1,b_1\}$ and $\{(x_2,y_2),a_2,b_2\}$, where (x_i,y_i) is the coordinate of one point i on the boundary of the first Fresnel zone, $a = \sqrt{(d_1 + d_2)^2 + b^2}$, and b is the maximum value of r. To simplify, the first Fresnel zone of link 2 has been turned into a standard form.

Definition 1 : If $\exists\ x_1$ and $\exists\ y_1$ s.t. $\frac{x_1^2}{a_2^2} + \frac{y_1^2}{b_2^2} < 1$, these two links have one intersection.

According to **Definition 1**, the intersections of all affected links are extracted.

3.2 Core Intersection Determination

Some intersections are formed by the links affected by environmental noise and multipath. In order to reduce their impact on localization performance, we need to eliminate them. Two related definitions are given as follows.

Definition 2 : $(x_{l,1}, y_{l,1}), \ldots, (x_{l,n}, y_{l,n})$ are the intersections in order of link l with all other links. If $(x_{l,i-1}, y_{l,i-1})$, $(x_{l,i}, y_{l,i})$ and $(x_{l,i+1}, y_{l,i+1})$ are all intersections of affected links, $(x_{l,i}, y_{l,i})$ is a midpoint.

Definition 3 : If $(x_{l,1}, y_{l,1})$ is the midpoint of all wireless links passing through it, it is a core intersection.

In this paper, the core intersection is highly related to the target. According to **Definition 2** and **Definition 3**, all core intersections are determined.

3.3 Targets' Quantity Statistic

We modify the gap statistic method to obtain the number of targets. The coordinates of core intersections are represented in matrix form $\{(x_i, y_i)\}, i = 1, 2, \ldots, n$, where n is the number of core intersections. Let $d_{i,j}$ denote the distance between interactions i and j: $d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$.

Step 1: Let $D_r = \sum_{i,j \in C_r} d_{i,j}$ be the sum of pairwise distances of all intersections in cluster r, where C_r is the set of all intersections in cluster r. Calculate $W_K = \sum_{r=1}^{K} \frac{1}{2n_r} D_r$,where K is the number of clusters and n_r is the number of intersections in cluster r.

Step 2: Within the range of the coordinate (x,y), use the Monte Carlo model to generate B sets of reference points. Then, calculate $W_{K,b}$, where $b = 1, 2, \ldots, B$. We define $Gap_K = (1/B) \sum_{b=1}^{B} \log(W_{K,b}) - \log(W_K)$.

Step 3: Let $\bar{Q} = (1/B) \sum_{b=1}^{B} \log(W_{K,b})$. We define

$$Sd_K = \left[(1/B) \sum_{b=1}^{B} \left(\log(W_{K,b}) - \bar{Q} \right)^2 \right]^{1/2}$$

and $S_K = Sd_K \sqrt{1 + 1/B}$. If $Gap_K \geq Gap_{K+1} - S_{K+1}$, the method is end and K is the number of targets. Otherwise, $K = K + 1$ and the method jumps **Step 2**.

3.4 Targets' Location Acquistion

We use the K-means algorithm to acquire the locations of targets, that is, the centers of each cluster. The advantages of the K-means algorithm are as follows: its time complexity is linearly related to the number of samples, and it is very efficient and has good scalability. Therefore, it can meet the requirements of fast, stable, and accurate localization of the IGWL scheme. In order to eliminate the interference of outliers on the classification results, the IGWL scheme uses graph theory (in Sect. 3.2) to remove outliers and intersections with low respect to targets. The gap statistic method determines the number of clusters K (in Sect. 3.3). The IGWL scheme randomly selects core interactions as initial core points. Since all samples are core intersections of wireless links, the classification result of the K-means algorithm will not be locally optimal. The Algorithm 1 details how the K-means algorithm obtains the centers of all clusters, that is, all centers are the targets' locations.

Algorithm 1 The K-means algorithm for obtaining the centers of all clusters

Require: $\{(x_i, y_i)\}$: the set of core intersections, $i = 1, 2, \ldots, n$;
K: the number of clusters.
Ensure: the centers of all clusters $\{(x_k, y_k)\}$, $k = 1, 2, \ldots, K$.
randomly select K cluster centers:$(x_1^{[0]}, y_1^{[0]}), (x_2^{[0]}, y_2^{[0]}), \ldots, (x_K^{[0]}, y_K^{[0]})$;
repeat
  initialize the cluster's set i: $C_i = \varnothing (1 \le k \le K)$;
  for i=1,2,...,m **do**
    calculate the Euclidean distance between sample (x_i, y_i) and k cluster center (x_k, y_k) $(1 \le k \le K)$: $d_{i,k} = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}$;
    determine the cluster label of (x_i, y_i) according to the nearest cluster center: $\lambda_k = \arg\min \left\{ d_{i,1}, d_{i,2}, \ldots, d_{i,K} \right\}$;
    the sample (x_i, y_i) is added to the cluster C_{λ_k}: $C_{\lambda_k} = C_{\lambda_k} \cup \{(x_i, y_i)\}$;
  end for
  for k=1,2,...,K **do**
    calculate new cluster center: $x_k^{'} = \frac{1}{|C_k|} \sum_{x \in C_k} x$ and $y_k^{'} = \frac{1}{|C_k|} \sum_{y \in C_k} y$;
    if $(x_k^{'}, y_k^{'}) \neq (x_k, y_k)$ **then**
      update the cluster center (x_k, y_k) to $(x_k^{'}, y_k^{'})$;
    end if
  end for
until the centers of all clusters remain unchanged.

4 Performance Evaluation

To verify the proposed IGWL scheme, we perform many simulations and compare it with the traditional RTI. The computer configuration for the simulation is that the processor is AMD Ryzen 5 5600 CPU at 3.90 GHz, RAM is 16.00 GB, and the

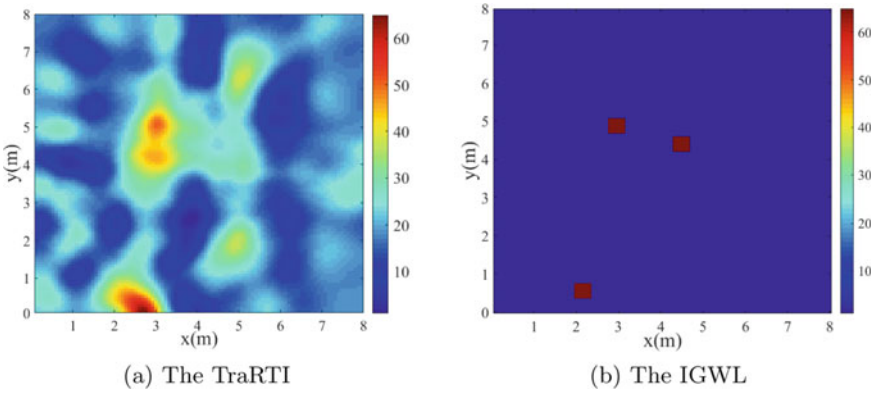


Fig. 1 The reconstruction images for three targets

operating system is Windows 10 64-bit. MATLAB 2018b is the simulation platform. The localization area 8 m×8 m is divided equally into 1600 squares. The target is considered as a square with 0.25 m^2. There are four reader antennas and forty tags around the localization area, and the distance between adjacent tags is 0.8 m. The transmitting power of the reader antenna is 32.5 dBm, the transmitting frequency is 922.625 MHz, and the beamwidth of the reader antenna is 60^o. The locations (2.33, 0.31), (2.82, 5.10), and (4.38, 4.27) of the three targets are examples to show the localization results of these two methods visually in Fig. 1. The traditional RTI cannot accurately localize three targets, while the proposed IGWL scheme can localize three targets and determine the number of targets.

Moreover, the Cumulative Distribution Function (CDF) of localization errors [16] is used to quantitatively analyze the performance of these two methods. We have done more than 300 simulations to obtain CDF curves shown in Fig. 2. Overall, the IGWL scheme has higher localization accuracy than the traditional RTI. Explain the performance of these two methods with nodes at an error of 0.5 m as follows. The probability that the traditional RTI localizes one, two, three, and four targets is 97, 37, 20, and 13% at the error, respectively, while the IGWL scheme is 100, 96, 85, and 64%, respectively. The probability that the error of the IGWL scheme is less than 0.5 m has increased by more than 44% on average.

5 Conclusion

We propose an IGWL scheme that directly utilizes wireless links to localize multiple targets. First, extract the intersections of target-affected links according to the first Fresnel zone. Then, determine core intersections with high respect to targets using graph theory. After, count the number of targets using the gap statistic method.

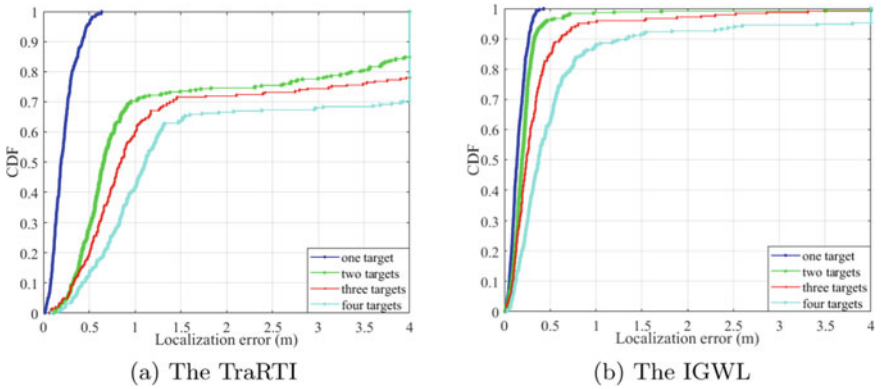


Fig. 2 The CDF of localization errors

Finally, acquire the targets' locations using modified K-means algorithm. The simulation shows that the localization accuracy of the proposed IGWL scheme is significantly higher than the traditional RTI.

Acknowledgment This work is supported by the Natural Science Foundation of China under Grant 61972279 and the Key Laboratory of Medical Electronics and Digital Health of Zhejiang Province under MEDH 2022014.

References

1. Shit RC, Sharma S, Puthal D, James P, Pradhan B, Moorsel AV, Zomaya AY, Ranjan R (2019) Ubiquitous localization (UBILOC): a survey and taxonomy on device free localization for smart world. IEEE Commun Surv Tutor 21(4):3532–3564
2. Kilic Y, Wymeersch H, Meijerink A, Bentum MJ, Scanlon WG (2014) Device-free person detection and ranging in UWB networks. IEEE J Select Top Signal Process 8(1):43–54
3. Seifeldin M, Saeed A, Kosba AE, El-Keyi A, Youssef M (2013) Nuzzer: a large-scale device-free passive localization system for wireless environments. IEEE Trans Mob Comput 12(7):1321–1334
4. Ma Y, Wang B, Gao X, Ning W (2020) The gray analysis and machine learning for device-free multi-target localization in passive UHF RFID environments. IEEE Trans Ind Inform 16(2):802–813
5. Tian C, Ma Y, Wang B (2023) Cooperative localization for passive RFID backscatter networks and theoretical analysis of performance limit. IEEE Trans Wireless Commun 22(2):1388–1402
6. Zhang J, Xiao W, Li Y (2021) Data and knowledge twin driven integration for large-scale device-free localization. IEEE Internet Things J 8(1):320–331
7. Ma Y, Ning W, Wang B, Liang X (2021) A data augmentation-based method for robust device-free localization in changing environments of passive radio frequency identification system. IEEE Trans Instr Measur 70:1–13
8. Wilson J, Patwari N (2011) See-through walls: motion tracking using variance-based radio tomography networks. IEEE Trans Mobile Comput 10(5):612–621
9. Zhao Y, Patwari N (2015) Robust estimators for variance-based device free localization and tracking. IEEE Trans Mobile Comput 14(10):2116–2129

10. Ma Y, Wang B, Ning W, Zhang Y (2020) PRSRTI: a novel device-free localization method using phase response shift based radio tomography imaging. IEEE Trans Vehicul Technol 69(11):13812–13820
11. Wagner B, Patwari N, Timmermann D (2012) Passive RFID tomographic imaging for device-free user localization. In: Proceedings of the 9th workshop on positioning navigation and communication (WPNC), Dresden, Germany, pp 120–125
12. Wagner B, Timmermann D (2013) Approaches for device-free multi-user localization with passive RFID. In: Proceedings of the international conference on indoor positioning and indoor navigation, Montbeliard-Belfort, France, pp 1–6
13. Ma Y, Zhang Y, Wang B, Ning W (2021) SCLA-RTI: a novel device-free multi-target localization method based on link analysis in passive UHF RFID environment. IEEE Sens J 21(3):3879–3887
14. Wang Q, Yigitler H, Jantti R (2016) Localizing multiple objects using radio tomographic imaging technology. IEEE Trans Vehicul Technol 65(5):3641–3656
15. Rappaport TS (2002) The wireless communications: principles and practice, 2nd edn. Prentice Hall, New York
16. Ju W, Xiaojiang C, Dingyi F, Qishi WC, Zhe Y, Tianzhang X (2015) Transferring compressive-sensing-based device-free localization across target diversity. IEEE Trans Ind Electr 62(4):2397–2409

Energy-Efficient Power Allocation for Multi-user D2D Underlay Communications in Distributed Antenna System

Guangying Wang, Tao Liu, Jinjin Zhu, and Xiangbin Yu

Abstract In this paper, the energy-efficient power allocation (PA) for multi-user Device-to-Device (D2D) underlay communications in distribu-ted antenna system is presented. The optimization problem of maximizing the energy efficiency (EE) of all D2D links is formulated with the minimum rate requirements and maximum transmit power constraints. To solve the non-linear and non-convex optimization problem with low complexity, the original problem is decomposed into two subproblems using the block coordinate descent algorithm. The non-convex subproblems are solved by the concave-convex procedure and successive convex approximation methods. Simulation results show that the EE performance improvement is achieved by the proposed PA scheme compared with the benchmark algorithm.

Keywords Device-to-Device (D2D) communication · Energy efficiency · Distributed antenna system · Power allocation

1 Introduction

Recently, the dramatic increase in the number of communication applications and smart terminal devices has led to explosive growth in data traffic. Energy Efficiency (EE) as one of the key performance indicators of next-generation communication systems, has attracted extensive attention and research from academia and industry [1].

Device-to-Device (D2D) communication has been considered one of the key technologies for future communication networks. It enables two devices to communicate directly which greatly improves the spectrum efficiency and EE [2, 3]. Distributed antenna system (DAS) has a very important role and research value in reducing com-

G. Wang (✉) · T. Liu · J. Zhu · X. Yu
College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China
e-mail: nuaawgy@foxmail.com

X. Yu
Key Laboratory of Wireless Sensor Network and Communication, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China

W. Wanget al. (eds.), *Communications, Signal Processing, and Systems,* Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_45

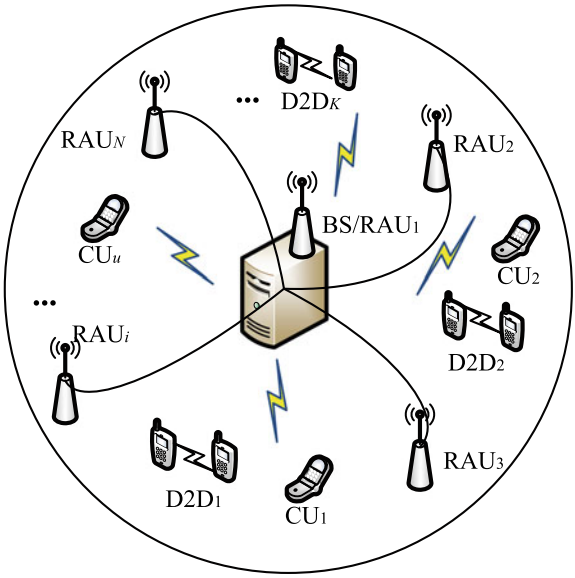
munication energy consumption and improving EE [4]. The EE optimization [5, 6] and power allocation (PA) [7] in D2D communications had been investigated. In [8], the energy-efficient PA scheme of single user was proposed for EE optimization.

In this paper, we propose the near-optimal PA scheme for multi-user D2D underlay communications in DAS. We consider the case of multiple CUs (cellular users) and multiple DUs (D2D users) in the system. The optimization problem of maximizing the EE of all D2D links is formulated under the minimum rate requirements and maximum transmit power constraints. To efficiently solve the formulated nonconvex problem, the original problem is decomposed into two subproblems using the block coordinate descent (BCD) algorithm. For the subproblem of DUs, the problem is solved by the concave-convex procedure (CCCP) method and fractional programming theory. For the subproblem of CUs, we utilize successive convex approximation (SCA) method to solve it. The effectiveness and superiority of the proposed PA algorithm are verified through simulation experiments.

2 System Model and Problem Formulation

We consider the system model including multi-CUs and multi-DUs underlay communications in uplink DAS, which consists of $U(u = 1, 2, ..., U)$ CUs and $K(k = 1, 2, ..., K)$ DUs. As shown in Fig. 1, the CUs intend to communicate with $N(i = 1, 2, ..., N)$ remote antenna units (RAUs). For simplicity, D2D-T represents the transmitter of DU, D2D-R stands for the receiver of DU.

Fig. 1 System model



In the uplink transmission, CUs transmit the data to the RAUs while D2D-T transmits data to D2D-R. Based on the above analysis, we can get the achievable rate of u-th CU

$$R_{c,u} = \log_2\left(1 + \sum_{i=1}^{N} \frac{p_u \left|g_{u,i}^c\right|^2}{\sum_{l=1,\neq u}^{U} p_l \left|g_{l,i}^c\right|^2 + \sum_{k=1}^{K} q_k \left|g_{k,i}^d\right|^2 + \sigma^2}\right), \tag{1}$$

The achievable rate of the k-th D2D pair is

$$R_{d,k} = \log_2\left(1 + \frac{q_k \left|g_{k,k}^d\right|^2}{\sum_{k'=1,\neq k}^{K} q_{k'} \left|g_{k',k}^d\right|^2 + \sum_{u=1}^{U} p_u \left|g_{u,k}^c\right|^2 + \sigma^2}\right), \tag{2}$$

where p_u is the transmit power of the u-th CU, q_k is the transmit power of the$k$-th D2D-T. The composite fading channel with path loss and Rayleigh fading is considered. Accordingly, the channel power gains are listed as follows

$$\begin{aligned} &g_{u,i}^c = h_{u,i}\sqrt{d_{u,i}^{-\beta}},\ g_{l,i}^c = h_{l,i}\sqrt{d_{l,i}^{-\beta}}, g_{k,i}^d = h_{k,i}\sqrt{d_{k,i}^{-\beta}}, \\ &g_{k,k}^d = h_{k,k}\sqrt{d_{k,k}^{-\beta}},\ g_{u,k}^c = h_{u,k}\sqrt{d_{u,k}^{-\beta}},\ g_{k',k}^d = h_{k',k}\sqrt{d_{k',k}^{-\beta}}, \end{aligned} \tag{3}$$

where $g_{u,i}^c$, $g_{l,i}^c$ and $g_{k,i}^d$ represent the channel gains from the u-th CU, l-th CU and the k-th transmitter of D2D to i-th RAUâĹŤs receiver, respectively. $g_{k,k}^d$ is the channel gain from$k$-th D2D-T to the k-th D2D-R. $g_{u,k}^c$and $g_{k',k}^d$ are the channel gains from the u-th CU and the k'-th transmitter of D2D to the D2D receiver, respectively. $h_{u,i}$ represents the small-scale fading coefficient between the u-th CU and i-th RAU, $h_{l,i}$ and $h_{k,i}$ are the small-scale fading coefficients between l-th CU and i-th RAU and between k-th D2D-T and i-th RAU, respectively. $h_{k,k}$ is the small-scale fading coefficient between k-th D2D-T and k-th D2D-R. $h_{u,k}$and $h_{k',k}$ are the small-scale fading coefficients between u-th CU and k-th D2D-R and between k'-th D2D-T and k-th D2D-R, respectively.Correspondingly, $L_{u,i}$, $L_{l,i}$, $L_{k,i}$, $L_{k,k}$, $L_{u,k}$, $L_{k',k}$ are the path loss, $d_{u,i}$, $d_{l,i}$, $d_{k,i}$, $d_{k,k}$, $d_{u,k}$, $d_{k',k}$ are the distances. β is the path loss exponent.

Based on the definition of EE, the EE of all D2D users can be defined as the ratio of total rate to power consumption. Thus, the objective problem for EE maximization can be written as

$$\begin{aligned} \max_{\mathbf{p},\mathbf{q}}\ & \eta_{EE} = \frac{\sum_{k=1}^{K} R_{d,k}}{\sum_{k=1}^{K} q_k + P_c} \\ \text{s.t. } & C_1 : R_{c,u} \geq R_{\min,c}, \forall u \\ & C_2 : R_{d,k} \geq R_{\min,d}, \forall k \\ & C_3 : 0 \leq p_u \leq P_{\max,c}, \forall u \\ & C_4 : 0 \leq q_k \leq P_{\max,d}, \forall k \end{aligned} \tag{4}$$

where $\mathbf{p} = [p_1, ..., p_u, ..., p_U]$ and $\mathbf{q} = [q_1, ..., q_k, ..., q_K]$ represent the power allocation vectors of the CUs and the DUs. P_c means the static circuit power consumption with a constant value. The observations reveal that the optimization variables $\mathbf{p}$ and $\mathbf{q}$ are coupled with each other. Both the objective function and the minimum rate constraints are non-linear and non-convex, so it is very challenging and intractable to solve the optimization problem.

3 Power Allocation Schemes for EE Maximization

The near-optimal scheme based on the BCD algorithm is designed in this section to alternatively optimize the PA for different users. We optimize the transmit power $\mathbf{q}$ of D2D users with given $\hat{\mathbf{p}}$. Without considering the inequality constraint C_1, the optimization problem (4) can be rewritten as

$$\begin{aligned} \max_{\mathbf{q}} \ & \eta_{EE}(\hat{\mathbf{p}}, \mathbf{q}) \\ \text{s.t.} \ & R_{d,k} \geq R_{\min,d}, \forall k, \\ & 0 \leq q_k \leq P_{\max,d}, \forall k. \end{aligned} \tag{5}$$

The obtained optimization problem (5) is still a difficult non-convex problem. We use CCCP method to convexify the non-convex function. Accordingly, we can obtain

$$\begin{aligned} R_{d,k} &= f_1(\mathbf{q}) - f_2(\mathbf{q}) \\ &= \log_2 \left(q_k \left| g_{k,k}^d \right|^2 + \sum_{k'=1,\neq k}^{K} q_{k'} \left| g_{k',k}^d \right|^2 + \sum_{u=1}^{U} p_u \left| g_{u,k}^c \right|^2 + \sigma^2 \right) \\ &\quad - \log_2 \left(\sum_{k'=1,\neq k}^{K} q_{k'} \left| g_{k',k}^d \right|^2 + \sum_{u=1}^{U} p_u \left| g_{u,k}^c \right|^2 + \sigma^2 \right) \end{aligned} \tag{6}$$

We adopt the first-order Taylor expansion to linearize the $f_2(\mathbf{q})$ at the initial value $\mathbf{q}_0$. Then the approximation $f_2(\mathbf{q})$ can be expressed as

$$f_2(\mathbf{q}|\mathbf{q}_0) \approx f_2(\mathbf{q}_0) + \nabla f_2(\mathbf{q}_0)(\mathbf{q} - \mathbf{q}_0)^T \tag{7}$$

where$\nabla f_2(\mathbf{q}_0)$ is the gradient of $f_2(\mathbf{q})$ at $\mathbf{q}_0$,

$$\nabla f_2(\mathbf{q}_0)(\mathbf{q} - \mathbf{q}_0)^T = \frac{\sum_{k'=1,\neq k}^{K} \left| g_{k',k}^d \right|^2 (q_{k'} - q_{k',0})}{\left(\sum_{k'=1,\neq k}^{K} q_{k'} \left| g_{k',k}^d \right|^2 + \sum_{u=1}^{U} p_u \left| g_{u,k}^c \right|^2 + \sigma^2 \right) \ln 2} \tag{8}$$

Let $F_1(\mathbf{q}) = \sum_{k=1}^{K} f_1(\mathbf{q})$, $F_2(\mathbf{q}) = \sum_{k=1}^{K} f_2(\mathbf{q})$, we use the first-order Taylor expansion to linearize the $F_2(\mathbf{q})$

$$F_2(\mathbf{q}|\mathbf{q}_0) \approx F_2(\mathbf{q}_0) + \nabla F_2(\mathbf{q}_0)(\mathbf{q}-\mathbf{q}_0)^T \tag{9}$$

where$\nabla F_2(\mathbf{q}_0)$ is the gradient of $F_2(\mathbf{q})$ at $\mathbf{q}_0$,

$$\begin{aligned}&\nabla F_2(\mathbf{q}_0)(\mathbf{q}-\mathbf{q}_0)^T \\ &= \tfrac{1}{\ln 2}\sum_{k=1}^{K}\left[\left(\sum_{j=1,j\neq k}^{K}\frac{\left|g_{k,j}^{d}\right|^2}{\sum_{k'=1,\neq j}^{K} q_{k'}\left|g_{k',j}^{d}\right|^2+\sum_{u=1}^{U} p_u\left|g_{u,k}^{c}\right|^2+\sigma^2}\right)(q_k-q_{k,0})\right]\end{aligned} \tag{10}$$

In terms of the fractional programming theory, the problem (5) can be transformed

$$\begin{aligned}&\max_{\mathbf{q}}\ \eta_2 = F_1(\mathbf{q}) - F_2(\mathbf{q}_0) - \nabla F_2(\mathbf{q}_0)(\mathbf{q}-\mathbf{q}_0)^T - \omega_1\left(\sum_{k=1}^{K} q_k + P_c\right)\\ &s.t.\quad f_1(\mathbf{q}) - f_2(\mathbf{q}_0) - \nabla f_2(\mathbf{q}_0)(\mathbf{q}-\mathbf{q}_0)^T \geq R_{\min,d},\ \forall k,\ k',\ u,\\ &\qquad 0 \leq q_k \leq P_{\max,d}, \forall k.\end{aligned} \tag{11}$$

where ω_1 is non-negative fractional programming factor

$$\omega_1 = \frac{F_1(\mathbf{q}) - F_2(\mathbf{q}_0) - \nabla F_2(\mathbf{q}_0)(\mathbf{q}-\mathbf{q}_0)^T}{\sum_{k=1}^{K} q_k + P_c} \tag{12}$$

Since the transformed optimization problem (11) is convex, we adopt the convex optimization software CVX for obtaining the optimal value.

In the following, we optimize the transmit power $\mathbf{p}$ of CUs with given $\hat{\mathbf{q}}$, for which the problem can be transformed as

$$\begin{aligned}&\max_{\mathbf{p}}\ \eta_{EE}(\mathbf{p}, \hat{\mathbf{q}})\\ &s.t.\quad R_{c,u} \geq R_{\min,c},\ \forall u,\\ &\qquad 0 \leq p_u \leq P_{\max,c}, \forall u.\end{aligned} \tag{13}$$

For the minimum rate constraint, we first introduce the auxiliary variable $\mathcal{T}_i$

$$\begin{aligned}&\sum_{i=1}^{N}\frac{p_u\left|g_{u,i}^{c}\right|^2}{\sum_{l=1,\neq u}^{U} p_l\left|g_{l,i}^{c}\right|^2+\sum_{k=1}^{K} q_k\left|g_{k,i}^{d}\right|^2+\sigma^2} \geq \gamma_{th}=2^{R_{\min,c}}-1\\ &\sum_{i=1}^{N}\mathcal{T}_i \geq \gamma_{th},\quad \forall i.\end{aligned} \tag{14}$$

We can further obtain

$$\frac{p_u\left|g_{u,i}^{c}\right|^2}{\sum_{l=1,\neq u}^{U} p_l\left|g_{l,i}^{c}\right|^2+\sum_{k=1}^{K} q_k\left|g_{k,i}^{d}\right|^2+\sigma^2} \geq \mathcal{T}_i \tag{15}$$

The optimization problem (13) can be converted to

$$\begin{aligned}
\max_{\mathbf{p}} \ & \frac{1}{Q}\sum_{k=1}^{K}\log_2\left(1+\frac{q_k\left|g_{k,k}^{d}\right|^2}{\sum_{k'=1,\neq k}^{K}q_{k'}\left|g_{k',k}^{d}\right|^2+\sum_{u=1}^{U}p_u\left|g_{u,k}^{c}\right|^2+\sigma^2}\right) \\
s.t. \ & \sum_{i=1}^{N}\mathcal{T}_i \geq \gamma_{th}, \ \ \forall i, \\
& \frac{p_u\left|g_{u,i}^{c}\right|^2}{\sum_{l=1,\neq u}^{U}p_l\left|g_{l,i}^{c}\right|^2+\sum_{k=1}^{K}q_k\left|g_{k,i}^{d}\right|^2+\sigma^2} \geq \mathcal{T}_i, \ \forall i, \ k, \ l, u, \\
& 0 \leq p_u \leq P_{\max,c}, \forall u.
\end{aligned} \tag{16}$$

where $Q=\sum_{k=1}^{K}q_k+P_c$.

Furthermore, by making $p_u=\exp(x_u)$, $\phi_1=\sum_{k=1}^{K}q_k\left|g_{k,i}^{d}\right|^2+\sigma^2$, we can get $\frac{\exp(x_u)\left|g_{u,i}^{c}\right|^2}{\sum_{l=1,\neq u}^{U}\exp(x_l)\left|g_{l,i}^{c}\right|^2+\phi_1}\geq\mathcal{T}_i$, i.e. $\frac{1}{\mathcal{T}_i}\geq\frac{\sum_{l=1,\neq u}^{U}\exp(x_l)\left|g_{l,i}^{c}\right|^2+\phi_1}{\exp(x_u)\left|g_{u,i}^{c}\right|^2}$.

We linearize $1/\mathcal{T}_i$ by first order Taylor expansion in every iterative step at the point $\tilde{\mathcal{T}}_i$ yields

$$\frac{1}{\tilde{\mathcal{T}}_i}-\frac{1}{\tilde{\mathcal{T}}_i^2}\left(\mathcal{T}_i-\tilde{\mathcal{T}}_i\right)\geq\frac{\sum_{l=1,\neq u}^{U}\exp(x_l-x_u)\left|g_{l,i}^{c}\right|^2+\phi_1\exp(-x_u)}{\left|g_{u,i}^{c}\right|^2} \tag{17}$$

The optimization problem (16) can be further transformed to

$$\begin{aligned}
\max_{x_u,\mathcal{T}_i} \ & \frac{1}{Q}\sum_{k=1}^{K}\log_2\left(1+\frac{q_k\left|g_{k,k}^{d}\right|^2}{\sum_{u=1}^{U}\exp(x_u)\left|g_{u,k}^{c}\right|^2+\sum_{k'=1,\neq k}^{K}q_{k'}\left|g_{k',k}^{d}\right|^2+q_k\varepsilon_{k,k}^{d}+\sigma^2}\right) \\
s.t. \ & \sum_{i=1}^{N}\mathcal{T}_i\geq\gamma_{th}, \ \forall i, \\
& \frac{1}{\tilde{\mathcal{T}}_i}-\frac{1}{\tilde{\mathcal{T}}_i^2}\left(\mathcal{T}_i-\tilde{\mathcal{T}}_i\right)\geq\frac{\mathcal{F}_l}{\left|g_{u,i}^{c}\right|^2}, \ \forall i,u, \\
& \exp(x_u)\leq P_{\max,c}, \forall u.
\end{aligned} \tag{18}$$

The optimization subproblem (18) is still a non-convex optimization problem since the objective function is non-concave. We next use the SCA method [9] to solve it. According to the SCA principle, for any $x\geq 0, \bar{x}\geq 0$, the tight lower bound of SCA for a non-convex problem can be expressed as

$$\log_2(1+x)\geq a\log_2(x)+b \tag{19}$$

where $a=\frac{\bar{x}}{1+\bar{x}}$, $b=\log_2(1+\bar{x})-a\log_2(\bar{x})$. The optimization problem (18) is transformed into

$$\begin{aligned}
&\max_{x_u,\mathcal{T}_i} \ \frac{1}{Q}\sum_{k=1}^{K}\left[-a_k\log_2\left(\frac{1}{q_k\left|g_{k,k}^{d}\right|^2}\left[\sum_{u=1}^{U}\exp(x_u)\left|g_{u,k}^{c}\right|^2+\phi_2\right]\right)+b_k\right]\\
&s.t. \quad \sum_{i=1}^{N}\mathcal{T}_i \geq \gamma_{th}, \ \forall i,\\
&\quad \frac{1}{\tilde{\mathcal{T}}_i}-\frac{1}{\tilde{\mathcal{T}}_i^2}\left(\mathcal{T}_i-\tilde{\mathcal{T}}_i\right) \geq \frac{\mathcal{F}_l}{\left|g_{u,i}^{c}\right|^2}, \ \forall i,u,\\
&\quad \exp(x_u) \leq P_{\max,c}, \forall u.
\end{aligned} \tag{20}$$

where $\phi_2=\sum_{k'=1,\neq k}^{K} q_{k'}\left|g_{k',k}^{d}\right|^2+q_k\varepsilon_{k,k}^{d}+\sigma^2$. It can be found that the converted optimization problem (20) is a convex one, which can be solved by standard optimization tools.

4 Simulation Results

In this section, we present simulations to examine the performance of our proposed algorithm. The main parameters are set as:$N=5$, $\beta=3$, $P_c=1.5\,\mathrm{W}$, $\sigma_z^2=-104\,\mathrm{dBm}$, U=2, K=2 and $r=1000\,\mathrm{m}$. There are N RAUs distributed uniformly in a cell. The polar coordinate of BS/RAU1 is (0, 0), and other RAUs are $\left(\sqrt{3/7}r,\ \pi i/3\right)$, $i=1,\cdots,N-1$. The minimum rate requirement is set as 1 bit/s/Hz. For convenience, we assume that $P_{\max,d}=P_{\max}$.

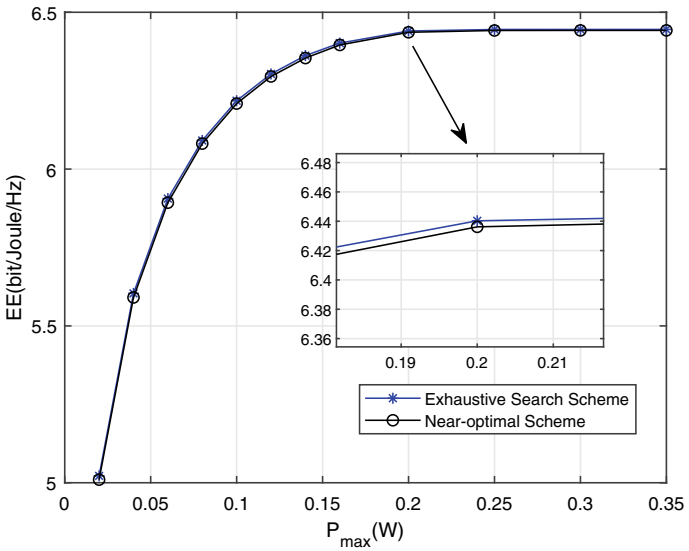


Fig. 2 EE with different schemes

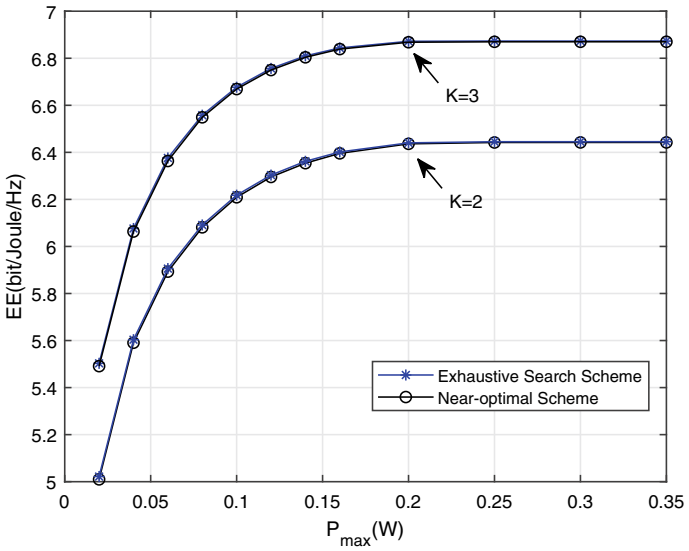


Fig. 3 EE under different number of DUs

Figure 2 shows EE performance of proposed near-optimal algorithm and exhaustive search algorithm. The EE value increases with the increase of the transmit power, and it tends to flatten out and stabilize at a fixed value eventually. When the transmit power of D2D users is small, the EE performance increases faster. The EE performance of near-optimal scheme is almost the same as that of the exhaustive search method. The results prove the effectiveness of the PA algorithm designed in this paper.

Figure 3 compares EE performance when there are different numbers of DUs in the system. The simulation results clearly show that the energy efficiency has been improved through the proposed power allocation algorithm. We give the uplink EE performances of two DUs and three DUs, respectively. With the increase of DUs, the performance of EE has been significantly improved. This is because more D2D users have accessed the communication system through multiplexing, which increases the overall data transmission rate. The increase in D2D user pairs has significantly improved EE performance. These results confirm the advantages of D2D communications in improving system capacity and energy efficiency.

5 Conclusion

This paper studies the EE optimization and PA algorithm for multi-user DAS-D2D communication systems. The EE maximization problem based on minimum rate constraints and maximum power constraints is presented. The formulated problem

is a non-linear and non-convex problem, which usually has no feasible solution. To tackle this issue, the original problem is decomposed into two PA subproblems by BCD algorithm. For the non-convex subproblems, we utilize CCCP method and SCA method to obtain the near-optimal PA solutions. The effectiveness of proposed scheme is analyzed through simulation experiments.

Acknowledgements This work was supported by Open Research Fund Key Laboratory of Wireless Sensor Network and Communication of Chinese Academy of Sciences (2017006).

References

1. Buzzi S, Chih-Lin I, Klein TE, Poor HV, Yang C, Zappone A (2016) A survey of energy-efficient techniques for 5G networks and challenges ahead. J IEEE Sel Areas Commun 34(4):697–709
2. Zhang S, Liu J, Guo H, Qi M, Kato N (2020) Envisioning device-to-device communications in 6G. J IEEE Network 34(3):86–91
3. Doppler K, Rinne M, Wijting C, Ribeiro CB, Hugl K (2009) Device-to-device communication as an underlay to LTE-advanced networks. J IEEE Commun Mag 47(12):42–49
4. He C, Sheng B, Zhu P, You X (2012) Energy efficiency and spectral efficiency tradeoff in downlink distributed antenna systems. J IEEE Wirel Commun Lett 1(3):153–156
5. Su Z, Feng W, Tang J, Chen Z, Fu Y, Zhao N, Wong KK (2022) Energy efficiency optimization for D2D communications underlaying UAV-assisted industrial IoT networks with SWIPT. J IEEE Internet Things J 10(3):1990–2002
6. Li X, He C, Huang L, Zhang C, Zhang J (2018) Energy efficient power allocation for co-located antenna systems with D2D communication. J AEU-Int J Electron Commun 83:100–105
7. Chen P, Zhou X, Zhao J, Shen F, Sun S (2022) Energy-efficient resource allocation for secure D2D communications underlaying UAV-enabled networks. J IEEE Trans Veh Technol 71(7):7519–7531
8. Wang G, Yu X, Teng T (2021) Energy-efficient power allocation scheme for uplink distributed antenna system with D2D communication. J Mob Netw Appl 26:1225–1232
9. Razaviyayn M (2014) Successive convex approximation: analysis and applications. University of Minnesota, Diss

UAV Autonomous Landing Pose Estimation Using Monocular Vision Based on Cooperative Identification and Scene Reconstruction

Xinyan Zhao, Lin Ma, and Danyang Qin

Abstract Nowadays UAVs are often employed in a weak or GPS signals unavailable environment which requires high demand for the UAV landing autonomously. In this paper, propose a monocular vision UAV autonomous landing pose estimation method based on cooperative identification and scene reconstruction. A rotating target detection algorithm specific to aerial images is used to identify and locate the target with the apron in aerial images. When the UAV lands to a height where the detail information of the cooperative identification on the apron can be extracted, firstly, the key points are extracted from the single frame images acquired by the airborne monocular camera and matched with the key points saved in the airborne database for geometric verification to filter out the wrong matching relationships. Then, the 3D coordinates of feature points saved in the onboard database are used to obtain the 2D-3D matching relationship and perform co-visual relationship screening to obtain stable matching relationships. Finally, the PnP problem is solved by BA optimization method, and the position and yaw angle of the UAV relative to the mobile apron are calculated according to the similar transformation matrix saved in the airborne database. The experiment results indicate that the proposed method improves the accuracy of UAV pose estimation and can be adopted as an alternate UAV autonomous landing technology in a narrow mobile apron.

Keywords UAV · Autonomous landing · Pose estimation · Monocular vision

X. Zhao · L. Ma (✉)
School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, China
e-mail: malin@hit.edu.cn

D. Qin
School of Electronics Engineering, Heilongjiang University, Harbin, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_46

1 Introduction

Currently, unmanned aerial vehicle (UAV) is widely used by a self-contained program and radio remote control [1]. Image data acquisition by drones can complete ground exploration in low altitude areas. Combining ground communication technology, satellite communication technology and remote sensing technology, UAVs can achieve intelligent monitoring of space-air-ground integration. With the increasing maturity of UAV technology, there is a growing demand for UAV autonomous landing technology. Autonomous drone landing technology makes drones more intelligent and saves labor costs. In addition, the use of autonomous drone landing technology in conjunction with mobile aprons improves the flexibility of drone usage and greatly expands the range of applications for drones.

Nowadays, autonomous landing technologies for UAVs include Inertial Navigation System (INS), Instrument Landing System (ILS), Microware Landing System (MLS) and Global Navigation Satellite System (GNSS) [2]. Among them, the INS has accumulated errors over time due to integration calculation; the ILS and MLS have strong dependence on ground equipment, leading to the limited use environment; the GNSS positioning accuracy is limited to give the exact location of the apron. With the increasingly advanced production process of vision sensors and the booming development of computer vision technology, the pose estimation of UAVs using the vision-based method which has a high accuracy of position estimation, especially in short-range measurement tasks, has become a research hotspot in the field of UAVs. Currently, UAV autonomous landing techniques generally use GPS signals to guide the UAV to a specified apron when it is far from the apron, and use vision-based methods for UAV position and attitude estimation relative to the landing platform when the UAV is close to the apron. In order to realize the autonomous and precise landing of UAVs to mobile platforms, Guo et al. proposed a hierarchical landing strategy using BeiDou satellite positioning signal to guide, machine vision positioning and ultrasonic ranging [3]. To address the problem of restricted landing area, Zhao et al. proposed a cooperative UAV/boat autonomous landing technology based on Differential Global Navigation Satellite System (DGNSS) signal guidance and landing landmarks detection [4]. However, when the UAV is located in special environments where the GPS signal is blocked, the receiver cannot receive the positioning signal to obtain the exact location of the target with the apron [5, 6]. When the GPS signal is interfered by electromagnetic waves, the receiver cannot correctly decode the positioning information in the GPS signal [7]. In addition, existing UAV position estimation methods cannot achieve safe and accurate UAV landings when the size of the apron is limited. Wang et al. proposed an algorithm based on ArUco markers for estimating the position, attitude and velocity of UAVs during the landing phase [8]. To improve the rate and accuracy of estimating attitude angle, Bo et al. designed an algorithm for relative attitude measurement based on color square identification [9], which enables accurate estimation of attitude angle within 3 m from the identification. Gao et al. implemented an UAV landing algorithm based on AprilTag code, with a positioning error of 0.2 m [10]. In order to achieve safe and accurate

landing of UAVs under the conditions of weak GPS signals, Ye et al. proposed a method based on H-pattern marking, which obtained a maximum estimation error of 0.52 m and 6.22° for position and yaw angle, and the method basically met the landing requirements [11].

Therefore, in this paper we propose a UAV pose estimation method based on cooperative identification and motion reconstruction with monocular vision. A rotating target detection algorithm for aerial image is used to identify and locate the target with the apron. When the UAV can obtain detailed information about the cooperative identification, a monocular vision UAV pose estimation algorithm based on the cooperative identification and scene reconstruction is used to calculate the pose of the UAV relative to the mobile apron, enabling accurate pose estimation during the autonomous landing of the UAV.

The remainder of this paper is organized as follows. In Sect. 2, we present the monocular vision-based autonomous UAV landing method, giving the framework of UAV autonomous landing method we propose and the coordinate system model used. In Sect. 3, the problem formulation and the solution of the method are given. In Sect. 4, we give the algorithm implementation and performance analysis. Finally, the conclusions are given in Sect. 5.

2 System Model

2.1 Overview

The UAV autonomous landing pose estimation method based on cooperative identification and scene reconstruction proposed in this paper has two parts, which are image target detection and UAV pose estimation shown in Fig. 1. The first part is to identify and locate the target containing the apron in aerial images. The second part is further divided into offline phase and online phase, whose main task is to reconstruct the 3D scene structure of the 2D identification and estimate the position and angle of the UAV relative to the 2D identification.

For the image target detection, we proposed to use YOLOv5 network and DOTA-v1.0 to train the for the detection of targets with aprons in the input aerial images, where DOTA-v1.0 is a large dataset suitable for image target detection. For the UAV pose estimation, we directly estimate the pose of the camera as the UAV's pose. In the offline phase, reconstruction of the assisted landing markers is carried out to fund the UAV real-time pose solution in the online phase. In the offline phase, the 3D scene structure of the 2D identification is reconstructed with the method of incremental reconstruction [12]. The modelling results, the images used for reconstruction and the pixel coordinates and the spatial coordinates of the key points are saved in the database. In addition, the similar transformation matrix is calculated and saved. In the online phase, firstly, key points are extracted from the query image captured by monocular camera and matched with the key points saved in the airborne

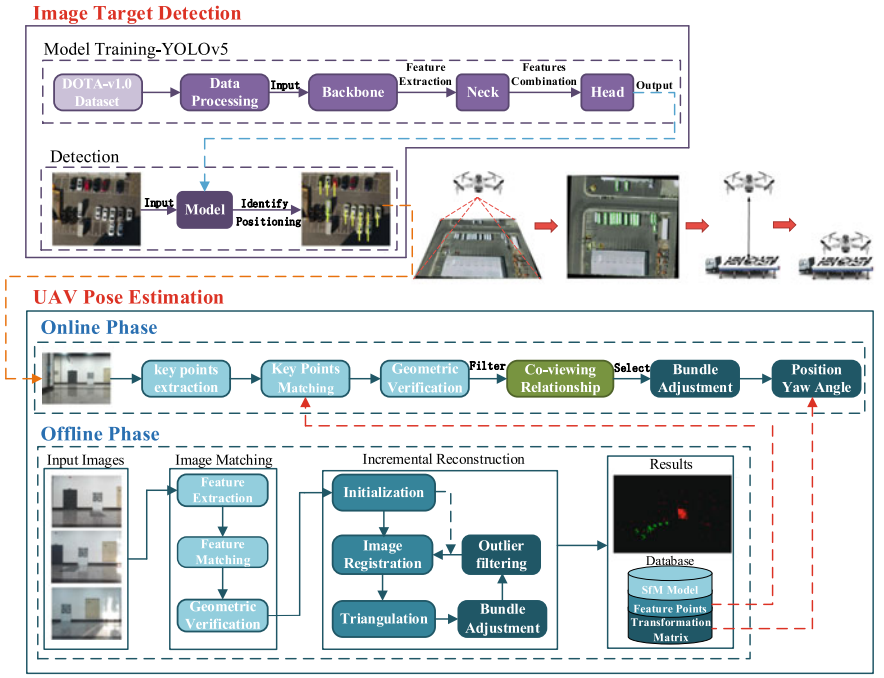


Fig. 1 Proposed method overview

database to obtain a 2D-2D matched pairs. The matched pairs are filtered using the RANSAC algorithm to remove false matches. The 2D-3D matched pairs can then be obtained from the spatial point coordinates saved in the database and the 2D-2D match matched pairs described above, and the stable 2D-3D matched pairs can be obtained by selection using the co-viewing relationship. Bundle Adjustment is used to solve for the camera pose in the SfM model coordinate system. Finally, the camera position and yaw angle in the 2D identification coordinate system can be calculated based on the similar transformation matrix in the airborne database.

2.2 *Coordinate System*

The UAV pose estimation involves the transformation of spatial point coordinates between multiple coordinate systems. The relation of these coordinate systems used in this paper are shown in Fig. 2.

The apron coordinate system O_w-x_w-y_w-z_w is based on the top left corner of the apron as the origin O_w, the plane in which the 2D identification is located as the $O_w x_w y_w$ face, and the z_w axis perpendicular to the $O_w x_w y_w$ face upwards, with the

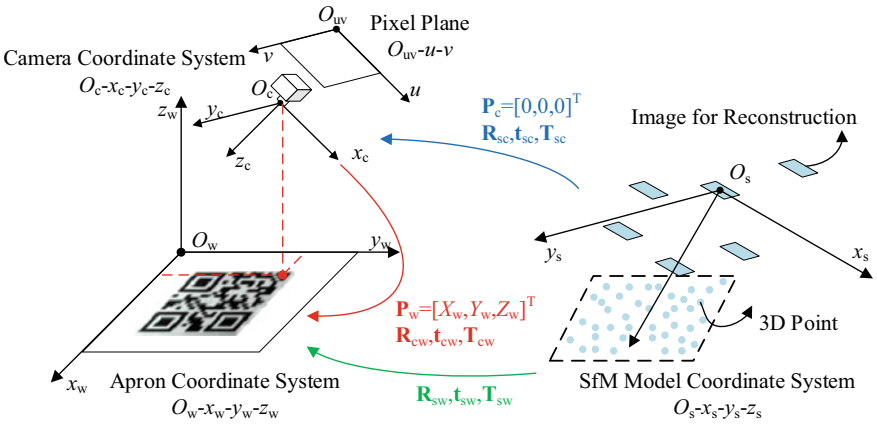


Fig. 2 Relation of different coordinate system

x_w axis and the y_w axis in the 2D identification plane and parallel to the border of the identification, this coordinate system is measured in metres.

The SfM model coordinate system O_s-x_s-y_s-z_s, which is the coordinate system established during the reconstruction with the pose of one of the images undergoing initialization as the origin.

The camera coordinate system O_c-x_c-y_c-z_c is based on the origin of the coordinate system with the camera optical center O_c, the z_c axis outwards perpendicular to the camera plane, the x_c and y_c axes in the camera plane, the x_c axis to the right and the y_c axis downwards. This coordinate system represents the real-time poses of the camera. The transformation relationships between the above coordinate systems are shown in Table 1.

Table 1 Coordinate system conversion

Coordinate system conversion	Rotation matrix	Translation vector	Transformation matrix
SfM model → apron	$\mathbf{R}_{sw}$	$\mathbf{t}_{sw}$	$\mathbf{T}_{sw}$
SfM model → camera	$\mathbf{R}_{sc}$	$\mathbf{t}_{sc}$	$\mathbf{T}_{sc}$
Camera → apron	$\mathbf{R}_{cw}$	$\mathbf{t}_{cw}$	$\mathbf{T}_{cw}$

3 Proposed Method

3.1 Camera Model, Pose Representation and Similar Transformation Matrix Solution

During an autonomous landing, a drone needs to estimate its attitude relative to a 2D identification. The attitude is represented by a rotation matrix, a displacement vector or a transformation matrix. The relationship between **R**, **t** and **T** can be expressed as follows:

$$\mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0^{\mathrm{T}} & 1 \end{bmatrix} \tag{3.1}$$

A spatial point P is defined as $\mathbf{P}_1 = [x_1, y_1, z_1]^{\mathrm{T}}$ in coordinate system 1, and is defined as $\mathbf{P}_2 = [x_2, y_2, z_2]^{\mathrm{T}}$ in coordinate system 2. The transformation relationship between the coordinates of point P in the two coordinate systems can be expressed in terms of the rotation matrix and the displacement vector or transformation matrix as follows:

$$\begin{cases} \mathbf{P}_2 = \mathbf{R}\mathbf{P}_1 + \mathbf{t} \\ \mathbf{P}_2 = \mathbf{T}\mathbf{P}_1 \end{cases} \tag{3.2}$$

In vision-based methods for UAV pose estimation, the imaging principle of the camera needs to be described. In this paper, images are acquired using a monocular camera, the imaging principle of which is the process by which light emitted or reflected from objects in the three-dimensional world is projected across the camera optical center onto the imaging plane. The relationship between the three-dimensional coordinates of a spatial point in the camera coordinate system and the coordinates of its projection in the pixel plane is expressed as follows:

$$Z_{\mathrm{c}} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_{\mathrm{u}} & 0 & c_{\mathrm{u}} \\ 0 & f_{\mathrm{v}} & c_{\mathrm{v}} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{\mathrm{c}} \\ Y_{\mathrm{c}} \\ Z_{\mathrm{c}} \end{bmatrix} \triangleq \mathbf{K}\mathbf{P}_{\mathrm{c}} \tag{3.3}$$

where **K** is the internal parameter matrix of the camera, calculated by the camera calibration.

To calculate the camera's pose in the apron coordinate system, the transformation matrix $\mathbf{T}_{\mathrm{sw}}$ which is expressed as (3.4) between the SfM model coordinate system and the apron coordinate system needs to be calculated in the offline phase.

$$\mathbf{T}_{\mathrm{sw}} = \begin{bmatrix} s_{\mathrm{sw}}\mathbf{R}_{\mathrm{sw}} & \mathbf{t}_{\mathrm{sw}} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \tag{3.4}$$

The coordinates of the spatial point P_i in the SfM model coordinate system and the apron coordinate system are $\mathbf{P}_{si}$ and $\mathbf{P}_{wi}$. Modeling the problem of solving $\mathbf{T}_{sw}$ as assuming that $\mathbf{P}_{wi}$ is the projection coordinate of $\mathbf{P}_{si}$ into the apron coordinate system, the projection error is expressed as follows:

$$\min_{s_{sw},\mathbf{R}_{sw},\mathbf{t}_{sw}} \frac{1}{2}\sum_{i=1}^{n} \|\mathbf{P}_{wi} - (s_{sw}\mathbf{R}_{sw}\mathbf{P}_{si} + \mathbf{t}_{sw})\|_2^2 \tag{3.5}$$

The transformation matrix $\mathbf{T}_{sw}$ is the value that minimizes the projection error. An analytical solution to the above problem of minimizing the projection error exists. The control points under the SfM model coordinate system and the apron coordinate system are expressed as follows:

$$\begin{cases} \mathbf{P}_{sc} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \mathbf{P}_{si} \\ \mathbf{P}_{wc} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \mathbf{P}_{wi} \end{cases} \tag{3.6}$$

The decentered coordinates of the spatial points under the two coordinate systems are $\mathbf{P}'_{si} = \mathbf{P}_{si} - \mathbf{P}_{sc}$ and $\mathbf{P}'_{wi} = \mathbf{P}_{wi} - \mathbf{P}_{wc}$. The problem of minimizing the projection error is deformed as follows:

$$\begin{aligned} \min_{s_{sw},\mathbf{R}_{sw},\mathbf{t}_{sw}} \sum_{i}^{n} \|e_i\|_2^2 = & \min_{s_{sw},\mathbf{R}_{sw},\mathbf{t}_{sw}} \sum_{i}^{n} \left\|\mathbf{P}'_{wi} - s_{sw}\mathbf{R}_{sw}\mathbf{P}'_{si}\right\|_2^2 \\ & + \min_{s_{sw},\mathbf{R}_{sw},\mathbf{t}_{sw}} \sum_{i}^{n} \|\mathbf{P}_{wc} - s_{sw}\mathbf{R}_{sw}\mathbf{P}_{sc} - \mathbf{t}_{sw}\|_2^2 \\ & + \min_{s_{sw},\mathbf{R}_{sw},\mathbf{t}_{sw}} 2(\mathbf{P}_{wc} - s_{sw}\mathbf{R}_{sw}\mathbf{P}_{sc} - \mathbf{t}_{sw})^{\mathrm{T}} \\ & \sum_{i}^{n} \left(\mathbf{P}'_{wi} - s_{sw}\mathbf{R}_{sw}\mathbf{P}'_{si}\right) \end{aligned} \tag{3.7}$$

where $\mathbf{t}_{sw}$ is only associated with the second term, $\mathbf{t}_{sw}$ should be expressed as:

$$\mathbf{t}_{sw} = \mathbf{P}_{wc} - s_{sw}\mathbf{R}_{sw}\mathbf{P}_{sc} \tag{3.8}$$

The above problem can be reduced as follows:

$$\min_{s_{sw},\mathbf{R}_{sw}} \sum_{i=1}^{n} \left\| \mathbf{P}'_{wi} - s_{sw}\mathbf{R}_{sw}\mathbf{P}'_{si} \right\|^2$$
$$= \min_{s_{sw},\mathbf{R}_{sw}} \frac{1}{n}\left(\sum_{i=1}^{n} \left\| \mathbf{P}'_{wi} \right\|^2 + \sum_{i=1}^{n} \left\| s_{sw}\mathbf{R}_{sw}\mathbf{P}'_{si} \right\|^2 - 2s_{sw} \sum_{i=1}^{n} \mathbf{P}'^{\mathrm{T}}_{wi}\mathbf{R}_{sw}\mathbf{P}'_{si} \right) \quad (3.9)$$

where the first and second terms are the variances of the control points in the apron coordinate system and the SfM model coordinate system, and the third term is the covariance of the projected control points, denoted as σ_1, σ_2, and D respectively. The above equation can be expressed as follows:

$$\min_{s_{sw},\mathbf{R}_{sw}} \sum_{i=1}^{n} \left\| \mathbf{P}'_{wi} - s_{sw}\mathbf{R}_{sw}\mathbf{P}'_{si} \right\|_2^2 = \min_{s_{sw},\mathbf{R}_{sw}} \left(s_{sw}\sqrt{\sigma_1} - D/\sqrt{\sigma_1} \right)^2 + \sigma_2 - D^2/\sigma_1 \quad (3.10)$$

where s_{sw} occurs only in the first term, so the optimal scale factor can be expressed as:

$$s_{sw} = D/\sigma_1 \quad (3.11)$$

The above problem is further simplified as:

$$\min_{\mathbf{R}_{sw}} \left(\sigma_2 - D^2/\sigma_1 \right) = \max_{\mathbf{R}_{sw}} D = \max_{\mathbf{R}_{sw}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{P}'^{\mathrm{T}}_{wi}\mathbf{R}_{sw}\mathbf{P}'_{si} = \max_{\mathbf{R}_{sw}} \mathrm{tr}(\mathbf{R}_{sw}\mathbf{Q}) \quad (3.12)$$

where $\mathbf{Q}$ is expressed as follows:

$$\mathbf{Q} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{P}'_{si}\mathbf{P}'^{\mathrm{T}}_{wi} \quad (3.13)$$

Performing the SVD decomposition on $\mathbf{Q}$, D can be further expressed as follows:

$$\mathrm{D} = \mathrm{tr}\left(\Sigma \mathbf{V}^{\mathrm{T}}\mathbf{R}_{sw}\mathbf{U} \right) \quad (3.14)$$

When D achieves its maximum value, $\mathbf{R}_{sw}$ can be expressed as follows:

$$\mathbf{R}_{sw} = \mathbf{V}\mathbf{M}\mathbf{U}^{\mathrm{T}} \quad (3.15)$$

where $\mathbf{M} = \mathrm{diag}(1, 1, |\mathbf{VU}|)$. At this point, the similar transformation matrix $\mathbf{T}_{sw}$ is completely solved.

3.2 Online Pose Resolution

Firstly, key points need to be extracted from the single frames and matched with the key points saved in the airborne database. In this paper, the SIFT feature extraction algorithm is adopted with scale and rotation invariance in order to obtain stable key points. Using the Lowe ratio test to match key points, for each key point $kp_{t,m}$ ($m = 1, 2, \ldots, M$) in image frame I_t, the feature descriptor distances to all feature points $kp_{t+1,n}$ ($n = 1, 2, \ldots, N$) in image frame I_{t+1} are calculated and ranked, and the ratio of the nearest descriptor distance to the second closest distance is calculated. The key points for which the ratio is less than a threshold are treated as a matched pair of points. The Euclidean distance is adopted to measure the similarity between two key point descriptors.

Since incorrect matching relations can reduce the accuracy of UAV pose estimation or even lead to the failure, the matched 2D-2D relations need to be selected to filter out incorrect matching relations. The RANSAC algorithm is used to calculate the homologous matrix H and to estimate the best model that maximizes the number of matched point pairs. Assuming that the relationship between the matched pairs of key points satisfying the model and the homologous matrix H is expressed as:

$$\mathbf{s}_{\mathrm{h}}\mathbf{P}_{\mathrm{uv2}} = \mathbf{H}\mathbf{P}_{\mathrm{uv1}} \tag{3.16}$$

The projection error sum for all key points is calculated as:

$$e_{\mathrm{sum}} = \sum_{i}^{n_{\mathrm{p}}} \left(\left(u_i' - \frac{h_{11}u_i + h_{12}v_i + h_{13}}{h_{31}u_i + h_{32}v_i + h_{33}} \right)^2 + \left(v_i' - \frac{h_{21}u_i + h_{22}v_i + h_{23}}{h_{31}u_i + h_{32}v_i + h_{33}} \right)^2 \right) \tag{3.17}$$

where $(u, v)^{\mathrm{T}}$ and $(u', v')^{\mathrm{T}}$ denote the pixel coordinates of the two feature points, s_{h} is the scale parameter and n_{p} is the number of matched pairs. When the projection errors reach a minimum and the number of matching pairs satisfying the model reaches a maximum, the solution is accomplished.

Then, the 2D-3D matching pairs can be obtained according to the coordinates of 3D points saved in the database. When a 3D point can be observed in different images, the point is considered stable. The co-viewing relationship is used to obtain stable 2D-3D pairs, which can improve the accuracy.

Bundle Adjustment is used to solve for the camera pose. Assuming that there are n stable 2D-3D pairs, the coordinate of the 2D point is $\mathbf{P}_{\mathrm{uv}i\_\mathrm{true}} = [X_{\mathrm{uv}i\_\mathrm{true}}, Y_{\mathrm{uv}i\_\mathrm{true}}, Z_{\mathrm{uv}i\_\mathrm{true}}]^{\mathrm{T}}$ ($i = 1, 2, \ldots, n$) and the coordinates of the 3D point is $\mathbf{P}_{si} = [X_{si}, Y_{si}, Z_{si}]^{\mathrm{T}}$ ($i = 1, 2, \ldots, n$). The projection coordinates of a spatial point in the pixel plane calculated according to the camera imaging model and the transformation matrix $\mathbf{T}_{\mathrm{sc}}$ can be expressed as $\mathbf{P}_{\mathrm{uv}i} = \mathbf{K}\mathbf{T}_{\mathrm{sc}}\mathbf{P}_{si}/s_i$. To find the optimal transformation matrix $\mathbf{T}_{\mathrm{sc}}$, calculate the sum of the reprojection error for n points. Construct the least squares problem as follows:

$$\mathbf{T}_{\mathrm{sc}}^{*} = \arg\min_{\mathbf{T}_{\mathrm{sc}}} \frac{1}{2} \sum_{i=1}^{n} \left\| \mathbf{P}_{\mathrm{uv}i\_\,\mathrm{true}} - \frac{1}{s_i} \mathbf{K} \mathbf{T}_{\mathrm{sc}} \mathbf{P}_{si} \right\|_2^2 \tag{3.18}$$

where, $\mathbf{T}_{\mathrm{sc}}^{*}$ is the transformation matrix with the smallest sum of reprojection errors.

Finally, the camera position and yaw angle are calculated. The coordinate of the camera in the camera coordinate system is $\mathbf{P}_{\mathrm{c}} = [0,\ 0,\ 0]^{\mathrm{T}}$, and the coordinate of the camera in the SfM model coordinate system can be calculated as $\mathbf{P}_{\mathrm{s}} = -\mathbf{R}_{\mathrm{sc}}^{-1}\mathbf{t}_{\mathrm{sc}} = -\mathbf{R}_{\mathrm{sc}}^{\mathrm{T}}\mathbf{t}_{\mathrm{sc}}$, according to the monocular camera imaging principle and the transformation relationship of the coordinate system. Based on the similar transformation matrix $\mathbf{T}_{\mathrm{sw}}$, the position of the camera in the apron coordinate system can be calculated as:

$$\mathbf{P}_{\mathrm{w}} = -\mathbf{R}_{\mathrm{sw}}\mathbf{R}_{\mathrm{sc}}^{\mathrm{T}}\mathbf{t}_{\mathrm{sc}} + \mathbf{t}_{\mathrm{sw}} \tag{3.19}$$

Based on the transformation relationship between the three coordinate systems, the transformation matrix $\mathbf{T}_{\mathrm{wc}}$ from the camera coordinate system to the apron coordinate system can be calculated as $\mathbf{T}_{\mathrm{wc}} = \mathbf{T}_{\mathrm{sc}}\mathbf{T}_{\mathrm{sw}}^{\mathrm{T}}$, which in turn leads to the matching rotation matrix $\mathbf{R}_{\mathrm{cw}}$. Decompose a rotation described by the rotation matrix into 3 rotations around different rotation axes: the angle of rotation about the z_{c} axis is the yaw angle θ_{yaw}, about the y_{c} axis after the above rotation is the pitch angle θ_{pitch} and about the x_{c} axis after the above rotation is the roll angle θ_{roll}. Based on the rotation matrix $\mathbf{R}_{\mathrm{cw}}$, the yaw angle can be calculated as:

$$\theta_{\mathrm{yaw}} = \arctan(r_{21}/r_{11}) \tag{3.20}$$

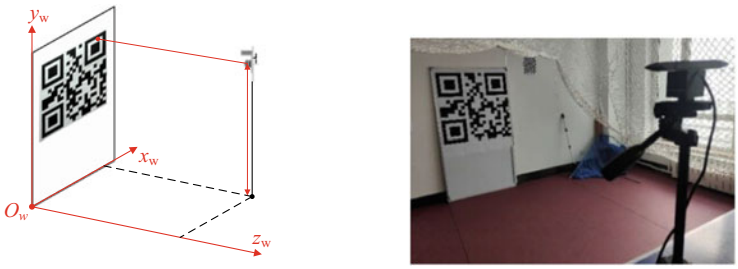
where r_{11} and r_{21} are the elements of the rotation matrix $\mathbf{R}_{\mathrm{cw}}$ at the corresponding positions. The camera position and yaw angle are now solved.

4 Implementation and Performance Analysis

4.1 Experiment Setup

The training of the rotating target detection algorithm is based on a hardware environment supported by a CPU (Intel® Xeon® Gold-5118 CPU@2.30 GHz × 12) and a GPU (Quadro P4000) and a software environment configured with Ubuntu 18.04. The model was trained for 50 rounds with the training set of DOTA-v1.0, a large dataset suitable for aerial image target detection.

To verify the performance of the UAV pose estimation method in this paper, a simulation test was conducted on the ground, and the test environment was modelled as shown in Fig. 3a, and the actual experimental scenario was shown in Fig. 3b. A board with a 1 m × 1 m QR code marker was fixed perpendicular to the ground, with the lower left corner of the board as the origin of the apron coordinate system,

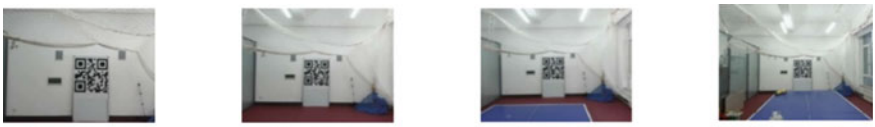


(a) Experiment scenario model (b) Actual experiment scenario

Fig. 3 Experiment environment

Table 2 Calibration results

Intra-camera parameters				Camera distortion parameters			
f_u	450.780505	c_u	344.397952	k_1	0.010261	p_1	− 0.006592
f_v	450.120812	c_v	275.155640	k_2	0.000645	p_2	0.000282



(a) Position one (b) Position two (c) Position three (d) Position four

Fig. 4 Part of the images used for reconstruction

perpendicular to the board outward as the z_w axis, and the x_w and y_w axes in the QR code marker plane, upward as the y_w axis and to the right as the x_w axis.

A calibrated monocular camera was used for data acquisition and the results of calibration are shown in Table 2. The actual position of the camera was measured using a laser rangefinder and the actual yaw angle was measured using an inclinometer.

Ensure that the camera yaw angle is 0° to take a series of images for reconstruction, as shown in Fig. 4. At different positions, the camera is rotated by 0, ± 45, ± 90 and ± 135° to capture a series of images for verification, as shown in Fig. 5, and the pose of the camera is recorded.

4.2 Performance Analysis

The model performance was evaluated using the validation set in DOTA-v1.0. The average detection accuracy of various targets is shown in Table 3.

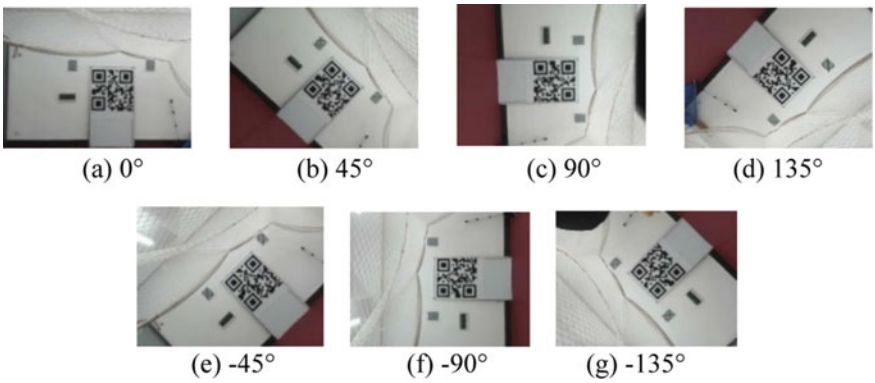


Fig. 5 Part of validation images

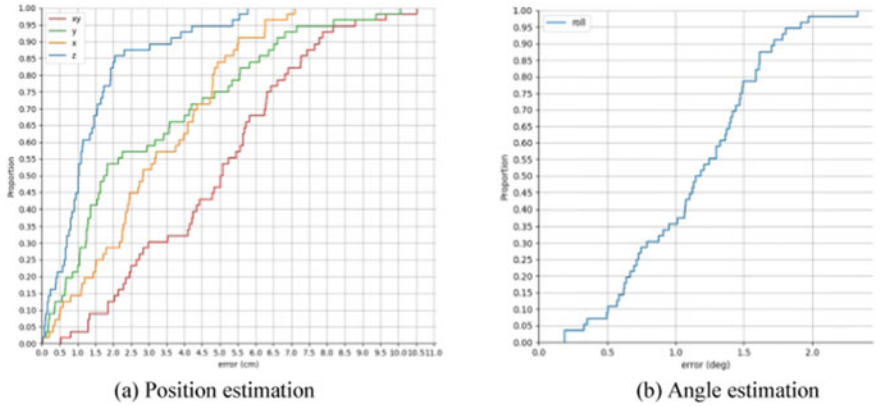
Table 3 AP accuracy for each type of target test

Category	AP	Category	AP	Category	AP	Category	AP
Plane	0.911	Small-vehicle	0.642	Basketball-court	0.600	Harbor	0.711
Baseball-diamond	0.615	Large-vehicle	0.674	Storage-tank	0.362	Swimming-pool	0.537
Bridge	0.499	Ship	0.873	Soccer-ball-field	0.378	Helicopter	0.144
Ground-track-field	0.516	Tennis-court	0.950	Roundabout	0.402	All	0.588

The model performs well on the categories more likely to be used as mobile apron carriers: aircraft, small vehicles, large vehicles and ships. Tests were conducted using actual UAV aerial images to detect vehicles in the images and the results are shown in Fig. 6, where the model can successfully detect vehicles in aerial images.

For the UAV pose estimation method we propose, the errors of the position and yaw angle of the validated image estimated by the algorithm are calculated with

Fig. 6 Test results



(a) Position estimation (b) Angle estimation

Fig. 7 Pose estimation performance

reference to the true values. The error cumulative distribution curves and angular error cumulative distribution curves are plotted as shown in Fig. 7a, b.

The maximum offset error produced on the x_w axis is approximately 7.2 cm, on the y_w axis is approximately 9.7 cm, on the z_w axis is approximately 5.8 cm and on the $O_w x_w y_w$ plane is 10.6 cm. The maximum error in yaw angle is 2.3°. The simulation experimental results show that the algorithm we proposed performs better on yaw angle estimation than the colour square feature image based attitude estimation algorithm designed by Bo et al. performs better on position estimation than the UAV pose estimation algorithm based on the AprilTag code used by Gao et al. and performs better on both position and yaw angle estimation than the UAV autonomous landing method based on the H-marker proposed by Ye et al.

5 Conclusion

In this paper, a UAV position estimation method using monocular vision based on cooperative identification and motion reconstruction was proposed. The method can be used as a general UAV precision landing technique, breaking the limitations of conventional fixed aprons and enabling safe landing of UAVs in small or restricted mobile aprons. Experimental results indicate that the method proposed in this paper can improve the accuracy of the pose estimation of UAV.

Acknowledgements This paper is supported by National Key R&D Program of China (2022YFC3801100) and National Natural Science Foundation of China (61971162).

References

1. Lee K, Kim K, Kim T, Kwon S (2011) Fuel cell system with sodium borohydride as hydrogen source for unmanned aerial vehicles. J Power Sources 196(21):9069–9075
2. Da D, Deng W, Liu X, Tang F (2018) Analysis of vision-based navigation for autonomous landing of unmanned aerial vehicle. Ord Ind Autom 37(4):23–27
3. Guo J, He D (2022) Research on autonomous landing technology for mobile platforms for agricultural drones. J Agric Mech Res 44(7):33–38
4. Cheng Z, Gao F, Li D, Zhao L (2020) Some key technologies for cooperative autonomous landing of drones/boats. Shipbuild China 61(1):156–163
5. Chen Q, Fan H, Huang C, Xu Z, Xu T, Yan L, Yin W (2022) Design of autonomous navigation scheme for UAV transmission line inspection. J Electron Test 21:64–66
6. Guo G, Liu J, Sun X (2023) Secure robust precise vehicle localization with 5G/GNSS fusion: advances and prospects. Control Decis 38(2):289–303
7. Leslie M (2019) What if the global positioning system didn't work. Engineering 5(6):985–986
8. Wang B, Tang C, Xu G, Wang Z (2020) Pose and velocity estimation algorithm for UAV in visual landing. In: 2020 39th Chinese control conference (CCC), pp 3713–3718
9. Guangxuan C, Bo H, Qiang L (2020) UAV attitude angle measurement system based on machine vision. In: 2020 IEEE 5th information technology and mechatronics engineering conference (ITOEC), Chongqing, China, pp 1552–1560
10. Gao J, Jing X, Li Y, Yuan S (2020) Method of UAV position based on cooperative two-dimension code. Mod Navig 11(1):20–25
11. Yan B, Tao C, Ye R, Yang T (2020) Research on vision-based autonomous landing of unmanned aerial vehicle. In: 2020 IEEE 3rd international conference on automation, electronics and electrical engineering (AUTEEE), Shenyang, China, pp 348–354
12. Frahm J-M, Schönberger JL (2016) Structure-from-motion revisited. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), Las Vegas, NV, pp 4104–4113

3D Wind Field Construction with Multiple Wind Profilers

Taofeng Gu, Jiamin Wang, Feng Shen, Haiyan Yue, Guangsheng Wu, Hao Wu, and Haijiang Wang

Abstract The construction of three-dimensional wind fields has important implications in various fields. Currently, the construction of three-dimensional wind fields is predominantly achieved through the integration of single or multiple Doppler radars. While this paper presented an algorithm for 3D wind field construction with multiple wind profiler radars. The algorithm considered the influence of the Earth curvature on the wind field construction results, meanwhile, it could be compatible with the wind profile radar data of different length of range bins, which improved the data utilization and the accuracy of results. This paper validated the algorithm by utilizing five radars in Guangzhou City, and it achieved favorable results.

Keywords Wind profiler radar · Construction of three-dimensional wind fields · Coordinate system transformation

1 Introduction

The three-dimensional wind field is an important meteorological product that can be used for diagnosing mesoscale weather systems, displaying the three-dimensional structure of atmospheric flow fields [1], monitoring and alerting for hazardous

T. Gu · G. Wu
Guangzhou Emergency Warning Information Release Center, Guangzhou 511430, China
e-mail: wgs@gd121.cn

J. Wang · F. Shen · H. Wu · H. Wang (✉)
Chengdu University of Information Technology, Chengdu 610225, China
e-mail: whj@cuit.edu.cn

H. Wu
e-mail: wuhao@cuit.edu.cn

H. Yue
Guangzhou Meteorological Observatory, Guangzhou 511430, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_47

weather events [2]. Additionally, there is a strong demand for wind field observations in areas such as meteorological support for large-scale events and weather support for airports.

There are various methods for constructing a three-dimensional wind field, with the commonly used approach being the inversion using single or multiple Doppler radars in a network. The algorithms often employed are based on the three-dimensional variational method (VAR). Currently, there are also many improved methods based on this approach. Wu proposed the utilization of spatial recursive filters [3], which enables the construction of variational analysis in physical space, thereby allowing for more degrees of freedom in adaptive error statistical definitions [4]. Xie and MacDonald used statistical analysis and numerical experiments to study the theoretical and numerical differences in actual three-dimensional variational analysis [5]. Additionally, Bousquet proposed the use of the Multiple-Doppler Synthesis and Continuity Adjustment Technique (MUSCAT) to efficiently solve for the three Cartesian wind speed components simultaneously, overcoming the drawbacks of iterative analysis techniques in handling airborne Doppler radar data [6]. This method has been widely applied [7–9]. These methods are relatively mature, and the constructed three-dimensional wind fields have achieved good results. However, the real-time synthesis of three-dimensional wind fields is challenging to achieve due to slow data transmission and limited scanning strategies, which restrict the real-time measurement of Doppler radar. As a result, it is difficult to apply real-time synthesized three-dimensional wind fields in operational settings.

Wind profiler radar utilizes microwave remote sensing to detect upper-level wind fields by receiving and processing the information returned by the electromagnetic beam under the influence of the vertical inhomogeneity of the atmosphere [10]. It can monitor atmospheric motion velocity, direction, and structure constants, with high vertical resolution. It is considered a new generation of high-performance atmospheric monitoring equipment [11].

Although some provinces or cities have multiple wind profiler radars at certain distances from each other, facilitating the network-based construction of large-scale wind fields [12], there are still some shortcomings. Existing algorithms ignore the impact of Earth's curvature on the results of wind field construction; many algorithms require that the range bins of radars be equal, resulting in low utilization of wind profiler radar data. The algorithm used in this paper overcomes the shortcomings. It can eliminate the impact of Earth's curvature on the results of wind field construction. And it is also compatible with wind profiler radar data of lengths of different range bins.

2 Data Preprocessing

2.1 Uniform Geographic Coordinates

The wind speed vectors measured at different heights by each wind profiler radar are referenced to the local coordinate system of that radar. This leads to different references for the wind speed vectors of each radar on the spherical surface. But the construction of a three-dimensional wind field requires a consistent coordinate system. Therefore, before the wind field construction work, it is necessary to unify the coordinate systems of each radar.

2.1.1 Specification of Coordinate System

1. Geocentric Coordinate System: With the center of the Earth as the origin, the unit vector pointing from the center of the Earth to the North Pole as the Z-axis, the direction from the center of the Earth to the intersection of the 0° longitude line and the equator as the X-axis, and the Y-axis in the direction of $X \otimes Z$.
2. Local Coordinate System of the Wind Profiler Radar Location: Define a local coordinate system at any point on the ground with the X-axis from west to east (denoted as X'), and the Y-axis from south to north (denoted as Y'). The Z-axis of the local coordinate system (denoted as Z') is exactly in the direction of the cross product of the local coordinate system X-axis and local coordinate system Y-axis, i.e., perpendicular to the ground surface at that point and pointing upwards.

2.1.2 Coordinate System Transformation

Let the unit vectors corresponding to the three axes of the geocentric coordinate system be $\vec{x}_{earth}$, $\vec{y}_{earth}$ and $\vec{z}_{earth}$. One of the radar coordinate systems in the radar network was selected as the reference coordinate system for the construction of the three-dimensional wind field, and the data from the other radars needed to be transformed according to this reference coordinate system.

For the radar selected as the reference coordinate system, its latitude and longitude were denoted as φ and θ, respectively. The unit vectors corresponding to the three axes in this coordinate system were $\vec{x}$, $\vec{y}$ and $\vec{z}$. The coordinates need to be transformed to the corresponding coordinates of the radars to be converted, which had latitudes φ' and longitudes θ'. The unit vectors corresponding to the three axes in the coordinate system of the radars to be converted are $\overrightarrow{x'}$, $\overrightarrow{y'}$ and $\overrightarrow{z'}$. The conversion relationships between these unit vectors and the unit vectors of the geocentric coordinate system were given by Eqs. (1) and (2).

$$\begin{bmatrix} \vec{x} \\ \vec{y} \\ \vec{z} \end{bmatrix} = \begin{bmatrix} -\sin\varphi & \cos\varphi & 0 \\ -\cos\varphi\sin\theta & -\sin\varphi\sin\theta & \cos\theta \\ \cos\theta\cos\varphi & \cos\theta\sin\varphi & \sin\theta \end{bmatrix} \begin{bmatrix} \vec{x}_{earth} \\ \vec{y}_{earth} \\ \vec{z}_{earth} \end{bmatrix} = B_1 \begin{bmatrix} \vec{x}_{earth} \\ \vec{y}_{earth} \\ \vec{z}_{earth} \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} \overrightarrow{x'} \\ \overrightarrow{y'} \\ \overrightarrow{z'} \end{bmatrix} = \begin{bmatrix} -\sin\varphi' & \cos\varphi' & 0 \\ -\cos\varphi'\sin\theta' & -\sin\varphi'\sin\theta' & \cos\theta' \\ \cos\theta'\cos\varphi' & \cos\theta'\sin\varphi' & \sin\theta' \end{bmatrix} \begin{bmatrix} \vec{x}_{earth} \\ \vec{y}_{earth} \\ \vec{z}_{earth} \end{bmatrix} = B_2 \begin{bmatrix} \vec{x}_{earth} \\ \vec{y}_{earth} \\ \vec{z}_{earth} \end{bmatrix} \tag{2}$$

Therefore, the transformation matrix $B_2 B_1^{-1}$ was required to convert the local coordinate system of the radar to the reference coordinate system. The coordinate system transformation process was described by Eq. (3).

$$\begin{bmatrix} \overrightarrow{x'} \\ \overrightarrow{y'} \\ \overrightarrow{z'} \end{bmatrix} = B_2 B_1^{-1} \begin{bmatrix} \vec{x} \\ \vec{y} \\ \vec{z} \end{bmatrix} \tag{3}$$

2.2 Wind Speed Correction

Taking the wind profile data measured by one of the wind profiler radars that requires wind vector correction as an example, let V' represent the wind vector at a certain range bin in the measured wind profile data. The unit vectors corresponding to the three axes of the local coordinate system at the wind profiler radar location were denoted as $\overrightarrow{x'}$, $\overrightarrow{y'}$, $\overrightarrow{z'}$. This could be expressed as Eq. (4).

$$V' = \begin{bmatrix} u' & v' & w' \end{bmatrix} \begin{bmatrix} \overrightarrow{x'} \\ \overrightarrow{y'} \\ \overrightarrow{z'} \end{bmatrix} \tag{4}$$

where $\begin{bmatrix} u' & v' & w' \end{bmatrix}$ represented the wind speed vector components to be transformed. By using Eq. (5), Eq. (6) was derived. Finally, Eq. (6) yielded the components $\begin{bmatrix} u & v & w \end{bmatrix}$ of the wind speed vector V in the reference coordinate system.

$$V' = \begin{bmatrix} u' & v' & w' \end{bmatrix} B_2 B_1^{-1} \begin{bmatrix} \vec{x} \\ \vec{y} \\ \vec{z} \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} u & v & w \end{bmatrix} = \begin{bmatrix} u' & v' & w' \end{bmatrix} B_2 B_1^{-1} \tag{6}$$

3 Construction of Three-Dimensional Wind Field

3.1 Mapping of Effective Wind Speed

It is necessary to map the valid wind speeds in the output wind profile data of each radar, ensuring that the wind speeds at any desired height above each radar are valid within a certain range. For a single wind profiler radar, taking the measured wind speed component u as an example, the corresponding fitting function is shown in Eq. (7).

$$u = \alpha_{u0} + \alpha_{u1} h \tag{7}$$

where h represented the height; α_{u0} and α_{u1} were the parameters to be solved for resampling the u component. These height values are organized into a matrix H following the format shown in Eq. (8). The vector components corresponding to the heights were $\vec{u}$, which could be expressed as Eq. (9); m represented the number of height levels in the wind profiler radar's detected profile.

$$H = \begin{bmatrix} 1 & h_1 \\ 1 & h_2 \\ \vdots & \vdots \\ 1 & h_m \end{bmatrix} \tag{8}$$

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \tag{9}$$

The fitting functions for the components and their corresponding parameters $\overrightarrow{\alpha_u}$ could be derived from Eq. (10).

$$\overrightarrow{\alpha_u} = \left(\mathrm{H}^T H\right)^{-1} \mathrm{H}^T \vec{u} \tag{10}$$

whereby, based on the definition of the height matrix H, $\overrightarrow{\alpha_u} = \begin{bmatrix} \alpha_{u0} & \alpha_{u1} \end{bmatrix}$. The wind speed components v and w could be obtained in the same way.

3.2 Wind Field Construction

After the set height grid, the wind speed components of the three vertices were linearly fitted by dividing the triangular area, and the corresponding fitting functions

were obtained before sampling to complete the construction of the 3D wind field. Taking the wind speed component u as an example, the linear function model is shown in Eq. (11).

$$u = a_u x + b_u y + c_u \tag{11}$$

$$a_u = \frac{1}{g}[u_1(y_3 - y_2) + u_2(y_1 - y_3) + u_3(y_2 - y_1)] \tag{12}$$

$$b_u = \frac{1}{g}[u_1(x_3 - x_2) + u_2(x_1 - x_3) + u_3(x_2 - x_1)] \tag{13}$$

$$c_u = \frac{1}{g}[u_1(x_3 y_2 - x_2 y_3) + u_2(x_1 y_3 - x_3 y_1) + u_3(x_2 y_1 - x_1 y_2)] \tag{14}$$

where x_i $(i = 1, 2, 3)$ and y_i $(i = 1, 2, 3)$ were the relative coordinates of the three radars; u_i $(i = 1, 2, 3)$ was the wind speed component measured corresponding at position (x_i, y_i); and $g = x_1(y_3 - y_2) + x_2(y_1 - y_3) + x_3(y_2 - y_1)$. Similarly, coefficients for the fitting equations of v and w components could be obtained.

4 Results

Figure 1 demonstrates the effectiveness of the proposed algorithm in mapping the horizontal wind component data from the wind profiler radar located in Conghua District, Guangdong Province. After mapping the wind components using the effective wind speed mapping method and sampling with a height grid of 1 m, it could be observed that the sampled wind speed components align with the original wind profile trends. Furthermore, the extrapolation of wind speed components beyond the range of the actual measured wind profile data, particularly above 5500 m, also yielded accurate results.

This paper conducted experiments using five wind profiler radars located in different districts of Guangzhou City, including Conghua District, Nansha District, Zengcheng District, Huadu District, and Huangpu District. Figure 2 shows the 3D wind field above the entire city at heights of 100–3000 m. The height layers involved in the construction of the 3D wind field are spaced at intervals of 500 m, and the horizontal grid spacing is 2000 m.

In addition, if the horizontal grid was divided more finely, more details could be seen in the constructed 3D wind field. As shown in Fig. 3, the figure shows a selected local 3D wind field. The interval between the height grids in the 3D wind field construction in the figure is 50 m, while the horizontal grid spacing is 100 m. It can be seen from the figure that the specific wind speed and direction are consistent with the real wind field properties, and the constructed 3D wind field results have high credibility.

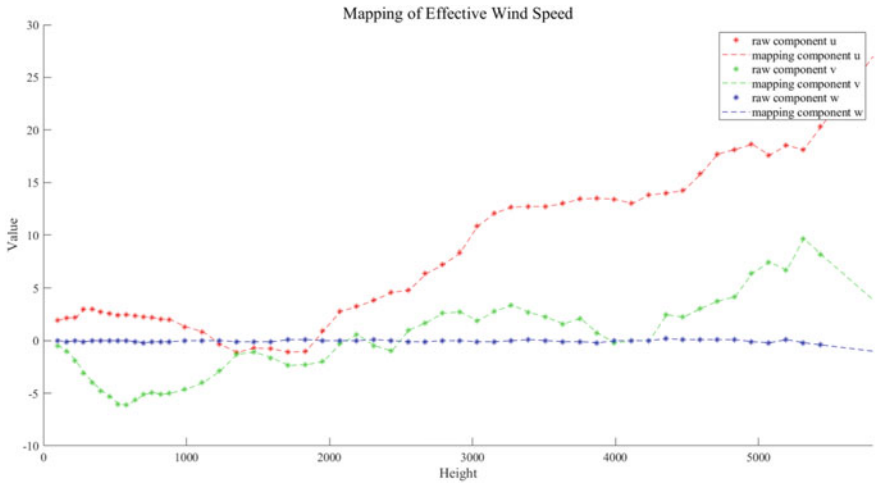


Fig. 1 Results of effective wind speed mapping

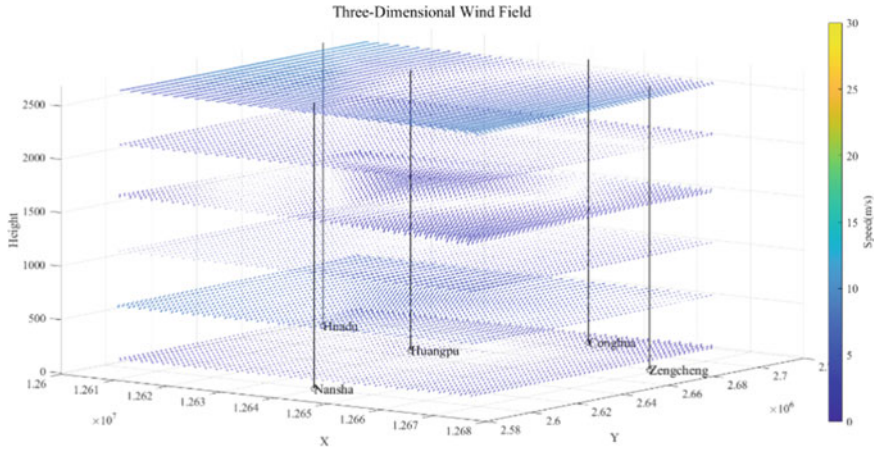


Fig. 2 Results of wind field construction over the entire city

5 Conclusions

The method proposed in this paper can effectively construct a 3D wind field in the coverage area of wind profiler radars. It can effectively overcome errors caused by Earth curvature and improve the utilization rate of wind profiler radar data through effective wind speed mapping. It solves previous restrictions on constructing 3D wind fields using wind profiler radars that require differences in geographical elevation where radars are located and also reduces errors between constructed and real wind

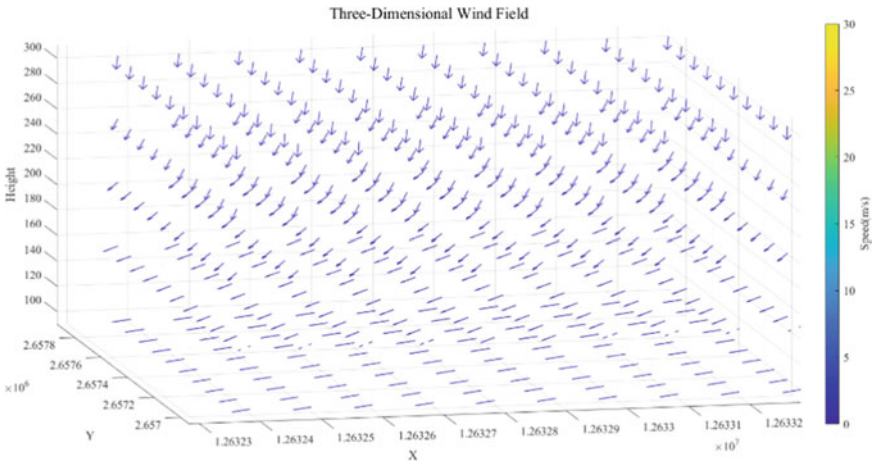


Fig. 3 The local wind field results after dividing the horizontal and vertical grids more finely

fields caused by earth curvature. Future research can further expand and optimize this method to better meet practical needs and promote development in related fields.

Acknowledgements This study was supported by the Guangzhou Science and Technology Bureau (No. 202206010016).

References

1. Liou YC, Chang YJ (2009) A variational multiple-Doppler radar three-dimensional wind synthesis method and its impacts on thermodynamic retrieval. Mon Weather Rev 137(11):3992–4010
2. Bousquet O. 6B.2 Dynamical and microphysical properties of high impact orographic mesoscale convective systems from high resolution operational multiple-Doppler and polarimetric radar data
3. Gao J, Xue M, Brewster K, Droegemeier KK (2004) A three-dimensional variational data analysis method with recursive filter for Doppler radars. J Atmos Oceanic Technol 21(3):457–469
4. Wu WS, Purser RJ, Parrish DF (2002) Three-dimensional variational analysis with spatially inhomogeneous covariances. Mon Weather Rev 130(12):2905–2916
5. Xie Y, MacDonald AE (2012) Selection of momentum variables for a three-dimensional variational analysis. Pure Appl Geophys 169(3):335–351
6. Bousquet O, Chong M (1998) A multiple-Doppler synthesis and continuity adjustment technique (MUSCAT) to recover wind components from Doppler radar measurements. J Atmos Oceanic Technol 15(2):343–359
7. Bousquet O (2012) Retrieving high resolution 3-D wind vector fields from operational radar networks. In: Bech J (ed) Doppler radar observations—weather radar, wind profiler, ionospheric radar, and other advanced applications [Internet]. InTech. [Cited 10 May 2023].

Available from: http://www.intechopen.com/books/doppler-radar-observations-weather-radar-wind-profiler-ionospheric-radar-and-other-advanced-applications/retrieving-high-resolution-3-d-wind-vector-fields-from-operational-radar-networks
8. Park SG, Lee DK (2009) Retrieval of high-resolution wind fields over the southern Korean Peninsula using the Doppler weather radar network. Weather Forecast 24(1):87–103
9. Goh YK, Holt AR. Analysis of three-dimensional wind fields from two operational Doppler radars
10. Lehmann V, Brown W (2021) Radar wind profiler. In: Foken T (ed) Springer handbook of atmospheric measurements [Internet]. Springer International Publishing, Cham, pp 901–933. [Cited 22 May 2023]. Available from: https://doi.org/10.1007/978-3-030-52171-4_31
11. Rogers RR, Ecklund WL, Carter DA, Gage KS, Ethier SA (1993) Research applications of a boundary-layer wind profiler. Bull Am Meteor Soc 74(4):567–580
12. Liu B, Guo J, Gong W, Shi L, Zhang Y, Ma Y (2020) Characteristics and performance of wind profiles as observed by the radar wind profiler network of China. Atmos Meas Tech 13(8):4589–4600

Research on Non-reference Text Image Blur Assessment System

Xin Li, Di Lin, Zixu Tao, Jikang Mo, Zongbo Hao, and Peirui Wang

Abstract The non-reference image blur assessment (NR-TIBA) system is significant in text image processing. Due to the high cost of subjective evaluation and the unavailability of reference images in the natural environment, an objective NR-TIBA system is essential. Quantitatively, some traditional methods are affected by the richness of image content, while deep learning methods are too expensive to label; qualitatively, research on text image blur evaluation is scarce. This paper proposes a simple evaluation system for quantitative analysis using deep learning qualitative analysis and the combination of the spatial domain and frequency domain. It has good quantitative performance and only requires a small amount of complex labeling data. Experiments show that the proposed system performs well and can be further applied.

Keywords Non-reference assessment · Text images · Blur assessment system

1 Introduction

Text image blur assessment (TIBA) is widely used in text images, such as text detection, text recognition, etc. The TIBA system consists of qualitative analysis and quantitative analysis. Qualitative analysis can reduce the input of low-quality images. Quantitative analysis can determine the model for processing text images and prevent lengthy text image processing pipelines. Therefore, a sound TIBA system can reduce downstream task calculations, saving costs and improving overall work efficiency. Obviously, human beings are more accurate in subjective evaluation of the blurriness of text images, which can be evaluated by Mean Subjective Score (MOS) or Mean Subjective Score Difference (DMOS), but the considerable workload and evaluation cost make this scheme challenging to implement. Therefore, in practical applications, an objective automated assessment system is required.

Generally, IBAs can be divided into three categories according to their dependence on the reference image: full-reference IBA (FR-IBA), reduced-reference IBA

X. Li · D. Lin (✉) · Z. Tao · J. Mo · Z. Hao · P. Wang
University of Electronic Science and Technology of China, Chengdu, Sichuan, China
e-mail: lindi@uestc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_48

(RR-IBA), and NR-IBA. The FR-IBA and RR-IBA have performed well, but the reference image is often unavailable in the natural environment. Therefore, NR-IBA has become a hot research topic in recent years. In recent years, the performance of NR-IBA based on deep learning has been improved, but the shortcomings are also significant. This type of method is highly dependent on accurate MOS or DMOS. Traditional NR-IBA methods are mainly based on the spatial domain and frequency domains, their computational costs are small, and their effects are within acceptable limits. However, they also have an identical drawback: the comparison results of images are strongly influenced by the image content.

In this paper, we design an assessment system that uses a deep learning approach to qualitative analysis of text image blurring and improves on the traditional frequency and spatial domain methods for quantitative analysis, reducing the impact of image content volume and providing good performance in both qualitative and quantitative image blurring while not requiring large amounts of complex annotation data.

2 Related Work

2.1 Traditional Methods

The traditional methods are mainly based on spatial and frequency domain methods.

In the spatial domain approach, more research has focused on using image edge information for IBA. The sharper the image, the sharper its content edges; the more blurred the image, the smoother the edges will become, which is in line with the human eye's intuitive visual perception of blurred images. Therefore, edge detection is an effective method for IBA. Edge detection of images is generally achieved by calculating the image gradient value. Commonly used gradient functions are the Brenner gradient function, the Tenengrad gradient function, and the Laplacian gradient function [1]. Brenner is relatively simple and only examines the grey scale difference between the pixel point being judged and one of its neighbors, but the edge extraction results are general. The Tenengrad function, which uses the Sobel operator to extract the horizontal and vertical gradients, respectively, has a particular ability to suppress noise but does not strictly simulate the physiological characteristics of human vision, so the edge extraction ability is limited. The Laplacian gradient function uses the Laplacian operator instead of the Sobel operator to improve the edge extraction ability but will be affected by noise. Many scholars have adopted the above operators for edge detection due to their real-time computational performance. However, all these purely spatial domain methods have severe drawbacks, and the assessed value may vary with the number of objects within the image. See Table 3 for the scores of Fig. 1.

The frequency domain commonly uses the discrete Fourier transform (DFT), discrete cosine transform (DCT) methods, and wavelet transform. Chetouani et al. [2] used the DFT to evaluate the degree of additive blurring by adding a blur. Saad

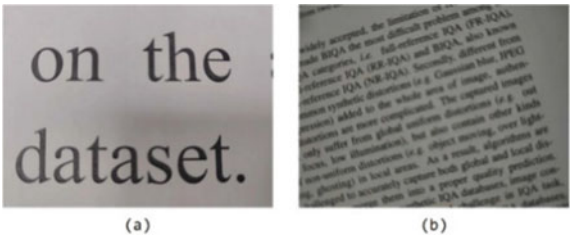


Fig. 1 **a** Little content but clear, **b** much content but blurred

et al. [3] used statistical features of the DCT to estimate image quality. Tong et al. [4] combined wavelet transform and energy [5, 6] methods to evaluate image sharpness. Gvozden et al. [7] calculated the mean root square of the image while using the diagonal wavelet coefficients for ranking and weighting to obtain the evaluation results. This type of spatial domain evaluation algorithm has higher accuracy and robustness but also suffers from the problem of high scores for high content.

2.2 Deep Learning-Based Methods

TIBA is a subproblem of image quality assessment (IQA). Kang et al. [8] accurately evaluated image quality using a 5-layer CNN, which performed the best performance at that time on the LIVE dataset. Gu et al. [9] designed a 3-layer DNN to evaluate the quality of reference-free images. Bosse et al. [10] modified the VGG network, which showed good generalization on publicly available datasets. Liu et al. [11] proposed the RankIQA model, which surpassed the partial full-reference approach. Su et al. [12] proposed a hyper-network-based IQA model to improve the evaluation of field models. Using deep meta-learning, Zhu et al. [13] achieved good generalization results and evaluation accuracy on both publicly available datasets. The deep learning (DL) approach achieved good results in quantitative analysis but still relied on a large amount of labeled data.

3 Methods

In practical applications, image blurring is usually analyzed qualitatively first and then quantitatively on this basis. In our proposed assessment system, we use a neural network to construct a qualitative text blur judging network and then use a modified traditional method for quantitative assessment.

3.1 Qualitative Assessment Method

Current research on deep learning models in IBA has focused on scoring, i.e., quantitative analysis. However, determining whether an image is blurred is a rather important issue. Although it is a simple binary problem, it still has an important guiding role in the TIBA system. ResNet has shown promising results in classification tasks with the introduction of the concept of Residual Block [14], which allows "skipping" some of the layers, allowing the network to be deeper and better trained.

The qualitative assessment of text images is a binary classification problem, so we only set two classes for the network: high-blur and low-blur. For less running time and training cost, we only use the simplest ResNet18 to train and achieve satisfactory accuracy. In theory, using a deeper ResNet (e.g., ResNet50, etc.) would yield a more accurate model, but accordingly, the model would be much more massive.

3.2 Quantitative Assessment Method

To avoid using complex labeled data, we used a combination of image spatial domain and frequency domain methods to evaluate TIBA quantitatively. For less running time, we chose to use the LAP operator as the basis for edge detection.

The LAP operator is isotropic and can extract edges in any direction, so only once is edge detection needed to extract edges in X and Y directions. The two-dimensional convolution form of the LAP operator is defined as follows:

$$g(i, j) = \sum_{r=-k}^{k} \sum_{s=-l}^{l} f(i - r, j - s)H(r, s), \;\; i, j = 0, 1, 2, \sim N - 1 \tag{1}$$

where f denotes the original input, r and s are denoted offsets, and $H(r, s)$ denotes an operator representing the environment in question. In particular, when both k and l values are taken as 1, $H(r, s)$ can be specialized to the following form, which is the most commonly used LAP operator for 2D images.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{2}$$

For digital images, the definition of an image edge based on the LAP operator can be simplified to the following form:

$$edge_{space} = D(f) = \sum_{y} \sum_{x} |g(x, y)| \quad (g(x, y) > T) \tag{3}$$

where T is the given edge detection threshold and $g(x, y)$ is the convolution of the LAP operator at pixel point (x, y). However, directly using this method for assessment will be an anomaly in that the score increases with the increase of image content. Therefore, the above formula needs to be further improved.

We do further calculations using the Fourier transform. The formula for the Fourier positive transform of a two-dimensional discrete function $f(x, y)$ is as follows:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}, \quad u \in [0, M-1]; \quad v \in [0, N-1] \tag{4}$$

where u is the horizontal value, v is the vertical value, and $f(x, y)$ is the point in the two-dimensional image. To improve the computational efficiency of the above equation, we used the fast Fourier transform (FFT) to extract the frequency domain information of the image. We removed only a small portion of the low-frequency information and we found the remainder can be understood as a measure of the amount of information in the image. Therefore, we summed the obtained high-frequency components as a correction to the results of the spatial domain evaluation.

$$Sharpness_{img} = \frac{edge_{space}}{F(u, v)} \tag{5}$$

This method uses an average-like concept and calculates the edge mean of the unit image content. Through experiments, our method corrects the bias due to the amount of image information to a certain extent. Using this value for quantitative analysis of image assessment is more effective.

4 Experiments

We use the BMVC text dataset [15] for training, where the sharp images are used as positive samples and the blurred images are negative samples. These data will be randomly cropped to 224×224 sizes and fed into the model, where the learning rate is set to $1e{-}3$, the weight attenuation is set to $1e{-}4$ to prevent overfitting, and the momentum is set to 0.9 for weight optimization. The images in Fig. 2 were fed into the qualitative model, giving discriminatory results that matched those of the human eye.

We processed the sharp images from the BMVC text, adding different degrees of blur to the random images, and the accuracy obtained is shown in Table 1. The qualitative model performed well in all these cases.

We scored parts of the BMVC dataset in quantitative experiments to obtain text image MOS. The Pearson linear correlation coefficient (PLCC) and the Spearman rank-order correlation coefficient (SRCC) is used to measure the monotonicity of the predictions and accuracy. Both criteria range from 0 to 1, with higher values

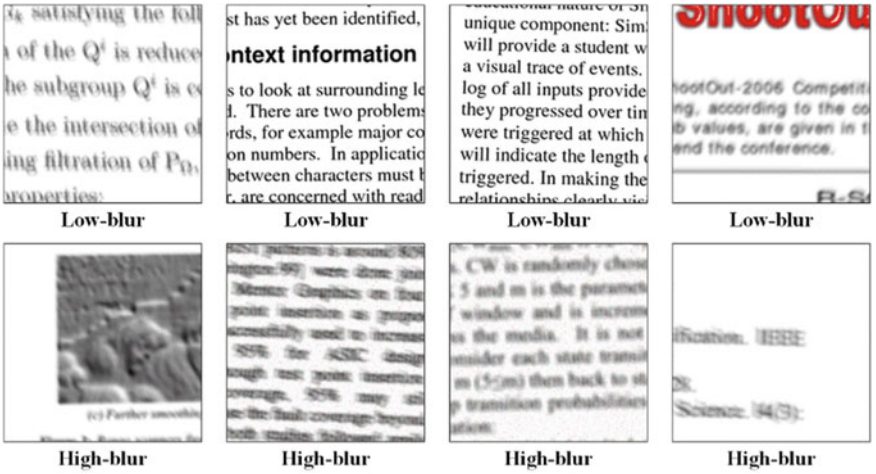


Fig. 2 Text images for qualitative testing

Table 1 Accuracy in data mixed with varying degrees of blur

Blur	Accuracy (%)
Low	0.798
Mid	0.814
High	0.835
Mixed	0.803

indicating better performance. We have used our method to compare with several mainstream traditional methods. See Table 2 for the quantitative results. Most of the metrics of our method are higher than those of the pure spatial domain method. Furthermore, we also compare with some deep learning based IQA methods, such as HyperIQA [12] and MetaIQA [13]. The gap between our method and the DL method is tiny. However, our method does not require labeled data, which saves labeling and training costs, and has certain universality, which other applications such as contour detection can apply.

Table 2 SRCC and PLCC by different methods

Method	SRCC	PLCC
FFT	0.117	0.208
Tenengrad	0.731	0.621
Laplacian	0.714	0.606
HyperIQA (DL)	0.766	0.685
MetaIQA (DL)	0.739	0.702
Ours	0.756	0.637

Table 3 Test results for images of Fig. 1 in real-life scenarios

Image	Blurredness ground truth	LAP score	Proposed system	
			Qualitative	Quantitative
a	Sharp	18.2	Low-blur	34
b	Blur	27.9	High-blur	30

The problem of blurred text images scoring higher than sharp images as the content increases have also been improved. We prepared 500 pairs of sharp-blurred text image pairs and used the LAP algorithm (to make the comparison more intuitive, we standardized the scores to the interval (0, 50)) and our algorithm to calculate the scores. Figure 3 shows a scatter plot of the quantification scores, where the abscissa is the serial number of the text image, and the ordinate is the quantitative score. Ideally, the scatterplot of blurred images should be distinguishable from sharp images. The scatter points in the plot of the proposed method are denser, and the demarcation line is more apparent, fully indicating that the proposed algorithm is less affected by the amount of text image content.

The results obtained by applying the system to realistic pictures (Fig. 1) are shown in Table 3. Clearly, (a) is more apparent than (b), and our assessment system matches the subjective opinion of the human eye in qualitative and quantitative terms.

Figure 4 is a flowchart of the proposed system. We do the qualitative assessment first, and the high-blur image can be discarded or deblurred according to requirements. Then we can perform the quantitative assessment on low-blur images and select different applications according to the score. Taking OCR as an example, images with high scores use quick OCR with less resource usage, while images with

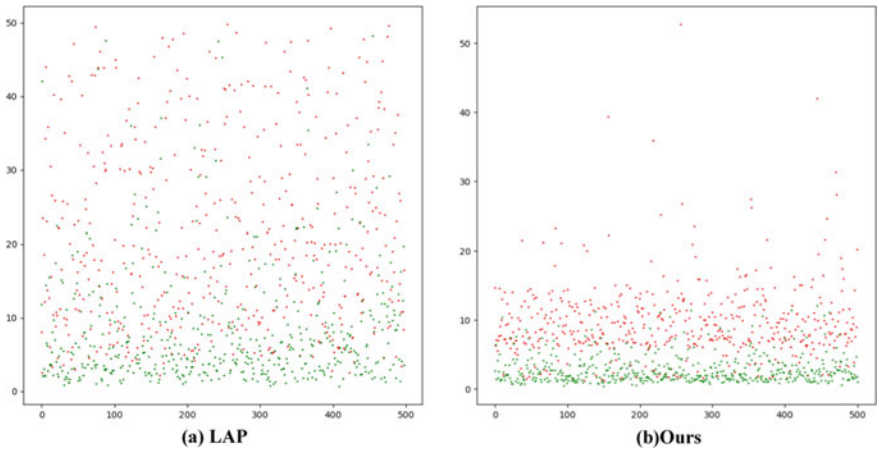


Fig. 3 Scatterplot of the LAP method and our method

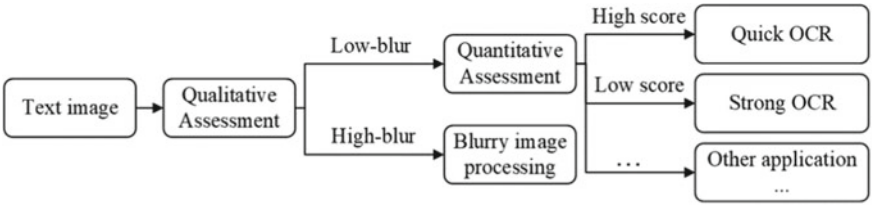


Fig. 4 Flowchart of the proposed system

low scores may be blurred so we use some robust OCR methods with a higher recognition rate. Our TIBA system can save computing resources and further improve the recognition rate. Of course, other related fields can also apply the proposed system.

5 Conclusion

In this paper, we proposed a simple and lightweight FR-TIBA system using a quantitative method for image ambiguity that combines the space and frequency domains with a qualitative image ambiguity that is easy to train and has a high accuracy rate. The system has been evaluated in experiments with good results. The system is a meager cost, simple to use, and can guide research and development programs with related needs. In addition, using a deeper model and increasing the training data set size could yield better results.

Acknowledgements Partially funded by Natural Science Foundation of Sichuan Province (2023NSFSC0479) and partially funded by Grant SCITLAB-20005 of Intelligent Terminal Key Laboratory of Sichuan Province.

References

1. Osibote OA, Dendere R, Krishnan S, Douglas TS (2010) Automated focusing in bright-field microscopy for tuberculosis detection. J Microsc 240(2):155–163
2. Chetouani A, Beghdadi A, Deriche M (2009) A new reference-free image quality index for blur estimation in the frequency domain. In: 2009 IEEE international symposium on signal processing and information technology (ISSPIT), Dec 2009. IEEE, pp 155–159
3. Saad MA, Bovik AC, Charrier C (2012) Blind image quality assessment: a natural scene statistics approach in the DCT domain. IEEE Trans Image Process 21(8):3339–3352
4. Tong H, Li M, Zhang H, Zhang C (2004) Blur detection for digital images using wavelet transform. In: 2004 IEEE international conference on multimedia and expo (ICME) (IEEE Cat. No. 04TH8763), June 2004, vol 1. IEEE, pp 17–20
5. Wu W, Hu S, Lin D, Wu G (2022) Reliable resource allocation with RF fingerprinting authentication in secure IoT networks. Sci China Inf Sci 65(7):170304

6. Lin D, Hu S, Wu W, Wu G (2023) Few-shot RF fingerprinting recognition for secure satellite remote sensing and image processing. Sci China Inf Sci. https://doi.org/10.1007/s11432-022-3672-7
7. Gvozden G, Grgic S, Grgic M (2018) Blind image sharpness assessment based on local contrast map statistics. J Vis Commun Image Represent 50:145–158
8. Kang L, Ye P, Li Y, Doermann D (2014) Convolutional neural networks for no-reference image quality assessment. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1733–1740
9. Gu K, Zhai G, Yang X, Zhang W (2014) Deep learning network for blind image quality assessment. In: 2014 IEEE international conference on image processing (ICIP), Oct 2014. IEEE, pp 511–515
10. Bosse S, Maniry D, Müller KR, Wiegand T, Samek W (2017) Deep neural networks for no-reference and full-reference image quality assessment. IEEE Trans Image Process 27(1):206–219
11. Liu X, Van De Weijer J, Bagdanov AD (2017) RankIQA: learning from rankings for no-reference image quality assessment. In: Proceedings of the IEEE international conference on computer vision, pp 1040–1049
12. Su S, Yan Q, Zhu Y, Zhang C, Ge X, Sun J, Zhang Y (2020) Blindly assess image quality in the wild guided by a self-adaptive hyper network. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 3667–3676
13. Zhu H, Li L, Wu J, Dong W, Shi G (2020) MetaIQA: deep meta-learning for no-reference image quality assessment. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 14143–14152
14. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
15. Hradiš M, Kotera J, Zemcık P, Šroubek F (2015) Convolutional neural networks for direct text deblurring. In: Proceedings of BMVC, Sept 2015, vol 10, no 2

Research on IR-RGB Image Fusion Method Based on Target Perception and Hue Correction

Qianjun Jiang, Di Lin, Zongbo Hao, He Chang, and Xin Li

Abstract In this paper, a hue-corrected IR-RGB fusion method is proposed based on object perception. We use the salient object detection network to detect the original IR image to obtain the salient regions in the image and then use the object bounding box in the dataset label to remove the noise information and retain the object area of interest. In order to implement different fusion schemes for the target area and the background area in the model, this paper trains a generator to synthesize the target fusion image, trains the target discriminator to identify the pixel intensity and hue information of the target area.And trains the background discriminator to identify Gradient information for the background. To improve the performance of downstream object detection tasks, an object detection network is added to the generative model, and an object detection loss is used to guide image fusion.

Keywords Image fusion · Generative adversarial network · Target perception · Infrared and visible light

1 Introduction

Single-modal images often cannot meet the needs of many complex scenes. In contrast, multi-modal images, such as IR-RGB images, have better performance in scenes with complex backgrounds and insufficient light because the advantages of the two modes can be combined to complement each other. Infrared light is more prominent for the imaging of heat source targets, less dependent on illumination and has high anti-interference, but because of its longer wavelength, the resolution is lower, and there is greater noise. On the other hand, visible light imaging has high resolution and rich texture details but has weak anti-interference ability and is easily affected by environmental factors [1]. Therefore, combining the advantages of the two modal

Q. Jiang · D. Lin (✉) · Z. Hao · H. Chang · X. Li
University of Electronic Science and Technology of China, Chengdu, Sichuan, China
e-mail: lindi@uestc.edu.cn

Q. Jiang
URL: https://www.uestc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_49

images, the research on generating IR-RGB fusion images has practical application significance [2–4].

It is generally believed that IR images have better performance than RGB images in highlighting heat source targets, but texture details and colors in RGB images is more than IR images. Therefore, in the target area, the pixel intensity of the IR image and the color of the RGB images are preserved to improve the target recognition performance, and the texture details and color fidelity of the RGB image are preserved in the background area to improve the human visual experience and generate multi-modal The fusion of images of advantages is the work of this paper.

2 Related Work

2.1 Traditional Image Fusion Methods

Traditional image fusion methods are relatively mature. It mainly includes multi-scale decomposition [5], sparse representation [6], subspace learning [7], image-based saliency [8], etc. The shortcomings of traditional image fusion methods are also obvious: (1) Feature extraction needs to be manually designed, which is complex and challenging to adapt to images in various scenarios. (2) Fusion rules need to be manually designed, which is complicated and cannot make good use of various effective information, and cannot adapt to different modes.

2.2 Image Fusion Method Based on Deep Learning

Deep learning can primarily address the limitations of traditional methods. The convolutional neural network has a good performance in image feature extraction, and many studies have proposed end-to-end fusion algorithms based on CNN [9]. Ma et al. apply masks to image fusion, enabling the network to focus on different regions [10]. The fusion scheme proposed by Xu et al. can unify multiple fusion tasks into one model [11]. Sun et al. utilize an attention mechanism and an object detection loss to guide image fusion [12]. Tang et al. introduced illumination perception into the image fusion strategy, enabling image fusion to adapt to different illumination conditions [13]. A recent study utilizes a GAN network to fuse images and uses a two-layer optimization scheme to train the network [14].

3 Methods

This section details an image fusion method based on object awareness and hue correction presented in this paper. Begin with the model structure, the role of each part of the model, and the relationship between each part are elaborated. Then starting with each part, explain the loss function and principle of each part in detail and gradually obtain the overall loss function.

Figure 1 shows the whole process of the method in this paper. The whole model is mainly composed of two parts, the fusion module, and the detection module.

Object detection module. In order to improve the performance of downstream tasks, such as target detection, the target detection network can be added to the training process of the model to guide the fusion of images to achieve better target detection performance. The target detection network used in this paper is YOLOV5s. The loss for object detection is as follows:

$$L_{det} = L_{confidence} + L_{cls} + L_{box} \tag{1}$$

where L_{det} is the object detection loss, $L_{confidence}$ is the confidence loss, $L_{cls}$is the classification loss, and L_{box} is the bounding box loss.

Image fusion module. The fused network uses a generative adversarial network model. It consists of a fused image generator and two discriminators. We generally believe that IR images have more obvious objects, while RGB images have more texture details. Therefore, this paper believes that a good fusion image should have a significant target, as well as rich texture details and color information. Two discriminators perform object and background discrimination on the fused image, respectively.

Fusion Image Generator. The quality of image fusion can be considered from several aspects, structural similarity, pixel intensity, and hue consistency. Structural Similarity (SSIM) measures how similar two images are. Compared with indicators

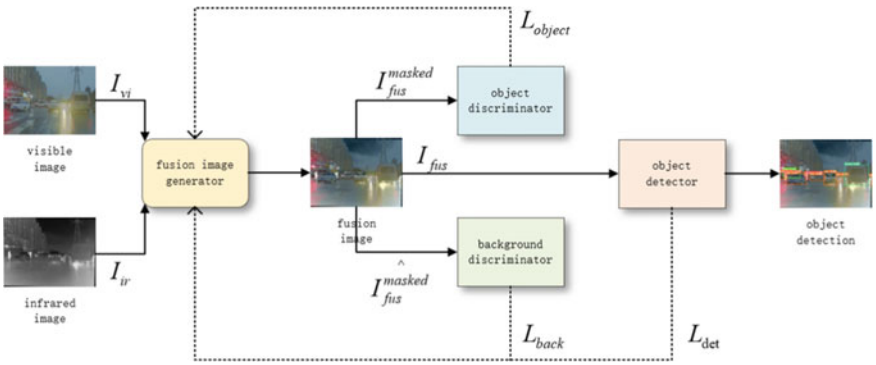


Fig. 1 Model architecture

such as peak signal-to-noise ratio (PSNR), SSIM can better express the structural similarity of images, and is more in line with human visual characteristics. We refer to the HSV color space model, propose hue consistency, and introduce hue loss as a correction for image fusion. Therefore, this paper uses SSIM, pixel intensity, and hue consistency as constraints for the fused image generator. The loss of the generator is defined as follow:

$$L_{gen} = L_{struct} + L_{pixel} + L_{hue} \tag{2}$$

where L_{gen} is the generation loss, L_{struct} is the structure loss, L_{pixel} is the pixel loss, and L_{hue} is the hue loss.

L_{struct} is defined as follows:

$$L_{struct} = [1 - SSIM(I_{fus}, I_{vi})] + [1 - SSIM(I_{fus}, I_{ir})] \tag{3}$$

where $SSIM$ computes the structural similarity of two images, I_{fus} is the fusion image, $I_{vi}$is the RGB image, and I_{ir} is the IR image. SSIM is a value with a value range of [0, 1], and the closer 1-SSIM is to 1, the more different two images are.

L_{pixel} is defined as follows:

$$L_{pixel} = ||I_{fus} - I_{vi}|| + ||I_{fus} - I_{ir}|| \tag{4}$$

L_{hue} is defined as follows:

$$L_{hue} = MSE(Hue(I_{fus}), Hue(I_{vi})) \cdot \lambda_1 \tag{5}$$

Among them, MSE calculates the mean square error of two images, Hue is the value of the hue channel (hue) in the HSV color space, and λ_1 is a hyperparameter. By adjusting the proportion of the hue loss in the entire generator loss, we can find to an appropriate value for both pixel intensity and hue.

Target discriminator. This paper believes that in the target detection of fused images, more attention should be paid to the pixel value of the target. As shown in the Fig. 2, in complex scenes such as insufficient light and smog occlusion, objects such as people are more prominent in infrared than in visible light. Therefore, The pixel intensity of the target area in the IR image should be preserved as much as possible. The role of the object discriminator is to identify the pixels and colors of the object area in the fused image.

It can be seen from the figure that if only the pixel value of the target area is considered, the fused image will lose color and be distorted in the target area. To address this issue, this paper introduces a hue correction loss to constrain the object discriminator so that the resulting fused image maintains hue consistency with the original RGB image.

The target discriminator loss is defined as follows:

$$L_{object} = E_{x \sim \tilde{p}(\mathcal{R}(I_{ir}^{masked}))}[D(x)] - E_{\tilde{x} \sim \tilde{p}(\mathcal{R}(I_{fus}^{masked}))}[D(\tilde{x})] \tag{6}$$

Fig. 2 **1** Left: IR image **2** middle: RGB image **3** right: fusion image

$$\mathcal{R}(I^{masked}) = I_{ir}^{masked} + Hue(I_{vi}^{masked}) \cdot \lambda_2 \tag{7}$$

$$\mathcal{R}(I_{fus}^{masked}) = I_{fus}^{masked} + Hue(I_{fus}^{masked}) \cdot \lambda_2 \tag{8}$$

Where L_{object} is the object discriminator loss, I_{fus}^{masked} is the object region of the fused image, and I_{ir}^{masked} is the object region of the IR image. I_{fus}^{masked} and I_{ir}^{masked} are obtained by multiplying I_{fus} and I_{ir} by mask, respectively, and the mask will be introduced in detail later. $Hue(I_{fus}^{masked})$ is the hue value of the target region of the fused image, and $Hue()$ calculates the hue value of the image. λ_2 is a hyperparameter that can be used to adjust the proportion of hue loss in the overall target discriminator loss.

Background discriminator. A good fused image should have both the advantages and disadvantages of multi-modal images. While ensuring the pixel intensity and tone consistency of the target, it is also necessary to ensure that the texture details of the background are not lost. The role of the background discriminator is to identify the background region of the fused image. In general, the gradient information of an image can be used to describe the texture details of the image, so the background discriminator loss is defined as follows:

$$L_{back} = E_{x\sim\tilde{p}(\hat{\mathcal{R}}(\hat{I}_{vi}^{masked}))}[D(x)] - E_{\tilde{x}\sim\tilde{p}(\hat{\mathcal{R}}(\hat{I}_{fus}^{masked}))}[D(\tilde{x})] \tag{9}$$

$$\hat{\mathcal{R}}(\hat{I}_{vi}^{masket}) = grad(\hat{I}_{vi}^{masked}) \tag{10}$$

$$\hat{\mathcal{R}}(\hat{I}_{fus}^{masked}) = grad(\hat{\mathcal{R}}(I_{fus}^{masked})) \tag{11}$$

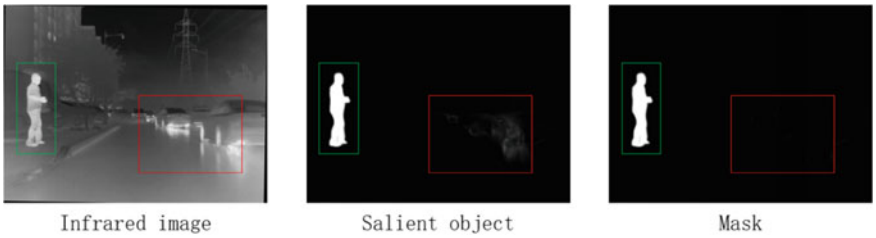


Fig. 3 The green box is the target area, and the red box is the interference information

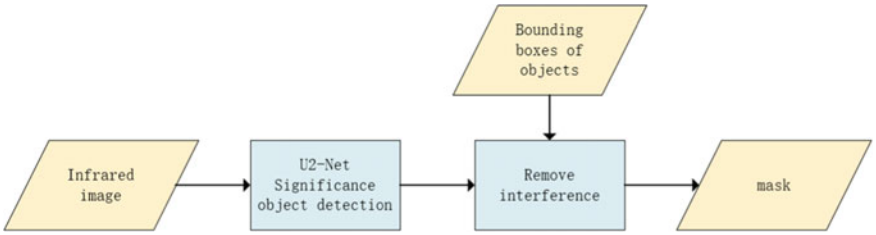


Fig. 4 Mask obtaining process

Where L_{back} is the background discriminator loss, $grad$ is the gradient value of the image, $\hat{I}_{fus}^{masked}$ represents the background region of the fused image, and $\hat{I}_{vi}^{masket}$ represents the background region of the RGB image (Fig. 3).

Image mask. In order to obtain the target area and background area in the image, this paper uses the saliency detection network U2-Net to preprocess the original IR image to obtain the mask information of the target area. The specific method is, firstly, using U2-Net to perform saliency detection on the IR images in the training set to obtain the saliency regions of the IR images. Since heat source targets such as people have high salience in IR images, the area where the target is located can be obtained through saliency detection. In order to further reduce the interference of other heat sources (such as automobile exhaust, etc.) on saliency detection. This paper proposes to use the prior knowledge, such as the object detection box of the dataset to eliminate the interference information in the saliency detection results, and only retain the useful object area information. The image mask acquisition process is as follows (Fig. 4).

The mask is used as follows:

$$I_{ir}^{masked} = I_{ir} \odot mask \tag{12}$$

$$I_{fus}^{masked} = I_{fus} \odot mask \tag{13}$$

$$\hat{I}_{fus}^{masked} = \hat{I}_{fus} \odot (1 - mask) \tag{14}$$

$$\hat{I}_{vi}^{masked} = \hat{I}_{vi} \odot (1 - mask) \tag{15}$$

Where $mask$ is the mask and $\odot$ is the pointwise multiplication.

In summary, the loss of image fusion is as follows:

$$L_{fus} = L_{gen} + L_{dis} \tag{16}$$

The total loss of the model is as follows:

$$L_{tot} = \alpha L_{fus} + \beta L_{det} \tag{17}$$

Where α and β are hyperparameters that we can use them to control the weight of fusion loss and object detection loss on image fusion, respectively.

4 Experiments

This paper uses two multimodal datasets, LLVIP and M3FD datasets. Among them, the M3FD dataset covers a wide range of scenes of various environments, lighting, seasons, and weather, including 4200 IR-RGB image pairs. LLVIP provides 15488 image pairs under low-light conditions. This data set is very suitable for verifying the effectiveness of image fusion algorithms. In this paper, the pictures in the data set are cut into fixed-size pictures by cropping to train the model. The model is trained for a total of 300 epochs. We set the batch size to 8 and the learning rate to 0.001. Our experimental conditions are 4 NVIDIA GeForce GTX 1080Ti graphics cards.

4.1 Qualitative Comparison

It can be seen from the above figure that from the perspective of highlighting the target, the TarDAL model and the model in this paper combine the advantages of conspicuous heat source targets in IR images and rich texture details and colors in RGB images. However, the images fused by the TarDAL model have problems such as the overall color being off-white, the color of heat sources, such as people, being lost, and the details being distorted. In contrast, the image fused by the model proposed in this paper, under the premise of highlighting the target, dramatically retains the texture details and color of the target and has a higher hue consistency with the original image. Therefore, the model proposed in this paper has a better human visual experience than TarDAL (Fig. 5 and Table 1).

Table 1 Fusion image evaluation indicators

	PSNR/RGB	PSNR/IR	SSIM/RGB	SSIM/IR	EN	mAP
TarDAL	13.14	12.56	0.61	0.23	5.73	0.88
Ours	24.82	18.05	0.88	0.33	5.84	0.88

Fig. 5 The color on the left is distorted overall, while the color on the right has more color information

Table 2 Detection performance improvement

	mAP	Improved
RGB	0.773	0.107
IR	0.787	0.093
Fus	0.880	

4.2 Quantitative Comparison

The fusion image generated by the our model has higher PSNR and SSIM, indicating that the image has higher image quality, and higher information entropy (EN) shows that the image has more information. In addition, the similar higher level of mAP shows that our model has higher image quality and better human visual perception without losing the performance of downstream object detection tasks, and retains more color information (Table 2).

According to the data in the table, we can see that the performance of target detection has been greatly improved by fused images, and the usability of this model has also been verified.

5 Conclusion

We propose a novel IVIF scheme and train a fusion model. This model can generate an image with both IR image target pixel intensity and RGB image background texture details, and at the same time, it also solves the problem of loss of target details and color distortion in fusion images of other models. And compared with other models, under the guarantee of a high level of mAP, the PSNR, SSIM, EN and other indicators have been greatly improved. It not only improves the performance of downstream target detection tasks, but also has the human visual experience.

Acknowledgements Partially funded by Natural Science Foundation of Sichuan Province (2023NSFSC0479) and partially funded by Grant SCITLAB-20005 of Intelligent Terminal Key Laboratory of Sichuan Province. This project is supported by the fund of Science and Technology of Sichuan Province (No.2021YFG0330).

References

1. Kim JU, Park S, Ro YM (2021) Uncertainty-guided cross-modal learning for robust multispectral pedestrian detection. IEEE Trans Circ Syst Video Technol 32(3):1510–1523
2. Liu T, Lam KM, Zhao R, Qiu G (2021) Deep cross-modal representation learning and distillation for illumination-invariant pedestrian detection. IEEE Trans Circ Syst Video Technol 32(1):315–329
3. Lin D, Hu S, Wu W, Wu G (2023) Few-shot RF fingerprinting recognition for secure satellite remote sensing and image processing. Sci China Inf Sci. https://doi.org/10.1007/s11432-022-3672-7
4. Wu W, Hu S, Lin D, Wu G (2022) Reliable resource allocation with RF fingerprinting authentication in secure IoT networks. Sci China Inf Sci 65(7):170304
5. Li H, Wu XJ, Kittler J (2020) MDLatLRR: a novel decomposition method for infrared and visible image fusion. IEEE Trans Image Process 29:4733–4746
6. Zhang Q, Liu Y, Blum RS, Han J, Tao D (2018) Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: a review. Inf Fusion 40:57–75
7. Fu Z, Wang X, Xu J, Zhou N, Zhao Y (2016) Infrared and visible images fusion based on RPCA and NSCT. Infrared Phys Technol 77:114–123
8. Bavirisetti DP, Dhuli R (2016) Two-scale image fusion of visible and infrared images using saliency detection. Infrared Phys Technol 76:52–64
9. Zhang H, Xu H, Xiao Y, Guo X, Ma J (2020, April) Rethinking the image fusion: a fast unified image fusion network based on proportional maintenance of gradient and intensity. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, no 07, pp 12797–12804
10. Ma J, Tang L, Xu M, Zhang H, Xiao G (2021) STDFusionNet: an infrared and visible image fusion network based on salient target detection. IEEE Trans Instrum Measur 70:1–13
11. Xu H, Ma J, Jiang J, Guo X, Ling H (2020) U2Fusion: a unified unsupervised image fusion network. IEEE Trans Pattern Anal Mach Intell 44(1):502–518
12. Sun Y, Cao B, Zhu P, Hu Q (2022, Oct) Detfusion: a detection-driven infrared and visible image fusion network. In: Proceedings of the 30th ACM international conference on multimedia, pp 4003–4011
13. Tang L, Yuan J, Zhang H, Jiang X, Ma J (2022) PIAFusion: a progressive infrared and visible image fusion network based on illumination aware. Inf Fusion 83:79–92

14. Liu J, Fan X, Huang Z, Wu G, Liu R, Zhong W, Luo Z (2022) Target-aware dual adversarial learning and a multi-scenario multi-modality benchmark to fuse infrared and visible for object detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5802–5811

A Study of RF Fingerprint Feature Dimensionality Expansion for I/Q Signal Data Extraction

Han Zhou, Di Liu, MengJuan Wang, Di Lin, BingTao Li, and Jun Yang

Abstract Security research based on RF fingerprints has received extensive attention in recent years. Traditional RF fingerprint identification technology is first based on feature engineering method to extract signal key features and then uses machine learning method to identify and classify, but the method relies on expert experience and high computational complexity and lacks universality in feature selection and judgment criterion; while the RF fingerprint identification research based on deep learning that has emerged in recent years is usually obtained by extracting statistical features from signal sample sequences and often require colossal storage and arithmetic power. This paper proposes an RF fingerprint feature expansion dimensional recognition method to save storage and computational power and improve the universality and robustness of RF fingerprint recognition technology. The method preserves some results of traditional RF fingerprint recognition feature engineering and then builds a dimensional data table using the time-frequency map image features converted from the sample sequence of the extracted signal as the extended dimension. Finally, using the feature dimension data as input, machine learning and deep learning models are used for fingerprint recognition, respectively.

Keywords RF fingerprint recognition · Time-frequency map · Image feature recognition · Machine learning · Deep learning

1 Introduction

With the widespread deployment of wireless networks and the successful development of the Internet of Things (IoT), people are paying more and more attention to security issues. The security mechanism based on RF fingerprints has received wide attention. Its principle is to use the unique fingerprints of RF devices formed by the intrinsic characteristics of the hardware or hardware production defects of the RF devices [1]. These unique RF fingerprint characteristics of the devices can

H. Zhou · D. Liu (✉) · M. Wang · D. Lin · B. Li · J. Yang
University of Electronic Science and Technology of China, Chengdu, Sichuan, China
e-mail: lindi@uestc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_50

be extracted through a series of processing of the signals of wireless devices. The use of RF fingerprints for identification belongs to the physical layer of the security authentication mechanism, and the feature is that it is pretty challenging to imitate and replicate the characteristics of RF fingerprints through software. Therefore, RF fingerprinting is considered a potential enabler to mitigate and even solve the security-related topics of future wireless networks [2]. Traditional research on RF fingerprinting techniques is based on a feature engineering approach, which requires experts in the field to first select the signal's critical features based on the signal's actual situation, calculate these features, and finally classify them based on fixed rules or machine learning methods. This method relies on manual experience and lacks universality regarding feature selection and judgment criteria [3]. In contrast, in recent years, with the integration of artificial intelligence-related technologies with RF fingerprinting, deep learning algorithms can be used to automatically learn the pattern features of radio waves, bypassing the need for experience-based manual feature extraction [4]. However, the current deep learning-based RF fingerprinting technology mainly uses raw I/Q signal data as the input to the model, which has two problems: first, the I/Q signal contains only signal time domain information and noise, while other implicit features are not evident in the time domain signal data. Second, when using I/Q signal data as input for deep learning model training, a large amount of storage space resources are required for I/Q data storage, and the training will also be difficult and time-consuming because of the enormous input.

To comprehensively solve the above problems, this paper proposes to characterize the original I/Q data and expand the feature dimension, convert the sample sequence of the signal into an image, extend the dimension by extracting the features of the image as the standard RF features, and establish the feature dimension data table. Instead of using the original I/Q data sequence of the infinite transmitter device, the feature data table formed after the characterization process is used as the input for subsequent identification, thus reducing the pressure of storage and model training.

2 Methods

2.1 Feature Table Dimension Selection

In the feature table dimension selection for the original data characterization, the work related to traditional RF fingerprinting feature engineering is retained: some features of two steady-state signal RF fingerprinting techniques based on unguided and guided structures are used as feature dimensions [5], and the image features of time-frequency images drawn by fusing wireless network signals are used as expanded dimensions.

2.2 Steady-State Signal RF Fingerprint Characteristics

The RF fingerprint identification technology based on the frequency-free structure is mainly applied to the complex electromagnetic environment, such as the modern information battlefield. In order to improve the universality of the feature data sheet dimension and adapt the signal data based on such an environment, some features, such as fractal features, high-order moment features, and phase noise spectrum features that do not require high accuracy for the signal itself, are selected.

With the massive popularity of wireless communication devices, especially the maturity and broad application of 5G technology, RF fingerprinting technology is no longer limited to the military field. In daily life, a large amount of communication data is with a guide frequency header, and the RF fingerprinting technology based on the guide frequency structure can effectively use the frame structure of communication signals. Therefore, the research in this paper selects the generation mechanism and calculation method of carrier frequency phase precision estimation features, constellation diagram features, I/Q offset features, and power spectrum difference statistics features, which are challenging to calculate accurately under the lead-free structure.

2.3 Time-Frequency Diagram Characteristics

This study finds that the RF fingerprint extraction method based on signal sample sequences requires enormous data storage space on the one hand when the data volume of signal samples is substantial and increases the computation and complexity of feature extraction on the other hand. When the signal is converted into an image, the image size is fixed so that the complexity of subsequent processing and computation is not increased, thus achieving light weight. By analyzing several time-frequency transformation algorithms, four of the time-frequency transformed images are shown in (Fig. 1).

It can be seen that the time-frequency map obtained by using the Choi-Williams distribution transformation [6] can effectively suppress the interference of the cross terms on the time-frequency distribution and also has a high resolution in the time-frequency representation, which can be used for image feature extraction.

In this study, the time-frequency images are preprocessed first, then the shape and texture features of the time-frequency maps are extracted separately.

Shape characteristics: Zernike moments are better in describing the image shape, and its low-order moments and high-order moments represent the image's overall shape and the image's details [7], respectively. At the same time, Zernike moments belong to the statistical properties of pixel grayscale in the image, which can better suppress the effect of noise. Assuming that the grey imagery scale function is denoted as $f(x, y)$, the Zernike moments are defined as

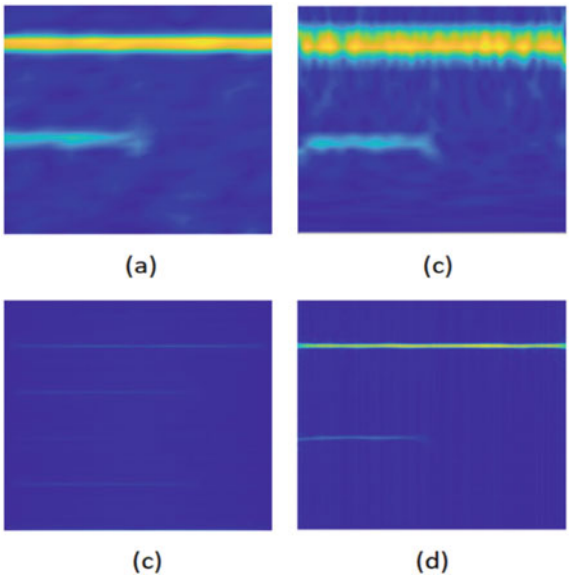


Fig. 1 **a** Short-time fourier transform time-frequency diagram; **b** wavelet transform time-frequency diagram; **c** WVD time-frequency diagram; **d** CWD time-frequency diagram

$$Z_{nm} = \frac{n+1}{\pi} \iint V_{nm}^{*}(x, y) f(x, y) dx dy \tag{1}$$

Texture characteristics: In 1994, Timo Ojala et al. combined traditional statistical and structural methods and proposed using Local Binary Pattern (LBP) to extract image texture features. However, to adapt the time-frequency map and better describe the texture information of the time-frequency map, the enhanced uniform local binary pattern EULBP is proposed by improving the uniform local binary pattern.

Step 1: Firstly, the pre-processed time-frequency map is divided into 8x8 sub-regions.

Step 2: Calculate the LBP values of the pixel points in each sub-region. As shown in (Fig. 2), a local neighbourhood block of size 3×3 is taken as the centre of each pixel point, and the grey value of the centre pixel point is compared with the grey value of each pixel point in the neighbourhood block, and the value more significant than the centre pixel point is recorded as 1, and the value less than the centre pixel point is recorded as 0. Thus, the 8-bit LBP value of each pixel point can be obtained.

Step 3: Obtain the histogram vector values. During the binarization coding process in step 2, the ULBP value, which is the LBP value with 0,1 jump count less than 2, is recorded and retained. Then the frequency of this ULBP value in the corresponding sub-region is calculated and represented as a histogram, which is then normalized and denoted by $x_0, x_1, x_2 \cdots x_j$ as the vector value of each histogram.

Step 4: Calculate the EULBP feature vector of the time-frequency map. The vector values obtained in step 3 are scaled up, and the histograms of each subregion are concatenated to obtain the EULBP eigenvectors of the entire time-frequency map.The EULBP eigenvalues are defined as

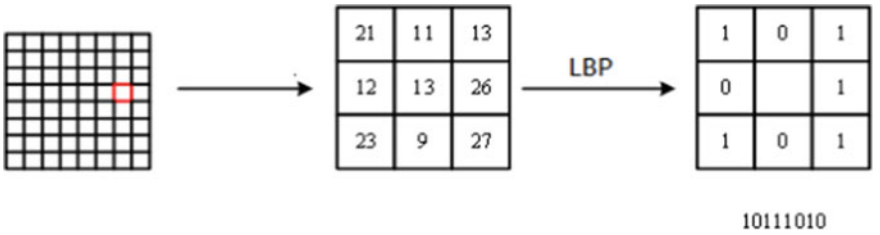


Fig. 2 Sub-region pixel point LBP coding process

$$EULBP = \sqrt{\frac{x_i}{\|M\|}}$$
$$\|M\| = \sqrt{x_0^2 + x_1^2 + \cdots + x_j^2} \tag{2}$$

2.4 Identification Model

Two basic device recognition schemes are proposed. The first is based on the deep learning wireless device recognition scheme, while the second chooses the traditional machine learning model.

Based on traditional machine learning models, the first recognition scheme experimentally demonstrated that the recognition rate of decision tree and SVM is 90.12% and 89.17%, respectively.

The second scheme is a deep learning model based on residual networks, and this paper is based on the ResNet18 network, an improved CSAM-ResNet network model more suitable for recognizing the time-frequency map of wireless device signals, which was found to have inherent advantages and good results in the field of image recognition [8]. After experiments, it was proved that the average recognition effect of this model could reach about 98.6%.

3 Experiment

3.1 Experimental Environment Configuration

This experiment is programmed in a Ubuntu environment using Python language, the deep learning framework used is Pytorch, and the experiment uses CUDA parallel computing framework. The hardware configuration is shown in (Table 1).

Table 1 Experimental equipment

Hardware name	Hardware specifications
Central processing unit	Intel xeon silver 4110
Graphics processor	NVIDIA GeForce GTX1080Ti
Main memory	SKHynix DDR4 2666V
Solid state drives	INTEL SSD SATA3.0

Table 2 Devices similar frequency

Central frequency point (MHz)	Bandwidth (KHz)	Equipment number	Data volume
1028	25	Device1	1,664,852
		Device2	1,335,246
		Device3	1,224,565
		Device4	1,552,456
1040	25	Device5	1,254,002
		Device6	1,365,421
1065	28	Device7	1,524,684
		Device8	1,258,746

3.2 Experimental Data

This paper uses a self-built dataset as the data source for training and testing the model. The self-built dataset comes from the RF signal data collected by the State Key Laboratory of Anti-Interference, which contains I/Q signal data of 8 devices in 3 similar frequency bands, as shown in (Table 2).

The time-frequency diagram of some device signals is shown in (Fig. 3). Device 1 in the self-built dataset.

3.3 Experimental Design and Analysis of Results

The experimental process is shown in Fig. 4, using the I/Q signal data of 8 devices in the similar frequency band as the original input, the data was first pre-processed to make it compound the requirements of the Python code for extracting features, and then the original data stream was extracted in two processes in parallel: A process is the steady-state signal RF fingerprint extraction process, and B process is the signal data drawn with time-frequency. The extended dimension extraction process is used to extract the image features, and finally, all the dimensional data are aggregated

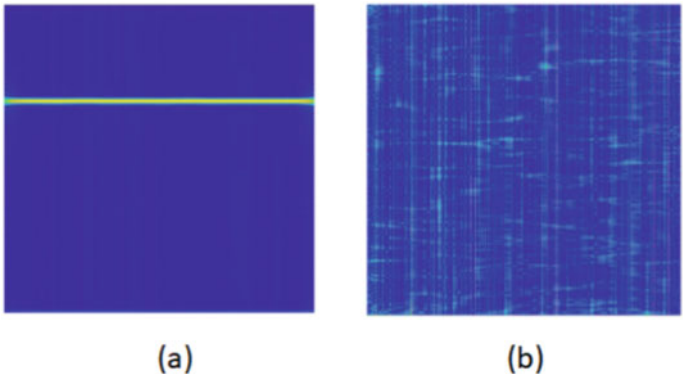


Fig. 3 Self-built signal data device1

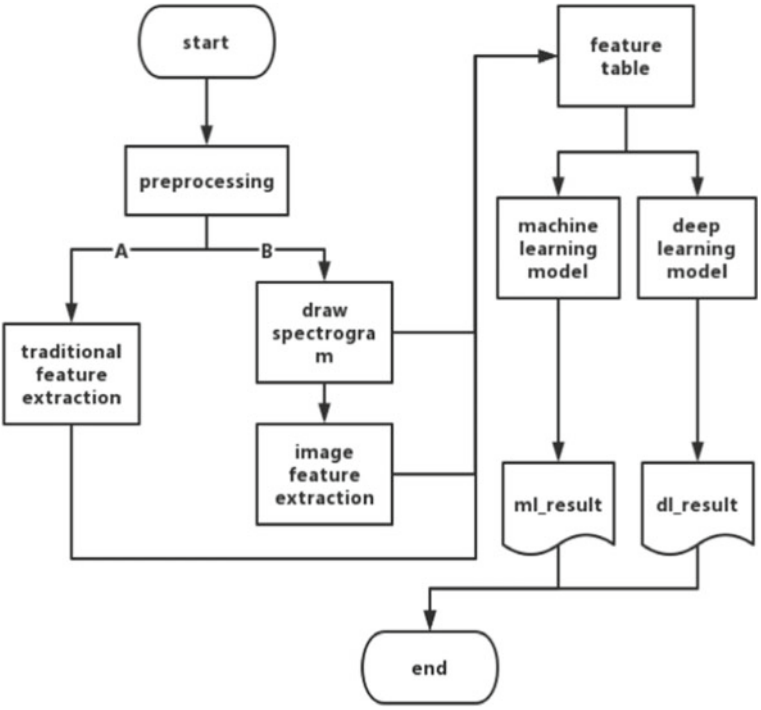


Fig. 4 Experimental flow chart

to form a dimensional data table, which is then put into the subsequent machine learning and deep learning models to learn and validate the results.

The model identification results are shown in (Table 3).

Table 3 Model identification results

Algorithm	Training time (s)	EAverage recognition accuracy (%)
Zernike and EULBP + dicision tree	11,520	90.12
Zernike and EULBP + svm	12,360	89.17
CSAM-ResNet	16,280	98.61

4 Conclusion

In this paper, we innovatively propose to use the image features of time-frequency maps drawn from signal data as extended dimensions to build feature data tables for recognition. Compared with the recognition scheme of pure feature engineering and the deep learning scheme that directly uses I/Q data as input, the arithmetic power and storage requirements for model training are reduced, and the accuracy rate is maintained at a high level. This study utilizes techniques in image recognition to provide new perspectives and ideas for studying RF fingerprint technology. Follow-up work can start from data pre-processing, improve the dimension selection of feature engineering, and improve or replace the image feature extraction algorithm; the page can work on the recognition model to further improve the recognition rate and data adaptability.

Acknowledgements Partially funded by Natural Science Foundation of Sichuan Province (2023NSFSC0479) and partially funded by Grant SCITLAB-20005 of Intelligent Terminal Key Laboratory of Sichuan Province.

References

1. Rehman SU, Sowerby KW, Alam S, Ardekani I (2014) Radio frequency fingerprinting and its challenges. In: 2014 IEEE Conference on communications and network security. San Francisco, CA, USA, pp 496–497. https://doi.org/10.1109/CNS.2014.6997522.
2. Wu W, Hu S, Lin D, Wu G (2022) Reliable resource allocation with RF fingerprinting authentication in secure IoT networks. Sci China Inf Sci 65(7):170304
3. Merchant K, Revay S, Stantchev G, Nousain B (2018) Deep learning for RF device fingerprinting in cognitive communication networks. IEEE J Sel Topics Signal Process 12(1):160–167. https://doi.org/10.1109/JSTSP.2018.2796446
4. Wu W, Lin D, Hu S, Yang G Secure IoU in adversarial environments: user authentication and performance optimization. In: IEEE transactions on vehicular technology. https://doi.org/10.1109/TVT.2023.3266770.
5. Lin D, Hu S, Wu W, Wu G (2023) Few-shot RF fingerprinting recognition for secure satellite remote sensing and image processing. Sci China Inf Sci. https://doi.org/10.1007/s11432-022-3672-7

6. Zhao P, Xu L (2021) Research on time domain filtering based on choi-williams distribution about time-phase modulation. In: 2021 7th annual international conference on network and information systems for computers (ICNISC). Guiyang, China, pp 708–715. https://doi.org/10.1109/ICNISC54316.2021.00132.
7. Khotanzad A, Hong YH (1990) Invariant image recognition by Zernike moments. IEEE Trans Pattern Anal Mach Intell 12(5):489–497. https://doi.org/10.1109/34.55109
8. Uzundurukan E, Dalveren Y, Kara A (2020) A database for the radio frequency fingerprinting of bluetooth devices. Data 5(2):55

Research on Wavelength Demodulation Algorithm for Fiber Bragg Grating Weak Signal Based on Hilbert Transform

Tao Li, Chuan Dong, Ying Geng, Tianmin Zhang, Jiejun Lin, Weijie Jiang, and Yinguo Huang

Abstract In order to improve the demodulation accuracy of the tunable F-P filter, the Hilbert transform is introduced into the Gauss fit algorithm to solve the problem of sidelobe or peak distortion in the low power and low signal-to-noise ratio. First, digital low-pass filtering is used to complete the data pre-processing, then using the peak-seek algorithm designed in this paper to confirm the peak position of the two signals, and finally using the linear interpolation algorithm to demodulate the FBG center wavelength. Demodulation experiments show that the demodulation method used in this paper has the advantages of small computation, high demodulation accuracy and less than 1 pm peak seeking error.

Keywords Fiber Bragg grating · Wavelength demodulation · Tunable F-P filter · Hilbert transform · Gauss fit algorithm

1 Introduction

Fiber Bragg Grating (FBG), as a wavelength modulated sensor, reflects the changes of measured physical quantities by wavelength modulation of the sensing signal [1]. Its central wavelength can be used to measure parameters such as strain [2], temperature [3], pressure [4], humidity [5], vibration [6], viscosity [7], etc. But in practical applications, the FBG center wavelength is obtained by the corresponding peak-seek method, so the detection accuracy of FBG sensor networks is closely related to the peak-seek accuracy of its demodulation algorithm.

Commonly used peak-seek algorithms include direct peak-seek algorithm, power weighted peak-seek algorithm, general polynomial fitting, and Gaussian fitting algorithm, etc. Among which Gaussian fitting algorithm has the advantages of good noise

T. Li · J. Lin · W. Jiang
Shanghai Marine Equipment Research Institute, Shanghai, China

C. Dong · Y. Geng · T. Zhang · Y. Huang (✉)
Tianjin University, Tianjin, China
e-mail: hyg@tju.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_51

immunity and small error, and the related improved algorithm is also the most. Hu et al. [7] optimized the fitting coefficients of Gauss curve by LM algorithm, which reduced the peak-seek error of FBG demodulation algorithm effectively; Jiang et al. [8] proposed a Weighted Least Squares fitting combined with Asymmetric Gauss correction (WLS-AG) algorithm to achieve high precision peak-seek for the sampled ultra weak fiber Bragg gratings with interference noise; Zhai and Li [9] designed a Gauss curve fitting method (WTG) based on the Wavelets Transform to improve the demodulation accuracy at low sampling rates. Djurhuus et al. [10] proposed a signal processing method for FBG temperature measurements based on the machine learning tool Gauss Process Regression (GPR). Experiments show that the method has lower error and higher detection accuracy.

However, when the FBG reflected light signal has low power, low signal-to-noise ratio, and multiple peaks, traditional demodulation algorithms suffer from low accuracy, poor noise resistance, and false peak detection. This paper combines Hilbert transform with Gauss fitting algorithm to achieve multi peak detection, while reducing the amount of demodulation algorithm operations and improving the demodulation accuracy.

2 Principles of Tunable F-P Filter Demodulation System

The tunable F-P filter demodulation method has the advantages of high accuracy, wide demodulation range, and the ability to achieve multiplexing of large-scale FBG sensor networks. The principle of tunable F-P filter demodulation technology based on F-P etalon is shown in Fig. 1, where the function of F-P etalon is to perform nonlinear calibration on tunable F-P filter.

A broad-band optical signal from ASE broadband light source enters Tunable F-P Filter, which is controlled by Drive circuit to transmit narrow-band optical signals

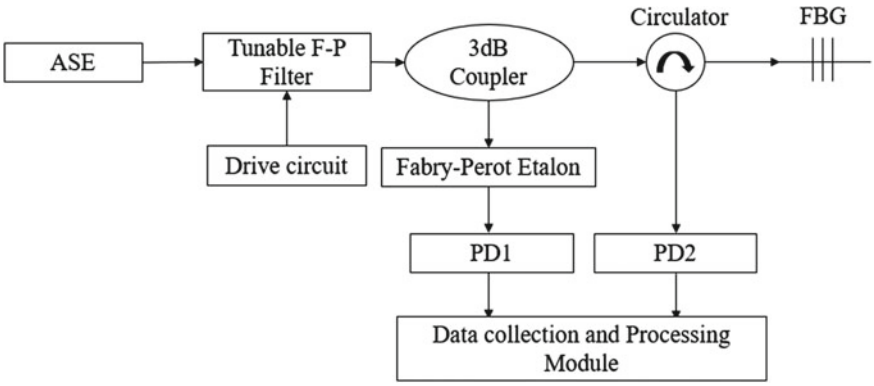


Fig. 1 Schematic diagram of tunable F-P filter demodulation principle based on F-P etalon

with different central wavelengths. The narrow-band optical signal enters the F-P Etalon one way through a 3 dB Coupler, its transmitted signal is detected by PD1, the other way through Circulator into the FBG, and the reflected optical signal is detected by PD2 when the narrow-band light transmitted by the Tunable F-P Filter meets the FBG Bragg condition. After the Data collection and Processing module, the peak-seek algorithm, curve fitting and other data processing are completed in the computer, thus wavelength demodulation of FBG is achieved.

3 Research on FBG Demodulation Algorithms

The overall flow of the center wavelength demodulation algorithm in this article is shown in Fig. 2.

First read the two signals from the FBG and F-P Etalon, the second is digital low-pass filtering is performed to complete data pre-processing, the next is to confirm the peak position of the two signals using the peak-seek algorithm designed in this paper, finally, the FBG center wavelength is demodulated using a linear interpolation algorithm.

3.1 Digital Low-Pass Filter Processing

First, digital low-pass filtering is performed on the acquired raw data to filter out part of the high-frequency noise introduced during the acquisition. Digital filters are divided into Finite Impulse Response Filter (FIR) and Infinite Impulse Response Filter (IIR), the differences between which are shown in Table 1. In order to reduce

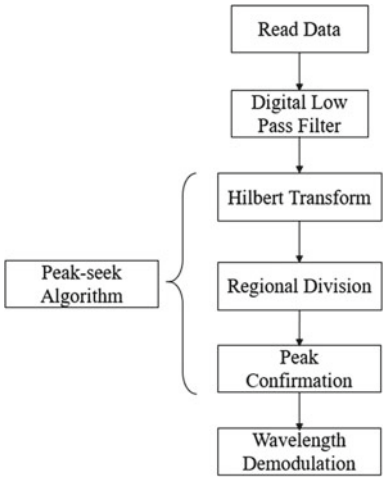


Fig. 2 FBG demodulation algorithm flow

Table 1 Performance comparison of FIR and IIR filters

	Characteristic	Applications
FIR	Good linearity, strong stability, but computational complexity	System of image signal processing and other waveforms carrying information
IIR	Simple design, but pole offset can cause stability problems	Systems with low phase requirements such as voice communications

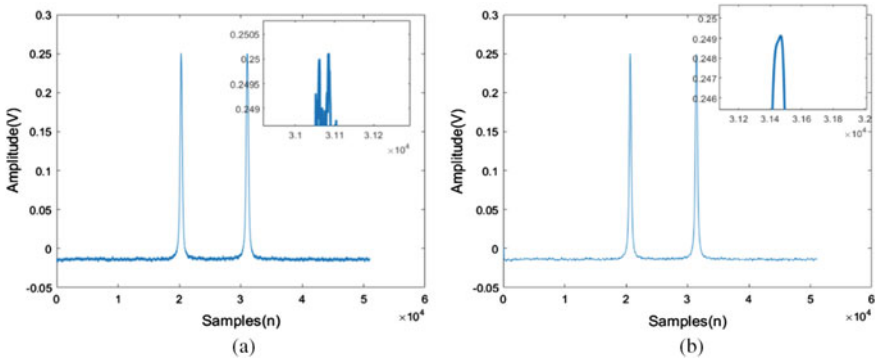


Fig. 3 **a** FBG reflectance spectrum raw data. **b** FBG digital filtered output

the loss of information as much as possible, FIR digital filters are used to process the data.

The raw data are the FBG reflection spectrum and the F-P etalon transmission spectrum, where all the sampling points in one scan period of the drive circuit are selected as the raw data. The time-domain waveform of the FBG raw data is shown in Fig. 3a.

The two peak signals in Fig. 3a correspond to the two FBG reflection spectra of the up and down travel of the tuned F-P filter driving voltage delta wave, respectively, and are evident from local amplification at the upper right corner, with some high-frequency noise in the original signal. Using the digital filter module in the Matlab toolbox, a low-pass filter with a cutoff frequency of 1 kHz was designed using a Hamming window, and the results are shown in Fig. 3b, where it can be found that the higher order harmonics near the FBG peak are filtered out, which will aid in the next step of the peak-seek algorithm processing. The same treatment was applied to the F-P etalon transmission spectrum, and the results are shown in Fig. 4.

3.2 Peak-Seek Algorithm Design

The peak-seek algorithm is to confirm that the optical signals of the F-P etalon and FBG are scanned by a tunable F-P filter, and the peak position of the electrical signal

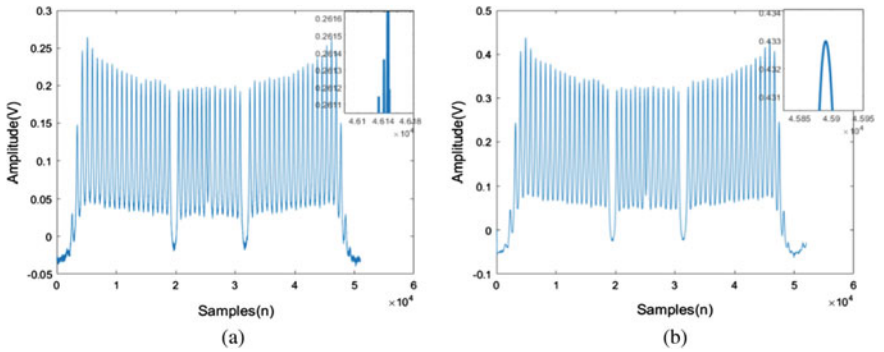


Fig. 4 **a** F-P etalon transmission spectrum raw data. **b** F-P etalon digital filtered output

after the optoelectronic conversion, with the peak value of each electrical signal corresponding to the central wavelength value of the optical signal. There are four commonly used peak-seek algorithms: direct peak-seek algorithm, power weighted peak-seek algorithm, general polynomial fitting, and Gaussian fitting algorithm.

The characteristics of various commonly used algorithms are shown in Table 2.

From the table, it can be seen that the Gaussian fitting algorithm has high peak finding accuracy and strong anti noise performance, so there are also the most improved algorithms based on the Gaussian algorithm (such as the Gaussian polynomial fitting algorithm). In order to reduce the calculation amount of the Gaussian fitting algorithm and adapt to the characteristic that the transmission spectrum of the F-P etalon contains multiple peak signals, this paper combines the Hilbert transform with the Gaussian fitting algorithm to complete the determination of the peak signal.

Table 2 Comparison of common peak-seek algorithms

	Accuracy	Noise immunity	Characteristic
Direct peak-seek algorithm	Low	Low	High sampling rate requirement
Power weighted peak-seek algorithm	Low	Low	High peak finding accuracy under low noise conditions
General polynomial fitting	Middle	Low	If the peak value is not within the sampling point, the error will be significant
Gaussian fitting algorithm	High	High	High accuracy and strong noise resistance, but complex calculations
Gaussian polynomial fitting algorithm	Middle	Low	Similar to general polynomial fitting methods, but with better accuracy and noise resistance

(1) Hilbert transform

The function of Hilbert transform is to realize the predetermined bit and region segmentation of the peak value of the signal after low-pass filtering denoising, which can not only reduce the data volume of Gaussian fitting algorithm, but also eliminate the interference of other components in the signal to the fitting. Hilbert transform of time domain signal x(t) can be expressed as:

$$\hat{x}(t) = H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{+\infty} x(\tau) \frac{1}{t-\tau} d\tau \tag{1}$$

where x(t) and $\hat{x}(t)$ are linear, and Hilbert transform can be expressed in convolution form:

$$\hat{x}(t) = H[x(t)] = x(t) * \frac{1}{\pi t} \tag{2}$$

Thus, the complexity of calculation $\hat{x}(t)$ is reduced. As shown in Fig. 5, Hilbert transform is performed on the filtered FBG reflection spectrum and F-P etalon transmission spectrum.

The signal peak after low-pass filtering is located between the two poles after Hilbert transform, so the sampling point interval corresponding to the two adjacent positive and negative poles can be used as the initial estimation of the peak position and peak interval. Compared with the constant threshold partition, Hilbert transform can overcome the problem of false peak detection caused by side lobe or peak distortion.

(2) Gaussian fitting algorithm

Assuming the sampling data is $(x_i, y_i)(i = 1, 2, \ldots)$, the Gaussian fitting algorithm can be represented by Eq. (3):

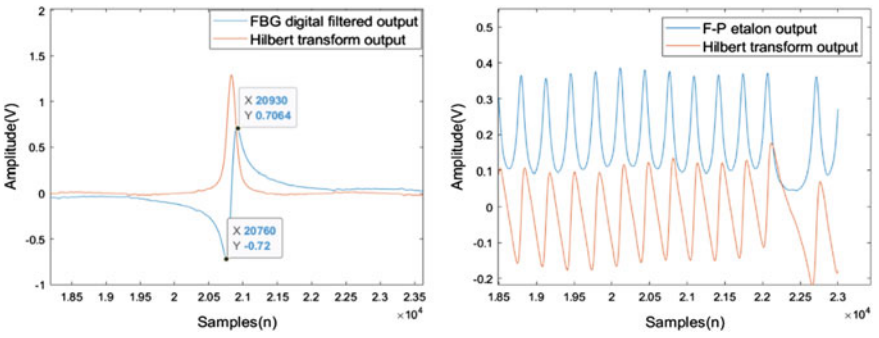


Fig. 5 Hilbert transform of FBG data and F-P etalon data

$$y_i = y_{\max} \times \exp\left[\left(\frac{x_i - x_{\max}}{\sigma}\right)^2\right] \tag{3}$$

In the formula, y_{max}, x_{max}, and σ are the peak values of the sampled data, the sampling points corresponding to the peak values, and the 3 dB bandwidth. For

$$z_i = \mathrm{In} y_i,\ b_0 = \mathrm{In} y_{\max} - \frac{x_{\max}^2}{\sigma^2},\ b_1 = \frac{2x_{\max}}{\sigma^2},\ b_2 = -\frac{1}{\sigma^2}$$

then Eq. (3) can be simplified as:

$$z_i = b_0 + b_1 x_i + b_2 x_i^2 \tag{4}$$

Substitute all data into Eq. (4), which can be expressed as Z = XB by matrix, Expand to:

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \tag{5}$$

The least square solution of matrix B can be calculated by the least square principle

$$B = \left(X^T X\right)^{-1} X^T Z \tag{6}$$

Through the above analysis, the peak position can be expressed as:

$$x_{\max} = -\frac{b_1}{2b_2} \tag{7}$$

Gaussian fitting curves can accurately calculate the center wavelength position. Due to its good mean square error, strong stability, and excellent noise resistance, it can effectively fit the spectral data of FBG and F-P etalon. The Gaussian fitting results are shown in Fig. 6.

3.3 Linear Interpolation Algorithm

Since the data acquisition module is used to simultaneously collect the signals of FBG and F-P etalon at the same sampling rate, the data of the two channels can be combined to solve the central wavelength value of FBG. The method is as follows: Confirm the sampling point corresponding to the center wavelength of each narrow-band spectrum through the negative peak marker point of the F-P etalon, and set the

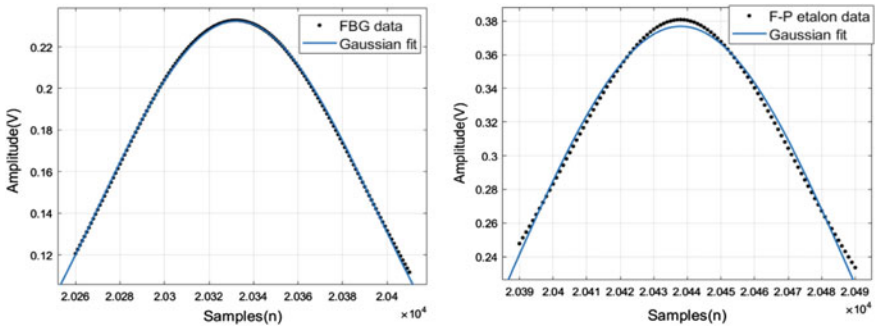


Fig. 6 Gaussian fitting of FBG data and F-P etalon data after threshold division

Table 3 Comparison of two peak-seek algorithms

Peak-seek algorithms	FBG1 error (pm)	FBG2 error (pm)
Gaussian fitting algorithm	5.14	5.08
Our peak-seek algorithm	0.56	0.63

center wavelengths of any two narrowband spectra of the F-P etalon as λ_1 and λ_2. The sampling points are x_1 and x_2 respectively, so the wavelength value corresponding to any sampling point λ_x can be obtained by Eq. (8)

$$\lambda_x = \frac{\lambda_2 - \lambda_1}{x_2 - x_1}(x - x_1) + \lambda_1 \tag{8}$$

The FBG center wavelength measured with a demodulator (SmartScan 08-80 XR, repeatability less than 1 pm) is the agreed true value, the comparison between the Gaussian fitting algorithm and the peak-seek algorithm designed in this article is shown in Table 3.

4 Conclusion

In this paper, firstly, digital low-pass filtering is performed on the collected original signal to remove high-frequency noise, and then Hilbert transform is introduced into the Gaussian fitting algorithm to solve the defect that the constant threshold division needs to predict the range of each peak, reducing the calculation amount of the Gaussian fitting algorithm. Finally, combined with the F-P etalon spectrum, the central wavelength of the FBG is extracted through the linear interpolation algorithm. The experiment shows that the peak finding error of the demodulation algorithm designed in this article is below 1 pm, which has certain reference value for engineering needs, especially for multi peak detection in distributed sensor networks.

References

1. Chen J, Liu B, Zhang H (2011) Review of fiber Bragg grating sensor technology. Front Optoelectron China 4:204–212
2. Shiratsuchi T, Imai T (2021) Development of fiber Bragg grating strain sensor with temperature compensation for measurement of cryogenic structures. Cryogenics 113:103233
3. Rosolem JB, Penze RS, Floridia C et al (2020) Dynamic effects of temperature on FBG pressure sensors used in combustion engines. IEEE Sens J 21(3):3020–3027
4. Duan Z, Jiang Y, Tai H (2021) Recent advances in humidity sensors for human body related humidity detection. J Mater Chem C 9(42):14963–14980
5. Li T, Guo J, Tan Y et al (2020) Recent advances and tendency in fiber Bragg grating-based vibration sensor: a review. IEEE Sens J 20(20):12074–12087
6. Song L, Fang F, Zhao J (2012) Study on viscosity measurement using fiber Bragg grating micro-vibration. Meas Sci Technol 24(1):015301
7. Hu Z, Pang C, Cheng F (2017) Application of Gaussian-LM algorithm in fiber Bragg grating reflection spectrum peak search. Laser Optoelectron Prog 54(1):013001
8. Jiang H, Guo Y, Zheng X et al (2020) Research on an improved spectral peak seeking algorithm with high precision. Study Opt Commun 46(1):33
9. Zhai X, Li Z (2020) Application of wavelet transform-based Gauss curve fitting in peak positioning. Opt Commun Technol 44(2):10–13
10. Djurhuus MSE, Werzinger S, Schmauss B et al (2019) Machine learning assisted fiber Bragg grating-based temperature sensing. IEEE Photon Technol Lett 31(12):939–942

BDF-YOLOV5: Improved YOLOV5 Based on Bi-directional Fusion Network for Dense Pedestrian Detection

Yuhui Xu and Ruian Liu

Abstract Pedestrian detection is a challenging task in the field of computer vision and plays a crucial role in downstream tasks, such as video surveillance and autonomous driving.Despite significant progress over the past two decades, scale variance and occlusion remain prominent issues. To address these problems, we propose BDF-YOLOv5 in this paper. Based on YOLOv5, we replace the original FPN with the BDF network structure. Furthermore, to further improve our BDF-YOLOv5, we additionally improved the loss function for bounding box regression and proposed weighted-CIOU. Extensive experimental results on the Crowdhuman dataset demonstrate the feasibility of our method. Compared to the baseline model (YOLOv5), BDF-YOLOv5 achieves an improvement of approximately 4.0%.

Keywords BDF-YOLOv5 · Pedestrian detection · YOLOv5 · Weighted-CIOU

1 Introduction

Human-centric computer vision tasks, such as pedestrian detection, face recognition, pose estimation, assisted driving, and intelligent robotics, have made great strides in the past decade. Among these tasks, pedestrian detection is one of the most basic and widely applicable. In addition to its important role in some application scenarios, such as video surveillance and traffic safety, pedestrian detection is also a foundational task for several other visual tasks. This paper aims to improve the performance of pedestrian detection and provide convenience for the aforementioned application scenarios.

In recent years, deep convolutional neural network-based object detection algorithms have been widely applied and made significant progress in natural scene object detection tasks, such as two-stage detectors including R-CNN [1], Faster R-CNN [2], SPPNet [3], and single-stage detectors such as SSD [4] and RetinaNet [5]. However,

Y. Xu · R. Liu (✉)
College of Electronic Information and Communication Engineering,
Tianjin Normal University, Tianjin 300387, China
e-mail: ruianliu@sina.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_52

Fig. 1 Illustrates with intuitive examples the three main issues in dense pedestrian detection tasks: occlusion, scale variation, and pedestrian diversity

most previous object detectors are designed for natural scene images. When directly applying these models to pedestrian object detection tasks in dense crowds, there are three main issues, as visually illustrated in Fig. 1. First, pedestrians are too dense, resulting in occlusions between individuals. Second, pedestrian detection scenes are often macroscopic images, leading to significant changes in target scale. Finally, the varied postures of pedestrians also pose challenges for detection.

In object detection tasks, the YOLO series [6–9] plays an important role in one-stage detectors due to its fast and efficient characteristics. In this paper, we propose an improved model, BDF-YOLOv5, based on YOLOv5 [10], to address the aforementioned three issues. The main contributions of this study are as follows:

- The backbone network still adopts the original version of CSPDarknet53, while we propose a novel feature network called BDF (Bi-Directional Fusion Network) for the neck part. By bi-directionally fusing high- and low-level semantic information of feature maps, BDF can effectively retain feature information and alleviate scale variation, thus improving the accuracy of multi-scale object detection.
- We propose a novel (BBR) bounding box regression loss function,Weighted CIOU, which balances the contribution of high-quality and low-quality regression samples. This effectively reduces false positives and mitigates occlusion issues in dense crowd scenes.
- For feature fusion, we abandoned the commonly used C3 module and instead adopted CSPNet.

2 Related Work

2.1 Pedestrian Detection Pipeline

Most pedestrian detectors, including hand-crafted feature-based methods and current mainstream deep learning models, consist of three main components: proposal generation, feature extraction, and classification and regression. Traditional hand-crafted methods such as Histograms of Oriented Gradients (HOG) analyze the local gradients of an image to capture the texture and shape information, and then send the extracted HOG features to downstream classifiers (*e.g.* SVM or AdaBoost [11]) for pedestrian detection.

With the rapid development of Convolutional Neural Networks (CNN), the research field of general object detection has also made great progress. The R-CNN series of algorithms have transformed pedestrian detection into an object detection problem, and by adopting a combination of region proposal extraction and classifiers, have gradually improved the detection accuracy. ALF, based on Single Shot MultiBox Detector (SSD) [4], stacks multiple predictors together and inherits the high efficiency of single-shot detectors, learning better detection from default anchor boxes. To improve the detection performance for occluded pedestrians, MGAN utilizes additional information from the visible boundary boxes as the guidance for attention masks [12].

2.2 Loss Functions for BBR

The bounding box regression loss function is critical to determining the performance of target localization. YOLOv1 [7] uses mean squared error to reduce the influence of large targets on the bounding box. YOLOv3 [8] proposed creating a penalty term to reduce the competitiveness of large boxes. To address the gradient disappearance problem of IOU loss, GIOU [13] incorporates the minimum bounding box into the penalty term, while DIOU [14] considers the distance between targets and anchors. CIOU [14] adds a scale and aspect ratio penalty to DIOU. SIOU [15] additionally considers angle cost,which has a faster convergence rate and better performance.

2.3 Object Detector Architecture

Based on convolutional neural networks, object detectors can generally be divided into two types: (1) one-stage detector, such as YOLOv5 [10] and SSD [4]; (2) two-stage detector, such as R-CNN [2] and CenterNet2 [16]. However, both one-stage and two-stage detectors consist of three parts, namely, backbone for feature extraction, neck for feature processing and fusion, and head for class and bounding box prediction. Next, we will introduce the neck structure in detail.

Neck. The design of neck is to better utilize the features extracted by the backbone. It performs further processing and reasonable utilization on the features maps extracted at different stages. The earliest neck structure only adopted up-sampling and down-sampling modules, which is simple and does not require feature aggregation operations. Currently, the commonly used neck structures include: FPN [17] that aggregates features from top-down paths, PANet [18] that adds additional bottom-up paths based on FPN, BiFPN [19] that repeatedly performs top-down and bottom-up operations while merging single-node inputs, and NAS-FPN [20] that uses neural architecture search to automatically design a feature network.

3 Proposed Method

3.1 About YOLOV5

Among the YOLOv5 series, four models have been proposed. They are YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. YOLOv5s is the smallest model with only 7.2 M parameters, which is very suitable for deployment on mobile devices. In our approach, we consider using the S model for real-time performance, although the X model may show better performance. Typically, YOLOv5 adopts the architecture of CSPDarknet53 and the SPP layer as the backbone, PANet as the neck, and YOLO detection head.

3.2 BDF-YOLOV5

BDF. The structure of BDF-YOLOv5 is shown in Fig. 2, due to the limitations of PANet, which only facilitates a single cross-layer connection and fails to fully exploit the exchange of high-level and low-level semantic information. Furthermore, on datasets focused on dense pedestrian detection, instances of the targets exhibit significant scale variations and include a considerable number of small objects. To overcome these challenges, we have modified the neck portion of YOLOV5 with the proposed BDF Net for improved performance.This method employs a bidirectional fusion strategy that combines bottom-up and top-down network pathways, enabling bidirectional cross-layer information exchange in intermediate feature layers. Although this incurs additional computational burden, the model accuracy is significantly improved. (P2 and P5 layers are located at the lowest and highest levels, respectively, and only support unidirectional input).

Fusion Block. As shown in Fig. 3, we abandoned the original 3×3 fusion module C3 and instead adopted CSPNet in the feature fusion block. In the residual part, we used ResConv, which first passes through a 3×3 convolutional layer, a 1×1 convolutional

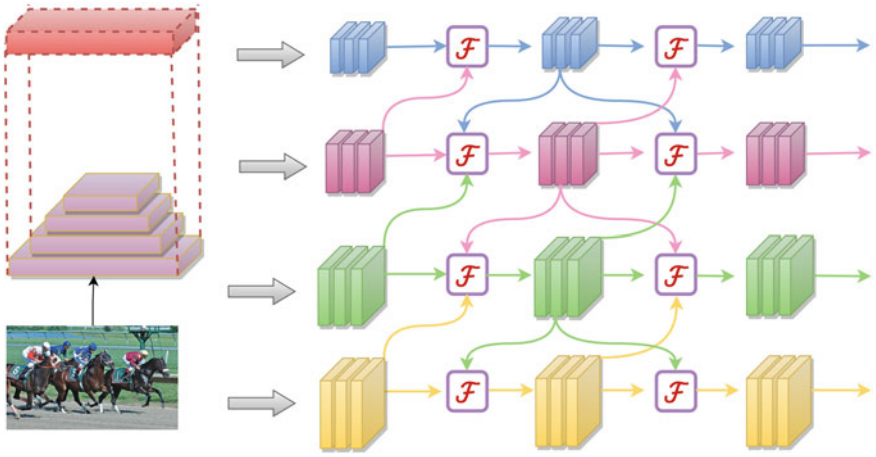


Fig. 2 The architecture of the BDF-YOLOv5. The BDF network as the neck to refine and fuse high-level semantic and low-level spatial features. In addition, a new fusion method is adopted

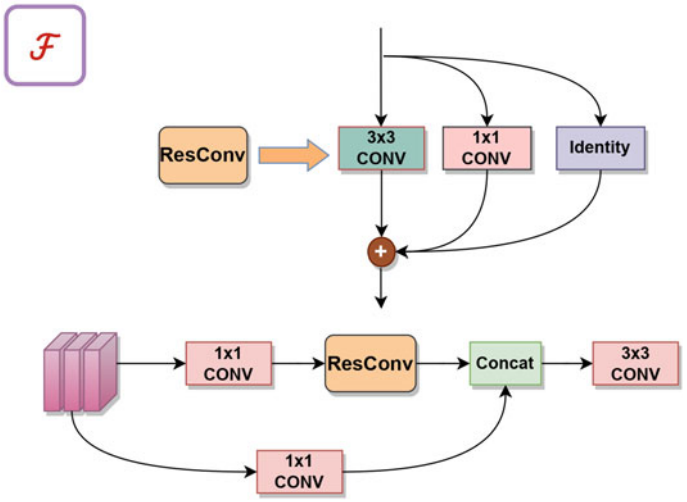


Fig. 3 The architecture of the fusion block

layer, and an identity layer, and finally adds them together directly along the channel dimension to achieve cross-layer connections.

Weighted CIOU. Most of the previous loss functions (GIOU, DIOU, CIOU) ignored the imbalance issue in Bounding Box Regression (BBR). As mentioned before, there inevitably exist a large number of low-quality samples (i.e., anchor boxes with little overlap with the target box) during the training process, which contribute the majority and thus reduce the sensitivity to high-quality regression samples during

optimization. Therefore, it is necessary to strengthen the contribution of high-quality regression boxes. In this paper, we added a weighting factor IOU^{γ} on the basis of the CIOU loss function,as the Eq. 1 shown. By judging the quality of regression boxes based on overlapping degree, we optimized the imbalance issue.

$$\phi_{CIOU} = IOU^{\gamma} \times (1 - IOU + \frac{\rho^2(b, b_{gt})}{C^2} + \alpha\upsilon). \tag{1}$$

The variables used in the Eq. 1 are defined as follows: b, b_{gt} denotes the center points of the predicted box and ground truth box, ρ denotes the Euclidean distance between the two center points, C represents the diagonal distance of the smallest closed area that can contain both boxes, υ is the similarity between the aspect ratios of the predicted box and ground truth box, and α represents the weight coefficient.

4 Experiments

4.1 Experimental Setting

The experiments were conducted in Python version 3.8 and all the models were trained and tested on an NVIDIA RTX3090 GPU. To ensure the reliability of the experiments, no pre-trained weights were used during training. The CrowdHuman [21] dataset was utilized in our experiments, which is a benchmark dataset to better evaluate detectors in crowd scenarios. The CrowdHuman dataset is large, richly-annotated, and contains high diversity.

4.2 Comparison Experiment

To demonstrate the advantages of the proposed method in this paper, we compared it with the original YOLOv5, YOLOv3, YOLOv5-lite, and SDD in terms of map0.5, map0.5:0.95, and detection speed (FPS). The results are shown in Table 1.

4.3 Ablation Studies

To further substantiate the effectiveness of our proposed improvements, we conducted several ablation experiments. Table 2 lists the effects of individual components as well as their combined impact, providing a comprehensive analysis of our network‘s performance.

Effect of the loss function. After optimizing the loss function, it was observed that the detection speed decreased from the original 107 FPS to 95 FPS, which is attributed to the introduction of additional computational steps. However, there was a remarkable improvement of 1.6 in mAP0.5, justifying the trade-off in terms of detection speed.

Effect of the BDF. By adopting our proposed BDF network, the number of layers of the original YOLOv5s has increased from 157 to the current 218, and the parameter count has also increased by 2.0 M.The sacrifice of model complexity has resulted in higher precision in detection results, indicating that our network plays a certain role in addressing the scale variance problem in objects detection.

Effect of model ensemble We integrated all our proposed improvements into a single model, as shown in Table 2. The final version achieved an improvement of 3.9 in testing performance, while maintaining a real-time detection speed of 86 FPS. These results demonstrate that our approach is capable of meeting the requirements for real-time detection applications.

5 Conclusion

In this paper, we proposed an advanced dense pedestrian detector based on YOLOv5. Throughout the experimentation process, we explored numerous feature extraction techniques, including but not limited to network architecture modifications. The results were satisfactory, with an improvement of nearly 4.0 over the original network.

Table 1 The comparison of the performance in crowdhuman dataaset

Methods	mAP0.5 (%)	mAP0.5:0.95 (%)	FPS
SSD	72.7	42.6	39
YOLOv3	83.6	50.6	51
YOLOv5-Lite	67.4	28.2	**138**
YOLOv5s	80.5	47.2	111
BDF-YOLOv5	**84.4**	**51.5**	86

The bold values are the best network model for the comparison of indicators in the ablation experiment

Table 2 Ablation Study:the impact of each added component

Methods	Parameters	mAP0.5 (%)	FPS
YOLOv5s	7.2M	80.5	111
YOLOv5 + weight-CIOU	7.2M	82.1 (↑ 1.6)	107
BDF-YOLOv5 + weight-CIOU	9.2M	83.6 (↑ 1.5)	95
BDF-YOLOv5 + weight-CIOU + fusion block	9.5M	84.4 (↑ 0.8)	86

We hope that this report will be helpful to the field of dense pedestrian detection. In the future, we plan to explore domain adaptation issues and further improve the performance of our proposed detector.

Acknowledgements Fund Project This work was supported by Tianjin Normal University Graduate Research Innovation Project Funding (2023KYCX005Y).

References

1. Girshick RB, Donahue J, Darrell T, Malik J (2013) Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE conference on computer vision and pattern recognition, pp 580–587
2. Girshick RB (2015) Fast r-cnn. In: 2015 IEEE international conference on computer vision (ICCV), pp 1440-1448
3. He K, Zhang X, Ren S, Sun J (2014) Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell 37:1904–1916
4. Liu W, Anguelov D, Erhan D, Szegedy C, Reed SE, Fu C-Y, Berg AC (2015) Ssd: single shot multibox detector. In: European conference on computer vision
5. Lin T-Y, Goyal P, Girshick RB, He K, Dollár P (2017) Focal loss for dense object detection. In: 2017 IEEE international conference on computer vision (ICCV), pp 2999–3007
6. Redmon J, Farhadi A (2016) Yolo9000: better, faster, stronger. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 6517–6525
7. Redmon J, Divvala SK, Girshick RB, Farhadi A (2015) You only look once: unified, real-time object detection. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 779–788
8. Redmon J, Farhadi A (2018) Yolov3: an incremental improvement. ArXiv, vol. abs/1804.02767
9. Bochkovskiy A, Wang C-Y, Liao H-YM (2020) Yolov4: optimal speed and accuracy of object detection. ArXiv, vol. abs/2004.10934
10. Jocher G, Stoken A, Borovec J, NanoCode012, Chaurasia A, TaoXie, Changyu L, Laughing AV, tkianai, yxNONG, Hogan A, lorenzomammana, AlexWang1900, Hajek J, Diaconu L, Marc, Kwon Y, oleg, wanghaoyang0106, Defretin Y, Lohia A, ml5ah, Milanko B, Fineran B, Khromov D, Yiwei D, Durgesh D, Ingham F (Apr. 2021) Ultralytics/yolov5: v5.0 - YOLOv5-P6 1280 models, AWS, Supervise.ly and YouTube integrations. [Online]. Available: https://doi.org/10.5281/zenodo.4679653
11. Lin Z, Pei W, Chen F, Zhang D, Lu G (2021) Pedestrian detection by exemplar-guided contrastive learning. IEEE Trans Image Process 32:2003–2016
12. Hasan I, Liao S, Li J, Akram SU, Shao L (2022) Pedestrian detection: domain generalization, cnns, transformers and beyond. ArXiv, vol abs/2201.03176
13. Rezatofighi SH, Tsoi N, Gwak J, Sadeghian A, Reid ID, Savarese S (2019) Generalized intersection over union: a metric and a loss for bounding box regression. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 658–666
14. Zheng Z, Wang P, Liu W, Li J, Ye R, Ren D (2019) Distance-iou loss: faster and better learning for bounding box regression. In: AAAI conference on artificial intelligence
15. Gevorgyan Z (2022) Siou loss: more powerful learning for bounding box regression. ArXiv, vol abs/2205.12740
16. Zhou X, Koltun V, Krähenbühl P (2021) Probabilistic two-stage detection. ArXiv, vol abs/2103.07461
17. Lin T-Y, Dollár P, Girshick RB, He K, Hariharan B, Belongie SJ (2016) Feature pyramid networks for object detection. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 936–944

18. Liu S, Qi L, Qin H, Shi J, Jia J (2018) Path aggregation network for instance segmentation. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 8759–8768
19. Tan M, Pang R, Le QV (2019) Efficientdet: scalable and efficient object detection. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 10 778–10 787
20. Ghiasi G, Lin T-Y, Pang R, Le QV (2019) Nas-fpn: learning scalable feature pyramid architecture for object detection. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 7029–7038
21. Shao S, Zhao Z, Li B, Xiao T, Yu G, Zhang X, Sun J (2018) Crowdhuman: a benchmark for detecting human in a crowd. ArXiv, vol abs/1805.00123

Research on Triple-Module Redundancy Computer with Reconfigurable Capacity

Yukun Chen, Jiangkang Wang, Dezhi Zhang, Gang Rong, and Yanchen Zhao

Abstract On-board computer system is the crucial component in spacecraft electronic system, and redundancy techniques can provide high reliability for on-board computer running. To ensure computer still work normally under fault condition, system architecture, principle and key technology of traditional and degraded triple-module redundancy computer were introduced. Some methods were adopted in order to reliably detect fault computer, on-board computer will work degradedly only which could not be renovated. Fault computer read current computer's pointer and crucial data at the beginning of every course, and sent synchronization require to current computer at the end of the course, then presented a design scheme of reconfigurable triple-module redundancy space on-board computer and designed reconfigurable flow. The scheme can make the fault computer has the ability of recovery. Practice indicates that the reconfigurable scheme can effectively improve the reliability of space on-board computer system, and the paper has engineering application value for design and implementation of space on-board computer system with high reliability.

Keywords Redundancy · Reconfigurable · Triple-module

1 Introduction

The reliability and security of On-board computer plays an important role in spacecraft because it provides measures to control aerospace. On-board computer system in space orbit has the feature of unmaintainability except for space station. Aircraft mission may be result in fail when On-board computer system has failure. Fault tolerance technology has become a urgent topic for on-board computer system to increase reliability.

Y. Chen (✉) · J. Wang · D. Zhang · G. Rong
Beijing Institute of Astronautical Systems Engineering, Beijing 100076, China
e-mail: cyk99811@163.com

Y. Zhao
Xi'an Aerospace Propulsion Institute, Xi'an 710100, Shanxi, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_53

2 The Triple-Module Redundant Architecture and Strategy

The architecture of triple-module redundancy can adopt three on-board computers or more than three on-board computers, and consist of three on-board computers at least. Redundancy capacity can be enhanced through two out of three principle, and the two out of three system base on the principle of majority. The correct results can be achieved when only one compute has failure in the redundant architecture, as depicted in Fig. 1. Isolation and reconfiguration are not carried out when the system has failure [1]. The system will be disabled when more than one computer has failure.

Assume the reliability of each machine is R_M, the reliability of voter is R_V, and the reliability of system is

$$R = \left(3R_M^2 - 2R_M^3\right) \times R_V$$

The key of the architecture is how to get two out of three. Figure 2 shows the principle block diagram for two out of three algorithm.

Voter can be implemented by hardware or software methods. Hardware voter has the future of simplicity and rapidness. The voter failure will result in system output failure because voter is a single point [2]. The voter will become complex as output channels increase, then the reliability of voter will reduce. Software voting and hardware gating are introduced to solve the problem. The method increases mutual communication between systems to exchange voting message, besides the correct

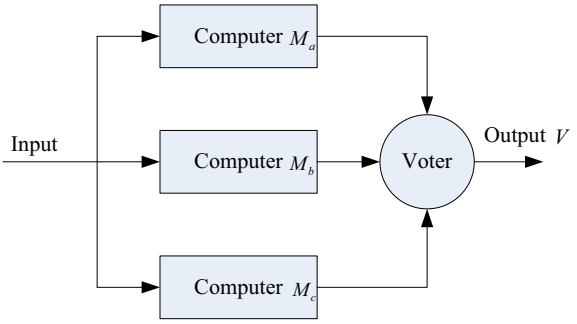


Fig. 1 The typical fault tolerance architecture of triple-module

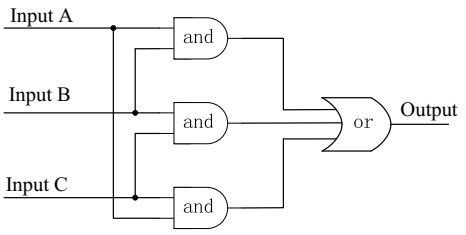


Fig. 2 The principle block diagram for two out of three algorithm

state voting strategy circuit is necessary [3]. Software voting increases time cost, and synchronization is also a indispensable problem.

3 The Triple-Module Redundant Architecture with Degraded Function

Triple-module redundant architecture is widely applied in control system. To increase system reliability and resource availability, the triple-module redundant architecture with degraded function can be presented according to system redundant strategy. Based on the design idea that triple-module computers voting when system work normally and one computer voting when system work abnormally, it can increase system reliability and enhance operating life [4]. The main feature is the architecture of triple-module, failure test and judgement is accomplished by arbitration and message exchange, and the normal computer result is the final output of whole system.

Triple-module redundant architecture with degraded function consists of three identical hardware computers running identical programme. Each computer has the same processor. The input of triple-modulecomputers is one to three, and the output is arbitration management control circuit consisted of hardware and software. Arbitration management control circuit allows that only one computer is output each time [5]. Figure 3 illustrates triple-module redundant architecture with degraded function.

Except for three identical redundant computers, the triple-module redundant architecture with degraded function has the following components and crucial technique:

(1) Communication among triple-module computers. Communication mode is full duplex by serial port.
(2) System synchronization. Synchronization has two levels, one is macroscopic period synchronization, the other is microscopic synchronization. Macroscopic synchronization is achieved by adopting unified clock frequency time circuit. The reliability of circuit has no relation with triple-module computers, and it also has redundant measures to ensure reliability. Microscopic synchronization is the synchronization in a period to ensure current computer is not influenced by the other two computers.
(3) State output. Each computer output its state message to arbitration management control circuit, then computer on duty can be present through logic judgement.
(4) Arbitration management control circuit. Basing on control instruction, arbitration management control circuit votes on duty signal through state signal. The on duty computer is system output by on duty signal.

Arbitration management control circuit releases one computer output through all state message or control instruction. Triple-module computers output state message through two out of three voting. One computer is choosing as output autonomously through triple-module computers state message, and other two computers is shut

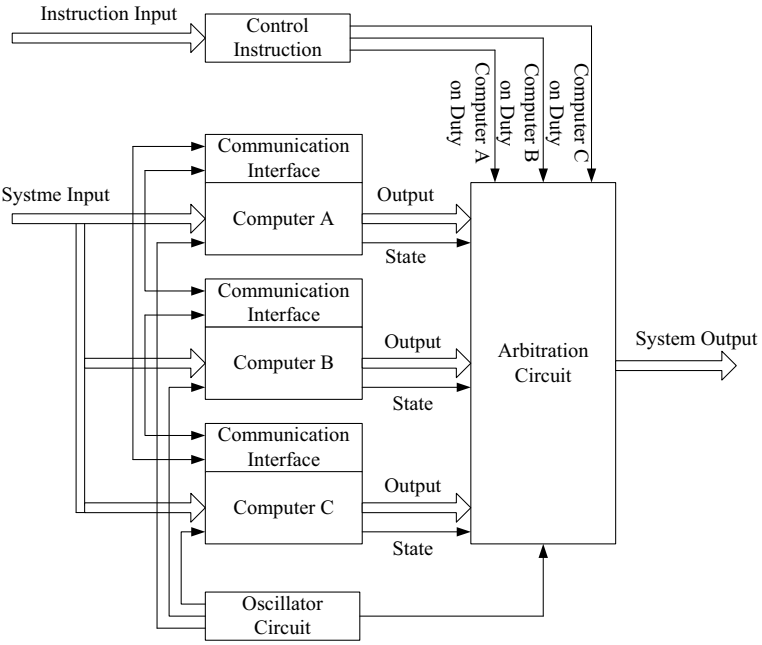


Fig. 3 The triple-module redundant architecture with degraded function

down [6]. Control instruction can also achieve the goal of releasing one computer output, at the same time the other two computers is shut down. Microprocessor on computer is responsible for data exchange among triple-module computers, and the physical link for data exchange is accomplished by serial port.

4 The Design of Triple-Module Redundancy Space On-Board Computer with Reconfigurable Capacity

Figure 3 displays that triple-module computers has the same and equal architecture, and any two computers between three computers has communication interface. To make full use of system resource, triple-module redundant architecture should not only degrade to one computer under definite conditions. System configuration can be achieved through adopting some strategy to recover fault computer. Static redundancy can be transferred to dynamic redundancy through measures of fault detection, fault location and fault recovery, then system resume to work normally and enhance system reliability and service life further.

4.1 The Fault Detection Mechanism

The symbol of triple-module computers is respectively A computer, B computer and C computer. The input signal and output signal of triple-module computers is independent, so the fault of one computer will not influence the other computers, and the on duty computer still can achieve all functions [7]. System has not single point failure from the view of interface and function. To reliably detect fault computer, single computer can still work normally under fault condition, and control right is acquired by self inspection, commutative inspection, other inspection.

(1) Self inspection. All output signal is marked through software. The computer will send abnormal feedback if it finds abnormal output.
(2) Commutative inspection. The voting FPGA of A computer monitors the real-time heartbeat signal of B computer's voting FPGA, C computer's voting FPGA and itself. The voting FPGA of B computer monitors the real-time heartbeat signal of A computer's voting FPGA, C computer's voting FPGA and itself. The voting FPGA of C computer monitors the real-time heartbeat signal of A computer's voting FPGA, B computer's voting FPGA and itself. State signal can be exchanged among three voting FPGA. Three CPU can acquire state signal and synchronization information of the other two CPU through respective interactive buffer, so commutative inspection can be achieved among three CPU.
(3) Other inspection. Three CPU send respectively data and state parameter to dual RAM of three voting FPGA, so that they can read triple-module computers data from corresponding voting FPGA to vote based on two out of three principle, then they dispatch vote results to three voting FPGA unit. Three voting FPGA feed back to three CPU after voting again based on two out of three principle, so that they can monitor respectively CPU's state.

If one voting FPGA cannot receive corresponding CPU data, it estimate the CPU has failure and send information to the other two CPU, and the other two CPU will degrade to dual hot standby mode. Three computer's CPU monitor power state signal of the two computer. If power has failure, triple-module computer estimate one computer has fault and transfer to dual computer mode.

4.2 The Reconfigurable Design of Fault Computer

If triple-module computer estimates one computer has software failure, the normal two computer will dominate output. At the same time it send reset signal to fault computer through reset pin on the internal bus, so that fault computer CPU can be reset [8]. The normal two computer degrade to dual hot standby mode, one computer is recognized to on duty. Triple-module computers are reconfigured based on the duty computer. If the fault computer cannot recover, the CPU will shut off all the output signal of the fault compute, and triple-module computer degrade to dual hot

standby mode. Figure 4 illustrates the reconfiguration flow diagram of triple-module redundancy mode.

As depicted in Fig. 4, if one computer's hardware fault happen many times in the process of triple-module computer voting based on two out of three principle, on duty computer will shut off power and restart power to resume fault computer, at the same time, on duty computer launches dual hot standby mode according to power-off state. After fault computer resumes, the method of implementing triple-module mode is as follows: at the beginning of each process on duty computer sends PC pointer and critical parameter to voting FPGA of fault computer, and after initialization fault computer begins to read data from on duty computer and

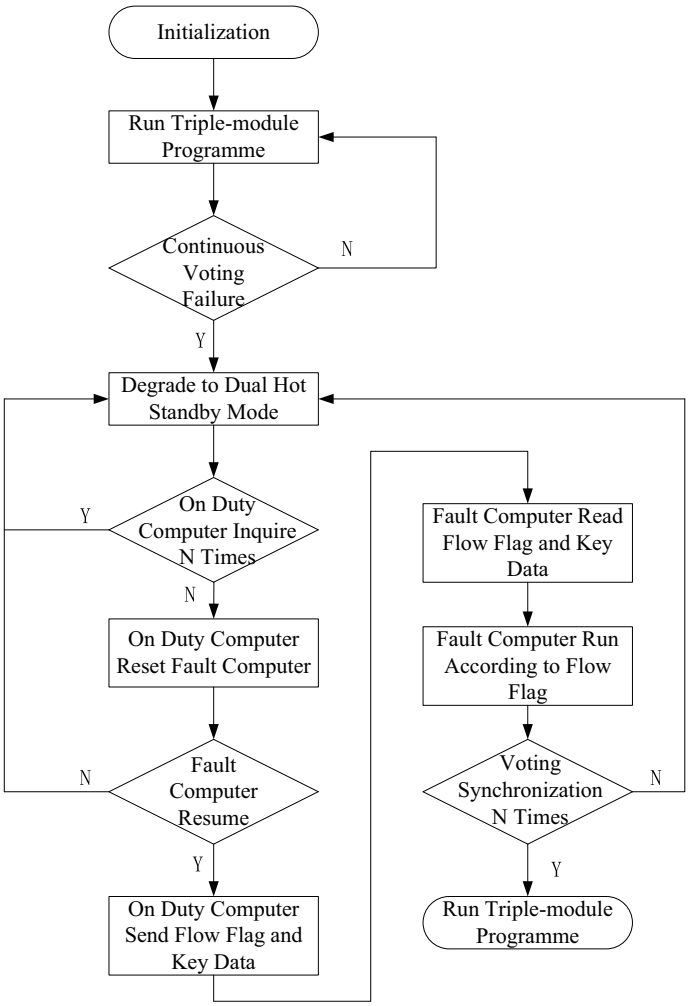


Fig. 4 The reconfiguration flow diagram of triple-module redundancy

immediately enters synchronization mode, at the same time fault computer sends operating data to voting FPGA of the other two computer. After reading on duty computer synchronization data, at the end of each process fault computer sends synchronization request to on duty computer through heartbeat monitor channel. After receiving synchronization request of fault computer, on duty computer reads data of triple-module computer and labels on data of fault computer, then it votes based on two out of three principle. If voting has the same result, on duty computer sends resumed triple-module synchronization signal through heartbeat signal, and it recovers triple-module computer mode.

When on duty computer has failure during dual hot standby, control right can be switched between dual on-board computer through remote control and switching autonomously [9]. When flight control centre estimates that current computer has failure according to telemetry data, control right can be switched between dual redundant computer by remote control command. When remote control mode takes into effect, autonomous switching is shut down, then output of the dual redundant is determined only by remote command. To shut down autonomous switching, permitting or forbidden time window of autonomous switching is set by remote command. Only when aircraft is in autonomous switching state, autonomous switching is permitted for on-board computer. In autonomous switching state, backup will take into effect when host has failure. Autonomous switching right is achieved by integral circuit to avoid accomplishing only by a piece of command. Switching command must be sent continuously many times, a certain level of integral circuit must be achieved to drive relay switching, and then backup computer will be on duty.

5 Conclusions

The reliability and security of on-board computer is the key component for aircraft. The paper displayed architecture of triple-module fault-tolerant computer system with degraded function, and introduced a design method of triple-module redundancy on-board computer with reconfigurable capacity, finally presented dynamic recovery flow diagram. Practice indicates that the triple-module computer with reconfigurable capacity can effectively enhance the reliability of space on-board computer system. The measure has engineering application value for design and implementation of space on-board computer system with high reliability under space harsh environment.

References

1. Yang M, Hua G, Feng Y (2014) Fault tolerance techniques for spacecraft control computer. National Defense Industry Press, Beijing
2. Duan Y, Wang J, Sun L (2022) Redundancy and fault-tolerant design of atmospheric data for UAV autonomous flight. J Proj Rocket Missiles Guid 42(6):75–78

3. Sun X, Chen Z, Gu Y (2018) Research on fault-tolerant flight control computer system based on dynamic reconfiguration. J Syst Simul 30(10):3957–3963
4. Wang Z, Wang M, An S (2022) Design and implementation of a fault-tolerant computer with 2×2 architecture. Ind Control Comput 25(5):44–45
5. Wang J, Wang S, Wang X (2018) Fault mode probability factor based fault-tolerant control for dissimilar redundant actuation system. Chin J Aeronaut 31(5):965–975
6. Zhang Z, Du J, Fan C (2021) The redundancy management in redundant fly control computer of launch vehicle. Flight Control Detect 4(3):48–56
7. Wang Z, Cheng S, Ma X (2020) Design and implementation of highly reliable fault-tolerant computer with integrated multi-task. Aeronaut Comput Techn 50(4):110–112
8. Lv Y (2020) A fault-tolerant method for space computer memory with low-cost and high-reliability. Aerosp Control Appl 46(3):66–70
9. Jiang B, Zhang K, Yang H (2021) Fault-tolerant control of satellite attitude control systems. Acta Aeronaut Astronaut Sin 42(11):524662

Research on Design and Application Methods of Digital Satellite

Chen Gang, Liang Jian, Liu Shuai, and Ruan Jiangdong

Abstract As the most commonly used spacecraft, satellite has developed rapidly in recent years, but it is also faced with a series of technical problems. In order to achieve the high quality, high efficiency, high reliability and low cost manufacturing of satellites, digital twins can be combined to realize the digitalization of satellites, and provide guidance and early warning during the design, manufacturing and use of satellites. This paper introduce the research status and key technologies of digital satellite, and forecast the application prospect of digital satellite.

Keywords Digital satellite · Simulation · Model · MBSE

1 Introduction

Recently, satellite has become the most widely used aerospace products, and it shows multi-dimensional trends such as low cost, diversified functions, intelligence, batch production and batch deployment. At the same time, satellites also faced the development problems of high-performance, low-cost, high efficiency, high quality, smarter on-orbit independent operation and maintenance, evaluation, life prediction and life extension.

With the development of new technologies, such as digital twins, big data, intelligent interconnection, it provides a promising solution to solve the problems. The combination of digital twins and satellite will have unique development advantages. Digital twins is a high-fidelity simulation for complex products with multi-physics, multi-scale and multi-fields, which can faithfully map and interact with the physical world to improve the performance and operation performance of its corresponding

C. Gang
Lunar Exploration and Space Engineering Center, 65 Zhichun Road, Beijing, China

L. Jian (✉) · L. Shuai · R. Jiangdong
China Academy of Space Technology, 104 Youyi Road, Beijing, China
e-mail: liangjiancast@163.com

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_54

physical entities [1]. Digital satellite is a multi-disciplinary, multi-physical and multi-scale twin model, and maps the whole life state of the physical satellite in digital space in real time [2]. It is a collection of digital models with the same origin, data sharing, co-evolution, real-time interaction and extended survival.

Due to the lack of system-level test and verification environment, some system defects will be moved back to the operation stage [3]. Thus, it is necessary to provide a digital test environment that can meet the combination of software and hardware. By means of design verification, fault simulation, digital virtual testing and iterative optimization, problems in satellite development can be found and solved, greatly improving development efficiency and saving R&D costs. Based on the simulation characteristics of digital twins, the input of different external scenes can be connected to realize the simulation output of different scenes. Therefore, typical application scenarios such as fault location, digital flying companion, health management, mission planning rehearsal can be achieved by digital twin. When the digital satellite model meets the delivery requirements, it can be embedded into the system simulation system as a plug-in to assist the system-level simulation application [4, 5].

2 Composition of Digital Satellite

The composition of the digital satellite system is shown in Fig. 1, which is divided into four levels. The basic layer is composed of relevant standards, specifications, software and hardware facilities. Relying on the basic layer, we build the satellite professional model base, basic supporting model base, algorithm base and database, and the digital satellite base can be formed through model inheritance and management. In order to achieve effective scheduling of models, a satellite digital twins basic support platform should be built, including simulation modeling parameter configuration subsystem, data integrated management subsystem, data access and drive subsystem, two-dimensional/three-dimensional visualization subsystem and application scenario simulation subsystem. Finally, the function of each subsystem is integrated to achieve the scene simulation, digital flying, test verification and mission effectiveness evaluation of digital satellite.

2.1 Digital Satellite Simulation Platform

Digital satellite simulation platform includes hardware platform and simulation engine. Hardware platform provides basic hardware resources such as computing, storage and network. Simulation engine provides simulation control service, model communication service, time unification service, simulation monitoring service, and various simulation models. At the same time, through user rights management, model resource management, design resource management, operation data management and

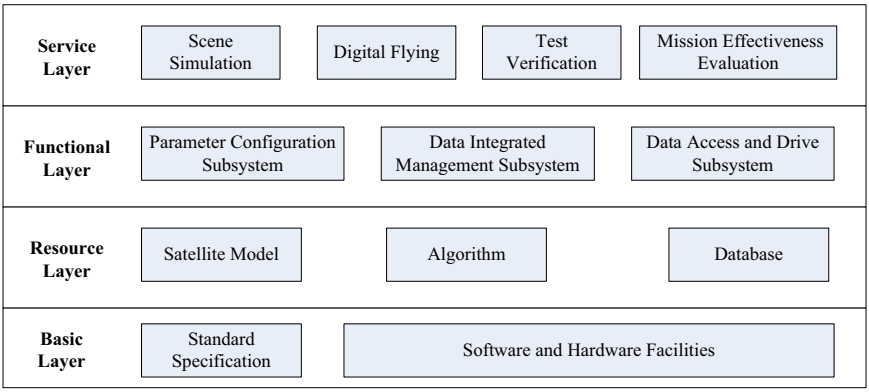


Fig. 1 The composition of the digital satellite

computing node management, platform provide users with a convenient and efficient simulation environment.

Hardware foundation is divided into two parts: computing resources and storage resources, which are used to complete the efficient calculation of complex data, as well as the storage of simulation process data and final calculation results, and finally realize the function of high-performance parallel task processing. Typical hardware products include standard rack server or blade server, high-density server, GPU accelerator card, SSD accelerator card, compression accelerator card, etc.

Simulation engine completes the generation, collection, feature processing, training support and other processes of sample data. It is the core support for training, evaluation, deployment and application. It also provides the management, monitoring and scheduling of hardware system resources, and supports the development and application of digital satellites.

2.2 Satellite Model

The precise mapping of satellite physical entities to digital satellites should base on models and data, and take full advantage of interactive feedback and iterative operation of the multi-disciplinary high-precision modeling, mechanism data fusion and other methods to achieve the synchronous presentation of satellite physical entities in the digital space. The following models need to be established for satellite system.

Attitude and orbit control system model provide initialization and startup simulation methods, simulate satellite ultra-high precision attitude determination ability and high precision attitude control ability, and consider the influence factors of orbit perturbation, to achieve the precise modeling and prediction of satellite orbit parameters.

Storage and data balance system model provide the constraints of satellite imaging dynamic generation data and on-board storage data.

Energy balance constraint model provide the dynamic balance process of energy supply, consumption and battery storage in the whole process of satellite on-orbit operation.

Environmental model is composed of land, sea, day and night, clouds and used to simulate the space environmental system.

Target model include sea and land scenes and typical targets.

Ground TT&C node model: ground TT&C node model to simulate the operation status and capability of ground TT&C station nodes.

2.3 Simulation Scenario Design and Effectiveness Evaluation

Simulation scenario design is the specific setting of the operation process which based on the simulation purpose and the corresponding domain knowledge. It is the description of the envisaged region, environment, objectives and events within the specified time range related to the events concerned. Extensive Markup Language (XML) is widely used as the carrier of simulation scenarios. It is a standardized, structured, formal expression method and has good scalability. The simulation scenario is the input file of a single simulation task. It specifies the status and task information of each subsystem of the digital satellite in the form of standardized XML tags to determine the simulation task and boundary conditions, so that the simulation can be carried out smoothly.

Simulation scenario design mainly completes the loading and editing of scene files, the setting of scene basic parameters, meteorological information and natural environment information, the selection of battlefield reconnaissance system, target working parameters, and the input of airborne platform and target motion parameter information through 3D visualization. Scenarios build a function to realize scenario management and editing in a visual way, set attribute parameters such as environment, platform, link, track and task in the scenario task, write scenario information into the XML scenario file, and the simulation platform parses the XML to quickly generate the simulation, which greatly improves the initialization of task verification. The simulation scenario design system is composed of scenario management module, terrain loading and geographic information management module, environment setting module, network planning module and task planning module.

The effectiveness evaluation system is mainly composed of effectiveness evaluation subsystem, multi-dimensional evaluation algorithm model library, and experimental data from multiple sources, it provides the construction method of visual indicator system for system-of-system capability and typical mission effectiveness evaluation, and provides classic hierarchical analysis method, fuzzy comprehensive evaluation method, gray correlation analysis equivalent energy evaluation model.

System evaluation results can be visualized in multiple dimensions, and provide effect analysis and system optimization scheme design to support the top-level design and optimization of the system. Digital satellite application.

3 Application of Digital Satellites

3.1 Satellite Mission Planning and Arrangement

Faced with the increasingly complex operational environment, structure, professional coupling and other system engineering problems of satellite systems, even a small mistake may cause a chain reaction and cause huge losses. Therefore, digital satellite can provide the modeling, scheduling, simulation verification and evaluation environment for mission planning in advance to ensure the implementation and correctness of the physical object, and improve the safety and reliability of the satellite.

Digital satellite supports the combination of mission planning results and system function models, and carries out comprehensive simulation verification to further judge the feasibility and evaluation level of the planned mission.

3.2 Digital Accompanying Flight

Based on the satellite operation data acquisition and real-time driving digital twins, the operation status of the physical system that has not been collected can be reproduced. Thus, we could comprehensive monitor the physical system and expand the boundary of the physical system. At the same time, based on the synchronous simulation data and physical data, the operation status of the system can be monitored and compared in real time.

4 Conclusion

This paper summarizes the current research status of digital satellite technology, introduces the composition of digital satellite system and related key technologies, including digital satellite simulation platform, satellite professional model, digital satellite simulation scenario design and effectiveness evaluation. Finally, this paper forecast the application of digital satellite.

References

1. Shangguan D, Chen L, Ding J (2020) A digital twin-based approach for the fault diagnosis and health monitoring of a complex satellite system. Symmetry 12(8):1307. https://doi.org/10.3390/sym12081307
2. Wang J, Jing F, Xia Y et al (2020) Study on model-based system engineering in satellite layout. MATEC Web Conf 316:04001. https://doi.org/10.1051/matecconf/202031604001
3. Reil RL (2014) Improved traceability of a small satellite mission concept to requirements using model based system engineering. In: Annual AIAA/USU conference on small satellites
4. Anyanhun AI, Edmonson WW (2018) An MBSE conceptual design phase model for inter-satellite communication, pp 1–8. https://doi.org/10.1109/SYSCON.2018.8369575
5. Fischer PM, Lüdtke D, Lange C et al (2016) Implementing model based system engineering for the whole lifecycle of a spacecraft. In: Deutscher Luft-und Raumfahrtkongress

Learning Social Constraints for Human Trajectory Prediction

Jianglin Zhou, Qi Xue, Jie Ren, Shuang Liu, Zhong Zhang, and Peng Guo

Abstract Human trajectory prediction is essential for ensuring the safe navigation and scenario interaction for autonomous vehicles and mobile robots. Human trajectory is influenced not only by the human itself but also by the constraints of the surrounding objects. Hence, modeling the exact social constraints is necessary. In this paper, we classify the social constraints methods into three kinds, i.e., LSTM-based methods, transformer-based methods and GCN-based method, to predict human trajectories. In addition, we compare their performance on ETH/UCY and SDD, and the experimental results display the superior performance of the transformer-based methods.

Keywords Human trajectory prediction · Social constraints · Transformer-based methods

1 Introduction

Human trajectory prediction [1–3] aims to predict the future positions based on the previously observed positions, which is usually considered as the time series problem [4–6]. It can effectively mitigate potential collision risks and has a broad application in intelligent autonomous vehicles [7] and mobile robots [8], etc.

Unlike other object movements with rule constraints, humans exhibit highly stochastic behavior as they have the freedom to stop, turn, or interact with their neighbors. In order to model human trajectories, some methods [9–11] learn spatio-temporal information, i.e., the social constraints among humans. For instance, Zhang et al. [9] utilize state-refined Long Short-Term Memory (LSTM) [12] network to

J. Zhou · Q. Xue · J. Ren · S. Liu · Z. Zhang (✉)
Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
e-mail: zhong.zhang8848@gmail.com

P. Guo (✉)
Catarc (Tianjin) Automotive Engineering Research Institute Co,. Ltd, Tianjin, China
e-mail: guopeng@catarc.ac.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_55

extract the social constraints among different humans. Zhong et al. [10] exploit the spatio-temporal transformer [13] to capture the spatio-temporal information of humans separately. Bhujel et al. [11] introduce Disentangled GCN (DGCN), a method that captures social constraints by separating spatial and temporal factors through Graph Convolutional Networks (GCNs) [14].

In this paper, we classify the existing methods for human trajectory prediction by capturing the social constraints. More precisely, these methods are broadly categorized into three groups, i.e., LSTM-based methods, transformer-based methods, and GCN-based methods. Afterwards, we compare the performance of the above methods on ETH/UCY [15] and Stanford Drone Dataset (SDD) [16].

2 Social Constraints Methods

In this section, we review the existing methods for capturing the social constraints, including LSTM-based methods, transformer-based methods, and GCN-based methods. At any time step t, the i-th human is denoted by the xy-coordinates (x_i^t, y_i^t). Moreover, we observe the positions of N humans from $t = 1, ..., T_{obs}$ and predict their future positions from $t = T_{obs} + 1, ..., T_{pred}$.

2.1 LSTM-Based Methods

The LSTM-based methods [17–19] utilize LSTM to learn temporal information, and apply the pooling mechanism to capture the social constraints in crowded scenes so as to predict human future positions. The detailed processes of LSTM-based methods are shown in Fig. 1.

Firstly, LSTM is employed to encode the input human trajectories. Meanwhile, the temporal information across different time steps is captured by LSTM to obtain the hidden state:

$$h_i^t = \text{LSTM}\left(h_i^{t-1}, c_i^{t-1}; e_i^t; W_l\right) \tag{1}$$

where c_i^t denotes the cell state and W_l represents parameter matrix. At time step t, the position of the i-th human is embedded as $e_i^t = \phi\left(x_i^t, y_i^t; W_e\right)$, where ϕ is the embedding function parameterized by W_e. Then, LSTMs are connected to each other through a pooling layer. The pooling layer captures the spatial information among different humans and integrates it with the temporal information to obtain the social constraints H_i. Afterwards, H_i is applied to renew the hidden state of the LSTM, thus obtaining the future position $\left(\hat{x}_i^{t+1}, \hat{y}_i^{t+1}\right)$ of the target human:

$$(\hat{x}_i^{t+1}, \hat{y}_i^{t+1}) \sim \mathcal{N}\left(\mu_t^i, \sigma_t^i, \rho_t^i\right) \tag{2}$$

where μ_t^i, σ_t^i and ρ_t^i denote the mean, standard deviation and correlation coefficient of the bivariate Gaussian distribution, respectively.

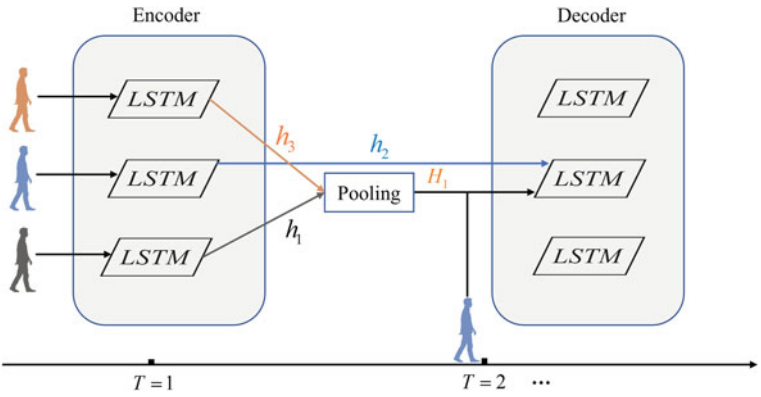


Fig. 1 The flowchart of LSTM-based methods

2.2 *Transformer-Based Methods*

Transformer achieves remarkable results in Natural Language Processing(NLP) [20–22] field, and naturally, researchers apply transformer to human trajectory prediction. Yu et al. [23] utilize interleaved temporal transformer and spatial transformer to learn the social constraints so as to predict human trajectories. Yin et al. [24] employ a multimodal transformer network based on pure attention mechanism to extract the social constraints. Moreover, Zhang et al. [25] combine the social constraints with human intentions to predict future positions. The flowchart of these methods is summarized as Fig. 2.

The transformer-based methods adopt an encoder-decoder structure as shown in Fig 2. During the encoding stage, the human trajectory is first mapped to higher dimensions via a fully connected (FC) layer. After that, the spatio-temporal information are extracted by the spatial and temporal transformers, respectively. This process can be formulated as follows:

$$S = ATT(E^t) = \frac{\text{softmax}[E^t W_q (E^t W_k)^T]}{\sqrt{d}} E^t W_v \tag{3}$$

$$T = ATT(E_i) = \frac{\text{softmax}[E_i W_u (E_i W_y)^T]}{\sqrt{d}} E_i W_z \tag{4}$$

where S and T represent the learned spatial and temporal information, respectively. Here, $E^t = \{e_i^t, e_2^t, ..., e_N^t\}$, $E_i = \{e_i^1, e_i^2, ..., e_i^{T_{obs}}\}$, $\sqrt{d}$ is used for normalization to ensure stability of the values, and W_q, W_k, W_v, W_u, W_y, and W_z denote the different parameter matrices. Afterwards, S and Z are concatenated and used as input to the decoder. During the decoding stage, the future position is obtained by the FC layer:

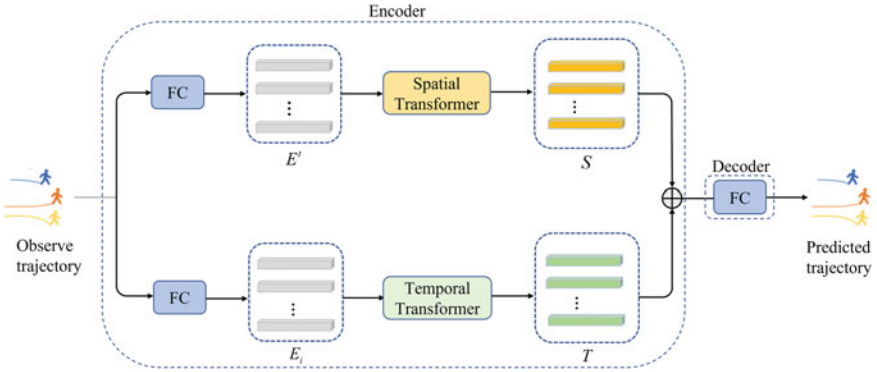


Fig. 2 The flowchart of transformer-based methods

$$(\hat{x}_i^{t+1}, \hat{y}_i^{t+1}) = \phi\left([Z||S]; W\right) \tag{5}$$

where $[\cdot||\cdot]$ indicates concatenate operation, and ϕ is the FC layer parameterized by W.

2.3 GCN-Based Methods

Treating the human positions as nodes and the complex social constraints among humans as edges of the graph can replace the conventional aggregation method for modeling human trajectories. For example, Mohamed et al. [26] construct a spatio-temporal graph to capture the social constraints and temporal dynamics among humans, enabling the prediction of future trajectories at once. Li et al. [27] introduce a spatial-temporal consistency network (STC-Net) based on the graph method, where the extended temporal convolution is utilized to mine the temporal information while the graph convolution is employed to capture the spatial information. Furthermore, Lian et al. [28] combine the spatial graph convolutional network (S-GCN) with the temporal transformer network (T-transformer) to mine the social constraints and use the temporal extrapolator convolutional neural network (TXP-CNN) to decode and obtain human feature positions.

Figure 3 generalizes the GCN-based method flowchart. Firstly, the input trajectory is constructed as a graph, denoted as $G = (V, E)$, in which V represents the set of nodes and E represents the set of edges. In addition, the adjacency matrix A is the set of $\{A_1, A_2, ..., A_{T_{obs}}\}$. Then, S-GCN extracts the spatial information from the input graph, followed by the temporal convolutional network (TCN) [29]. Specifically, the monolayer S-GCN [27] can be implemented by the following equation:

$$f(V^{l+1}, A_t) = \psi\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}V^l \mathrm{W}^l\right) \tag{6}$$

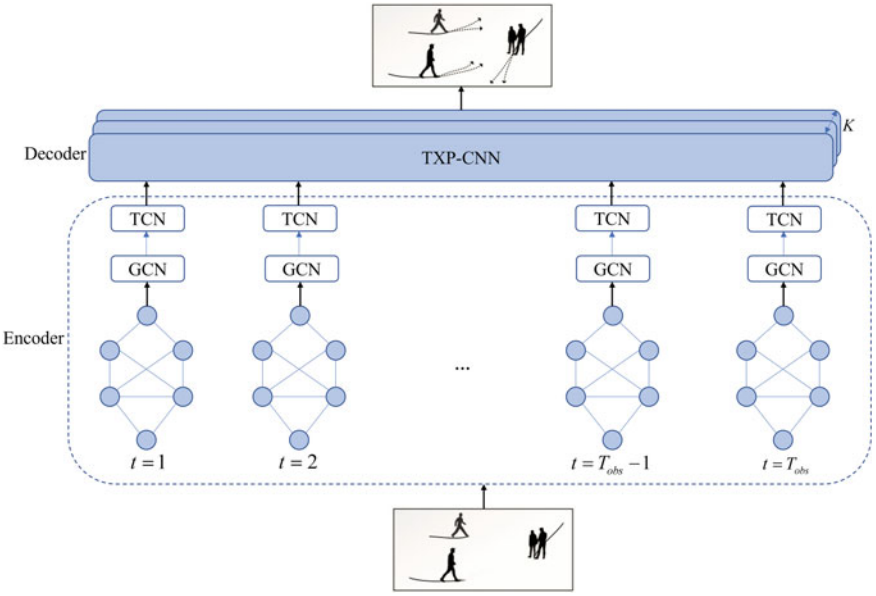


Fig. 3 The flowchart of GCN-based methods

where $\tilde{A} = I + A_t$ is the adjacency matrix, $\tilde{D}$ is the degree matrix, W^l is the parameter matrix of the l-th layer, and ψ is the activation function. Finally, TXP-CNN is used to decode and predict the future trajectory.

3 Experiments

In this section, we provide a brief introduction to the evaluation indicators and the human datasets. Afterwards, we evaluate the performance of the aforementioned methods.

We utilize the Average Displacement Error (ADE) and the Final Displacement Error (FDE) as the evaluation indicators:

$$ADE = \frac{\sum_{i=1}^{N} \sum_{t=T_{obs}+1}^{T_{pred}} \left\| (x_i^t, y_i^t) - (\hat{x}_i^t, \hat{y}_i^t) \right\|_2}{N \times [T_{pred} - (T_{obs} + 1)]} \tag{7}$$

$$FDE = \frac{\sum_{i=1}^{N} \| (x_i^{T_{pred}}, y_i^{T_{pred}}) - (\hat{x}_i^{T_{pred}}, \hat{y}_i^{T_{pred}}) ||_2}{N} \tag{8}$$

We make some evaluations on ETH/UCY [15] and SDD [16]. ETH/UCY contains five different scenes, which contain human trajectories in different environments.

Table 1 Compare different methods on ETH/UCY. Lower value (in meters) is better

Evaluation metrics (ADE ↓ / FDE ↓)						
	ETH	HOTEL	UNIV	ZARA1	ZARA2	AVG
Social-LSTM [17]	1.09/2.35	0.79/1.76	0.67/1.40	0.47/1.00	0.56/1.17	0.72/1.54
SR-LSTM [19]	0.63/1.25	0.37/0.74	0.51/1.10	0.41/0.90	0.32/0.70	0.45/0.94
STAR [23]	0.36/0.65	0.17/0.36	0.31/0.62	0.26/0.55	0.22/0.46	0.26/0.53
ForceFormer [25]	0.36/0.52	0.09/0.14	0.21/0.42	0.15/0.22	0.12/0.20	0.19/0.30
Social-STGCNN [26]	0.64/1.11	0.49/0.85	0.44/0.79	0.34/0.53	0.30/0.48	0.44/0.75
STC-Net [27]	0.64/1.18	0033/0.54	0.39/0.74	0.29/0.49	0.26/0.45	0.38/0.68
PTP-STGCN [28]	0.36/1.04	0.34/0.45	0.48/0.87	0.37/0.61	0.30/0.46	0.42/0.68

Table 2 Compare different methods on SDD. Lower value (in pixels) is better

Evaluation metrics (ADE ↓ / FDE ↓)			
	Social-LSTM [17]	VIKT [10]	STC-Net [27]
ADE	57.00	12.59	11.88
FDE	31.20	23.15	20.08

SDD is a comprehensive dataset that captures the movement of human crowds from a bird's-eye perspective. We observe the first 8 frames of the human trajectory and predict the future trajectory of the next 12 frames.

We can draw some conclusions from Tables 1 and 2. Firstly, the transofmer-based methods achieve better performance compared to the remaining two kinds of methods. Specifically, ForceFormer achieves 0.19/0.30 on ADE/FDE. Secondly, the LSTM-based method underperforms, because LSTM treats pedestrians within neighborhood areas equally and cannot focus on highlighting the social constraints among different humans. Thirdly, the graph-based method achieves superior results on SDD, which indicates the great potential of the graph-based methods.

4 Conclusion

In this paper, we have classified the social constraints, including LSTM-based methods, transformer-based methods, and GCN-based methods. Moreover, we have evaluated the performance of the above methods on the ETH/UCY and SDD datasets. The experimental results show that the transfomer-based method achieves higher performance.

Acknowledgements This work was supported by National Natural Science Foundation of China under Grant No. 62171321, Natural Science Foundation of Tianjin under Grant No. 22JCQNJC00010, the Scientific Research Project of Tianjin Educational Committee under Grant No. 2022KJ011, and the Tianjin Research Innovation Project for Postgraduate Students under Grant No. 2022SKY256.

References

1. Nasr Esfahani H, Song Z, Christensen K (2023) A deep neural network approach for pedestrian trajectory prediction considering flow heterogeneity. Transp Transp Sci 19(1):2036262
2. Zhou H, Ren D, Yang X, Fan M, Huang H (2023) CSR: cascade conditional variational auto encoder with socially-aware regression for pedestrian trajectory prediction. Pattern Recog 133:109030
3. Huynh M, Alaghband G (2023) Online adaptive temporal memory with certainty estimation for human trajectory prediction. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision, pp 940–949
4. Wu Y, Bighashdel A, Chen G, Dubbelman G, Jancura P (2023) Continual pedestrian trajectory learning with social generative replay. IEEE Rob Autom Lett 8(2):848–855
5. Xu D, Shang X, Peng H, Li H (2023) MVHGN: Multi-view adaptive hierarchical spatial graph convolution network based trajectory prediction for heterogeneous traffic-agents. IEEE Trans Intell Trans Syst 1–10.https://doi.org/10.1109/TITS.2023.3248090.
6. Hasan F, Huang H (2023) MALS-Net: a multi-head attention-based LSTM sequence-to-sequence network for socio-temporal interaction modelling and trajectory prediction. Sensors 23(1):530
7. Ballan L, Castaldo F, Alahi A, Palmieri F, Savarese S (2016) Knowledge transfer for scene-specific motion prediction. In: Proceedings of the European conference on computer vision, pp 697–713
8. Marchetti F, Becattini F, Seidenari L, Bimbo AD (2020) Mantra: memory augmented networks for multiple trajectory prediction. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7143–7152
9. Zhang P, Ouyang W, Zhang P, Xue J, Zheng N (2019) SR-LSTM: state refinement for LSTM towards pedestrian trajectory prediction. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 12085–12094
10. Zhong X, Yan X, Yang Z, Huang W, Jiang K, Liu RW, Wang Z (2023) Visual exposes you: pedestrian trajectory prediction meets visual intention. IEEE Trans Intell Transp Syst 1–11. https://doi.org/10.1109/TITS.2023.3266762.
11. Bhujel N, Yau W-Y (2023) Disentangling crowd interactions for pedestrians trajectory prediction. IEEE Rob Autom Lett 8(5):3078–3085
12. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
13. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Y, Polosukhin I (2017) Attention is all you need. In: Proceedings of the advances in neural information processing systems, pp 6000–6010
14. Veličković P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y (2017) Graph attention networks. arXiv preprint arXiv: 1710.10903
15. Lerner A, Chrysanthou Y, Lischinski D (2007) Crowds by example. Comput Graphics Forum 26(3):655–664
16. Robicquet A, Sadeghian A, Alahi A, Savarese S (2016) Learning social etiquette: Human trajectory understanding in crowded scenes. In: Proceedings of the European conference on computer vision, pp 549–565
17. Alahi A, Goel K, Ramanathan V, Robicquet A, Fei-Fei L, Savarese S (2016) Social LSTM: human trajectory prediction in crowded spaces. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 961–971
18. Xue H, Huynh DQ, Reynolds M (2018) SS-LSTM: a hierarchical LSTM model for pedestrian trajectory prediction. In: Proceedings of the IEEE winter conference on applications of computer vision, pp 1186–1194
19. Bhaskara R, Chiu M, Bera A (2023) SG-LSTM: Social group LSTM for robot navigation through dense crowds. arXiv preprint. arXiv: 2303.04320
20. Devlin J, Chang M-W, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv: 1810.04805

21. Lan Z, Chen M, Goodman S, Gimpel K, Sharma P, Soricut R (2019) Albert: a lite bert for self-supervised learning of language representations. arXiv preprint arXiv: 1909.11942
22. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Y, Polosukhin I (2017) Attention is all you need. In: Proceedings of the advances in neural information processing systems, pp 6000–6010
23. Yu C, Ma X, Ren J, Zhao H, Yi S (2020) Spatio-temporal graph transformer networks for pedestrian trajectory prediction. In: Proceedings of the European conference on computer vision, pp 23–28
24. Yin Z, Liu R, Xiong Z, Yuan Z (2021) Multimodal transformer networks for pedestrian trajectory prediction. In: Proceedings of the international joint conferences on artificial intelligence, pp 1259–1265
25. Zhang W, Cheng H, Johora FT, Sester M (2023) ForceFormer: exploring social force and transformer for pedestrian trajectory prediction. arXiv preprint arXiv: 2302.07583
26. Mohamed A, Qian K, Elhoseiny M, Claudel C (2020) Social-stgcnn: a social spatio-temporal graph convolutional neural network for human trajectory prediction. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 14424–14432
27. Li S, Zhou Y, Yi J, Gall J (2021) Spatial-temporal consistency network for low-latency trajectory forecasting. In: Proceedings of the IEEE international conference on computer vision, pp 1940–1949
28. Lian J, Ren W, Li L, Zhou Y, Zhou B (2023) Ptp-stgcn: pedestrian trajectory prediction based on a spatio-temporal graph convolutional neural network. Appl Intell 53(3):2862–2878
29. Lea C, Flynn MD, Vidal R, Reiter A, Hager GD (2017) Temporal convolutional networks for action segmentation and detection. In: Proceedings of the IEEE international conference on computer vision, pp 156–165

An RF Fingerprint Data Enhancement Method Based on WGAN

Bingtao Li, Di Liu, Jun Yang, Han Zhou, and Di Lin

Abstract RF fingerprinting is an emerging technology in the field of IoT security and is widely used in many areas, such as management, wireless device authentication, and interference source determination (Hall et al. in IEEE Trans Dependable Secure Comput 201–206, 2004 [1]; Wu et al. in Sci China Inf Sci 65(7):170304, 2022 [2]; Lin et al. in Sci China Inf Sci 2023 [3]). Most of these application scenarios rely on recognition methods for devices. Most of the mainstream recognition methods are based on a large amount of data for training. In case of insufficient sample size, the mainstream recognition methods are not applicable. Generative adversarial networks (GANs), with their adversarial properties, are well-suited and effective for applications in scenarios where the amount of data is insufficient. In this paper, we propose an RF fingerprint data enhancement method based on Wasserstein Generative Adversarial Network (WGAN). The experimental results show that the method can effectively improve the accuracy of RF fingerprint recognition in the same and limited data set.

Keywords WGAN · Radio frequency fingerprint · Data augmentation

1 Introduction

With the widespread use of wireless devices, the field of IoT security is gaining more and more attention. There are problems with end-to-end information transmission and mutual authentication between devices. Traditional methods use IP addresses or cookies and other methods for authentication, which also require complex encryption algorithms for encryption, and for some devices with poor arithmetic power, identifying the device becomes a burden on the processor instead.

B. Li · D. Liu (✉) · J. Yang · H. Zhou · D. Lin
University of Electronic Science and Technology of China, Chengdu, Sichuan, China
e-mail: 202121090302@std.uestc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_56

The recognition technology based on RF fingerprints has received much attention. It is favored for its uniqueness, short-term invariance, and measurability and is considered a promising solution to the recognition problem. There are various defects in the electronic components of wireless transmitters. These defects create a signal signature that varies from transmitter to transmitter, just like a fingerprint for each person. Conventional RF fingerprinting techniques select different parts of the signal for fingerprinting and use various mathematical methods to extract fingerprint information. The I/Q imbalance feature is one of the RF information features. It is a phenomenon that occurs when digital baseband signals are processed and modulated to RF, and residual mirror and carrier leakage signals are mixed into RF due to errors in the quadrature modulator. The I/Q imbalance is also a unique property of each transmitter and can be measured and calculated [4].

Currently, using neural networks to identify the RF fingerprint signals of devices is very popular. However, some devices have a long signal emission period, complex data sample acquisition, and insufficient data volume, which makes it challenging to train the neural network for the problem of varying sample sizes of different categories of signals existing in the acquired signals. The commonly used signal data enhancement methods use deep neural networks to construct signal samples. The sample distribution refers to the current distribution of existing signal features, but subject to the signal sample size, the distribution of existing sample features cannot represent the distribution of features of the actual category of signals, which will lead to the enhanced data not representing the actual signal feature data, and overfitting and other phenomena occur in the signal recognition process. They need help to achieve the generalization of the model. Generating adversarial networks with generators and discriminators that gradually fit the actual distribution in the adversarial and generate data can solve this problem well.

In this paper, we use Wasserstein Generative Adversarial Network (WGAN) to augment the data to solve the problem of insufficient data volume and improve the dataset's recognition accuracy. The main work of this paper is as follows:

- Use OFDM to eliminate the multipath effect of the signal.
- Improved WGAN model is used to generate RF fingerprint signal data, which improves the performance of the RF fingerprint recognition model.

2 Related Work

Due to the unbalanced sample size of signals from different categories and the small amount of actual data, the neural network model must be adequately trained. Otherwise, it is prone to overfitting. The unstable distribution of training data and test data is also a problem. Data Augmentation (DA) method effectively alleviates the above problems. Data augmentation methods can compensate for the lack of training samples to a certain extent and improve the generalization ability of deep learning by giving data diversity, alleviating the overfitting problem, and reducing the sensitivity

to certain confounding factors. Data augmentation methods are a means of regularizing deep learning models. Unlike explicit means such as Dropout and weight decay, data augmentation methods do not reduce the capacity of network models and increase the computational complexity. Data augmentation methods in radio signal modulation recognition include superimposing Gaussian noise and imposing channel fading effects. Unlike signal modulation features, RF fingerprint features are more susceptible to confounding factors such as the channel, and it is worthwhile to investigate how to perform data augmentation and improve the robustness of RF fingerprints under confounding factors.

In 2021, He et al. [5] addressed the problem of insufficient data volume for deep learning by using an improved auxiliary classifier-generative adversarial network (AC-GAN) to learn different image feature sets and extend the enhanced dataset to meet the training requirements. The results show that the method has higher recognition accuracy and performs better at a low signal-to-noise ratio than other methods. Shen et al. [6] used both a channel fading model and Gaussian white noise for data augmentation, where the channel fading model considered both multipath and Doppler effects and improved the correct rate from 68.6% to over 80% for high-speed scenes (Doppler frequency shift of 100 Hz). Cekic et al. [7] used a data augmentation method that used both a channel fading model and a carrier frequency offset (CFO) data augmentation method and found that as the number of training samples containing different periods in the training dataset increases, the effectiveness of this data augmentation method becomes closer to that of the combined method using carrier frequency offset compensation and data augmentation, i.e., the absence of prior knowledge can be compensated by enriching data diversity to compensate for the prior knowledge deficiency.

In 2022, Wang and Gan [8] smoothed and filtered the signal to perform noise reduction in a statistical sense to enhance the recognizability of RF fingerprints and improve recognition accuracy using more sample point accumulation. The results are better under low signal-to-noise ratio and minor sample conditions. Wang et al. [9] used the Mixup algorithm for data augmentation of time-series data by simple linear interpolation to obtain and train the augmented data. Up to 4.79% improvement in classification accuracy was achieved.

3 Method

3.1 Removal of Multipath Effect

The signal transmitted by the wireless channel reaches the receiver from different paths through the reflection of several objects during the transmission. The different paths have different propagation distances, and there is a sequence of arrival times at the receiver. The superposition of multiple signals may lead to distortion of the

signal. This is the multipath effect. We eliminate the multipath effect on the signal's time delay by the circular prefix.

3.2 Normalization

RF fingerprint recognition does not differentiate devices based on power differences. The signal power is susceptible to interference from various factors. Normalization reduces the effect of power and enhances the signal characteristics and is, therefore, a standard process for RF fingerprinting. The normalization equation of the signal [10] is as in Eq. (2):

$$s[n] = \frac{r[n]}{x_{rms}} \tag{1}$$

where x_{rms} is the mean root square of the signal amplitude, $r[n]$ can be considered as the received signal, and $s[n]$ is the signal after processing.

3.3 Wasserstein GAN

Goodfellow et al. [11] proposed Generative adversarial networks in 2014. Generative adversarial networks are widely used in many fields with their powerful learning ability, e.g., computer vision and natural language processing. As a kind of time-domain data, RF signals can be expressed as images and have some similarities with speech. Therefore, it is feasible to use generative adversarial networks to recognize RF fingerprints.

Generative adversarial networks consist of generators and discriminators and have achieved some results using a non-traditional combined training model. The generator continuously tries to generate signals with similar characteristics to the original signal and use them as part of the discriminator's input, trying to trick the discriminator and obtain feedback from the discriminator to get training indirectly. In contrast, the discriminator trains the recognition ability by judging the actual signal and the spoofed signal generated by the generator and passes the result to the generator. The structure of the generative adversarial network is illustrated as shown in Fig. 1.

WGAN is an improved version of GAN, which solves the problems of GAN training instability and pattern collapse, the network structure of WGAN is roughly the same as GAN, with some modifications in the objective function and some algorithms.

The objective function of GAN is as in Eq. (2):

$$\min_G \max_D E(G, D) = \mathbb{E}_{x\sim p_r}\left[\log D(x)\right] + \mathbb{E}_{x\sim p_g}\left[\log(1 - D(G(z)))\right] \tag{2}$$

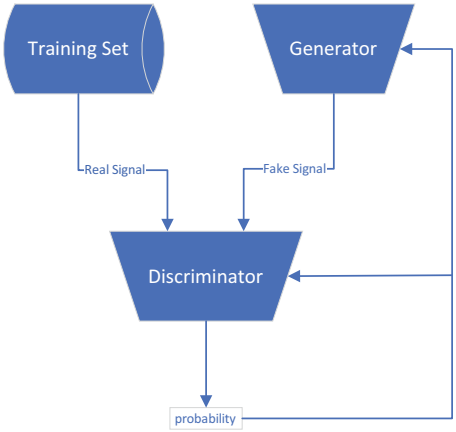


Fig. 1 The structure of GAN

p_r is the real data distribution, and p_g is the data distribution generated by the generator. This objective function is essentially an optimization function common to both the generator and the discriminator.

In two papers [12, 13], Arjovsky analyzed where the problems of GAN lie and gave an improved algorithm implementation process. GAN has two problems: first, the better the discriminator is, the more serious the generator gradient disappears. Second, the unreasonable distance measure leads to unstable gradients and insufficient diversity, so the stability and efficiency of the neural network are inferior.

GAN uses KL and JS distance to measure the degree of data fit generated. Due to the high randomness of the noise generated by the generator at the beginning, there is no overlapping part between the generated and original data. The KL distance, as in Eq. (3), is expressed as positive infinity for the region without overlap, which does not guide the model well for adjustment. The JS distance, as in Eq. (4), minimizes the loss of the generator equivalent to minimizing the JS distance between p_r and p_g under the optimal discriminator, and since it is almost impossible to have non-negligible overlap between the real data and the generated data, the JS distance is always constant, which eventually results in the generator's gradient approximating to 0 and the gradient vanishing.

$$D_{KL}(P_1|P_2) = E_{x\sim P_1} \log \frac{P_1}{P_2} \tag{3}$$

$$D_{JS}(P_1|P_2) = \frac{1}{2} D_{KL}\left(P_1|\frac{P_1+P_2}{2}\right) + \frac{1}{2} D_{KL}\left(P_2|\frac{P_1+P_2}{2}\right) \tag{4}$$

Based on these issues, Martin has also proposed changes that can be summarized in three points:

- Remove the sigmoid function from the last layer of the discriminator

- Use Wasserstein distance instead of KL and JS distance to measure the difference between the real and false data distribution
- Do not use momentum-based optimization algorithms, such as momentum and Adam. RMSProp is recommended.

The improved objective function is as in Eq. (5):

$$\min_{\theta} \max_{w} D(G_\theta, f_w) = E_{x\sim p_r(x)}[f_w(x)] - E_{z\sim p_{\text{noise}}(z)}[f_w(G_\theta(z))] \tag{5}$$

A discriminator gradient g_w is calculated by deriving the objective function. Next, the optimization is performed using the RMSProp function, which is a modification of AdaGrad, and instead of directly accumulating g_w^2, a decay parameter w is added to control the amount of historical information acquired, which makes the gradient decreases more smoothly and rapidly. The new weight w is calculated from the RMSProp function, supplemented by the previous weight w and g_w. Since the discriminator in GAN does a true–false binary classification task, where the result is 1 or 0, the sigmoid function is used at the end. But now the discriminator in WGAN does an approximate fit to the Wasserstein distance, so delete the sigmoid function. Also, to ensure that the gradient always satisfies Lipschitz continuity and limits the range of parameter variation, the algorithm uses the clip function for weight cropping.

4 Experiment

4.1 Lab Experiment and Dataset

The experimental hardware environment is AMD 5600X CPU, NVIDIA GeForce RTX 3070 GPU, and 16 GB RAM; the software environment is Windows 11 64-bit, python 3.6.5. Pytorch is used to build the neural network in this paper.

The RF fingerprint data set in this paper is generated by MATLAB simulation with QPSK modulation, a sampling frequency of 20 Mbps, a sampling interval of 0.05 us, a signal-to-noise ratio of 8, and the number of subcarriers of 52, including four pilot signals and the rest are data signals. The number of devices to be identified in the simulation is 4. After data pre-processing, a total of nearly 50,000 data are generated, including 40,000 training data and 10,000 test data, each corresponding to one of the devices in 0–3.

4.2 Experiment Design

As shown in Fig. 2, in the experimental phase, we divided the dataset into a training set and a test set, with 40,000 training data and 10,000 test data. Then, we input

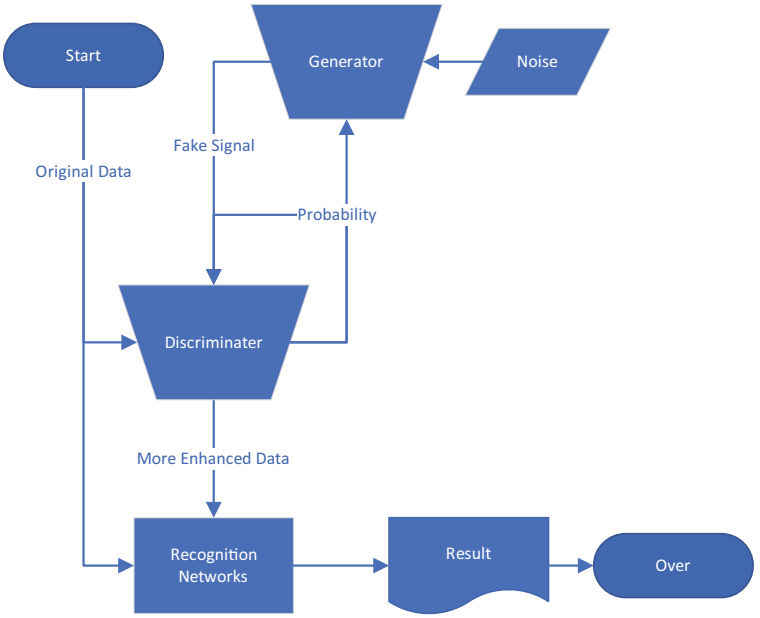


Fig. 2 Flow chart of experiment

the training set into WGAN for data augmentation, generating 40,000 pieces of data for each label, with a total of 160,000 pieces of data. Then, the original dataset and the enhanced dataset are input into DNN, CNN, and RNN-LSTM, respectively, for training with no change in hyperparameters. And test data are used for validation.

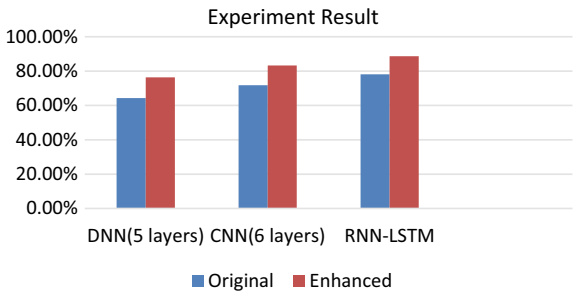
4.3 Experiment Result

The experimental results from Table 1 show that DNN improves by 11.9% accuracy, CNN improves by 14.6% accuracy, and RNN improves by 9.4% accuracy. Using the enhanced dataset for training effectively improves the accuracy of the neural network model. Side by side, it is demonstrated that WGAN generates new data with quality that matches the distribution of the original dataset (Fig. 3).

Table 1 The result of experiment

Dataset	Method	Acc (%)
Original	DNN (5 layers)	64.30
Original	CNN (6 layers)	71.80
Original	RNN-LSTM	78.10
Enhanced	DNN (5 layers)	76.40
Enhanced	CNN (6 layers)	83.3
Enhanced	RNN-LSTM	88.7

Fig. 3 The result of experiment



5 Conclusion

WGAN generates and confronts features as an effective method for data enhancement. WGAN uses RMSProp and clip for gradient optimization to make the training process more stable and solve the problem that GAN is prone to pattern collapse.

This thesis solves the feature recognition problem in the case of insufficient data by using the features of WGAN. WGAN can enhance the quality of generated data in the adversarial process and has low dependence on the amount of data. However, the training is stable but slow because WGAN uses weight pruning to force the discriminator to satisfy Lipschitz continuity and prevent gradient explosion. Later, we can introduce gradient penalty, upgrade the network to WGAN-GP to improve the training speed, and further investigate the classification problem of RF fingerprints.

Acknowledgements Partially funded by Natural Science Foundation of Sichuan Province (2023NSFSC0479) and partially funded by Grant SCITLAB-20005 of Intelligent Terminal Key Laboratory of Sichuan Province.

References

1. Hall J, Barbeau M, Kranakis E (2004) Radio frequency fingerprinting for intrusion detection in wireless networks. IEEE Trans Dependable Secure Comput 201–206

2. Wu W, Hu S, Lin D, Wu G (2022) Reliable resource allocation with RF fingerprinting authentication in secure IoT networks. Sci China Inf Sci 65(7):170304
3. Lin D, Hu S, Wu W, Wu G (2023) Few-shot RF fingerprinting recognition for secure satellite remote sensing and image processing. Sci China Inf Sci. https://doi.org/10.1007/s11432-022-3672-7
4. Zou S, Liu J, Yang H (2013) Research of compensating zero-IF modulator IQ imbalance in digital domain. Video Eng 23:163–166. https://doi.org/10.16280/j.videoe.2013.23.030
5. He Z, Hou S, Zhang W, Zhang Y (2021) Multi-feature fusion classification method for communication specific emitter identification. J Commun 02:103–112
6. Shen G, Zhang J, Marshall AJ, Cavallaro J (2021) Towards scalable and channel-robust radio frequency fingerprint identification for LoRa. IEEE Trans Inf Forensics Secur 17:774–787
7. Cekic M, Gopalakrishnan S, Madhow U (2020) Robust wireless fingerprinting: generalizing across space and time. arXiv:2002.10791
8. Wang W, Gan L (2022) Radio frequency fingerprinting improved by statistical noise reduction. IEEE Trans Cogn Commun Netw 8:1444–1452
9. Wang T, Bian Y, Zhang Y, Hou X (2022) Using artificial intelligence methods to classify different seismic events. Seismol Res Lett
10. Shen G, Zhang J, Marshall AJ, Peng L, Wang X (2020) Radio frequency fingerprint identification for LoRa using spectrogram and CNN. In: IEEE INFOCOM 2021—IEEE conference on computer communications, pp 1–10
11. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville AC, Bengio Y (2014) Generative adversarial nets. NIPS
12. Arjovsky M, Bottou L (2017) Towards principled methods for training generative adversarial networks. arXiv:1701.04862
13. Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. In: International conference on machine learning

Deep Image Retargeting Network with Multi-loss Functions

Xiaoting Fan, Long Sun, and Zhong Zhang

Abstract Image retargeting aims at displaying an image on a serious of device screen with different sizes, which has been widely applied in computer graphics and other fields. At present, many deep learning-based image retargeting methods implement an encoder-decoder retargeting network to resize attention map. Then, various types of loss functions are presented to preserve salient contents and reduce structure distortions. In this survey, we first review three types of loss functions utilized in deep image retargeting that consists of pixel-based, probability-based, and perception-based. Furthermore, we explore the baseline encoder-decoder retargeting network with three types of loss functions, and conduct the experiments on two public datasets, *Retargetme* and *Pascal Voc 2007* datasets to verify their impact on the deep image retargeting.

Keywords Image retargeting · Deep learning · Multi-loss functions

1 Introduction

With the rapid commercialization of display technologies, the demand for image retargeting techniques increases gradually. Image retargeting is an emerging research area, which aims to resize an image into arbitrary resolutions intelligently [1, 2]. It plays a fundamental role in many multimedia applications, such as mobile device, commercial advertisement, magazine production and so on.

X. Fan · Z. Zhang (✉)
Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
e-mail: zhong.zhang8848@gmail.com

X. Fan
e-mail: xtfan@tjnu.edu.cn

L. Sun
School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China
e-mail: sun_long@tju.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_57

In the last decades, image retargeting has attracted numerous research scholars attention and many image retargeting methods have been proposed. In general, traditional image retargeting include discrete-based and continuous-based. Discrete-based methods change the image size by removing or inserting pixels in an image, i.e. image cropping [3], seam-carving [4]. In contrast, continuous-based methods perform the mesh warping with different optimization functions, such as uniform scaling [5], and letterboxing [6].

Due to the outstanding performance in feature representation, deep learning techniques have been employed in many image retargeting tasks. Compared with traditional image retargeting methods, most deep learning-based methods aim to implement an encoder-decoder retergeting model to resize the attention map firstly, then design different types of loss functions to preserve visually salient regions and remove unimportant pixels. At present, three types of loss functions are proposed to constraint the deep image retargeting network, i.e., pixel-based, probability-based, and perception-based loss functions. For example, Cho et al. [7] and Lin et al. [8] applied a pixel-based mean absolute error loss to measure the patch similarity of original and target feature maps. Similarity, Danon et al. [9] compared target feature maps with output feature maps by using the pixel-based mean squared error loss. Furthermore, Cho et al. [7] considered the image retargeting as a classification task, where a probability-based cross entropy loss was introduced to constrain the output image to have a similar classification result as the input image. In addition, Tan et al. [10] introduced an unsupervised deep cyclic image retargeting method without any explicit user annotations, in which a cyclic perception coherence loss was designed to evaluate the cycle coherence.

In the survey, we review three types of loss functions for deep image retargeting, i.e., pixel-based mean absolute error loss [7, 11–13], probability-based cross entropy loss [9, 11], and perception-based cyclic perception coherence loss [10]. The mean absolute error loss is the most widely loss function in deep image retargeting, it reduces the shape distortions of retargeted image by calculating the average distance between the predicted value and true value of input image. Different from the pixel-based loss, the cross-entropy loss uses the image-level annotations to constrain the classification result of retargeted image closer to the classification result of input image. In addition, the cyclic perception coherence loss aims to compare the deeper feature maps of input image with those of retargeted image, which can improve the human perceptual quality. Furthermore, we explore the baseline encoder-decoder retargeting network with different loss functions and conduct the experiments on two public datasets, *Retargetme* [14] dataset and *Pascal Voc 2007* [15] dataset to verify their impact on the deep image retargeting. Meanwhile, three loss functions are combined together by an optimization strategy to obtain a high-quality retargeted image.

2 Methods

In this section, the attention map extraction module of deep image retargeting method is presented firstly. Then, the mean absolute error loss, cross-entropy loss, and cyclic perception coherence loss are introduced in details. The overall optimization strategy on the multi-loss functions is described finally.

2.1 Attention Map Extraction Module

The overview architecture of the deep image retargeting network with multi-loss functions is illustrated in Fig. 1, which includes two parts: an attention map extraction module and multi-loss functions. Firstly, the attention map extraction module based on encoder-decoder is designed to extract the visual important objects. Afterwards, a shift layer [7] is utilized to retarget the attention map. Finally, multi-loss functions are applied to constrain the deep image retargeting network.

The attention map extraction module mainly consists of two stages, i.e., an encoder stage and a decoder stage, where the encoder stage is applied to extract high-level deep semantic features and the decoder stage is utilized to generate an attention map. To be specific, an input image I is first fed into a pre-trained VGG-16 Network [16] to extract a feature map. Then, the feature map is sent into an inverse VGG-16 Network to generate the attention map. Meanwhile, fully connected layers of VGG-16 Network are removed and the Rectified Liner Unit (ReLU) layers in the decoder stage are replaced with Exponential Linear Unit (ELU) [17] layers.

2.2 Mean Absolute Error Loss

To make the image retargeting result have a similar geometry structure to the original image, a pixel-based mean absolute error loss is designed as a supervision for geometry structural similarity. Specially, if the geometry structure of the image retargeting result is preserved well, the neighbourhood of each pixel in the image retargeting

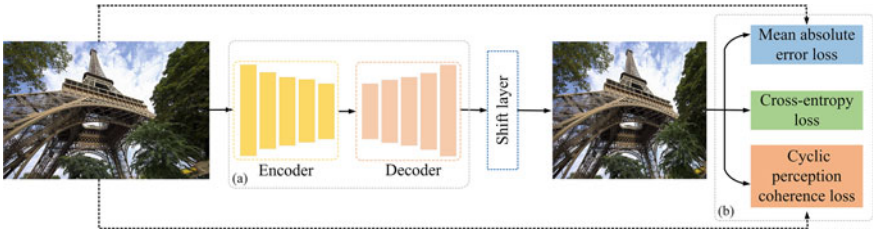


Fig. 1 The architecture of the deep image retargeting network with multi-loss functions. **a** Attention map extraction module, **b** multi-loss functions

result should be similar to the neighbourhood of each corresponding pixel in the original image. The mean absolute error loss [12] is defined as follows.

$$L_{Mean} = \|I_t - I_r\| \tag{1}$$

with

$$I_r = I_o(w + S(w, h), h) \tag{2}$$

where w and h are the pixel coordinate in the original image. I_o, I_t, and I_r are the original image, target image, and reconstructed image. $S(w, h)$ is the shift map obtained by the shift layer.

2.3 Cross Entropy Loss

In general, cross entropy loss is designed for classification tasks. It measures the difference between the expected distribution and the real distribution by calculating the difference between two probability distributions. For image retargeting, if the main objects or salient regions is preserved intact after retargeting, the retargeted image should have the same classification score as the original image. The cross entropy loss [7] is defined as follow.

$$L_{Cross} = \frac{1}{CS} \sum_{c=1}^{C} \sum_{s=1}^{S} [p_{cs} \mathbf{log} q_{cs} + (1 - p_{cs}) \mathbf{log}(1 - q_{cs})] \tag{3}$$

where C and S denote the number of classes and images. p_{cs} and q_{cs} denote the ground truth label and sigmoid output.

2.4 Cyclic Perception Coherence Loss

In most cases, the mean absolute error is inconsistent with human visual perception, which cannot meet the accuracy requirements of deep image retargeting task. Therefore, a cyclic perception coherence loss is designed to evaluate the visual perceptual consistency between original image and reconstructed image. It can enable the deep image retargeting network to extract the significant regions and learn the same distribution between original image and target image. The cyclic perception coherence loss [10] is defined as follow.

$$L_{Cycle} = \frac{1}{N} \sum_{n=4}^{N} [f_n(I_t \times \eta_n - f_n(I_r))] \tag{4}$$

where N is the convolutional layers. $f_n(\cdot)$ is the pre-trained VGG-16 Network. η is the parameter.

2.5 Optimization Strategy

To demonstrate the effectiveness of different loss functions in deep image retargeting network, the experiments are conducted to compare the performance of three types of loss functions. Specifically, mean absolute error loss, cross-entropy loss, and cyclic perception coherence loss are combined together to optimize the deep image retargeting network. The total multi-loss function is defined as:

$$L_{Total} = \alpha L_{Mean} + \beta L_{Cross} + \gamma L_{Cycle} \tag{5}$$

where, α, β and γ are the relative weights.

3 Experimental Validations

3.1 Experimental Setup

Dataset: In order to evaluate the impact of different loss functions on deep image retargeting method, we test the deep image retargeting network with three loss functions on two datasets. The first is the *Retargetme* dataset, which is one of the most commonly used datasets in the image retargeting. It contained 80 images with different cases, such as foreground objects, texture elements, people, and so on. The second is the *Pascal Voc 2007* dataset, which is widely used in the visual object classification, recognition and detection. It contains 9963 labeled images and 24640 objects, in which the training set and verification set contain 5012 images, and test set contains 4915 images.

Implementation details: The adaptive moment estimation optimizer method [18] is applied to train and update the parameters of deep model. The initial learning rate is set as 10^{-4}, the batch size is set to 4, and the momentum is set to 0.9. During the training, the weight α, β and γ are set to 0.5, 0.5 and 1. η_4 and η_5 are set to 1 and 3. The input target aspect ratios are randomly generated within $(\frac{W}{4} \sim \frac{W}{2}) \times (\frac{H}{4} \sim \frac{H}{2})$, where W and H denote the width and height of the input image.

Quantitative metrics: In addition to the qualitative comparison, the average gradient score [12] is used to quantitative assess the quality of the image retargeting results, which can reflect the clarity and texture changes of an image. If the image retargeting results is smooth and clear, the average gradient score is high. The average gradient score *AGS* is defined as follows.

$$AGS = \frac{1}{WH} \sum_{w=1}^{W} \sum_{h=1}^{H} \sqrt{0.5 \times [(\frac{\partial G(w,h)}{\partial x})^2 + (\frac{\partial G(w,h)}{\partial y})^2]} \tag{6}$$

where $G(w, h)$ is the gray value of (w, h).

Fig. 2 Six test data. **a** Original, **b** the case of MAE loss, **c** the case of CS loss, **d** the case of CPC loss, **e** the case of MAE and CE losses, **f** the case of MAE and CPC losses, **g** the case of multi-loss functions

3.2 Comparison Evaluation

In this section, we conduct the qualitative and quantitative experiments on the *Retargetme* dataset and *Pascal Voc 2007* dataset to evaluate the effect of different types of loss functions on the deep model. MAE, CE, and CPC represent the attention map extraction module with the mean squared error loss, cross-entropy loss, and cyclic perception coherence loss, respectively.

Figure 2 indicates some visual images on the *Retargetme* dataset and *Pascal Voc 2007* dataset that are shrank by 25%. Compared with the case of MAE loss and the case of CPC loss, the case with CS loss deforms the salient objects more serious, because it only considers the classification of the image, while ignoring the image content details. In addition, it can be seen that the case of MAE and CE losses has evident shape distortions, but the case of MAE and CPC losses provides slight visual distortions. It proves that the perception-based loss is more effective than the probability-based loss in preserving visual saliency objects. Furthermore, the case of multi-loss functions obtains the same structure as the original image. It demonstrates that the multi-loss functions with optimization strategy can retain the foreground and uniform the background.

Table 1 Average gradient score of different loss functions for deep image retargeting

Test data	MAE	CS	CPC	MAE + CE	MAE + CPC	Multi-loss
Car	0.2789	0.2456	0.3059	0.3108	0.3578	0.3867
Child	0.4643	0.4354	0.4876	0.5087	0.5267	0.5564
Johanneskirche	0.2789	0.2564	0.2978	0.3067	0.3245	0.3793
LakeVillage	0.2465	0.2356	0.2643	0.2956	0.3134	0.3496
Pencils	0.2798	0.2503	0.2987	0.3546	0.3896	0.4045
Flamingo	0.3792	0.3086	0.4029	0.4372	0.475	0.5021
Avgerage	0.3213	0.2887	0.3429	0.3689	0.3978	0.4298

The quantitative comparisons of different loss functions for deep image retargeting are illustrated in Table 1, where the average gradient score is calculated on six test datasets. In the case of using a single loss function, the case of CPC loss obtains the best result on average gradient score, because it considers the impact of human visual perception on salient objects. Additionally, the average gradient score of the case of MAE and CPC losses is better than the case of MAE and CE losses. It demonstrates that the combination of pixel-based MAE loss and perception-based cyclic perception coherence loss could promote the deep image retergeting network to extract more accurate salient features. Compared with other cases, the case of multi-loss function gives the best performance, which proves that the combination of three loss functions could fully learn feature similarity before and after image retargeting.

4 Conclusion

In this survey, we first review the mean absolute error loss, cross-entropy loss, and cyclic perception coherence loss, which are widely applied in the deep image retargeting. Then, a deep image retargeting network with multi-loss functions is introduced in details. Finally, experimental comparisons and analysis of the deep image retargeting with different loss functions are conducted. Experimental results on two common datasets have illustrated that multi-loss functions can obtain high-quality retargeted images in preserving salient contents and reducing structure distortions.

Acknowledgements This work was supported in part by the Natural Science Foundation of Tianjin under Grant No. 22JCQNJC00010, in part by the Scientific Research Project of Tianjin Educational Committee under Grant No. 2022KJ011, and in part by the National Natural Science Foundation of China under Grant No. 62171321.

References

1. Tan W, Yan B, Li K, Tian Q (2016) Image retargeting for preserving robust local feature: application to mobile visual search. IEEE Trans Multimedia 18(1):128–137
2. Zhou Y, Chen Z, Li W (2021) Weakly supervised reinforced multi-operator image retargeting. IEEE Trans Circ Syst Video Technol 31(1):126–139
3. Yan J, Lin S, Kang SB, Tang X (2013) Learning the change for automatic image cropping. In: IEEE conference on computer vision and pattern recognition, vol 971–978
4. Avidan S, Shamir A (2007) Seam carving for content-aware image resizing. ACM Trans Graph 26(3):10–19
5. Rubinstein M, Shamir A, Avidan S (2009) Multi-operator media retargeting. ACM Trans Graph 28(3):23–33
6. Mishiba K, Oyamada Y, Kondo K (2017) Content-aware image retargeting incorporated with letterboxing. IEICE Trans Inf Syst 100(4):865–873
7. Cho D, Park J, Oh T-H, Tai Y-W, Kweon IS (2017) Weakly- and self-supervised learning for content-aware deep image retargeting. In: IEEE international conference on computer vision (ICCV), pp 4568–4577
8. Lin J, Zhou T, Chen Z (2019) DeepIR: a deep semantics driven framework for image retargeting. In: IEEE international conference on multimedia and expo workshops (ICMEW), pp 54–59
9. Danon D, Arar M, Cohen-Or D, Shamir A (2021) Image resizing by reconstruction from deep features. Comput Vis Media 4:453–466
10. Tan W, Yan B, Lin C, Niu X (2020) Cycle-IR: deep cyclic image retargeting. IEEE Trans Multimedia 22(7):1730-C1743
11. Liu S, Wei Z, Sun Y, Ou X, Lin J, Liu B, Yang M-H (2018) Composing semantic collage for image retargeting. IEEE Trans Image Process 27(10):5032–5043
12. Fan X, Lei J, Liang J, Fang Y, Ling N, Huang Q (2021) Stereoscopic image retargeting based on deep convolutional neural network. IEEE Trans Circ Syst Video Technol 31(12):4759–4770
13. Fan X, Lei J, Liang J, Fang Y, Ling N (2021) Unsupervised stereoscopic image re-targeting via view synthesis and stereo cycle consistency losses. Neurocomputing 447(11):161–171
14. Rubinstein M, Gutierrez D, Sorkine O, Shamir A (2010) A comparative study of image retargeting. ACM Trans Graph 29(6):160:1–160:10
15. Everingham M, Gool L, Williams C, Winn J, Zisserman A (2007) The pascal visual object classes challenge 2007 (voc2007) results. http://www.pascalnetwork.org/challenges/VOC/voc2007/workshop/index.html
16. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. CoRR (abs/1409.1556, 2015)
17. Clevert D-A, Unterthiner T, Hochreiter S (2015) Fast and accurate deep net-work learning by exponential linear units (elus). arXiv preprint arXiv:1511.07289
18. Kingma D, Ba J (2014) Adam: a method for stochastic optimization. arXiv:1412.6980, [Online]. Available: https://arxiv.org/abs/1412.6980

RF Fingerprinting Based on Contrastive Learning and Convolutional Neural Networks

Jun Yang, Di Liu, Bingtao Li, Han Zhou, and Di Lin

Abstract Hardware differences create a unique fingerprint of the radiation source and are attached to the radio signal, and this unique property of the radiation source can be used for RF fingerprinting. The RF fingerprinting method based on expert experience has excessive prior knowledge and poor robustness in different environments. The RF fingerprint recognition method based on deep learning, especially the method that can directly process Raw I/Q shows great potential, but most current deep learning-based RF fingerprint recognition methods require manual annotation of I/Q data. In this paper, the SimSiam model is used to process the data in the form of self-supervised comparative learning, which greatly reduces the labor cost while ensuring the accuracy. The backbone network uses an optimized convolutional neural network (CNN) for classification recognition, which saves manpower and time while ensuring recognition accuracy.

Keywords RF fingerprinting · Self-supervised comparative learning · CNN

1 Introduce

There are specific gaps in generating and assembling devices composed of radiation sources [1]. This hardware difference makes the radiation source of the same model and batch have an inherent attribute that is different from other individuals [2], which will act on the radio signal and produce unintentional modulation. However, it will not affect the radiation source to achieve its original communication or detection function. By characterizing this property by a particular method, the association between the radio signal and the radiation source can be inferred to achieve individual identification, called radio frequency fingerprinting (RFF) [3–5]. According to whether to convert I/Q signal into expert features or not, there are two categories of RF fingerprint recognition methods based on expert features and deep learning and RF fingerprint recognition methods based on raw I/Q and deep learning. According

J. Yang · D. Liu (✉) · B. Li · H. Zhou · D. Lin
University of Electronic Science and Technology of China, Chengdu, Sichuan, China
e-mail: 202121090303@std.uestc.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_58

to whether the I/Q signal is converted into expert features, there are two main categories of RF fingerprinting methods based on expert features and deep learning and RF fingerprinting methods based on raw I/Q and deep learning.

1.1 Expert Features and Deep Learning for RF Fingerprinting Methods

The deep learning method based on expert features needs the support of a priori knowledge, and its classification effect greatly depends on the expert's knowledge and understanding of the domain to which it belongs. The first step is to perform some pre-processing operations on the signal to mitigate the adverse effects of the wireless channel. Then expert feature transform operations include short-time Fourier transform, wavelet transform, and constellation diagram. Convert I/Q data into expert feature representations such as time–frequency feature maps, bispectral feature maps, and constellation maps. Finally, the transformed expert features feed into the DNN model for feature extraction and classification recognition.

Deep learning methods based on expert features have been extensively studied in RF fingerprinting and achieved good results. However, this method relies too much on prior knowledge, is less robust in different environments, and needs more generalized.

1.2 RF Fingerprint Recognition Methods Based on Raw I/Q and Deep Learning

Raw I/Q [6, 7] refers to unprocessed I/Q data, which not only indicates that the I/Q data output by the receiver has not lost information but also means that there is no suppression of the confounding factors that are not conducive to radio frequency fingerprint recognition. Under non-cooperative conditions, it is challenging to perform preprocessing, such as carrier frequency offset compensation and time synchronization, and almost only I/Q data can be used. From the existing research, RF fingerprinting based on raw I/Q and deep learning has at least the following advantages:

1. Does not rely on domain-specific and prior knowledge, has end-to-end characteristics, and uses data-driven automatic learning features.
2. It has the potential to learn protocol-independent RF fingerprints, and extracted RF fingerprints are more difficult to forge.
3. Better scalability and flexible deployment.
4. Potential for identification of large-scale and ultra-large numbers of individuals.

In this paper, RF fingerprinting based on raw I/Q and deep learning is implemented, and its main contributions are as follows:

1. Reduce labor costs and use contrastive learning models to label data. A simple Siamese (SimSiam) [8] model is pre-trained in an unsupervised manner by comparing samples from different recipients.
2. Improve the convolutional neural network [9], optimize the results through a small amount of labeled data.

2 Related Work

2.1 Contrastive Self-supervised Learning

The supervised information of self-supervised learning is not manually labeled. However, the algorithm automatically constructs supervised information in large-scale unsupervised data for supervised learning or training, so it can avoid a large number of label labeling of the dataset, significantly reducing human labor.

In self-supervised learning, pseudo-supervised tasks are set up to replace human annotations with specific data properties. For example, rotate a picture by 0/90/180/270 degrees and then train a model to predict the rotation angle instead of labeling the image as cat/dog. If marking an image as cat/dog requires human participation, you can write a script to rotate the image and record the angle of its rotation as a tag. After learning feature representations from millions of images, we can use transfer learning to fine-tune some supervised tasks, such as image classification of cats and dogs, as shown in Fig. 1.

Comparative learning is mainly through an auxiliary task, constructing positive and negative samples, and learning by comparing the distance difference between positive and negative samples. The core idea is to have a sample anchor and then construct its positive sample pos, and negative sample neg, and then anchor and pos are much more similar than anchor and neg, which is formulated as

$$\text{score}\left(f(x), f\left(x^{+}\right)\right) \gg \text{score}\left(f(x), f\left(x^{-}\right)\right) \tag{1}$$

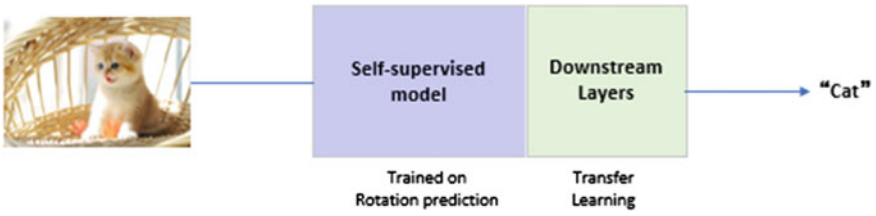


Fig. 1 Contrastive self-supervised learning

2.2 *Convolutional Neural Networks*

Convolutional neural networks have representation learning capabilities [10], can perform shift-invariant classification of input information according to their hierarchical structure, can perform supervised learning and unsupervised learning, and its implicit convolution kernel parameter sharing within layers, and the sparsity of connections between layers make convolutional neural networks grid-like with a small amount of computation topology features.

As shown in Fig. 2, the tangible result is divided into two parts: forward propagation and backpropagation. First, select the training sample (x, y) and input x into the network. The weight is randomly initialized (usually, the decimal is selected), the information is extracted and transformed from the input layer by feature and finally reaches the output layer to obtain the output result. The output result is compared with the ideal result, and the global loss is calculated. The obtained error is transmitted to neurons in different layers in reverse, and the weights and biases are adjusted according to the "iterative method" to find the overall optimal results.

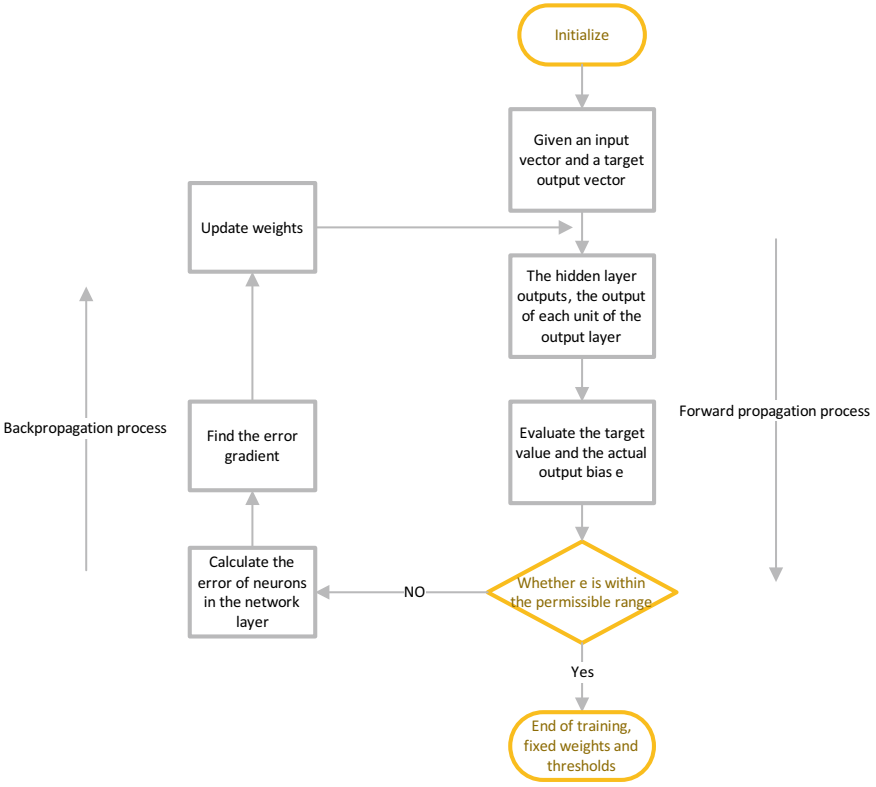


Fig. 2 The training process of CNN

3 Method

This is because samples with high SNR reflect fingerprint features more intuitively than samples with low SNR. We choose high SNR samples for unsupervised learning to prevent model collapse. In this paper, we use SimSiam [11] to perform self-supervised comparison learning and replace its backbone network with an improved CNN network that is more suitable for RF recognition, and the improved model can distinguish the differences between I/Q signal samples more efficiently, thus ensuring the model recognition accuracy.

The structure of the whole model is shown in Fig. 3, first the I/Q Signal x is enhanced to obtain x_1, x_2. The enhanced data is used as input; x_1 and x_2 are encoded by the same encoder (CNN) to obtain two representation vectors $z_1 = f(x_1)$, $z_2 = f(x_2)$. After that, the z_1 is passed through an MLP map to obtain the formula $p_1 = h(z_1)$, the similarity between z_1, p_1 is calculated, and learned as an objective function. SimSiam's input is a cheerful sample pair, so we need to maximize the similarity between the two. Because the formula for cosine similarity is to find the cosine value of the angle between two vectors. The larger the cosine, the smaller the angle between the two vectors, indicating that the two vectors are closer together and vice versa. So, the loss function is as follows:

$$l = \frac{1}{2}\left(-\frac{p_1}{||p_1||_2} * \frac{z_2}{||z_2||_2} + \left(-\frac{p_2}{||p_2||_2} * \frac{z_1}{||z_1||_2}\right)\right) \tag{2}$$

The final loss function is obtained by making a simple modification to the above formula:

$$L = \frac{1}{2}D(p_1, z_2) + \frac{1}{2}D(p_2, z_1) \tag{3}$$

The previous is the forward process, and in the backward propagation, will stop the gradient propagation to one side of the model, the reaction formula then becomes

$$L = \frac{1}{2}D(p_1, stopgrad(z_2)) + \frac{1}{2}D(p_2, stopgrad(z_1)) \tag{4}$$

My CNN framework is partially inspired by CNN [9] and performs well in RF fingerprint recognition. The structure of the described CNN is shown in Fig. 4, specifically: convolutional layer + first convolutional module + second convolutional module + third convolutional module + Flatten layer + Dropout layer + first fully connected layer + second fully connected layer + third fully connected layer; where the first convolutional module, second convolutional module and third convolutional module have the same structure and all include two convolutional layers connected sequentially and a maximum pooling layer.

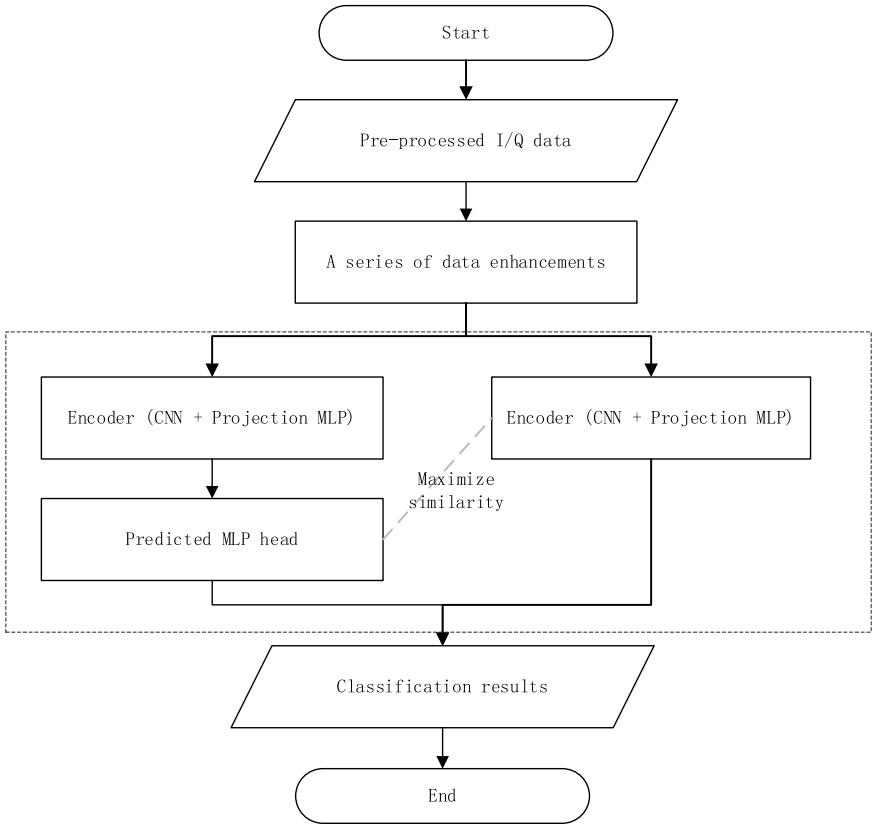


Fig. 3 The entire model structure

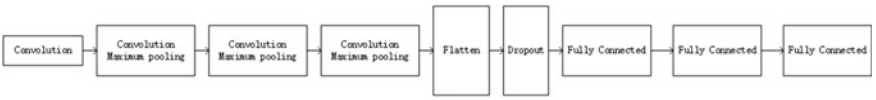


Fig. 4 CNN

4 Experiment Evaluation

4.1 Experiment Setting

We selected about 5 million I/Q signal from eight devices. These signal data are generated under the AWGN channel, and we select an SNR range of 18–20 dB and use all samples for the next training phase.

Table 1 Experiment results

Net	Acc (%)	Training time (minutes)/epoch
ResNet	79.77	270.22
SimSiam + CNN	77.64	140.23
SVM	70.98	433.12

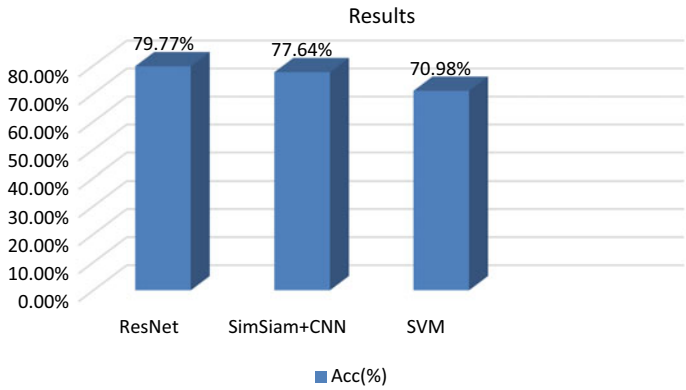


Fig. 5 Experiment results

4.2 Experiment Results

The results are shown in Table 1. From the table, we can see that the accuracy of the model with self-supervised learning is almost catching up with ResNet, but the training time is much less than that of ResNet. The accuracy rate is a little higher than that of SVM, but the training time is three hours less (Fig. 5).

5 Conclusions

In this paper, we propose a new scheme to reduce the manual workload and improve the RFF recognition rate. We use contrast learning to label the data and an improved convolutional neural network for classification recognition with good portability and operability.

Acknowledgements Partially funded by Natural Science Foundation of Sichuan Province (2023NSFSC0479) and partially funded by Grant SCITLAB-20005 of Intelligent Terminal Key Laboratory of Sichuan Province.

References

1. Brik V, Banerjee S, Gruteser M, Oh S (2008) Wireless device identification with radiometric signatures. In: Proceedings of the annual international conference on mobile computing and networking, MOBICOM, pp 116–127. https://doi.org/10.1145/1409944.1409959
2. Klein RW, Temple MA, Mendenhall MJ (2009) Application of wavelet-based RF fingerprinting to enhance wireless network security. J Commun Netw 11:544–555
3. Soltanieh N, Norouzi Y, Yang Y, Karmakar NC (2020) A review of radio frequency fingerprinting techniques. IEEE J Radio Freq Identif 4:222–233
4. Wu W, Hu S, Lin D, Wu G (2022) Reliable resource allocation with RF fingerprinting authentication in secure IoT networks. Science China Inf Sci 65(7):170304
5. Lin D, Hu S, Wu W, Wu G (2023) Few-shot RF fingerprinting recognition for secure satellite remote sensing and image processing. Science China Inf Sci. https://doi.org/10.1007/s11432-022-3672-7
6. Shen G, Zhang J, Marshall AJ, Valkama M, Cavallaro J (2021) Radio frequency fingerprint identification for security in low-cost IoT devices. In: 2021 55th Asilomar conference on signals, systems, and computers, pp 309–313
7. Gritsenko A, Wang Z, Jian T, Dy JG, Chowdhury KR, Ioannidis S (2019) Finding a 'new' needle in the haystack: unseen radio detection in large populations using deep learning. In: 2019 IEEE international symposium on dynamic spectrum access networks (DySPAN), pp 1–10
8. Chen X, He K (2020) Exploring simple Siamese representation learning. In: 2021 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 15745–15753
9. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. Commun ACM 60:84–90
10. Kim J, Park H, Lee J, Kang U (2018) SIDE: representation learning in signed directed networks. In: Proceedings of the 2018 world wide web conference
11. Zha X, Li T, Qiu Z, Li F (2023) Cross-receiver radio frequency fingerprint identification based on contrastive learning and subdomain adaptation. IEEE Signal Process Lett 30:70–74

A Knowledge Graph for Network Security

Huikang Zhang, Youyun Xu, Jian Chen, Wenyu Zhou, and Liangliang Cheng

Abstract In order to enable decision-making for more in-depth analysis, the network security knowledge graph can perform semantic analysis and understanding of multi-source, heterogeneous, and fragmented huge datasets. However, there are still issues with building network security knowledge graphs, such as limited datasets and ineffective entity extraction. To address this issue, this work proposes a BERT-based entity extraction model. Experiments have demonstrated that our model not only outperforms competing methods on publicly available datasets but also achieves roughly 88 percent performance on the Chinese network security dataset in terms of a variety of criteria. Finally, a network security knowledge graph is created based on knowledge extraction done on gathered Chinese network security texts.

Keywords Network security · Knowledge graph · Entity extract

1 Introduction

The quick advancement of Internet technology has greatly contributed to societal progress and wealth, and it has had a profound effect on both the economy and society. However, as a result, the Internet environment has grown more complicated and harsh, and Trojans, ransomware, and other sorts of network attacks have posed substantial risks to cyberspace. The National Information Security Vulnerability Database reportedly gathered around 26,000 vulnerabilities in the most recent year, demonstrating an annual pattern of growth. Frequent network security events have had a significant negative impact on the nation, its businesses, and its citizens'

H. Zhang (✉) · Y. Xu · J. Chen · W. Zhou · L. Cheng
School of Communication and Information Engineering Nanjing University of Posts and Telecommunications, Nanjing 210003, China
e-mail: 1021010314@njupt.edu.cn

J. Chen
e-mail: chenjian1980@njupt.edu.cn

National Engineering Research Center of Communications and Networking, Nanjing 210003, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_59

capacity to flourish economically and socially. Because of this, network security has taken on increased importance, and we now need to take action to combat the challenges it faces.

Network security knowledge graphs, which represent network knowledge with a graph-based data model, provide holistic approaches for processing massive volumes of complex network security data derived from diverse sources [12]. In network security, building a knowledge graph enables the effective extraction of crucial security insights from large and fragmented data. It improves network security measures by looking at how these insights are related to one another. In the field, this research area is quite important. With the use of knowledge graph technology, vital information infrastructure, threat intelligence, security reports, network attack events, and other security data may be thoroughly analyzed and explored.

The task of constructing a knowledge graph usually involves several parts such as ontology design, collection, cleaning, and pre-processing of text data, manual annotation to build datasets, knowledge extraction, and classification. In the current network security entity extraction technology, we also face some problems and aspects that need improvement, such as the scarcity of Chinese datasets in the relevant network security field, which are mostly dominated by English, and the need to improve the effectiveness of entity extraction.

In response to the above issues, we collect a large amount of Chinese network security data and annotated them as a dataset. We propose and apply an entity extraction model to efficiently and accurately extract entities with network security semantics from unstructured text data, achieving the automated construction of a knowledge graph.

2 Related Work

With the steady rise in network attacks in recent years, the creation of knowledge graphs in the area of network security has steadily gained attention.

A history-based structured learning method that simultaneously extracts the entities and relations in a sentence is proposed by Miwa and Sasaki [11]. Bekoulis et al. [3] uses a CRF (Conditional Random Fields) layer to represent the named entity identification task and approached the relation extraction task as a multi-class classification problem. On several datasets with diverse languages and situations, they run trials to show the efficiency of their model. Eberts and Ulges [6] and others propose an attention model SpERT based on span-based joint entity and relation extraction. The model is trained using entity negative samples and relation negative samples, which helps enhance the model's robustness.

Liao [9] proposes a method that utilizes regular expressions and similarity comparison to automatically extract Indicators of Compromise (IOCs) from blog forums, such as malware signatures and botnet IPs. The advantage of this method is that it uses contextual security keywords to locate potential entities and relations. Balduccini et al. [2] combines security ontologies with regular expressions to extract

semi-structured security data and achieves good F1 performance, but this method is not suitable for extracting unstructured text data.

Jia [7] and others propose a five-tuple network security ontology model centered around three dimensions: assets, vulnerabilities, and attacks, and deduce new inference rules using the path ranking algorithm. Pastuszuk et al. [1] designs an integrated ontology model for network security data and dynamic knowledge sources, taking into account the dynamic and highly variable nature of IT systems.

Du [5] and his team integrate multiple sources of vulnerability intelligence data, design a vulnerability knowledge graph framework, and use a deep learning model to extract entities and attributes. Based on rules, they extract relationships to construct a vulnerability knowledge ontology for inspection and analysis, resulting in a multi-source knowledge graph. Wang et al. [13] proposes a knowledge graph completion method based on ensemble learning and adversarial training. They use operations such as projection and rotation to model the relationships between entities. At the same time, in order to address the issue of poor robustness of entity extraction models, they use cooperative adversarial training methods to improve the robustness and generalization ability of the extraction model.

3 Knowledge Graph Construction

This section designs a program to automate the creation of knowledge graphs. From semi-structured and unstructured data, it is able to extract entities, relationships, and properties related to network security. Ontology modeling, data acquisition, entity extraction, and knowledge storage are the primary steps of the procedure.

3.1 Ontology Modeling

An ontology is a formal description that combines concrete data and domain knowledge. The deep connections between network security entities can be described normatively, methodically, and thoroughly by building a complete network security ontology using knowledge and data. This work builds a network security ontology model based on knowledge about the subject, as illustrated in Fig. 1, and saves it using the Protégé tool.

3.2 Data Acquisition

For the purpose support its investigation of entity extraction methods and related knowledge graph creation in the area of network security, this paper relies on a sizable amount of Chinese and English text data as well as dictionaries. However, there aren't

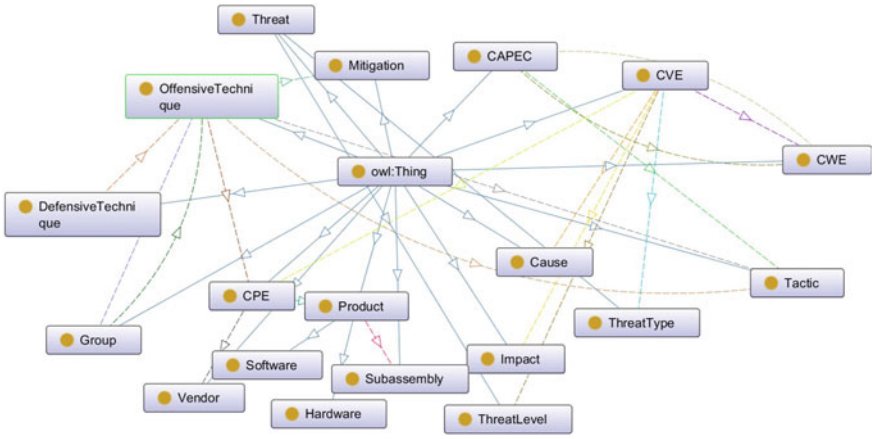


Fig. 1 Network security ontology model

many publicly accessible dictionaries and datasets for Chinese and English network security right now. This article collects 46,018 semi-structured and unstructured Chinese network security data from publicly available security websites such as the National Information Security Vulnerability Sharing Platform and the National Internet Emergency Center, as experimental data required for research.

3.3 Entity Extraction

A pre-training fine-tuning model called BERT [4] was proposed in 2018. The bidirectional TransFormer, on which its general architecture is based, theoretically outperforms the static Word2Vec language model, the dynamic LSTM-based ELMo model, and the unidirectional TransFormer GPT model in terms of performance.

With the goal do tasks like entity recognition and classification down the road, LSTM may extract characteristics from textual terms like position, context information, entity kind, etc.

To fully comprehend the sequence data, Bi-LSTM (bidirectional LSTM), an extension of LSTM, may model sequence data bidirectionally by using information from before and after the current time step. Other models can be integrated with LSTM. Entity extraction may effectively represent input text by combining BERT and Bi-LSTM. Representative feature learning and sequence modeling can combine the benefits of both models, improving the results of the extraction process.

The basis of the model presented in this research is a pre-trained BERT model. A Bi-LSTM model is then used to further extract text features from the sequence produced by the BERT model. Finally, the most representative features and entity width features are produced using the max-pooling and width encoding procedures, respectively. The overall structure of the model is shown in Fig. 2.

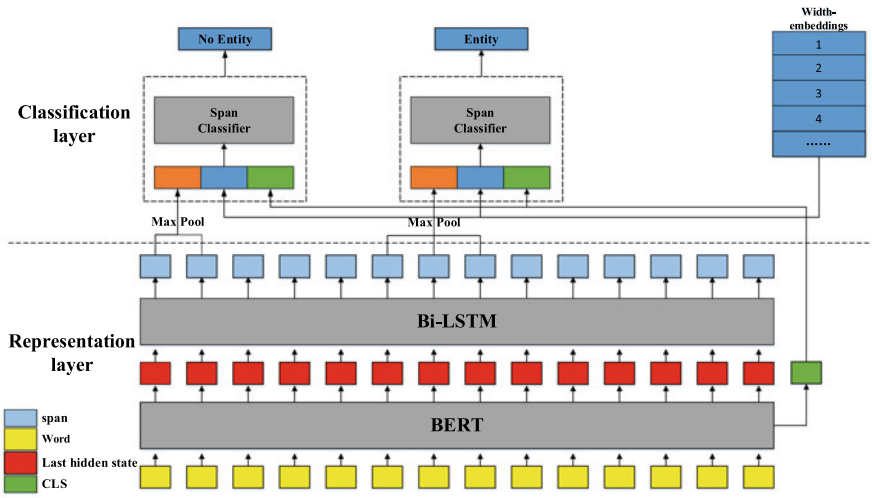


Fig. 2 Entity extraction model

3.3.1 Representation Layer

BERT and Bi-LSTM models make up the model's input representation section. First, a sequence of word vectors is created by applying the BERT model to create representations for each word in the text sequence. The BERT model also adds a few special token embeddings, like "[CLS]", to specify information like the beginning and end positions in the sequence. "[CLS]" is primarily used to capture the semantic information of the entire sentence and generate a semantic vector for the entire sentence, known as the "CLS token embedding".

The word vectors generated by the BERT model are input into a Bi-LSTM model to effectively model contextual information and the mutual influence between words in the sentence. The Bi-LSTM model extracts the word's vector representation-corresponding output state for each word. A span classifier uses this output state, or vector sequence, as input to generate labels for potential spans.

BERT is largely in the role of capturing contextual information and semantic linkages in the input representation layer, whereas Bi-LSTM is the duty of modeling sequence information and word dependencies. Their benefits can be fully leveraged by combining the two to enhance entity extraction. In particular, BERT can more precisely model word sequences and capture dependencies and semantic linkages between words, whereas Bi-LSTM can more accurately capture contextual information of words and spread it among vocabulary.

3.3.2 Classification Layer

The span classifier in the classification layer can classify the entity type of the candidate span. Its input is any candidate span from the Bi-LSTM. For example, $E = (e_i, e_{(i+1)}, e_{(i+2)}, \ldots, e_{(i+k)})$represents an entity span consisting of the character set from i to $i + k$.

An efficient method for improving the model's robustness when the sample is negative. In contrast to conventional random sampling techniques, our negative samples primarily concentrate on the entity span surrounding the positive samples, which helps to increase the accuracy of entity span border prediction in real-world applications.

In order to extract the most essential information from the subvector features produced by Bi-LSTM, we employ the max-pooling method. We use P_{max} to represent the max-pooling operation, and Eq. (1) represents the feature representation after the max-pooling operation.

$$f(E) = [P_{\max}(e_i, e_{(i+1)}, e_{(i+2)}, \ldots, e_{(i+k)})] \tag{1}$$

Let w_k represent the encoding feature of the length k, then Eq. (2) represents the feature representation after the operation of width embeddings, where the symbol ";" denotes the vector concatenation operation.

$$f(E) = [P_{max}(e_i, e_{(i+1)}, e_{(i+2)}, \ldots, e_{(i+k)}); w_{(k+1)}] \tag{2}$$

The main function of "[CLS]" is to encode the entire sentence. By combining the candidate entity span with "[CLS]", the contextual information of the entity can be more comprehensively expressed. Eq. (3) represents the feature representation obtained by concatenating the "[CLS]" token vector, where represents the "[CLS]" vector.

$$f(E) = [P_{max}(e_i, e_{(i+1)}, e_{(i+2)}, \ldots, e_{(i+k)}); w_{(k+1)}; c] \tag{3}$$

Then $f(E)$ is used as the input of the span classifier, which performs label classification using fully connected layers.

$$y = f(E)A^T + b \tag{4}$$

Finally, the cross-entropy loss function is used to calculate the loss for the classification results.

$$l_n = -W_{y_n} \log \frac{\exp(x_{n,y_n})}{\sum_{c=1}^{C} \exp(x_{n,c})} \bullet 1\{y_n \neq ignore\_index\} \tag{5}$$

In Eq. (5), $x_{(n,c)}$ represents the score of sample n belonging to category c, y_n represents the true category of sample n, W is a weight vector, and C is the total

Fig. 3 Display of a partial network security knowledge graph

number of categories. The probability of sample n being assigned to each category is calculated by the softmax function, and then the cross-entropy loss is computed using the probability vector and the one-hot vector composed of the true labels. If y_n equals $ignore\_index$, the loss of the sample is 0, which means that the loss of the sample is ignored.

3.4 Knowledge Storage

This paper creates a Python program to extract entities and relationships, and maps them to the associated idea nodes in the ontology model, to create a knowledge graph from semi-structured data and unstructured text. For unstructured text, we use sentence segmentation to break the article into sections. This increases the model's efficacy while lowering the computing cost needed for enumerating entity spans and categorization. A knowledge graph is then created by concluding the connections between concept nodes and creating the interconnections between entities, as seen in Fig. 3.

4 Experiments

In this section, we compare our model with other entity extraction models on different datasets, including ADE, CoNLL04, and SciERC public datasets, as well as a network security dataset that we collected and annotated ourselves.

ADE: The ADE dataset refers to a dataset for the task of extracting adverse drug events, which was published in 2009. The dataset is mainly used to evaluate natural language processing algorithms for identifying drug side effects from bio-medical texts. The dataset contains 42,720 sentences and two entity types: "Adverse-Effect" and "Drug".

CoNLL04: The CoNLL04 dataset refers to a named entity recognition task dataset that was released at the Conference on Computational Natural Language Learning (CoNLL) in 2004. The dataset is mainly used to evaluate whether named entity recognition models can correctly extract entities from news articles. The dataset contains 1336 sentences and four entity types: "Location", "Organization", "People", and "Other".

SciERC: The SciERC dataset is a dataset released for the task of entity relation extraction in scientific literature. It is based on the CoNLL-2003 dataset and is especially expanded and improved for scientific literature. The sentences in the SciERC dataset come from 16 different fields of scientific literature. The dataset contains 2578 sentences and six entity types: "Generic", "Metric", "OST", "Material", "Method", and "Task".

Securitydataset: The Securitydataset is a dataset we collected and labeled ourselves. The data comes from text data on publicly available websites such as the National Information Security Vulnerability Sharing Platform and the National Internet Emergency Center. We randomly selected 3000 sentences for labeling as the data required for the experiment. The entity types include seven types: "Vendor", "Product", "Version", "Impact", "Vulnerability", "Subassembly" and "Cause" (Table 1).

Our model exhibits exceptional performance in entity extraction tasks on the ADE dataset, surpassing other models in all performance metrics. This success can be attributed to several crucial characteristics of the ADE dataset. Firstly, the dataset specifically focuses on drug-induced adverse events, resulting in a relatively narrow domain, simple data structure, and a limited number of entity types, namely "Adverse-Effect" and "Drug". This simplification facilitates the model's acquisition of domain-specific knowledge and features. Secondly, the ADE dataset contains a substantial number of samples with high-quality annotations, enhancing the model's ability to generalize, and improving performance on unfamiliar data.

Regarding the CoNLL04 dataset, our model achieves the highest Precision in entity extraction tasks. Additionally, on the SciERC dataset, our model outperforms other models across all metrics, including a 3.11% increase in Precision, a 2.27% increase in Recall, and a 2.37% increase in the F1 score.

Overall, these results demonstrate that our model is highly effective in entity extraction tasks, particularly on large-scale datasets. We believe that this model can be applied to various natural language processing fields to provide strong support for entity extraction tasks.

We compare the performance of our model with other models on a Chinese dataset collected and annotated from the network security domain. Experimental data show that while the BERT+Bi-LSTM model combines two useful techniques, its per-

Table 1 Test set results. The experimental data of the models on ADE, CoNLL04, and SciERC are cited from the reference experimental data

Dataset	Model	Precision	Recall	F1
ADE	Multi-head [3]	84.72	88.16	86.40
	BiLSTM+SDP [8]	82.70	86.70	84.60
	SpERT [6]	88.99	89.59	89.28
	Ours	**95.72**	**95.91**	**95.63**
CoNLL04	Multi-head [3]	83.75	84.06	83.90
	Table-filling [11]	81.20	80.20	80.70
	SpERT [6]	85.78	**86.84**	**86.25**
	Ours	**85.91**	84.58	85.79
SciERC	SciIE [10]	67.20	61.50	64.20
	SpERT [6]	70.87	69.79	70.33
	Ours	**73.98**	**72.06**	**72.70**
Securitydataset	BERT+Bi-LSTM(without max pool and width embedding)	59.22	63.15	60.61
	SpERT	87.57	87.14	86.98
	Ours(avg pool)	86.44	86.44	85.92
	Ours(max pool)	**88.03**	**88.12**	**87.75**

formance is relatively low. In contrast, the SpERT model performs better in entity extraction tasks.

It is worth mentioning that all performance metrics of our model are higher than those of SpERT, indicating that adding a Bi-LSTM layer on the basis of SpERT can further improve entity extraction performance.

To verify that max-pooling performs better than average-pooling in entity extraction tasks, we also conduct entity extraction experiments in average-pooling mode. Through the results of two sets of experiments, Ours (max pool) and Ours (avg pool), it is shown that compared to average pooling, the max pooling method better preserves the important information of entities, with each metric improving by more than 1.5%. Therefore, it performs better in entity extraction tasks. Max pooling is better at capturing the most representative information in entities while ignoring the effects of other information, thus enabling more accurate entity recognition. On the other hand, average pooling cannot effectively distinguish between important and less significant feature points of entities, leading to information mixing and affecting the effectiveness of entity extraction.

5 Conclusions

This article focuses on the problem of a lack of Chinese datasets in the field of network security and proposes three main improvements for the construction of a network security knowledge graph. Firstly, a large amount of network security-related data is crawled through web crawler technology and annotated for subsequent research, providing a rich Chinese dataset. Secondly, a deep learning-based entity extraction model is designed and implemented based on this dataset and experimentally verified to have superior performance. Lastly, an entity relationship is extracted and presented based on the constructed network security knowledge graph using ontology concept nodes, which supplements and presents relationships between entities despite not using relationship extraction technology.

For future improvements, we will focus on the following directions: (1) further improve and expand the Chinese dataset to improve the model's generalization ability; (2) attempt to use relationship extraction technology to accurately extract and represent relationships between entities; (3) explore the application of network security knowledge graphs in other natural language processing tasks, such as text classification and sentiment analysis, to achieve cross-domain applications; (4) consider applying knowledge graphs to actual network security problems to help related industries and enterprises improve their network security capabilities.

Acknowledgements This research was supported by the National Natural Science Foundation of China (NO. 61936005) and the Key Laboratory Fund(NO. 61421010202).

References

1. Pastuszuk J, Burek P, Ksiezopolski B (2021) Cybersecurity ontology for dynamic analysis of IT systems—sciencedirect
2. Balduccini M, Kushner S, Speck J (2015) Ontology-driven data semantics discovery for cybersecurity. Springer, New York
3. Bekoulis G, Deleu J, Demeester T, Develder C (2018) Joint entity recognition and relation extraction as a multi-head selection problem. Exp Syst Appl 114:34–45
4. Devlin J, Chang MW, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding
5. Du L, Xu C (2022) Knowledge graph construction research from multi-source vulnerability intelligence. In: Lu W, Zhang Y, Wen W, Yan H, Li C (eds) Cyber security. Springer, Singapore, pp 177–184
6. Eberts M, Ulges A (2019) Span-based joint entity and relation extraction with transformer pre-training. *arXiv*
7. Jia Y, Qi Y, Shang H, Jiang R, Li A (2018) A practical approach to constructing a knowledge graph for cybersecurity. Engineering 4(1, SI):53–60
8. Li F, Zhang M, Fu G, Ji D (2017) A neural joint model for entity and relation extraction from biomedical text. BMC Bioinf 18
9. Liao X, Yuan K, Wang X, Li Z, Xing L, Beyah R (2016) Acing the ioc game: toward automatic discovery and analysis of open-source cyber threat intelligence. In: ACM SIGSAC conference on computer and communications Security

10. Luan Y, He L, Ostendorf M, Hajishirzi H (2018) Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction
11. Miwa M, Sasaki Y (2014) Modeling joint entity and relation extraction with table representation. In: Conference on empirical methods in natural language processing
12. Sikos LF (2023) Cybersecurity knowledge graphs. Knowl Inf Syst 2023
13. Wang P, Liu J, Hou D, Zhou S (2022) A cybersecurity knowledge graph completion method based on ensemble learning and adversarial training. Appl Sci 12(24)

MINEPAT: Mine Path and Tail Entity by Reinforce Learning and Graph Attention Mechanism

Zhiqiang Teng, Youyun Xu, Xuan Chen, and Jian Chen

Abstract Although Knowledge Graph (KG) has a wide range of applications, most KGs suffer from incompleteness, which seriously hinders practical application. Knowledge Reasoning (KR) is one of the key methods to solve this problem. However, many existing models cannot do completion tasks, or the completion accuracy is too low. Therefore, MINEPAT is proposed to make up for it, which transforms reasoning into a sequential decision-making problem. It uses Reinforce Learning (RL) to choose the next relation and Graph Attention Mechanism (GAT) to choose the next entity. Besides, a stepwise reward strategy is proposed to alleviate the effect of sparse reward values. Experimental results show that the prediction accuracy of MINEPAT outperforms other models in completion tasks, especially in HITS@1.

Keywords Knowledge graph · Knowledge reasoning · Graph attention mechanism · Reinforce learning

1 Introduction

Massive volumes of data have been generated due to the rapid expansion of the Internet of Things (IoT) and Cloud Computing, and how to organize and utilize these data has become a challenge. In 2012, Google gave their answer—Knowledge Graph (KG), whose concept originated from the Semantic Web. Afterward, a large number of KGs emerged, including NELL [3], Freebase [1], and WordNet [9]. However, both manually and automatically constructed KGs suffer from incompleteness. Some of these are caused by incomplete information during construction, while others are

Z. Teng (✉) · Y. Xu · X. Chen · J. Chen
School of Communication and Information Engineering Nanjing University of Posts and Telecommunications, Nanjing 210003, China
e-mail: 1021010311@njupt.edu.cn

J. Chen
e-mail: chenjian1980@njupt.edu.cn

National Engineering Research Center of Communications and Networking, Nanjing 210003, China

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_60

implicit relations between the entities. Although the former cannot be completed through the existing triples of self, some of the latter can be completed through Knowledge Reasoning (KR).

KR, as one of the important methods for completing KG, has received widespread attention in recent years. A large number of excellent models have emerged. It includes three technical routes: rule-based KR [15], embedding-based KR [2], and path-based KR [7], to which MINEPAT belongs. Path-based KR has produced a batch of models based on reinforcement learning, such as the DeepPath [14] which first introduced Deep Reinforce Learning (RL) into KR, and the AttnPath [11] that introduced Long Short Term Memory (LSTM) and Graph Attention Mechanism (GAT) into KR. But both focus on finding reasoning paths for specific relations to assist in judging the truthfulness of given triples and cannot do completion tasks. Meanwhile, both only focus on the relations and ignore the entities during reasoning. Although the MINERVA [4], can give the tail entity, it fixed the searching step length and restricted the next-hop range of each entity, resulting in the loss of graph information during reasoning. And like DeepPath, it adopts the strategy of unified reward for the entire path, which results in sparse reward value. The above points lead to low completion accuracy.

Therefore, in this paper, we propose a new model, MINEPAT, which aims to make up for the shortcomings of the models mentioned above to improve the accuracy in completion tasks. Our contributions can be mainly divided into three parts as follows:

Firstly, our model can not only give a tail entity for an incomplete triple but also can mine potential reasoning paths;
Secondly, our model considers not only the relations but also the entities and graph structure information during reasoning;
Thirdly, we propose a new reward mechanism that effectively mitigates the impact of sparse reward values.

The rest of the paper is organized as follows. The relevant work of KR is introduced in the next section. In Sect. 3, the MINEPAT is introduced in detail. Section 4 shows the experimental results and conducts qualitative analysis. Finally, Sect. 5 concludes and looks forward.

2 Related Works

Bordes et al., first proposed the Embedding-Based model TransE in 2013 [2]. The main idea is that if a triple holds, then in the embedding space, the Euclidean Distance between the sum of the head entity embedding and the relation embedding, and the tail entity embedding should be smaller, otherwise it is larger. Ideally, the Euclidean Distance should be 0 if it holds. However, in KGs, there are many 1-to-N, N-to-1, and N-to-N relations, which can cause conflicts between entities. To mitigate this problem, Wang et al., proposed the TransH [12], of which each relation not only has a representation embedding but also has a projection embedding that translates entities to the hyperplane. But The entities and relations are still in the same space. Lin

et al., proposed the TransR [8], which represents entities and relations in separate spaces, and each relation has a matrix that projects entities to the corresponding space. Although it has a stronger representation ability, the sharp increase in model parameters makes the convergence speed of the model much slower than that of TransE and TransH. To simplify the complexity of the TransR and improve its training speed, Ji et al., proposed the TransD [6]. However, the above models only use the information of entities and relations, without the structure information. Xiong et al., proposed the DeepPath [14] in 2017, which uses an agent to mine the inference path. It designs a reward mechanism that integrates accuracy, diversity, and efficiency. Wang et al. [11] believed that DeepPath doesn't consider the influence of the taken paths during reasoning, so they incorporate the LSTM network to encode the information of taken paths and proposed the AttnPath. However, it has two common problems like DeepPath: the input to the agent requires the tail entity embedding, which makes them fail to predict the tail entity for incomplete triple; secondly, both models use a random selection strategy for multiple next entities after selecting the next relation, ignoring the entity information during reasoning. Based on DeepPath, Das et al. [4] proposed the MINERVA, which uses an LSTM network to assist in predicting the tail entity, but it reduces the connection relations of each entity and fixes the searching step, resulting in losing KG information during reasoning. Hildebrandt et al. [5] proposed the R2D2 model for judging truthfulness for a triple, which uses two agents to collect evidence chains and integrates all evidence for final judgment, but it also cannot do completion tasks.

3 Methodology

In this section, firstly, we will introduce the basic elements of the RL Agent and the Environment. Then, we will introduce the application of GAT in utilizing the information of entities and structure information. Finally, we will briefly introduce the training process of MINEPAT and the gradient of its loss function.

3.1 Environment and RL Agent

In MINEPAT, we transform RL reasoning into sequential decision-making. We will introduce Environment, State, Action, and Reward respectively. Among them, the Reward part will introduce a new reward mechanism.

The structure of our model is shown in Fig. 1. **Left**: The KG environment, which includes the triples of the KG. **Top Right**: The structure of RL Agent, consists of three fully connected layers. The output of the last layer is through the Softmax layer to normalize the probability of actions. **Bottom Right**: The influence factor of the neighbor entities of the current entity are probabilistically normalized by the Softmax layer for choosing the next entity.

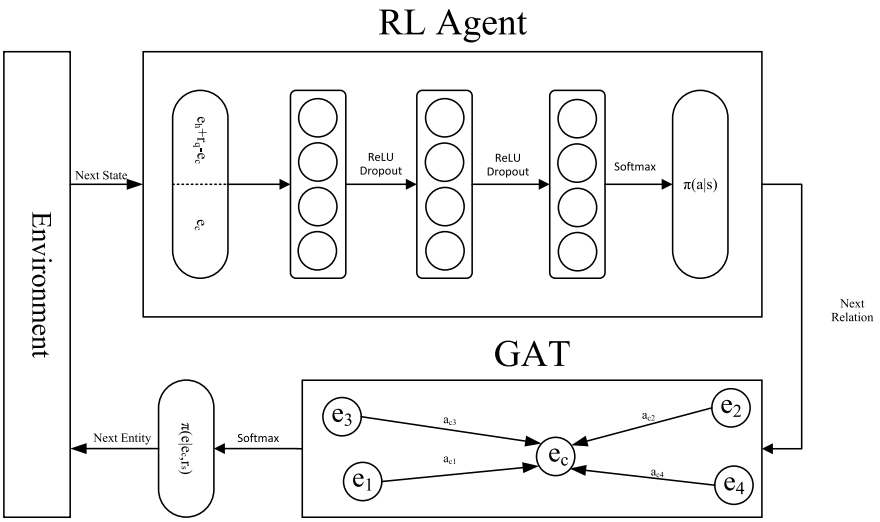


Fig. 1 Overview of our model

Environment: In our model, Environment refers to the entire KG. We also add reverse triples to the Environment like others. Of course, during training or testing, the Environment does not include the current triple and its reverse triple.

State: State is the input of the Agent. Intuitively, when choosing the next action, we need to consider the head entity e_h, query relation r_q, the current entity e_c, and the distance from the target tail entity e_t. In completion tasks, we don't know the specific tail entity because this is exactly what we want to reason. Inspired by the TransX series models, we sum the e_h and the r_q to represent the ideal tail entity embedding. Therefore, the State is divided into two parts: the current entity embedding and the distance to the ideal target tail entity, $s = [e_c; e_h + r_q - e_c]$.

Action: Essentially, our model transforms reasoning into a sequential decision-making problem. The decision at each step is divided into two parts: choosing the next relation, for which is the RL Agent responsible, and then choosing the next entity, for which is the GAT responsible. Therefore, the Action here refers to the next relation. The RL Agent outputs a normalized probability distribution of all relations, on which the agent chooses the next action. However, sometimes there may be no next entities connected under the next action selected by the RL Agent for the current entity e_c, which we call an invalid action. When encountering that, we will force a choice among valid actions.

Reward: The core of the RL is how the Agent takes a series of Actions in the Environment to achieve maximum cumulative rewards. The final reward is given based on whether the agent reaches the target tail entity e_t. Different from the existing models, we adopt a stepwise reward strategy. Intuitively, for a successful path, steps closer to the target tail entity are more reliable and should be rewarded more, while

steps closer to the head entity may contain errors that can be corrected in later decisions, so they have lower reliability and deserve a lower reward; the reward strategy for not reaching the target tail entity is completely opposite. To achieve the stepwise effect of the reward value, we used an isometric sequence.

3.2 Graph Attention Mechanism

The RL Agent only takes into account the relations. Besides, we also need to consider the entities and the rich information inherent in the structure of the KG. Therefore, we use the GAT [10] to assist us in choosing the next entity. Meanwhile, we believe that the structure information under different relations is different, so we trained a mapping matrix W and a weight vector $\mathbf{a}$ for each relation separately. The a_i, denoting the influence factor between the current entity e_c and the i-th entity e_i connected by the next action, is:

$$a_i = LeakyReLU\left(\mathbf{a}\left[We_c^T; We_i^T\right]\right) \tag{1}$$

Finally, the influence factors of all neighbor entities are normalized through a Softmax layer, and a random selection is made based on the output probability distribution.

3.3 Training Pipeline

We believe that the attention points of entity embedding are inconsistent for different r_q tasks. To balance the effectiveness and complexity, we use the TransH as pre-trained embeddings. In each r_q task, we project all entities to corresponding hyperplanes. Our model is divided into two parts: RL and GAT. Although the two parts can be combined into one model, for the convenience of comparing the effectiveness of GAT, we train them separately. For each training sample, the parameters of the RL Agent and GAT policy networks are updated using the REINFORCE algorithm [13] after rewarded. In the following formulas, the θ represents the parameters of the policy network and the T is the length of the path. The gradient of the RL Agent is as follows:

$$\nabla_\theta J(\theta) \approx \nabla_\theta \sum_{t=1}^{T} log\pi_\theta (a = r_t \mid s_t) \tag{2}$$

Where r_t is the relation, s_t the state, and $\pi_\theta (a = r_t \mid s_t)$ the probability of the relation of the path.

The gradient of the GAT is as follows:

$$\nabla_\theta J(\theta) \approx \nabla_\theta \sum_{t=1}^{T} log\pi_\theta (e = e_t \mid e_{t-1}, r_t) \tag{3}$$

Where r_t is the relation, e_t the entity, and $\pi_\theta (e = e_t \mid e_{t-1}, r_t)$ the probability of the entity of the path. Especially, e_0 represents the head entity.

Like DeepPath and AttnPath, during RL, we save the successful paths found by the RL Agent after removing loops. These reasoning paths can help identify whether a given triple is true or false.

4 Experiments

In this section, we will compare the performance of MINEPAT against other models through three tests: HITS@N, LinkPrediction, and FactPrediction. Although MINEPAT is mainly used for completion tasks, it can also mine potential reasoning paths like DeepPath and AttnPath, so we conduct LinkPrediction and FactPrediction on it. The HITS@N will compare the success rate in reaching valid tail entities, and the LinkPrediction and the FactPrediction will compare the MAP.

4.1 Dataset and Settings

We test on the NELL-995. Table 1 shows the basic information. It's created by Xiong [14], which is more suitable for testing the performance of multi-hop inference. It removes some of the auxiliary illustrative relations that are helpless in reasoning, such as ***haswikipediaurl***. Meanwhile, to further improve the convergence speed of RL Agent, it deletes the relations with low frequency and their triples from the KG and only retained the top 200 relations and their triples. Meanwhile, we also add the reverse triples like others, so there are 400 relations. When training and testing, we remove the current triple and its reverse triple from the KG.

We set the embedding size to 50. The dimensions of the three fully connected layers of the RL Agent are set to 512, 1024, and 400 respectively. The attention of the GAT is set to 50, and the negative slope of the leakyReLU is set to 0.1. The learning rate of the optimizer is set to 0.001, of which the weight_decay is set to 0.005. The dropout is set to 0.2. We used one RTX 4090 to train the model.

Table 1 Base Information of the NELL-995.

Dataset	Entity	Relation	Relation-1to1	Relation-1toN	Triple
NELL-995	75492	400	59	341	308426

Table 2 HITS@N results. The results of ComplEx, ConvE, DistMult, Path-Baseline, and MINERVA are cited from Das [4]

Metric	ComplEx	ConvE	DistMult	Path-Baseline	MINERVA	MINEPAT
HITS@1	0.612	0.672	0.610	0.300	0.663	**0.787**
HITS@3	0.761	**0.808**	0.733	0.417	0.773	0.795
HITS@10	0.827	**0.864**	0.795	0.497	0.831	0.813

Each row represents a metric, and the bold font in each row indicates that the model in its column achieved the best performance in that metric

4.2 HITS@N

We compare MINEPAT against several other models using HITS@1, 3, 10 respectively, which are standard metrics for KG completion tasks. We take all the positive triples in the test set used by Xiong[14] for FactPrediction and LinkPrediction to build the test set for HITS@N. We let the MINEPAT try ten times and limit the searching step length to 50. We evaluate the average performance of numerous trained relations against other models because our model trains each relation separately. The specific performance for each relation can be seen in Sect. 4.5.

From Table 2, we can see that MINEPAT is substantially ahead of other models by about 10% points in HITS@1, but slightly inferior to the best, ConvE, in HITS@3. And MINEPAT performs worse in HITS@10. On the whole, the performance of HITS@1 and HITS@10 of our model has the smallest difference. The main reason is that our RL Agent has a better ability to learn paths with higher frequency in training, which is strengthened under the influence of the stepwise reward strategy, but it's not good at mining diverse paths. According to the analysis of the experimental results, although ten attempts, the same tail entity tends to be given, and when the first completion is wrong, it is difficult to jump out of it for the remaining times. But in practice, we pay more attention to the performance of HITS@1. At this point, our model outperforms other models by a large margin and even exceeds other models' performance of HITS@3, which is encouraging.

4.3 FactPrediction

The FactPrediction is mainly used to test the ability of the model to judge the truthfulness of a given triple. We used the same test set as Xiong [14], which constructed

Table 3 FactPrediction MAP.

TransE	TransH	TransR	TransD	DeepPath	MINEPAT
0.382	0.386	0.411	0.409	0.495	**0.556**

Each row represents a metric, and the bold font in each row indicates that the model in its column achieved the best performance in that metric

Table 4 LinkPrediction MAP

TransE	TransR	DeepPath	AttnPath	MINEPAT
0.728	0.776	0.821	**0.842**	0.825

Each row represents a metric, and the bold font in each row indicates that the model in its column achieved the best performance in that metric

ten negative triples for each positive triple. For each sample, we use reasoning paths found in training to validate. If it can reach the target tail entity, it gets 1 point, otherwise 0 points. At the same time, we use the reciprocal of the path length as the weight, and the final score is the sum of all the scores. Finally, the scores of all samples are sorted in descending order, and the reciprocal of their rank is the AP. The results are shown in Table 3, from which we can see that MINEPAT is better than the DeepPath and significantly ahead of the TransX series models. When DeepPath encounters an invalid action, it will stay in the same place until the next action, while we force the selection among valid actions, so more reasoning paths can be mined.

4.4 LinkPrediction

The DeepPath and AttnPath cannot directly give the tail entities for the incomplete triples, so Hits@N does not apply to them. In this test, we use the same test set as for FactPrediction. For each relation, the samples with the same head entity but different tail entities are taken as a group. The feature vector containing whether or not each path is reachable is fed into the fully connected neural network for true or false binary classification. The score of the positive label is sorted in descending order, and the reciprocal of their rank is the AP. The results are shown in Table 4. We can see that the performance of MINEPAT is comparable to DeepPath, but there is a significant gap compared with AttnPath. The analysis shows that AttnPath uses forced next hop and weighted reward, which can encourage the model to discover more efficient inference paths, so the overall performance is better than ours. Although DeepPath also uses weighted reward, there is no forced next hop when encountering invalid action, so the performance is similar to ours.

Table 5 Results of some relations in HITS@N

Tasks	AT	Our model without GAT			Our model with GAT		
		HITS@1	HITS@3	HITS@10	HITS@1	HITS@3	HITS@10
athleteHomeStadium	1	0.754	0.754	0.764	0.764	0.764	0.764
athletePlaysInLeague	1	0.636	0.636	0.655	0.706	0.709	0.724
athletePlaysSport	1	0.817	0.823	0.837	0.857	0.883	0.913
orgHeadquarterCity	1	0.837	0.846	0.855	0.846	0.846	0.864
orgHiredPerson	1.671	0.749	0.780	0.810	0.782	0.799	0.820
personBorninLocation	1	0.616	0.638	0.984	0.644	0.655	0.701
personLeadsOrg	1.417	0.863	0.868	0.868	0.863	0.872	0.881
teamPlaysInLeague	1	0.786	0.792	0.814	0.801	0.807	0.830
worksFor	1.319	0.796	0.796	0.800	0.820	0.822	0.822
Overall	–	0.762	0.770	0.786	0.787	0.795	0.813

4.5 Qualitative Analysis

GAT enables both entity information and structured information to be taken into account during reasoning. To validate its effectiveness, we analyze the performance of different relations with and without GAT in the completion tasks. The results are shown in Table 5. AT means the average tail entities of the relation per head entity. We can see that GAT improves the performance of the model by 2.5% points on average. 2.5% points may seem trivial. After analysis, we found that the number of (head, relation) is 137,662 out of a total of 308,426 triples. On average, each (head, relation) has 2.24 tail entities. On the whole, 110814 (80.5%) (head, relation) has only 1 tail entity, so in most cases, we only have one next entity to choose from, and it does not matter whether GAT is present or not. When encountering multiple entities, the prediction accuracy of the model is improved by the GAT.

5 Conclusion and Future Work

In this paper, we propose a novel model, MINEPAT, which utilizes RL and GAT to do completion tasks and simultaneously mines potential reasoning paths for each relation. We use not only relations but also entities and structural information for reasoning. In addition, we also propose a stepwise reward strategy to alleviate the impact of sparse rewards. The experimental results and qualitative analysis show the excellent performance of MINEPAT against other models, especially in HITS@1.

In the future, we are interested in making one model apply to all relations simultaneously without training separately. At the same time, we would like to make the model predict multiple tail entities for the 1toN type relation.

Acknowledgements We gratefully acknowledge the support of the National Natural Science Foundation of China (NO. 61936005) and the Key Laboratory Fund (NO. 61421010202).

References

1. Bollacker K, Evans C, Paritosh P, Sturge T, Taylor J (2008) Freebase: a collaboratively created graph database for structuring human knowledge. In: Proceedings of the 2008 ACM SIGMOD international conference on management of data, SIGMOD '08, New York, NY, USA, 2008. Association for Computing Machinery, pp 1247–1250
2. Bordes A, Usunier N, Garcia-Duran A, Weston J, Yakhnenko O (2013) Translating embeddings for modeling multi-relational data. In: Proceedings of the 26th international conference on neural information processing systems, vol 2, NIPS'13, Red Hook, NY, USA, 2013. Curran Associates Inc, pp 2787–2795
3. Carlson A, Betteridge J, Kisiel B, Settles B, Hruschka E, Mitchell T (2010) Toward an architecture for never-ending language learning. In: Proceedings of the twenty-fourth AAAI conference on artificial intelligence, AAAI'10. AAAI Press, pp 1306–1313
4. Das R, Dhuliawala S, Zaheer M, Vilnis L, Durugkar I, Krishnamurthy A, Smola A, McCallum A (2017) Reasoning over paths in knowledge bases using reinforcement learning, go for a walk and arrive at the answer
5. Hildebrandt M, Serna JA, Ma Y, Ringsquandl M, Joblin M, Tresp V (2020) Reasoning on knowledge graphs with debate dynamics
6. Ji G, Liu K, He S, Zhao J (2016) Knowledge graph completion with adaptive sparse transfer matrix. In: Proceedings of the thirtieth AAAI conference on artificial intelligence, AAAI'16. AAAI Press, pp 985–991
7. Lao N, Mitchell T, Cohen W (2011) Random walk inference and learning in a large scale knowledge base. In: Proceedings of the conference on empirical methods in natural language Processing, EMNLP '11, USA. Association for Computational Linguistics, pp 529–539
8. Lin Y, Liu Z, Sun M, Liu Y, Zhu X (2015) Learning entity and relation embeddings for knowledge graph completion. In: Proceedings of the twenty-ninth AAAI conference on artificial intelligence, AAAI'15. AAAI Press, pp 2181–2187
9. Miller GA (1995) Wordnet: a lexical database for English. Commun ACM 38(11):39–41
10. Velickovic P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y (2017) Graph attention networks
11. Wang H, Li S, Pan R, Mao M (2019) Incorporating graph attention mechanism into knowledge graph reasoning based on deep reinforcement learning, pp 2623–2631
12. Wang Z, Zhang J, Feng J, Chen Z (2014) Knowledge graph embedding by translating on hyperplanes. In: National conference on artificial intelligence
13. Williams RJ (1992) Simple statistical gradient-following algorithms for connectionist reinforcement learning. Mach Learn 8(3–4):229–256
14. Xiong W, Hoang T, Wang WY (2017) Deeppath: a reinforcement learning method for knowledge graph reasoning
15. Yang F, Yang Z, Cohen WW (2017) Differentiable learning of logical rules for knowledge base reasoning. In: Proceedings of the 31st international conference on neural information processing systems, NIPS'17, pp. 2316–2325, Red Hook, NY, USA, 2017. Curran Associates Inc

Speech Command Recognition Algorithm Based on Improved MFCC Features

Ziyi Song and Lin Ma

Abstract In recent years, deploying low-power speech recognition algorithms on resource-constrained hardware devices has become a research focus. However, traditional algorithms rely on extensive data and complex models, which pose challenges for low-power devices with limited resources. This paper proposes an improved MFCC feature extraction method combined with the DTW algorithm, which enhances computational speed and reduces resource consumption in speech recognition. Experimental results validate the effectiveness and feasibility of the proposed method.

Keywords Mel frequency cepstral coefficient · Speech recognition · Dynamic time warping · Feature extraction

1 Introduction

In recent years, Speech command recognition holds significant importance in an increasing number of application domains. For instance, intelligent assistants, voice control systems, voice interactive devices, and voice navigation systems are prominent examples. In the field of speech recognition, traditional algorithms have often relied on methods such as Gaussian Mixture Models (GMM) and Hidden Markov Models (HMM) [1]. These methods' dependence on large-scale data and complex com-putational models has resulted in high demands for computational resources. For example, the GMM-HMM algorithm requires significant amounts of training data and computational power for model training and matching, posing challenges for implementing speech recognition on low-power devices. Low-power devices typically have limited computational capabilities and storage space, making it difficult to meet the high requirements of traditional algorithms.

Z. Song · L. Ma (✉)
School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, China
e-mail: malin@hit.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_61

Linear Predictive Coding (LPC) [2] is a modeling technique that simulates the human vocal tract and predicts future speech signals based on past signal frames. However, it has a limitation of assuming a linear relationship in speech production, failing to capture the full complexity of the vocal system. In comparison, Perceptual Linear Predictive (PLP) [3] analysis aims to minimize prediction error while modeling and predicting speech signals. Nevertheless, PLP's performance is diminished in the presence of non-stationary noise, despite its superior performance in stationary noise environments [4].

Mel Frequency Cepstral Coefficients (MFCC) [5] is a widely used front-end feature extraction technique in speech recognition [6]. It mimics the human auditory system by incorporating physiological sensitivity to different frequencies and psychological perception of human voices [7]. MFCC effectively reduces noise interference and enables both semantic recognition and speaker identification through its low-dimensional and high-dimensional features [8].

To address the demand for efficient and low-power speech recognition algorithms on low-power devices, researchers are dedicated to finding solutions that adapt to resource-constrained hardware environments. One key objective is to enhance computation speed and reduce resource consumption to meet the practical needs of low-power devices. Consequently, the development of speech recognition algorithms suitable for low-power hardware devices has become a crucial focus of current research.

The remainder of this paper is organized as follows. Section 2 introduces the system model. Section 3 presents the proposed robust tem-plate matching method based on enhanced MFCC. Section 4 provides experimental validation. Lastly, Sect. 5 presents the conclusions.

2 System Models

In order to accomplish speech command recognition, this paper adopts two steps: feature extraction and feature classification. The speech recognition algorithm utilizes an approach based on feature tem-plate matching, as illustrated in Fig. 1.

As shown in Fig. 1, the feature extraction process involves several steps, including pre-emphasis, framing, windowing, power spectrum computation, mel filtering, cepstral analysis, and differentiation. The feature classification step employs the DTW algorithm, which includes initialization, iterative cumulative difference computation, alignment path recording, and template training.

The process of voice command recognition is as follows: Firstly, the features are extracted from the input speech command, which represents a temporal characteristic that varies with the duration of the command. Secondly, the extracted temporal features of the input command are compared with the temporal features of known speech commands in the database using a matching algorithm. Each comparison generates a difference value, and this process is iterated for all known speech commands in the template library. Finally, by comparing all difference values, the command with the smallest difference value is selected as the final recognition result.

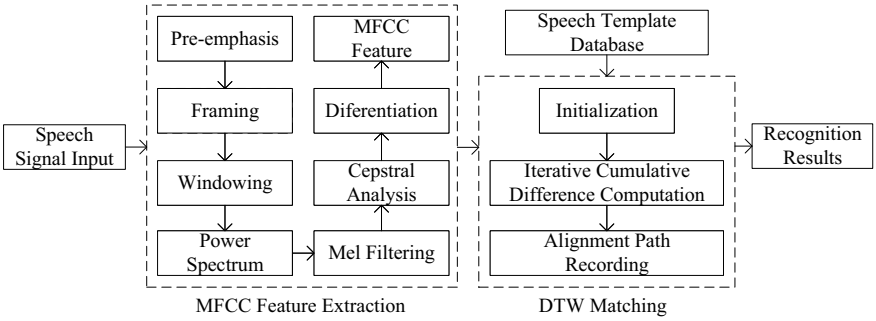


Fig. 1 Overall research framework

In detail, the features extracted from the speech signal are MFCC features, whose core essence is to model the human ear's auditory perception, but the human ear has different degrees of perception for different frequencies. Furthermore, after undergoing FFT transformation, speech commands still contain redundant information. To address this, a set of filters, known as Mel filters, can be employed to process the energy spectrum of speech signals, enabling compression and concentration of energy.

3 Proposed Method

3.1 Improved MFCC Features

In the time domain, high-pass filtering is typically implemented for preprocessing using Eq. 1:

$$y(n) = x(n) - ax(n-1) \tag{1}$$

where $0 \leq a \leq 1$ represents the pre-emphasis coefficient, and it has been found that $a = 0.97$ yields the best results in previous studies [5]. $x(n)$ represents the nth sampled point, while $y(n)$ represents the nth point after preprocessing.

According to Eq. 1, the entire pre-emphasis step involves N multiplications, where N is the number of points in the time domain of the speech command. The pre-emphasis coefficient is usually set to 0.97, which can be approximated by the hardware-friendly coefficient 31/32, as it is close to 0.968 [8]. This coefficient is referred to as hardware-friendly because it can be computed using shifts instead of multiplications.

The next step is frame segmentation, with frame step and frame length. After framing, a windowing operation is performed on the sequence. This paper employs the Hamming window function. The window function is multiplied element-wise with each frame in the time domain of the speech command, replacing the rectangular window.

The next step involves extracting the spectral information. The calculation formula for FFT is given by Eq. 2:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi nk}{N}}, 0 \leq k < N \tag{2}$$

where N is the number of points for FFT, which is set to 512 in this paper.

After obtaining the spectral representation of the speech signal in the time domain, the power spectral density can be calculated using Eq. 3:

$$Y(k) = \frac{|X(k)|^2}{N} \tag{3}$$

According to the Parseval's theorem, as shown in Eq. 4, the energy of one frame can be computed:

$$\sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X(k)|^2 \tag{4}$$

where $x(n)$ represents the time domain signal, $X(k)$ represents the frequency domain signal, and N is the number of FFT points.

By applying Eq. 5, the energy of one frame can be obtained:

$$E_i = \sum_{k=0}^{N-1} Y_i(k) \tag{5}$$

where E_i represents the energy of the ith frame, Y_i represents the energy spectral density of the ith frame, and N represents the total number of points in the frequency domain of one frame, which is the same as the number of points used in FFT.

Mel frequency is a nonlinear frequency scale that corresponds to the human auditory system. The conversion from Mel scale to analog frequency is given by:$m = 2595\lg(1 + f/700)$, $f = 700(10^{m/2595} - 1)$ where f represents the analog frequency in Hz, and m represents the Mel frequency in Mel scale.

Mel filters are commonly represented as triangles, and several filters form a filter bank, as shown in Fig. 2.

Each filter in the filter bank is a triangular filter which we simplify as triangular filterbanks. The filter calculation is given by Eq. 6:

$$H_i(k) = \begin{cases} \dfrac{1}{k_{i+1} - k_{i-1}} & k_{i-1} \leq k \leq k_{i+1} \\ 0 & \textbf{otherwise} \end{cases} \tag{6}$$

where $i = 1, \cdots, N_{\text{mel}}$ represents the ith Mel filter and k_i denotes the center frequency of the Mel filter transformed to digital frequency.

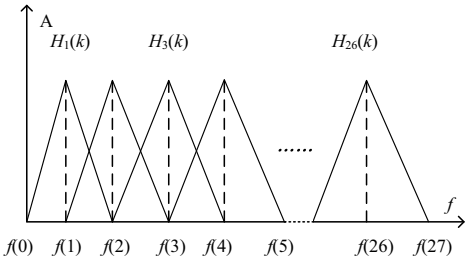


Fig. 2 Mel frequency

After performing the FFT transformation to the frequency domain, the product of the frequency domain function of the excitation signal and the frequency domain function of the transfer function is obtained. The use of logarithmic computation allows for the transformation of the multiplication relationship into an addition relationship, facilitating the separation of the two functions. Therefore, a logarithmic operation is applied to the energy spectrum: $\mathbf{L} = \ln(\mathbf{S})$. After that, do the discrete cosine transform:

$$c_i(m) = \sqrt{\frac{2}{N_{mel}}} \sum_{j=1}^{N} l_{ij} \cos\left(\frac{\pi m}{N_{mel}}(j - 0.5)\right) \tag{7}$$

where $i = 1, 2, \ldots, N_{\text{MFCC}}$, after applying the DCT operation, N_{MFCC} feature components are obtained. l_{ij}represents the ith row and jth column point of matrix $\mathbf{L}$.

3.2 Feature Classification

Firstly, initialize a difference matrix $\mathbf{D}$ for the two sequences, as defined in Eq. 8:

$$\mathrm{D}(i, j) = \sum_{k=0}^{N_{\text{MFCC}}-1} \left|\mathbf{X}_1{}^{(i)}(k) - \mathbf{X}_2{}^{(j)}(k)\right| \tag{8}$$

where $\mathbf{X}_1^{(i)}$ represents the ith frame of feature sequence $\mathbf{X}_1$, $\mathbf{X}_2^{(j)}$ represents the jth frame of feature sequence $\mathbf{X}_2$, and N_{MFCC} is the number of MFCC feature points per frame.

The meaning of each point in the difference matrix $\mathbf{D}$ corresponds to the distance difference between the frames of the two different sequences represented by the corresponding row and column. The difference matrix is $\mathbf{D}$ where each element of the matrix represents a point in the dashed grid.

Furthermore, a state matrix $\mathbf{F}$ and weighted matrix $\mathbf{W}$ are needed to record the minimum difference and weighting values to reach the current point. The initial

values of the first row and column of the state matrix **F** and are **W** set as follows:

$$\begin{cases} \mathrm{F}(i,0) = \mathrm{F}(i-1,0) + \mathrm{D}(i,0) \\ \mathrm{F}(0,j) = \mathrm{F}(0,j-1) + \mathrm{D}(0,j) \\ \mathrm{F}(0,0) = 0 \end{cases}, \begin{cases} \mathrm{W}(i,0) = \mathrm{W}(i-1,0) + W_1 \\ \mathrm{W}(0,j) = \mathrm{W}(0,j-1) + W_3 \\ \mathrm{W}(0,0) = 0 \end{cases} \tag{9}$$

where $i = 1, 2, \cdots, M_1 - 1$ and $j = 1, 2, \cdots, M_2 - 1$. W_1 and W_3 are the distance weighting coefficients, typically set as $W_1 = 1$ and $W_3 = 1$. These coefficients represent the weighting values for moving left and moving up, respectively.

Subsequently, the remaining points in the state matrix **F** are recursively computed according to Eq. 10:

$$\mathbf{F}(i,j) = \min\{F_1, F_2, F_3\}, \begin{cases} F_1 = \mathrm{F}(i-1,j) + W_1\mathrm{D}(i,j) \\ F_2 = \mathrm{F}(i-1,j-1) + W_2\mathrm{D}(i,j) \\ F_3 = \mathrm{F}(i,j-1) + W_3 D(i,j) \end{cases} \tag{10}$$

where $i = 2, 3, \cdots, M_1 - 1$, $j = 2, 3, \cdots, M_2 - 1$, W_1, W_2 and W_3 are distance weighting coefficients typically set as $W_1 = 1$, $W_2 = 2$ and $W_3 = 1$.

According to the transition rules, we recursively obtain $\mathrm{F}(M_1 - 1, M_2 - 1)$, which represents the minimum difference between the two time sequences of speech commands after time warping. In addition, it is necessary to normalize $\sum_{W_i \in path} W_i$ based on the weighted values along the path. The final difference value is given by:

$$\delta_{1,2} = \frac{\mathrm{F}(M_1 - 1, M_2 - 1)}{\sum_{W_i \in path} W_i} \tag{11}$$

where $\delta_{1,2}$ represents the difference value between sequence 1 and sequence 2.

After the above steps, we obtain only the minimum difference and do not have the alignment path. To facilitate subsequent processing for the average template, it is common to record the alignment path. When backtracking the path, we start from $(M_1 - 1, M_2 - 1)$ and move towards $(0, 0)$. The definition of the path matrix **P** given by Eq. 12:

$$\mathrm{P}(i,j) = \begin{cases} 1\ , \mathrm{F}(i,j) = \mathrm{F}(i-1,j) + \mathrm{D}(i,j) \\ 2\ , \mathrm{F}(i,j) = \mathrm{F}(i-1,j-1) + 2\mathrm{D}(i,j) \\ 3\ , \mathrm{F}(i,j) = \mathrm{F}(i,j-1) + \mathrm{D}(i,j) \end{cases} \tag{12}$$

There are two methods for template training: Casual template; Robust template training: To address the issue of poor recognition performance for different speakers using casual templates, a robust template training algorithm is employed. This algorithm relies on the alignment path, where the two training sequences are subjected to the DTW algorithm and the alignment path is recorded. Along this alignment path, the two sequences are averaged. The alignment path is recorded in the record matrix R, which keeps track of the coordinates of the current point on the alignment path:

Table 1 Accidental template difference values (bold indicates errors)

Speakers	Input Command	Template speech commands			
		Forward	Backward	Left	Right
A	Forward	61.53	64.11	63.39	63.94
	Backward	69.47	57.50	64.15	64.50
	Left	78.16	64.10	63.92	68.11
	Right	49.46	56.53	27.54	24.71
B	Forward	68.97	**60.28**	61.80	61.39
	Backward	67.35	54.42	59.52	60.81
	Left	71.23	64.60	62.85	63.16
	Right	66.94	**58.63**	61.49	66.03

$$\begin{cases} R(0,i) = p_x(i) \\ R(1,i) = p_y(i) \end{cases} \tag{13}$$

where $i = 1, 2, \ldots, N_p$, N_p representing the endpoint of the alignment path, with the maximum value of $M_1 + M_2$. p denotes the optimal alignment path. Using 14, we can obtain a new template y of length N_p:

$$y(i) = \frac{(x_1(R(0,i)) + x_2(R(0,i)))}{2} \tag{14}$$

where $i = 1, 2, \ldots, N_p$ and $y(i)$ represents the ith point of the newly obtained averaged template. x_1 and x_2 are the two feature sequences to be averaged. R is the record matrix obtained from Eq. 13. After performing arithmetic averaging on the two feature sequences modified by the best path, and further averaging with other training templates based on Eqs. 13 and 14, the resulting averaged template is more reliable than the casual training template.

4 Experiment Results

4.1 Experiment Setup

There are two speakers, A and B. The accidental template was generated using the speech of speaker A, while the robust template was created by combining the speech of both speakers. Each speaker had four different speech commands. The raw data generated from the experiment is presented in Table 1.

In Table 1, each data represents the difference value between two speech samples, denoted as δ in Eq. 11. A smaller difference value indicates greater similarity. The cells highlighted in bold in the table indicate erroneous recognition results, such as

Table 2 Recognition rates of different testers for different templates

Tester	Accidental template (%)	Robust template (%)
A	81.3	87.5
B	68.8	81.3

the fifth row where the recognized command is Backward but the actual meaning of the tested speech is Forward.

4.2 Simulation Result

The final recognition rates are presented as Table 2.

It can be observed that using the speech of Tester A alone as the random template yields poor recognition performance when applied to the test set of Tester B. However, by considering the robust template that incorporates training data from all participants, not only is the recognition rate improved for Tester B, but it also enhances the recognition rate for Tester A. This is because the training templates take into account the speech commands of Tester A to a greater extent.

5 Conclusion

This paper focuses on the analysis of speech signals using the MFCC feature extraction and DTW template matching processes. By enhancing the features and refining the templates, significant improvements have been achieved in accurately recognizing speech commands. Experimental results indicate a remarkable increase in recognition accuracy, with a notable improvement of 20% compared to previous methods. These findings highlight the effectiveness and reliability of the proposed approach in achieving robust and accurate speech command recognition.

Acknowledgements This paper is funded by NSF of China (61971162) and National Key R&D Program of China (2022YFC3801100).

References

1. Palaz D, Magimai-Doss M, Collobert R (2019) End-to-end acoustic modeling using convolutional neural networks for HMM-based automatic speech recognition. Speech Commun 37(108):15–32

2. Chowdhury A, Ross A (2020) Fusing MFCC and LPC features using 1D triplet CNN for speaker recognition in severely degraded audio signals. IEEE Trans Inf Forensics Secur 15(15):1616–1629
3. Hermansky H (1990) Perceptual linear predictive(PLP) analysis of speech. J Acoust Soc Am 28(4):357–366
4. Yu D (2013) Robust feature extraction based on auditory mechanism and its application in speaker recognition (doctoral dissertation). Harbin Institute of Technology, pp 5–6
5. Davis S, Mermelstein P (1980) Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Trans Acoust Speech Sig Process 28(4):357–366
6. Deng L (2004) Switching dynamic system models for speech articulation and acoustics. Math Found Speech Lang Process 115–133
7. Deng M, Meng T, Cao J et al (2020) Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. Neural
8. Shan W, Yang M, Xu J et al (2020) 14.1 A 510nW 0.41V low-memory low computation keyword-spotting chip using serial FFT-based MFCC and binarized depthwise separable convolutional neural network in 28nm CMOS. In: 2020 IEEE international solid-state circuits conference, San Francisco, CA, USA: IEEE, pp 230–232; Networks 33(130):22–32

Optimization Method of Multipath OLSR Protocol in Flying Ad Hoc Network

Zizhi Wo, Lin Ma, and Shouming Wei

Abstract Data transmission in flying Ad Hoc network (FANET) depends greatly on routing protocol. Multi-Path optimized link state routing (MPOLSR) is a popular one. However, due to frequent topology changes in the FANET, there are some disadvantages in applying the original protocol. Aiming at the low packet delivery rate, large delay and low throughput caused by MPOLSR protocol in FANET, a node selection and routing update algorithm based on link transmission quality is proposed in this paper. The algorithm combines link quality and node connectivity to select a MultiPoint Relay node set, and distinguishes link weights based on link transmission quality. At the same time, we replace the original protocol algorithm with the shortest path faster algorithm. Simulation results show that the proposed protocol improves the packet delivery rate and throughput performance, and reduces the end-to-end delay.

Keywords FANET · MPOLSR · MPR · Link transmission quality

1 Introduction

With the continuous development of technology, the application fields of Unmanned Aerial Vehicle (UAV) are becoming more and more extensive. The network formed by information exchange between multiple UAVs is called Flying Ad Hoc Network (FANET), which is an extended application of Mobile Ad-hoc Network (MANET) [1]. It has more and more applications in large-scale cargo transportation, electric power transmission, and target tracking in complex and unknown environments [2–5]. Node information interaction in FANET requires appropriate routing protocol, and it is necessary to design an applicable routing protocol according to the specific environmental characteristics of the FANET [6].

The topology of FANET is complex, and the direct application of the routing protocol in the MANET will cause huge interference to the communication between

Z. Wo · L. Ma (✉) · S. Wei
School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, China
e-mail: malin@hit.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_62

UAVs, which needs to be improved [6]. Many scholars have carried out in-depth analysis and research on this. Reference [7] introduced the concept of hazy vision algorithm and applied it to OLSR. Simulation results show that the algorithm reduces the routing overhead of the network, but the complexity is high. Reference [8] proposed a trust-based optimal path algorithm to replace the original Dijkstra algorithm. The simulation results show that the anti-jamming performance of the flight ad hoc network has been greatly improved. Reference [9] proposed the MPR-deep algorithm to optimize the position and forwarding probability of FANET relay drones. The simulation results show that the algorithm has strong generalization ability, but it takes a long time to train and consumes a lot of resources. Reference [10] applied the firefly algorithm to multipath routing calculations. The results show that the larger the population size, the better the performance of the algorithm, but the execution time increases significantly.

As mentioned above, for the improvement of FANET routing protocol, we need to clear the specific situation at the beginning. Simultaneously, it is necessary to minimize the consumption of computational resources while improving metrics such as delivery rate and throughput. Based on the aforementioned considerations, this paper proposes corresponding optimizations for routing protocols, aiming to enhance key network performance metrics and adapt them better to UAV swarm scenarios characterized by frequent network topology changes. The remainder of this paper is organized as follows. Section 2 presents the system framework. In Sect. 3, specific implementation schemes for the optimization algorithm are discussed. Section 4 conducts the performance simulation. And finally, the conclusion is drawn in Sect. 5.

2 System Model

2.1 Main Research Content

The proposed process flow diagram for optimizing the MPOLSR protocol in FANET is illustrated in Fig. 1. This approach incorporates both node reachability and link transmission quality for the selection of MPR nodes. Additionally, it calculates the shortest path based on the expected link transmission as the edge weight for nodes and replaces the original protocol shortest path algorithm.

The selection of the MPR set in the original protocol solely relies on the node connectivity. However, in the context of frequent mobility in FANET, the stability of links between nodes decreases. Nodes with high connectivity may quickly move out of the communication range of other nodes, making it inappropriate to use a single criterion of node connectivity for MPR selection. To address this issue, this paper proposes a network state-aware MPR set selection algorithm that considers the link transmission quality between two nodes. This approach aims to avoid selecting nodes with high connectivity, thereby enhancing network communication performance.

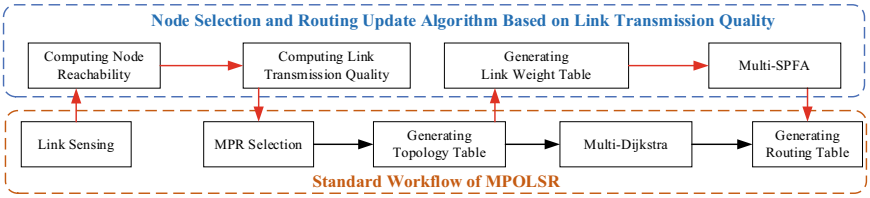


Fig. 1 Framework diagram of the research content in this paper

2.2 Mobility Model

Currently, the mobility models used in FANET can be primarily categorized into two types: physical mobility models and group mobility models. When faced with a scenario requiring a large number of UAV nodes in complex and unknown environments, employing physical mobility models may impose significant network communication burdens and may not accurately represent the actual movement trajectories of multiple UAVs during task execution, thus having limitations. The most widely applied mobility model in aerial ad hoc networks is the Reference Point Group Mobility (RPGM) model, depicted in Fig. 2. In this type of mobility model, mobile nodes are divided into different groups based on certain rules, and their final displacement vectors are the sum of their individual vectors and group displacement vectors.

The RPGM mobility model not only preserves the random motion characteristics of individual nodes but also reflects the overall movement characteristics of groups, which closely resembles real-world scenarios involving intelligent motion of UAV swarms. There is significant data interaction within each group, and data sharing is also required among different groups, indicating the presence of information exchange. Therefore, this model exhibits more realistic and FANET-relevant node movement behaviors, making it the chosen mobility model in this paper.

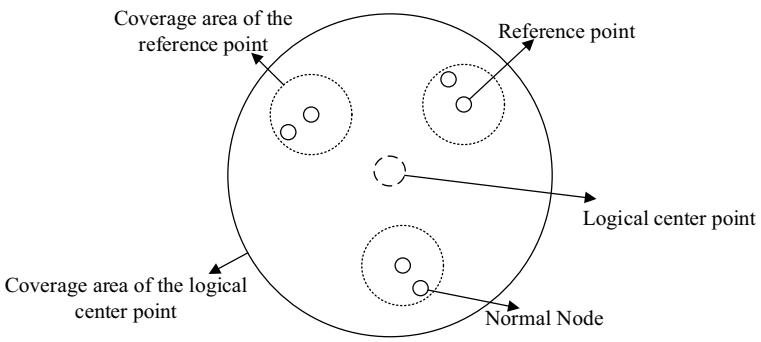


Fig. 2 RPGM mobility model

3 Optimization Algorithms

3.1 Network State-Aware MPR Set Selection Algorithm

In ad hoc networks, the link transmission quality between nodes can be characterized using the ETX [11]. ETX is calculated based on the ratio of received HELLO messages from neighboring nodes within a certain period, representing the forward Link Quality (LQ) and the reverse link quality, also known as Neighbor Link Quality (NLQ). A smaller ETX value indicates higher link quality in network communication. However, it is not possible to calculate LQ, NLQ, and consequently, ETX. Therefore, this paper introduces certain modifications to the parameters of the ETX, exchanging the numerators of LQ and NLQ, resulting in new parameters LQ_I and NLQ_I. This modification does not affect the calculation of ETX. The formula for calculating the link quality ETX is as follows:

$$LQ\_I = \frac{R_{ij}}{S_{ij}}, \quad NLQ\_I = \frac{R_{ji}}{S_{ji}}, \quad ETX' = \frac{1}{LQ\_I \times NLQ\_I} \tag{1}$$

where R_{ji} represents the number of HELLO messages received by node j from node i, S_{ij} represents the number of HELLO messages sent by node i to node j, and R_{ji} and S_{ji} have the similar definition but the opposite direction from j to i. Therefore, by taking into account both the node connectivity and the link transmission quality between nodes, this paper utilizes a new weight metric to calculate MPR nodes, replacing the standard MPR selection algorithm. The node weight metric W is calculated as follows:

$$W = \alpha \frac{1}{ETX'} + \beta D \tag{2}$$

where D represents the connectivity between 1-hop and 2-hop nodes. This metric takes into account both link quality and node connectivity. α and β are weight metrics, satisfying the condition $\alpha + \beta = 1$. The optimized MPR selection is shown in Fig. 3.

After the optimization and improvement of the aforementioned algorithm, the MPOLSR protocol will be better suited to adapt to the scenarios in FANET where there are frequent node movements leading to drastic changes in the network topology. It will be able to more effectively select suitable MPR nodes, reduce end-to-end latency and improve packet delivery rates.

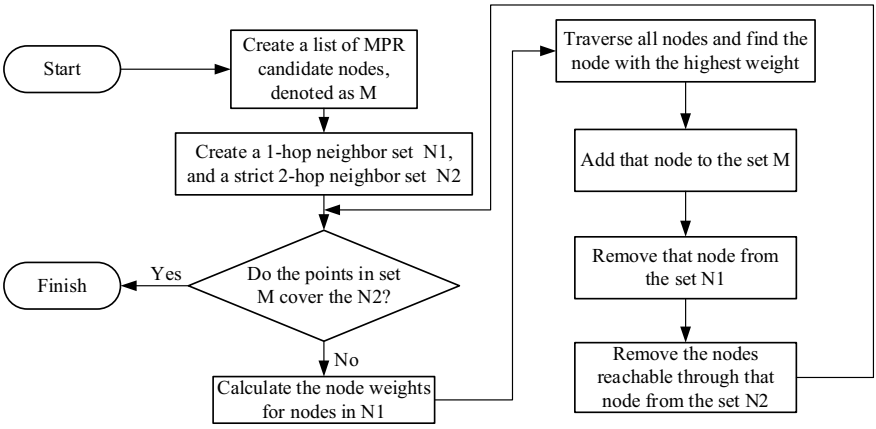


Fig. 3 MPR selection algorithm diagram

3.2 Shortest Path Optimization Based on Link Transmission Expectation

Regarding the change in the calculation method of edge weights, the traditional MPOLSR routing protocol does not differentiate between link weights. However, in FANET, UAVs move more frequently, leading to decreased link stability between nodes. For links with higher transmission expectations, lower weights need to be assigned. Conversely, higher weights should be assigned to links with lower transmission expectations. To implement this mechanism, the source node needs to obtain the ETX values of other edges. This information can be acquired by adding relevant fields in the Topology Control (TC) packets, and additional fields need to be added in the local tables of nodes to store this information. Taking into account both the optimization of algorithm complexity and the edge weight of the shortest path, this paper proposes a multipath routing algorithm that uses link quality as the weight as shown in Fig. 4.

The core idea is inspired by the SPFA algorithm. This approach reduces computational complexity and distinguishes the weights of different links. Nodes perceive the network status by exchanging HELLO messages and TC messages with each other. Customized link weight metrics are used as the weights of each edge, which differentiates it from the traditional MPOLSR protocol that uses hop count as the edge weight in the Dijkstra algorithm. After executing this algorithm, the computational complexity is lower compared to the original MPOLSR protocol's Dijkstra algorithm, resulting in shorter execution time, which can effectively avoid the phenomenon of decreased packet delivery rates and increased transmission delays caused by frequent changes in the topology of the FANET.

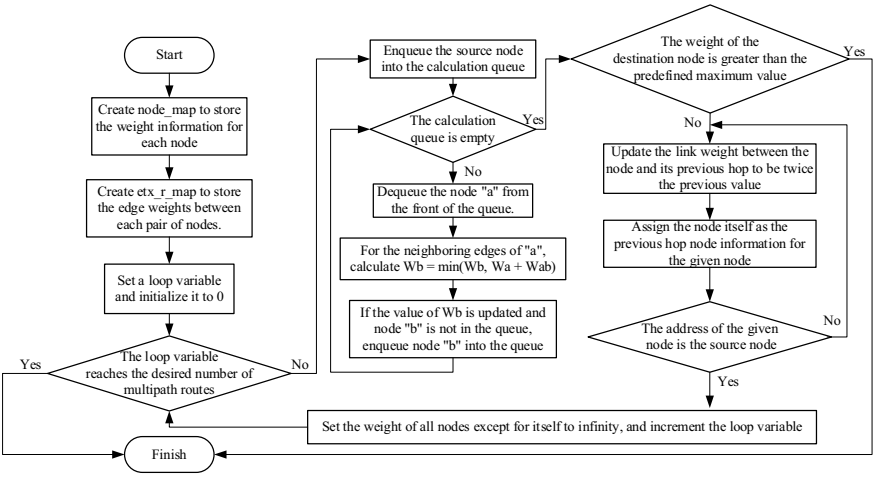


Fig. 4 Flow chart of the shortest path optimization algorithm

4 Protocol Performance Simulation

4.1 Simulation Metrics and Scenario Settings

The simulation platform used is NS2. The node mobility scenario is set as RPGM. Three metrics, namely throughput, packet delivery ratio, and average end-to-end delay, are measured under different node speeds to evaluate the performance of the optimized algorithm in comparison with the original MPOLSR protocol.

Throughput refers to the amount of data successfully transmitted within a unit of time and is commonly used to measure the communication capacity of a network or a link. The calculation is as follows, N_{th} means the total data received, T_{th} means the total time taken for receiving messages:

$$th = \frac{N_{th}}{T_{th}} \tag{3}$$

Packet delivery rate represents the proportion of successfully received data packets at the destination node compared to the total number of data packets sent by the source node. The calculation is as follows, *recv* means the number of packets received, *send* means the number of packets sent by the source node:

$$pdr = \frac{\Sigma recv}{\Sigma send} \tag{4}$$

Table 1 Simulation parameter settings

Parameter	Meaning
Simulation range	2500 m × 2500 m
Mobility model	RPGM
Propagation model	Two ray ground
MAC layer protocol	802.11
Routing protocol	MPOLSR, LQWD-MPOLSR, LQWS-MPOLSR
Group center coverage radius	200 m
Average number of groups	10
Nodes number	50, 100
Movement speed	5–30 m/s
Communication distance	300 m
Maximum number of connections	20, 40
Packet interval	0.1 s
Packet size	512 Bytes

The average end-to-end delay is the average time consumed by all destination nodes to successfully receive data packets. The calculation is as follows, $delay_i$ means the delay to destination node i, N means the total number of data packets:

$$ave\_delay = \frac{1}{N} \sum_{i=1}^{N} delay_i \tag{5}$$

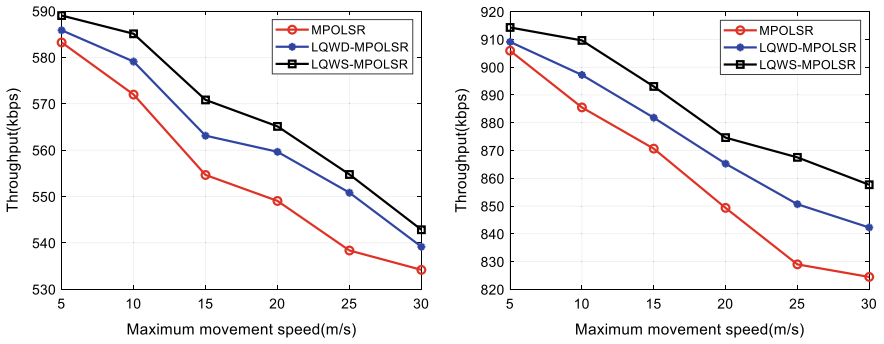
The simulation scenario parameters and their specific settings are presented in Table 1. In this paper, the MPOLSR protocol, which has been modified with the aforementioned optimization algorithm, is named the LQWS-MPOLSR protocol. Compared to the original MPOLSR protocol, the LQWD-MPOLSR protocol optimizes the MPR selection mechanism and incorporates link transmission quality as edge weights but retains the original protocol's shortest path algorithm. The simulation is conducted with 50 nodes and a set of 20 pairs of communication connections, as well as with 100 nodes and a set of 40 pairs of communication connections.

4.2 Simulation Results and Performance Analysis

As shown in Fig. 5, it represents the throughput curves for different node speeds. It can be observed that the network throughput decreases as the node speed increases. The original MPOLSR protocol exhibits the lowest throughput, indicating that the improved MPR selection algorithm and shortest path algorithm are more suitable for scenarios with faster node speeds. The LQWS-MPOLSR protocol demonstrates

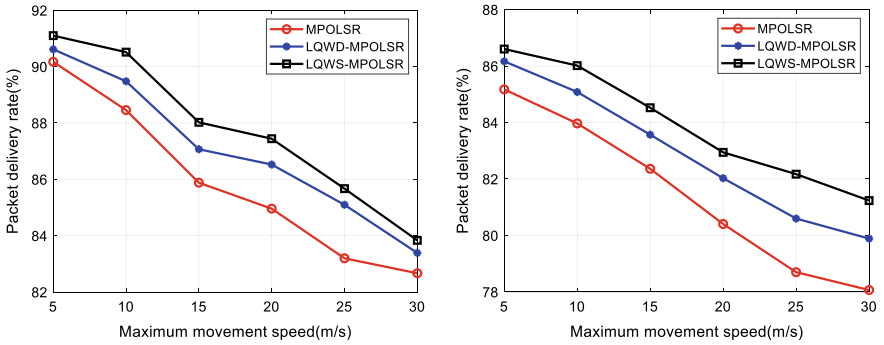
a higher throughput than the LQWD-MPOLSR protocol, particularly noticeable in scenarios with a larger number of communication connections, indicating the transmission of a greater amount of information flow and further emphasizing the lower packet loss rate of the SPFA algorithm.

As shown in Fig. 6, it illustrates the packet delivery ratio curves for different node speeds and different numbers of nodes. It can be observed that the packet delivery ratio decreases as the node speed increases. Within the same scenario, the packet delivery ratio for 100 nodes is lower than that for 50 nodes. When the number of nodes is higher, the overall network becomes denser, and each node has to handle a larger number of message packets, leading to message processing queue congestion and packet loss, thus reducing the packet delivery ratio. The LQWS-MPOLSR protocol, which incorporates the SPFA algorithm, achieves the best packet delivery ratio, while the original MPOLSR protocol performs the worst. This indicates that the algorithm optimization improves the packet delivery ratio.



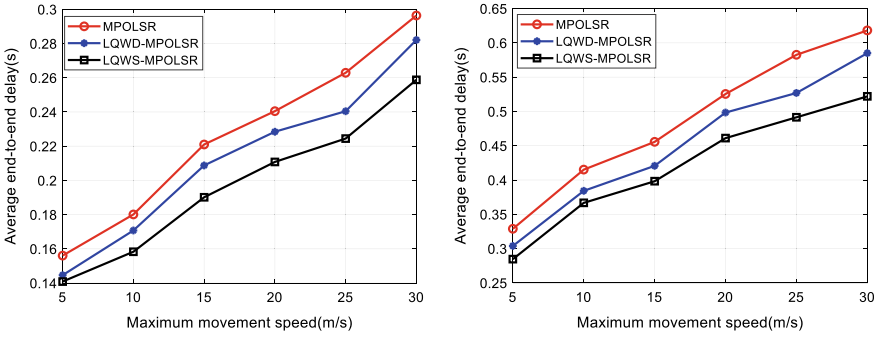
(a) 50 nodes curve (b) 100 nodes curve

Fig. 5 Variation curve of throughput with speed



(a) 50 nodes curve (b) 100 nodes curve

Fig. 6 Variation curve of packet delivery rate with speed



(a) 50 nodes curve (b) 100 nodes curve

Fig. 7 Variation curve of average end-to-end delay with speed

As shown in Fig. 7, it illustrates the average end-to-end delay curves for different node speeds and different numbers of nodes. It can be observed that the average end-to-end delay increases as the node speed increases. Within the same scenario, the average end-to-end delay for 100 nodes is higher than that for 50 nodes. This is partly due to the complex network topology, which requires more time for shortest path calculations. Additionally, a larger number of communication connection pairs and intermediate forwarding nodes increase the processing time for messages, leading to longer end-to-end delays. The original protocol exhibits the highest delay, while the LQWS-MPOLSR protocol achieves lower delay compared to the LQWD-MPOLSR protocol.

5 Conclusion

This paper proposes a node selection and routing update algorithm based on link transmission quality. It introduces the mechanism of transmission expectation and modifies the formats of HELLO and TC messages to calculate node reachability and link transmission quality for selecting the MPR node set. Then, using the link weights, it employs the lower-complexity multi-path SPFA algorithm to compute multiple shortest paths. Based on the simulations and comparisons, we can know that the optimized protocol improves the performance in terms of packet delivery ratio, average end-to-end delay, and network throughput. Furthermore, the performance of the SPFA algorithm is superior to that of the Dijkstra algorithm, confirming the effectiveness of the proposed algorithm in this paper.

Acknowledgements This research is funded by National Key R&D Program of China (2022YFC3801100) and National Natural Science Foundation of China (61971162).

References

1. Wu Q, Zhang M et al (2022) Routing protocol for heterogeneous FANETs with mobility prediction. China Commun 19(1):186–201
2. Perreault M, Behdinan K (2021) Delivery drone driving cycle. IEEE Trans Veh Technol 20(7):1146–1156
3. Cui M (2018) Application of low-altitude photogrammetry technology for unmanned aerial vehicles in geological disaster monitoring in China. World Nonferrous Met 15:194–195
4. Wang W, Cao W et al (2017) Application of quadcopter unmanned aerial vehicle in environmental impact assessment of power transmission and transformation projects. Environ Impact Assess 39(02):48–51
5. Zhu D, Liu H et al (2023) Research on cooperative target tracking of multiple unmanned aerial vehicles in a designated area. Comput Meas Control 31(3):201–207
6. Abdulhae OT, Mandeep JS et al (2022) Cluster-based routing protocols for flying ad hoc networks (FANETs). IEEE Access 10:32981–33004
7. Zhang M (2018) Research and optimization of link-state routing protocol for D2D self-organizing networks. Master's Thesis in Engineering at Beijing University of Posts and Telecommunications
8. Shcherba EV, Litvinov GA et al (2019) A novel reputation model for trusted path selection in the OLSR routing protocol. In: International conference on information science and communications technologies (ICISCT), Tashkent, Uzbekistan, pp 1–5
9. Wang Q, Liu J et al (2021) MPRdeep: multi-objective joint optimal node positioning and resource allocation for FANETs with deep reinforcement learning. In: IEEE 46th conference on local computer networks (LCN), Edmonton, Canada, pp 315–318
10. Perepelkin D, Nguyen T (2022) Research of multipath routing processes in software defined networks based on firefly algorithm. In: International Russian automation conference (RusAutoCon), Sochi, Russian Federation, pp 476–483
11. Mohapatra S, Tripathy T (2016) MM-OLSR: multi metric based optimized link state routing protocol for wireless ad-hoc network. In: Signal processing, communication, power and embedded system (SCOPES). IEEE, pp 153–158

UAV Trajectory Planning Based on Improved Quantum Particle Swarm Optimization

Rangang Zhu, Jing Wang, Jian Wang, and Lin Ma

Abstract Trajectory planning is important for unmanned aerial vehicle (UAV) to smoothly execute tasks, and it is also a focus of attention for scholars throughout the world. UAV seeks an optimal or suboptimal path from the starting point to the ending point, with the goal of meeting the predetermined target requirement, while satisfying its own mobility and external interference conditions. Although a lot of researches improved the trajectory planning method, its convergence is still a hot topic, since the algorithm complexity is high, and it is easy to fall into local optima. Hence this paper proposes an improved quantum particle swarm optimization algorithm for UAV trajectory planning to address the problem of premature convergence and local optima in trajectory planning. The simulation results show the feasibility of our improved quantum particle swarm optimization algorithm, which can reduce the number of iterations and improve the planning success rate.

Keywords UAV · Trajectory planning · Quantum particle swarm

1 Introduction

Unmanned Aerial Vehicle (UAV) is playing an extremely important role in high-risk and high demand mission scenario. It has many advantages such as low cost, high efficiency, low risk of injury or death and good maneuverability [1]. UAV is very suitable for some dangerous and repetitive tasks. Usually, it is operated by radio remote control equipment and independent program control equipment [2]. It involves a wide range of technology fields, including communication [3], intelligent control [4], sensor network [5], aviation power propulsion [6] and information processing [7]. With the gradual maturity of UAV research and development technology, the manufacturing

R. Zhu · J. Wang
College of Electronic Engineering, National University of Defense Technology, Changsha, China

J. Wang · L. Ma (✉)
School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, China
e-mail: malin@hit.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_63

costs of UAV continues to decline, and it have been widely used in many fields. At present, UAV can serve as a flight platform to conduct many important works, such as plant protection, power inspection, disaster rescue, aerial photography and data collection [8]. Therefore, as an important component of unmanned aerial vehicle planning system, UAV trajectory planning has attracted more and more attention.

UAV trajectory planning refers to find the optimal or suboptimal trajectory of a UAV from its starting point to its target point under certain constraints, which can meet both the maneuverability and the mission objectives. Trajectory planning is of great significance for improving the operational efficiency, implementing reconnaissance, military strikes, etc. [9]. With the continuous progress of air defense technology, many countries have conducted extensive researches on the trajectory planning under different operational requirements. The essence of trajectory planning is the process to find the optimal solution in various possible trajectory. There are several common research methods, such as dynamic programming [10], mixed integer linear programming [11], Voronoi diagram [12], programming methods based on graph theory [13], A* algorithm [14] and Dijkstra algorithm [15].

In this paper, we aim to solve the trajectory planning problem based on improved quantum particle swarm optimization. Our proposed method can not only improve the accuracy and real-time performance of UAV trajectory planning, but also improve the global search ability and smoothing the trajectory while meeting the constraints of UAV maneuverability, terrain environment and many other threat constraints. The remainder of this paper is organized as follows. Section 2 will introduce the system model. The proposed method will be demonstrated in Sect. 3. In Sect. 4, implementation and performance analysis will be provided. Conclusion will be drawn finally in Sect. 5.

2 System Model

In this paper, we propose an improved quantum particle swarm optimization (MQPSO) algorithm to solve the problem that the traditional UAV trajectory planning algorithm has a weak ability to explore global optimization and is prone to fall into local optimization. In quantum particle swarm optimization, a population division strategy is implemented to expand the search space, which can avoid falling into local optima, and increase the diversity of the population. Subsequently, an optimization strategy is adopted, which effectively combined the quantum particle swarm optimization algorithm with evolutionary programming algorithm to resulting in a better optimization. The schematic diagram of the overall research route of this paper is shown in Fig. 1.

The trajectory planning problem of UAVs in a 3D environment is relatively complex and requires processing a large amount of different information. First, we collect terrain information and threat information in the environment, and model them mathematically. Then, according to the actual requirements of the flight task,

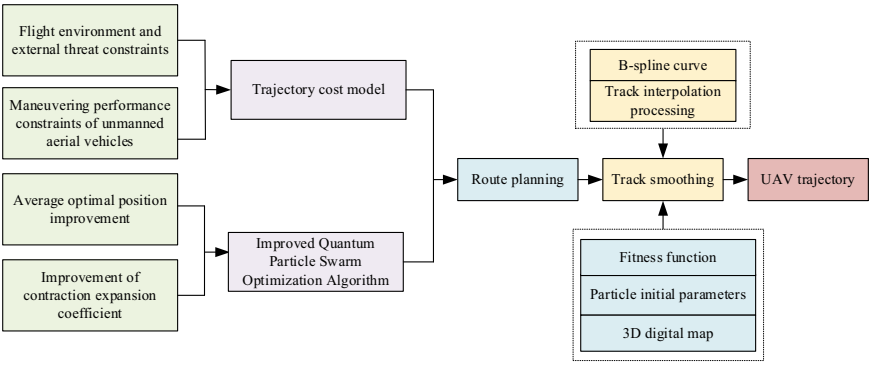


Fig. 1 Schematic diagram of the overall research route of this paper

we choose an optimization algorithm based on quantum particle swarm optimization. Finally, the initial trajectory is smoothed and the optimal trajectory that met the requirements is generated.

3 Proposed Method

3.1 Modeling of UAV Trajectory Planning

In the trajectory planning, there are many ways to construct the 3D map. This paper adopts a function simulation method which is widely used in many literature to simulate the real terrain, which is expressed as:

$$z_1(x, y) = \sin(y + a) + b \sin x + c \cos\left(d\sqrt{x^2 + y^2}\right) + e \cos y + f \sin\left(g\sqrt{x^2 + y^2}\right) \tag{1}$$

where x and y refer to the coordinates of the points on the horizontal line, and z_1 is the height. The parameters from a to g are constant and used to control the slope of the reference terrain on the electronic map. On this basis, a reference terrain model suitable for UAVs is formed.

The main obstacle in the space when the UAV executes the trajectory planning task is the mountain. exponential function is proposed to model the mountain:

$$z_2(x, y) = \sum_{i=1}^{n} h_i \exp\left[-\left(\frac{x - x_i}{x_{si}}\right)^2 - \left(\frac{y - y_i}{y_{si}}\right)^2\right] \tag{2}$$

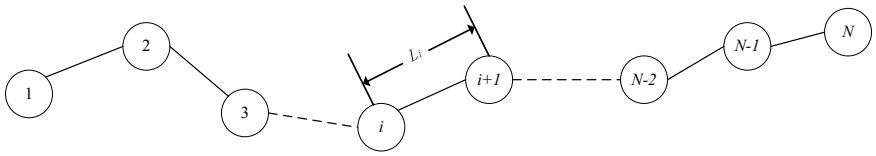


Fig. 2 Schematic diagram of UAV trajectory

where i is the ith peak, n is the total number of peaks, (x_i, y_i) is the coordinates of the center position of the peak, and z_2 is the terrain height of the corresponding coordinates. and h_i is the peak slope parameters, and are used to represent the peak slope height.

The threat distribution model in this paper is expressed as:

$$z_3(x, y) = \frac{a_i}{\left(b_i + c_i(x - x_i)^2 + d_i(y - y_i)^2\right)^n} \tag{3}$$

where the parameters from a_i to d_i are threat coefficients.

The maximum total length of the flight trajectory is the sum of the lengths of each flight trajectory segment, $L_{\max}$ represented by the maximum range. The UAV flight trajectory is shown in Fig. 2.

L_i represents the length of the i flight trajectory in the first segment, and the total length L of the trajectory needs to meet the conditions as shown:

$$\sum_{i=1}^{N} L_i \le L_{\max}, \quad i = 1, 2, \ldots, N \tag{4}$$

Assuming UAV has the maximum yaw angle θ and maximum pitch angle φ, these two constraints can be expressed as:

$$\cos\theta \le \frac{a_i^T a_{i+1}}{\|a_i\| * \|a_{i+1}\|}, \quad i = 2, \ldots, N-1 \tag{5}$$

$$\tan\varphi \ge \frac{|z_i - z_{i-1}|}{|a_i|}, \quad i = 2, \ldots, N-1 \tag{6}$$

The maximum yaw angle θ constraint is shown in Fig. 3.

Assuming the angle of attack as γ, the attack calculation formula between two trajectory nodes is:

$$\frac{|Z_i - Z_{i-1}|}{\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}} \le \tan\gamma, \quad (i = 2, \ldots, N) \tag{7}$$

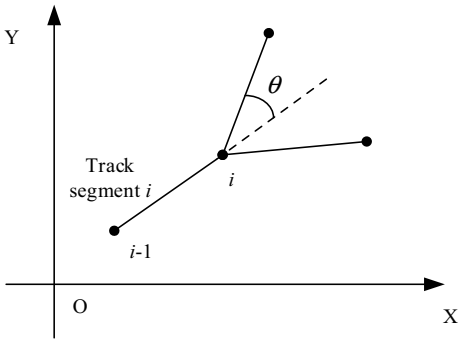


Fig. 3 The maximum yaw angle constraint

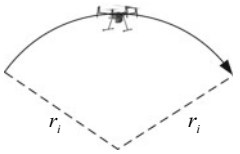


Fig. 4 Flight turning radius

To verify the effectiveness of UAV trajectory planning, it is necessary to meet the specific requirements for turning radius $r_i \geq r_{\min}$, as shown in Fig. 4.

The B-spline function in three-dimensional space is:

$$p(\boldsymbol{u}) = \sum_{i=0}^{n} d_i N_{i,k}(\boldsymbol{u}) \tag{8}$$

where $N_{i,k}$ is defined as:

$$\begin{aligned} N_{i,0}(u) &= \begin{cases} 1, & u_i \leq u \leq u_{i+1} \\ 0, & \text{other} \end{cases} \\ N_{i,k}(u) &= \frac{u - u_i}{u_{i+k} - u_i} N_{i,k-1}(u) + \frac{u_{i+k+1} - u}{u_{i+k+1} - u_{i+1}} N_{i,k-1}(u) \end{aligned} \tag{9}$$

3.2 Improved Quantum Particle Swarm Optimization

Assuming that the fitness of particles i is f_i, the average fitness of the population is $\overline{f}$, and the variance of fitness is s^2, then its expression is:

$$s^2 = \sum_{i=1}^{M} \left(f_i - \overline{f}\right)^2 \tag{10}$$

where M represents the number of particles. The fitness variance can reflect the aggregation degree of particles of the whole population. If s^2 is smaller, it indicates all particles gather more tightly. On the contrary, all particle clusters become less compact with the algorithm iterating, and the degree of particle aggregation in the later stage of the population will become tighter and smaller. When s^2 reaches a threshold σ, the algorithm enters the later stage of iterative search, resulting in precocity.

In order to ensure that the algorithm is not limited by local optima, avoid falling into the above situation, and continue to expand the search area, this paper propose to add an evolution factor when obtaining the average optimal position:

$$\lambda_t = \mu_1[K_t(0,1) + \mu_2 N_t(0,1)] \tag{11}$$

where $K_t(0, 1)$ and $N_t(0, 1)$ are the interval random number generated by the standard Cauchy distribution and standard Gaussian distribution respectively, μ_1 and μ_2 is the interference coefficient, which are defined as:

$$\begin{cases} \mu_1 = \mu_{1\min} + (\mu_{1\max} - \mu_{1\min})\frac{t}{t_{\max}} \\ \mu_2 = \mu_{2\max} - (\mu_{2\max} - \mu_{2\min})\frac{t}{t_{\max}} \end{cases} \tag{12}$$

where $\mu_{1\min}$ and $\mu_{1\max}$ is the minimum and maximum of μ_1, $\mu_{2\min}$ and $\mu_{2\max}$ is the minimum and maximum of μ_2. t and $t_{\max}$ are current and the largest iteration respectively.

Based on the idea of dynamic parameter changes in particle swarm optimization and in response to practical problems, a nonlinear adaptive method is proposed to control this parameter, and a new dynamic reduction method is proposed for the contraction expansion coefficient to improve the optimization effect of model parameters. The specific update method is as follows:

$$\alpha_2 = \alpha_{\max} - (\alpha_{\max} - \alpha_{\min}) \tan\left(\frac{t}{t_{\max}} \cdot \frac{\pi}{4}\right) \tag{13}$$

where $\alpha_{\max}$ and $\alpha_{\min}$ represent the initial and the final values of the contraction expansion coefficient respectively.

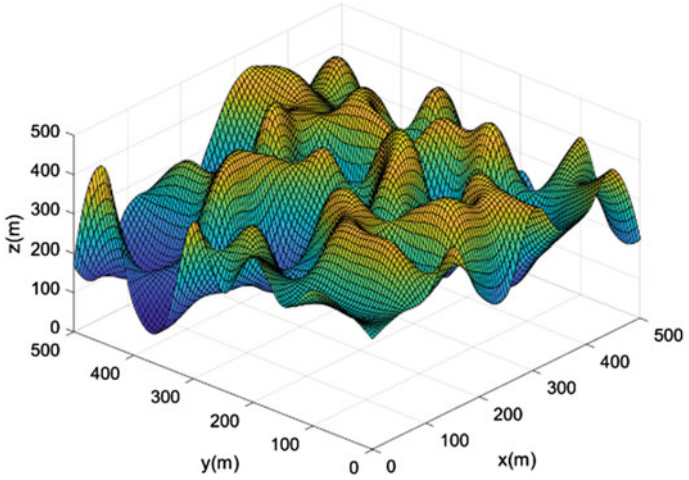


Fig. 5 Simulated three-dimensional mountainous terrain map

4 Implementation and Performance Analysis

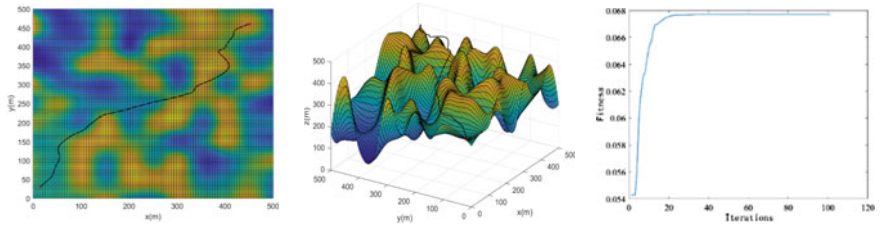
4.1 Experiment Setup

In order to validate the UAV trajectory planning performance based on our proposed quantum particle swarm optimization in this paper, we establish a three-dimensional terrain based on a two-dimensional plane and simulate a three-dimensional mountainous area for simulation. There are a total of 15 peaks, with randomly generated peak positions, heights, and ranges. Here is a random terrain map shown in Fig. 5.

4.2 Simulation Results

Considering the single UAV trajectory planning, we set the starting point coordinate as the coordinate origin and the height as the ground height. The endpoint height is the ground height. Set the number of trajectory points $N = 10$, particle swarm population size is 100, iteration number is 100, and the inertia weight $\omega = 0.9$. During the simulation process, the reciprocal of the trajectory cost function is taken as the fitness value, and the lower limit of the fitness value is set to 0.05. If it is less than this lower limit, the trajectory planning is considered unsuccessful. The trajectory planning results of a single UAV obtained by improving the quantum particle swarm optimization algorithm are shown in Fig. 6.

In Fig. 6, it can be seen from the two-dimensional trajectory results in Fig. 6a that a smooth trajectory is obtained from the starting point to the endpoint, with the



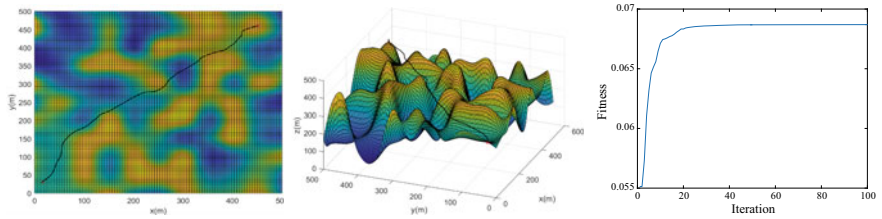
a) 2D trajectory b) 3D trajectory c) Fitness

Fig. 6 Single UAV trajectory planning (particle swarm population size is 100)

trajectory points all in the flying area, avoiding various peaks in the flying area and meeting the obstacle avoidance requirements. The planned trajectory is relatively smooth, with no obvious turning points. Figure 6b are three-dimensional trajectory from different angles. By switching angles, it can be seen that the planned flight trajectory avoids various peaks in the mountainous terrain, and the flight trajectory is mostly close to the ground between valleys, reducing the probability of colliding with mountains. The rise and fall curves are both smoothed. Figure 6c shows the fitness curve. It can be seen that after 40 iterations, the algorithm basically converges, and the fitness value finally stabilizes around 0.06, which is higher than the lower limit of the fitness value. It can be considered that the planning is successful.

Set the population size $N = 200$ and ensure that the remaining parameters remain unchanged. Perform trajectory planning simulation on a single UAV again, and the results are shown in Fig. 7.

In Fig. 7, it can be seen that the trajectory planning can obtain a smooth trajectory from the starting point to the endpoint, avoiding obstacles in mountainous areas, and the curve is smooth at all points without excessive turning angles. Comparing Figs. 6a and 7a, it can be seen that the larger the population size, the higher the accuracy of the search, and the larger the search range, reducing unnecessary corners in the trajectory, allowing the UAV to fly along the shortest trajectory. From Fig. 7c, it can be seen that the algorithm basically converges after 30 iterations. Comparing Figs. 6a and 7a, it can be found that the population size is negatively correlated with the number of



a) 2D trajectory b) 3D trajectory c) Fitness

Fig. 7 Single UAV trajectory planning results (particle swarm population size is 200)

iterations. The larger the population size, the fewer iterations the algorithm has. That is, as the population size increases, the convergence speed of the algorithm increases.

In the simulation experiment, the improved quantum particle swarm optimization algorithm and the improved quantum particle swarm optimization algorithm were used in the form of randomly established terrain maps each time. Under the same basic parameter settings, the trajectory planning simulation of a single UAV is carried out with a fitness convergence accuracy of 0.05. 1000 simulations were conducted using both algorithms, and the success rates obtained are shown in Table 1. In Table 1, although each planned map has randomness, the number of peaks is the same and the overall complexity is similar. Through comparison, it can be seen that after various improvements to the quantum particle swarm algorithm, the improved quantum particle swarm algorithm has a higher success rate in trajectory planning under the same conditions.

In the simulation, the parameter settings in Table 1 are fixed, and 1000 trajectory planning are conducted based on the traditional particle swarm algorithm and the improved particle swarm algorithm, respectively. Due to the randomness of the terrain, the convergence speed of the iteration fluctuates within a certain range, and the convergence speed are also related to the population size. When the population size is 100, the average convergence speed before improvement is about 80 iterations, and the probability of planning failure is high. After improvement, it can converge stably within 40 iterations. When the population size is 200, the average convergence speed before improvement is about 20 to 40 iterations, and the improved algorithm can stably converge within 20 iterations. In addition, the calculation time of trajectory planning depends on the complexity of the map. The time of 100 trajectories is statistically calculated, and it can be found that the time after improvement is slightly lower than that before improvement, and the average fitness remains higher, which is not affected by the size of the initial fitness. Therefore, the improved algorithm proposed in this paper is effective, greatly improving the calculation success rate and reducing the time loss caused by calculation failures. The convergence speed is accelerated, ensuring an improved success rate without significant computational time consumption.

5 Conclusion

This paper focuses on the trajectory planning problem of UAVs based on improved quantum particle swarm optimization. The quantum particle swarm optimization algorithm is utilized to improve the accuracy and real-time performance of UAV trajectory planning by addressing the issues of premature algorithm and limited search space. Simulation results show that our proposed method can reduce the number of algorithm iterations and improve the planning success rate of the algorithm.

Table 1 Comparison of single planning success rates

Population size	Quantum-behaved particle swarm optimization (%)	Improved quantum particle swarm optimization algorithm (%)
$N = 100$	79.7	89.6
$N = 200$	82.3	92.1
$N = 500$	88.9	98.4

Acknowledgements This research is funded by National Natural Science Foundation of China (61971162) and National Key R&D Program of China (2022YFC3801100).

References

1. Ving K, Gebreyohannes S, Karimoddini A (2019) An area-decomposition based approach for cooperative tasking and coordination of UAVs in a search and coverage mission. In: 2019 IEEE aerospace conference, pp 1–8
2. Li B, Fei Z, Zhang Y (2019) UAV communications for 5G and beyond: recent advances and future trends. IEEE Internet Things J 6(2):2241–2263
3. Jiang Y, Wu Y, Wang M (2022) Research on unmanned vehicle path planning based on improved bat algorithm. In: 2022 international symposium on distributed computing and applications for business engineering and science, pp 206–209
4. Huang C, Fei J (2018) UAV path planning based on particle swarm optimization with global best path competition. Int J Pattern Recognit Artif Intell 32(6):245–270
5. Bing X, Youwei Z, Xueyan Z, Xuekai S (2021) An improved artificial bee colony algorithm based on faster convergence. In: 2021 IEEE international conference on artificial intelligence and computer application, pp 776–779
6. Ma H, Shen S, Yu M et al (2018) Multi-population techniques in nature inspired optimization algorithms: a comprehensive survey. Swarm Evol Comput 44:365–387
7. Zhao Y, Zheng Z, Liu Y (2018) Survey on computational-intelligence-based UAV path planning. Knowl-Based Syst 158:54–64
8. Bansal JC, Gopal A, Nagar AK (2018) Stability analysis of artificial bee colony optimization algorithm. Swarm Evol Comput 41:9–19
9. Saadaoui H, El Bouanani F (2021) Communication and energy optimization of local PSO-assisted multi-UAVs for moving targets exploration. In: 2021 4th international conference on advanced communication technologies and networking (CommNet), pp 1–7
10. Xia X, Xing Y, Wei B et al (2018) A fitness-based multi-role particle swarm optimization. Swarm Evol Comput 44:349–364
11. Lin A, Sun W, Yu H et al (2018) Global genetic learning particle swarm optimization with diversity enhancement by ring topology. Swarm Evol Comput 44:571–583
12. Zhang D, Xian Y, Li J et al (2015) UAV path planning based on chaos ant colony algorithm. In: International conference on computer science & mechanical automation. IEEE, pp 163–178
13. Lee W, Kim DE (2018) Adaptive approach to regulate task distribution in swarm robotic systems. Swarm Evol Comput 44:1108–1118
14. Yang P, Tang K, Lozano JA et al (2015) Path planning for single unmanned aerial vehicle by separately evolving waypoints. IEEE Trans Robot 31(5):1130–1146
15. Yu Z, Sun F, Lu X (2021) Overview of research on 3D path planning methods for rotor UAV. In: 2021 international conference on electronics circuits and information engineering, pp 368–371

A Noval Image Sensing Bionic Pre-encoder

Jindong Li, Xiao Hu, Zishen Huang, and Mingjin Zhu

Abstract Image preprocessing is important for deep learning image recognition tasks, and this paper introduces a new method called a bionic precoder, inspired by the human retinal visual system, to improve deep learning image recognition models. The precoder simulates a non-uniform distribution of retinal photoreceptors and iterates with trainable parameters. The trained precoder performs retinal-like preprocessing to improve the quality of the input data. Significant improvements in the accuracy/recall curve and receiver operating characteristic curve were observed experimentally. This study presents a promising technique for incorporating precoders into deep learning, leading to more accurate and efficient image recognition models.

Keywords Retinal · Deep learning · Image convolution · Pre-processing

1 Introduction

In recent years, the field of image recognition has witnessed significant advancements due to the rapid development of parallel computing, which has facilitated the application of deep learning. Deep learning revolves around the fundamental concept of emulating the functionality of neurons in the human brain. This is achieved by constructing neural networks with multiple layers to extract and classify features from images. Unlike traditional recognition methods that rely on handcrafted features, multi-layer neural networks have the capability to learn from large volumes of training data. Consequently, these neural networks can automatically extract features from images and utilize them in classifiers for classification tasks [1, 2].

J. Li · X. Hu (✉) · Z. Huang · M. Zhu
School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China
e-mail: huxiao@gzhu.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_64

1.1 Pre-processing

To achieve high recognition accuracy and stability in multi-layer neural networks, it is often necessary to stack multiple layers. However, this can lead to complex network structures and extremely high algorithmic complexity. Simply increasing the complexity of the network structure is not a wise approach to improving recognition rates. Instead, starting with the preprocessing of the model can have a cost-effective and high-yield effect.

Data preprocessing plays a crucial role in making the input data more compatible with the model's requirements. This, in turn, improves the training efficiency and generalization ability of the model [3]. For image data, preprocessing techniques aim to standardize the range of pixel values, remove irrelevant information, simplify data, and highlight image features [4]. Common image preprocessing methods include cropping, scaling, and normalization [5, 6].

In this paper, a novel pre-encoder inspired by the mammalian retina is proposed for retinal perceptual nuclei. This pre-encoder performs image preprocessing specifically tailored to retinal data. By incorporating insights from the mammalian retina, this pre-encoder enhances the compatibility of retinal images with subsequent neural network layers, ultimately improving the overall performance of the model.

1.2 Retina

In the human retinal visual system [7], image data received by the cerebral cortex is preprocessed by the retina, and retinal preprocessing improves the performance of the entire visual system: Photoreceptor cells in the retina sense and respond to the intensity, color, and direction of light for initial processing and understanding of images; bipolar cells compress and transmit neural signals to the cerebral cortex; retinal pigment epithelial cells in the retina sense the intensity and duration of light to adapt to different light environments.

The neural structure of the human retina primarily consists of ganglion cells and bipolar cells [8]. Ganglion cells, including optic rod cells and cone cells, convert light stimuli into neural signals, which are then transmitted to the cerebral cortex for analysis. The visual system's efficiency is attributed to the non-uniform distribution of photoreceptor cells in the retina. The central concavity of the retina has over 5 million optic cone cells, which gradually decrease towards the periphery. The central region of the retina provides the highest visual acuity, while the periphery has lower acuity. This preprocessing role of the retina reduces the processing load on the cerebral cortex, resulting in increased efficiency and accuracy of information processing [9].

Taking inspiration from the efficient preprocessing performed by the retina, we propose the use of retina-like pre-encoders. These pre-encoders aim to provide high-quality data inputs to multilayer neural networks, thereby improving their performance from a different perspective, without incurring excessive overhead. This work includes the following:

1. Inspired by retinal preprocessing properties, we propose a plug-and-play retinal-like precoder that can preprocess image data and improve the quality of model input data, thus improving the performance of image neural networks;
2. Design of ablation experiments to verify the effectiveness of this pre-encoder;

2 Method

2.1 Pre-encoder

In image convolution, the pre-encoder plays a crucial role in preprocessing the input image. It performs various operations such as feature extraction and dimension reduction, thereby reducing the computational complexity of image convolution and improving the efficiency of the convolution operation. The pre-encoder not only improves the efficiency, but also contributes to the accuracy of the convolution operation.

2.2 Bionic Pre-encoder

In alignment with neural networks, the parameters of the retina-like pre-encoder can be learned, allowing it to adapt to input data through training on a large amount of data. This eliminates the need for manual feature selection and enables the pre-encoder to possess high generalization capability.

As previously mentioned, the design of the bionic pre-encoder follows the principle of "dense center and sparse periphery" based on the non-uniform distribution of photoreceptor cells in the retina.

For instance, consider a pre-encoder of size 15×15, with empty parameters at 5×5, 11×11, and 13×13. These regions belong to the sparse dispersion area of the pre-encoder, while the central 3×3 area represents the central dense area. Each pre-encoder's central dense area is responsible for feature extraction within its corresponding block. The sparse dispersion area of the pre-encoder sparsely overlaps with other areas, enabling it to jointly extract sparse features with the surrounding kernels. This arrangement allows the pre-encoder to focus on central features and retain spatial semantic information in the image.

Unlike pooling layers that can reduce image resolution and compromise spatial semantic information, the bionic pre-encoder preserves spatial semantics to a certain extent. It avoids introducing excessive redundant information, which can lead to information redundancy and the inclusion of misleading content. Such redundancies can negatively impact the overall model performance.

By carefully balancing dense and sparse regions and avoiding information redundancy, the bionic pre-encoder enhances the performance of the model, preserving important spatial semantic information and improving its overall efficiency and accuracy (Fig. 1).

Fig. 1 Main workflow diagram under the bionic pre-encoder

3 Experimental Results and Analysis

The experiment was conducted on a system running Ubuntu 18.04.6 LTS with Python 3.7.11 and PyTorch (1.10.1 + cu113). A 3080Ti GPU was utilized for acceleration. The objective of the experiment was to assess the effectiveness of the pre-encoder. To achieve this, LeNet5 was chosen as the backbone network, and the performance of the model was observed with and without the inclusion of the pre-encoder.

Two distinct datasets were employed to evaluate the model's performance: the CIFAR-10 dataset and the ODIR-5k dataset.

The ODIR-5k dataset consists of color fundus photographs from the left and right eyes of 5000 patients, along with corresponding diagnostic keywords provided by doctors. This dataset involves fine-grained classification, where the image features exhibit similarity in structure and layout. The differences between various types of images within this dataset are relatively small.

3.1 Results of Experiments

Figure 2 depicts the PR and ROC curves for the model without the Bionic pre-encoder on the CIFAR-10 dataset. On the other hand, Fig. 3 shows the PR and ROC curves for the model with the Bionic pre-encoder on the same dataset.

The curves show that the model with the Bionic pre-encoder outperforms the model without the Bionic pre-encoder. The PR curve is significantly shifted to the upper right corner, and the Average Precision (AP) value increases from 0.38 to 0.56, reflecting a significant improvement of 47%.

Furthermore, the area under the ROC curve (AUC) shows an improvement in both the micro-average and macro-average cases. In the micro-average case, the AUC increases from 0.85 to 0.93, indicating improved performance.

However, when applied to the specific classification task of ODIR, the model fails to converge and cannot function effectively as a classifier. Figures 4a and 5a display the performance of the LeNet5 model during training on the ODIR dataset. Based on the PR curve and the area under the ROC curve, it is evident that the model is inadequate for the ODIR classification task.

Figure 6 illustrates the training-time process and the mean area under the ROC curve (mAUC). It is apparent that the training process diverges significantly, preventing the network from converging and achieving satisfactory performance.

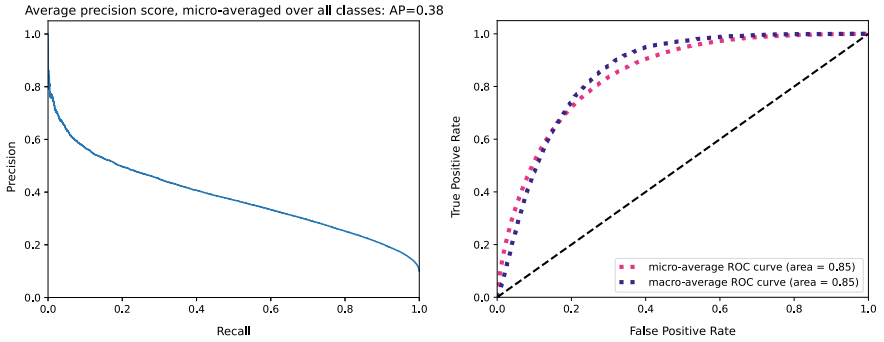


Fig. 2 PR curve(left) with ROC curve(right) on the CIFA10 without bionic pre-encoder

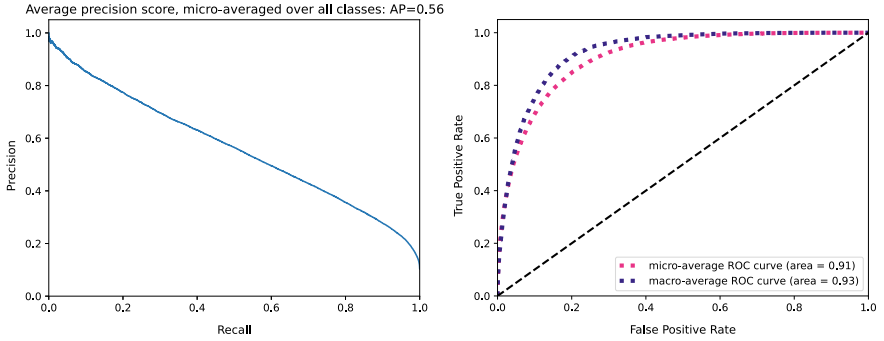


Fig. 3 PR curve(left) and ROC curve(right) on the CIFA10 with bionic pre-encoder

Figures 4b and 5b show that the performance of the model can be effectively improved with the Bionic pre-encoder for the same ODIR dataset, with the AP value increasing from 0.29 to 0.33. The ROC curve shows that the Macro mAUC can be increased from 0.50 to 0.58. Although the model itself is still poor at classification, the Fig. 7 shows that the model gradually converges during the training process and the mAUC can reach about 0.61. Thus, the model's ability to represent the data improves, and thus the mAUC improves.

3.2 Convolution Kernel Distribution

Figures 8, 9, 10 and 11 show images of the distribution of weights for the different convolution layers. All the weights are flattened into a one-dimensional array at the time of counting and then counted up in a histogram.The horizontal axis indicates the range of distribution of these weight values and the vertical axis indicates the epoch.

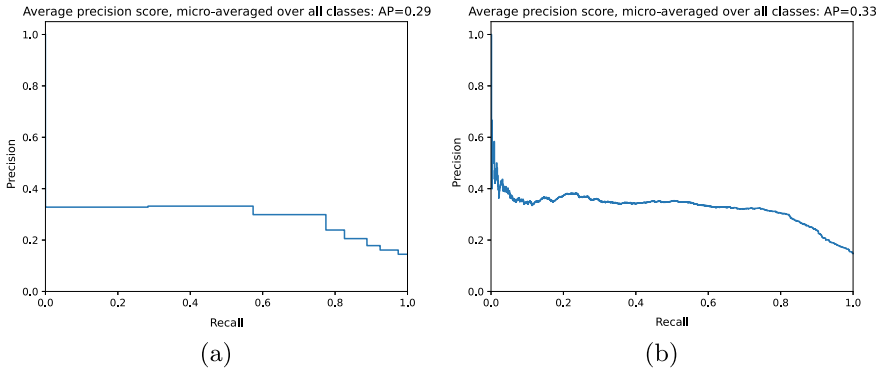


Fig. 4 PR Curve without pre-encoder (**a**) and with (**b**)

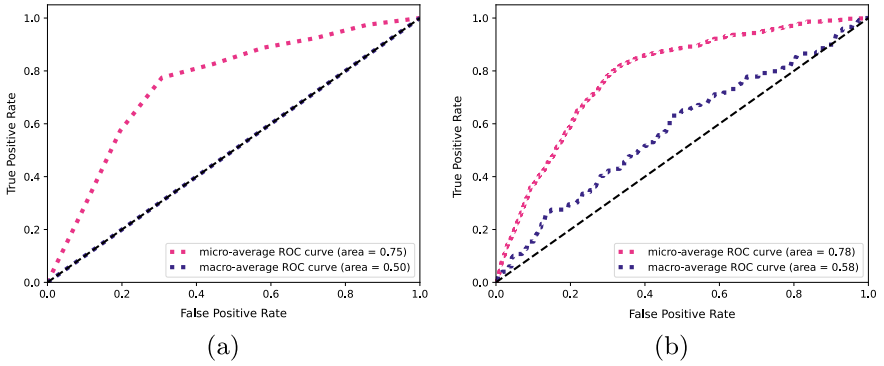


Fig. 5 ROC Curve without pre-encoder (**a**) and with (**b**)

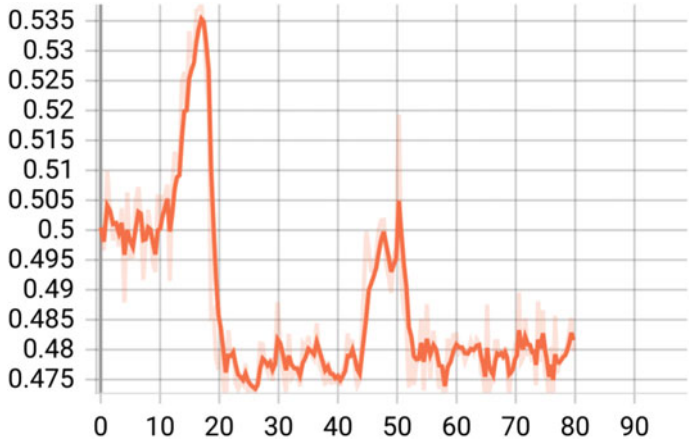


Fig. 6 mAUC without pre-encoder

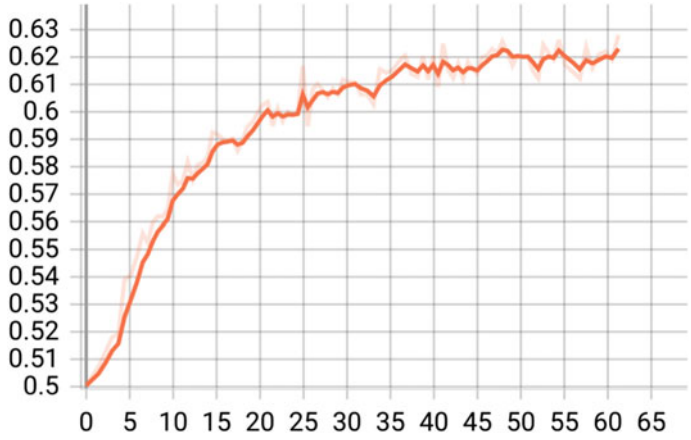
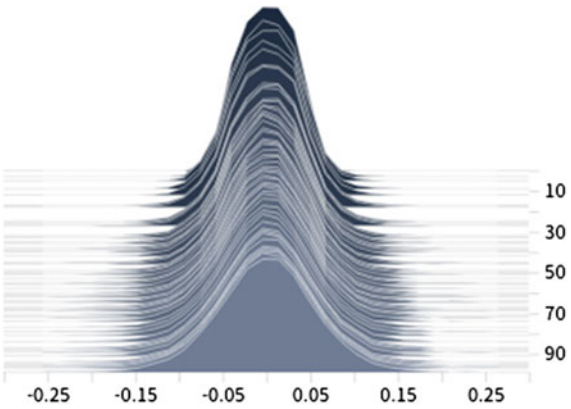


Fig. 7 mAUC with pre-encoder

Fig. 8 Conv1 without pre-encoder



The convolutional kernels in the figure follow a standard distribution, which enhances the stability and convergence of the model. Figures 8 and 10 display the histogram distributions of the weight values for the first and second layer convolutional kernels, respectively, trained without the Bionic pre-encoder. It can be observed that in the absence of the pre-encoder, the kernel distributions are relatively sharp, concentrated, and have smaller variances. These types of kernels assist the model in capturing features with distinct boundaries more effectively.

On the other hand, Figs. 9 and 11 exhibit smoother and broader distributions with wider peaks. This indicates that the kernel values are more dispersed and have larger variances. Such distributions are often suitable for scenarios where the data features are more scattered and have a larger range of variations. The more dispersed kernel values with larger variances contribute to the model's robustness and generalization capabilities.

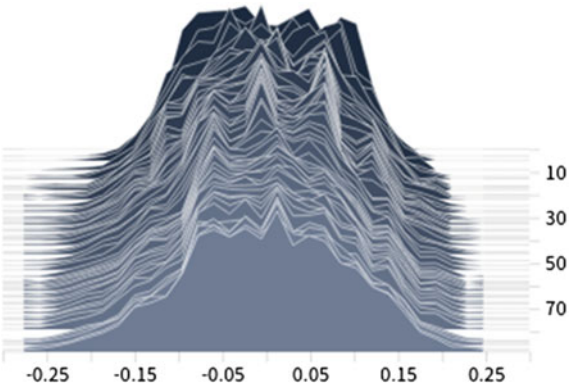


Fig. 9 Conv1 with pre-encoder

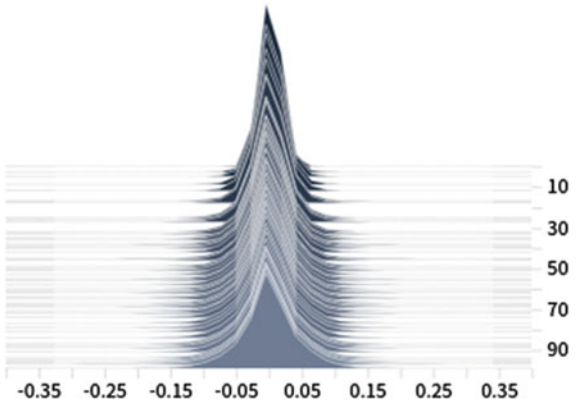


Fig. 10 Conv2 without pre-encoder

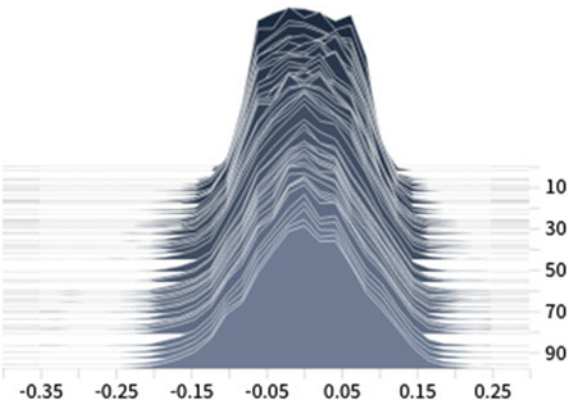


Fig. 11 Conv2 with pre-encoder

4 Conclusion

In this study, we introduced a bionic pre-encoder inspired by the retina of the mammalian visual system, presenting a novel approach to optimize deep learning algorithms. Through a comprehensive comparison experiment, we demonstrated the effectiveness of the proposed precoder in feature extraction, feature dimension reduction, and overall computational efficiency enhancement for image convolution tasks. The results showcased significant improvements in model performance, as evidenced by the precision-recall curves and receiver operating characteristic curves. The precoder exhibited a pivotal role in image convolution, leading to higher efficiency and accuracy. This innovative precoder concept holds great potential to advance the field of image processing technology by enabling more efficient and precise convolution operations. It opens up avenues for further research and applications in various domains, ultimately benefiting a wide range of industries that rely on image recognition and analysis.

References

1. Ding J, Hu X, Zhong X (2021) A semantic encoding out-of-distribution classifier for generalized zero-shot learning. IEEE Sig Process Lett 28:1395–1399
2. Xiang J et al (2021) Cross refinement network with edge detection for salient object detection. IET Sig Process 15(7):425–436
3. Ioffe S , Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. JMLR.org
4. Yun S, Han D, Oh SJ, et al (2019) CutMix: regularization strategy to train strong classifiers with localizable features
5. He K, Girshick R, Dollár P (2018) Rethinking ImageNet pre-training
6. Zhang H, Cisse M, Dauphin YN et al (2017) Mixup: beyond empirical risk minimization
7. Carnevale NT , Hines ML (2006) The neuron book. Cambridge University Press
8. Polyak SL (1941) The retina. Univ, Chicago Press
9. Yan Q, Zheng Y, Jia S, Zhang Y, Yu Z, Chen F, Tian Y, Huang T, Liu JK (2022) Revealing fine structures of the retinal receptive field by deep-learning networks. IEEE Trans Cybern
10. Hanley JA, Mcneil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve

Cross Laser Stripe Center Extraction Based on Gray Center of Gravity Method

Xiaoqing Gao, Wenbin Tan, Zelong Ren, Haoyu Li, and Yinguo Huang

Abstract In the photoelectric self-collimating three-dimensional angle measurement system, the extraction of the center point of the reflected light is a prerequisite for accurate angle measurement. Using cross laser instead of traditional point laser and extracting the center of cross laser stripe is an effective way to improve the accuracy of center point measurement. To address this problem, this paper proposes a method for extracting the center of cross laser stripes based on the gray center of gravity method. Firstly, the working principle of the photoelectric self-collimating three-dimensional angle measurement system is analyzed, and the mathematical relationship between the rotation angle of the mirror and the movement of the center point of the reflected light is determined. Then, the basic principle of applying the gray center of gravity method to obtain the center of the cross stripe is introduced. Finally, the stability of this paper's method under different cross laser cases is verified by experiments, and the accuracy of this paper's method under scattered light interference conditions is verified by comparison with other commonly used methods.

Keywords Gray center of gravity method · Laser stripe · Central extraction

X. Gao · W. Tan · Z. Ren · H. Li
School of Mechanical Engineering, Tianjin University of Commerce, Tianjin 300134, China
e-mail: gxq@stu.tjcu.edu.cn

Z. Ren
e-mail: renzelong@stu.tjcu.edu.cn

H. Li
e-mail: 2667301077@qq.com

W. Tan (✉)
Tianjin University of Commerce, North Star District. 409, Tianjin 300134, China
e-mail: twb@tjcu.edu.cn

Y. Huang
School of Precision Instrument and Opto-electronics Engineering, Tianjin University, Tianjin 300072, China

Y. Huang (✉)
Tianjin University, Weijin Road. 92, Tianjin 300072, China
e-mail: hyg@tju.edu.cn

W. Wang et al. (eds.), *Communications, Signal Processing, and Systems*, Lecture Notes in Electrical Engineering 1032, https://doi.org/10.1007/978-981-99-7505-1_65

1 Introduction

The photoelectric self-collimating three-dimensional angle measurement system is a precision optical measuring instrument [1, 2], which is based on the principle of optical self-collimation, and the angle measurement values are obtained by the optical self-collimator [3]. The traditional three-dimensional angle measurement system uses CCD as a light source to receive and extract the center point of the laser spot and measure it accurately [4], spot edges need to be detected during center point extraction. The image is affected by noise and sharpness, which can easily produce errors and thus affect the accuracy of extracting the center point. Therefore, in this paper, the three-dimensional angle measurement system applies two cross-shaped laser stripe center point extraction. Compared with laser spot center extraction, cross laser stripe has the advantages of simple extraction and higher accuracy.

At present, the extraction laser stripe center is constantly developing in the direction of higher precision and higher stability [5, 6]. There are many error sources that affect the accuracy of laser center extraction, improve the laser center extraction error is important to improve the accuracy of three-dimensional angle measurement system [7]. Pan et al. [8], In order to reduce the noise interference in the laser stripes, similar laser structures in the structured light image were found through block matching, and their coordinate mean was obtained as the final laser center, but the operation efficiency was not high, and the anti-noise interference effect of the straight line was not good [9]. Zhou et al. [10], proposed to use the Euclidean distance and gray value energy of pixels to construct a function and a shortest path search algorithm to centrally extract the laser line with interference, which improves the operation efficiency and accuracy, but the accuracy of this method can only reach the pixel level, which is not suitable for scenarios with high accuracy. Hu and Fang [11], improved the running speed by proposing a method combined with the gray center of gravity method, but only relying on boundary points with large gradient amplitude to participate in the calculation, the accuracy of extraction depends largely on boundary pixels, so it is easily affected by boundary noise.

In actual three-dimensional measurement, the crosshair laser stripes collected by CCD will have strong light noise interference. This paper, apply a gray center of gravity method to extract the central coordinates of the laser stripe image, and firstly, by bilateral filtering and median filtering preprocessing of the image, the interference can be effectively mitigated and extracted to the pixel point set, and calculate the center coordinate of the cross laser stripe. Finally, the algorithm in this paper verifies the stability and accuracy by conducting experimental tests.

2 Photoelectric Self-collimating Three-Dimensional Angle Measurement Principle

As shown in Fig. 1, the laser emitted by laser 1 expands the diameter through the beam expander 2, it increases the diameter of the outgoing beam and reduces the divergence angle through the concentrator 3, and then passes through the cross-shaped plate 4 to make the laser into a cross laser, collimated through the collimator 6, so that the laser is focused to the plane mirror 8, when the plane mirror 8 undergoes an angle change, the reflected laser light is collected by 7 through a beam splitting prism 5, 7 is a linear array CCD.

At this point, if the plane mirror turns the θ angle, It can be known from the knowledge of geometric optics, the deflection of the reflected light relative to the incident light is 2θ. at this point, the gathered crosshairs will produce a displacement t, set the focal length of the lens to f, and the formula can be known according to the trigonometric relationship:

$$\tan 2\theta = \frac{t}{f} \tag{1}$$

When the angle θ of rotation is small, can use approximate relationships $\tan\theta \approx \theta$, the above formula can be reduced to:

$$\theta = \frac{t}{2f} \tag{2}$$

As a result, the collimator lens selected has a focal length of 500 mm, the laser type is SLD, the wavelength is $\lambda = 850$ nm, the spectral width is 15 nm, the planar mirror size is φ50 mm. The coordinates of the center of the laser stripe cross on the CCD can be obtained.

$$x_b = \frac{2\cos^2\alpha\cos\beta\sin\beta}{2\cos^2\alpha\cos^2\beta - 1} \cdot f \tag{3}$$

$$y_b = \frac{2\sin\alpha\cos\alpha\cos\beta}{2\cos^2\alpha\cos^2\beta - 1} \cdot f \tag{4}$$

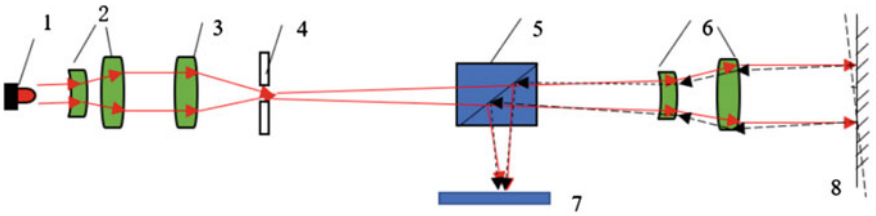


Fig. 1 Self-collimating system schematic

where x_b, y_b A is the center point coordinates, α, β is the angle of rotation of the plane mirror, f is the focal length. Therefore, as long as the coordinates of the center point are measured, the transformation angle of the target object can be derived from the formula.

3 Extraction Principle of the Center of the Crossed Laser Stripe

In the cross-laser stripe, the gray value distribution state of the laser strip is consistent with the Gaussian distribution [12], the point with the largest gray value is the center of the laser stripe, the grayscale distribution function of laser stripes is as follows:

$$G(x) = \frac{a}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{5}$$

where a is the magnitude of the grayscale distribution, σ is standard deviation, μ is the average of the image column coordinates, it is also the central point of the request.

In this paper, the gray center of gravity method is used to process the gray distribution features in the cross-section of each laser stripe line by line, by the direction of the row coordinates, the gray center of gravity points within the laser fringe area are calculated and extracted column-by-column, and these points are used to represent the laser fringe center position of the cross-section. Finally, all center points are fitted to form a centerline. In this paper, the grayscale center of gravity calculation of the horizontal laser stripe is first performed, form the appropriate center of gravity point into the center line, then transpose the image, calculate the gray center of gravity point of the other laser stripe to form the center line, calculate the intersection point of the center line of the two laser stripes, and find the center point of the cross laser stripe. All pixels whose gray value exceeds the threshold of the center of gravity participate in the calculation of the grayscale center of gravity, the calculation formula is as follows:

$$y_{center} = \frac{\sum_{j\in T} j \times g(j)}{\sum_{j\in T} g(j)} \tag{6}$$

where j is the pixel position of each column, T is the collection of pixels participating in the calculation, g(j) is the grayscale value of the pixel, y_{center} is the grayscale center of gravity point.

4 Experiments

This article is written in Visual Studio 2022 and OpenCV4.5.5 platform to complete. Firstly, the image is preprocessed by bilateral filtering and median filtering to reduce the interference of astigmatism noise.

4.1 Stability Experiment

The gray barycenter method was used to extract the center points of different cross laser stripe images to test whether the algorithm could stably calculate the center coordinates. The center point coordinates of different cross laser stripe images were extracted as shown in Fig. 2.

The fitting points in the figure form blue and green lines, and the red point is the center point coordinate. Under different cross laser stripe images, the gray center of gravity method is used to extract the center point of the four groups of images respectively. According to the extraction results of the above figure, the algorithm in this paper maintains stable extraction when extracting different cross laser stripe images, and the red dot in the figure is the coordinates of the extracted center point.

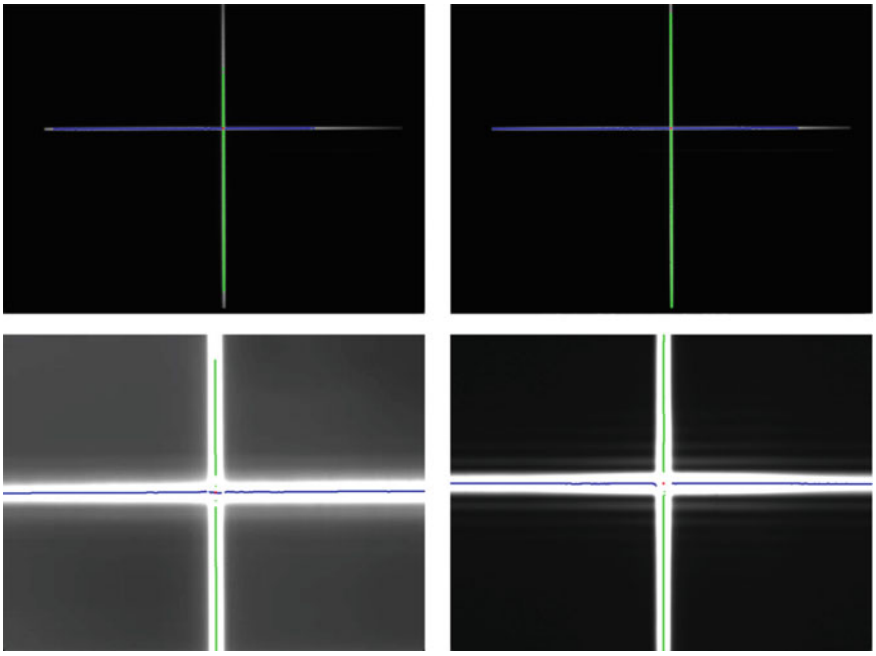


Fig. 2 Extract different cross-laser stripe center points

4.2 Accuracy Test

In this experiment, the center coordinates were extracted when the cross laser fringe carried noise by running different program algorithms. In the actual situation, the strong light noise was randomly generated when the mirror reflected the laser. Firstly, the standard image with the known center of the cross fringe is designed, and the noise interference is superimposed, as shown in Fig. 3.

The coordinates of the center point in Fig. 3 are known, and the coordinates are (360.800, 329.700). The results of gray center of gravity method and other algorithms are shown in Table 1.

Compared with the known central coordinates, the extreme value method will have a great influence due to noise, and its error value does not exceed 48.35 pixels. The maximum error values of the threshold method and the edge method do not exceed 7.8 pixels and 21.2 pixels, respectively, but the extraction results are inaccurate. The overall error value of the grayscale center of gravity method does not exceed 0.09 pixels, and the extracted coordinate error is small. Through comparison, the accuracy of the grayscale center of gravity algorithm in this paper is high.

(a) Single stripe (b) Parallel double stripe (c) Cross-striped

Fig. 3 Noise interference image

Table 1 Different algorithms extract the center point of the laser stripe

Algorithm type	a image		b image		c image	
	Coordinate	Error	Coordinate	Error	Coordinate	Error
Gray center of gravity method	X = 360.859 Y = 329.767	−0.059 −0.067	X = 360.71 Y = 329.777	0.09 −0.077	X = 360.748 Y = 329.771	0.052 0.071
Extreme value method	X = 358.72 Y = 329.05	2.08 0.65	X = 312.45 Y = 325.754	48.35 3.946	No result	
Threshold method	X = 353 Y = 322	7.8 7.7	X = 353 Y = 322	7.8 7.7	X = 353 Y = 322	7.8 7.7
Edge method	X = 382 Y = 316	−21.2 13.7	X = 382 Y = 316	−21.2 13.7	X = 382 Y = 316	−21.2 13.7

5 Conclusion

In this paper, a gray center of gravity method based on cross laser fringe is proposed. Firstly, a straight line is synthesized by calculating the gray center of gravity point of each point of horizontal laser fringe. Secondly, transpose the image matrix to compute another laser stripe gray center of gravity point, the center points fit together into a straight line. Then the center point of the cross laser fringe was calculated and the center point was extracted. Finally, the stability of the proposed algorithm is verified by extracting the center points of different cross laser fringe images, and the accuracy of the proposed algorithm is verified by comparing with other algorithms.

Funding: This work was funded by the Tianjin Research Innovation Project for Postgraduate Students (2022SKYZ027); Tianjin University of Commerce College Students Innovation and Entrepreneurship Training Program Fund Project (202310069062).

References

1. Li Y-G, Jiang S-Y, Feng W-L, Liu Y, Wang Z-T, Wang L (2022) Analysis of horizontal angle measurement error of photoelectric autocollimating theodolite. Aerosp Measuring Technol 42(02):43–49
2. Ren W, Cui J, Tan J (2022) A three-dimensional small angle measurement system based on autocollimation method. Rev Sci Instrum 93(5):055102
3. Chiu JS, Wang SF, Wang WJ et al (2015) Application of total internal reflection and heterodyne interferometry in electrical conductivity measurements. IEEE Sens J 16(2):336–342
4. An Z-J (2020) Research on high precision two-dimensional autocollimator. Huazhong University of Science and Technology
5. Liu T-C, Liu G-H, Hu L et al (2023) Laser center line extraction method based on normal guidance. Appl Opt 44(01):211–218
6. Ji Z-Y, S X-J, Fu W-J et al (2021) Research review of laser strip center line extraction. Meas Control Technol 40(06)
7. Zhao YP, Peng CHQ, Wang J (2018) Analysis of error source for large aperture thin beam auto-collimating measurement system. Semicond Optoelectron 39(3):414–419
8. Pan S, Shao X-J, Li X et al (2021) Subpixel center extraction of laser fringe based on block matching. Laser Infrared 51(06):814–819
9. Wang L, Chen N-N, Wu L et al (2016) Extraction of laser fringe subpixel center under high noise background. Appl Opt 37(02):321–326
10. Zhou Y, Meng X-Q, Jiang D-B et al (2020) Extraction method of fringe center of structured light with complex interference. Chin J Lasers 47(12):172–180
11. Hu Y, Fang S-P (2021) Fringe center extraction method of linear structured light. Adv Laser Optoelectron 58(01):196–200
12. Ke F, Xie J, Chen Y (2016) A flexible and high precision calibration method for the structured light vision system. Optik 127(1):310–314