Lecture Notes in Electrical Engineering 872

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

Communications, Signal Processing, and Systems

Proceedings of the 11th International Conference on Communications, Signal Processing, and Systems, Vol. 1

Springer

Lecture Notes in Electrical Engineering 872

Series Editors

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

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

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

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

China

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

India, Japan, Rest of Asia

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

Southeast Asia, Australia, New Zealand

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

USA, Canada

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

All other Countries

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

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

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

Communications, Signal Processing, and Systems

Proceedings of the 11th International Conference on Communications, Signal Processing, and Systems, Vol. 1

Springer

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

Xin Liu
Dalian University of Technology
Dalian, China

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

Wei Wang
Tianjin Normal University
Tianjin, China

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

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-99-2652-7 ISBN 978-981-99-2653-4 (eBook)
https://doi.org/10.1007/978-981-99-2653-4

Contents

MUSIC-Based Sound Source Localization Algorithm from UVA-Embedded Microphone Array

Sheng Guan[1(✉)], Ruo Jia[1], Libo Qiao[2], Guohui Gu[1], Jianhong Kang[1], and Yujia Song[1]

[1] Jilin Earthquake Agency, Changchun 130117, China
gs-sunny@foxmail.com
[2] Faculty of Psychology, Tianjin Normal University, Tianjin 300387, China

Abstract. The application of an unmanned aerial vehicle (UAV) equipped with microphone array for field search and rescue has demonstrated good potentials. One of the main issues in using a multirotor UAV for sound source localization is that the ego noise of the UAV's rotors interferes with the audio observation and degrades the sound source localization performance. This paper introduces a variant of the Multiple signal classification (MUSIC) technique to audio processing embedded in the UAV scenarios, suppressing background noise by employing the noise datasets to reconstruct the covariance matrix. As simulation results are described, the performance improvement is achieved while using the new approach compared with the generalized cross-correlation with phase transform (GCC-PHAT), non-linear generalized cross-correlation (GCC-NONLIN) and conventional MUSIC method.

Keywords: MUSIC algorithm · sound source localization · unmanned aerial vehicle

1 Introduction

Unmanned aerial vehicles are an emerging technology with numerous of current and potential applications in a wide range of applications and industries. UAV's could be configured with other non-intrusive sensors (cameras, gyroscope, inertial measurement module, motor controllers, compass, motor controllers, etc.) which may deliver helpful extra information. In search and rescue missions, a search work in rugged outdoor terrain can be extremely time-consuming and physically demanding, and tracking of the location of several rescue groups over a wide area where there is no communications infrastructure may be an additional issue. Drones can aid with rescuers in locating victims more quickly in a wide search area. Besides, they can fly over obstacles and inaccessible terrain in a coordinated manner to reach potential target regions directly, and ultimately, cover a large-scale area more rapidly than ground units do. For example: a drone search for a lost or an injured person who can call for help, but cannot move or use any communication

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 1–8, 2023.
https://doi.org/10.1007/978-981-99-2653-4_1

equipment [1]. To deal with such circumstances, several drones embedded instruments have been created by utilizing the video camera, acoustic signals (moaning, shouting) have obvious vital signs, UAV-based acoustic application scenarios have been given far less attention in research. For all that, drones configured with a microphone array may provide some strengths in emergency scenarios, particularly when visual feedback is lacking because of poor lighting conditions (dust, night, fog, smoke, dense forests, or other clutter, etc.) or obstacles that restrict the visibility.

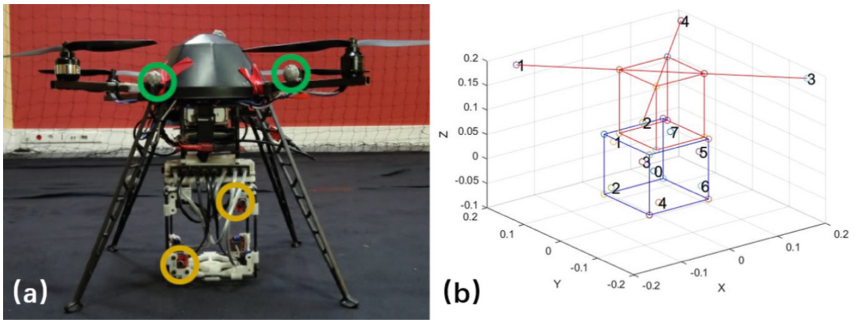


Fig. 1. Data acquisition system architecture

(a) The four-rotor unmanned aircraft applied to the DREGON data collection, configurated with a 3D-printed 8-microphone array embedded on the base. Green loops emphasize two of the passive tags applied for movement acquisition, yellow loops emphasize two of the microphones.
(b) The relative position of the rotors of the drone in the array's coordinate reference frame is shown in the figures.

To push ahead the research on acoustic characteristics during unmanned aircraft flight, and development of novel acoustic source position approaches for both search and rescue, 2019 IEEE Signal processing Cup [2, 3] presents a novel publicly available dataset named DREGON for DRone EGonoise and localization. The DREGON datasets consists of both in-flight and static acoustic recordings of speech or white-noise signals from a loudspeaker, a four-rotor drone configured with an eight-channel stereo microphone array was utilized to record sound for diverse scenes both with or without a desired sound source emitted by a loudspeaker. The main task is to estimate the direction of a sound source (azimuth, elevation) using audio measurements from an 8 microphones array embedded in a flying UAV (Fig. 1). In this paper, we focus on the performance of speech sound source position algorithm in stable hovering conditions datasets. Table 1 presents an introduction to the static database.

With reference to sound source position on drones, MUSIC [4, 5] seems at present the perfect choice available because of its multi-source positioning ability and to noise robustness [6]. In this study, we have studied the spectral properties of the self-noise of a multi-rotor UAV through audio signals captured by 8-channel microphones, and

improve the robustness of MUSIC algorithm using individual motor recordings, the performance of diverse prevalent auditory localization approaches, i.e., GCC-PHAT and GCC-NONLIN algorithms [7], are compared on the DREGON database with pay special attention to real time capabilities. The rest of this paper is organized as follows: In the Sect. 2 the signal model and acoustic feature will be introduced. Then the sound source localization method will be depicted based on MUSIC algorithm in Sect. 3. Experimental results are given in Sect. 4. Conclusion and discussion are in Sect. 5.

Table 1. Parameter of static datasets

Parameter name	Description
Number of sources	1
Source signal	Speech (males or females)
Sample frequency	44.1kHz
Block size	2.5 s (110250 points)
Additional information	Individual motor speed Individual motor recordings 50 60 70 80 90 (turns/s)

2 Signal Model and Acoustic Analysis

2.1 Signal Model

M microphones embedded on the multi-propeller UAV picks up the sound delivered by a static loudspeaker under far field assumption from an unknown (azimuth θ, elevation φ) direction, arriving at different array elements with different time delays τ, time delays collapse to simple phase shifts, the localization of the microphone array in a spherical coordinate system, taking the first microphone signal, $x_0[n]$, as reference, the observation can be written as

$$X_M[n] = \begin{bmatrix} x_0[n] \\ x_0[n-\tau_1] \\ \vdots \\ x_0[n-\tau_{M-1}] \end{bmatrix} = \begin{bmatrix} 1 \\ e^{j2\pi f \tau_{p,1}} \\ \vdots \\ e^{j2\pi f \tau_{p,M-1}} \end{bmatrix} \otimes s_p[n] = A \otimes S + V \tag{1}$$

where ⊗ denotes convolution operator, $s_p(n)$ is the p^{th} source signal, A, S and V is the narrowband steering vector, the target sound, and the self-noise, respectively. The time delays τ can be modelled as is

$$\tau_{p,m} = \frac{d_m}{c} \begin{bmatrix} \cos(\varphi_p)\cos(\theta_p) \\ \cos(\varphi_p)\sin(\theta_p) \\ \sin(\varphi_p) \end{bmatrix} \tag{2}$$

Following a short-time Fourier transform (STFT), the convolution in the time domain is approximately characterized as an instantaneous mixing operation at each time-frequency (TF) point [k, f]:

$$X[k,f] = A \times S[k,f] + N[k,f] \tag{3}$$

where f and k are frequency and frame indices, respectively.

2.2 Acoustic Analysis of UAV

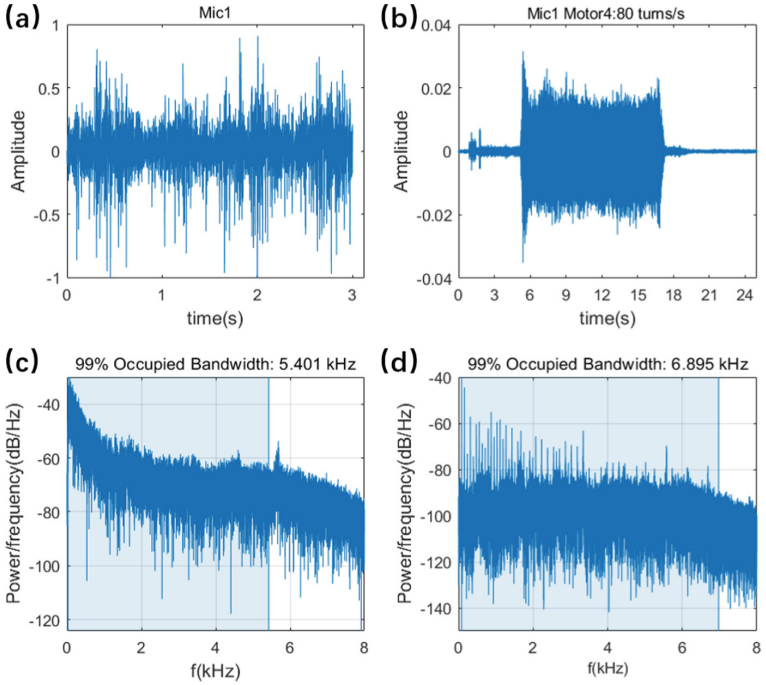


Fig. 2. Acoustic analysis of observation signal(a)-(c) Time-domain waveform and occupied bandwidth of the 1.wav. (b)-(d) Time-domain waveform and occupied bandwidth of the motor4_80.wav.

As reference in [9], three major components can be confirmed: harmonic noise related to the propeller's rotation; mechanical noise; wind noise due to the air flow. Figure 2 shows spectrograms noise and speech from one of the microphones, majority of the noise energy is in the frequencies under 3 kHz, and can be interpreted by the reality that the energy of sound and drone noise signals is primarily situated in a parallel frequency domain. As the microphones mounted on the drones are more closely to the propellers and motors than target acoustic source, self-noise severely masks the target source, resulting in a very low signal-to-noise ratio (e.g., S/N ratio less than −15 dB).

3 Sound Source Localization

The MUSIC algorithm is one of the most widely-used subspace-based approach known as more robust to noise in positioning. The frequency-domain correlation matrices of the observed signal R_x and noise signal R_n, defined as

$$R_x(k,f) = E[x(k,f)x^H(k,f)] = \sum_{k=1}^{K} x(k,f)x^H(k,f) \tag{4}$$

$$R_n(k,f) = E[v(k,f)v^H(k,f)] = \sum_{k=1}^{K} v(k,f)v^H(k,f) \tag{5}$$

where E[•], $[]^H$ denote expectation operator, complex conjugation, respectively. One major issue of estimating the direction (θ, φ) of a sound source given audio recordings from a microphone array is the noise produced by the UAV, noise reduction is therefore necessary to estimate the target sound before further processing. Although the noise generated by UAV shows time-varying characteristics in the spectrum, it is approximately white noise in each frequency bins, applying generalized eigenvalue decomposition to $R_x(R_n)^{-1}$, this decomposition suppresses the noises

$$\begin{aligned} R_s(k,f) &\approx R_x(k,f)R_n(k,f)^{-1} \\ &\approx [\, U_s(k,f)\ U_n(k,f)\,][\begin{matrix} V_s(k,f)/V_n(k,f) & \\ & I \end{matrix}][\begin{matrix} U_s(k,f) \\ U_n(k,f) \end{matrix}] \end{aligned} \tag{6}$$

where U_s, U_n denote signal and noise subspaces matrices, V_s, V_n are the corresponding eigenvalue matrices, I is identity matrix. According to the orthogonality between noise and signal subspaces, $U_s \perp U_n$, spatial power spectrum can be obtained from

$$\Phi(k,f,\tau) = \frac{1}{A(f,\tau)U_n^H(k,f)U_n(k,f)A^H(f,\tau)} \tag{7}$$

Pooling function can be chosen between max and sum function.

$$\Phi^{sum}(\tau) = \sum_{k=1}^{K}\sum_{f=1}^{F} \Phi(k,f,\tau) \tag{8}$$

$$\Phi^{\max}(\tau) = \max_{k} \sum_{f=1}^{F} \Phi(k,f,\tau) \tag{9}$$

Note that, one could prefer can max function for sources that are only active within a few time frames in analyzed signal due to integration of irrelevant information when the sources are inactive.

The flowchart of the MUSIC-based sound source localization as shown in Fig. 3, The main steps of the extend MUSIC are summarized as follows:

Step1: Perform the short-time Fourier transform of observed multichannel signal (Eq. 3);

Step2: Compute frequency domain covariance matrix R_x (Eq. 4) and R_n (Eq. 5);
Step3: Perform eigenvalue decomposition on the covariance matrix $R_x(R_n)^{-1}$ (Eq. 6) to obtain the noise subspace matrix;
Step4: Compute the corresponding spatial spectrum (Eq. 7) and sum pooling (Eq. 8);
Step5: Choose P local peaks of Eq. 8 as the estimates of the sound source localization.

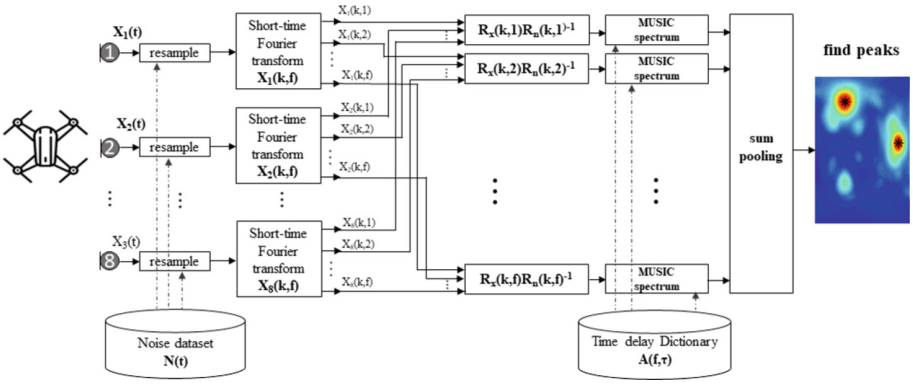


Fig. 3. The architecture of MUSIC-based sound source localization

4 Experiments Results

In this section, we compare the GCC-PHAT, GCC-NONLIN, standard MUSIC and extend MUSIC algorithm on hovering scenarios. Two evaluation indexes are considered for both azimuth and elevation: a source localization is considered correct if and only if both the estimated azimuth and elevation angles are within 5° of the true azimuth and elevation angles, and the average of the absolute angular errors of successful estimates. The analysis window is fixed to 250 ms (4000 points) and 500 ms (8000 points), the short time Fourier transform window is 1024 with 512 overlap, audio signals are down-sampling to 16 kHz. Noise suppression covariance matrix was constructed with motor1_70 and motor4_70 for extend MUSIC method. The results are showed in Table.2, the extend MUSIC method more robust to noise than GCC-PHAT GCC-NONLIN and MUSIC. The figure also shows that the size of the analysis window has a significant impact on the positioning accuracy, with a larger window implying better performance.

Table 2. Performance of different algorithms

Methods/Parameter	250 ms			500 ms		
	Accuracy (%)	Azimuth error (°)	Elevation error (°)	Accuracy (%)	Azimuth error (°)	Elevation error (°)
GCC-NONLIN	59.60%	22.59	8.09	76.60%	7.46	4.90
GCC-PHAT	60.60%	25.89	9.15	82.00%	8.51	4.83
MUSIC	64.00%	24.36	9.22	79.00%	10.21	5.18
This paper	**71.00%**	**23.05**	**9.18**	**83.30%**	**5.08**	**3.90**

5 Conclusions

In this paper, we exploit the narrow band MUSIC method to address sound source localization from static speech datasets, and static noise datasets are used to improve the robustness of the MUSIC algorithm, the experimental results basically satisfy the positioning requirements. Ego-noise reduction has been a technological hot button in the robotic sector for several years [8, 9]. In summary, many time-varying non-stationary and stationary noise components are observed in the speech recordings. It is necessary to develop a more accurate noise model to address sound source localization for search and rescue with drone, and additional modalities such as rotating speed of motor could be employed for this objective.

In our future work, we will extend the MUSIC algorithm to the application of in-flight recordings, which includes several subtasks such as source localization and tracking, signal extraction, and enhancement.

Acknowledgement. The authors would like to express our warm gratitude to the open-source database (http://dregon.inria.fr/) staff who made it possible, thanks to anonymous reviewers. This work supported by CEA_EDEM202106, NORSCBS21–07, Youth Research Project of Tianjin Normal University (52XQ2101).

References

1. Zimroz, P., Trybała, P., Wróblewski, A., et al.: Application of UAV in search and rescue actions in underground mine—A specific sound detection in noisy acoustic signal. Energies **14**(13), 3725 (2021)
2. Strauss, M., Mordel, P., Miguet, V., et al.: DREGON: Dataset and methods for UAV-embedded sound source localization In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 1–8 (2018)
3. Deleforge, A., Di Carlo, D., Strauss, M., et al.: Audio-based search and rescue with a drone: highlights from the IEEE signal processing cup 2019 student competition [SP competitions]. IEEE Signal Process. Mag. **36**(5), 138–144 (2019)
4. Schmidt, R.: Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **34**(3), 276–280 (1986)

5. Liang, Q., Zhang, B., Zhao, C., et al.: TDoA for passive localization: Underwater versus terrestrial environment. IEEE Trans. Parallel Distrib. Syst. **24**(10), 2100–2108 (2012)
6. Salvati, D., Drioli, C., Ferrin, G., et al.: Beamforming-based acoustic source localization and enhancement for multirotor UAVs In: 2018 26th European Signal Processing Conference (EUSIPCO). IEEE, 987–991 (2018)
7. Blandin, C., Ozerov, A., Vincent, E.: Multi-source TDOA estimation in reverberant audio using angular spectra and clustering. Signal Process. **92**(8), 1950–1960 (2012)
8. Sinibaldi, G., Marino, L.: Experimental analysis on the noise of propellers for small UAV. Appl. Acoust. **74**(1), 79–88 (2013)
9. Wang, L., Cavallaro, A.: Ear in the sky: Ego-noise reduction for auditory micro aerial vehicles in 2016 In: 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 152–158 (2016)

Routing Protocol Based on Q-Learning in WBAN

Jiaxuan Chen[1,2(✉)] and Jiasong Mu[1,2(✉)]

[1] College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
gs-sunny@foxmail.com

[2] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China

Abstract. In order to solve the problems of telemedicine and human health monitoring, wireless body area network (WBAN) came into being. WBAN is a communication network based on human body, which is composed of network elements related to human body to monitor and maintain human life information. Its sensors throughout the body, make people can always get the required physiology data. Energy consumption is one of the most challenging problems in the application of WBAN due to the energy limitation of body area network devices. In order to improve the working life of nodes, this paper proposes an energy-saving routing protocol based on Q-learning. By exchanging the information between Neighbor nodes, the residual energy, communication energy consumption and routing loss are sent to the source node as feedback information, so that the source node can learn the routing strategy. The current device sets the node with the highest Q value as the destination node to find a path with less energy loss. The simulation results show that the route loss is reduced by about 15% and the average lifetime of the system is prolonged by about 10% by adding the consideration of routing loss and second-step transfer.

Keywords: Q-learning · Ad-Hoc · WBAN · Routing protocol · Routing loss · Working life

1 Introduction

In 2001, the concept of WBAN was first proposed [1]: WBAN is a communication network composed of several micro smart sensors and an sink node. Its original purpose was to better monitor human health indicators and provide real-time and effective data for medical treatment. With the deepening of the research, WBAN in telemedicine, movement and the military training, electronic entertainment areas such as display own superiority also gradually.

In order to adapt to different application requirements, WBAN sensors form is also different. For example, implantable sensors designed to measure internal body data such as blood pressure and blood sugar; Attachment sensor for measuring heartbeat, pulse and

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 9–17, 2023.
https://doi.org/10.1007/978-981-99-2653-4_2

other physiological data; And wearable or hand-held sensors designed to serve electronic entertainment. For these kinds of sensors, their common ground is that the sensors are located near the human body and move with the human body's movement. This makes THE WBAN sensor face two common challenges: 1. The topology of the network may change at any time, and the traditional routing protocol is not applicable in WBAN [3]; 2. Electromagnetic wave will be affected by human body during transmission, resulting in greater losses [4, 5]. And applied in the field of electronic entertainment sensor is usually wear simple, the requirements of the rechargeable high transmission rate sensor. Sensors are small, difficult to install, and energy supply is insufficient, which makes the sensor monitoring human health has been facing the challenge of energy problems since its birth. This paper will focus on meeting the challenges of energy problems and reducing the energy loss in the system.

At present, in order to solve the problem of limited sensor energy in the medical field and the problem of electromagnetic wave energy loss in the transmission of WBAN sensor, researchers commonly use the following measures: 1. Research the energy-saving MAC layer protocol; 2. Design efficient transmission antenna; 3. Design energy-efficient routing protocols.

In this paper, a routing protocol based on Q-learning will be designed to comprehensively solve the problem of sensor energy problem and energy loss in the transmission process. Compared with other energy-saving routing protocols, routing protocols based on Q-Learning can better consider the relationship between residual energy of nodes and energy consumption of nodes. The system can predict and estimate the future energy by Q value. Meanwhile, the update mode of Q value can effectively help the system to maintain routing and select the shortest path [6]. However, the loss of energy transmitted between nodes is not considered. In this paper, in the process of acquiring Q value, consideration will be given to the transmission loss and the energy received by the destination node. So that the system can consider the electromagnetic loss when choosing the path, choose the path with less loss. At the same time, in order to reduce the problem of repeated operation that may occur in route maintenance, this paper increases the consideration of one step in the process of Q value update to two steps, and sets the weight at the same time, so that the system will not only consider the second step but ignore the first step. This reduces the possibility of repeated maintenance caused by link interruption during route maintenance and prolongs the working life of the system to a certain extent.

This paper will start from WBAN routing protocol, improve the existing routing protocol, so as to achieve the purpose of increasing the working life of WBAN system. The second chapter of this paper will describe the current routing protocol research on WBAN system energy, the third chapter will describe the principle of routing protocol, the fourth chapter to simulate the comparison of algorithms, in the fifth chapter to summarize the whole paper.

2 Related Work

For WBAN system, its network topology may change at any time, so essentially, the routing protocol of WBAN belongs to the self-grouping routing protocol. Meanwhile,

for the traditional ad-Hoc routing protocol, indexes such as delay and bandwidth that can improve QoS are considered in the design, but this increases the energy consumption burden of nodes. Therefore, for the limited energy OF WABN system, the traditional routing protocol is not directly applicable.

To solve this problem, researchers proposed a class of energy-aware routing protocols. This kind of protocol is designed on the basis of considering the total energy consumption of nodes and prolonging the working life of network. Therefore, in the design of this kind of routing protocol, it is necessary to know the cost of node sending information and give the evaluation criteria. For example, the routing protocol in literature [7] takes into account the distance between the node and the sink node and the remaining energy of the node to construct a function. The selected criteria are that the next hop node has more residual energy and is close to the sink node. However, such routing protocols do not take into account the balance of power consumption. In a multi-hop network, nodes near the center of the network frequently undertake trunks and heavy forwarding tasks. As a result, some nodes on the network die prematurely and the network connectivity is affected. In order to prevent the occurrence of such a situation, literature [8] added the index of communication times on the basis of the above factors, and set weights for each parameter when constructing the selection function, so that the same node will not be frequently selected in the routing selection of the network and the premature death of the node will be prevented.

For WBAN system, the biggest loss occurs in the electromagnetic wave transmission between nodes. In order to reduce the influence brought by such loss, the routing protocol proposed in literature [9] comprehensively considers the residual energy of nodes, path loss and total energy consumption of the system, so as to select the optimal path. In literature [10], indicators such as link quality and transmission hop count are taken into additional consideration. The above two routing protocols take into account the electromagnetic loss between nodes, but do not take into account the possibility of premature death of nodes, which makes the system prone to link fracture in real work.

In order to meet the above situation, this paper proposes an improved routing protocol, which comprehensively considers the residual energy of nodes and the energy loss during transmission, and enables the system to consider the next two steps during routing maintenance, so as to prolong the working life of the system and reduce the system energy consumption.

3 Routing Protocol

For routing protocols to solve energy problems, the fundamental purpose is to find the optimal path between the source node and the destination node when sending information. This process can be regarded as a Markov process. To implement this process, you must determine return values, states, and actions. However, if the action selection of the system is determined only by the return value, the future will not be considered enough. Therefore, Q-learning mechanism is introduced to strengthen the consideration of the future in the system and comprehensively maintain the energy consumption of the system.

Q-learning regards nodes in the network as Agent, who choose the next action according to the status of the current environment. After performing the action, the environment

returns a reward value signal to Agent. By interacting with the environment, the Agent learns the optimal strategy from it. And learning the optimal strategy for the whole process in the data is how to obtain the maximum return value. The specific flow of the algorithm is shown in Fig. 1.

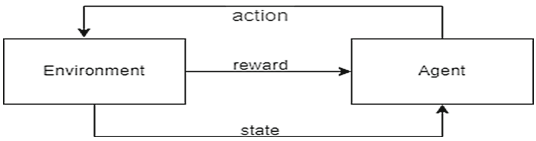


Fig. 1. Basic principle of node action

3.1 Return Value Acquisition

The purpose of q-learning in WBAN system is to optimize the strategy π, so that the Agent can obtain the maximum return value by executing action A under state S, that is, π(S) = A. While the accumulated return value R of Agent in the process of executing action at from state S_t to reaching target strategy π is:

$$R^n(S_t) = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} \cdots = r_t + \gamma R^{\pi}(S_{t+1}) = \sum_{i=0}^{\infty} \gamma^i r_{t+i} \tag{1}$$

where γ is the discount factor, reflecting the proportion of future action return value. R_t is the return value of the Agent node from state S_t to state S_{t+1}. In the current study, R_T did not take into account the loss of electromagnetic waves as they travel across the surface of the human body. Formula (2) is the transmission loss formula of electromagnetic wave near human body, where n is the attenuation factor,

$$PL(dB) = PL(d_0) + 10 * n * \log_{10}(\frac{d_0}{d}) \tag{2}$$

When the electromagnetic wave passes through different positions of the human body, the n will be greatly different. WBAN system sensors are all over the body, so different path selection will cause different electromagnetic wave loss during transmission. In actual use, multiple transmission will occur to meet the accuracy of data, resulting in energy waste. Therefore, in this paper, the energy acquired by the next hop node and the transmission loss of the node are added into the calculation of the return value to obtain a more accurate return value. R_t can be obtained as follows:

$$r_t = \frac{E_R}{E_C} + \frac{E_G}{PL} \tag{3}$$

where E_R refers to the remaining energy of the node, E_C refers to the energy consumed by node transmission, E_G refers to the energy acquired by the next hop node, and PL refers to the transmission loss.

3.2 Optimal Route Selection

After obtaining a more accurate return value for system evaluation, the return value cannot be directly applied to route selection. If only the return value is used for route selection, the future prediction is not enough, which is not conducive to improving the lifetime of the system. In order to improve the future prediction and estimation of nodes in routing selection, a Q value function Q (s, a) is defined, whose specific expression is shown in (4).

$$\begin{aligned} Q^{\pi}(s,a) =& E_{\pi}[R_t|S_t = s, A_t = a] \\ =& E_{\pi}[\sum_{i=0}^{\infty} \gamma^i r_{t+i}|S_t = s, A_t = a] \end{aligned} \tag{4}$$

The Q value represents the expected return immediately after performing action A in state S and the discounted return after using the best strategy π.

In practical applications, the Sink node device is considered as a device with computing capability. In this way, the Sink node can calculate Q value according to the above mechanism and update the Q value table after receiving the return value feedback from other nodes, and finally select the path according to the size of Q value. In this way, the system can issue instructions to nodes from the perspective of long-term continuity to achieve the purpose of improving the lifetime.

3.3 Routing Updates

In practical application, in order to meet various requirements, the Sink node will set some thresholds for each node in the system, such as energy consumption threshold and loss threshold. If the data of a node in the route selection path does not reach the threshold, the sink node will be sent messages along the path in reverse direction. Meanwhile, the path with a large Q value will be rearranged to send information and the Q value table will be updated. Update rules are as shown in (5):

$$Q_{t+1}(s_t, a) = (1-\alpha)Q(s_t, a) + \alpha[r_t + \gamma \max Q(s_{t+1}, a^{'}) + (1-\omega)Q(s_{t+2}, a^{''})] \tag{5}$$

In this paper, a new learning factor ω ($\omega \in (0.5, 1]$) is introduced. After joining ω system can consider the two steps of routing in the future. With the addition of ω, the system allows for step 1 and Step 2 updates. The system can predict routing problems earlier to maintain the stability of routes, improve the stability of the system, and improve the life cycle of the system and the working efficiency of nodes. At the same time, $\omega > 0.5$ guarantees the weight of the system to the prediction of the first step, so that the system will not ignore the problem of the first step because there is no problem in the second step.

4 Simulation Comparison

4.1 Simulation Parameters

In order to verify the WBAN routing algorithm based on Q-learning, the route is simulated. Its simulation parameters are shown in Table 1:

Table 1. Simulation parameters

network environment	120cm*200cm
Transmission energy consumption	100pJ
Min receiving energy consumption	40pJ
n	3.5–9.4
γ	0–1
ω	0.5–1

4.2 Simulation Process

1. Verify whether Q-learning can improve the working life of the network;
2. By comparing the energy obtained by nodes and routing loss, verify whether the improved calculation method of return value is beneficial to improve the working quality of nodes;
3. By comparing the node life cycle, verify whether the improved Q value update rule is conducive to improving system stability.

4.3 Simulation Results

4.3.1 Comparison Between Q-learning and Non-q-Learning:

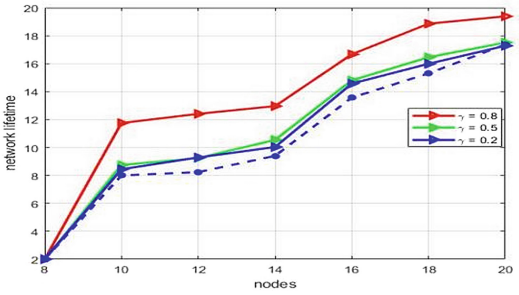


Fig. 2. Q-learning and no Q-learningS

As shown in Fig. 2, the solid line part represents the lifetime of Q-learning under three γ parameters, and the dotted line part represents the lifetime of no Q-learning. By comparison, the lifetime of systems without Q-learning is smaller than that of systems with Q-learning, which indicates that the addition of Q-learning is conducive to improving the lifetime of system nodes.

4.3.2 Quality of Work

As shown in Fig. 3, the left system only has the basic Q-learning algorithm, while the right system is an improved Q-learning algorithm. It can be found that when the path

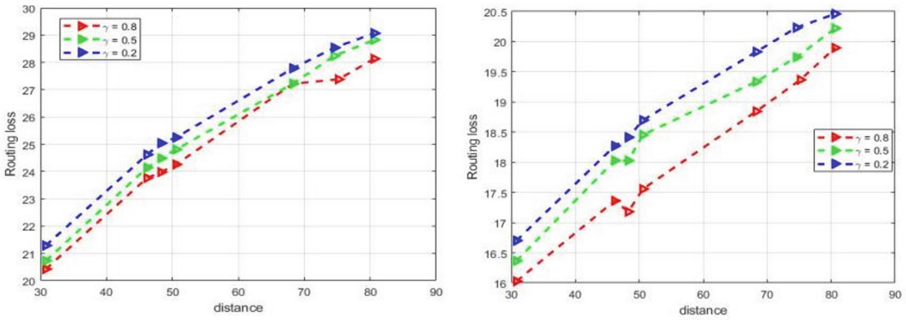


Fig. 3. Routing loss

distance is the same, the energy loss caused by the improved algorithm is smaller. It is proved that the improved algorithm can better consider the electromagnetic loss and reduce the electromagnetic loss during transmission. In Fig. 4, receiving energy from nodes also reflects the improvement in system efficiency.

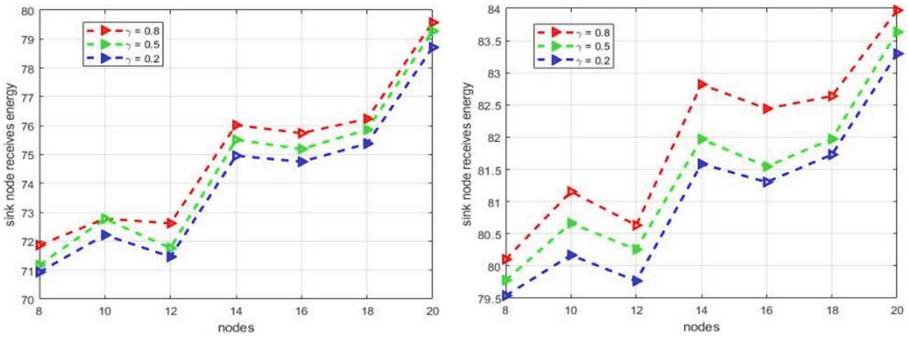


Fig. 4. Received energy

4.3.3 Average Lifetime

As shown in Fig. 5, the right figure is the improved algorithm, and the left figure is the basic algorithm. It can be found that when the number of nodes is small, the node life cycle is almost the same, because when the number of nodes is small, the system path is relatively fixed. However, when the number of nodes is large, the system has more paths to choose, but the improved algorithm takes more into account when updating paths, so the repeated updates that may occur are reduced, thus improving the network lifetime.

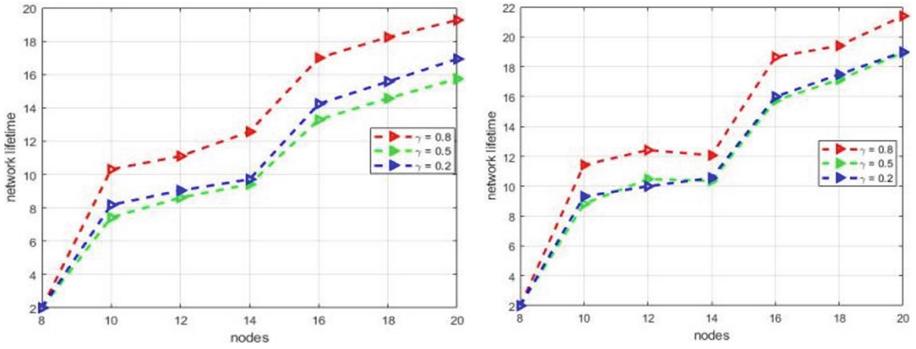


Fig. 5. Network lifetime.

5 Conclusion

This paper introduces the WBAN routing protocol algorithm based on Q-learning, and designs the corresponding simulation experiment. Through experiments, the method of obtaining return value and updating Q value introduced in this paper can effectively improve the lifetime of the system. When the number of nodes is large, the life cycle of the system can be increased by about 10%. At the same time, the routing loss of transmission is reduced, and the overall working quality of the system is improved.

References

1. Dam B, Pitchers S, Barnard M. From PAN to BAN: Why body area networks," presented at the Wireless World Res. Forum 2[J]
2. Roopali Punj, Rakesh Kumar. Technological aspects of WBANs for health monitoring: a comprehensive review[J]. Wireless Networks,2019,25(3)
3. WangYi Dan, Liu Yi. Analysis and Comparison of WSN Routing Protocols Oriented to WBAN[J]. Applied Mechanics and Materials,2012,241–244(241–244)
4. Ullah, S., et al.: A Comprehensive Survey of Wireless Body Area Networks[J]. J. Med. Syst. **36**(3), 1065–1094 (2012)
5. Lee, G., Garner, B., Li, Y.: Simulation and measurement of electromagnetic wave propagation on dynamic human bodies[J]. IET Microwaves Antennas Propag. **11**(10), 1347–1353 (2017)
6. Zhang Lieping,Tang Liu,Zhang Shenglan,Wang Zhengzhong,Shen Xianhao,Zhang Zuqiong. A Self-Adaptive Reinforcement-Exploration Q-Learning Algorithm[J]. Symmetry,2021,13(6)
7. Khanna A, Chaudhary V, Gupta S H. Design and Analysis of Energy Efficient Wireless Body Area Network (WBAN) for Health Monitoring[J]. Transactions on Computational Science XXXIII. 2018: 25–39
8. Sangwan, A., Bhattacharya, P.P.: Reliable Energy Efficient Multi-hop Routing Protocol for Heterogeneous Body Area Networks[J]. International Journal of Sensors Wireless Communications and Control. **8**(1), 47–56 (2018)
9. Hojjatollah, E., Minaei, B.B.: EMRP: Evolutionary-based multi-hop routing protocol for wireless body area networks[J]. Aeu-International Journal of Electronics and Communications. **93**, 63–74 (2018)

10. Anwar, M., Abdullah, A.H., Altameem, A., et al.: Green Communication for Wireless Body Area Networks: Energy Aware Link Efficient Routing Approach[J]. Sensors. **18**(10), 3237 (2018)
11. Kok-Lim Alvin Yau et al. Application of reinforcement learning to wireless sensor networks: models and algorithms[J]. Computing, 2015, 97(11) : 1045–1075
12. Shayesteh Tabatabaei and Mohammad Teshnehlab and Seyed Javad Mirabedini. A New Routing Protocol to Increase Throughput in Mobile Ad Hoc Networks[J]. Wireless Personal Communications, 2015, 83(3) : 1765–1778

A Novel Two Stage Hybrid Multi-user Precoding Design for Massive MIMO System

Liuwei Zhang, Xin Zhang(✉), Xiu Zhang, and Yue Wang

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
ecemark@tjnu.edu.cn

Abstract. Massive multi-user multi-input multi-output (MIMO) with a hybrid structure is considered a promising technology for the future wireless communication network. In this paper, we propose a new two stage hybrid precoding method for multi-user massive MIMO system. Firstly eliminating the user-interference with block diagonalization method, then through calculating the optimal hybrid precoding for each user, the proposed hybrid precoding achieves maximization the sum of the sum rate of all users when ensures maximization of the sum rate of each single user. Simulation results show that the proposed scheme achieves better performance than the state-of-the-art schemes under the same hardware cost.

Keywords: Hybrid precoding · Massive MIMO · Block diagonalization · Multi-user · Manifold optimization

1 Introduction

Hybrid precoding (HP) design is recently recognized as a key technology in millimeter-wave communication systems. To improve spectrum efficiency (SE) and energy efficiency, many hybrid precoding algorithms have been studied, including orthogonal matching pursuit (OMP) algorithm [1], manifold optimization (MO) algorithm [2] and the Dinkelbach method and weighted minimum mean square error (WMMSE) [3].

However, the high computational complexity and the realization of the optimal solution are the main problems of the above technologies. In [4], a low complexity HBF algorithm based on maximum minimum phase difference was proposed to improve sum rate. A hierarchical-agglomerative-clustering (HAC) based fully-connected hybrid precoding (FHP) scheme to suppress inter-user interference and improve SE is proposed in [5]. Reference [6] also proposes a low complexity grouping optimization based hybrid beamforming (HBF) algorithm to maximize sum rate. To optimize the analog precoder and digital precoder pairs, [7] proposes a novel criterion two-stage HBF design scheme for the MU-MIMO scenario. A novel virtual end-to-end mutual information (EEMI) assisted two-stage hybrid block diagonalization (HBD) scheme have been judiciously devised in [8] to balance the processing complexity and the design flexibility.

In this paper, we propose a novel two-stage approach to design hybrid multi user precoding. Unlike the traditional hybrid multi user precoding, the proposed scheme maximizes the sum of the SE of all users by maximizing the SE of each user. Simulation

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 18–26, 2023.
https://doi.org/10.1007/978-981-99-2653-4_3

results show that the performance of the proposed low-cost and low-computational complexity scheme is close to that of the full-complexity BD scheme in [9]. Furthermore, the proposed scheme is examined in mmWave channel and Rayleigh fading channel and compared with spatial sparse precoding/combining method [1] and low-complexity PE-AltMin algorithm [2], which was originally proposed for SU-MIMO but extended to MU-MIMO in this paper.

2 System Model

2.1 System Model

In this study, we consider a downlink of multi-user massive MIMO system with hybrid transceiver structure shown in Fig. 1. Similar hybrid transceiver structure has been studied in [7]. It is assumed that a base station with N_b antennas and M_b RF chains simultaneously communicates with K user terminals in a cell, and each user has N_u antennas and M_u RF chains to support N_s data streams. To ensure the effectiveness of the large-scale MIMO communication, the transmitter uses its $KN_s \leq M_b \ll N_b$ RF chains and the receiver uses its $N_s \leq M_u \leq N_u$ RF chains for each user.

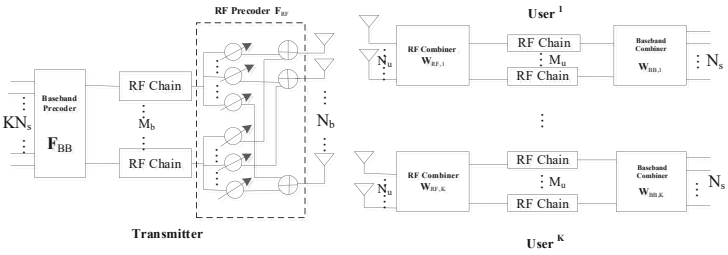


Fig. 1. Hybrid transceiver structure of multi-user Massive MIMO system

After the process of combining, the received signal of the k-th user is given as:

$$y_k = W_{BB,k}^H W_{RF,k}^H H_k F_{RF} F_{BB} \sqrt{P} s + W_{BB,k}^H W_{RF,k}^H n_k \tag{1}$$

where $s = \left[s_1^T, \ldots, s_K^T\right]^T$ is the $KN_s \times 1$ symbol vector for total transmitted symbols of K user, satisfying $E\left\{ss^H\right\} = \frac{1}{KN_s} I_{KN_s}$, P denotes the average transmitting power, $F_{BB} = \left[F_{BB,1}, F_{BB,2}, \ldots, F_{BB,K}\right]$ is the $M_b \times KN_s$ digital precoder, F_{RF} is the $N_b \times M_b$ analog precoder, and its entries have the same amplitude. To satisfy the total transmitted power constraint, F_{BB} is constrained to $\|F_{RF} F_{BB}\|_F^2 = KN_s$ where $\| \bullet \|_F$ is the Frobenius norm. $W_{RF,k}$ is the $N_u \times M_u$ analog combiner whose entries have the same amplitude and $W_{BB,k}$ is the $M_u \times N_s$ digital combiner for the k-th user. H_k is the channel matrix between the base station and the k-th user. n_k is the additive Gaussian white noise vector satisfying the independent and identically distributed complex Gaussian distribution with zero mean and variance σ_k^2 at k-th user.

Assume that when the base station transmits Gaussian signal, the achievable sum rate is expressed as:

$$R_{sum} = \sum_{k=1}^{K} log_2(1 + SINR_k) \tag{2}$$

where $SINR_k$ represents the signal-to-interference-noise ratio at the k-th user, which can be written as:

$$SINR_k = \frac{\frac{P}{KN_s}\left|W_{BB,k}^H W_{RF,k}^H H_k F_{RF} F_{BB,k}\right|^2}{\sum_{i\neq k}\frac{P}{KN_s}\left|W_{BB,k}^H W_{RF,k}^H H_k F_{RF} F_{BB,i}\right|^2 + \sigma_k^2 \|W_{BB,k}^H W_{RF,k}^H\|_2^2} \tag{3}$$

2.2 Channel Model

In this paper, without loss of generality, we discuss two kinds of fast fading channel models, i.e., Rayleigh fading channels and the mmWave channels.

The normalized channel matrix is H_k for the k-th user in Rayleigh fading channels, whose each entry follows i.i.d $CN(0,1)$. In the mmWave channels, we use the Saleh-Valenzuela model as in [1–3, 5], which can be described as:

$$H_k = \sqrt{\frac{N_b N_{u,k}}{N_{c,k} N_{p,k}}} \sum_{c=1}^{N_{c,k}} \sum_{p=1}^{N_{p,k}} \alpha_{c,p}^k \partial_u^k\left(\theta_{c,p}^k\right) \partial_b^k\left(\varnothing_{c,p}^k\right)^H \tag{4}$$

where $N_{c,k}$ is the number of scattering clusters and $N_{p,k}$ is the number of propagation path in each clusters at the k-th user, $N_{u,k}$ is the receiving antenna configured by the k-th user, $\alpha_{c,p}^k$ is the complex gain of the p-th path in the c-th scattering cluster for the k-th user, $\partial_u^k\left(\theta_{c,p}^k\right)$ and $\partial_b^k\left(\varnothing_{c,p}^k\right)^H$ denote the receive and transmit array response vectors at the azimuth angles of $\theta_{c,p}^k$ and $\varnothing_{c,p}^k$ respectively. When the transmitter and receiver antennas configure uniform linear arrays (ULAs) with N-element, the array response vector can be defined as:

$$\partial_{ULA}(\theta) = \frac{1}{\sqrt{N}}\left[1, e^{-j\frac{2\pi d}{\lambda}\sin\theta}, \ldots, e^{-j(N-1)\frac{2\pi d}{\lambda}\sin\theta}\right]^T \tag{5}$$

where λ is the carrier wavelength and d is the antenna spacing, and we will use this channel model in simulations.

2.3 Problem Formulation

To maximize the system sum-rate, it is imperative to find the precoder for the BS and combiners for each MSs in a massive MU-MIMO system with hybrid structure. The optimization model can be formulated as:

$$\left(F_{RF}, \left(F_{BB,k}, W_{RF,k}, W_{BB,k}\right)_{k=1,2,\ldots,K}\right) = \max_{(F_{RF}, F_{BB}, W_{RF}, W_{BB})} R_{sum} \tag{6-a}$$

$$s.t.|F_{RF}(i,j)| = \frac{1}{\sqrt{N_b}},\ \left|W_{RF,k}(i,j)\right| = \frac{1}{\sqrt{N_u}} \tag{6-b}$$

$$\sum\nolimits_{k=1}^{K} \left\|F_{RF}F_{BB,k}\right\|_F^2 = KN_s \tag{6-c}$$

3 The Proposed Hybrid Multiuser Precoding Design Scheme

Since both the objective function and constraints in the problem (6) are nonconvex, and it is difficult to find global optima. In this paper, we introduce a two-stage hybrid multi-user precoder algorithm. Firstly, the BD algorithm is used to eliminate user interference and obtain the optimal full-digital hybrid precoder. Then, a manifold optimization method is adopted to design the precoder and combiner for each user.

3.1 Block Diagonalization Algorithm

In this section, we first use BD algorithm to eliminate user-interference and decompose the system into K parallel single-user MIMO sub-systems. To obtain full digital optimal precoder at the k-th user, we define $\overline{\overline{H}}_k \triangleq \left[H_1^T, H_2^T, \ldots H_{k-1}^T, H_{k+1}^T, \ldots H_K^T\right]^T$, since the $F_{opt,k}$ should lie in the null space of $\overline{\overline{H}}_k$, , performing the singular value decomposition (SVD) of $\overline{\overline{H}}_k$, i.e.:

$$\overline{\overline{H}}_k = \overline{\overline{U_k \Sigma_k}}\left[\overline{V_k^{(1)}}, \overline{V_k^{(0)}}\right]^H \tag{7}$$

where $\overline{\overline{V_k^{(1)}}} = \overline{\overline{V_k}}(:, 1:(K-1)M_u)$ is the orthogonal basis of the subspace of $\overline{\overline{H}}_k$, and $\overline{\overline{V_k^{(0)}}} = \overline{\overline{V_k}}(:, (K-1)M_u + 1 : end)$ stands for the orthogonal basis of the null space of $\overline{\overline{H}}_k$, that is to say, if a signal is sent in the direction of $\overline{\overline{V_k^{(0)}}}$, other users except the k-th user cannot receive the signal at all, that is, if:

$$H_k \overline{\overline{V_i^{(0)}}} = 0, i \neq k \tag{8}$$

the inter-user interference of different users can be eliminated. The singular value decomposition of $H_k \overline{\overline{V_k^{(0)}}}$ is further conducted to eliminate the intra-stream interference of each user, so that the symbols of the target user can be transmitted in parallel, namely $H_k \overline{\overline{V_k^{(0)}}} = U \Sigma V^H$. Define $U^{(N_s)}$ and $V^{(N_s)}$ as the first N_s columns of U and V, respectively. Thus, the optimal full-digital precoding and combining matrix of the k-th user can be expressed as:

$$F_{opt,k} = \overline{\overline{V_k^{(0)}}} V^{(N_s)}, W_{opt,k} = U^{(N_s)} \tag{9}$$

3.2 Manifold Optimization Method Based on Hybrid Precoding

In this section, we solve hybrid precoder design problem for each user under the premise of fixing the optimal full-digital precoder. The problem of (6) for the k-th user can be reformulated as:

$$\min_{f_{RF,k},F_{BB,k}} \left\| F_{opt,k} - f_{RF,k}F_{BB,k} \right\|_F^2 \tag{10-a}$$

$$s.t. \left| f_{RF,k}(i,j) \right| = 1, \left\| f_{RF,k}F_{BB,k} \right\|_F^2 = N_s \tag{10-b}$$

according to the least squares structure of the objective function, the closed-form solution of the optimal $F_{BB,k}$ for the k-th user is given by $\widehat{F_{BB,k}} = f_{RF,k}^{\dagger}F_{opt,k}$. Substituting the optimal $\widehat{F_{BB,k}}$ into (10), the result can be rewritten as:

$$\min_{f_{RF,k}} \left\| F_{opt,k} - f_{RF,k}\widehat{F_{BB,k}} \right\|_F^2 \tag{11-a}$$

$$s.t \left| f_{RF,k}(i,j) \right| = 1 \tag{11-b}$$

Next, we utilize the manifold optimization (MO) method [2, 10] to obtain a local optimal $f_{RF,k}$ at the k-th user. The basic idea is to define a Riemannian manifold for $f_{RF,k}$ with the constant modulus constraint, then iteratively update the optimization variable on the direction of the Riemannian gradient (i.e., the Euclidean conjugate gradient is projected onto the tangent space of a point on the Riemannian manifold), which is similar to the traditional Euclidean gradient descent algorithm (please refer to [2] for specific details). However, the derivation part of the conjugate gradient in the Euclidean space is extremely complex in [2]. For simplicity, we according to the knowledge of partial derivation in Mathematics in this paper to calculate the Euclidean gradient [10], which is expressed as:

$$\nabla f(x) = vec\left(G\left(f_{RF,k}\right)\right) = vec\left(-2\left(F_{opt,k} - f_{RF,k}\widehat{F_{BB,k}}\right)\widehat{F_{BB,k}}^{H}\right) \tag{12}$$

With this understanding, the MO method can be used to solve the problem with the constant modulus constraints. The same procedures can also be applied for the decoder. Lastly, we summarize the entire algorithm procedures in Algorithm 1.

Algorithm 1 The proposed HP design scheme for massive MU-MIMO system

The first stage:
input: H
  for k=1 to K **do**
   Compute $F_{opt,k}$ and $W_{opt,k}$ by using BD algorithm
  end for
End stage
The second stage:
  for k=1 to K **do**
   input $F_{opt,k}$, initialize $f_{RF,k}^{0}$ randomly and set n=0
    repeat
      Compute $\widehat{F_{BB,k}^{n}} = \left(f_{RF,k}^{n}\right)^{\dagger} F_{opt,k}$
      Compute the Euclidean gradient $\nabla f\left(f_{RF,k}^{n}\right)$ using (12)
      Compute $f_{RF,k}^{n+1}$ using the manifold optimization method
      $n \leftarrow n+1$
    until a stopping condition is satisfied
    output $f_{RF,k}$ and $F_{BB,k} = \frac{\sqrt{N_s}}{\left\| f_{RF,k} \widehat{F_{BB,k}^{n}} \right\|_F} \widehat{F_{BB,k}^{n}}$
  end for
End stage
Get $W_{RF,k}$, $W_{BB,k}$, $k \in \{1, \ldots, K\}$ by the same procedure as the second stage except $W_{BB,k}$ does not require satisfying the power constraint
Obtain $F_{RF}, \left(F_{BB,k}, W_{RF,k}, W_{BB,k}\right)_{k=1:K}$

3.3 Complexity Analysis

In this section, we analyze and compare the complexity of several approaches. Obviously, as shown in Table1, the processing speed of the proposed algorithm is comparable to or much faster than the state-of-the-art algorithms, while its performance is close to optimal.

Table 1. Complexity comparison

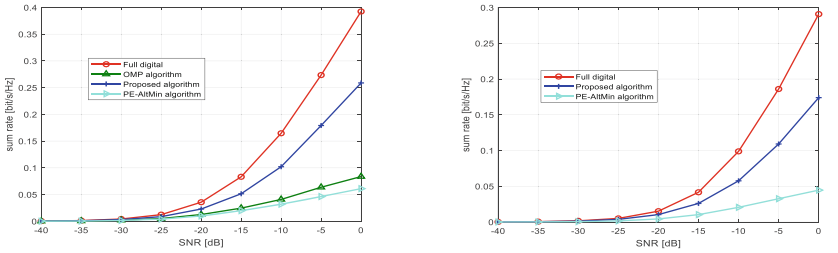
Algorithm	Complexity ($N_b \gg M_b, M_b \geq KN_s$)
The proposed Algorithm	$O(M_b N_b K N_s)$
PE-AltMin Algorithm [2, Section IV]	$O\left(min\left\{KN_s^2 M_b, KN_s (M_b)^2\right\}\right)$
OMP Algorithm [1]	$O\left((N_b)^3 M_b K\right)$

4 Simulation Results

In this section, to prove the effectiveness of the proposed HP design scheme, we evaluate the sum rate of different hybrid multi-user precoding schemes in different configurations. In addition, Rayleigh fading channel and mmWave channel are considered in this paper. The channel configuration in all simulations is the same as that in [2], and the range of transmission SNR in simulation is set as -40dB ~ 0dB.

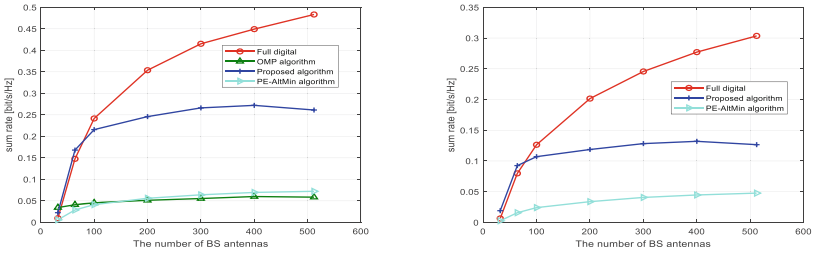
Figure 2 shows the sum-rate achieved by the proposed HP scheme, BD-OMP [1] and BD-PE-AltMin [2] in both the Rayleigh and mmWave channels, where the BD-OMP scheme is only discussed in millimeter wave channel. Note that the traditional high-dimensional full-digital BD precoding scheme is drawn as the upper bound of the sum rate. We can observe from both Fig. 2(a) and Fig. 2(b) that the proposed HP design scheme is superior to BD-OMP and BD-PE-AltMin scheme. Meanwhile, the proposed HP design scheme has lower computational complexity. Therefore, in Rayleigh fading and millimeter-wave channels, the proposed HP design scheme is the optimal hybrid multiuser precoding design scheme.

Figure 3 further evaluates the sum-rate with the number of BS antennas for different HP schemes in both the Rayleigh fading channels and mmWave channels. We assume that the number of base station antennas ranges from 32 to 512, and the $SNR = 0dB$. It can be seen from the figure that whether in Rayleigh fading channel or millimeter wave channel, when the number of BS antennas increases, the performance of the proposed HP scheme is closer to that of the full-digital optimal scheme.



(a) Performance in mmWave channels (b) Performance in Rayleigh channels

Fig. 2. Sum-rate versus SNR in a 256×64 8-user MU-MIMO system where $N_s = 4$, $M_b = 32$, $M_u = 4$ for different schemes.



(a) Performance in mmWave channels (b) Performance in Rayleigh channels

Fig. 3. Sum-rate versus N_b in a 8-user MU-MIMO system where $N_s = 4$, $N_u = 64$, $M_b = 32$, $M_u = 4$, $SNR = 0dB$.

5 Conclusion

In this paper, we propose a novel two stage hybrid precoding scheme for multiuser massive MIMO system. Unlike the traditional hybrid multi-user precoding scheme, the proposed hybrid precoding scheme aims to maximize the sum rate of each single-user system in order to maximize the sum of the sum rate of all users. Simulation results show that the proposed hybrid multi-user precoding scheme outperforms the state of the art for HBF systems in the literature and significantly improves achievable rate in Rayleigh fading channel and mmWave channel.

Acknowledgements. This research was supported in part by the Scientific Research Plan Project of the Committee of Education in Tianjin (No. JW1708).

References

1. Ayach, O.E., Rajagopal, S., Abu-Surra, S., Pi, Z., Heath, R.W.: Spatially sparse precoding in millimeter wave MIMO systems. IEEE Trans. Wireless Commun. **13**(3), 1499–1513 (2014)
2. Yu, X., Shen, J., Zhang, J., Letaief, K.B.: Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems. IEEE J. Sel. Top. Signal Process. **10**(3), 485–500 (2016)
3. Jian, R., Chen, Y., Liu, Z., et al.: Energy efficiency optimization for multi-user multi-stream massive MIMO hybrid precoding. Int. J. Wireless Inf. Networks **28**(3), 319–331 (2021)
4. Ding, Y., Hu, A.: Maximizing minimum phase difference based hybrid beamforming for multiuser mmWave massive MIMO systems. In: 2018 IEEE 4th International Conference on Computer and Communications (ICCC), IEEE 403–408 (2018)
5. Zhang, L., Gui, L., Ying, K., Qin, Q.: Clustering based hybrid precoding design for multi-user massive MIMO systems. IEEE Transactions on Vehicular Technology, 68 (12), 12164–12178 (2019)
6. Ding, Y., Hu, A.: Grouping optimization based hybrid beamforming for multiuser mmWave massive MIMO systems. 2019 IEEE 2nd International Conference on Computer and Communication Engineering Technology (CCET), IEEE 203–207 (2019)
7. Wu, X., Liu, D., Yin, F.: Hybrid beamforming for multi-user massive MIMO systems. IEEE Trans. Commun. **66**(9), 3879–3891 (2018)

8. Gao, S., Cheng, X., Yang, L., Hybrid multi-user precoding for mmWave massive MIMO in frequency-selective channels.: IEEE Wireless Communications and Networking Conference (WCNC). IEEE **2020**, 1–6 (2020)
9. Spencer, Q.H., Swindlehurst, A.L., Haardt, M.: Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. IEEE Trans. Signal Process. **52**(2), 461–471 (2004)
10. Kasai, H.: Fast optimization algorithm on complex oblique manifold for hybrid precoding in millimeter wave MIMO systems. In: 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE 1266–1270 (2018)

Diffeomorphic Image Registration Based on Medical Image

Zhao Zhang, Baoju Zhang(✉), Hongwei Liu, and Bo Zhang

Tianjin Normal University, TianJin, China
wdxyzbj@163.com

Abstract. Medical image registration plays an important role in image surgical guidance, image information fusion, disease diagnosis and other clinical tasks. It can correct images of different periods, different modes and different imaging principles through certain spatial transformations to help doctors better understand patients' physical conditions. In the past, image registration methods based on deep learning often ignore the conversion between images and can not avoid the overlap of deformation fields, that is, the differential diffeomorphism can not be guaranteed. Therefore, this paper proposes a differential diffeomorphic image registration method based on medical images, which can ensure the topological preservation and reversibility of the transformation. We use brain MRI image data set for training and testing, experimental results show that our method has advanced registration accuracy, has good application value.

Keywords: Image registration · medical images · Variability models · Deep networks · diffeomorphism

1 Introduction

Image registration is a process in which a floating image is mapped to a fixed image through one or a series of spatial transformations, and the corresponding positions in space correspond to each other. Image registration algorithm optimization is to find a suitable spatial transformation to maximize the similarity between two images.

Traditional image registration methods are varied and widely used in medical field. Its classification methods include classification based on image dimension, image mode, image region, image feature and transformation type [1].

With the development of deep learning, image registration method based on convolutional neural network shows its speed and accuracy, so it develops rapidly in the field of medical image registration [2]. According to the deep learning model, scholars divided medical image registration methods into three categories, including supervised, weakly supervised and unsupervised medical image registration methods. In supervised learning, real transformation parameters (i.e. deformation field) need to be obtained in advance as labels. In the training process of the network, the objective function is propagated back, so as to update the parameters of each layer. The methods of obtaining

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 27–34, 2023.
https://doi.org/10.1007/978-981-99-2653-4_4

transformation parameters mainly include obtaining by traditional methods and generating by model. Weak supervision adopts the similarity measure of tags as the objective function to update the weight of the network in reverse iteration [3]. Compared with traditional registration methods, the disadvantage of learn-based registration method is that it can only deal with small and low-complexity deformation. Although the effect of supervised learning and weakly supervised learning registration methods has made great progress compared with traditional registration methods, but it is difficult to obtain labels [4]. Therefore, unsupervised learning medical image registration is more popular among scholars.

In order to further improve the registration accuracy, this paper proposes a diffeomorphic image registration method based on medical images. We learn the registration function from 3d data sets and output a pair of diffeomorphic maps. A forward mapping from input image to output image can be obtained by synthesizing the inverse of the output map and another map. The results show that this method can ensure the topological preservation and reversibility of image transformation.

2 Learning-Based Registration Methods

The purpose of the registration of variable deformation is to find a nonlinear relation to align two images and perform a nonlinear transformation. We use A and B respectively to represent the fixed image and the floating image, and F represents the deformation field, so the deformable registration optimization process can be expressed as.

$$F^* = \arg_F L_{sim}\left(\mathrm{A},\mathrm{B}^*\right) + L_{reg}(F) \tag{1}$$

Where F^* represents the optimal deformation field, $L_{sim}(,)$ represents the similarity function, $L_{reg}()$ represents the smooth regularization function, and B^* represents the floating image after deformation.

In traditional deep learning registration methods, displacement field Y is usually parameterized, and there is a certain transformation relationship between displacement field Y and deformation field F. However, the disadvantage of this method is that it cannot deal with the registration problem with large deformation. In this paper, we add the velocity field V[5] to the differential homeomorphism transformation model. The relationship between velocity field and deformation field can be expressed as

$$\frac{d_{F_t}}{d_t} = v^t\left(F^t\right) \tag{2}$$

where v^t represents velocity field at time t, $t \subset (0, 1)$.

The concept of a group is an algebraic structure of a set plus an operation. Lie groups are groups with continuous properties. In our method, we express the deformation field as a Lie algebra, and time 1 represents the time of the whole registration process. According to the conceptual formula $R(t) = \exp(\phi^t)$ of Lie algebra, we can exponential it to the deformation field F^1 at time 1, i.e. $F^1 = \exp(\mathrm{v})$. From the conclusion of lie algebraic group theory, $F^{(1/2^{t-1})} = F^{(1/2^t)} \circ F^{(1/2^t)}$, Where $\circ$ represents the compound operator. We are given the initial deformation field $F^{(1/2^T)} = x + v(x)/2^T$, where, 2^T represents

the average amount of time 1 divided. In our experiment, we set the value of T to 7. So we can get $F^{1/2}$ by repetition, i.e. $F^{(1/2^6)} = F^{(1/2^7)} \circ F^{(1/2^7)} \ldots$ end up with $F^{(1/2)} = F^{(1/4)} \circ F^{(1/4)}$.

We define F^1_{PQ} as the deformation field mapped from image P to image Q. Similarly, F^1_{QP} is the deformation field mapped from image Q to image P. Therefore, the deformable registration problem is parameterized as equation $f(P, Q) = \left(F^1_{PQ}, F^1_{QP}\right)$. We train P and Q to the intermediate mapping deformation field $F^{0.5}_{PQ}$ and $F^{0.5}_{QP}$ respectively. After the model converges, we use the intermediate mapping deformation field to form F^1_{PQ} and F^1_{QP}.

3 Registration Network

3.1 Registration Network Structure

Our registration structure flow is shown in Fig. 1, which mainly includes a U-net, scaling and square layer, spatial transformer network. The image P and Q are input, and the speed field v_{PQ} and v_{QP} are output through the registration network. The speed field generates the intermediate mapping deformation field $F^{0.5}_{PQ}$ and $F^{0.5}_{PQ}$ through the scaling and square layer, and finally generates the final registration image $F^1_{PQ}(P)$ and $F^1_{QP}(Q)$ through the spatial transformer network.

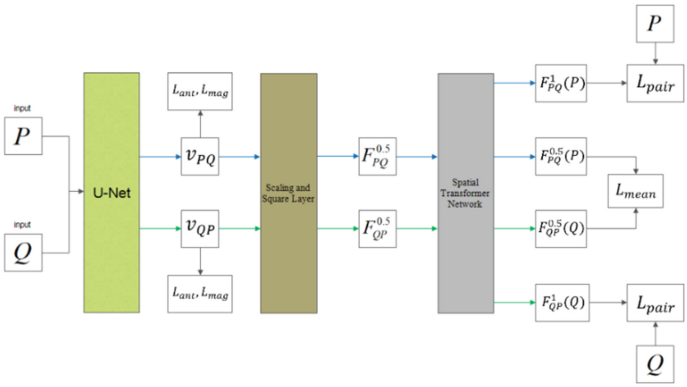


Fig. 1. Flow chart of registration structure

U-net is a five-layer encoder-decoder structure with jump connections. As shown in Fig. 2, in the encoder part, each step shall undergo convolution operation with convolution kernel size of 3*3*3, a stride of 1 and convolution kernel size of 2*2*2 and a stride of 2. The convolution operation with a stride of 2 is used as down-sampling. In the decoder part, each step goes through a splicing process of the corresponding feature graph of the encoder and decoder, and then goes through the convolution operation with the convolution kernel size of 3*3*3, a stride of 1 and the convolution kernel size of 3*3*3,

a stride of 2. The convolution operation with a stride of 2 is used as up-sampling. After each convolution operation described previously, a LeakyRelu activation function will be used, where the coefficient of the LeakyRelu activation function a is 0.2. In the last two layers of decoder, we use convolution operation with convolution kernel size of 5*5*5 and step size of 1, and use soft sign activation function to generate speed field. Finally, we multiply the output speed field by a constant c, In order to limit the range of speed field. In actual operation, our value c is 100.

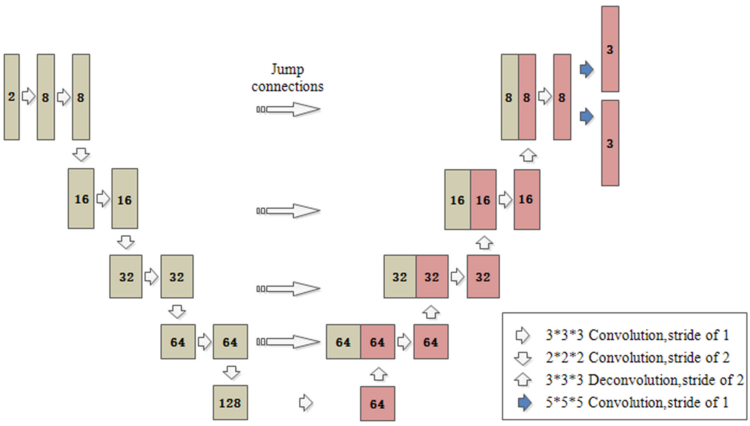


Fig. 2. Schematic diagram of registration network structure

3.2 Loss Function

In the field of image registration, there are many criteria for similarity measurement. In this paper, NCC(Normalized Cross Correlation) is used to calculate symmetric similarity loss term. P and Q represent the two input images. $\overline{P}(x)$ and $\overline{Q}(x)$ respectively represent the local mean values of P and Q at position x. The local mean values are calculated by taking x as the center and window size as ω^3,where ω is set to 7. NCC is defined as follows:

$$NCC(P,Q) = \sum_{x} \frac{\sum_{x_i}\left(P(x_i)-\overline{P}(x)\right)\left(Q(x_i)-\overline{Q}(x)\right)}{\sqrt{\sum_{x_i}\left(P(x_i)-\overline{P}(x)\right)^2 \sum\left(Q(x_i)-\overline{Q}(x)\right)^2}} \tag{3}$$

where x^i represents the position in the window centered on x.

We define L_{mean} as the similarity loss of the intermediate mapping, which is then used to measure the similarity between ' P and Q' and 'intermediate mapping $F_{PQ}^{0.5}(P)$ and $F_{QP}^{0.5}(Q)$'. L_{pair} is defined as the similarity loss of real transformation mapping, which is used to measure the similarity between ' P and Q' and ' $F_{PQ}^{1}(P)$ and $F_{QP}^{1}(Q)$ after deformation'. Symmetric similarity loss can be expressed as $L_{sim} = L_{mean} + L_{pair}$, where

$$L_{mean} = -\,\mathrm{NCC}\left(P\left(F_{PQ}^{0.5}(P)\right), Q\left(F_{QP}^{0.5}(P)\right)\right) \tag{4}$$

$$L_{pair} = - \mathrm{NCC}\left(P\left(F_{PQ}^{1}(P), Q\right)\right) - \mathrm{NCC}\left(Q\left(F_{QP}^{1}(Q), P\right)\right) \tag{5}$$

Global regularization will reduce the registration accuracy to a certain extent, but also can not maintain the topological preservation of the transformation, so this paper proposes to add Jacobian determinant regularization on the basis of global regularization to further optimize the local direction consistency. The Jacobian regular loss can be expressed as

$$L_{J\,\mathrm{det}} = \frac{1}{N} \sum \max(0, -|J_F(z)|) \tag{6}$$

where N represents the total number of elements in J_F;$|J_F()|$ represents the Jacobian determinant of the deformation field F at position z.

In addition to the local directional consistency loss, we also retain the global gradient regularization loss in the traditional method, which is common to enhance the smoothness of the velocity field. The global gradient regularization loss can be expressed as

$$L_{reg} = \sum_{z} (\left\| \nabla F_{PQ}(s) \right\|_2^2 + \left\| \nabla F_{QP}(s) \right\|_2^2) \tag{7}$$

where,$\nabla F_{PQ}(s)$ represents the gradient at voxel s in deformation field F_{PQ}.

We define $L_{mag} = \frac{1}{N}\left(\left\| v_{PQ} \right\|_2^2 - \left\| v_{QP} \right\|_2^2 \right)$ as the constraint function of the magnitude of deformation field, mainly to ensure that the predicted velocity fields v_{PQ} and v_{QP} will not have low registration accuracy due to the large difference of magnitude.

We define

$$L_{ant} = \sum_{h} (\delta\left(\nabla F_{PQ}(h) + 1\right)\left|\nabla F_{PQ}(h)\right|^2 + \delta\left(\nabla F_{QP}(h) + 1\right)\left|\nabla F_{QP}(h)\right|^2) \tag{8}$$

as anti-overlap constraint loss. Where, $\nabla F_{PQ}(h)$ represents the gradient at voxel h in deformation field F_{PQ}, and $\delta()$ represents the indicator function of overlapping positions in the punishment deformation field. The function of this loss is to reduce the overlap of deformation fields and ensure the topological preservation of deformation fields.

Therefore, the total loss function of this experiment is

$$L(X, Y) = L_{sim} + \lambda_1 L_{J\,\mathrm{det}} + \lambda_2 L_{mag} + \lambda_3 L_{ant} + \lambda_4 L_{reg} \tag{9}$$

4 Test and Result Analysis

This paper uses the Open OASIS data set, which is called Open Access Series of Imaging Studies. The dataset consisted of MRI brain images of 414 subjects from different populations. We used FREESURFER software to preprocess the dataset with skull removal, cerebellar removal, deviation correction, and radiological alignment, and cropped the dataset to size 160*192*224. In our experiment, we used 300 images as the training set, 20 images as the verification set, and the rest as the test set.

Adam optimizer was selected for this training, the initial learning rate was set as $1e^{-5}$, the parameter of loss function was set as $\lambda_1 = 1$, $\lambda_2 = 0.001$, $\lambda_3 = 1000$, $\lambda_4 = 3.5$, and the training times were set as 20,000 times. The experiment proved that the model had the best effect when the number of traversal times of training set reached 10,000 times.

Dice similarity coefficient (DSC) [6] is used as the index to measure spatial folding similarity of images in this paper, and its definition is $d = \frac{2|F \cap F^1|}{|F|+|F^1|}$. In our experiment, the value range of Dice is (0, 1), and the larger the value of Dice is, the better our registration effect is.

Our method is compared with some of the most advanced deep learning registration methods such as SYMNET, ICNET and FAIM. Experimental results show that the average DSC similarity coefficient of our algorithm reaches 0.659, which is improved by one to two percentage points compared with other algorithms. Therefore, the registration effect will be more prominent in practical operation. Specific experimental data are shown in the Table 1.

Table 1. Compared with SYMNET, ICNET and FAIM data

Method	Our method	SYMNET	ICNET	FAIM
Avg.DSC	0.659	0.648	0.636	0.640

Table 2. The impact of different λ_1 on data

λ_1	0	1	10	100	1000
Avg.DSC	0.654	0.659	0.654	0.653	0.652

We also conducted several experiments according to different Values of Jacobian determinant coefficient λ_1, and finally determined that when the value of λ_1 is 1, the registration effect is the best.

Figures 3, 4, 5 and 6 are the registration effect pictures of this experiment. Figures 3 and 4 are the input images, Fig. 5 is the registration effect picture from image 3 to image 4, and Fig. 6 is the registration effect picture from image 4 to image 3.

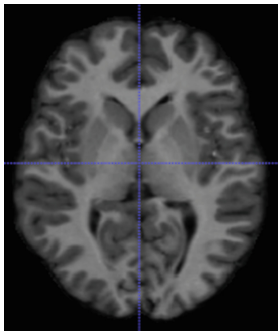


Fig. 3. Input P

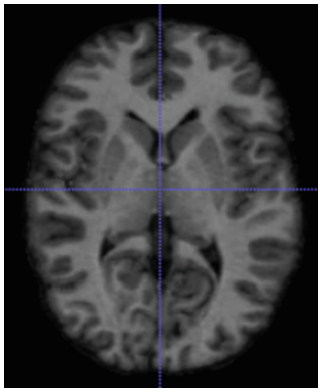


Fig. 4. Input Q

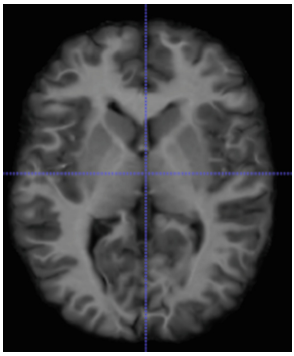


Fig. 5. P to Q registration graph

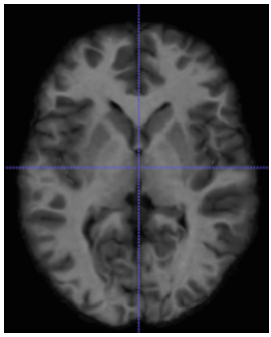


Fig. 6. Q to P registration graph

5 Conclusion

In this paper, a method of diffeomorphic image registration based on medical images is introduced. The method completes the mapping from image to registration image by learning the diffeomorphic deformation field. We also introduce jacobian regularization loss as local consistency loss and verify the effect of different coefficients. We evaluated our model using a large number of brain MRI images from the OASIS dataset and compared it with three advanced registration methods. Experimental results show that the proposed method has better registration accuracy and diffeomorphism than other algorithms.

References

1. Ma, L., Luo, F.: Advances in deep medical image registration: Towards unsupervised learning. J. Image Graph. (2021)
2. Chee, E., Wu, Z.: AIRNet: Self-Supervised Affine Registration for 3D Medical Images using Neural Networks. Biomedical Image Registration (2018)

3. Zhou, Q., Wang, Y.: Research progress of medical image registration based on deep learning. J. Univ. Shanghai Sci. Technol. (2021)
4. Huang, P., Guo, C., Chen, H., Zhang, H.: A review of image registration methods based on deep learning. J. Hangzhou Dianzi University (Natural Science Edition) (2020)
5. Mok, T., Chung, A.: Fast Symmetric Diffeomorphic Image Registration with Convolutional Neural Networks. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (2020)
6. Zhao, A., Guttag, J.: VoxelMorph: A Learning Framework for Deformable Medical Image Registration. IEEE Transactions on Medical Imaging (2019)

Neural Network for UWB Radar Sensor Network-Based Sense-Through-Wall Human Detection

Dheeral Bhole(✉), Qilian Liang, and Chenyun Pan

Department of Electrical Engineering, University of Texas at Arlington, Arlington, Tarrant, TX 76019, USA

dheeral.bhole@mavs.uta.edu, {liang,chenyun.pan}@uta.edu

Abstract. Human detection through wall is a very critical and challenging task. In this paper, we present a neural network-based method using Ultra-wideband radar to detect a human through wooden door, gypsum wall and brick wall. We consider the tough problem of stationary human detection through wall and door. We employ various machine learning models for this task and achieve an excellent average accuracy of **92.67%** with neural network-based method.

Keywords: Machine learning · ultra-wideband · support vector machines · neural networks · random forest · k-nearest neighbors

1 Introduction

Human detection through walls and doors is critical in applications such as hostage rescue situation, surveillance, activity recognition, etc. Such a crucial task also comes with a lot of diverse engineering problems. Good penetration and range resolution have enabled Ultra-wideband (UWB) radars to become a primary choice for target detection in many outdoor and indoor scenarios [1, 2]. Large bandwidth of UWB signals enable it to achieve high range resolution and detect multiple nearby objects [3]. An efficient configuration of Wireless Sensor Network is also important for higher energy efficiency and performance [4].

Machine Learning (ML) has been utilized in various fields like natural language processing, 5G communications, Internet of Things (IoT), reinforcement learning, social networks, economics, and medical sciences [5]-[8]. In this paper,

we use WSN and ML to solve a tougher problem of stationary human detection through wall as compared to the previous work which considered human detection through wall based on movement [9]. With movement the radar receives different echoes for different timestamps, making detection easier. The success of detection depends on the ability of UWB signals to penetrate through building materials. Different materials have different dielectric properties thus offering varied challenging scenarios. We have considered typically occurring building structures like gypsum wall, brick wall and wooden door.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 35–43, 2023.
https://doi.org/10.1007/978-981-99-2653-4_5

The remainder of this paper is structured as follows. Data collection process is presented in Sect. 2. In Sect. 3, dataset preparation and modeling are put forward. In Sect. 4, we introduce results and conclude this paper in Sect. 5.

2 Data Collection

The human detection through wall measurement task was carried out at the University of Texas at Arlington and the details can be found in [9]. 3 different locations were used to sense human through brick wall, wooden door, and gypsum wall. P220 UWB radar antenna was used for this experiment. It is operated in monostatic mode where pulses are transmitted from a single omni-directional antenna. Scattered pulse waves are received by a collocated omni-directional antenna [3]. The radar parameters in each of the cases are Integration: Hardware Integration = 512, Software Integration = 2, Pulse Repetition Frequency: 9.6 MHz Step Size: 13 bin, Window Size (ft): 10 ft.

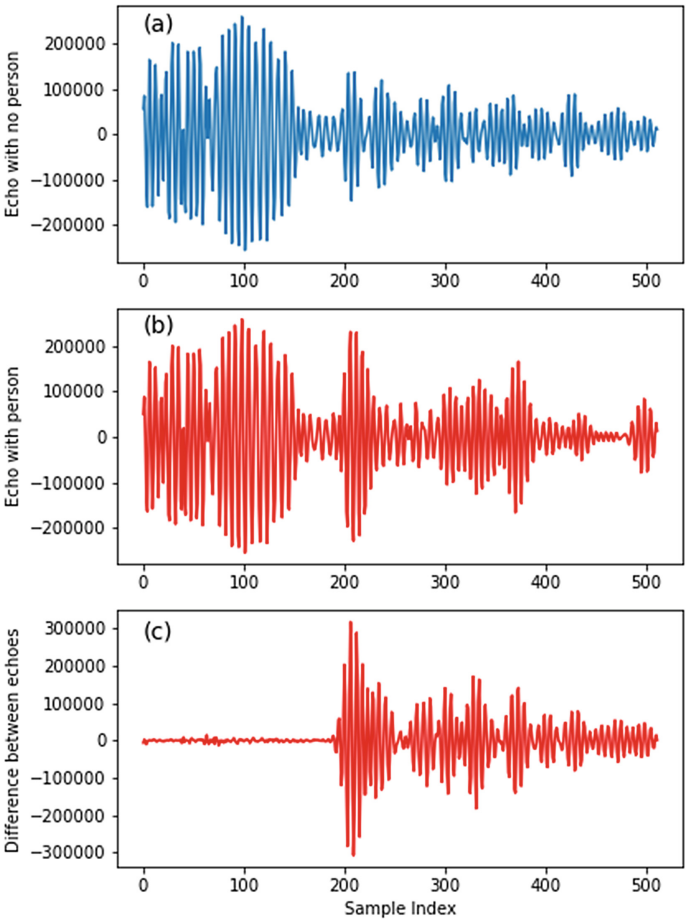


Fig. 1. Reflected echoes for two cases with: (a) no person (b) person and their difference in (c)

Figure 1 displays the reflected echoes with (a) no person and (b) person through gypsum wall. Observing Fig. 1a or Fig. 1b, it is tough to decide whether a target is present or not. The reflected echo in Fig. 1a can be considered as response from the surroundings since there is no person in range. So, the difference in Fig. 1a and Fig. 1b is the response from human detected through wall which is shown in 1c.

3 Dataset Preparation and Modeling

Preparing a good, clean, noiseless, outlier free, model friendly dataset is at the heart of the ML process. The dataset we handled had inconsistent received echoes with different number of samples per signal for multiple locations (gypsum wall, brick wall & wooden door) with person and no person cases. We used 512 samples for each signal (see Fig. 1) to have uniform amount of data for all cases. So, each signal had 512 samples and 766 such signals were recorded for both person and no person case repeated over 3 different locations. The entire dataset had $3*2*766 = 4596$ signals with 512 samples each.

Figure 2 shows the flow of data. Received UWB data is reshaped as explained above, then sent to the training and testing phase. We used 80–20 split for training and testing on the entire data. The model classifies new data in the testing phase and gives final predictions.

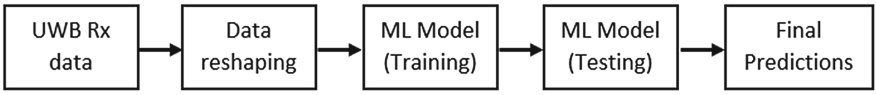


Fig. 2. Data pipeline

The received waveforms are time domain 1-D data. After exploratory data analysis, we conclude that this dataset is not enough for deep learning models. So, we employ algorithms like k-nearest neighbors (KNN), support vector machines (SVM), random forest (RF) and shallow neural networks (NN) which work very well with less data.

KNN is a simple non-parametric supervised learning algorithm where the output decision for a test sample is based on its k neighbors. K in KNN is based on feature similarity and finding the correct value of k is called parameter tuning. Smaller k value can produce noisy output and have larger influence on the result. Larger k value will be computationally expensive and might result in lower variance but increased bias. We tried 5, 25, 50, 100, 200 and 500 number of neighbors. The performance is best for 200 neighbors which are shown in Sect. 4.

SVM is a popular classification method used in applications where less data is available. SVM solves the problem of maximizing the distance or margin separating two classes in input space. The problem turns out to be convex and any local solution is also a global solution [10]. We used different kernel methods such as linear, radial basis function (RBF), polynomial, sigmoid and got the best performance for RBF for all cases which are shown in Sect. 4.

RF is an ML method consisting of multiple decision trees. Each tree receives different subset of data obtained by bootstrapping and random feature selection thus creating an

uncorrelated forest of trees. This leads to a more accurate ensemble than any of the individual tree. Typically, the more the number of trees in the forest the better the performance of the ensemble. We tried 5, 25, 50, 100, 200, 500, 750 and 1000 number of trees. The performance plateaus after 500 trees and is the optimal choice for all cases which are shown in Sect. 4.

NN are a set of ML models which allow computers to learn from experience and understand the world in terms of a hierarchy of concepts, with each concept defined through its relation to simpler concepts. NN are made up of neurons arranged as an input layer, one or more hidden layers, and an output layer. The artificial neurons from one layer are connected to the next layer with an associated weight and threshold. The node is activated if its output is above the specified threshold. We experimented with a lot of NN configurations with 1, 2 and 3 hidden layers with 50, 100 and 128 neurons in each layer. We got the best performance for NN with 1 hidden layer having 50 neurons in the hidden layer for all cases which are shown in Sect. 4.

4 Results

In this section, we conduct a comparative analysis of the 4 ML model's performance for different test conditions. The considered performance metrics are area under the curve (AUC), accuracy and testing time. For consistent performance, we conducted trials 10 times and averaged accuracy, and testing time for each model for all scenarios.

4.1 AUC

AUC is an important performance evaluation metric used to analyze a classifier. ROC is a probability curve and AUC is the degree of separability. The model is better able to classify if its AUC is higher.

4.1.1 Wooden Door Scenario

The ROC curve in Fig. 3 shows AUC metric for considered models for the wooden door scenario. False positive rate and false negative rate can be extracted from these curves. SVM has the best AUC of 0.99.

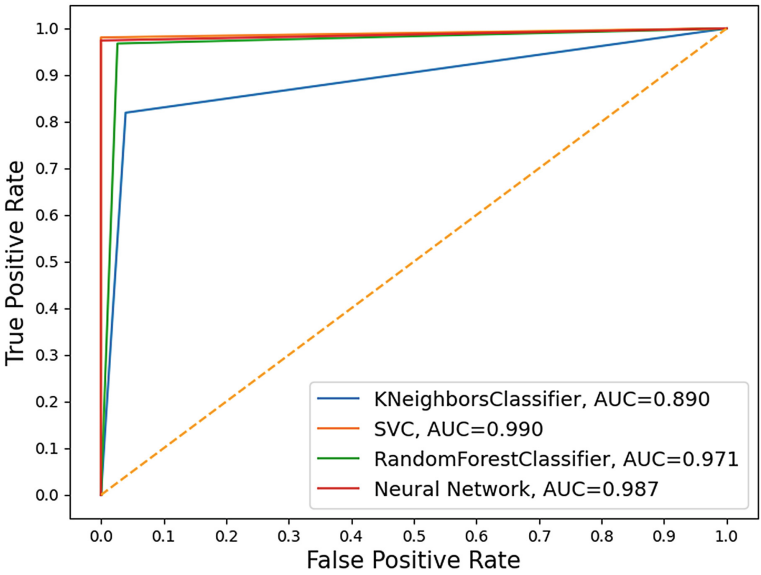


Fig. 3. Receiver Operating Characteristics (ROC) analysis for wooden door scenario

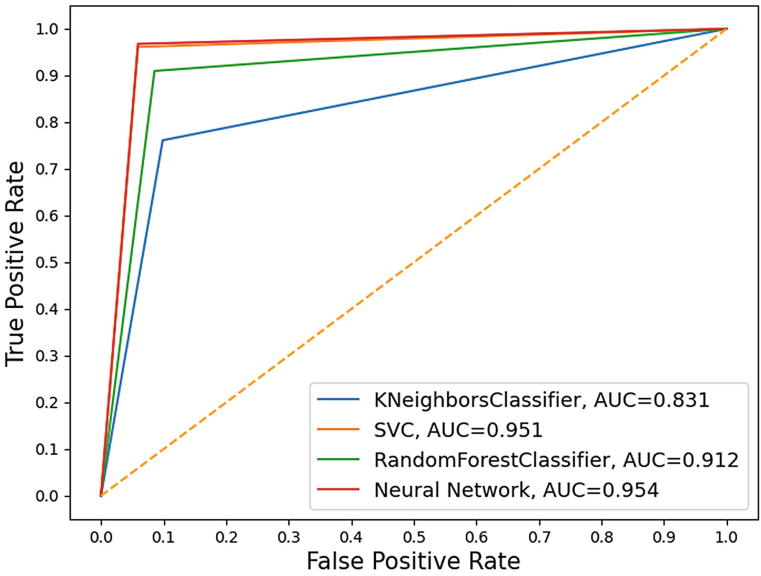


Fig. 4. Receiver Operating Characteristics (ROC) analysis for gypsum wall scenario

4.1.2 Gypsum Wall Scenario

The ROC curve in Fig. 4 shows AUC metric for considered models for the gypsum wall scenario. NN has the best AUC of 0.954.

4.1.3 Brick Wall Scenario

The ROC curve in Fig. 5 shows AUC metric for considered models for the brick wall scenario. SVM has the best AUC of 0.86.

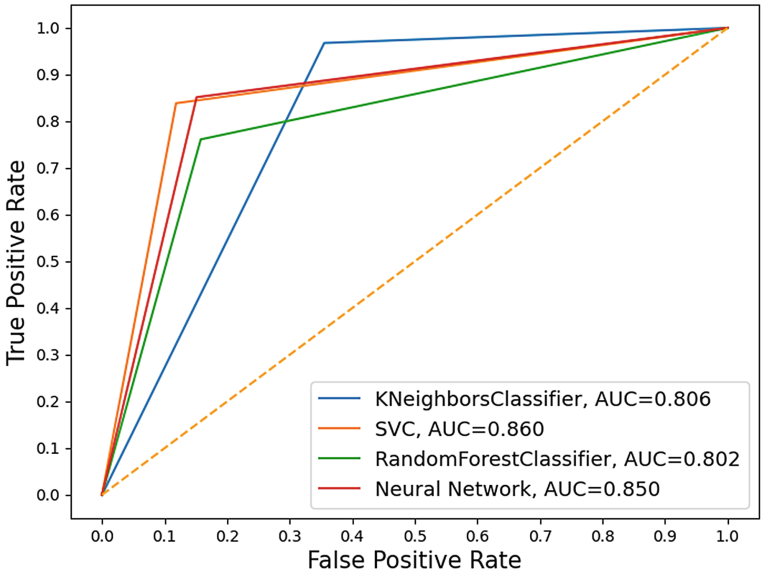


Fig. 5. Receiver Operating Characteristics (ROC) analysis for brick wall scenario

4.2 Accuracy

Accuracy is another important metric which encompasses performance for true positive, false positive, true negative and false negative detections. Figure 6 shows the accuracy performance of all models for all scenarios.

- For wooden door, SVM performs the best with **99%** accuracy. NN and RF also perform well while KNN performs the worst. As expected, the ML models find it easiest to detect human through wooden door as the door does not block any of the UWB signals.
- For gypsum wall, NN performs the best with **95%** accuracy. SVM and RF also perform well while KNN performs the worst. ML models find gypsum wall case easier than brick wall case but tougher than wooden door case. This performance arises from the fact that gypsum wall blocks UWB signal more than the wooden door but less than the brick wall.

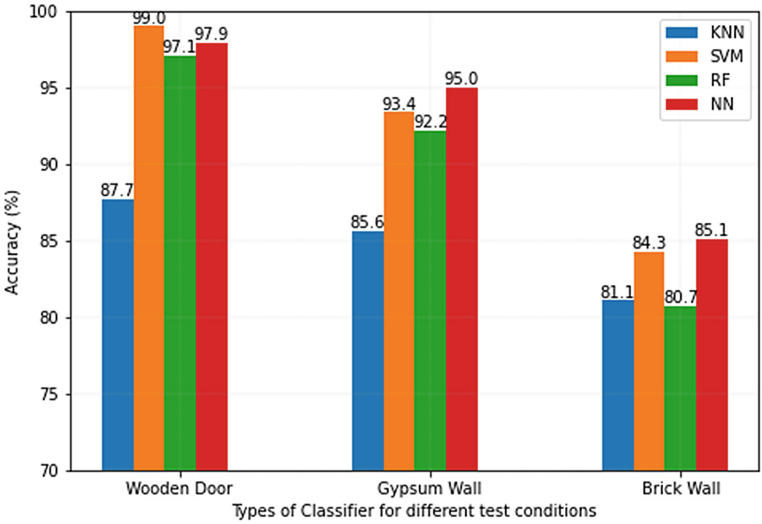


Fig. 6. Accuracy of ML models for different walls

- For brick wall, NN performs the best with **85.1%** accuracy. SVM also performs well while RF and KNN perform the worst. As expected, the ML models find it toughest to detect human through brick wall as the brick wall blocks the most UWB signal comparatively.

4.3 Testing Time

Testing time is a critical performance evalution metric denoting how much time a classifier takes to predict the testing dataset. This is an even more important metric because once a model is deployed, the performance of an application depends on the testing time it takes to make new predictions. We measured the total testing time for a model and divided it by the number of samples used for testing, giving average testing time per data sample (μs). In this paper for all scenarios as seen in Fig. 7, KNN has the least testing time requirement. Comparatively, SVM, RF and NN have higher testing time requirement.

Overall NN performed the best with a combined accuracy of **92.67%** for all scenarios. Neural network have a lower testing time per sample requirement for similar performing models. It would become an even better choice if more data is available for training, provided the model does not overfit the data.

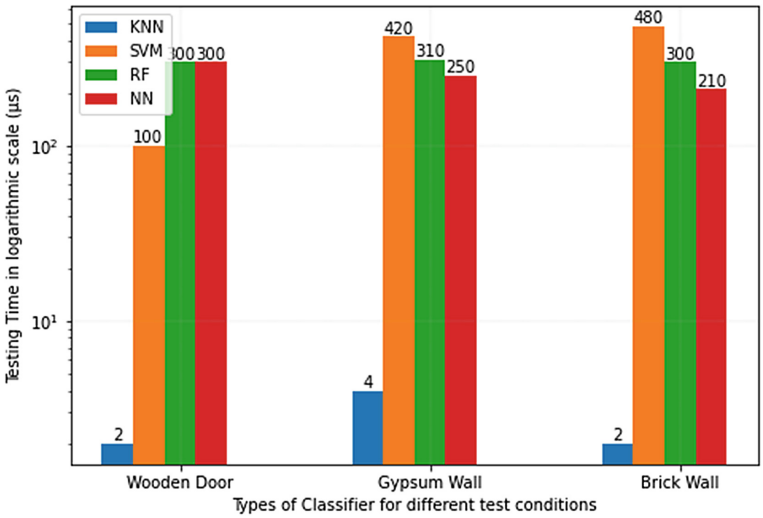


Fig. 7. Testing time per data sample

5 Conclusion

In this paper, we present a neural network-based method for human detection through wall using UWB radar. ML models like k-nearest neighbors, support vector machines, random forest and neural networks are considered for this task. Real world UWB radar data is used to train and test the system. Results show that neural network model achieved the best combined accuracy of **92.67%** for all scenarios. Also, accuracy is best for wooden door scenario then followed by gypsum wall and brick wall scenarios. Use of ML for supervised classification task is justified by excellent performance and very low computational needs.

Acknowledgement. This work was supported by U.S. National Science Foundation (NSF) under Grant CCF-2219753.

References

1. Ghavami, M., Michael, L., Kohno, R.: Ultra-wideband signals and systems in communication engineering. John Wiley & Sons (2007)
2. Bhole, D., Liang, Q.: Machine Learning Enabled Sense-Through-Foliage Target Detection Using UWB Radar Sensor Network. In: International Conference in Communications, Signal Processing, and Systems 2022 (pp. 1239–1246). Springer, Singapore https://doi.org/10.1109/ICC.2012.6364984
3. Oppermann, I., Hämäläinen, M., Iinatti, J.: UWB: theory and applications. John Wiley & Sons 15 Oct 2004
4. Di, C., Li, F., Li, S.: Sensor deployment for wireless sensor networks: A conjugate learning automata-based energy-efficient approach. IEEE Wirel. Commun. **27**(5), 80–87 (2020Oct 28)

5. Guo, H., Li, S., Li, B., Ma, Y., Ren, X.: A new learning automata-based pruning method to train deep neural networks. IEEE Internet Things J. **5**(5), 3263–3269 (2017Jun 6)
6. Yan, Z., Li, G., Tian, Y., Wu, J., Li, S., Chen, M.: Poor HV. Dehib: Deep hidden backdoor attack on semi-supervised learning via adversarial perturbation. InProceedings of the AAAI conference on artificial intelligence. 35(12), pp. 10585–10593) (2021)
7. Di, C., Liang, Q., Li, F., Li, S., Luo, F.: An efficient parameter-free learning automaton scheme. IEEE Transactions on Neural Networks and Learning Systems. **32**(11), 4849–4863 (2020Oct 5)
8. Zhao, Y., Li, S., Jin, F.: Identification of influential nodes in social networks with community structure based on label propagation. Neurocomputing **19**(210), 34–44 (2016Oct)
9. Singh, S., Liang, Q., Chen, D., Sheng, L.: Sense through wall human detection using UWB radar. EURASIP J. Wirel. Commun. Netw. **2011**(1), 1–1 (2011Dec)
10. Bishop, C.M., Nasrabadi, N.M.: Pattern recognition and machine learning. New York: springer; (2006)

Design and Implementation of Elevator Running with the Door Open Failure Monitoring System

Ruize Sun[1,2], Xinan Wang[1], Chuqiao Feng[2], Huancheng Xiao[2], Zhaorui Xu[2], and Jianmin Cao[2(✉)]

[1] School of Software and Microelectronics, Peking University, Beijing, China
[2] School of Big Data and Internet, Shenzhen Technology University, Shenzhen, China
caojianmin@sztu.edu.cn

Abstract. Elevator failures have continued to rise in recent years, and elevator running with the door open is a serious safety accident that requires prompt detection and intervention. This paper proposes a new elevator running with the door open monitoring system that extracts the original running data from the elevator CAN bus interface, sends it to the remote monitoring platform via NBIoT, and stores it in a private database. The basic states of the elevator could be obtained by analyzing the raw data and used to determine whether the elevator running with the door open and notifying the maintenance personnel in time. The monitoring system is not limited by elevator brand or model, which does not require additional sensors, and it only monitors without interfering with the original elevator system.

Keywords: Elevator failure · Running with the door open · CAN bus · Data analysis

1 Introduction

With the significant increase in elevator usage, higher and higher floors, and the growing time of elevator use, the phenomenon of elevator failure inevitably shows an increasing trend [1, 2]. However, the quality and control of elevator maintenance in China are still lacking effective means, and a large amount of data is still completed manually, which is far behind countries such as the United States and Japan [3]. Elevator safety management is still severe, and research on elevator monitoring and real-time failure alarm has become a community concern [4].

According to relevant data, the probability of elevator door system accidents is 80% of the total number of elevator accidents, indicating that door system-related failures are the most common type of elevator failure [5]. According to GB 7588-2003 "Safety Code for Elevator Manufacturing and Installation", if a floor door or any one of multiple

Sponsored by the Science and Technology Plan Project of Guangdong Administration for Market Regulation [2020ZT02], Scientific Research Capacity Improvement Project of Key Construction Disciplines in Guangdong Province of China [2021ZDJS109] and Industry University Cooperative Education Program of the Chinese Ministry of Education [202002321012].

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 44–51, 2023.
https://doi.org/10.1007/978-981-99-2653-4_6

floor doors is open, the elevator should not be started or kept running under normal condition [6]. However, elevator running with the door open failure still occurs from time to time. By analysis of previous failure causes, the majority of them are caused by short-circuiting of the electrical safety circuit, sometimes by artificial short-circuiting of the electrical safety circuit or the door lock circuit to assist in the repair of the elevator and it may also be caused by humid weather [5, 7, 8].

In the industry today, there are two primary methods for monitoring failures related to elevator running with the door open. One method is to install sensors on the elevator doors to determine the elevator's door status. A magnetron switch, for example, is used to detect the opening and closing status of the elevator door, and a human infrared sensor is added to detect whether someone is present when the elevator is running, and the combination of the magnetron switch's state and the human infrared sensor could determine whether the elevator is trapped [9]. However, this method is difficult to implement, costly, and prone to false alarms. Another method is to monitor elevator operation status using a "communication converter," but this method requires the elevator manufacturer to publish the elevator protocol, so it is not universal [10, 11].

This paper proposes a new monitoring system based on this, which includes the data collector, remote monitoring platform, and display interface (Web side, App client). The data collector collects raw elevator running data via the CAN bus interface and sends it to the remote monitoring platform via the NB-IoT module. It obtains the elevator's basic status through data analysis and makes logical decisions in order to notify elevator maintenance personnel via the App client when the elevator failure occurs. Moreover, the monitoring system includes a failure elimination alarm function. Since the system analyzes the elevator's original CAN bus data without installing sensors or converting the protocol, it is not limited by the elevator's brand or model, is not disturbed by sensor data errors or false alarms, and has a low cost, which are beneficial to the system's promotion.

2 The Monitoring System Design Program

The monitoring system includes the data collector, remote monitoring platform, and display interface. The data collector obtains the raw running data from the elevator's CAN bus interface, sends it to the remote monitoring platform via the NBIoT module, and stores it in our private database. The elevator's basic running status could be obtained by data analysis, and then the logical combination of the operating status is judged. Once it monitors the elevator running with the door open, it will promptly remind the maintenance personnel through the display interface and could eliminate the warning after the failure is lifted. The overall schematic diagram of the monitoring system is shown in Fig. 1.

Elevator floors are becoming progressively taller, and the connection between the elevator controller and the elevator car is typically made via the CAN bus interface with serial transmission. The data collector obtains running data from the CAN bus interface and sends it to the remote monitoring system. The following basic status of elevator could be obtained by analyzing raw CAN bus data.

Door signal: generally includes door closed signal, which means the elevator door is closed; door movement signal, which means the door is in the process of opening or

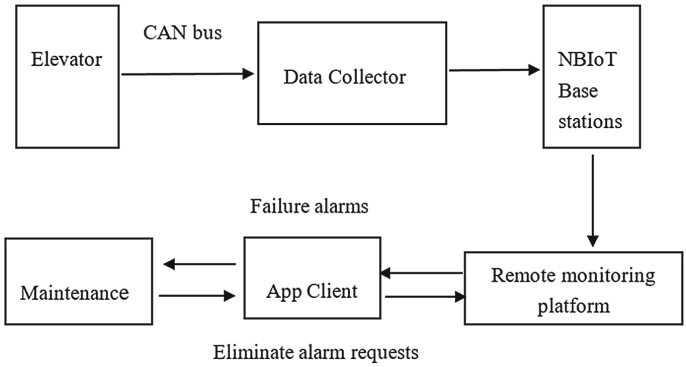


Fig. 1. Overall schematic diagram of the monitoring system

closing; door opened signal, which means the elevator door is in place and passengers can get on and off the elevator.

Floor signal: which indicates the number of floors served by the elevator and is generally judged by the elevator's built-in floor encoder.

Running signal: notifies the user when the elevator is running up or down. It is typically divided into upward and downward signals.

Leveling area signal: indicates that the elevator has arrived at the leveling area of a certain floor, as judged by the elevator's leveling sensor and need it all sensing generally. The elevator will only open the door normally at rest if the leveling area signal is received.

Car call signal: It indicates that passengers are in the elevator car to operate the elevator. Under normal circumstances, as long as there is a call signal, it means that someone is in the elevator.

To avoid impacting the original elevator system and increasing the fault nodes, photovoltaic isolation is used. Data could only be output from the elevator side in one direction and cannot be input. Thus, the data collector exhibits "monitoring but not controlling" characteristics.

The SN65HVD230 from Texas Instruments is used in the CAN module to receive CAN bus data from the elevator. The microcontroller is the STM32L433Rx. The NBIoT module BC35-G is powered by the Shanghai Yiyuan Communication Technology Co., Ltd. A power module and a backup power supply are also installed to ensure the data collector's normal operation.

The data collector is small in size, low in cost, and easy to install. Eelevator with one data collector (one node) could collect and transmit running data and thus meet the application requirements. The block diagram of the data collector is shown in Fig. 2.

3 The Elevator Failure Monitoring Method

3.1 Normal Signal Logic of Elevator and Formation of Door Failure Signal

When the elevator receives signals such as door signal, floor signal, etc. as described above, the remote monitoring platform could determine whether the elevator is running with the door open and at overspeeding based on the logical combination of these signals.

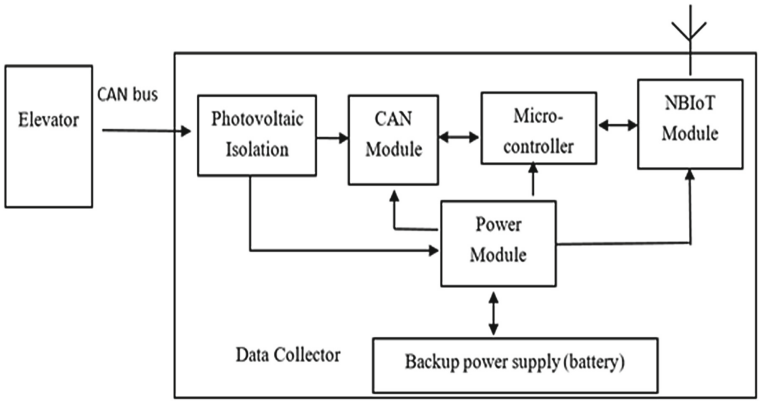


Fig. 2. Functional block diagram of data collector

The normal logic of the elevator is shown in Fig. 3. When the elevator arrives at a floor's leveling area, it sends out the leveling area signal; when the elevator controller receives it, the elevator switches from running to stationary, and the running signal becomes low level. After the elevator stops area, the door closed signal will change from a high level to a low level, indicating that the elevator door opens normally. If the target floor is still exist after the elevator finishes opening and closing, it will continue to run upward or downward, then leave the current leveling area N1 and go to the next leveling area N2 to repeat the above action.

However, if the elevator's electrical safety circuit is short-circuited, it is not always possible to get the signal for the elevator door to open when the elevator stops area. Under this condition, the door closed signal is always high, as shown in Fig. 3 by the dashed line of the door closed signal. Because the safety circuit has been short-circuited, the elevator could no longer determine whether the door is open or closed, so the elevator could still running upward or downward at this time, but the status of the elevator door is unknown.

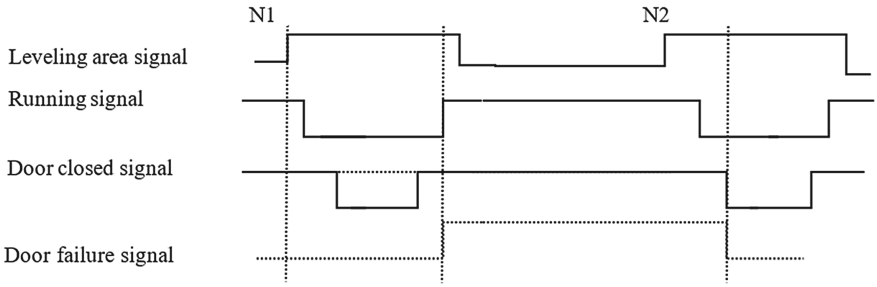


Fig. 3. Normal elevator signal logic, door failure signal is dashed line.

If the low level of the door closed signal (door opened signal) is not detected at the N1 level, the elevator will continue to run until the door is detected to open, indicating that

the elevator is running with the door failure. The door failure could appear as the elevator running with the door open, or it may also be manifested as the door being stuck closed. In reality, there is a common situation in which the elevator maintenance personnel, in order to facilitate the operation of a layer or a few layers of the door lock circuit, short the elevator only on the short floor performance door failure. Its performance and judgment logic are similar to those described above.

3.2 Elevator Door Area Signal

When an elevator has a door failure, in order to determine whether the elevator is running with the door open, a door area signal, i.e. the area where the door performs or is ready to perform the opening and closing action, is first combined. Generally, the door area signal and the leveling area signal overlap significantly. Some new elevators, in order to reduce passenger waiting time, will receive the signal before leveling area via additional sensors, allowing the door to open and close in advance. In short, the status of the door and the leveling are highly correlated.

The door area signal is shown in Fig. 4. Beginning with the leveling area signal, the elevator stops running, and the running signal changes from high to low. After the door opening and closing actions are completed, the elevator resumes running with the running signal changes from low to high. From the time the leveling signal changes from high to low to the time the running signal changes from low to high, indicating that the elevator is in the door area and the door area signal is formed.

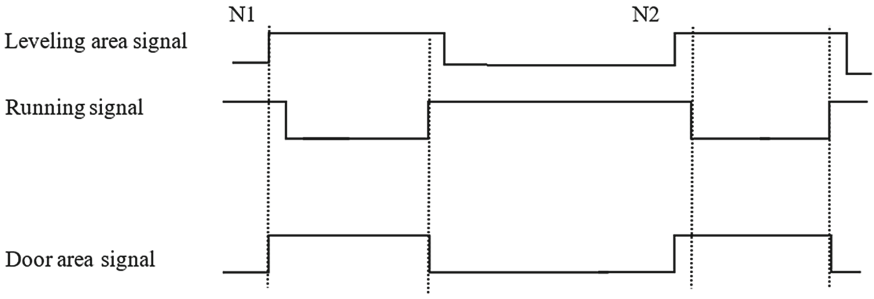


Fig. 4. Formation of the door area signal

Furthermore, the elevator's running speed could be calculated based on the duration of the running signal. As shown in Fig. 4, the elevator running signal begins at a low level, switches to a high level until it returns to a low level. The elevator running time is Δt while the elevator is running from N1 floor to N2 floor, the floor height is H, and the distance of the elevator operation is L = (N2-N1)*H. The elevator's running speed could be calculated using the above distance and time as follows.

$$v = f(L,\ \Delta t) \tag{1}$$

If this speed $v > v_m$ (maintenance speed), means that the elevator is not in maintenance and is running at a high speed. If there is a door failure at the same time, there is a serious safety hazard.

3.3 Elevator Failure Monitoring Method

Furthermore, it is possible to determine whether the elevator is running with the door open based on the door failure signal, door area signal, and car call signal, as shown in Fig. 5.

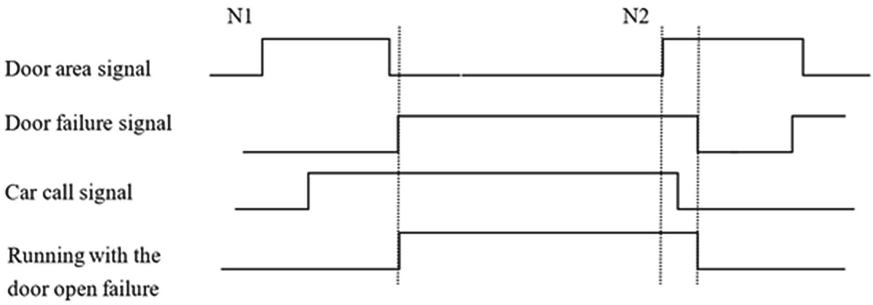


Fig. 5. Running with the door open failure formation

When the door failure signal is valid, that is, the elevator door is long open or long closed, and the elevator could still run upward or downward in the long open door state, if the elevator has a car call signal in the door area, it means that someone has walked into the elevator and pressed the call button for a floor, and the elevator door is open. The elevator is running with the door open from the time the signal in the door area changes from high level to low level (shown in Fig. 4), until the end of the car call signal and the elevator reaches the target floor. If the elevator is running at a high speed when combined with formula (1), it means that the elevator is running with the door open and at a high speed, posing a serious safety hazard. The relevant units need to be notified quickly for emergency maintenance.

When the elevator is running normally, the remote monitoring platform page is displayed as shown in Fig. 6. The elevator's basic status, such as floor and speed, can be displayed, and the elevator status shows normal, with the failure indication on the right side turned off. When the elevator has a door failure or a possible concurrent failure, such as overspeeding or trapped people, the elevator status will show failure and the corresponding failure indicator on the right side will light up. The maintenance people will be alerted in time by App Client.

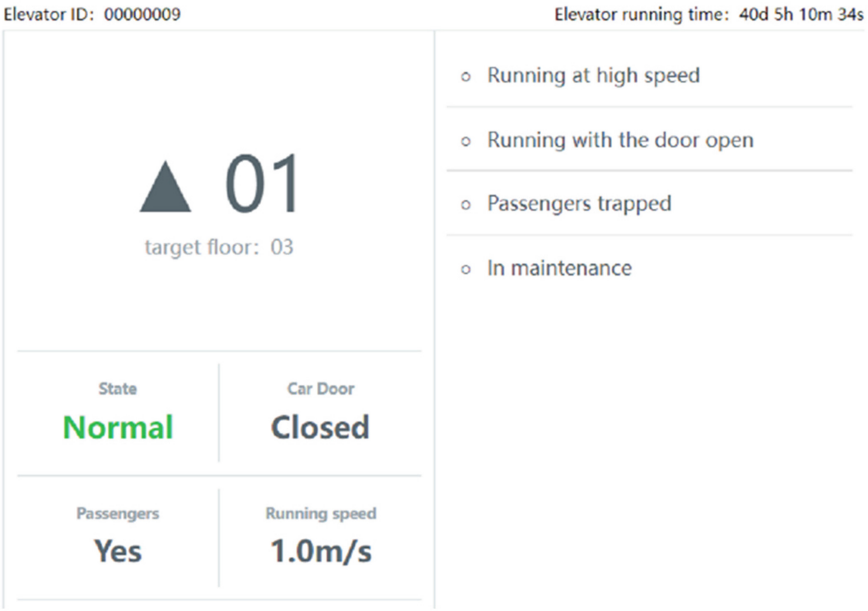


Fig. 6. Elevator under normal condition

4 Conclusion

This paper designs and implements a new type of elevator running with the door open failure monitoring system that obtains elevator signals, including floor signal, running signal, leveling in place signal, door signal, and car call signal, from the elevator CAN bus interface. It determines whether the elevator has a door failure, whether it is running with the door open, and, most importantly, whether it has trapped people. The relevant failure will be displayed once it has been detected, and maintenance personnel will be notified in time.

The system's benefits are reflected in that it is not limited by elevator brand and model because the CAN bus interface is commonly used regardless of the manufacturer or model of elevator. By analyzing the CAN bus data in advance, it is simple to achieve the unification of the monitoring system interface. With the use of photovoltaic isolation, data could only be output from the elevator and not reverse input, so the original elevator system is unaffected. Furthermore, since the signal is monitored via the CAN bus interface rather than sensor data, the error is low and there is no risk of sensor false alarm.

References

1. Dn, A., Lei, G.A., Xb, B., et al.: Preventive maintenance period decision for elevator parts based on multi-objective optimization method. J. Build. Eng. (2021)
2. Feng, W., Zhu, J., Xie, D.: The elevator upgraded the application research. In: IEEE International Conference on Cognitive Informatics & Cognitive Computing. IEEE (2012)

3. Xu, H.: Research on Cost Control of Elevator Installation and Maintenance. Zhejiang University, Zhejiang (2013)
4. Wu, G.: Research and Implementation of Elevator Remote Monitoring System and its Fault Diagnosis Technology. Southeast University, Jiangsu (2004)
5. Kong, W.: Analysis of the causes of elevator door walking accidents and preventive measures. China Elevator **26**(20), 40–41 (2015)
6. GB 7588-2003 Safety Code for the Manufacture and Installation of Elevators
7. Deng, W.: Example of design and inspection of grounding protection for electrical safety circuit of elevator door lock. China Elevator **28**(5), 38–40 (2017)
8. Kong, W.: Research on the causes and prevention analysis of elevator door opening operation. Qual. Tech. Supervision Res. **3**, 4 (2015)
9. Shi, H., Pu, H.: Intelligent elevator control and safety monitoring system. Autom. Instrument. **33**(3), 105–108 (2018)
10. An, Y.: Research on Remote Monitoring System of Elevator Without Interface. Tianjin University, Tianjin (2008)
11. Huang, Q., Cao, J., Sun, R.: Design and implementation of an elevator power failure warning system. J. Phys. Conf. Ser. **1621**(1), 012050 (7pp) (2020)
12. Salunkhe, A.A., Kamble, P.P., Jadhav, R.: Design and implementation of CAN bus protocol for monitoring vehicle parameters. In: 2016 IEEE International Conference on Recent Trends in Electronics (2016)

Research and Design of Frequency Quadrupler Based on CMOS

Hai Wang[1], Zhihong Wang[1], Guiling Sun[1,2], Xiaomei Jiang[1], Chaoran Bi[3], and Ying Zhang[1,4(✉)]

[1] Teaching Center for Experimental Electronic Information, College of Electronic Information and Optical Engineering, Nankai University, Tianjin 300350, China
caroline_zy@nankai.edu.cn
[2] Department of Electronic Information Science and Technology, College of Electronic Information and Optical Engineering, Nankai University, Tianjin 300350, China
[3] College of Media Design, Tianjin Modern Vocational Technology College, Tianjin 300350, China
[4] Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology, Tianjin 300071, China

Abstract. With the continuous development of wireless communication technology, the public has an increasingly strong demand for modern communication systems with better performance. As a vital component of RF front-end circuit, frequency multiplier is usually applied in the transmitter to generate high-frequency signal. CMOS integrated circuit, with the advantages of high integration, small volume and low cost, is widely used in communication, radar, civil, military and other fields. This paper focuses on the operating principle of frequency multiplier and the design method of triode frequency multiplier. Based on CMOS technology, a quadrupler with 4GHz output frequency is designed with ADS simulation software after debugging and optimization, which is applied to undergraduate teaching of professional experiment of electronic information science and technology.

Keywords: Frequency Multiplier · CMOS · Quadrupler · Simulation

1 Introduction

In recent years, the global information industry has developed rapidly [1, 2]. Radio frequency and microwave technology has also been widely utilized in the fields of cellular mobile system, private communication system, wireless LAN, vehicle anti-collision radar, all kinds of live broadcasting satellites and global positioning system [3–5]. Frequency multiplier, an essential part of microwave signal transmission system, can make the output signal frequency equal to an integral multiple of the input signal frequency, which directly affects the stability of the whole system and the quality of the output signal [6, 7]. Therefore, the research and design of frequency multiplier in RF microwave circuit is favored by scholars and researchers at home and abroad [8]. It is a challenging subject to design a frequency multiplier with excellent performance and stable operation.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 52–61, 2023.
https://doi.org/10.1007/978-981-99-2653-4_7

The crystal oscillator can generate a frequency source with high stability, but the frequency of the oscillation source is relatively low, generally not higher than 150MHz. The microwave electric vacuum tube device technology which can produce high-frequency oscillation source has the same disadvantage as the microwave semiconductor oscillation technology, that is lack of frequency stability. Instead, high frequency source with good stability can be obtained by frequency doubling technology [9, 10]. As the frequency multiplier is adopted in the transmitter, the main oscillator can oscillate at a lower frequency to improve the frequency stability, which means the low-frequency signal can be multiplied into a high-frequency signal while retaining the high stability [11]. The working frequency band of the system is widened by frequency doubling technology, such as the oscillation source in the frequency sweep instrument. Multiple frequencies in integer ratio can also be obtained from one oscillator, such as the frequency-marker oscillator in the frequency sweeper. In addition, the frequency multiplier is applied to increase the frequency offset in frequency modulation equipment. In phase keyed communication machine, the frequency multiplier is an indispensable component of carrier recovery circuit. It is developing rapidly and widely utilized in RF and microwave circuit design [12].

2 Principle of Frequency Multiplier

The frequency multiplier is a device for multiplying the signal frequency. If the input signal frequency is f_{in}, the output signal frequency after passing through the frequency multiplier is f_{out}, which is equal to $N \cdot f_{in}$. The coefficient N is a positive integer, which is called frequency doubling number. The frequency doubling circuit is generally divided into two parts. One part depends on the nonlinear circuit to make the input low-frequency signal produce rich harmonics, which is responsible for generating new frequency components. The other part is the filter circuit, mainly used to obtain the frequency component required by the designer.

The input-output matching method is usually applied to design the frequency doubler. The transistor with push-push structure is used to design frequency multipliers in order to obtain circuit modules with better performance [13]. When designing the tripler, engineers usually carry out the design work by combining injection locking, self-mixing structure and input-output matching. It is generally difficult to design a quadrupler or more high-order frequency multiplier directly. Therefore, researchers often cascade multistage low-order frequency multiplier to realize its function.

For a long time, chemical compound transistor process is the most commonly used in the design of frequency multiplier. In recent years, however, with the continuous development and maturity of CMOS tube technology, this design method has gradually become more popular among researchers [14]. There are different types of frequency multipliers according to various implementation methods. The working frequency is higher, and the frequency doubling number is greater. Microwave frequency multiplier has developed from the earliest nonlinear varistor diode frequency multiplier to varactor diode, step tube and avalanche tube frequency multiplier, and from bipolar transistor frequency multiplier to single gate and double gate microwave FET frequency multiplier.

2.1 Varactor Frequency Multiplier

This is a frequency multiplier that depends on the capacitance change of varactor diode, also known as variable reactance diode [15]. When PN junction is reverse biased, the change trend of junction capacitance and reverse bias voltage is opposite. When a stable signal is input to the varactor, the junction capacitance changes by different reverse bias voltages. The waveform distorted current is generated in the circuit, that is, the signal generates higher harmonics after passing through the varactor. At this time, the required frequency can be selected to realize signal frequency doubling by adding a suitable filter to the subsequent circuit. The frequency multiplication efficiency of this kind of frequency multiplier is inversely proportional to the frequency multiplication times N.

2.2 Step Recovery Diode Frequency Multiplier

Step recovery diode is an ideal nonlinear device. The main feature is that the falling time during shutdown is very short, almost equal to zero. If a sine wave signal is applied to the step diode, the output waveform will be quickly clamped off because the transfer time is limited. Many high-frequency harmonics will be generated in the rapid mutation process of reverse recovery current. If the reverse bias voltage is applied at both ends of the step recovery diode, the capacitance of the PN junction remains basically stable. On the contrary, if it is a forward voltage, the capacitance will become very large. The two states correspond to two equivalent circuits: the reverse bias is capacitance and the forward conduction is short circuit. In the field of microwave technology, step recovery diode is an excellent frequency doubling device. The output signal has the characters of high-efficiency and high-order frequency doubling, and the frequency doubling number N is relatively large.

2.3 Triode Frequency Multiplier

Triode frequency multiplier is a kind of nonlinear conductance frequency multiplier. The nonlinear resistance of PN junction is utilized to generate rich harmonics, which means the triode works in the amplification region in the frequency doubling circuit. Triode frequency multiplier consists of class C bias transistor, tuning circuit at output frequency and band-pass filter circuit. When the signal is input from the base of the triode, the triode is turned on for some time. The collector current can be approximately regarded as a half wave rectified waveform. The collector tuning circuit presents high impedance load characteristics on the required harmonic components. At the same time, the band-pass filter attenuates the unwanted harmonic components at the output of the tuning circuit, so as to realize the frequency doubling of the signal. This type of frequency multiplier is characterized by unidirectionality, good isolation and high gain. Moreover, field effect transistor has the advantages of high frequency, high input impedance, low noise, low driving power and strong stability. The frequency multiplier designed by field effect transistor can obtain frequency doubling signal with wide frequency band, and can effectively improve the isolation of input and output ports.

3 Design and Simulation of Frequency Multiplier Based on ATF54143

ATF54143, produced by Avago company, is an enhanced transistor with low noise and high electron mobility. Compared with other devices, it has lower noise, fewer power and simpler bias circuit design [16]. ATF54143 is suitable for designing frequency multiplier with high frequency and low frequency doubling times. The circuit includes four parts: input matching circuit, bias circuit, output matching circuit and band-pass filter circuit, as shown in Fig. 1. The first three parts constitute the amplification circuit of the frequency doubling module. It is the core circuit, which is mainly applied to amplify the input signal and produce many new harmonic waves. And the final band-pass filter circuit is added to obtain the required frequency.

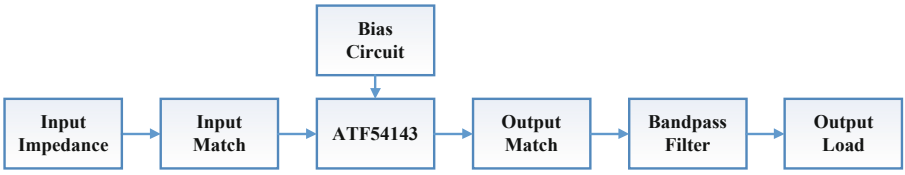


Fig. 1. Schematic Diagram of Frequency Multiplier Based on ATF54143

3.1 Amplifier Circuit Design

As the core component, the performance of amplifier circuit directly affects the stability of the whole system. Therefore, multiple indexes must be comprehensively considered in the design of amplification circuit. Amplifier circuit is the first stage of frequency doubling circuit. it is necessary to have the characteristics of low noise and high gain at the same time. In this way, the sensitivity of the whole system can be improved and the noise figure can be reduced. There are two common design methods of amplifier circuit, one is balanced amplifier structure, the other is single-stage amplifier structure. Both of them have their own advantages and disadvantages. The former consists of two single-stage amplifiers and two 3dB couplers, which is more convenient to achieve minimum noise figure and better input-output matching. The reliability and stability of balanced amplifier structure are better than that of single-stage amplifier structure, however, the constraints of cost and size cannot be ignored. When the input-output reflection coefficient is less than -10dB, the circuit matches well. The single-stage amplifier can easily achieve this achievement, so it is used in the circuit design process of this paper.

3.1.1 Stability Analysis

The amplifier circuit is generally regarded as a two port network with known S-parameter, as shown in Fig. 2. Γ_L and Γ_S are load and source reflection coefficients respectively, which are determined by load impedance and source impedance. In order to make the amplification circuit stable, both the mode of the input reflection coefficient $|\Gamma_{in}|$ and the

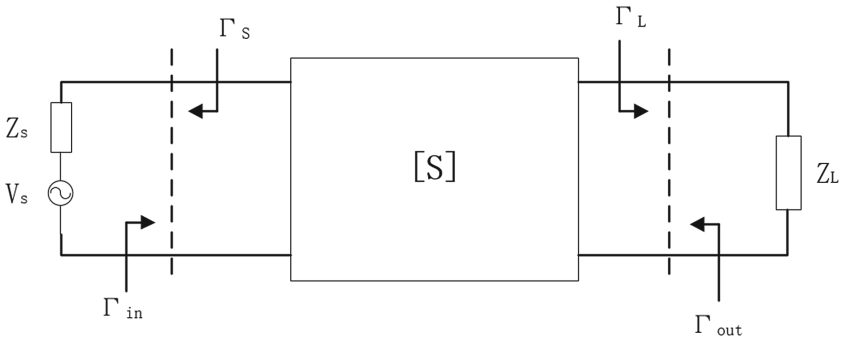


Fig. 2. The Equivalent Dual Port Network of Amplifier Circuit

mode of the output reflection coefficient $|\Gamma_{out}|$ are less than 1. It is usually required to meet the requirements of formula 1 and 2 at the same time to ensure the absolute stability of the designed circuit [17]. K is called the stability discrimination coefficient. Besides, a single parameter μ can be used to determine whether the circuit is in an absolutely stable state. For unstable circuits, a negative feedback circuit is usually added to the source of the transistor to improve the stability of the overall system.

$$K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} > 1 \tag{1}$$

$$|\Delta| = |S_{11}S_{22} - S_{12}S_{21}| \tag{2}$$

3.1.2 Matching Circuit

It is essential to design an impedance matching network with satisfactory performance, which can ensure the normal operation of the amplification circuit. There are generally three parts in the circuit: input matching, output matching and interstage matching. There are two main methods to design a matching circuit network. The first is the analytical method, that is, the numerical values of various components in the circuit are obtained by computer simulation. The results are very accurate, but the process is troublesome. Smith chart is applied for matching in the second method, called graphic design method as well. The gain and stability of the amplification circuit are mainly discussed in the frequency doubling circuit. Conjugate matching should be achieved between the transistor and the input-output circuit, so that the amplification circuit can meet the requirements.

3.1.3 Bias Circuit

In the transistor amplification circuit, the required voltage and current values need to be provided by the bias network. In other words, the suitable quiescent point for the transistor is supplied by the bias circuit. In order to make the bias circuit stable, a DC negative feedback circuit will be added in the design process, so that the quiescent point will remain unchanged. A resistor is connected in series with the emitter of the transistor

to form a current series negative feedback, which makes the quiescent point more stable. But the gain of the amplification circuit will be affected to some extent. A suitable capacitor can be connected in parallel with this resistor to form an AC signal path. In this way, the feedback of the circuit to the AC signal is avoided, and the influence of the emitter resistance on the gain of the amplification circuit is reduced. In order to avoid new noise and improve circuit performance, inductive negative feedback can be adopted. Because the resistance of the inductor will rise up with the increase of frequency, the interference and intermodulation component of high-frequency harmonic to the circuit can be suppressed.

3.2 Circuit Design of Bandpass Filter

As the last stage of frequency multiplier, filter circuit plays an indispensable role. When the preamplifier circuit amplifies the input signal and generates a large number of harmonics, it is necessary to use the filter circuit to obtain a pure single output signal according to the required frequency. In this design, the lumped parameter filter may make the final result inaccurate because of the high design frequency. The microstrip filter is applied, instead. On the basis of the filter theory, all types of filter design can start with the normalized low-pass filter, and then map to band-pass filter. The parameters of microstrip line can be devised with the help of simulation software such as ADS, or the combination of calculation and software in the design process. A qualified filter circuit should not only accurately screen out the target frequency, but also ensure that the transmission loss is as small as possible.

3.3 Frequency Quadrupler Simulation

The quiescent point is determined according to the data manual of ATF54143 and the initial bias circuit is generated with the help of ADS simulation software. Replace the resistance value of the theoretical calculation result with the similar conventional nominal resistance. Add the choke circuit to obtain the bias circuit as shown in Fig. 3, after the stability analysis of the circuit. The input and output signals are matched by the Smith Chart tool of ADS. Conjugate matching is adopted in this process, that is, ATF54143 should have a conjugate relationship with the impedance of the input and output circuit. Then, the filter circuit is further optimized. The schematic diagram is illustrated in Fig. 4. In the light of the simulation results shown in Fig. 5, the signal with 4GHz frequency can be successfully generated by this system, which means the designed circuit can screen out the required frequency.

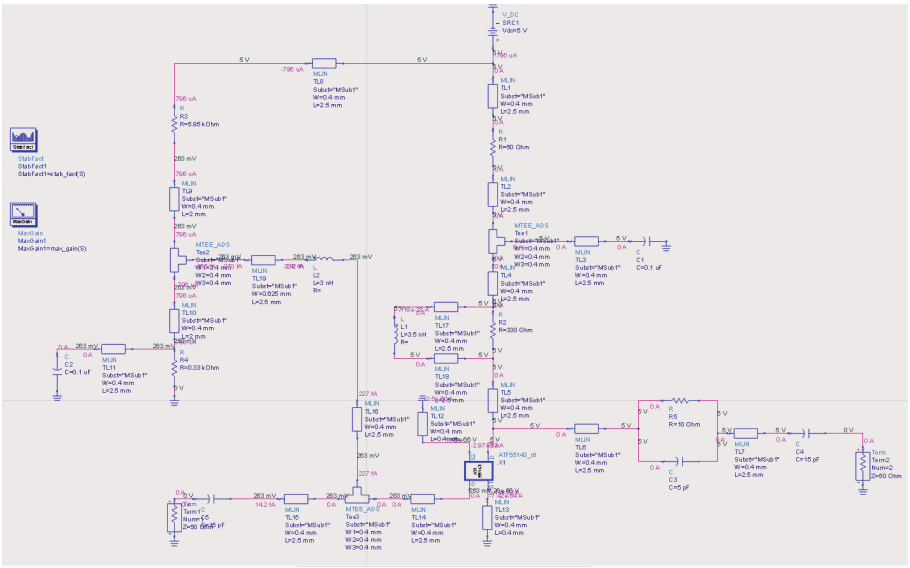


Fig. 3. Schematic Diagram of Bias Circuit

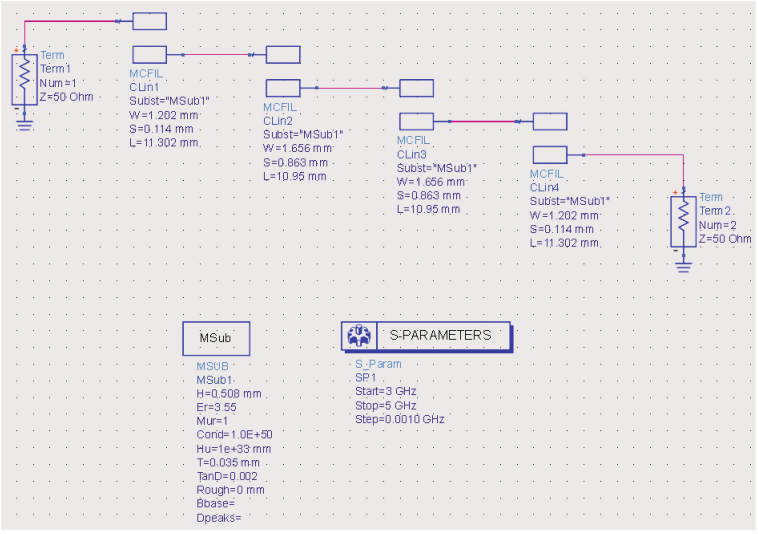


Fig. 4. Schematic Diagram of Filter Circuit

Ultimately, the amplification circuit and filter circuit are connected together to form a complete frequency quadrupler, as demonstrated in Fig. 6. As can be seen from Fig. 7, the output signal generated by the frequency doubling circuit meets the design requirements.

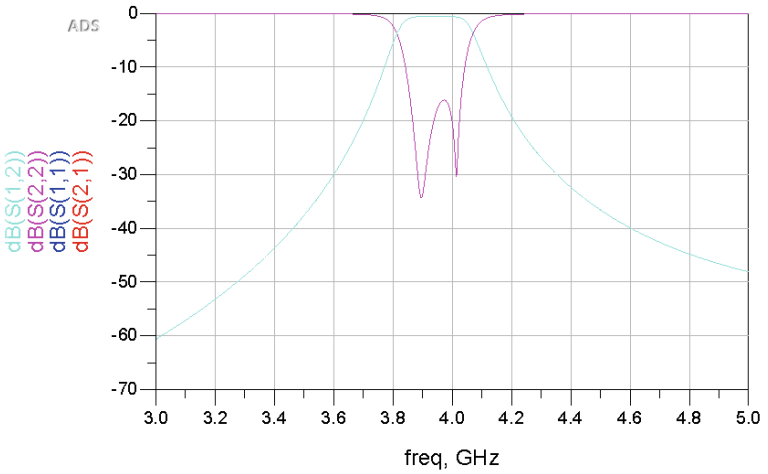


Fig. 5. Simulation Result of Filter Circuit

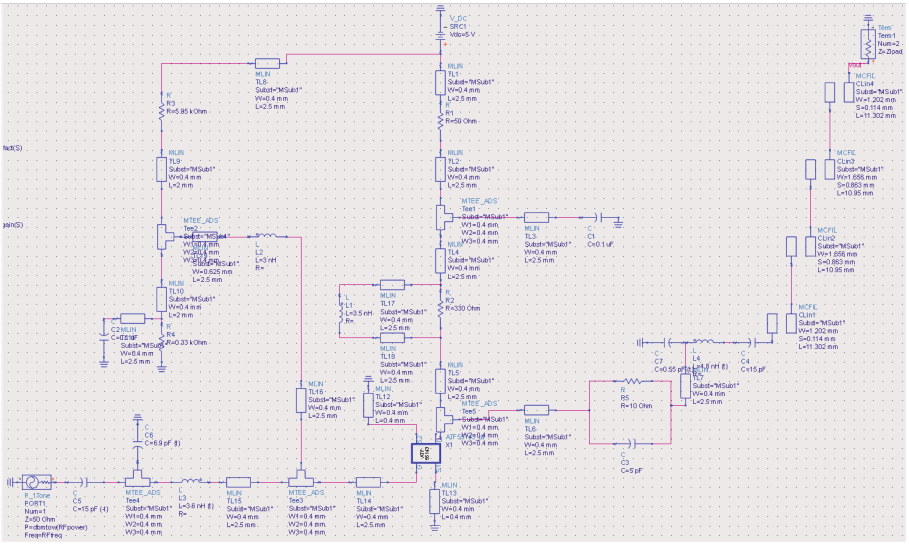


Fig. 6. Schematic Diagram of Frequency Quadrupler

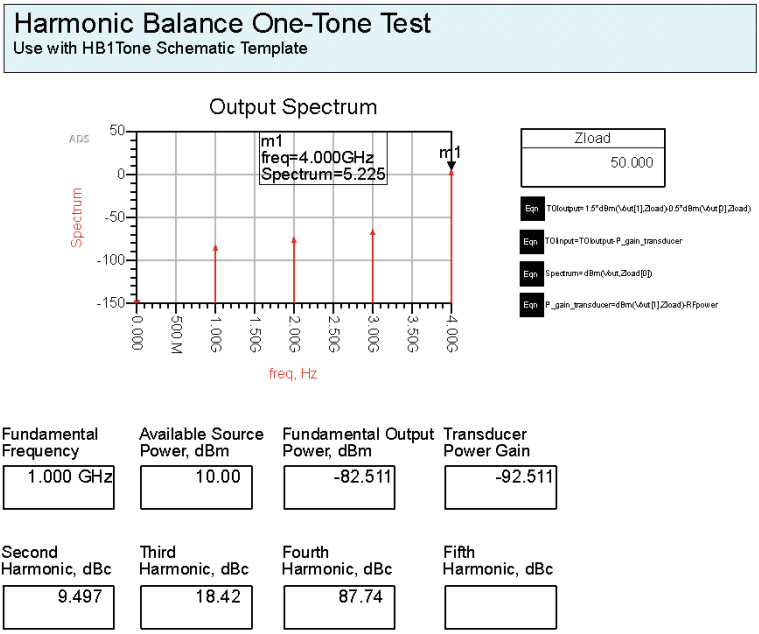


Fig. 7. Simulation Results of Frequency Quadrupler

4 Conclusion

The types and principles of various frequency multipliers, as well as the basic theory and design method of each part circuit, are summarized and analyzed in this paper. On the basis of these related research, a quadruple frequency multiplier with output frequency of 4GHz is designed with the aid of ADS software. By the schematic design and circuit simulation, the design requirements are eventually realized after debugging and optimization.

Acknowledgement. This work was supported by the 2nd batch of University-Industry Collaborative Education Program of Ministry of Education in 2022 (220805078291821 and 220902278263517); the 1st batch of University-Industry Collaborative Education Program of Ministry of Education in 2021 (202101186002); the 2nd batch of University-Industry Collaborative Education Program of Ministry of Education in 2021 (202102296002); the 2022 Undergraduate Experimental Teaching Reform Project Fund of Nankai University (22NKSYSX03); the 2022 Undergraduate Education Teaching Reform Project Fund of Nankai University and Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology.

References

1. Sugama, C., Chandrasekar, V.: Topside Radio Frequency System Consolidation with Gregorian Parabolic Antenna Stacking. Electron. Lett. **56**(24), 1295–1298 (2020)

2. Ali, M., Watanabe, A.O., Lin, T.H., et al.: Package-integrated, wideband power dividing networks and antenna arrays for 28-GHz 5G new radio bands. IEEE Trans. Compon. Packag. Manuf. Technol. **10**(9), 1515–1523 (2020)
3. Misra, S., Kocz, J., Jarnot, R., et al.: Development of an on-board wide-band processor for radio frequency interference detection and filtering. IEEE Trans. Geosci. Remote Sens. **57**(6), 3191–3203 (2019)
4. Draper, D.W.: Radio frequency environment for earth-observing passive microwave imagers. IEEE J. Sel. Topics Appl. Earth Observations Remote Sens. **11**(6), 1913–1922 (2018)
5. Megawer, K.M., Elkholy, A., Gamal, A.M., et al.: Design of crystal-oscillator frequency quadrupler for low-jitter clock multipliers. IEEE J. Solid-State Circuits **54**(1), 65–74 (2019)
6. Jaffri, I., Ayed, A.B., Darwish, A.M., et al.: Novel baseband equivalent model for digital predistortion of wideband frequency-multiplier-based millimeter wave sources. IEEE Trans. Microw. Theory Tech. **68**(9), 3942–3957 (2020)
7. Chung, A., Darwish, A.M., Viveiros, E., et al.: Analysis and compensation of nonidealities in frequency multiplier-based high-frequency vector signal generators. IEEE Trans. Microw. Theory Tech. **67**(6), 2270–2283 (2019)
8. Kast, J., Elsherbeni, A.Z.: Simulation of a nonlinear frequency multiplier using the FDTD technique. Appl. Comput. Electromagn. Soc. J. **35**(11), 1426–1427 (2020)
9. Liu, Y.J., Liu, G., Asbeck, P.M.: High-order modulation transmission through frequency quadrupler using digital predistortion. IEEE Trans. Microw. Theory Tech. **64**(6), 1896–1910 (2016)
10. Vu, T.A., Takano, K., Fujishima, M.: 300-GHz Balanced varactor doubler in silicon CMOS for ultrahigh-speed wireless communications. IEEE Microwave Wirel. Compon. Lett. **28**(4), 341–343 (2018)
11. Vehring S, Ding Y S, Scholz P, et al. A 3.1-dBm E-Band Truly Balanced Frequency Quadrupler in 22-nm FDSOI CMOS. IEEE Microwave Wirel. Compon. Lett. **30**(12), 1165–1168 (2020)
12. Ku, B.H., Chung, H., Rebeiz, G.M.: A Milliwatt-level 70–110 GHz frequency quadrupler with >30 dBc harmonic rejection. IEEE Trans. Microw. Theory Tech. **68**(5), 1694–1705 (2020)
13. Steinweg, L., Riess, V., Starke, P., et al.: A low-power 255-GHz single-stage frequency quadrupler in 130-nm SiGe BiCMOS. IEEE Microwave Wirel. Compon. Lett. **30**(11), 1101–1104 (2020)
14. De La Cruz-Blas, C.A., Thomas-Erviti, G., Algueta-Miguel, J.M., et al.: CMOS aAnalogue current-mode Multiplier/Divider circuit operating in Triode-saturation with Bulk-driven techniques. Integration **59**, 243–246 (2017)
15. Choi, S.H., Yi, C., Kim, M.: A 600 GHz varactor frequency multiplier using 65-nm CMOS technology. Microw. Opt. Technol. Lett. **60**(1), 114–117 (2018)
16. Du H, Li Y, Sun L, et al. Fault diagnosis of radio frequency low noise amplifier using GMM-HMM. In: 8th International Symposium on Test Automation and Instrumentation. pp. 159–163 (2020)
17. Wang H, Wang Z H, Sun G L, et al. Optimal design of an S-Band low noise amplifier. In: 2019 Communications, Signal Processing, and Systems. vol. 571, pp. 439-447 (2020)

Simulation Design of a Low Noise Amplifier with 2.5 GHz Operating Frequency

Hai Wang[1], Guiling Sun[1,2], Zhihong Wang[1], Yi Zhang[3], Xiaomei Jiang[1], and Ying Zhang[1,4(✉)]

[1] Teaching Center for Experimental Electronic Information, College of Electronic Information and Optical Engineering, Nankai University, Tianjin 300350, China
caroline_zy@nankai.edu.cn

[2] Department of Electronic Information Science and Technology, College of Electronic Information and Optical Engineering, Nankai University, Tianjin 300350, China

[3] Institute of Photoelectronic Thin Film Devices and Technology and Tianjin Key Laboratory of Thin Film Devices and Technology, Nankai University, Tianjin 300350, China

[4] Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology, Tianjin 300071, China

Abstract. Nowadays the wireless communication infrastructure industry is facing with a number of great challenges, such as providing better signal quality and wider coverage in a crowded spectrum. Low noise amplifier (LNA), a significant component of communication system, plays an important role in modern wireless communication, radar, electronic countermeasure system between countries and so on. The sensitivity performance of the base station receiver can be dramatically improved by a suitable amplifier, especially the first stage LNA. In this paper, a low noise amplifier with an operating frequency of 2.5 GHz is designed with the help of the simulation software. The minimum values of input return loss S_{11} and output return loss S_{22} are −19.232 dB and −29.422 dB respectively. The maximum gain is 14.567 dB at the center frequency. Besides, the stability coefficient K is greater than 1 and the noise figure (NF) is 0.662 dB.

Keywords: Wireless Communication · LNA · Operating Frequency · Stability coefficient · Noise Figure

1 Introduction

With the dramatic progress of global wireless communication industry, the relative technology and infrastructure of modern radio frequency (RF) network are extensively utilized in public service areas [1–3]. It not only has a positive impact on people's daily life in contemporary society, but also directly promotes the rapid development of RF communication system in China [4, 5]. People are no longer satisfied with the simple function of the RF front end in the wireless system, that is, receiving signals [6]. Gradually, consumers put forward higher requirements, such as expanding the distance and coverage of the signal, reducing power consumption and noise interference, enhancing receiver sensitivity and so on [7, 8].

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 62–68, 2023.
https://doi.org/10.1007/978-981-99-2653-4_8

Microstrip line is a microwave transmission line composed of a single conductor strip supported on a dielectric substrate [9, 10]. It is a planar transmission line suitable for manufacturing microwave integrated circuits, because it is convenient for the connection of the whole system and the installation of microwave active devices. Due to the compact structure, the whole circuit can be produced in two dimensions by microstrip circuit technology. Compared with metal waveguide, microstrip line has the advantages of small volume, light weight and low cost [11, 12]. However, the high loss and the small power capacity are challenging and unsatisfying. Various types of microstrip lines have emerged as a result of the development of microwave low-loss dielectric materials and microwave semiconductor devices [13]. Dielectric substrates are often made of materials with high dielectric constant and low microwave loss [14]. The conductor should have the characteristics of high conductivity, good stability and strong adhesion to the substrate. The second generation of microwave printed transmission line is widely used in microwave integrated circuits and high-speed pulse circuits. It has the advantages of small size and wide frequency band, as well as the high reliability, precision and sensitivity [15, 16].

Low noise amplifier, one of the indispensable modules of RF circuit receiver, is mainly used for digital high-frequency signal amplification [17, 18]. It can not only reduce the distortion of the signal transmitted by the antenna in the subsequent system, but also drive the later stage circuit. The leading performance indexes of LNA are usually measured by operating frequency, transmission gain and noise figure [19, 20]. In addition, the excellent system stability is also a basic requirement, which means the system should be able to work normally when the input signal or external condition change [21]. If the circuit is in an unstable state, it cannot be applied to practice regardless of other outstanding performance. In this paper, an S-band LNA with high gain, low noise figure and stable operation at 2.5 GHz is designed and optimized.

2 Design and Implementation of LNA

The RF input signal amplifier is mainly composed of five components, including radio frequency transistor, bias control circuit, stability control part, input and output matching circuit. The overall design step of the module is start with the selection of the appropriate transistor according to the design requirements. Based on the DC analysis, the engineers devise the bias control circuit and determine the conditions for stability. Then, they design the input and output matching circuit, optimize the performance and structural layout of the whole integrated circuit.

ATF54143, a high electron mobility transistor chip from Avago, is selected as the core device of LNA in this design. According to its bias curve in S-band, VDS and IDS of the transistor are separately 3.75 V and 0.011 mA, when the center frequency is 2.5 GHz. So the quiescent point can be determined by the bias circuit which is a resistive voltage divider circuit to provide DC power to the transistor. The AC signal leakage in this circuit is prevented by two pairs of bypass filter capacitors. The stability circuit with resistance and inductance in parallel is applied to improve the Q value of the whole system. Combined with negative feedback regulation, it has good stability in the operating frequency band, as shown in Fig. 1. It can be seen from Fig. 2 that the transmission gain at 2.5 GHz is approximately 15.274 dB, which is satisfactory.

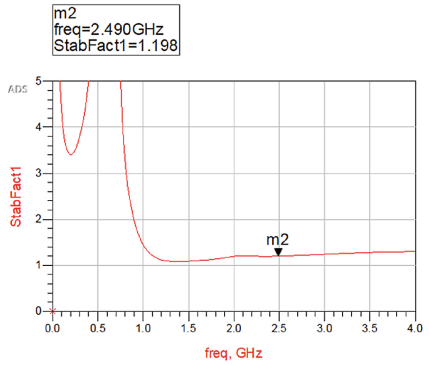


Fig. 1. Stability Curve of DC Bias Circuit

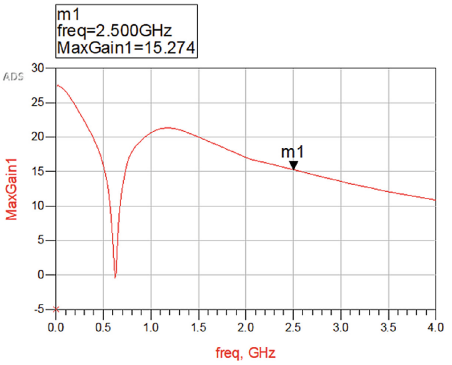


Fig. 2. Maximum Gain Curve of DC Bias Circuit

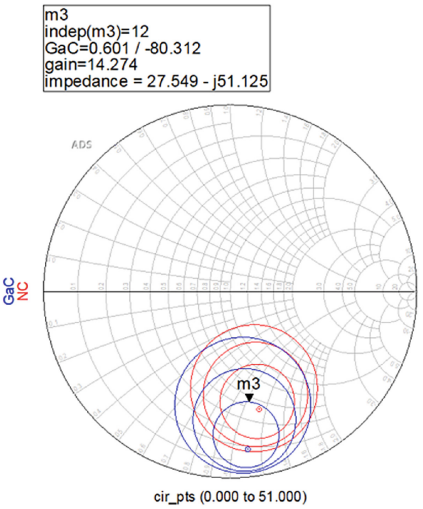


Fig. 3. Constant Noise Circles & Input Gain Circles

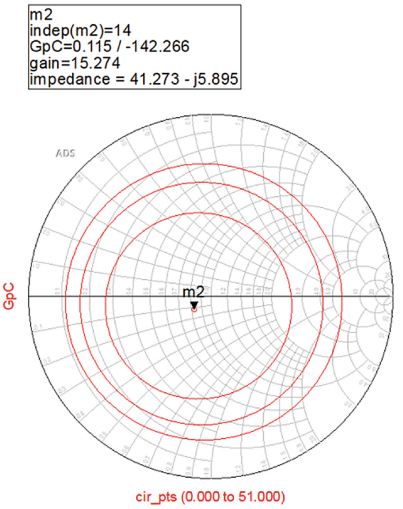


Fig. 4. Constant Output Gain Circles

As illustrated in Fig. 3, the positions of the minimum noise figure and the maximum gain at the input of the bias circuit do not coincide. Generally, it is necessary to take into account all the parameters of the low noise amplifier so that it could have smaller noise figure, larger transmission gain and lower power consumption. Point m3 is on the maximum gain circle and inside the minimum noise figure circle at the same time, which means it has excellent gain coefficient and anti-noise performance. In this case, the amplifier has the best performance and the equivalent input impedance is 27.549–j51.125 Ω. The maximum output gain of the bias circuit is at point m2 which is 15.274 dB, shown in Fig. 4. At this moment, the equivalent output impedance is 41.273–j5.895 Ω. The input and output matching circuits also play a significant role. After designed and optimized, the equivalent input and output impedance are respectively 26 + j51 Ω and

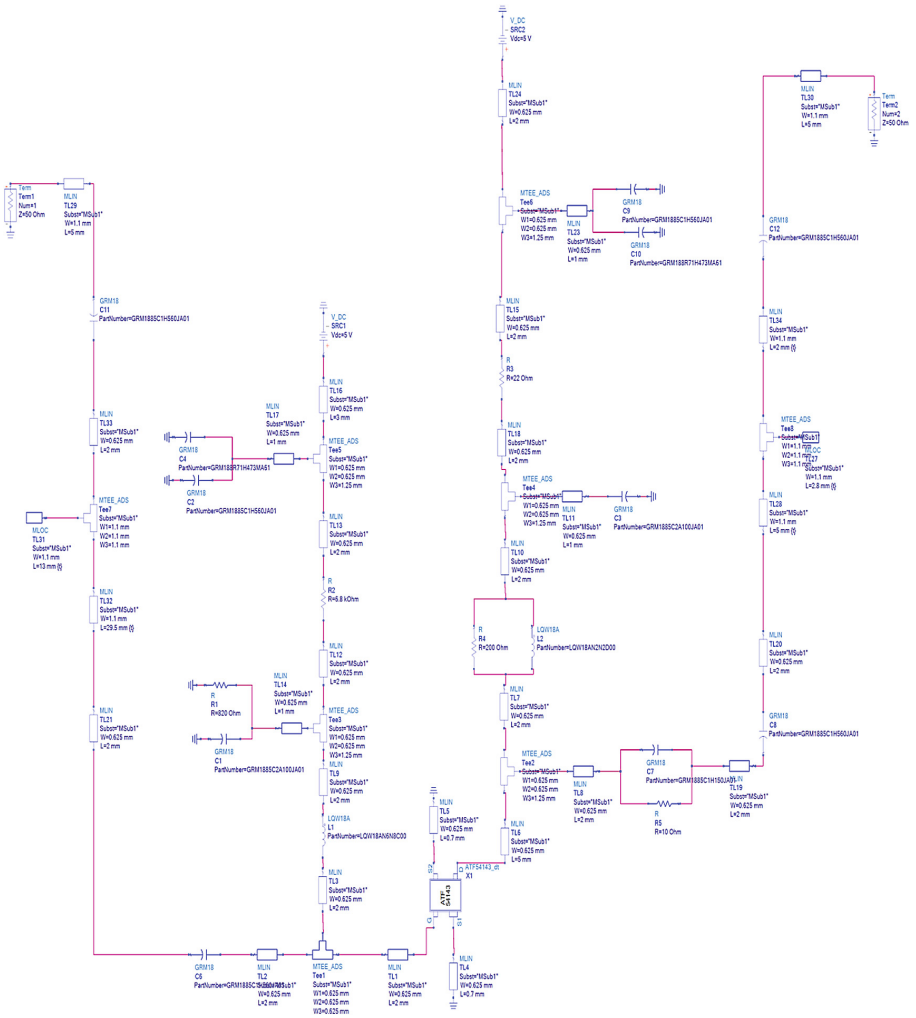


Fig. 5. Schematic Diagram of Low Noise Amplifier

40.5 + j5.5 Ω near the operating frequency. Thus, a low noise amplifier with a center frequency of 2.5 GHz is designed, as shown in Fig. 5.

The curve indicated in Fig. 6 shows that the maximum gain of LNA near the operating frequency band meet the design requirements. In the frequency range from 2.3 to 2.7 GHz, S_{21} is between 10 and 14.592 dB, shown in Fig. 7, which means the transmission efficiency is relatively high. S_{11} and S_{22} are simultaneously extraordinary small. In addition, this LNA also has the characteristics with delightful stability coefficient and noise figure. The input and output are both matched to the maximum gain circle. A small part of the noise figure is sacrificed for greater transmission gain.

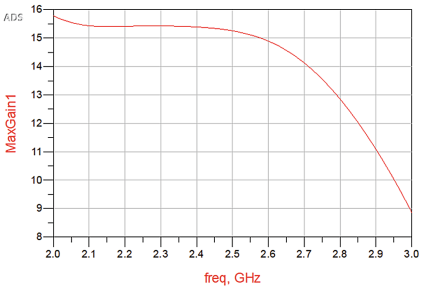


Fig. 6. Maximum Gain of LNA

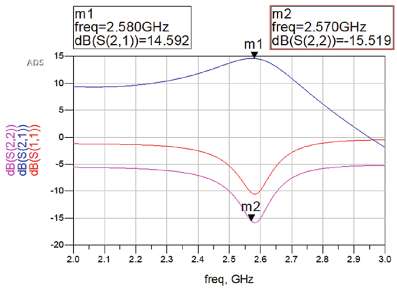


Fig. 7. Scattering Parameters of LNA

Manual wiring is carried out on the basis of the layout formed by the software. The distance between discrete devices shall be reduced as far as possible. And it shall be as close and convenient as possible to connect to the ground. Design grounding vias to further improve the layout and connect the circuits led by two DC power supplies into one in parallel. Remove discrete devices and add ports in the original position, and design grounding holes to further improve the layout. Connect the circuits led by two DC power supplies into one in parallel. The circuit layout of S-band low noise amplifier is ultimately obtained in Fig. 8.

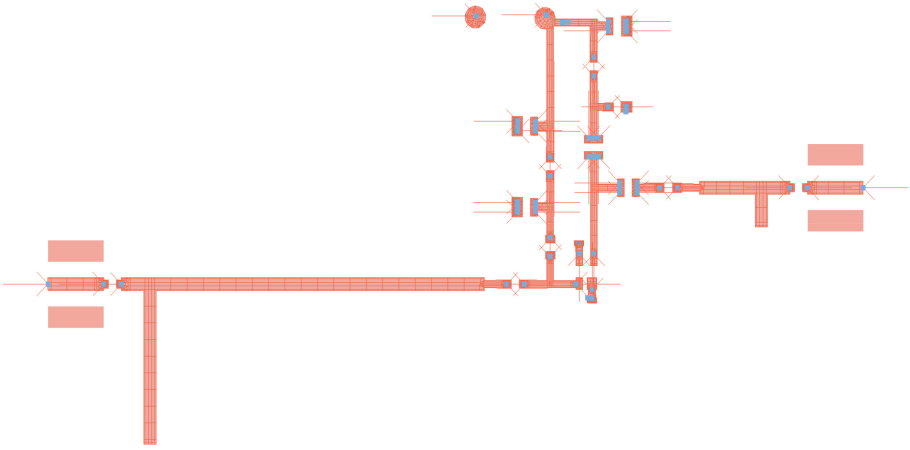


Fig. 8. Circuit Layout of LNA

With joint simulation of schematic diagram and circuit layout, the results are more consistent with the actual situation. As can be seen from Fig. 9(a), the maximum value of S_{21} is 14.567 dB at the frequency of 2.5 GHz, which means most signals can pass through the designed circuit module smoothly. The return loss S_{11} and S_{22} reach the minimum near the operating frequency, which are −19.232 dB and −29.422 dB respectively. Because the return loss is negative, the circuit is in a relatively stable working condition and it is not easy to oscillate. Noise figure and stability coefficient are 0.662

dB and 1.215 dB separately according to the simulation results illustrated in Fig. 9(b) and Fig. 9(c), that is, the LNA's indicators are excellent.

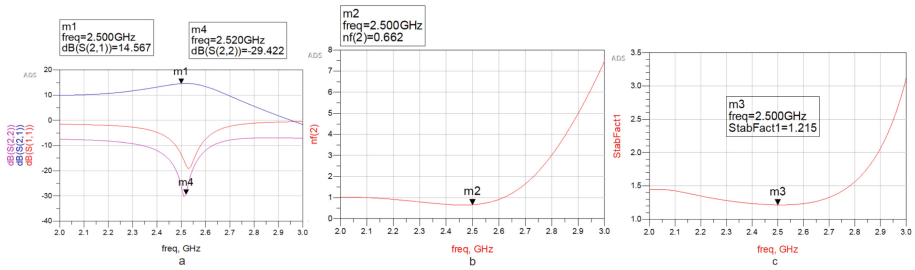


Fig. 9. Co-simulation results of LNA (a) Scattering Parameters (b) Noise Figure (c) Stability Coefficient

3 Conclusion

In this paper, a low noise amplifier working in S-band is designed by using ATF54143 chip. The matching circuit improves the gain on the basis of ensuring low noise performance. The center frequency is 2.5 GHz and the maximum gain is about 14.567 dB through several debugging and optimization by simulation software. Other performance indexes, such as return loss, stability coefficient and noise figure, also meet the design requirements. The devised amplification unit is good at improving transmission gain and eliminating signal noise, however, the bandwidth is a little narrow. Nowadays, further research needs to be carried out for broadband microwave signal and RF signal. For example, connect two chips of the same type in parallel by the bridge structure which can increase the bandwidth on the basis of ensuring the gain and noise performance. Engineers also can focus on the research of new integrated transistor chips with better function.

Acknowledgement. This work was supported by the 2nd batch of University-Industry Collaborative Education Program of Ministry of Education in 2022 (220805078291821 and 220902278263517); the 1st batch of University-Industry Collaborative Education Program of Ministry of Education in 2021 (202101186002); the 2nd batch of University-Industry Collaborative Education Program of Ministry of Education in 2021 (202102296002); the 2022 Undergraduate Experimental Teaching Reform Project Fund of Nankai University (22NKSYSX03); the 2022 Undergraduate Education Teaching Reform Project Fund of Nankai University and Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology.

References

1. Kumar, A.R.A., Dutta, A., Sahoo, B,D.: A low-power reconfigurable narrowband/wideband LNA for cognitive radio-wireless sensor network. IEEE Trans. Very Large Scale Integration (VLSI) Systems, 2020, 28(1): 212–223

2. Saied, A.M., Abutaleb, M.M., Eladawy, M.I., et al.: Ultra-low power LNA design technique for UWB applications. AEU-Int. J. Electron. C. **97**, 149–153 (2018)
3. Weinreb, S., Shi, J.: Low Noise Amplifier with 7-K Noise at 1.4GHz and 25°C. IEEE Transactions on Microwave Theory and Techniques **69**(4), 2345–2351 (2021)
4. Lin, C.Y., Huang, G.L., Lin, M.T.: Compact ESD protection design for CMOS low-noise amplifier. IEEE Trans. Electron Dev. **67**(1), 33–39 (2020)
5. Komatsu, K., Miyaji, Y., Uehara, H.: Iterative nonlinear self-interference cancellation for in-band full-duplex wireless communications under mixer imbalance and amplifier nonlinearity. IEEE Trans. Wireless Commun. **19**(7), 4424–4438 (2020)
6. Aneja, A., Li, X.J.: Design and analysis of a continuously tunable low noise amplifier for software defined radio. Sensors **19**(6), 1273–1279 (2019)
7. Wibisono, G., Wildan, M., Wahyudi, J., et al.: Co-design structure of dual-band LNA and dual-band BPF for radio navigation aid application. Wireless Pers. Commun. **116**(3), 1659–1681 (2021)
8. Guitton, G., De Souza, M., Mariano, A., et al.: Design Methodology Based on the Inversion Coefficient and its Application to Inductorless LNA Implementations[J]. IEEE Trans. Circuits Syst. I Regul. Pap. **66**(10), 3653–3663 (2019)
9. Eroglu, A., Ronnow, D.: Design of multilayer and multiline microstrip directional coupler with closed form relations. Progress In Electromagn. Res. C **83**, 1–14 (2018)
10. Ponchak, G.E.: Coplanar stripline coupled to planar double split-ring resonators for bandstop filters. IEEE Microwave Wirel. Compon. Lett. **28**(12), 1101–1103 (2018)
11. U-Yen, K., Rostem, K., Wollack, E.J.: Modeling strategies for superconducting microstrip transmission line structures. IEEE Trans. Appl. Superconduct. **28**(6), 1–5 (2018)
12. Aslam, K., Khanna, M., Veenugopal, V., et al.: Microwave monolithic filter and phase shifter using magnetic nanostructures. AIP Adv. **8**(5), 1–6 (2018)
13. Shu, S., Klimovich, N., Eom, B.H., et al.: Nonlinearity and wideband parametric amplification in an NbTiN microstrip transmission line. Phys. Rev. Res. **3**(2), 1–9 (2021)
14. Kacar, M., Weller, T., Mumcu, G.: Conductivity improvement of microdispensed microstrip lines and grounded coplanar waveguides using laser micromachining. IEEE Trans. Compon. Packag. Manuf. Technol. **10**(12), 2129–2132 (2020)
15. Zhou, M., Ren, H., Arigong, B.: A novel microstrip line balun with transparent port impedance and flexible open arm structure. IEEE Microwave Wirel. Compon. Lett. **30**(2), 160–163 (2020)
16. Wang, H., Wang, Z., Sun, G., et al.: Design of a third-order filter with 10 GHz center frequency. In: The 9th International Conference on Communications, Signal Processing, and Systems **654**, 225–233 (2020)
17. Singh, R., Mondal, S., Paramesh, J.: A millimeter-wave receiver using a wideband low-noise amplifier with one-port coupled resonator loads. IEEE Trans. Microw. Theory Tech. **68**(9), 3794–3803 (2020)
18. Fang, J., Zhang, C.M., Singor, F.W., et al.: A broadband CMOS RF front end for direct sampling satellite receivers. IEEE J. Solid-State Circuits **54**(8), 2140–2148 (2019)
19. Nawaz, A.A., Albrecht, J.D., Cagri Ulusoy, A.: A Ka/V band-switchable LNA with 2.8/3.4 dB noise figure. IEEE Microwave Wirel. Compon. Lett. **29**(10), 662–664 (2019)
20. Nasrollahpour, M., Romano, A., Zaeimbashi, M., et al.: Integration of a novel CMOS-compatible magnetoelectric antenna with a low-noise amplifier and a tunable input matching. Analog Integr. Circ. Sig. Process **105**(3), 407–415 (2020)
21. Chang, C.H., Onabajo, M.: Analysis and demonstration of an IIP3 improvement technique for low-power RF low-noise amplifiers. IEEE Trans. Circuits Syst. I Regul. Pap. **65**(3), 859–869 (2018)

Cross-Modal Two-Stream Target Focused Network for Video Anomaly Detection

Hong Ma, Zhe Sun, Yong Su(✉), Hongrui Wang(✉), Shunan Li, Zeyu Yu, Yajie Kang, and Hengpeng Xu

College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
suyong@tju.edu.cn, 1930310007@stu.tjnu.edu.cn

Abstract. Video anomaly detection (VAD) is a challenging computer vision task that aims at identifying the events that do not conform to expected behaviors in a video. Training an unsupervised learning model on normal data is one of the most prevalent VAD paradigms owning to rare abnormal snippets. However, anomaly related and irrelated information sharing feature space will degrade the representation ability of the model for abnormal events. In this article, we propose a novel unsupervised video anomaly detection method, named Cross-Modal Two-Stream Target Focused Network (CMTSTFN) to focus on the anomaly related target and alleviate the impact of anomaly irrelated information on the model. Firstly, we use the global-local optical flow as a priori knowledge to blur the anomaly irrelated information in the image domain by Gaussian blur kernel. More importantly, conditioned on the blurred video frame, we then employ a dual cross-modal attention module to reconstruct the fused image feature map and optical flow feature map, respectively. Therefore, both spatial and temporal cues of abnormal events could further deteriorate the quality of the final reconstructed frame, making the anomalies more detectable. Experimental results on UCSD Pred1 and UCSD Pred2 demonstrate the effectiveness of the proposed CMTSTFN.

1 Introduction

Video anomaly detection (VAD) [3,8,13] is an issue of great significance and challenges in computer vision and machine intelligence. Different from other detection tasks, a side effect of data imbalance would inevitably occur in the VAD task, since the amounts of normal and abnormal data are hugely imbalanced in real-life situations [16].

To overcome this issue, the most widely used VAD paradigm is training an unsupervised model describing normality without abnormal samples [2,4,8]. Specifically, a majority of recent video anomaly detection approaches utilize deep reconstruction models, which assume that a frame would be detected as an anomaly with a high reconstruction error.

H. Ma and Z. Sun—Contribute equally to this work.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 69–78, 2023.
https://doi.org/10.1007/978-981-99-2653-4_9

Fig. 1. The anomaly irrelated pixels are marked in green

However, reconstruction methods can not always guarantee a significant reconstruction error for anomalies, which means that both the normal and anomaly events may be well reconstructed. The key reason is that the anomaly-related and irrelated information always shares the common feature space, which would degrade the representation and discrimination ability for abnormal events. As shown in Fig. 1, in many videos, the volume of anomaly irrelated pixels, such as the background, is much larger than that of anomaly-related pixels, which will result in background loss playing a key role in model optimization. Even worse, the high-frequency feature of anomaly irrelated information will impose a high computational burden on the model.

To tackle the above challenges, we propose a novel unsupervised video anomaly detection method, named Cross-Modal Two-Stream Target Focused Network (CMTSTFN) to focus on the anomaly-related target and alleviate the impact of anomaly irrelated information on the model. Considering anomaly-related information of a given scenario, the proposed CMTSTFN use the global-local optical flow as a priori knowledge, to blur the anomaly irrelated pixels in the image domain by Gaussian blur kernel. Conditioned on the blurred video clip, we build two independent generators to encode latent features of the video clip and the optical flow [12], respectively. To model the explicit relationships between different models, we design a dual cross-modal attention module, to simultaneously search for discriminative appearance and motion features for anomaly-related targets. Finally, the input video frames and optical flow are reconstructed by two independent decoders, respectively. We test the proposed CMTSTFN on various public anomaly detection datasets. Extensive experiments prove the excellent generalization and high effectiveness of CMTSTFN.

The main contributions of this work are threefold: (1) We explore the optical flow as a priori knowledge and Gaussian blur kernel to make the CMTSTFN more focused on the anomaly-related information. (2) We design a dual cross-modal attention module for modeling relationships between appearance

and motion features. In this way, both spatial and temporal cues of abnormal events could further deteriorate the quality of the final reconstructed frame, making the anomalies more detectable. (3) The proposed CMTSTFN reduces the amount of anomaly irrelevant information to a great extent, thus alleviating the computational burden of the model in the training process, and the convergence speed of the model can be competitive with the state-of-the-art.

2 Methodology

2.1 Overview

As shown in Fig. 2, the proposed CMTSTFN model consists of three components: a preprocessing network (for blurring the anomaly irrelated pixels in the image domain and generating optical flow maps of the blurred video frames), a dual cross-modal attention module (for encoding latent features of the video clip and the optical flow and model the explicit relationships between the two modals) and a reconstruction module (with two independent decoders to reconstruct the input video frames and optical flow, respectively) In the training stage, the encoder and decoder are optimized to minimize the reconstruction error. The memory contents [5] are also updated to record the prototypical elements of the encoded data.

2.2 Global-Local Optical Flow

The traditional optical flow method [1] is to calculate the optical flow between two video frames. However, it only considers motion information between two frames, rather than the motion information of the video sequence, which lead to the blurring of some anomaly related information.

To tackle the above challenges, we propose the global-local optical flow. Given a video sequence $i(i = 1, 2, 3...N)$, each video has M frames. We use the trained PWC-NET [14] to calculate the optical flow between frame a and frame $b(speed_{ab})$. The local optical flow formula of each frame is as follows:

$$\hat{speed} = \sum_{a=j+1}^{x} speed_{js}/(x-1)(j = 1, 2, ..., M) \tag{1}$$

The average optical flow calculation of each video is as follows:

$$speed_i = \sum_{j=1}^{M} \hat{speed}_j/M(i = 1, 2, ..., N) \tag{2}$$

The global optical flow calculation formula is as follows:

$$SPEED = \sum_{i=1}^{N} seed_i/N \tag{3}$$

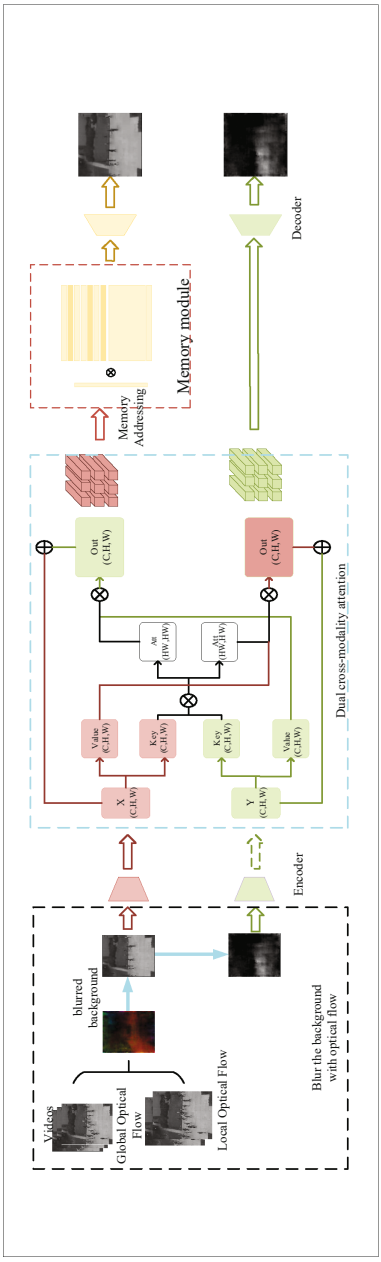


Fig. 2. The structure of the CMTSTFN

To calculate Mask ,we set λ as the threshold. Local Mask :

$$mask_{ij} = \begin{cases} 1 \; if \hat{speed}_{ij} < \lambda \\ 0 \; else \end{cases} \tag{4}$$

Global Mask:

$$mask_{ij} = \begin{cases} 1 \; if SPEED < \lambda \\ 0 \; else \end{cases} \tag{5}$$

If $mask_{ij}$ and $MASK_{ij}$==0, the pixel is abnormally unrelated. We convert it to blur background exploiting the Gaussian filter.

2.3 The Dual Cross-Modality Attention

Attention algorithms [15] have a wide range of applications in the field of computer vision. Inspired by some related works, we apply the dual cross-modality attention module (DCMA) [10] to model the interaction between the extracted appearance and motion feature maps.

Denoting input features extracted from the blurred video frames and the optical flow maps as T and T_h, respectively, the DCMA module convert them to T' and T'_h:

$$T', T'_h = f_{DCMA}(T, T_h) \tag{6}$$

Taking the conversion of T as an example, we first convert the input $T \in R^{C \times H \times W}$ into value component $V \in R^{C \times H \times W}$ which represents the domain-specific information, and the key component $K \in R^{C/r \times H \times W}$ which measures the correlation between two different domains. r is a scalar to reduce K's channel dimension for improving the computation efficiency of the model. Similarly, the optical flow features $T_h \in R^{C \times H \times W}$ are converted into two components V_h and the K_h. Then we measure the correlation $C \in R^{HW \times HW}$ between the two modalities through:

$$C = flt(K)^T \bigotimes flt(K_h) \tag{7}$$

and multiply the correlation C with weight matrix W to generate the attention map A through the softmax operation which is formulated as:

$$A = softmax(C \bigotimes W) \tag{8}$$

where $flt(\cdot)$ denotes the flatten operation and $\bigotimes$ denotes matrix multiplication. Similarly, we obtain the optical flow attention map:

$$A_h = softmax(C^T \bigotimes W_h) \tag{9}$$

Finally, we apply A to the corresponding value component V_h to obtain the refined features R which is formulated as:

$$R = flt(V_h) \bigotimes A \tag{10}$$

We then recover R to the input dimension, and output the fusion features T' by:

$$T' = T + R \tag{11}$$

The calculation for features T'_h in the optical flow domain is the same.

2.4 Encoder and Decoder

The encoder is used to obtain the latent encoding from the input while the decoder is to reconstruct the samples by taking the latent encoding as input. Given an input sample $x \in X$ which represents the domain of the data sample, the encoder converts it to an encoded representation as $z \in Z$ which represents the domain of the data sample, and the decoder is trained to reconstruct the latent representation z to the domain X:

$$z = f_e(x; \theta_e) \tag{12}$$

$$\hat{x} = f_d(z; \theta_d) \tag{13}$$

where θ_e and θ_d denote the parameters of the encoder $fe(\cdot)$ and decoder $fd(\cdot)$, respectively.

2.5 Memory Module

Memory-Based Representation. To improve the discriminative and representative ability of the module, we applied the memory module to our work. Given a latent encoding $z \in R^C$. The memory network obtains $\hat{z}$ relying on a soft addressing vector $w \in R^{1\times N}$ as follows:

$$\hat{z} = wM = \sum_{i=1}^{N} w_i m_i \tag{14}$$

where the weight vector w is computed according to z, and the matrix $M \in R^{N\times C}$ which consists of the row vector m_i contains N real-valued vectors of fixed dimension C.

Attention for Memory Addressing. The MemAE define the memory as a content addressable memory with an addressing scheme that computes attention weights w which is formulated as follows:

$$w_i = \frac{exp(d(Z, m_i))}{\sum_{j=1}^{N} exp(d(z, m_j))} \tag{15}$$

where $d(\Delta, \Delta)$ denotes a cosine similarity measurement:

$$d(z, m_i) = \frac{z m_i^T}{||z||||m_i||} \tag{16}$$

Hard Shrinkage for Sparse Addressing. The MemAE apply a hard shrinkage operation to promote the sparsity of w using the continuous ReLU activation function as:

$$\hat{w}_i = \frac{max(w_i - \lambda, 0) \cdot w_i}{|w_i - \lambda| + \epsilon} \tag{17}$$

where λ is a very small positive scalar. After that, we re-normalize the processed weights $\hat{w}_i$:

$$\hat{w}_i = \frac{\hat{w}_i}{||\hat{w}||_1} \tag{18}$$

Finally, we can get the reconstruction formula with the Memory module:

$$\hat{z} = \hat{x}M \tag{19}$$

2.6 Training

Firstly we obtain the loss function of VAE. Given a dataset $\{x^t\}_{t=1}^T$ which contains T samples. We use the $l2 - norm$ to measure the reconstruction error on each sample as follows:

$$L_1 = \frac{1}{T}\sum_{t=1}^{T}(||x_1^t - \hat{x}_1^t||_2^2) \tag{20}$$

where the $\hat{x}_1^t$ denote the reconstructed sample of the training sample x_1^t from the optical flow domain.

Nextly we construct the loss function of MemAE as:

$$L_2 = \frac{1}{T}\sum_{t=1}^{T}(||x_2^t - \hat{x}_2^t||_2^2 + \alpha E(\hat{w}^t)) \tag{21}$$

where α is a hyper-parameter in training and the $\hat{x}_2^t$ denote the reconstructed sample of the training sample x_2^t from the blurred frames domain.

By combining the loss functions above, we finally construct the training objective for CMTSTFN as:

$$L(\theta_e, \theta_d, M) = L_1 + L_2 \tag{22}$$

During training, the memory M and the paraments of both the encoders and decoders are all updated through optimization via back propagation and gradient descent.

3 Experiment

3.1 Implementation Details

For the applied datasets, we reshape each video frame to the size of 256×256 and take 16 frames of which as the input to the model. The encoder and decoder are set as follows: Conv2(3, 2, 64)-Conv2(3, 2, 128)-Conv2(3, 2, 128)-Conv2(3, 2, 256) and Dconv2(3, 2, 128)-Dconv2(3, 2, 128)-Dconv2(3, 2, 64)-Dconv2(3, 2, 3). After obtaining the reconstructed images, we use the MSE Loss to calculate its reconstruction error, and cross-entropy to calculate the memory loss, where

the Entropy Loss Weight is set to 0.0002. We used the Adam optimizer to train the CMTSTFN model on PyTorch with the learning rate set to 0.00001.

We evaluate our method (CMTSTFN) on two widely used datasets. The UCSD datasets are shot by a fixed camera overlooking the sidewalk, and it has two anomaly datasets, UCSDped1 and UCSDped2. The UCSDped1 is a video clip where pedestrians walk away from and into the scene. There are 34 video clips in the training set and 36 video clips in the test set. The UCSDped2 is a video clip where pedestrians walk parallel to the camera scene. The training set has 16 video clips and the test set has 12 video clips. The main anomalies included cycling, skateboarding, and driving on the sidewalk. We use the PWC-NET network to process the input video frame into a background blurred image and obtain its optical flow map at the same time.

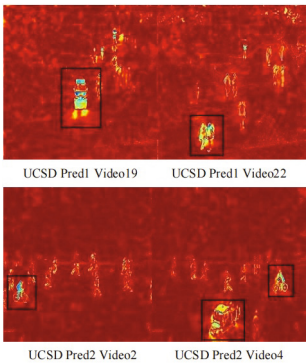


Fig. 3. The heat maps of the abnormal video frames

3.2 Comparison with Previous Methods

Table 1 illustrates the comparison of per class means the accuracy of our proposed CMTSTFN with other models. The results show that our method could achieve competitive results under the same protocol. Moreover, we obtain the reconstruction error by computing the gray value difference between the pixels of the original image and the corresponding pixels of the reconstructed image. After that, we use the reconstruction error to generate a heat map, where the region corresponding to the larger reconstruction error is brighter, while the region corresponding to the smaller reconstruction error is darker and appears dark red. As shown in Fig. 3, it is obvious that the abnormal parts such as bicycles, and cars are brighter than other normal regions. Therefore, compared with the discriminant model, our model can detect anomalies in each pixel.

Table 1. Comparison with other methods

Module	UCSD Pred 1	UCSD Pred 2
MIPPCA [7]	0.590	0.693
MIPPC+SFA [11]	0.668	0.613
MDT [11]	0.818	0.829
Conv-AE [6]	0.750	0.850
ConvLSTM-AE [9]	0.755	0.881
CMTSTFN	0.708	0.936

3.3 Ablation Studies

Table 2 shows the performance of our module in different settings. Compared with the CMTSTFN, the input video frame of CMTSTFN(B) is not blurred. The difference in accuracy between them indicates that blurring the irrelated information is crucial for VAD. And the input of the MEMAE(A) is only the blurred video frame while the input of the AE(B) is only the optical flow map. The comparison shows that both spatial and temporal cues of abnormal events could further deteriorate the quality of the final reconstructed frame.

Table 2. Ablation studies

Method	Auc	Input
CMTSTFN(B)	0.929	the video frames and the optical flow maps
MemAE(A)	0.914	the blurred viedo frames
AE(B)	0.336	the optical flow maps
CMTSTFN	0.936	the blurred viedo frames and the optical flow maps

4 Conclusion

In this paper, we proposed a Cross-Modal Two-Stream Target Focused Network (CMTSTFN) for video anomaly detection (VAD). We use the global-local optical flow as a priori knowledge to blur the anomaly irrelated information and employ a dual cross-modal attention module to reconstruct the fused image feature map and optical flow feature map, respectively. In this way, both spatial and temporal cues of abnormal events could further deteriorate the quality of the final reconstructed frame, making the anomalies more detectable.

Acknowledgement. This work was supported in part by the Doctoral Foundation of Tianjin Normal University under Grant No. 52XB2102.

References

1. Beauchemin, S.S., Barron, J.L.: The computation of optical flow. ACM Comput. Surv. (CSUR) **27**(3), 433–466 (1995)
2. Bergmann, P., Fauser, M., Sattlegger, D., Steger, C.: Mvtec ad-a comprehensive real-world dataset for unsupervised anomaly detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9592–9600 (2019)
3. Chalapathy, R., Chawla, S.: Deep learning for anomaly detection: A survey. arXiv preprint arXiv:1901.03407 (2019)
4. Eskin, E., Arnold, A., Prerau, M., Portnoy, L., Stolfo, S.: A geometric framework for unsupervised anomaly detection. In: Applications of Data Mining in Computer Security, pp. 77–101 (2002)
5. Gong, D., et al.: Memorizing normality to detect anomaly: memory-augmented deep autoencoder for unsupervised anomaly detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1705–1714 (2019)
6. Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A.K., Davis, L.S.: Learning temporal regularity in video sequences. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 733–742 (2016)
7. Kim, J., Grauman, K.: Observe locally, infer globally: a space-time mrf for detecting abnormal activities with incremental updates. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2921–2928 (2009)
8. Kiran, B.R., Thomas, D.M., Parakkal, R.: An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. J. Imaging **4**(2), 36 (2018)
9. Luo, W., Liu, W., Gao, S.: Remembering history with convolutional lstm for anomaly detection. In: 2017 IEEE International Conference on Multimedia and Expo (ICME), pp. 439–444 (2017)
10. Luo, Y., Zhang, Y., Yan, J., Liu, W.: Generalizing face forgery detection with high-frequency features. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 16317–16326 (2021)
11. Mahadevan, V., Li, W., Bhalodia, V., Vasconcelos, N.: Anomaly detection in crowded scenes. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1975–1981 (2010)
12. Maurer, D., Stoll, M., Bruhn, A.: Order-adaptive regularisation for variational optical flow: Global, local and in between. In: International Conference on Scale Space and Variational Methods in Computer Vision, pp. 550–562 (2017)
13. Mohammadi, B., Fathy, M., Sabokrou, M.: Image/video deep anomaly detection: a survey. arXiv preprint arXiv:2103.01739 (2021)
14. Sun, D., Yang, X., Liu, M.-Y., Kautz, J.: Pwc-net: Cnns for optical flow using pyramid, warping, and cost volume. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8934–8943 (2018)
15. Zhai, Y., Shah, M.: Visual attention detection in video sequences using spatiotemporal cues. In: Proceedings of the 14th ACM International Conference on Multimedia, pp. 815–824 (2006)
16. Zong, B., et al.: Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In: International Conference on Learning Representations (2018)

Fall Detection in the Wild: An Intelligent Emergency Assistance System

Yang Li[1], Ge Guo[1(✉)], Jixin Li[1], Yu Sun[1], Gaoming Ren[1], Lei He[1], and Yong Su[2]

[1] Shaanxi Institute of Technology, Shaanxi, China
jianjianbeng@qq.com
[2] Tianjin Normal University, Tianjin, China

Abstract. Recent years have witnessed increasing attention to fall detection, powered by the strong demands of *Intelligent Emergency Assistance Systems* (IEAS), as it is crucial to reduce incidents of serious injuries caused by falls. In this paper, we propose an IEAS, a framework that integrates deep learning-based fall detection and edge computing-based local processing. First, we introduce a novel challenging benchmark dataset, called WildFall, to better evaluate fall detection models in real-world environments. The WildFall consists of 2,000 various types of video clips, involving movies, TV series, and surveillance. Then, the conceptual design, modules, and practical implementation of IEAS are described in detail. Finally, we discuss and analyze the extensibility of IEAS and future work. We believe that both the proposed IEAS and the public availability of WildFall will attract more research attention, to provide timely emergency assistance for the injured.

1 Introduction

Falls are a common but often overlooked, cause of injury, which may end in disability or even death in a split second. Falls may occur indoors or outdoors (Fig. 1(a)), most falls may not cause serious consequences, but older people are more vulnerable, especially if they have a long-term health conditions. The impact of falls on older people is not limited to psychological problems, abrasions, fractures, but also injury-related death. Every year, more than 8 million people (including 2.8 million elderly people) are treated in emergency rooms for fall-related injuries[1]. In particular, for people over 85 years of age, falls are the leading cause of about two-thirds of all reported injury-related deaths. This situation will continue to worsen since the number of people over the age of 60 may increase from the current 1 billion to 1.2 billion by 2025.

As a developing country, many elderly people in China lack necessary care. We sent 600 questionnaires to participants covering 19 provinces in China and got 400 valid questionnaires (203 male,197 female) aged from 55 to 97[2] As shown

[1] https://www.cdc.gov/homeandrecreationalsafety/falls.
[2] https://github.com/usodonkey/Hci-group.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 79–85, 2023.
https://doi.org/10.1007/978-981-99-2653-4_10

in Fig. 1(b), the statistical results show that nearly half of the participants have experienced a fall within three years, and the probability of being injured under the age of 60 is the only 9.5%, while the probability that being injured over the age of 70 is an impressive 66.23%.

Moreover, for participants with a fall experience, the probability of people over 70 being injured indoors is 43%, and the probability of people under 60 is only 9.5%. Obviously, for the elderly without supervision, the consequences of injuries indoors are more serious than those outdoors. In addition, we found that for elderly people, falls have not only caused physical injuries but also produced a huge medical cost for fall-related injuries. The average medical expenses reached 22,491 RMB (3473 USD).

The above shocking numbers prompted researchers to design applications, algorithms, and wearable devices to detect falls for timely rescue [1–3]. *However, most methods are inefficient in common real-world environments, let alone harsh environments where falls are more likely to occur.* The major challenges are summarized as follows:

Lack of public large-scale real-world fall dataset: Most of the fall datasets are collected in constrained environments. The fall video clips captured in the laboratory could not match well real-world environments, which limits the generalization ability, predictive precision, and efficiency of existing models trained on these datasets. In a real-world surveillance network, the diversity of lighting, backgrounds, resolutions, and views of video data will obviously have an adverse effect on deep models. Many researchers have found that changing background would greatly affect the overall performance of the vision-based fall detection systems. Moreover, when encountering harsh environments such as extreme weather and excessive dust, the performance of vision-based systems will deteriorate drastically. Although many fall detection methods can achieve high performance in MobiAct, SisFall, MobiFall, UniMiB SHAR, and UP-Fall datasets, RGB-Depth, radar, or accelerometer sensors used in these datasets are constrained by many factors such as price, popularity, and privacy sensitivity [4].

The vision-based approach seems more feasible, but the bottleneck is also obvious: as fall events typically occur at a much lower probability, it is usually difficult to collect sufficient fall clips in real-world environments. Manually screening and extracting fall clips from surveillance videos requires massive labor costs. Therefore, it remains a very challenging problem due to the serious imbalance between normal and fall samples, and the lack of fine-grained fall labeling data.

High transmission, computation burden and Latency: Generally, an intelligent sensor system consists of hundreds of sensors. Developing a deep learning-based intelligent sensing system requires a significant amount of storage to store large media files and powerful computational devices. Therefore, computational cost, storage, and long transmission distance are also important issues that need to be considered. Traditional intelligent sensor systems rely on a centralized server to process large-scale distributed sensing data. The City-scale sensor network may produce a huge amount of data every minute, most of them are redundant and do not contain fall events. If the centralized server is far away from the sensors,

and sensors cannot provide the computational capabilities for further intelligent operations, the backbone network with limited capacity would inevitably lead to unpredictable delays. Therefore, for latency-sensitive tasks, how to reduce traffic and transmission delay is a key challenge. Simply increasing the number of servers to improve computing power not only can not fundamentally solve the problem of transmission delay but also needs more cost.

Accordingly, the above-mentioned challenges motivate us to collect a fall dataset in real-world environments and design an intelligent emergency assistance architecture to improve the reliability of fall detection.

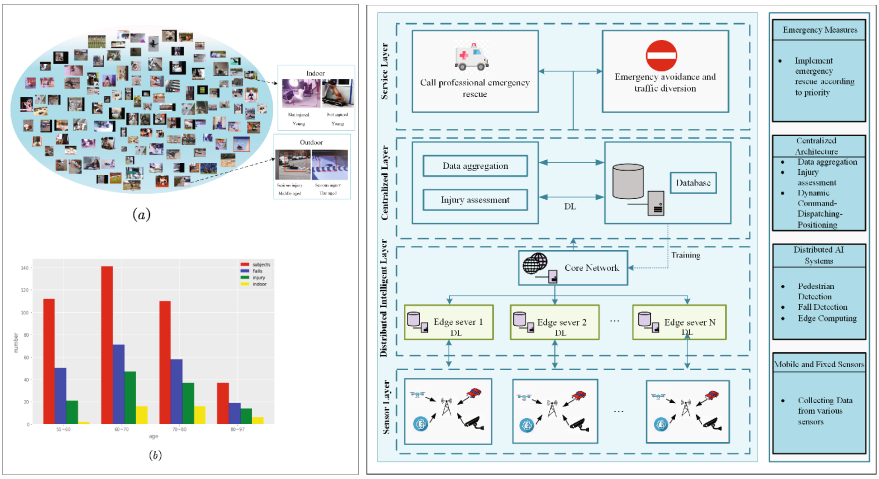


Fig. 1. WildFall and Questionnaire statistics. Left:(a) WildFall is a large-scale benchmark for general multi-label fall detection. This screenshot of a dataset depicts a subset of videos in the dataset labeled with the fall entity.(b) questionnaires statistics aged from 55 to 97. Right: The IEAS architecture.

2 The WildFall Dataset

Various sensor devices, such as accelerometers, gyroscopes, RGB cameras, and radars, have been used to detect falling events. Among them, computer vision technology is the most straight and valid to detecting falls. However, vision-based fall detection has not received much attention due to the lack of large-scale fall datasets. Up until recently, the vast majority of computer vision research tackles fall detection or fall recognition as a subproblem of action recognition. In that case, the task is reduced to classifying the sequence into one of the relevant categories of the action. The action recognition task attempts to find a model that has low loss, on average, across all examples, which is called empirical risk minimization. However, one of the challenging factors of the multi-class classification model for fall detection is the data imbalance, meaning that the

fall events are more difficult to capture than other regular events because of their scarcity in the real world. The above phenomena can be found in many motion recognition models, as the confusion matrix in [5] shows, fall instances are not only easily confused with their similar actions, such as staggering, but also misclassified into many categories with large semantic gap such as reading, tear up the paper and torch neck.

To this end, we introduce WildFall[3] a large-scale benchmark dataset with reasonable quality labels for multi-label fall detection. We gathered 50 trained volunteers for data collection. We assigned volunteers to 12 multimedia video platforms[4] to search for videos through given topics. After collecting the video, we extracted clips with complete fall events, and the maximum length of video clips is 8 s. To facilitate model training, in the pre-processing stage, we try to make the starting position of the target in the center of the image. In total, the WildFall dataset contains more than 2K video clips-over 269 min. For each fall video clip, the age range, scene (indoor or outdoor), and degree of injury are estimated and labeled. The multi-labeled data is very important to evaluate fall-related injuries, which could give the IEAS an initial warning priority. In addition to the fall video clips, we also collected five other challenging actions in real-world environments, including walking, jumping, going up, going downstairs, and creeping forward. Compared with the previous fall dataset collected in the laboratory scene, the proposed WildFall dataset is more realistic and challenging. The diversity of datasets is a necessary condition for the development of an intelligent emergency rescue system. We believe that a large scale and diverse dataset is a necessary condition for the development of an intelligent emergency rescue system.

3 Intelligent Emergency Assistance System and Implementation

Figure 1 shows the proposed IEAS architecture, which consists of four different layers: sensor layer, distributed intelligence platform, centralized layer, and service layer.

- The sensor layer is composed of carry-on smart devices (including mobile phones, UAVs, tachographs, etc.), wired or wireless security camera devices, and network elements, which are used to generate heterogeneous sensing data from various sensing devices and then transfer to edge computing servers.
- Distributed intelligent layer, as the core layer of IEAS, provides computing, storage, and communication resources between sensor devices and centralized mediation layers. This layer reflects the intelligent characteristics of the system, including deep learning-based pedestrian detection, fall detection, and edge processing.

[3] The raw labeled data is publicly available and can be achieved from https://pan.baidu.com/s/1SRE4aD9iNnrmJh7AybBD7w.

[4] YouTube, TikTok, Tencent, iQiyi, Bilibili, Youku, Weibo, TuDou, Kuaishou, Ku6, Mango, RED Video.

- The centralized layer includes: a cloud computing data center for aggregating perceptual data (video clips) from distributed intelligent layer; A database server that maintains historical sensory data, which can be used for incremental training of deep networks; Further, assess the injury based on the fall data provided by the distributed intelligence layer; If the sensing data comes from the mobile platform and the user cannot provide precise positioning, the centralized layer needs to request permission and use the wireless positioning technology based on the cellular network.
- The service layer includes receiving and controlling a series of emergency requests submitted by the centralized layer, such as emergency medical rescue, emergency traffic control, etc.

In crowded scenes, passers-by can directly call or implement rescue. Therefore, an intelligent emergency rescue system is more suitable for sparsely populated or isolated scenes, where fall events are always hard to be detected, and untimely rescue will aggravate the injury or even lead to death.

4 Numerical Results

In this section, a fall detection case will be conducted on the WildFall dataset to evaluate the performance of the IEAS framework quantitatively. Firstly, in the pedestrian detection stage, we use Fast R-CNN as the backbone network to detect and estimate the bounding box of pedestrians and then rescale the image according to the size of the bounding box. Then, we use the OpenPose toolbox [6] to estimate the pose composing 18 key points for each person in the scaled images. Given the pedestrian skeleton data, we need to design a deep network that can identify fall events. As we all know, most CNN with local receptive fields may limit the ability to model the spatio-temporal context. Although RNNs can capture global context information, they are not hardware friendly requiring more resources to train. Therefore, we propose a deep network [7], named L2C (local to context) self-attention augmented model, which aims to aggregate spatial-temporal saliency of joints and frame dependencies from local-to-global.

A raw skeleton sequence can be expressed as $X \in R^{C \times J \times F}$, where C, J, F denote the number of channels, joints, frames, separately. The coordinates of a joint in the $f-th$ frame can be denoted as $(x_{j,f}, y_{j,f}, z_{j,f})$. The overall pipeline of the proposed model is shown in Fig. 2. The proposed whole L2C self-attention augmented model contains three self-attention augmented modules including the spatial convolution operator for encoding local information and the self-attention operator for encoding context information. The spatial convolution operator can obtain spatial features $X_{Local} \in R^{C_2 \times J \times F}$ by aggregating local joint information. The self-attention mechanism models the global dependencies, by directly attending to all the positions in an input sequence, which can be utilized to compute the correlation between arbitrary joints of the input sequence. The input of the self-attention mechanism consists of queries $Q \in R^{C_k}$, keys $K \in R^{C_k}$ and values $V \in R^{C_v}$. $C_3 = C_v + 2C_k$ is the number of output channels of the

previous embedding module. Then a softmax function is applied to the dot products of queries with keys and obtains a weights matrix on the values.

The attention weights here represent the correlation of each dimension's features, which can capture long-range dependencies by explicitly attending to all the joints. In order to further complement the localness modeling and enhance the generalization ability of the model, we also use the multi-head mechanism in the self-attention block. After multi-head self-attention, we use the 1×1 convolution operator to encode the spatio-temporal features and then stack them with the output of the local convolution operator. In the experiment, the number of neurons (C_2) in the convolution operator is set to 68, and the embedding dimensional $c3 = 100$; In multi-head self-attention module, the number of heads is set to 4, and the dimensional value, key, and a query is set to $C_v = 60$, $C_k = 20$, respectively. Before the classifier, we added the max-pooling layer to select better features for classification. We randomly split the hyperedge set into training and test sets by a ratio of $4 : 1$. In addition to skeleton features, we use I3D to capture RGB features in videos. Finally, the skeleton feature and RGB feature are fused to produce the final result.

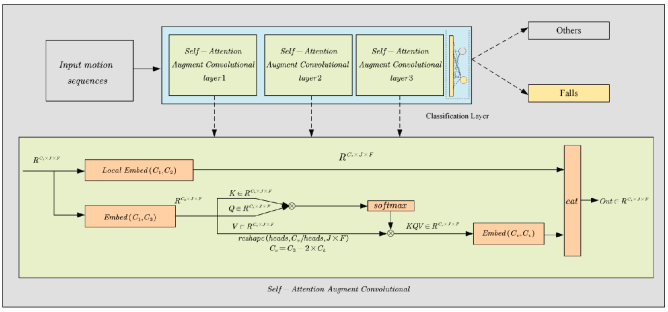


Fig. 2. The architecture of L2C self-attention augmented model.

We first tested the proposed L2C self-attention augmented model on the NTU RGB+D dataset [8], and the detection accuracy of fall events (946 samples) can reach 96.4%. Next, we evaluate the I3D, proposed model, and Multi-modal fusion model (Pose+I3D) on the WildFall dataset. The results are shown in Table 1, which illustrates that in real-world environments, due to the influence of sensor resolution, backgrounds, illuminations, and other factors, pose features perform much better than RGB features, and feature fusion strategy can further improve the detection performance, but it still exists a certain gap with the results obtained from laboratory data. In particular, the injury estimation module of the centralized layer can also be implemented using a model with the same structure.

Table 1. The fall events detection results on WildFall and NTU RGB+D dataset.

Method	WildFall	Fall (NTU)
Pose	72.8(182/250)	96.4
I3D	59.2(148/250)	-
Pose+I3D	74.8(187/250)	-

5 Conclusion

In this paper, we propose the IEAS to intelligently detect fall events in sparsely populated areas. First, we introduce a novel challenging benchmark dataset collected from the internet. The dataset named WildFall consists of 2,000 videos captured in real-world environments. Then, an intelligent architecture is proposed, including pedestrian detection, deep representation learning, and local processing. Finally, numerical experimental results show that the framework can intelligently detect fall events in real-world environments.

References

1. Fáñez, M., Villar, J.R., de la Cal, E.A., Suárez, V.M.G., Sedano, J., Khojasteh, S.B.: Mixing user-centered and generalized models for fall detection. Neurocomputing **452**, 473–486 (2021)
2. Wang, Y., Yang, S., Li, F., Wu, Y., Wang, Y.: Fallviewer: a fine-grained indoor fall detection system with ubiquitous wi-fi devices. IEEE Internet Things J. **8**(15), 12455–12466 (2021)
3. Paolini, C., Soselia, D., Baweja, H., Sarkar, M.: Optimal location for fall detection edge inferencing. In: IEEE Global Communications Conference, GLOBECOM, pp. 1–6 (2019)
4. Casilari, E., Santoyo-Ramón, J.A., Cano-García, J.M.: Analysis of public datasets for wearable fall detection systems. Sensors **17**(7), 1513 (2017)
5. Jiang, Y., Xu, J., Zhang, T.: View-independent representation with frame interpolation method for skeleton-based human action recognition. Int. J. Mach. Learn. Cybern. **11**(12), 2625–2636 (2020)
6. Cao, Z., Hidalgo Martinez, G., Simon, T., Wei, S., Sheikh, Y.A.: Openpose: realtime multi-person 2d pose estimation using part affinity fields. IEEE Trans. Pattern Anal. Mach. Intell. **43**(1), 172–186 (2021)
7. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, pp. 5998–6008 (2017)
8. Shahroudy, A., Liu, J., Ng, T., Wang, G.: NTU RGB+D: a large scale dataset for 3d human activity analysis. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR, pp. 1010–1019 (2016)

Research on the Development Method of High Reliability Downhole Communication System

Jia-hui Liang[1], Jin Chen[1(✉)], Jing- han Li[1], Sen Li[1], Bao- zhu Han[1], and Han-ling Wu[2]

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

[2] Beijing Institute of Astronautical Systems Engineering, Beijing 100076, China

Abstract. In order to meet the needs of downhole operations in oilfields and solve the problems of many cables, long transmission distances, and high cost of wiring and acquisition equipment caused by information acquisition in production sites, this paper proposes a design scheme for oil well underground highly reliable communication. The system uses the mature signal processing and communication technology, and uses a direct current(DC) carrier to transmit data. This paper expounds the principles of signal conversion, modulation/demodulation, coupling and reception in the system. Finally, the whole communication system is tested by using the computer to receive the data collected and sent by the sensor through the serial port debugging assistant, and during the period, several important signals are observed on the oscilloscope. The results show that the measured signal waveform is in line with expectations, and the computer can receive the data accurately, which verifies the accuracy of the system.The system effectively improves the intelligence level of downhole operations in oilfields, and is conducive to promoting the development of a new era of digital oil extraction.

Keywords: Communication Systems · Modulation and Demodulation · DC Carrier Communication Technology · Frequency-Shift Keying (FSK)

1 Introduction

The communication mode of oil field underground has always been a key research problem in the oil well mining industry [1]. Raditional oilfield downhole communication is mainly divided into two categories: wireless and wired. Wireless communication methods such as electromagnetic waves and sound waves have the advantages of high reliability and long transmission distance. However, the transmission rate is slow, the attenuation is large, and there are many interference in the transmission process, which make it have certain limitations. Wired communication methods include cables, optical cables, which have the advantages of fast transmission rate and low attenuation, but this requires a large amount of physical cable consumption, and the transmission distance is limited. Power line carrier communication technology is a popular wired

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 86–94, 2023.
https://doi.org/10.1007/978-981-99-2653-4_11

communication technology in recent years. It refers to the use of existing power lines to transmit information. Its dual-purpose method can greatly simplify the transmission channel and reduce costs. Power line carrier communication technology has the characteristics of low transmission rate and small capacity, so it has great practical value for communication places that do not require high information transmission rate. Therefore, based on the current research on downhole communication in oilfields, this paper designs a high-reliability communication system based on DC power line carrier communication technology, using Field Programmable Gate Array(FPGA) as the hardware core to achieve efficient modulation/demodulation, and facing the underground high temperature environment.

2 The Composition of the Communication System

The DC carrier communication system mainly consists of six modules, as shown in Fig. 1. Data transmission is divided into two parts: Downhole transmission and surface transmission.

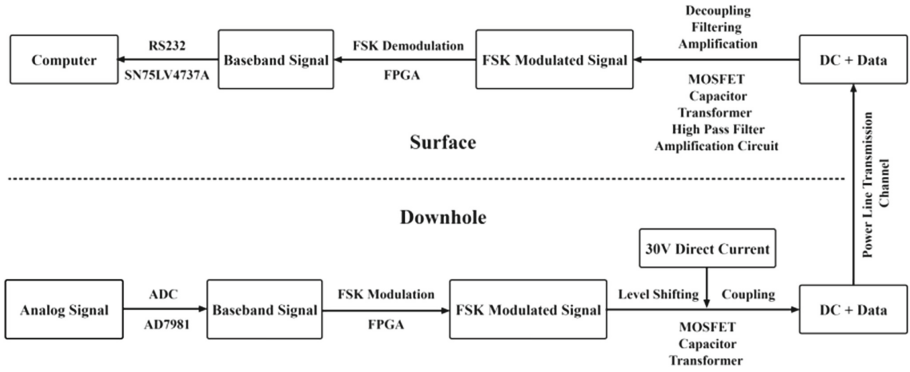


Fig. 1. DC Carrier Communication System

Downhole Transmission Part: First, the analog-to-digital converter(ADC) converts the original signal (temperature, pressure, flow) and other analog quantities into baseband signals, and then sends them to the FPGA for FSK modulation to generate FSK modulated signals. Finally, the FSK modulated signal is superimposed on the DC power line after the level conversion and coupling of the H bridge composed of MOSFET, the transformer and the coupling capacitor.

Surface Transmission Part: When the FSK modulated signal is transmitted to the ground with the DC power line, the first step is decoupling: Separate high-voltage DC and low-voltage AC signals. Then, in order to filter out the low frequency interference generated in the channel and improve the signal quality, the signal will pass through the filter circuit and the amplifier circuit. After signal conditioning, the FSK modulated signal will be transmitted to the FPGA, and the FPGA uses the zero-crossing detection technology to realize FSK demodulation to restore the baseband signal. Finally, it is

transmitted to the computer through the RS232 serial interface, so as to realize the reliable communication from the well to the ground.

3 Downhole Data Transmission

3.1 Analog-to-Digital Conversion Module

The downhole information measured by the sensor is usually an analog signal, which cannot be directly recognized by the processor and is not suitable for direct transmission. This requires the use of an ADC to convert it into a digital signal.

In the field of oil exploration, the special high temperature environment puts forward higher requirements for the ADC. The AD7981 used in this paper is a 16-bit high temperature precision ADC chip, which can work under extreme high temperature, and its rated operating temperature is $-55°C$ to $+175$ °C, the sampling rate can be up to 600KSPS. When the chip operates at 10KSPS, its power consumption is only 70 μW, which is very suitable for downhole applications. The hardware circuit is shown in Fig. 2. When the ADC module works, the analog signal collected by the sensor will be transmitted to the "IN-" after passing through the RC filter. The AD7981 chip performs data transmission with the FPGA through the Serial Peripheral Interface(SPI) bus, using a three-wire SPI mode with busy indication.

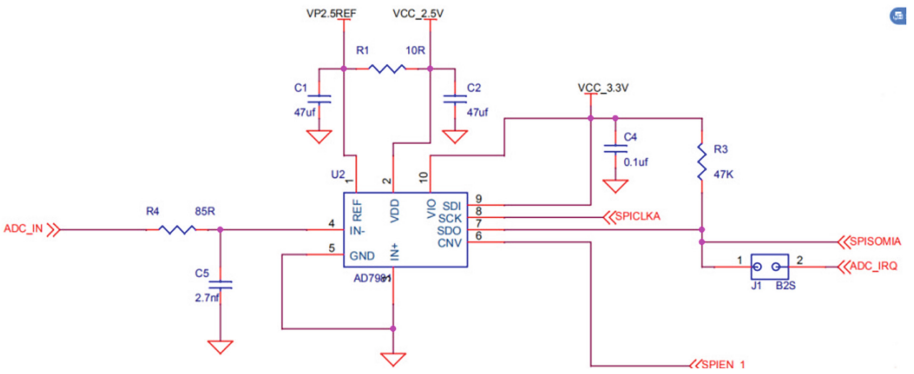


Fig. 2. ADC Module

3.2 FSK Modulation Module

FSK is a commonly used digital modulation method [2]. Because this method has the advantages of good anti-noise performance, long transmission distance and low bit error rate, it has been widely used in medium and low speed data transmission. With the rapid development of Electronic design automation technology, large-scale programmable logic devices have emerged. This paper adopts a method of using FPGA as the hardware core, employing VHDL language to design the system, and constructing the FSK modulator/demodulator model. This method has good portability of product upgrade.

The modulation principle of FSK is to gate two different independent frequency sources through a switch circuit under the control of a binary baseband pulse sequence, so that it can output a carrier signal with a frequency of f1 or f2 during each symbol period. The realization of this principle is shown in Fig. 3. In this paper, the system uses FPGA to divide the frequency of the external clock signal for obtaining carrier signals with frequencies f1 and f2, and then uses the baseband signal output by the ADC to control the two-to-one data selector to select different carrier signal outputs, which results in the FSK modulated signal.

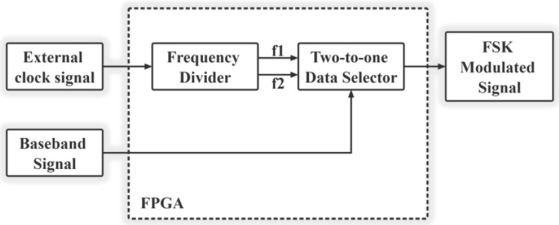


Fig. 3. FSK Modulation Module

3.3 Signal Coupling Module

The output stage of FSK modulation is composed of an H bridge composed of 4 MOSFET arrays, a current limiting resistor R48, a coupling transformer T, and a coupling capacitor C3, as shown in Fig. 4. The MOSFET arrays acts as a low impedance transformer to drive and isolate voltage transients. During system operation, voltage transients may be coupled from the power line to the transformer, and MOSFET arrays acts as saturated drive switches to minimize power dissipation.

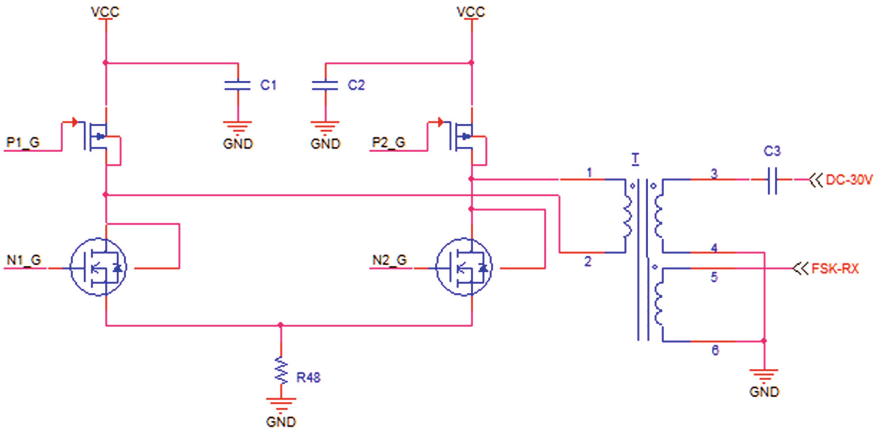


Fig. 4. Signal Coupling Module

The coupling technology in cable transmission is mainly divided into capacitive coupling, transformer coupling and composite coupling technology [3]. Capacitive coupling is a direct coupling, with high coupling efficiency, ideal transmission characteristics and low attenuation. However, it cannot couple the slowly changing carrier signal, as there will be phase distortion, which will affect the transmission signal. The biggest advantage of transformer coupling is that it can realize impedance transformation, has good insulation and high safety, but its transmission characteristics are relatively poor compared with capacitive coupling, and the working attenuation is about 2–4 dB larger than that of capacitive coupling. After comprehensively considering the characteristics of the above technologies, the composite coupling technology is used in the literature, which will achieve better results than a single coupling method. Figure 5 shows the effect of the FSK modulated signal superimposed on the DC power line, which can be measured at DC-30 V.

Fig. 5. The Effect after the Signal is superimposed with DC

4 Surface Data Transmission

4.1 Decoupling, Filtering and Amplification of Signals

Because the three-phase transformer used in the system can integrate the functions of sending, receiving and coupling in the same circuit structure, as shown in Fig. 4, the circuit structure used for the ground receiving signal is the same as that of the coupling module. This circuit can separate the high-voltage DC and low-voltage AC signals of the DC power line, and the obtained AC signal is transmitted to the filter circuit at the "FSK-RX".

Due to the poor frequency characteristics of the DC power line, large amplitude attenuation and serious interference, the separated FSK modulated signal needs to be filtered and amplified before being sent to the FPGA for demodulation. The signal conditioning mode diagram is shown in Fig. 6.

4.2 FSK Demodulation Module

In this system, the zero-crossing detection method is used to achieve demodulation. The principle is the number of times the signal passes through the zero point per unit time, which can be used to measure the frequency [4], as shown in Fig. 7. Therefore, the corresponding carrier frequency value can be obtained by detecting the number of zero-crossing points of the FSK modulated signal, and the baseband signal can be determined according to the known carrier frequency.

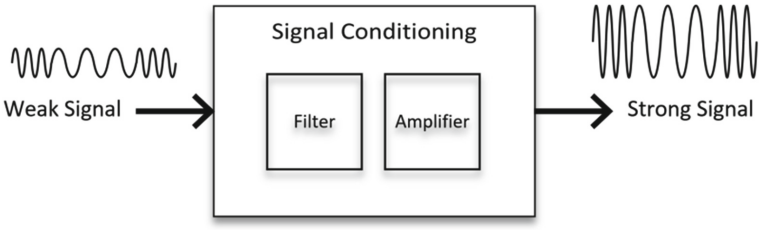


Fig. 6. Signal Conditioning Mode

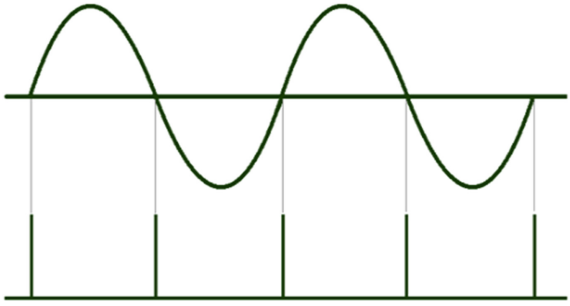


Fig. 7. Zero-crossing Detection

The specific demodulation process of FSK signal is shown in Fig. 8. The control signal is used as a synchronization signal to start the demodulation of the system. Since the number of zero crossings in one cycle of the FSK modulated signal needs to be calculated, the clock counter is used to calculate the period of the FSK modulated signal. If the ratio of the clock frequency to the frequency of the FSK modulated signal is 6:1, it means that an FSK modulated signal is sent every 6 clock cycles. The counter counts the number of zero-crossing points of the FSK modulated signal in one cycle. The role of the decider is to decide the value of the counter within one cycle of the modulated signal. If a FSK modulated signal symbol has n cycles of carrier f1 or m cycles of carrier f2 ($n < m$), the decision threshold x value can be selected as $n \leq x \leq m$. When the threshold of the value is $n < x < m$, the demodulator has a certain fault tolerance capability and can better demodulate the baseband signal.

4.3 RS232 Serial Communication Module

The computer communicates with external devices through the RS232 serial interface, which is a common communication solution in many systems [5]. In this paper, the SN75LV4737A level conversion chip adopts SSOP chip package, which is small in size and has an operating temperature range of −40 °C–+125 °C. It can realize the conversion from TTL to RS232 level standard, so as to realize the data transmission between the computer and FPGA.The SN75LV4737A serial communication circuit is shown in Fig. 9.

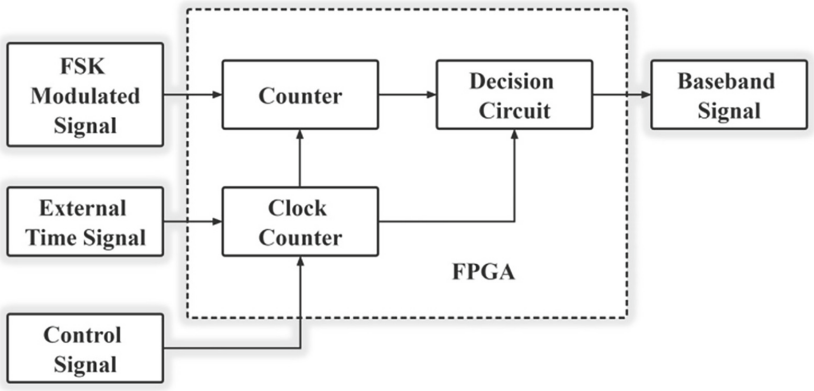


Fig. 8. FSK Demodulation Module

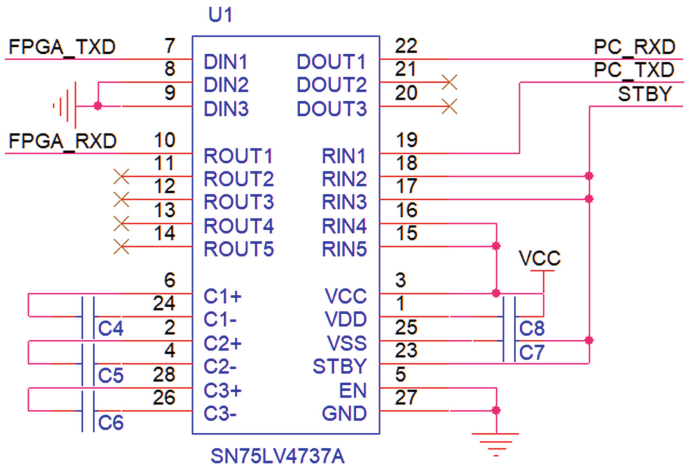


Fig. 9. RS232 Serial Communication Module

5 Experiments and Test

After the program is downloaded successfully, we started the power supply for testing. Figure 10 is the hardware test platform. The signal transmission path is as follows: the sensor starts to collect and return data, and the data is transmitted to the computer after ADC, modulation, channel transmission, filtering, amplification and demodulation. Figure 11 is the waveform transmitted in the channel, which is the result of the AC signal superimposed on the 30 V DC signal. In Fig. 12, the green waveform is the result of FSK modulation, and the orange waveform is the result of filtering at the receiving end. In this figure, the frequencies of the two signals are the same, and the signal waveform is not severely distorted, but only slightly decrease in amplitude, which is the inevitable result of signal loss during channel transmission. This experiment verifies the correctness and accuracy of the DC carrier system.

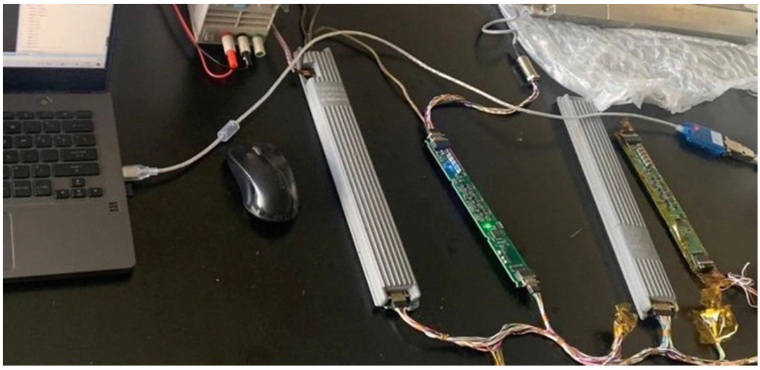


Fig. 10. Hardware Test Platform

Fig. 11. Signal Measured in the Channel

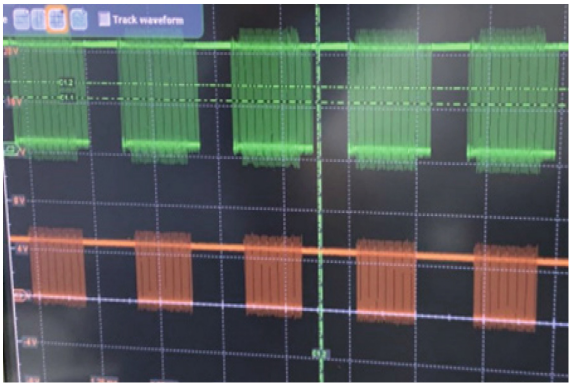


Fig. 12. Comparison before and after Signal Transmission

6 Summary

Oilfield downhole communication is a very important task in the process of efficient and safe oil exploitation. This paper designs a high-reliability communication system for oil wells, which is based on power line DC carrier technology.Then, this paper describes the working principle of the whole system, and verifies each functional module that constitutes the system. Finally, an experimental platform is built to simulate the communication process between the underground and the surface, and several key signals are observed. The experiment verifies the correctness of the system function.

Acknowledgment. The work was supported by the National Key Research and Development Plan Project: Robotic Systems for Agriculture(RS-Agri) NO. 2019YFE0125200.

References

1. Wait, J.R.: electromagnetic wave transmission in a well logging cable-theory. Microwave Theory Tech. **38**(10), 1518–1521 (1900)
2. Wu, Chang Fu (2013) Analysis, design and implementation of FSK modulate systems. Advanced Materials Research Vol. 765
3. Costa, daS., Guilherme, L.: Coupling for power line communications: a survey. J. Commun. Inf. Syst. **32**(1) (2017)
4. Lee, E.K.B.: Zero-crossing zero-IF demodulator. Electron. Lett. **31**(15), 1218–1219 (1995)
5. Han, X, Kong, X.: The designing of serial communication based on RS232. 2010 First ACIS International Symposium on Cryptography, and Network Security, Data Mining and Knowledge Discovery, E-Commerce and Its Applications, and Embedded Systems, IEEE (2010)

Research on Abnormal Heart Rate Detection Method of Athletes Based on Wireless Sensing Multi-source Information Fusion

Youchun Zhang(✉)

School of Application Engineering, Anhui Business and Technology College, Hefei 231131, China

zhangyouchun@ahbvc.edu.cn

Abstract. Most of the existing heart rate anomaly detection methods only use a single signal to determine the abnormal heart rate state of college athletes, which is easily disturbed, and when the actual detection environment is not so ideal, it is impossible to collect enough accurate data, which can lead to a large deviation in the test results. This paper proposes a heart rate anomaly detection method based on the fusion of random forest and multi-source information, and uses multiple miniature sensors to synchronously collect the breathing, heartbeat, pulse and grip signals of college athletes, when a signal acquisition is biased The other signal is not affected, and then all the collected signals are integrated to establish a multi-source fusion heart rate abnormal state dataset as the basis for detection, avoiding the defect that a single signal sometimes has a larger error, and the detection accuracy rate is higher. Compared with classification algorithms such as SVM and GBDT, random forests can balance the error of the data set, are insensitive to missing values, and the classification speed is fast, and the optimal detection accuracy of the model is 89.16%, which can effectively avoid abnormal heart rate of college athletes.

All along, the world has attached great importance to how to avoid abnormal heart rate, and more and more companies and scientific research institutions have designed various devices to avoid abnormal heart rate. The detection method of abnormal heart rate is generally divided into three methods: based on physiological signals, based on exercise information, and based on the faces of college athletes [3]. Physiological information detection method through the direct contact with the human body sensor to collect the physiological signals of college athletes to judge the abnormal state of heart rate in college athletes in sports, this method of obtaining reliable data, and physiological signals directly from the human body, accurate and reliable when analyzing abnormal heart rate states, but usually to install a device to detect physiological signals on college athletes, the interference of college athletes is larger. The Motion Information Detection Method detects abnormal heart rate states of college athletes based on signal characteristics such as detecting the driver's control of the steering wheel [4], the change of rotation speed

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 95–101, 2023.
https://doi.org/10.1007/978-981-99-2653-4_12

and slowness, whether the pressure on the accelerator pedal is smooth, and the driving trajectory of the movement. This detection method is simple to implement, but it is susceptible to environmental factors, such as the good environment of the driving road surface, whether the climate is harsh, etc., and the anti-interference ability of the detection system is poor [5]. In addition, some normal driving operations such as hyper-movement, parallel lanes, etc., may also lead to misjudgment.

The face information detection method of college athletes installs high-definition cameras in sports, and when college athletes start sports, the cameras begin to work. Through computer vision, the changes in facial expression characteristics of college athletes are detected, such as blinking frequency, eyelid closure PERCLOS, eye tracking, pupillary response [6], head rotation amplitude and height lifting, whether to yawn and other behaviors Assess whether college athletes have entered abnormal heart rate states, which is less costly and simple to achieve.

However, due to the fact that the clarity of the shooting image is determined by the ambient lighting conditions, this method has higher requirements for light, and the camera has been shooting at the university athletes, which will have a certain impact on the privacy of the college athletes, and the college athletes are prone to resistance and irritability psychologically Affect driving safety. Most of the existing heart rate anomaly detection methods only use a single signal to judge the abnormal state of heart rate of college athletes, and the signal collection during the actual exercise process is easily disturbed by various factors, and sometimes the collected data is not accurate enough, due to the single basis for discrimination, the trained model is poorly robust, prone to false alarms, etc., but affects the safety of college athletes. Therefore, this paper combines random forests with multi-source information fusion for detection. The multi-source information fusion detection method uses multiple micro-sensors with high accuracy [7] to synchronously collect signals such as breathing, heartbeat, pulse, grip and other signals of college athletes, and when the collection of one signal is deviated, the acquisition of another signal is not interfered with This information is then filtered and processed by Fourier transform, and a multi-source heart rate anomaly state dataset is established, which avoids the defect of poor anti-interference of a single signal [8], and greatly improves the accuracy of heart rate anomaly detection. Compared with classification algorithms such as SVM (Support Vector Ma-chines) and GBDT (Gradient Boosting Decision Tree), random forests can balance the error of the data set and are not sensitive to missing values. It can operate in parallel and has a high classification speed, which meets the requirements of signal acquisition with certain errors and high real-time requirements in heart rate anomaly detection.

1 Multi-source Signal Acquisition Platform for Abnormal Heart Rate

The experimental environment set up is shown in Fig. 1, and the actual line movement environment is simulated by simulating the driver to ensure the safety of the tester. The signal acquisition platform mainly includes three parts: Doppler radar module, flexible grip sensor module [9] and photoplethysmography pulse sensor module. In addition, the facial images of cameras recording the driving process of college athletes were used as

the basis for expert criticism of the data set signal classification [10]. The hardware design of the three modules all use smaller sensors to minimize contact with college athletes and prevent influence on the operation of college athletes. After the initial signal is collected, the physiological signals of college athletes collected by Doppler radar are filtered by a zero-phase ellipse filter to separate the breathing and heartbeat signals. The low-frequency noise in the pulse signal and the fixed power frequency interference at 50 Hz are filtered out by the zero-phase Butterworth band-pass filter and the Chebyshev II band-stop filter to obtain a clear and complete pulse signal and facilitate feature extraction.

In order to increase the collection efficiency, the experiment was selected in the afternoon, and the human body's activity level was generally low during this period, and it was easy to be sleepy, so that the test college athletes could better enter the abnormal state of heart rate. After the system is built and debugged, the experimental personnel are arranged to collect data.

2 Detection of Heart Rate Anomalies in the Fusion of Random Forests and Multi-source Information

The random forest is one of the classical ensemble learning algorithms, and its basic unit is the same as the gradient boost tree algorithm (GBDT), which is also a decision tree, but the type of the underlying decision tree can vary according to the actual data set. It is not necessary to use the same base learner for all. In addition, unlike GBDT, it adopts the idea of Bagging, which improves the stability of decision tree classification.

2.1 Random Forest Algorithm Model

At the heart of the random forest algorithm are two words, one is "forest" and the other is "random". "Forest" is because it integrates many decision trees, and the random forest algorithm uses the underlying decision tree model as a weak learner. In this paper, the Stochastic Forest mainly uses the Gini coefficient to calculate the impurity of the model nodes, and the smaller the Gini coefficient calculated from the node sample, the lower the impurity of the data The better the selection of representative features. Suppose a classification problem has a total of K different categories, and the probability of the *k-th* category is pk, then its probability distribution Gini coefficient expression is: the node of each decision tree in a random forest, select the values to split the dataset among the features of the input sample to minimize the division of Gini impurities.

Because the sample and feature selection composition of each base learner are different, the generalization ability is greatly enhanced. As shown in Fig. 2, the random forest is trained multiple times by generating different decision trees, and after the results are obtained separately, the categories of the test signal data are selected by voting method, and the result with the highest vote is used as the output. Assuming that a training set has k samples and samples it randomly, the probability that the same sample will be picked each time it is sampled is $1/k$. The probability that k random samplings did not choose it is: Bagging sampled the samples of the training dataset, and then selected them in turn, and about 36.8% of the samples in the training set will not be selected, which is called

“ Out of bag data”. These “out-of-bag data” do not have learning data as models, and can be used for “outsourced estimation” to analyze whether the learned classification model has good generalization performance. For random forest generation, trees are independent of each other, there are no dependencies, and can be trained in parallel, so the operation speed is fast. Its node division feature attribute set is randomly selected, without manual selection, compared to SVM and GBDT algorithms, it is not afraid of the loss of certain data features It is insensitive to outliers and has a good performance for high-dimensional data with many features.

2.2 Experimental Results and Analysis

Comparing the performance of the stochastic forest algorithm under the single signal dataset and the multi-source fusion dataset, it is verified that the multi-source information fusion has a good effect on improving the accuracy of the heart rate anomaly detection model, and then the established multi-source information fusion dataset is further input to the classification model based on SVM, GBDT and random forest for analysis and comparison. The superiority of the random forest algorithm as a classification algorithm model for heart rate anomaly detection is verified, and the classification model of heart rate anomaly detection under the optimal parameters is obtained after the detection method is adjusted.

3 Advantages of Multi-source Information Fusion Datasets

The experiment used the single signals of heartbeat, respiration, pulse and grip force to train the random forest classification algorithm model, and compared and analyzed the performance of the classification model trained by the multi-source information fusion dataset in the detection of heart rate abnormality. The optimal detection classification effect obtained in the test is shown in Fig. 1 and Fig. 2.

The detection accuracy of each heart rate anomaly status level can be seen that the heart rate anomaly detection model based on a single signal may have better classification accuracy in a certain category, but the overall detection accuracy is far less than that of the heart rate anomaly detection model based on multi-source information fusion. The best detection accuracy of the stochastic forest algorithm model after multi-source information fusion is about 89%, while the detection accuracy of a single dataset is only about 75%. Compared with a single data set, the random forest model based on multi-source information fusion increased the detection rate of abnormal heart rate states by 14%, which verified the superiority of multi-source information fusion.

The multi-source information fusion dataset is established to train SVM, GB-DT, and random forest algorithm models, respectively. The sample dataset is randomly divided into training sets and test sets, and the GBDT classification model and random forest classification model are trained with the training set, and the test set tests the classification model effect. The main adjustment parameters in the training process of random forest classification model are: the maximum number of features established by the decision tree (max-feature) and the maximum depth of the decision tree (max-depth), The number of decision trees generated (n_estimators). After multiple trainings, the classification

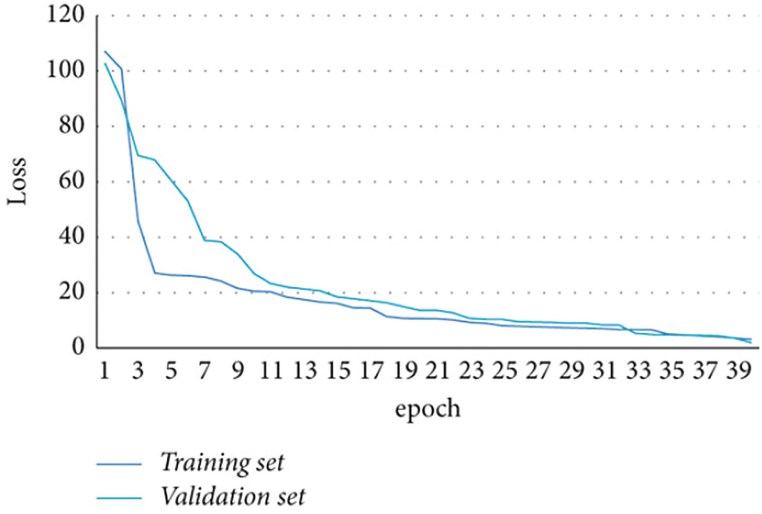


Fig. 1. Training process loss convergence curve.

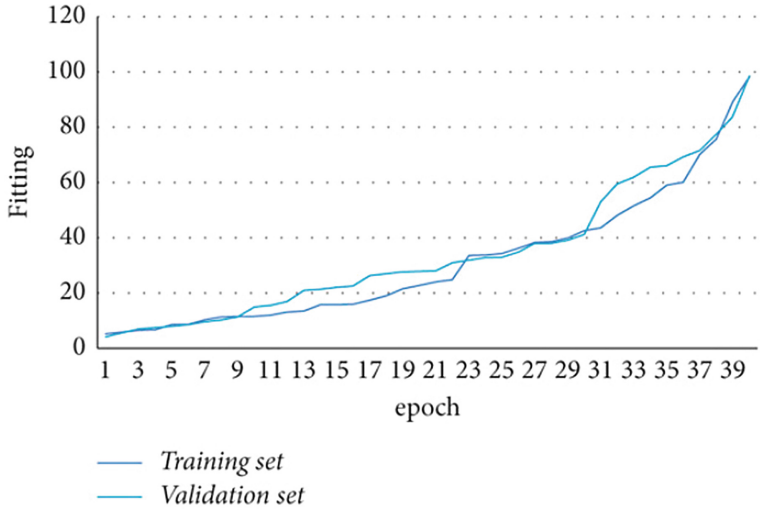


Fig. 2. Training process performance improvement diagram.

model is trained to obtain SVM, GBDT and random forest algorithm models under the multi-source information fusion dataset learning, and the entire data set is input into three abnormal heart rate driving detection models for testing. The output renderings of heart rate anomaly detection under different classification algorithms (see Fig. 3) and related accuracy parameters are compared.

Combined with the detection effect shown in Fig. 2, it can be seen that the random forest algorithm model has a good performance for the detection of abnormal heart rate

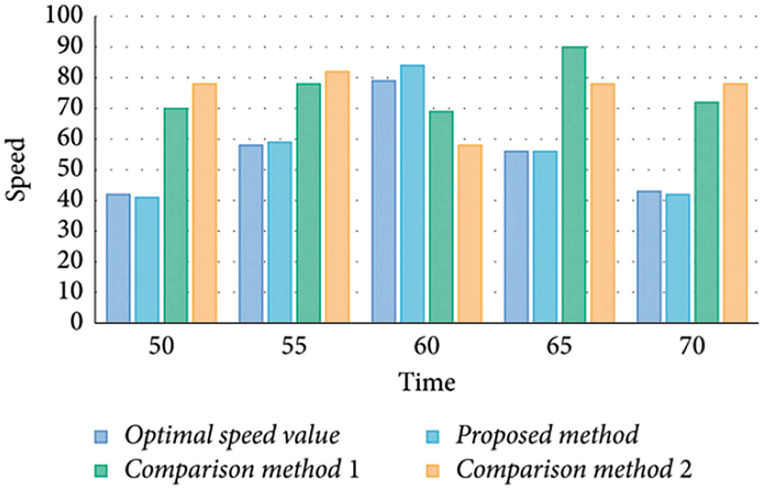


Fig. 3. Fatigue driving status detection effects of Different algorithm models

states at all levels, and the recognition accuracy is not much improved compared with SVM and GBDT. In addition, random forests can produce decision trees in parallel, and the detection model is also faster than SVM and GBDT for abnormal heart rate states. In order to avoid the deviation of the experimental effect caused by the accidental division of the test data set, the model automatically establishes the test set by random sampling multiple times, analyzes the detection accuracy of different classification algorithms, and takes the average of the detection accuracy of multiple detections in the experiment. The above experiments verify that the algorithm model based on random forest and multi-source information fusion can accurately detect abnormal heart rate states, and the best detection accuracy of the random forest algorithm model after multi-source information fusion is 89.18%. Compared with GBDT and classic SVM, its training speed is faster and more accurate, and it basically meets the requirements of real-time and high-precision detection of abnormal heart rate conditions.

4 Conclusion

Based on the heart rate anomaly detection method of random forest and multi-source information fusion, the signal acquisition module adopts a miniature sensor to reduce the deep contact with college athletes, and will not cause physical and mental interference to college athletes during exercise. After the collected signals are filtered wave processing, a fusion dataset with high reliability is established, and a random forest model is designed to learn and train the heart rate abnormal state data set, which realizes the high-precision and high-speed detection of heart rate abnormalities, and avoids the problem of large detection errors based on only a single signal. In addition, for the individual differences of athletes in different colleges and universities, the heart rate anomaly detection model is prone to deviations, and in the future, it is necessary to analyze more signal characteristics that can accurately reflect the abnormal heart rate state of college athletes, try to select

those feature values that are less affected by the driving environment and individual differences, and optimize the classification algorithm to make the classification model have self-learning ability, so as to achieve the optimal heart rate abnormal classification model that adapts to the individual through self-learning training.

Acknowledgment. The work of this paper was supported by Natural Science Foundation for Higher Education of Anhui Province of China (No. KJ2020A1095, KJ2021A1509); Academic Project for Top-notch Talents of Disciplines (Majors) for Higher Education of Anhui Province of China (No. GxbjZD2020105).

References

1. ISI. Network Simulator. Development was supported by DARPA through the VINT project at LBL, Xerox PARC, UCB, and USC/ISI in 2020. http://www.isi.edu/nsnam/ns/
2. Meng, M., Wu, X., Jeong, B.-S., Lee, S., Lee, Y.-K.: Energy efficient routing in multiple sink sensor networks. In: Proceedings of the 5th International Conference on Computational Science and Its Applications (ICCSA 2007), pp. 561–566, IEEE, Kuala Lumpur, Malaysia, August 2021
3. Wang, C., Wu, W.: A load-balance routing algorithm for multi-sink wireless sensor networks. In: Proceedings of the International Conference on Communication Software and Networks (ICCSN 2009), pp. 380–384, Macau, China, February 2019
4. Robert, S., Wayne, K.: Ghaps. In: Algorithms, chapter 4. Princenton Editorial, 4th edn. (2019)
5. Chang, X., Wu, J., He, Z., Li, D., Sun, H., Wang. Understanding user's travel behavior and city region functions from station-free shared bike usage data. Transp. Res. Part F: Psychol. Behav. (2020)
6. Xing, Y., Wang, K., Lu, J.J.: Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China. J Transp. Geography (2020)
7. Lee, D., Derrible, S., Pereira, F.C.: Comparison of Four Types of Artificial Neural Network and a Multinomial Logit Model for Travel Mode Choice Modeling. Transp. Res. Record (49) (2020)
8. Ma, X., Liu, C., Wen, H., Wang, Y., Wu, Y.-J.: Understanding commuting patterns using transit smart card data. J. Transp. Geography (2019)
9. Paulssen, M., Temme, D., Vij, A., Walker, J.L.: Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. Transp. (4) (2019)
10. Srinivasan, K.K., Mahmassani, H.S.: Analyzing heterogeneity and unobserved structural effects in route-switching behavior under ATIS: a dynamic kernel logit formulation. Transp. Res. Part B (9) (2021)

Application Research of Fluid Energy Harvester Based on Wireless Sensing Technology

Youchun Zhang(✉)

School of Application Engineering, Anhui Business and Technology College, Hefei 231131, China

zhangyouchun@ahbvc.edu.cn

Abstract. This paper proposes a airflow energy harvester based on wireless sensing, and conducts experimental research on it. The experimental results show that the optimal angle of incidence of the airflow is 83°, at which the minimum nitrogen flow can be used to drive the permanent magnet rotor, so that the energy harvester generates an induced voltage. Due to the presence of eccentricity, the energy harvester has a stable air flow rate of 137 ccm to 733 sccm and the energy harvester generates a maximum voltage of 160 mV. The improved energy harvester overcomes the shortcomings of eccentricity and can still work stably in high-speed airflow, with a stable air flow range of 137 sccm to 2748 sccm and a maximum output voltage of 3V. Subsequently, corresponding boost and management circuits can be designed to store and boost the acquired electrical energy to provide intermittent power to the wireless sensing network. Since the horizontally arranged double-ended nozzle has good output characteristics, the corresponding airflow harvesting device can be designed to optimize the energy harvester.

Keywords: wireless sensing · Airflow energy harvesters · Finite element analysis · Torque · Airflow angle

In recent years, obtaining energy from the operating environment of devices has become a hot topic of research, providing a very promising way to power wireless sensor nodes and replace traditional batteries. Some of the energy harvested by currently proposed energy harvesters mainly include: RF energy, vibration energy, air flow energy [and light energy] of which air flow as a universal energy has received widespread attention [1]. The most common and widely used device for harvesting airflow energy is a wind turbine. Although wind turbines can produce enormous amounts of energy, their drawbacks cannot be ignored. Wind turbines are huge and require a lot of space; It generates wireless sensing radiation and noise, so it should be installed away from residential areas. In addition, wind turbines are unable to collect energy from low-speed airflows, which contain considerable amounts of energy that have not yet been fully exploited [2].

A kind of airflow energy harvester consisting of housing, lifting magnet, upper pyrolytic graphite plate, coil, permanent magnet rotor and lower pyrolytic graphite plate was studied. The permanent magnet rotor [3], which is freely suspended between two pyrolytic graphite plates, rotates under the action of an external air stream and generates an induced voltage in the coil. The finite element software COMSOL Multiphysics5.3

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 102–108, 2023.
https://doi.org/10.1007/978-981-99-2653-4_13

and ANSYS Maxwell 16.0 are used to establish a simulation model, and the suspension, driving and output characteristics of the energy harvester are simulated and verified by experiments It was found that when the nozzle was 83°, the rotor was driven by the maximum air flow torque; Experimental test permanent magnet rotor suspension height of 56.5 mm, and the error with the simulation height of 57.5 mm is only 1.77%; The energy harvester operates stably with air flow rates ranging from 137 sccm to 733 sccm and a maximum output voltage of 160 mV [4, 5].

1 Theoretical Analysis

Compared with ferromagnetism and paramagnetism, the antimagnetism in daily life is relatively weak. In fact, antimagnetism plays a role in a variety of substances, such as water, copper[, bismuth, and even biology [6]. An object with antimagnetism is called an antimagnetic body, and its susceptibility (χm) is less than 0. This means that the anti-magnet will be subject to a repulsive force, called the anti-magnetic force (F) in the magnetic field, and $\mu 0$ is the relative permeability; H is the magnetic field strength; V is the volume of the anti-magnet. If the magnetic field is strong enough, the magnet will float above the magnet or wireless sensor. In the wireless sensing architecture proposed in this paper, the magnet is attracted by the lifting magnet and floats freely between the HOPG plates [7]. Excellent suspension results can be obtained with HOPG plates, as HOPG plates have the strongest magnetic resistance relative to all materials at room temperature (Fig. 1).

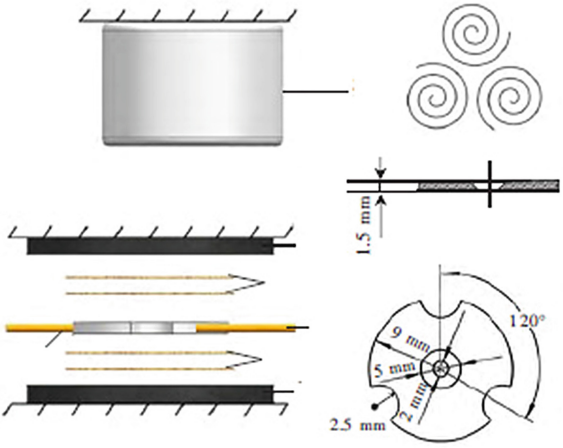


Fig. 1. Structure diagram of airflow energy collector

The rotor is subjected to an upward anti-magnetic force (FD), an upward magnetic attraction (FL), a downward anti-magnetic force (FE), and its own gravity (G) in the suspended state. In the condition of the rotor being able to stabilize suspension, the rotor and the nozzle are at the same level, and the pressure difference between the upper and

lower surfaces of the rotor has little effect on the stability of the rotor. Since the rotor rotates under the action of the air flow, only a part of the air flow acting on the side wall of the rotor blade is analyzed. The use of external nitrogen gas flow to blow the permanent magnet rotor, nitrogen gas flow nozzle and permanent magnet rotor on the same level, the incidence angle of the nitrogen gas flow is θ, the permanent magnet rotor is applied to a driving torque, so that the permanent magnet rotor can be rotated horizontally. In the figure, R is the radius of the rotor, R0 is the radius of the rotor bore, R1 is the distance from the center of the rotor at the bottom of the blade, and ω is the angular velocity of the rotor [8].

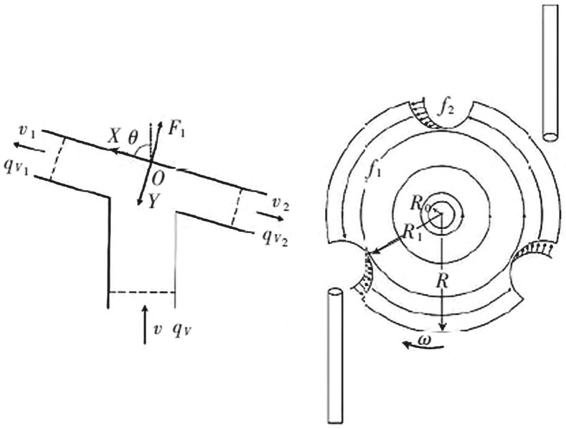


Fig. 2. Schematic diagram of action between light air flow and permanent magnet rotor

As shown in Fig. 2, the contact surface of the nitrogen fluid and the permanent magnet rotor is microstructured, taking this microstructured contact as the control body, and the gas flow acting on the permanent magnet rotor can be simplified to a hydrodynamic model of a jet acting on a diagonal flat wall after being divided into two jets. Assuming that the volume flow rate of the nitrogen gas stream is qV and the velocity is v, it is divided into two bundles of fluids after interacting with the permanent magnet rotor, the velocity is v1 and v2, and the volume flow rate is qV1 and qV2, respectively. The force of the fluid on the flat wall (F1) is perpendicular to the wall, and F1 is the direct driving force for the rotation of the permanent magnet rotor. The rotor is subjected to one viscous resistance (f1) and one windward resistance (f2) during rotation. When the volumetric flow of the air flow is qV [9], the radius of the nozzle is r0, and the volume flow of the air flow is converted to the flow rate (v) then the momentum equation is listed on the X and Y axes in the control body: ρ is the density of the fluid. The driving force F1 and the driving torque (T) are respectively: L is the distance from F1 to the center of the rotor; r is the radius of the rotor element. As can be seen from Fig. 2, the friction force received by the rotor during the rotation process is composed of two kinds of friction force, one is the viscous resistance equation: the μ is the dynamic viscosity of the fluid; A is the surface area of the rotor; h is the height of the rotor from the pyrolytic graphite plate. Both the upper and lower surfaces of the rotor are subject to viscous resistance,

and the frictional moment is calculated for the viscous resistance f2: b is a constant, which depends on the nature of the fluid and the size of the rotor. The 3 blades of the rotor are subject to the resistance of the windward surface, and the frictional moment When the driving torque and the frictional resistance moment of the rotor are balanced, the angular velocity of the rotor can be obtained, and M is the total frictional torque. With a coil with n turns, a magnetic field strength of B, and an area of the coil of S, the induced electromotive force (E) is the maximum output power (Pmax) of the energy harvester when the load resistance (RL) is equal to the coil internal resistance (RI).

2 Simulation Analysis

In order to obtain magnetic and anti-magnetic forces, a 3D model with actual dimensions is established by the software COMSOLMultiphysics 5.3. Numerical simulation results of the energy harvester's magnetic field lines are analyzed using simulation software, and the analysis results of the magnetic field lines are shown in Fig. 3.

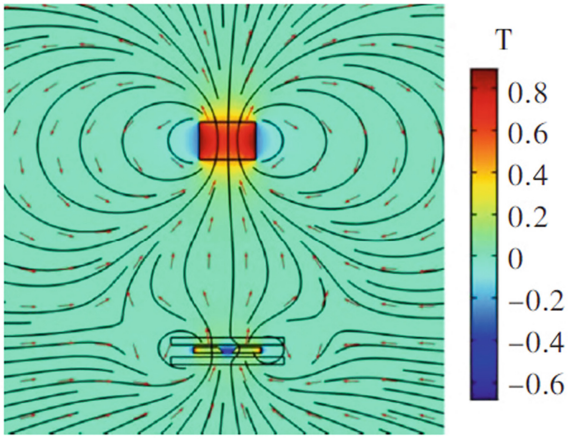


Fig. 3. Magnetic field line distribution of energy collector

Simulation of the convection field using the finite element software COMSOL Multiphysics 5.3. Figure 4 shows the streamlined diagram of the airflow acting on the permanent magnet rotor. In fact, the air flow of the rotor sidewall is the main driving force, in order to more clearly observe the flow of air and the sidewall of the rotor blade, the streamline diagram is a cross-sectional view of the rotor center plane. When the airflow acts on the blade of the permanent magnet rotor, the airflow is mainly divided into two streams of air. One stream of air moves clockwise along the arc of the blade and the other along the arc of the blade counterclockwise, which is consistent with the physical model of torque analysis.

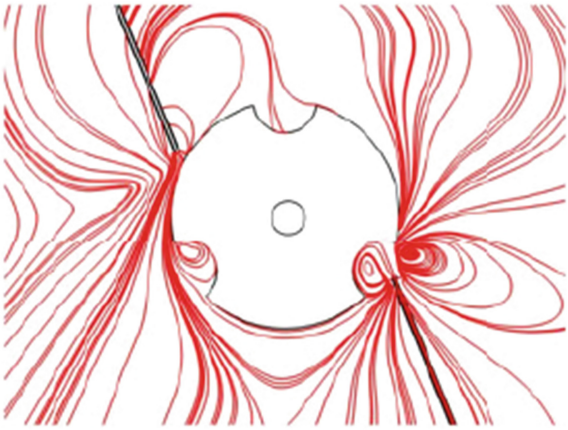


Fig. 4. Airflow streamline acting on permanent magnet rotor

The output voltage of the energy harvester is simulated by ANSYS Maxwell16.0. The rotational speed of the kinematic step is set to 20 000 r/min according to the experimental data. At the same time, set the stop time of the transient setting to 3 ms and the time step to 0.012 5 ms, and the simulation results are shown in Fig. 5.

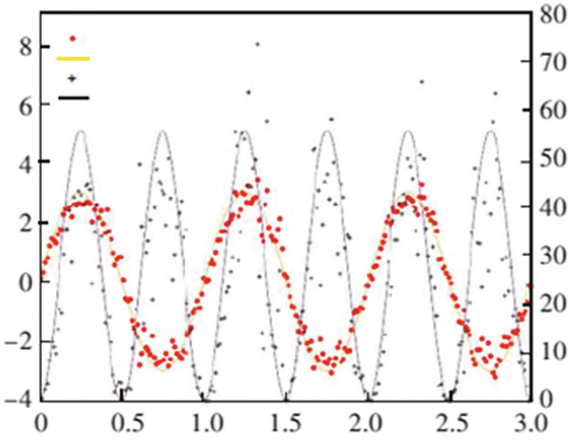


Fig. 5. Simulation results of output voltage and power

Since the rotor has a speed of 20 000 r/min, the time for one rotation is 3 ms. As can be seen from Fig. 5, rotating a cycle produces a total of three output voltage peaks, equal to the number of blades of the rotor. The output voltage has a frequency of 1 000 Hz, an output peak voltage of 3.1 V and an output power of up to 59.6 mV.

3 Airflow Drive Experiments and Analysis

The established airflow energy harvester experimental platform is shown in Fig. 6. In this experiment, nitrogen cylinders with constant pressure are used as an air source and connected to a centrally symmetrical nozzle. Using two mass flow controllers (AITOLY MFC300), the gas flow is controlled via computer software. The voltage and frequency of the induced current are measured with an oscilloscope (Tektronix TDS2012B). The frequency of the voltage signal of the photoelectric rotation measuring device is measured with an oscilloscope (TektronixTDS2012B) as a data processor (KEYENCE LV-NH35, 80 μs response time). The speed of the floating permanent magnet rotor can be calculated by calculating the frequency of the voltage signal. Experimental data of voltage and speed are processed to obtain a curve of peak voltage and speed as a relationship to the volume flow of the airflow (Fig. 6), and the two vertical axes in the figure correspond to the peak voltage and speed respectively. The two curves in the figure coincide to show that the relationship between peak voltage and rotational speed is linear and consistent with Eq. (12). Since the mass flow controller has a measuring range of 0 to 2 748 cm^3/mm, the maximum induced voltage reaches 3 V and the maximum output power reaches 55.8 mW. At the same time, the maximum speed of the rotor reaches 20 000 r/min.

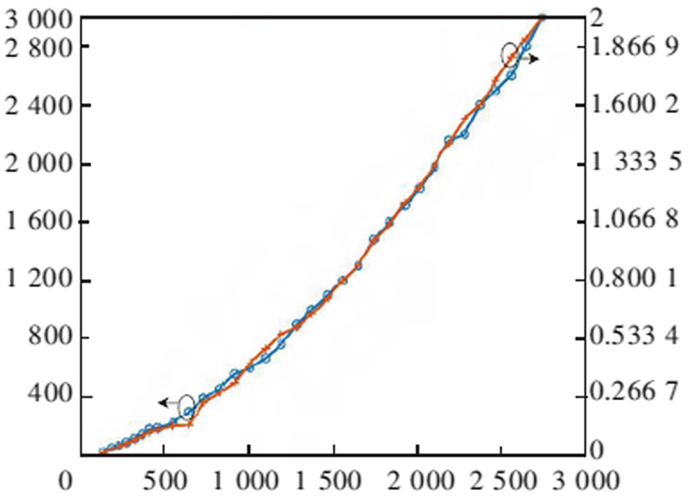


Fig. 6. Relationship curve between peak voltage and rotating speed and air flow volume flow respectively

The experimental waveform of the maximum output voltage of the airflow energy harvester is shown in Fig. 6. Since the rotor speed reaches 20 000 r/min, the waveform of the output voltage is very regular and is displayed as a sinusoidal curve on the oscilloscope. The frequency is 998.84 Hz, which is only 0.16% error compared to the simulation frequency of 1 000 Hz.

4 Conclusion

In this paper, an airflow energy harvester based on wireless sensing technology is proposed and studied. The output characteristics of the energy harvester were studied through theoretical analysis, simulation and experiments. It is found that the relationship between the output voltage and the airflow of the energy harvester is a quadratic function, and the relationship between the output voltage and the rotor speed is a linear function, which is verified by theoretical analysis and experiments. The maximum output voltage is 3 V, the maximum output power is 55.8 mW, and the maximum speed is 20 000 r/min. When the air flow reaches its maximum, the output voltage of the energy harvester still rises according to the trend of the quadratic function. Due to the limitations of the experimental equipment, the maximum output voltage at this time is not the limit of the energy harvester. This also reflects the advantages of using wireless sensing technology to reduce friction in airflow energy harvesters. Experimental results demonstrate the possibility of airflow energy harvesters based on wireless sensing technology. Future research efforts include optimizing rotor layout, coil size and nozzle position to improve the motion characteristics of magnetic rotors.

Acknowledgment. The work of this paper was supported by Natural Science Foundation for Higher Education of Anhui Province of China (No. KJ2020A1095, KJ2021A1510, KJ2021B002); Academic Project for Top-notch Talents of Disciplines (Majors) for Higher Education of Anhui Province of China (No. GxbjZD2020105).

References

1. Dehghan Hamani, I., Tikani, R., Assadi, H., Ziaei-Rad, S.: Energy harvesting from moving harmonic and moving continuous mass traversing on a simply supported beam. Measurement (C) (2020)
2. Yu, L., Tang, L., Yang, T.: Piezoelectric passive self-tuning energy harvester based on a beam-slider structure. J. Sound Vib. (2020)
3. Alameh, A.H., Gratuze, M., Nabki, F.: Impact of geometry on the performance of cantilever-based piezoelectric vibration energy harvesters. IEEE Sens. J. (22) (2019)
4. Zhou, G., Li, Z., Zhu, Z., Hao, B., Tang, C.: A new piezoelectric bimorph energy harvester based on the vortex-induced-vibration applied in rotational machinery. IEEE/ASME Trans. Mechatron. (2) (2019)
5. Wang, J., Zhou, S., Zhang, Z., Yurchenko, D.: High-performance piezoelectric wind energy harvester with Y-shaped attachments. Energy Conversion Manage. (2019)
6. Alameh, A.H., Gratuze, M., Elsayed, M.Y., Nabki, F.: Effects of proof mass geometry on piezoelectric vibration energy harvesters. Sensors (5) (2019)
7. Gupta, V.: Solar Energy Collector Having a Tree Structure. EP3837764A1 (2021)
8. Fu, H., Yeatman, E.M.: A methodology for low-speed broadband rotational energy harvesting using piezoelectric transduction and frequency up-conversion. Energy (2017)
9. Karadag, C.V., Topaloglu, N.: A self-sufficient and frequency tunable piezoelectric vibration energy harvester. J. Vib. Acoust. (1) (2020)

Visual Multi Scene Intelligent Car Service Platform Based on Internet of Things Technology

Guoqing Pan(✉)

School of Information Engineering, Suzhou University, Suzhou 234000, Anhui, China
guoqingpan@126.com

Abstract. The visual multi scene intelligent car service platform based on Internet of things technology is different from any previous machine idea. It is specially designed to serve an industry or a field. What we want to achieve here is a multi scene service state, and switching scenes without switching machines (only switching conversion method). At the same time, in order to meet this requirement, the traditional centralized design scheme can not be used to arrange all tasks in the machine. Therefore, the hardware level is separated from the software level for step-by-step processing, which will not only reduce the pressure of the hardware machine, but also improve its processing efficiency. Further, its processing capacity is no longer limited by the underlying hardware, but determined by the performance of the upper computer, After having a very powerful upper computer, it can realize the control mode of multi lower computer management. At the same time, it can adopt the scheme of time-sharing processing to deal with the multi lower computer, so as to meet its various performance requirements. In this way, we can complete the service mode of more scenarios under the current technical force, and only adopt a machine mode, a scalable and scalable smart car service platform whose efficiency can be improved at any time.

Keywords: Arduino mega · Arduino nano · raspberry PI 4b · upper computer

The smart car service platform will work in all walks of life in the future, and will not be limited by time and place. For example, when such a machine is put into a university campus, it can classify the garbage discarded by students next to the garbage can in the morning, remind students whether to wear masks, collect various physical data of the current location, and then go to the campus for face data recognition, Or to measure the body temperature of personnel. At noon, they do not need to rest like people. The operator can control them to patrol at all times in an operation room [1].

1 Overall Software and Hardware Framework

Arduino mega, Arduino nano and raspberry PI 4b are used for the hardware master control of visual multi scene intelligent car. Arduino Mega is responsible for the collection of most data, such as temperature, humidity, air particles, fire source detection, altitude,

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 109–115, 2023.
https://doi.org/10.1007/978-981-99-2653-4_14

air pressure, equipment voltage, vehicle inclination, as well as the control of vehicle body and the angle control of video acquisition equipment, as well as its own obstacle avoidance and line patrol functions. At the same time, the packaged data packets shall be sent up to raspberry PI 4B and summarized and sent to the upper computer [2, 3].

Arduino nano is only responsible for the collection of GPS data, because there will be a lot of time delay in the acquisition of GPS data. If it is processed in Arduino mega, the acquisition of other data will be blocked, resulting in the overall delay of data. Therefore, Arduino nano is used to obtain GPS data separately, transmit it to Arduino Mega in serial mode, and then package it for processing.

The task of raspberry PI 4b is to forward the data uploaded by Arduino Mega to the host computer, forward the instructions issued by the host computer to Arduino mega, collect the current video data in real time, and set up a forwarding platform by motion to realize its real-time viewing on the host computer and push the video output stream, so that the host computer can analyze the video data.

The upper computer mainly receives, analyzes, processes and displays the data. When it obtains the data forwarded by raspberry PI 4b, it will fully analyze the data packet to obtain the data, and conduct GUI display processing on the data in real time. This process can be deployed on different platforms. At the same time, the obtained data can be stored in the local database or uploaded to various Internet of things cloud platforms, It will refresh the data in real time and analyze the GPS data in real time to obtain the real-time position of the equipment.

2 Arduino MEAG Function Design

2.1 Displacement Function

This function is realized in Arduino Mega. The hardware adopts 12V power supply + ordinary motor (including reducer) + L298N motor drive + mcnamu wheel + aluminum alloy body. PWM speed control can realize forward and backward, in-situ steering and fixed angle steering at different speeds. Relying on the particularity of mcnamu wheel, it can also realize lateral translation. The specific movement function of the vehicle is controlled by issuing instructions [4].

2.2 Tracking Function

This function is realized in Arduino Mega. The hardware adopts five-way photoelectric tracking module to output the low level of low reflective material (black line). It can start the tracking mode by issuing an instruction. In the tracking mode, the vehicle relies on the photoelectric module located at the chassis to control the vehicle body. When there is only a low level in the middle position, start the vehicle to go straight. When the low level is output on the left or right, it can rotate accordingly, but pay attention to controlling the vehicle speed here, In order to avoid the vehicle's too late to respond, to improve the data, you can use the tracking module in front, or use the camera to collect image processing [5].

2.3 Obstacle Avoidance Function

This function is realized in Arduino mega, which is divided into left and right obstacle avoidance and front obstacle avoidance. HC is used for left and right obstacle avoidance_ The SR04 ultrasonic module realizes ultrasonic ranging. When the distance on one side is less than the set value and the distance on the other side is greater than the set value, slide and translate in the opposite direction until the vehicle distance is greater than the set value, but stop the displacement when the distance on the other side is less than the measured value. When the distance on both sides is less than the measured value, the vehicle will not move to prevent collision [6].

2.4 Temperature and Humidity Collection

This function is implemented in Arduino Mega. The DHT11 temperature and humidity module is adopted to import the DHT library. The library function can be used to read the data, but it should be noted that the digital interface should be used [7].

2.5 Altitude, Barometric Pressure Acquisition

This function is implemented in Arduino Mega. Gy-bme280 module is used to import BMP library and use library function to read data. However, it should be noted that the module should use I2C communication mode. At the same time, the altitude data of this module is easy to be affected by temperature and pressure, and the correction amount should be set according to local conditions. It also has temperature measurement function, but it has been realized in the front and will not be used here [8].

2.6 Ambient Gas Monitoring

We can use this module to measure the output of methane and dust (including 2.ppm methane and 2.ppm butane), and we can use this module to measure the output of methane and dust (including 2.ppm methane and 2.ppm butane), There is no need to use it here. In fact, digital output can also be used here, but it is not easy to use potentiometer to judge when to output high level. Because Mega has a large number of analog interfaces, it is most suitable to use analog interfaces.

2.7 Fire Source Detection

This function is realized in Arduino mega, which adopts flame (fire light) sensor module and steering gear module. The specific principle is that the flame will emit infrared. This module detects infrared. It has digital and analog interfaces, and still uses analog interfaces. At the beginning, the steering gear is used to scan within 0–180° in front of the vehicle. Once there is a fire source, the vehicle will be stopped temporarily, the steering gear will be suspended, the steering gear angle will be read, and the vehicle will be controlled to turn to the steering gear angle, Start the vehicle until the read analog quantity is less than the set value (it is determined that the fire source is very close), stop the vehicle and start the fire extinguishing program [9].

3 Raspberry PI Function Design

3.1 Instruction and Data Forwarding

This function is realized in raspberry PI, which is written in Python and is the basis for the communication between the upper computer and Arduino Mega. Raspberry PI and the upper computer are connected in socket mode in full duplex. At the same time, raspberry PI and Arduino Mega are connected in serial mode in full duplex. The data sent by the upper computer is encoded and transmitted to Arduino mega, and the data collected by Arduino Mega is encoded and sent to the upper computer, In this way, the host computer can control the whole equipment and collect information.

3.2 Video Data Acquisition

This function is realized in raspberry PI. The multi-functional night vision camera is connected to raspberry PI. Two steering gears are used in Arduino Mega to regulate the steering of steering gear in order to regulate the direction of camera. Motion is used to set up 8081 port, and the upper computer uses IP: port to access data to obtain image data.

3.3 GPS Function Design

Arduino nano is an intelligent hardware development board based on microchip Atmega 328p 8-bit micro control chip, with small size and powerful functions. The double row pin leads out, which can be conveniently stacked on the bread board, can be flexibly connected with other modules through DuPont terminals, and can properly hide in various designs. In short, there are a lot. Let's take a look at the performance profile of nano.

4 Host Computer Design of Intelligent Car Service Platform

4.1 Function Design of Upper Computer

Upper computer refers to the computer that can directly send control commands, generally PC/host computer/master computer/upper computer. Various signal changes (hydraulic pressure, water level, temperature, etc.) are displayed on the screen. The lower computer is a computer that directly controls the equipment and obtains the equipment status. It is generally PLC/single chip microcomputer/slave computer/lower computer and so on. The command sent by the upper computer is first given to the lower computer, which then interprets it into the corresponding timing signal according to this command and directly controls the corresponding equipment. The lower computer reads the equipment status data (generally analog quantity) from time to time, converts it into digital signal and feeds it back to the upper computer.

4.2 Instruction Issuance

This function is realized in the upper computer. The console is written in Python, and the corresponding coding instructions are sent in the console. It is connected with the raspberry PI in socket mode. The raspberry PI terminal receives it and forwards it to Arduino Mega for execution, such as forward and backward, left and right turns, tracking, etc. (Figs. 1 and 2).

4.3 System Software and Hardware System Display

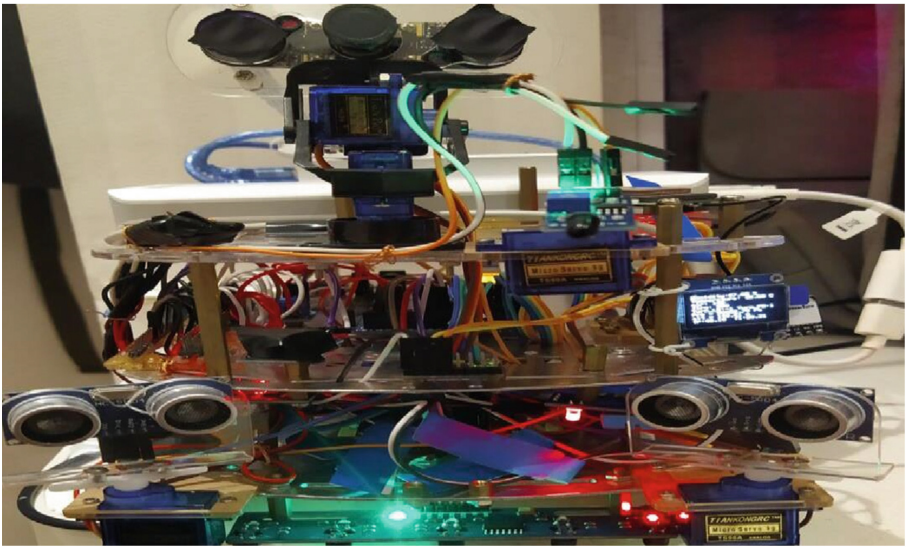


Fig. 1. Display of intelligent car hardware platform

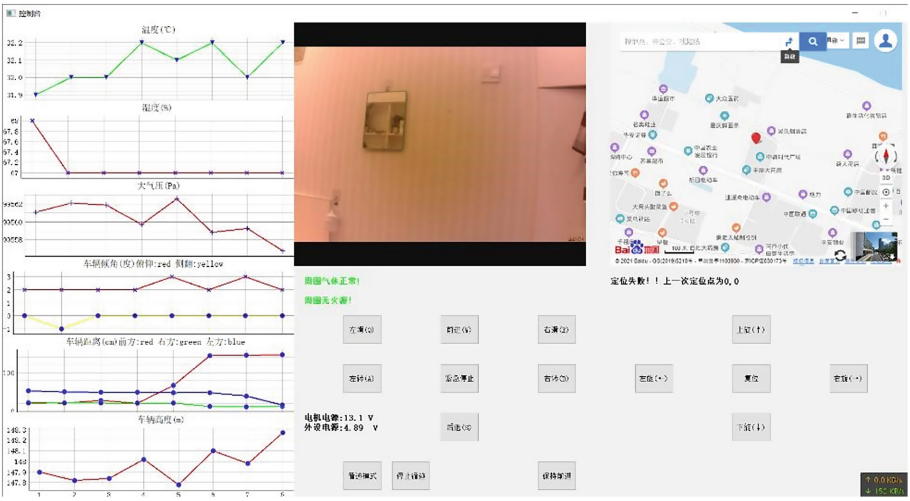


Fig. 2. Smart car software platform display

5 Conclusion

Multi scene, mixed type, combination of software and hardware, separation of software and hardware, timely expansion, unlimited efficiency, one lower computer and multiple upper computers, convenient management and labor liberation. Different from any previous machine idea, it is specifically designed to serve a certain industry or a certain field. What we want to achieve here is a multi scene service state, and switching scenes without switching machines (only switching algorithms). At the same time, in order to meet this requirement, the traditional centralized design scheme can not be used to arrange all tasks in the machine. Therefore, the hardware level is separated from the software level for step-by-step processing, which will not only reduce the pressure of the hardware machine, but also improve its processing efficiency. Further, its processing capacity is no longer limited by the underlying hardware, but determined by the performance of the upper computer, After having a very powerful upper computer, it can realize the control mode of multi lower computer management. At the same time, it can adopt the scheme of time-sharing processing to deal with the multi lower computer, so as to meet its various performance requirements.

Acknowledgment. School level Industrial College of Suzhou University (szxy2021cxxy04) top talent project of Anhui Universities (gxbjzd43), natural science research project of Anhui Provincial Department of Education (kj2018a0453), natural science research key project of Anhui Provincial Department of Education (kj2021a1110).

References

1. Li, F., Liu, J., Chen, Z., Huang, J., Liu, C., Qu, Z.: Navigating to urban environmental health: professionalized and personalized healthy living assistant based on intelligent health risk management. Urban Climate **40**, Article ID 101020 (2021)

2. Skar, S.L.G., PinedaMartos, R., Timpe, A., et al.: Urban agriculture as a keystone contribution towards securing sustainable and healthy development for cities in the future. Blue-Green Syst. **2**(1), 1–27 (2020)
3. Ogbonna, K.I., Bisong, T.L.: Strategies for the management of environmental hazards: implications for sustainable healthy living in Calabar Urban, Cross River State, Nigeria. Glob. J. Agric. Sci. **17**(1), 71–74 (2018)
4. Ali, M.J., Rahaman, M., Hossain, S.I.: Urban green spaces for elderly human health: a planning model for healthy city living. Land Use Policy **114**, Article ID 105970 (2022)
5. Eggen, M., Bouland, C.: Improving healthy living by urban walking and cycling: impacts of the built environment, green cover, air quality and noise in Brussels-Capital, Belgium. Environ. Epidemiol. **3**, 107 (2019)
6. Tresch, S., Frey, D., Le Bayon, R.-C., et al.: Litter decomposition driven by soil fauna, plant diversity and soil management in urban gardens. 7e Science of the Total Environment, **658**, 1614–1629, 2019
7. Lu, Y.: e association of urban greenness and walking behavior: using google street view and deep learning techniques to estimate residents' exposure to urban greenness. Int. J. Environ. Res. Public Health **15**(8), 1576 (2018)
8. Wang, R., Liu, Y., Lu, Y., et al.: Perceptions of built environment and health outcomes for older Chinese in Beijing: a big data approach with street view images and deep learning technique. Comput. Environ. Urban Syst. **78**, Article ID 101386 (2019)
9. Helbich, M., Yao, Y., Liu, Y., Zhang, J., Liu, P., Wang, R.: Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China. Environ. Int. **126**, 107–117 (2019)

Research and Design of Intelligent Monitoring System for Solar Energy Power Generation Based on GPRS Technology

Zhiyang Song(✉)

School of Information Engineering, Suzhou University, Suzhou 234000, Anhui, India
zhiyangsong@126.com

Abstract. In order to solve the problems of poor monitoring efficiency and untimely maintenance of traditional solar power generation system, a set of intelligent monitoring and detection system for solar energy power generation based on GPRS technology is designed. The main functional design of the system includes a three-layer architecture of data acquisition and perception layer, master array node network transmission and intelligent monitoring and control center. The main control array node hardware system mainly includes Ethernet communication and GPRS wireless communication hardware circuit. The sensing layer hardware system includes a light radiometer, a backplane and battery temperature sensing circuit, an ambient temperature and humidity sensing circuit, and a peripheral circuit connection design. The software design process includes data acquisition process and data transmission process, and the results of experimental test analysis show that the system can monitor and control the working parameters of the solar energy power generation system in real time, and provide real-time and efficient condition guarantee for the maintenance and repair of the solar energy power generation system.

Keywords: GPRS · Real-time monitoring · Data acquisition · Wireless communication

With the over-exploitation and use of traditional consumption-oriented energy and the increasingly serious environmental pollution, clean energy is favored, and solar energy generation has been more and more promoted and used as a new energy technology. The photovoltaic power generation system is a semiconductor diode that uses the technology of light energy and power conversion to realize the transmission and storage of electrical energy. When the photoelectric energy conversion diode deployed in the photovoltaic power generation system receives the illumination of sunlight, the internally deployed conversion circuit can realize the operation of converting solar energy into electrical energy, and then realize the storage and transmission of electrical energy [1]. GPRS-based intelligent monitoring and detection system of solar energy power generation applies wireless communication GPRS technology, wireless sensing data acquisition technology and integrated circuit technology to the intelligent monitoring and detection system of solar energy power generation, providing efficient, fast and efficient for the solar power generation industry. Stable and accurate new management methods have

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 116–124, 2023.
https://doi.org/10.1007/978-981-99-2653-4_15

greatly promoted the information development of smart grids and smart power systems [2].

1 Overall Design of Intelligent Monitoring and Detection System for Solar Energy Generation

The intelligent monitoring and detection control system of solar energy power generation mainly includes three parts: (1) data acquisition perception layer: This layer realizes the collection and storage of data such as the amount of sunlight radiation, the temperature value of the energy harvesting plate surface, the photoelectric conversion efficiency and the power generation rate through the data acquisition node. (2) Main control node network layer: This layer mainly realizes the data compression and upload of the underlying data acquisition layer, and transmits the packaged data to the total data server node for storage and analysis through wireless GPRS technology. (3) Intelligent monitoring and detection center application layer: This layer can realize the abnormal work of the data collection node, realize the reception and storage of the collected data through the data server, and warn the working state of the photoelectric conversion diode of the intelligent power generation system in real time, so as to realize the real-time intelligent monitoring and detection of the entire solar power generation system [3]. The data acquisition layer and the main control array node use the second-layer LAN to communicate data, the main control array realizes data upload through GPRS wireless communication, and the intelligent monitoring and detection center monitors the overall working status of the solar energy power generation system in real time through the server.

2 Master Node Network Layer Hardware System Design

The main control array of the power generation monitoring system adopts AM3352 embedded chip to realize the data storage collected by the sensing layer, and transmits the data to the data server of the monitoring center using wireless technology after compression, and its functions mainly include two parts [4]: (1) automatically collect the data of the sensing layer, and realize the data upload and distribution between the data acquisition node and the master array through Ethernet communication technology; (2) after the underlying collection data is preprocessed, Data is transmitted to the application layer data server via wireless GPRS transmission. The communication node of the master array includes three interface types: RS485 interaction with host computer, GPRS wireless communication upload data server, and data interaction between Ethernet mode and acquisition sensing layer.

2.1 Ethernet LAN Communication Circuit Design of the Host Controller

The main control chip and the perception layer realize data interaction through the LAN, and the host controller AM3352 and the LAN chip LAN8710A realize short-range communication and communicate with the host controller data through the RMII interface. The main controller and the perception acquisition layer upload and download data through ETH_MDIO and ETH_MDC [5]. The circuit connection of the main control chip Ethernet LAN is shown in Fig. 1.

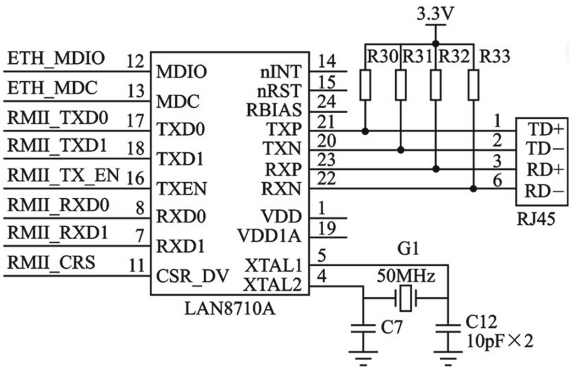


Fig. 1. Circuit diagram of the circuit connection between the main control chip and the sensing acquisition layer

2.2 GPRS Wireless Communication Circuit Design

The distance between the main control array node of solar energy generation and the intelligent monitoring center of solar energy power generation is far, and the two realize remote transmission and download of data through GPRS transparent wireless transmission technology. The GPRS module realizes the data docking of the main control array node and the data server of the monitoring center through the UART0RX and UART0TX serial ports [6, 7]. The control chip AM3352 controls the network status of GPRS and the transmission and storage of collected data through the serial port. The circuit connection design between the GPRS wireless communication module M22 and the SIM CARD is shown in Fig. 2.

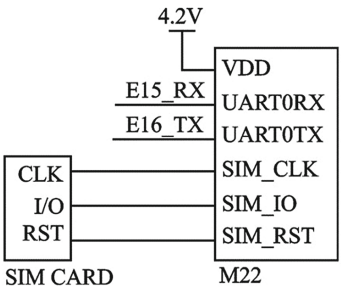


Fig. 2. GPRS wireless communication circuit module design

3 Data Acquisition Perception Layer Hardware System Design

The data acquisition and perception layer mainly includes two parts: (1) sensor module: data perception and collection uses modules such as light radiant, solar backplane and

battery storage temperature detection sensor and power generation environment temperature and humidity detection sensor to provide hardware support for the system; (2) data acquisition master control chip: embedded chip TM4C129 realizes data perception and collection of the solar energy generation monitoring system, including various types of temperature and humidity, light intensity and battery and other data parameters, and through the local area network to achieve data transmission [8].

3.1 Illuminatimeters

In the intelligent power generation monitoring system, the light radiation meter MS-711 measures the light energy irradiation intensity value by generating a linear proportional relationship between the light radiation value and the voltage signal through the temperature difference potential of the light perception point and the cold contact point. The voltage signal is transmitted through the A/D conversion interface, and the voltage value and light intensity are used to obtain the current time sunlight exposure value. The actual light radiometer MS-711 used in this system is shown in Fig. 3.

Fig. 3. Physical diagram of a light radiometer

3.2 Backplane and Battery Temperature Sensing Circuit

In the intelligent power generation monitoring system, the hardware temperature data acquisition includes two parameters: the temperature of the light backplane and the temperature of the power storage battery. The sensor hardware adopts the multi-threaded model DS18B20, which can be connected to the main control chip TM4C129 through the I/O interface to achieve multithreaded temperature data upload. As shown in Fig. 4, the DS18B20 hardware circuit connection diagram is supplied by a 5V external power supply.

3.3 Monitoring System Environmental Sensing Circuit Design

The environmental temperature and humidity data collection of the intelligent power generation monitoring system uses the model SHT10, which can simultaneously realize the realtime acquisition of temperature and humidity data in the solar power generation environment. The sensor is connected to the main control chip TM4C129 through the

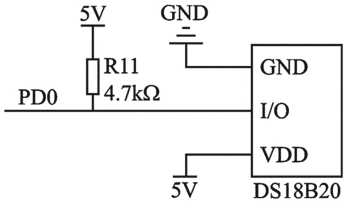


Fig. 4. Temperature sensor interface circuit diagram

SCK and DATA interface, and the temperature and humidity parameters in the power generation environment can be transmitted to the perception layer master control chip for processing in real time. The connection design of the ambient temperature and humidity data acquisition circuit is shown in Fig. 5.

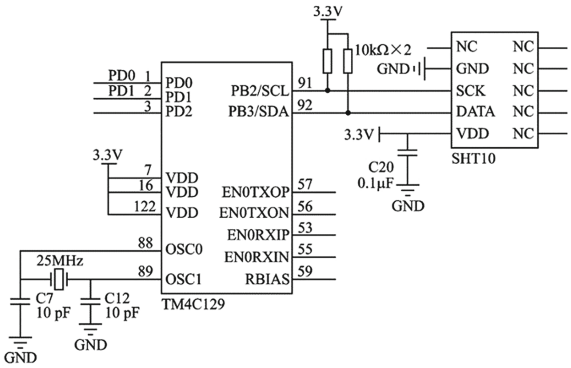


Fig. 5. Design of environmental sensing circuit for monitoring system

3.4 Design of Connection Between Peripheral Circuit and Ethernet Circuit of Main Control Chip

The internal circuit of the main controller TM4C129 includes a 10/100M Ethernet communication interface, which has its own Ethernet communication, and can complete the data transmission communication of the LOCAL area network without the need for an external Ethernet communication chip. The Ethernet communication circuitry of the host controller and the associated peripheral circuit connections are shown in Fig. 6.

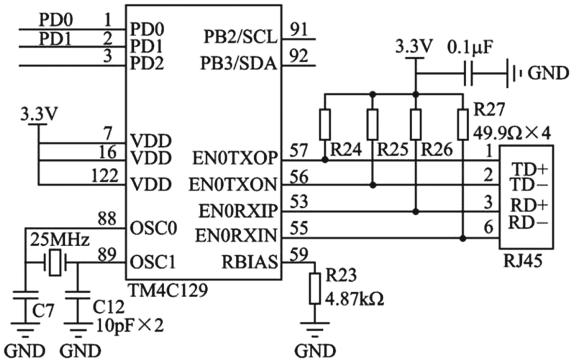


Fig. 6. The peripheral circuit of the main control chip and the Ethernet connection circuit

3.5 Distributed Communication Protocols

In order to improve the efficiency of data transmission, the communication protocol of the data management system uses ADP-based transmission technology with a data frame range of 1–64 bits.

The frame format is mainly designed with three types:

(1) data common format: mainly 1-byte frame type and 63-byte data length.
(2) Query command format: mainly including 1-byte frame type, 1-byte catetyte, 22-byte timestamp, 24-byte length detection and 18-byte data validation;
(3) storage data frame format: 1-byte frame type, 1-byte date tyte, 1-byte dataSIQ and 40-byte data. Three frame modes can meet the needs of data storage, checksum retrieval, etc. [10].

4 System Software Design and System Experimental Testing

4.1 Software Process Design of Intelligent Power Generation Monitoring System

The main program of the perception layer data acquisition node first completes the initialization and module initialization of the acquisition TM4C129 system, and coordinates the normal operation of various software modules. After initialization, the sensing layer data acquisition program realizes the acquisition of temperature parameters, ambient temperature and humidity parameters, and illumination radiation values, and realizes the processing and transmission of data to the master array node through the network layer. The main program of the main control array node mainly realizes the function of system initialization and receiving the underlying data, and realizes the pre-processing of data and the network transmission of data and other functions of the collected data, and the transmission mode mainly realizes the upload function of data through GPRS wireless transmission technology.

4.2 System Testing

Taking a photoelectric power generation array node as the object, the system data is monitored and detected for 12 h (6 am to 6 pm) in real time, and various parameters (light intensity, backplane temperature, battery temperature and ambient temperature and humidity) are displayed in the form of real-time data curves on the PC interface.

When one or more parameters fluctuate within the normal display range, the system issues an alarm to determine that the power generation system has failed. The system test results show that the intelligent monitoring system of solar energy power generation has completed the real-time monitoring and collection of power generation array data, and the application layer can judge the working conditions of the power generation system according to the real-time data graph, providing the maintenance efficiency of the power generation system. In order to verify the feasibility of the system design, the prototype system was implemented and tested in this code. The validation development board used for the test is Power837x, and the remaining hardware and software environments tested are listed in Table 1.

Table 1. System hardware and software configuration table

Hardware environment	PC	Development board
Hardware configuration	CPU:Intel(R)Core(TM) i7-8550UCPU 1.80GHz Memory:16384M HD:MZVLW256 256GSATA	CPUMPC837x DDR2SDRAM 512MB NOR Flash S29GL128M90TAIR1 (16MB) NAND Flash K9F2G08 (256MB)
Software configuration	OS:Windows10 IDE:Visual Studio 2017 QT:5.9.2(MSVC2017)	OS:Sylix-BH Unqlit
Network environment	100M LAN	100M LAN

This article uses multiple boards to simulate multiple on-board data sources, and different boards store different types of on-board data while running their own NoSQL-based data The data query module of database technology, and the integrated management module and the ground center are deployed in the PC simulation of the satellite environment to complete the external communication link. The PC and development board are connected via Ethernet on the switch to simulate the distributed internal communication link between the integrated management module and the on-board data query module. In the construction of the test environment, the data exchange equipment uses the switch to realize the ground-to-air network transmission environment business development board 1/2/3 analog satellite sensor data acquisition, which is responsible for simulating the perception and acquisition of different types of data. Database technology uses NoSQL to store, retrieve, and query data. The test system simulates satellite

stowage data stored in NoSQL and transmitted to the development board via the UDP protocol. The upload and download rates of the database are shown in Fig. 8. It can be seen from the figure that the more data the sensor collects, the lower the network transmission efficiency, and the upload and download rate is maintained at 1.5M/S.

As shown in Fig. 9, the trial operation diagram of the satellite stowage data management system can accurately realize the query, retrieval and transmission of data through the test of the system operation command. The test results show that the system can realize the storage, retrieval and air-to-ground transmission of satellite-mounted data.

5 Conclusion

Based on GPRS technology, the intelligent monitoring system of solar energy power generation realizes the real-time acquisition and upload of light radiation intensity, backplane and battery temperature and environmental temperature and humidity parameters through the hardware system design and software system design of the data acquisition and perception layer, and realizes the functions of data processing and remote wireless transmission to the application layer server through the wireless GPRS and Ethernet of the main control array node. The intelligent monitoring and control system realizes the collection of real-time data parameters of various parameters of the solar power generation system, provides real-time operation of repair and maintenance for solar power generation, and greatly improves the management level and maintenance efficiency of intelligent power.

Acknowledgment. School level Industrial College of Suzhou University (szxy2021cxxy04) top talent project of Anhui Universities (gxbjzd43), natural science research project of Anhui Provincial Department of Education (kj2018a0453), natural science research key project of Anhui Provincial Department of Education (kj2021a1110)School-Level Scientific Research Platform of Suzhou University (2021xjpt51).

References

1. Nasirin, C., Lionardo, A.: Effective implementation of marketing management of hospital pharmacy: a study on the impact of hospital service quality improvement in enhancing the of patients' satisfaction and loyalty. Syst. Rev. Pharm. **11**(5), 705–712 (2020)
2. Das, B., Ghosh, T.S., Kedia, S., et al.: Analysis of the gut microbiome of rural and urban healthy Indians living in sea level and high altitude areas. Sci. Rep. **8**(1), 10104–10115 (2018)
3. Ogbonna, K.I., Bisong, T.L.: Strategies for the management of environmental hazards: implications for sustainable healthy living in Calabar Urban, Cross River State, Nigeria. Glob. J. Agric. Sci. **17**(1), 71–74 (2018)
4. Tinti, D., Puccinelli, C., Mancini, L.: Urban green space and healthy living: an exploratory study among Appia Antica Parks users (Rome-Italy)[proceedings]: presented in the 19th International MESAEP Symposium, Rome-Italy from October 04–06, 2017. Fresenius Environ. Bull. **28**(6), 4984–4989 (2019)
5. Schram-Bijkerk, D., Otte, P., Dirven, L., Breure, A.M.: Indicators to support healthy urban gardening in urban management. 7e Sci. Total Environ. **621**, 863–871 (2018)

6. Fischer, L.K., Brinkmeyer, D., Karle, S.J., et al.: Biodiverse edible schools: linking healthy food, school gardens and local urban biodiversity. Urban Forestry Urban Greening **40**, 35–43 (2019)
7. Zhu, D., Chen, Q.-L., Li, H., et al.: Land use influences antibiotic resistance in the microbiome of soil collembolans Orchesellides sinensis. Environ. Sci. Technol. **52**(24), 14088–14098 (2018)
8. An, P., Wang, Z., Zhang, C.: Ensemble unsupervised autoencoders and Gaussian mixture model for cyberattack detection. Inf. Process. Manag. 59(2), Article ID 102844 (2022)

Science and Education Integration Mode Python-C Experimental Teaching Research Design of WIFI Line Avoidance Vehicle

Shiyuan Zhang(✉) and Bingqian Ni

School of Information Engineering, Suzhou University, Suzhou 234000, China
shiyuanzhang007@126.com

Abstract. The research purpose of WIFI patrol line obstacle avoidance trolley based on Python & C is to improve the quality of people's daily life, play a cornerstone role in automatic driving and cargo transportation, and lay the foundation for the development of future science and technology. In terms of intelligent vehicle environment detection technology, target recognition and detection technology based on convolutional neural network model and improved regional convolutional neural network. In order to optimize the project to a greater extent, multi-source data fusion technology based on deep learning network, radar sensor and camera sensor will be developed in the multi-sensor fusion technology of the car, and the deep learning technology is automatic The use of the perception layer, decision-making layer and control layer in the driving field in the research content continuously improves the accuracy of perception, detection and control, and strives to achieve certain success in this aspect.

Keywords: Line avoidance · C & Python · Technology

1 Introduction

The development of science and technology promotes the progress of the times, and people's living standards are constantly improving, which has a lot to do with the exploration of science and technology. President Xi stressed in his speech that development is the first priority, talent is the first resource, and innovation is the first driving force. Innovation and in-depth research on people's daily life needs and the direction of company development are strong support for national development. Recently, for example, in the express delivery industry, research is being carried out on drone delivery and smart car cargo handling, and in terms of life, most of the car manufacturing industry is scrambling to gather talents and technologies in order to study autonomous driving technology, and individual technical members are also racing against time to study algorithms, strive for first-class, compete for excellence, and jointly make big breakthroughs in science and technology, press the development fast forward button, achieve new results, and develop hard technology. In order to provide a force in scientific and technological innovation, our team planned a research program and achieved certain results. After only one year,

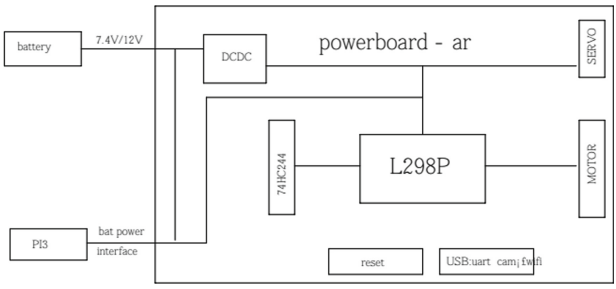
Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 125–133, 2023.
https://doi.org/10.1007/978-981-99-2653-4_16

the team's research mainly includes automatic car line patrol technology and human intervention control technology, remote video transmission, and route planning. After research, the basic function of pre-planning has basically been realized, and since the purpose of research is used as the foundation of the development direction, there is still a lot of room for development on this basis. In different adaptation sites, the car has room to modify itself according to the situation, whether it is in the improvement of the algorithm, the reconfiguration of the hardware or the stability. As a framework for technical research, the general technology has been basically realized, and it is believed that under the future in-depth research, it can open up a way for future research in this area to a greater extent.

2 System Architecture Design

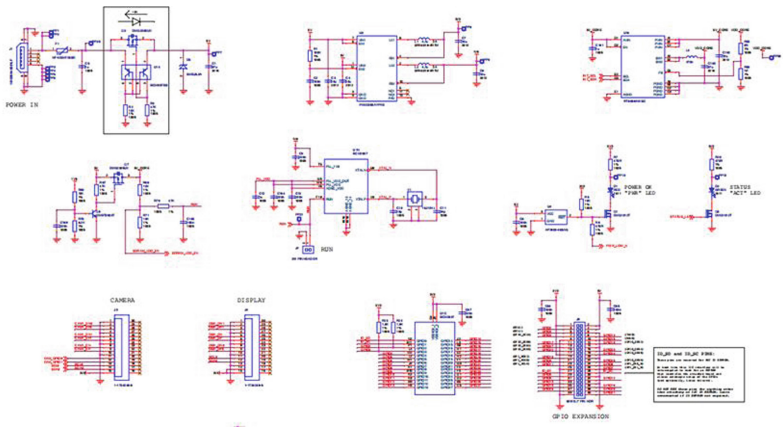
In terms of hardware, the main hardware module is to make the trolley diversified, so that it has the following basic functions:

1. Ultrasonic obstacle avoidance function, can self-identify obstacles and stop, to avoid collision with objects.
2. Automatic driving function, after adding ultrasonic and infrared rays to the four-wheel drive trolley, the trolley has the function of patrolling the line and automatically adjusting the direction, and the specific use of this function is mainly used in automatic driving during cargo transportation or road survey.
3. Camera function, which can record the real-time road conditions of the car and observe the surrounding situation from the perspective of the car. 4. Networking function, after the car joins the Internet module, it can obtain Internet related information, and derive the extended function of software from the network aspect.
4. Servo device, which allows the trolley to be connected to other functions, and can expand different additional hardware functions in different fields.



The module can be divided into: ultrasonic module, video transmission module, infrared sensor module, WIFI module and other basic modules required for research trolleys. The hardware of the trolley has been compatible and debugged for many times, and the problems that have occurred in the transmission of instructions between software and hardware have been corrected, and finally the software and hardware can be perfectly integrated, and the basic and preset implementation functions are consistent. The trolley

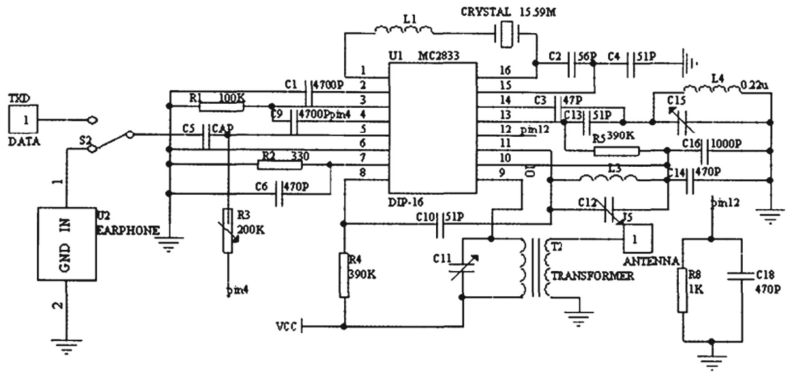
of this project is more intelligent and multi-functional than before. The development of intelligent tools is used in many fields, and it is also very popular in intelligent car driving, and the progress made in these aspects will pay more attention to the development of intelligence, and the research of intelligent vehicles has a very important role in the development of scientific and technological forces.



The software has the following basic functions: 1. Flash C language and Python environment, C language as the main grassroots language of the entire car, Python has a strong graphics processing function to make the car in the pattern recognition can have high efficiency. 2. The trolley adopts four-wheel drive, and the control instructions are written to the four motors respectively, which is realized by different wheel rotation and adjusts the direction in place. 3. Gravity sensor control, which can simulate the steering wheel through the gyroscope of the wireless device to manually control the trolley. 4. Write VR perspective control, through the gyroscope action of the mobile phone can change the camera angle of the car. 5. Route planning, draw the route to be driven on the control end, and the car can automatically drive according to the route set by the terminal. 6. Autonomous driving, based on deep learning, realizes the autonomous driving function at the execution layer.

3 Basic Features

The project trolley used the Raspberry Pi motherboard as the basis for the research design. Autonomous driving: As an experimental project of autonomous driving technology, it is used to simulate intelligent vehicle driving systems, and to study intelligent vehicle autonomous driving systems through this project. The smart car can drive according to the pre-set route, and can automatically avoid collision on the driving road. The research on the merger of autonomous driving with the Internet of Things intelligent car can greatly save manpower and maximize labor efficiency. The innovative development of intelligent cars has certain special functions, which is where most ordinary products in current life cannot have, and after development and design, the prospect of the development of this technology has immeasurable room for improvement.



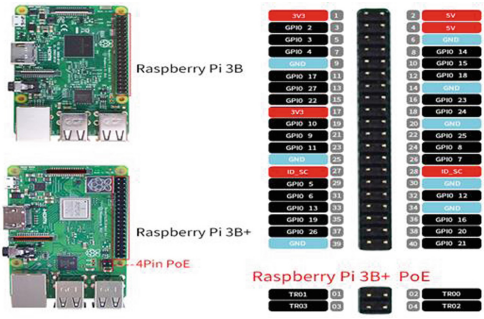
Signal transceiver function diagram

The trolley has a visual distance detection function, which is assisted by ultrasonic and infrared equipment, and obtains certain results after a certain algorithm processing. The trolley can still work well without human operation, and can drive on the original route or pre-planned route. The trolley avoids colliding with obstacles by scanning the surrounding situation with sensors to determine whether there are obstacles and the distance to them. There is a map drawing function on the control end, and the trolley will drive according to the drawn trajectory after the drawing is completed.

4 Procedure and Principle Implementation

```
void cruising_mod()// mode function toggle function
{
 if (Pre_Cruising_Flag != Cruising_Flag)
 {
        if (Pre_Cruising_Flag != 0)
        {
               MOTOR_GO_STOP;
        }
        Pre_Cruising_Flag =    Cruising_Flag;
 }
 switch (Cruising_Flag)
 {
        case 1: Follow(); return;// follow pattern
        case 2: TrackLine(); return; Line following mode
        case 3: Avoiding(); return;// obstacle avoidance mode
        case 4: AvoidByRadar(AvoidLength); return;//ultrasonic obstacle avoidance mode
        case 5: Send_Distance(); return; The ultrasonic distance is displayed on                the PC side
        default: return;
 }
}
```

5 Procedure and Principle Implementation



WiFi smart car, the firmware used has integrated python library, can directly run python programs. There are many reserved IO ports on the power board, which can be used directly (use BCM mode, chip internal pin sequence).

Due to its own accuracy problem, the actual robotic arm, gimbal, etc., cannot be directly driven by the IO port, otherwise there will be jitter. Therefore, a PWM generator is integrated on the power board, and the driver package is integrated into the Smbus library, and a simple API interface is set aside for calling, and the corresponding servo number is PWR The A53 motherboard corresponds to the servo number SER1-SER8:

```
from smbus import SMBus

XRservo = SMBus(1)


XRservo. XiaoRGEEK_SetServo(0x01,angle)
```

This function sets the servo angle and sets the angle value of the 0x01 servo to angle. Similarly, the angle value of the servo 1–8 can be set;

XRservo. XiaoRGEEK_SaveServo()

This function is the servo angle memory and stores the current steering gear values.

XRservo. XiaoRGEEK_ReSetServo().

This function restores all servos to stored angles.

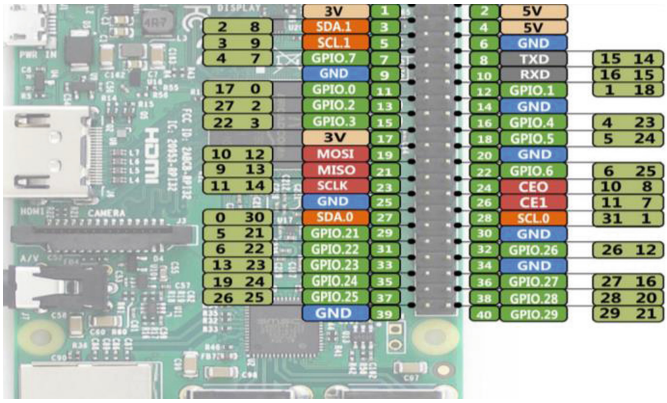
Code width algorithm:

Minimum width of a 16-bit address code: 1 12 $\times$ 16 = 18 ms Maximum width of a 16-bit address code: 2 24 ms $\times$ 16 = 36 ms.

It is easy to know the width and unchanged of the 8-bit data code and its 8-bit inverse code: (1 12 ms + 2. 24 ms) $\times$ 8 = 27 ms.

The width of the $\therefore$ 32-bit code is (18 ms + 27 ms) ~ (36 ms + 27 ms).

The key to decoding is how to identify “0” and “1”, and the definition of the bit we can find that “0” and “1” are both 0 The low level of 56 ms starts, the difference is that the width of the high level is different, “0” is 0 56 ms, “1” is 1 68 ms, so it is necessary to distinguish between “0” and “1” according to the width of the high level. If from 0 After 56 ms low, the delay begins, 0 After 56 ms, if the level read is low, it means that the bit is “0”, and vice versa, it is “1”, and for reliability, the delay must be greater than 0 56 ms is longer, but not more than 1 12 ms, otherwise if the bit is “0”, the reading is already the high level of the next bit, so take (1 12 ms + 0. 56 ms) /2 = 0. 84 ms is the most reliable, generally take 0 84 ms.



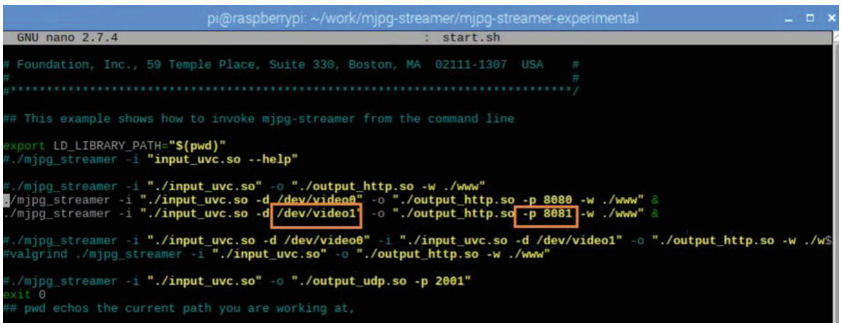
After logging in to the Raspberry Pi system with putty, enter the gpio readall command to view the corresponding GPIO comparison table.

```
pi@raspberrypi:~ $ gpio readall
 +-----+-----+---------+------+---+---Pi 3---+---+------+---------+-----+-----+
 | BCM | wPi |   Name  | Mode | V | Physical | V | Mode |  Name   | wPi | BCM |
 +-----+-----+---------+------+---+----++----+---+------+---------+-----+-----+
 |     |     |    3.3v |      |   |  1 || 2  |   |      | 5v      |     |     |
 |   2 |   8 |   SDA.1 | ALT0 | 1 |  3 || 4  |   |      | 5v      |     |     |
 |   3 |   9 |   SCL.1 | ALT0 | 1 |  5 || 6  |   |      | 0v      |     |     |
 |   4 |   7 | GPIO. 7 |   IN | 0 |  7 || 8  | 1 | ALT5 | TxD     |  15 |  14 |
 |     |     |      0v |      |   |  9 || 10 | 1 | ALT5 | RxD     |  16 |  15 |
 |  17 |   0 | GPIO. 0 |  OUT | 0 | 11 || 12 | 1 | IN   | GPIO. 1 |   1 |  18 |
 |  27 |   2 | GPIO. 2 |   IN | 1 | 13 || 14 |   |      | 0v      |     |     |
 |  22 |   3 | GPIO. 3 |   IN | 1 | 15 || 16 | 1 | IN   | GPIO. 4 |   4 |  23 |
 |     |     |    3.3v |      |   | 17 || 18 | 1 | IN   | GPIO. 5 |   5 |  24 |
 |  10 |  12 |    MOSI |  OUT | 1 | 19 || 20 |   |      | 0v      |     |     |
 |   9 |  13 |    MISO |  OUT | 1 | 21 || 22 | 1 | OUT  | GPIO. 6 |   6 |  25 |
 |  11 |  14 |    SCLK |  ALT | 0 | 23 || 24 | 0 | OUT  | CE0     |  10 |   8 |
 |     |     |      0v |      |   | 25 || 26 | 1 | OUT  | CE1     |  11 |   7 |
 |   0 |  30 |   SDA.0 |   IN | 1 | 27 || 28 | 1 | IN   | SCL.0   |  31 |   1 |
 |   5 |  21 | GPIO.21 |   IN | 1 | 29 || 30 |   |      | 0v      |     |     |
 |   6 |  22 | GPIO.22 |   IN | 1 | 31 || 32 | 0 | IN   | GPIO.26 |  26 |  12 |
 |  13 |  23 | GPIO.23 |  OUT | 1 | 33 || 34 |   |      | 0v      |     |     |
 |  19 |  24 | GPIO.24 |  OUT | 0 | 35 || 36 | 0 | OUT  | GPIO.27 |  27 |  16 |
 |  26 |  25 | GPIO.25 |  OUT | 0 | 37 || 38 | 1 | OUT  | GPIO.28 |  28 |  20 |
 |     |     |      0v |      |   | 39 || 40 | 0 | OUT  | GPIO.29 |  29 |  21 |
 +-----+-----+---------+------+---+----++----+---+------+---------+-----+-----+
 | BCM | wPi |   Name  | Mode | V | Physical | V | Mode |  Name   | wPi | BCM |
 +-----+-----+---------+------+---+---Pi 3---+---+------+---------+-----+-----+
```

Main thread: loop port monitoring, ensure the priority of instruction transmission, mainly run Communication_Decode(), instruction analysis function, and directly process manual control instructions, including trolley direction control, speed regulation, servo (gimbal, robotic arm) control, LED control, change Cruising_Flag value.

T1 thread: Loop execution of Cruising_Mod(), according to the Cruising value, switch to different automatic modes, including infrared follow mode, infrared line following mode, infrared obstacle avoidance mode, ultrasonic obstacle avoidance mode, ultrasonic distance PC display, Ultrasonic remote control obstacle avoidance, Path_Dect debugging mode (new), Path_Dect tracking mode (new).

T2 thread: loop through path_dect_img_processing(), and when entering camera tracking mode, start processing the camera data: intercept grayscale channel, binarization, output path _Dect_px value. Among them: the camera tracking program is: according to the Path_Dect_px value, to determine the direction of movement of the car.



The propagation speed of sound in the air is v, and according to the time difference △t between the transmission and reception of the echo recorded by the timer, the distance from the transmission point to the obstacle can be calculated, ie:

$$s = v * \Delta t/2$$

The unit of distance calculated above is meters (m), and centimeters (cm) are generally used in the code, so the result will be multiplied by 100.

In addition: the speed of sound propagation in the air is related to temperature, as follows:

$$v = 331 + 0.6T$$

thereinto

v is the speed of sound propagation in m/s.

T is the air temperature in °C.

Generally, the speed is 340 m/s at 15 °C.

6 Contributions and Results

Finally, the intelligent car based on the Raspberry Pi has been developed, and through the test, it can plan its own route, and according to the planned route form, this research is very closely related to unmanned driving, which is conducive to the development of driverless cars.

After continuous improvement of the system, it is finally connected to GPS navigation (Google Maps), and the path returned by the map can be read to form the path to be driven by the car, so that the car can automatically find the way between points. That is to say, in terms of driverless driving, the driver only needs to enter the place to reach to let the smart car automatically find the best path to drive autonomously.

At present, the intelligent car has entered the experimental stage, and the driverless car can ensure the safe operation of the driverless car through the reference to the car's pathfinding and automatic driving algorithms.

7 Prospect

The research of smart cars is currently in the stage of being taken seriously, most countries are convening talents, and many companies have invested a lot of money in the research of small cars, although the harvest is not optimistic, but the research on this technology has never stopped, and even more and more enthusiastic. Regarding the development of science and technology, according to "rejuvenating the country through science and education" and "strengthening the country with science", the level of science and technology is often related to the level of development of the country.

The application of intelligent transportation can connect the car to the Internet, obtain satellite maps through big data combined with local algorithms to detect the surrounding real-time road conditions, and at the same time, because the network control of the car can realize real-time acquisition of the instructions sent by the control end, to ensure that the next new command to be processed can be received after completing a command. This treatment makes the trolley more flexible to a large extent. In the case of situations that have not been considered in advance, it also has a certain temporary processing and correction function, so as to carry out remote human intervention and realize a greater degree of automatic processing work of the trolley.

The research of this project has important research value in both life and military aspects, and can follow up with different application fields for different requirements and upgrades.

Funding. Natural Science Foundation of Anhui Province:2008085QD192; Suzhou University University-level Quality Engineering Project SZXY2020XGJY05; Suzhou College Transversal Program (No 2020xhx087); Key Research Project of Suzhou University(2020YZD01); Excellent academic and technical backbone of Suzhou University(2020XJGG01); The second batch of collaborative education projects (202002154001) in 2020; University Student Research Project (KYLXYBXM21-098).

References

1. Yue, J., Li, X.: Automatic tracking and obstacle avoidance system of intelligent trolley based on machine vision. J. Hangzhou Normal Univ. (Nat. Sci. Ed.). **19**(107(02)), 91–98
2. Wang, F., Zhang, T., Chen, Z.: An intelligent variable speed patrol line obstacle avoidance trolley. CN105005299A[P]
3. Junfeng, Y.U.E., Xiumei, L.I.: Automatic tracking and obstacle avoidance system of intelligent trolley based on machine vision. J. Hangzhou Normal Univ. (Nat. Sci. Ed.) **19**(02), 200–207 (2020)
4. Liu, H., Wang, Z.: Design of intelligent trolley based on automatic tracking and obstacle avoidance. New Gen. Inf. Technol, **002**(019), 30–35 (2019)
5. Fu, X., Gao, X.: Design and implementation of multi-functional intelligent obstacle avoidance trolley. J. Longdong Univ. **29**(113(03)), 16–19 (2018)
6. Shi, N., Zhang, Y., Qin, X., et al.: A kind of obstacle avoidance trolley:, CN210310640U[P] (2020). [6] Liu Xiaogen. Application of computer software technology in the era of big data

Effect and Compensation of Wing Deformation on Conformal Uniform Linear Array STAP

Jun-Hao Cui(✉), Duo Lu, Yan-Chuan Wang, and Zhang-Xin Chen

University of Electronic Science and Technology of China, Chengdu, China
junhaocui@std.uestc.edu.cn

Abstract. This paper give a conformal wing uniform linear array deformation model, analysing the antenna direction chart and STAP performance based on the model. The spatial structure of the antenna array will change under the influence of wing deformation, thus changing the wave path difference of the array elements. After the wing deformation, the power of antenna direction chart will be offset, the main lobe became wider and the main-to-sidelobe ratio are increasing, and the STAP performance decreased. In addition, The different incident cone angle in the detection region result in different STAP performance losses. In order to improve STAP performance after wing deformation, a compensation method combining wing deformation model and RMVB algorithm is presented. The effectiveness of the proposed method is verified based on simulation.

Keywords: STAP · wing conformal array · wing deformation · wave path difference · compensation

1 Introduction

In airborne radar, it is often to detect the location of the target by radar echo signal. Airborne radar moves fast and the Doppler frequency of clutter will greatly expand, sometimes even reach twice the pulse repetition frequency (PRF) [1]. Therefore, clutter distributed in the whole Azimuth-Doppler plane will seriously affect the radar system to detect targets. Space-time adaptive processing (STAP) [2] can effectively suppress clutter and improve the detection performance of radar to moving targets through two-dimensional adaptive filtering in both spatial and time-domain.

Conformal array airborne radar is a research hotspot in the new generation airborne radar [3], wing conformal radar is one of the important design methods which based on the wing surface not only effectively reduce the Radar Cross Section (RCS) but also not affect the original aerodynamic performance of the aircraft [4,5]. However, as the wing is the main structure of the aircraft to generate lift, it is inevitable that the wing will have certain bending and torsional deformation due to aeroelastic factors during flight, which will change the

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 134–142, 2023.
https://doi.org/10.1007/978-981-99-2653-4_17

spatial structure of the conformal antenna array and affect the overall performance of the radar system [6]. At present, most of the research focuses on how to accurately measure the wing deformation curve in real time, which can be divided into photoelectric sensor measurement method and non-photoelectric sensor measurement method [7–9]. In 2010, Beihang first proposed a method to measure wing deformation in real time using a dual laser rangefinder [10]. The deformation curve function of airfoil is given by data fitting, and the error analysis is made to show the applicability of the deformation curve function.

At present, there is no literature research on the influence of deformation on STAP performance of wing conformal radar, and wing deformation is an inevitable problem in the actual flight process of aircraft. This paper will give the conformal wing uniform linear array deformation model, and give the influence of deformation on the antenna direction chart and STAP performance. Based on the model and RMVB algorithm, a compensation method for the target steering vector is proposed, which can effectively improve the STAP performance of the deformed wing.

2 Signal Model and Problem Formulation

2.1 Signal Model

Consider a side-looking airborne radar with an uniform linear array (ULA). The N channel antenna array with inner spacing d and each antenna transmits a coherent burst of M pulses in a coherent processing interval (CPI) with the constant pulse repetition frequency (PRF) f_{PRF}. The radar transmitting wavelength is λ, and the radar is located on the airborne platform with velocity v and altitude h. Deflection angle β refers to the angle between ULA and the movement direction of the airborne platform. The clutter is modeled as a large number N_c of independent clutter patches which are uniformly distributed in azimuth angles on a given range cell under test (CUT). Suppose there is a target in the detection region, radar echo data can be expressed as,

$$H_0 : x = x_t + x_c + n \qquad H_1 : x = x_c + n \tag{1}$$

where x_t, x_c and n are the target signal, the clutter and the noise, respectively. Each clutter patch can be described by the azimuth angle θ and the elevation angle φ of the lth CUT. The spatial frequency f_{si} and the normalized Doppler frequency f_{di} of the ith clutter patch are respectively expressed,

$$f_{si} = \frac{d \cos\theta_i \cos\varphi_i}{\lambda} \tag{2}$$

$$f_{di} = \frac{2v \cos(\theta_i - \beta) \cos\varphi_i}{\lambda f_{PRF}} \tag{3}$$

The target and clutter are modeled as $x_t = \xi v_{(f_{st}, f_{dt})}$ and $x_c = \sum_{i=1}^{N_c} \sigma_i v_{(fsi, fdi)}$, where ξ and σ_i denote the random complex amplitude, $v_{(f_{st}, f_{dt})}$ and $v_{(f_{si}, f_{di})}$

denote the spatial-temporal steering vector of the hypothetic target and the ith clutter patch of the lth range cell. The space-time steering vector of the ith clutter patch is expressed as $v_{(f_{si},f_{di})} = v_{(f_{si})} \otimes v_{(f_{di})}$, where $\otimes$ represents the Kronecker product. $v_{(f_{si})} = [1, e^{-jf_{si}}, \ldots, e^{-j(N-1)f_{si}}]^T$ and $v_{(f_{di})} = [1, e^{-jf_{di}}, \ldots, e^{-j(M-1)f_{di}}]^T$ respectively represent the temporal steering vector and the spatial steering vector.

In conventional STAP framework, the optimal adaptive filter weight is set according to the criterion of maximum the output signal-interference-noise-ratio (SINR), which leads to the following constrained power minimization problem,

$$\min_{w} E\left\{\left\|w^H x_l\right\|_2^2\right\} \text{s.t.} \mathrm{w}^{\mathrm{H}} \mathrm{v}_{(\mathrm{f_{st}},\mathrm{f_{dt}})} = 1 \tag{4}$$

where $v_{(fst,fdt)}$ is the space-time steering vector of the hypothetic target. Then by using the method of Lagrange multipliers, the optimal adaptive weight vector is obtained as,

$$w_{opt} = \frac{R^{-1} v_{(fst,fdt)}}{v_{(fst,fdt)}^H R^{-1} v_{(fst,fdt)}} \tag{5}$$

where $R = R_c + R_n$ is clutter plus noise covariance matrix (CNCM). Considering that the clutter patches are mutually independent, the CNCM can be written as $\widehat{R} = E\left\{xx^H\right\} = \frac{1}{L}\sum_{l=1}^{L} x_l x_l^H$, where L is the number of CUT. In order for the SINR-loss to be less than 3 dB, the number of samples greater than $2NM$ generally [11].

2.2 Wing Deformation

The wing is the main structure of the lift of the aircraft. In the flight process, the wing will inevitably produce certain bending deformation due to aeroelastic factors, which affects the spatial structure of conformal uniform linear array. For the sake of illustration, assume that the wing deformation occur only in 2D plane. The fitting curve of wing deformation is given by Beihang university [10],

$$z(x) = \frac{h_j}{3l^4}(x^4 - 4lx^3 + 6l^2x^2) \tag{6}$$

where l is wingspan, and h_j is the real-time wingtip displacement. In practice, the wing deformation rate $\eta = h_j/l$ is generally between 0% and 20%. Consider the relative position (x_n, z_n) of each array element on the deformed conformal wing array, we give the conformal wing uniform linear array deformation model as shown in Table 1,

Figure 1(a) shows the deformed wing and the relative position of each array element that the original length of wing $l_0 = 5$ m, and the real-time deformation rate $\eta = 0.05, 0.1, 0.15, 0.2$.

The wave path difference depends on the spatial structure of the array. For azimuth θ and pitch φ in airborne radar, the incident cone angle α is

Table 1. The conformal wing uniform linear array deformation model

Given:

the original length l_0

the real-time wing deformation rate η

Step. 1:

$z'(x) = \frac{\eta}{3l^3}(4x^3 - 12lx^2 + 12l^2x)$

$l_0 = \int_0^l \sqrt{1 + (z'(x))^2}dx$

get wingspan l

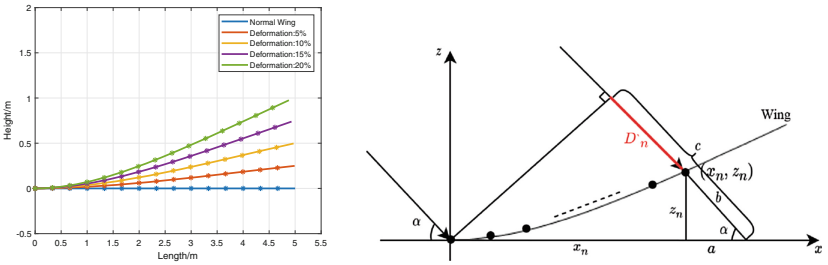
Step. 2:

$d = \int_{x_{n-1}}^{x_n} \sqrt{1 + (z'(x))^2}dx$

$z_n = \frac{\eta}{3l^3}(x_n^4 - 4lx_n^3 + 6l^2x_n^2)$

$x_0 = z_0 = 0$

get the relative position of each array element (x_n, z_n)



(a) (b)

Fig. 1. (a) Wing deformation simulation $l_0 = 5$ m (b) Diagram of wing deformation

$\cos\alpha = \cos\theta\cos\varphi$. After the wing deforms, Combined with the relative position of each element (x_n, z_n) from the conformal wing uniform linear array deformation model, the wave path difference between the array elements can be calculated by Fig. 1(b).

$$a = \frac{z_n}{\tan\alpha}, b = \frac{z_n}{\sin\alpha}, c = (x_n + a)\cos\alpha \tag{7}$$

Clearly, the wave path difference for each array element before and after the wing deformation respectively is.

$$D_n(\alpha) = (n-1)d\cos\alpha \tag{8}$$

$$D'_n(\alpha) = c - b = (x_n + \frac{z_n}{\tan\alpha})\cos\alpha - \frac{z_n}{\sin\alpha} \tag{9}$$

3 Influence of Wing Deformation

3.1 The Direction Chart

Wing deformation first affects the direction chart. The direction chart of ULA with N elements without deformation can be written as,

$$F(\theta) = \sum_{n=1}^{N} A_n f_n(\theta) e^{j\frac{2\pi}{\lambda}(D_n(\theta) - D_n(\theta_0))} \tag{10}$$

where θ_0 is the phase compensation factor when the beam points in a certain direction, and θ is the array scanning range, such as $\left[\frac{\pi}{6}, \frac{5\pi}{6}\right]$. In the ideal case, the antenna array element direction chart $f_n(\theta)$ satisfies the omnidirectional property, and its influence can be ignored in the beam scanning range, i.e. $f_n(\theta) = 1$. The complex weighting coefficient is set to $A_n = 1$ after channel consistency calibration. Thus, using (8), the ULA direction chart can be rewritten as,

$$F(\theta) = \sum_{n=1}^{N} e^{j\frac{2\pi}{\lambda}((n-1)d\cos\theta - (n-1)d\cos\theta_0)} \tag{11}$$

Consider the wing deforms, the direction chart becomes,

$$F'(\theta) = \sum_{n=1}^{N} e^{j\frac{2\pi}{\lambda}((x_n + \frac{z_n}{\tan\theta})\cos\theta - \frac{z_n}{\sin\theta} - (n-1)d\cos\theta_0)} \tag{12}$$

Given the number of array elements $N = 15$, make the beam point to $\theta_0 = 90°$ and $\eta = 0.05, 0.1, 0.15, 0.2$ get the direction chart as shown in Fig. 2. It is found the beam direction factor cannot accurately compensate for the phase offset, the major lobe becomes wider and the main-to-sidelobe ratio becomes higher, thus affecting the performance of radar.

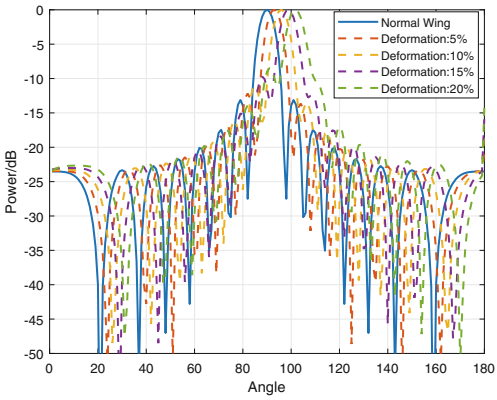


Fig. 2. Direction chart of antenna before and after deformation

3.2 STAP and Space-Time Steering Vector

The spatial frequency and the spatial steering vector can be rewrite in terms of the wave path difference between the array elements,

$$f_{si,n} = \frac{D_n(\alpha_i)}{\lambda} \tag{13}$$

$$v_{(f_{si})} = [1, e^{-i2\pi fsi,1}, \ldots, e^{-i2\pi fsi,N-1}]^T \tag{14}$$

Consider the wing deforms, the spatial frequency and the spatial steering vector becomes,

$$f'_{si,n} = \frac{D'_n(\alpha_i)}{\lambda} \tag{15}$$

$$v(f'_{si}) = [1, e^{-i2\pi f'_{si,1}}, \ldots, e^{-i2\pi f'_{si,N-1}}]^T \tag{16}$$

Using (8) and (9), can easy observe that,

$$v(fs_t) = Tv(f'_{si}) \tag{17}$$

where,

$$T = \begin{bmatrix} 1 & \cdots & 0 \\ 0 & e^{-i2\pi(fsi,1-f'_{si,1})} & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-i2\pi(fs_{i,N-1}-if'_{si,N-1})} \end{bmatrix} \tag{18}$$

So, the target spatial-temporal steering vector relation of conformal uniform linear array before and after deformation is,

$$v_{(f_{st},f_{dt})} = \left(Tv_{(f'_{st})}\right) \otimes v_{(f_{dt})} = (T \otimes I)\, v'_{(f_{st},f_{dt})} \tag{19}$$

The CNCM of the deformed conformal uniform linear array is R', thus the optimal STAP weight be obtained by,

$$w'_{opt} = R'^{-1} v'_{(f_{st},f_{dt})} \tag{20}$$

However in the real STAP system the target spatial-temporal steering vector is $v_{(f_{st},f_{dt})}$, so the resulting adaptive filter weight is,

$$w'_d = R'^{-1} v_{(f_{st},f_{dt})} = R'^{-1}(T \otimes I)v'_{(f_{st},f_{dt})} \tag{21}$$

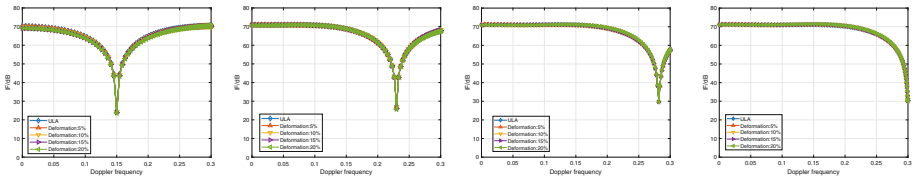
It is found that the influence of wing deformation on STAP processing can be equivalent to target spatial-temporal steering vector mismatch and the degree of mismatch is determined by the matrix T. T is determined by the deformation rate η of the wing and the echo incident cone angle α.

4 Simulations

In this section, experiments are conducted to show the influence of wing deformation and different incident angles α on STAP performance. The following improvement factor (IF) is employed as the metric for comparison,

$$IF = \frac{SINR_{out}}{SINR_{in}} = \frac{\left|w^H v_{(f_{st},f_{dt})}\right|^2 tr(R)}{w^H R w} \tag{22}$$

Unless otherwise specified, we use $N = 15$, $M = 8$ in following simulation. The number of clutter patches in a CUT is divided into $N_c = 180$ equally spaced by azimuth from 1° to 180°. The platform attitude $h = 10km$, the speed $v = 150\,\mathrm{m/s}$, the deflection angle $\beta = 90°$, and the pitching angle between the detection region and the platform is $\varphi = 30°$. The clutter-to-noise ratio (CNR) is set to 50 dB. The wingspan length $l_0 = 5\,\mathrm{m}$, and consider deformation rate η=0.05, 0.1, 0.15, 0.2. In addition, in order to discuss the influence of wing deformation on STAP, the change of antenna direction chart is not considered in the following simulation.



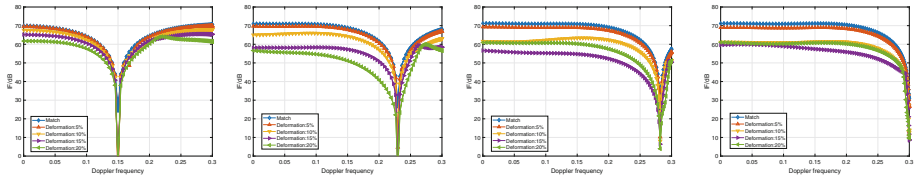
(a) $\theta = 30°$ (b) $\theta = 50°$ (c) $\theta = 70°$ (d) $\theta = 90°$

Fig. 3. The IF performance of the matched space-time steering vector before and after wing deformation

Figure 3 shows IF performance at different deformation rates η and different detection azimuth angles θ under the condition that there is no mismatch of the target steering vector. It is found that the STAP performance is no difference between ULA and the wing after deformation.

Figure 4 shows the IF performance at different deformation rates η and different detection azimuth angles θ under the condition that there is mismatch of the target steering vector. It is found that target steering vector mismatch can lead to STAP performance deteriorates. Compared with (a)–(d), θ is closer to $\theta = 30°$ the IF performance decreased slightly, and when θ is closer to $\theta = 90°$ the IF performance decreased serious. And IF performance decline is positively correlated with wing deformation rate. If η was less than 0. 05, IF performance almost did not deteriorate significantly. If η exceeds 0. 15, STAP performance decreased significantly.

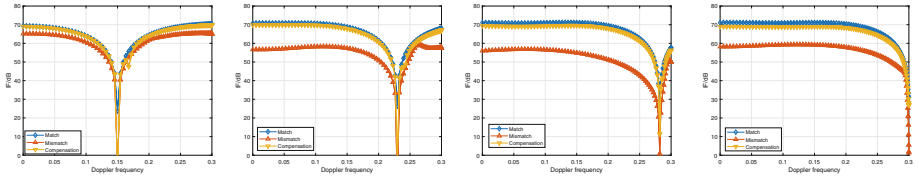
The real time wing deformation curve cannot be accurately obtained. But the original length l_0 of the wing is easy to know, so we give a compensation method



(a) $\theta = 30°$ (b) $\theta = 50°$ (c) $\theta = 70°$ (d) $\theta = 90°$

Fig. 4. The IF performance of the mismatched space-time steering vector before and after wing deformation

by the conformal wing uniform linear array deformation model in this paper and RMVB algorithm [12]. It is easy to give the uncertainty set of space-time steering vector $\Delta v' = conv\left\{v'_{(\eta=0)}, v'_{(\eta=\Delta\eta)}, \cdots, v'_{(\eta=k.\Delta\eta)}\right\}$. Where $v'_{(\eta=\Delta\eta)}$ represents the space-time steering vector at deformation rate $\Delta\eta$.



(a) $\theta = 30°$ (b) $\theta = 50°$ (c) $\theta = 70°$ (d) $\theta = 90°$

Fig. 5. The compensated IF performance

Figure 5 are the compensated IF performance based on the conformal wing uniform linear array deformation model in this paper, assume that the real wing deformation rate is $\eta = 0.145$ and $\Delta\eta = 0.01$. It is found that the STAP performance can be improved effectively by proposed method.

5 Conclusion

This paper give a conformal wing uniform linear array deformation model. Wing deformation will lead to the change of wave path difference between array elements. Experiments show that wing deformation will lead to beam pointing deviation, the major lobe becomes wider and main-to-sidelobe ratio decrease and STAP performance decreases with the increase of deformation rate. The more serious the deformation, the more serious the deterioration of STAP performance. Finally, a simple and effective compensation method is presented based on the deformation model and RMVB algorithm, which can effectively improve the STAP performance of wing deformation.

References

1. Klemm, R.: Comparison between monostatic and bistatic antenna configurations for STAP. IEEE Trans. Aerosp. Electron. Syst. **36**(2), 596–608 (2002)
2. Brennan, L.E., Reed, L.S.: Theory of adaptive radar. IEEE Trans. Aerosp. Electron. Syst. **AES-9**(2), 237–252 (1973)
3. Josefsson, L.: Conformal Array Antenna Theory And Design. Conformal Array Antenna Theory and Design (2006)
4. Hersey, R.K., Melvin, W.L., Mcclellan, J.H., Culpepper, E.: Adaptive ground clutter suppression for conformal array radar systems. IET Radar Sonar? Navigation **3**(4), 357–372 (2009)
5. Hersey, R.K., Melvin, W.L., McClellan, J.H.: Clutter-limited detection performance of multi-channel conformal arrays. Sig. Process. (2004)
6. Schippers, H., Tongeren, J., Knott, P., Deloues, T., Scherbarth, M.R.: Vibrating antennas and compensation techniques research in NATO/RTO/set 087/RTG 50. In: IEEE Aerospace Conference (2007)
7. AIAA. Shape estimation of deforming structures (2000)
8. Derkevorkian, A., Masri, S.F., Alvarenga, J., Boussalis, H., Bakalyar, J., Richards, W.L.: Strain-based deformation shape-estimation algorithm for control and monitoring applications. AIAA J. **51**(9), 2231–2240 (2013)
9. Peng, C., Zhang, W., Wang, Z., Ying, S., Fang, L.: Measurement of aircraft wing deformation using fiber Bragg gratings. In: 2017 16th International Conference on Optical Communications and Networks (ICOCN) (2017)
10. Li, X., Xu, L., Cheng, T., Tian, X.: Real-time measurement of aerodynamic deformation of wing by laser rangefinder. In: Instrumentation and Measurement Technology Conference (2010)
11. Reed, S.I., Mallett, D.J., Brennan, E.L.: Rapid convergence rate in adaptive arrays. IEEE Trans. Aerosp. Electron. Syst. **AES-10**(6), 853–863 (1974)
12. Lorenz, R.G., Boyd, S.P.: Robust minimum variance beamforming. IEEE Trans. Sig. Process. **53**(5), 1684–1696 (2005)

A Sequence Detection Technology Based on IPIX Dataset

Hao Wan, Xiaoqing Tian, and Jing Liang(✉)

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China
liangjing@uestc.edu.cn

Abstract. The traditional sea surface detector detects the weak target on the sea according to the target power, but the emergence of sea spikes and waves makes the performance of this kind of detector seriously degraded in the face of complex and changeable marine environment. We propose a sequence feature detector based on feature extraction. Firstly, three features are extracted from the time-frequency domain, phase domain and Doppler domain to distinguish sea clutter and targets. After normalization, the bidirectional long short-term memory network (LSTM) is constructed to train the data. Numerical experiments show that the performance of the feature detector is significantly better than that of other detectors, and it is more suitable for working in complex sea surface environment, but there is still room for improvement in false alarm rate.

Keywords: target detection · deep learning · IPIX Dataset · sea clutter

1 Introduction

Target detection in complex sea surface environment has always been a major challenge for radar remote sensing. The emergence of sea spikes and waves leads to three characteristics of non-Gaussian, non-stationary and non-uniform sea clutter, which seriously restricts the performance of target detection. Lately, feature-based detection technology is very popular. This type of technology attempts to mine the differentiated features between sea clutter and targets, and then classifies them by means of machine learning. On this basis, convolutional neural networks(CNN) based on deep learning are applied to target detection and recognition in complex sea environments. This type of technology focuses on high-dimensional feature extraction of echoes. In [1], Su et al. classified sea state based on radar polarization and CNN. In [2], Graph convolutional neural network and graph mapping are combined, and the performance of the proposed detector is greatly improved. A method for fast detection of clutter and interference is proposed by Zhang et al. Through the difference between clutter and interference in time frequency domain and target, the detection probability is greatly improved by using Faster R-CNN network.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 143–150, 2023.
https://doi.org/10.1007/978-981-99-2653-4_18

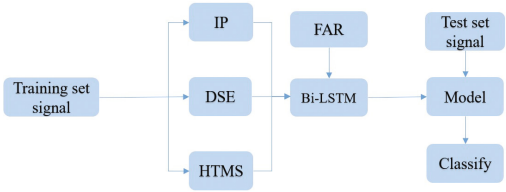


Fig. 1. Structure of the proposed method.

The above detectors have achieved good detection performance. However, the detection of radar echo signals through CNN network in the background of sea clutter needs to consider that the detection of the environment involves different characteristics, and only Using a single time-frequency feature is unreliable. Sea clutter and target become indistinguishable in time-frequency domain when the situation becomes complicated, it may even lead to misclassification. [4,5].

In view of the limitations of feature extraction, this paper extracts three sequence-features of sea clutter and target: Doppler spectral entropy (DSE), Hilbert time marginal spectrum (HTMS) and Instantaneous phase (IP) and introduces a sequence network called bidirection long short-term memory (Bi-LSTM) network, which can be regarded as a combination of multiple transition gates, bypassing certain units to memorize information for longer time steps, and has certain advantages in learning the correlation of multiple feature sequences. The features are learned and trained through the Bi-LSTM network, and then the target detection is completed (Fig. 1).

The rest of this article is structured as follows. Section 2 reviews the mathematical model of target detection on complex sea surface. Section 3 describes the proposed method in detail. The experimental results are exhibitded in Sect. 4. Finally, we summarize this article in Sect. 5.

2 Problem Modeling

The radar emits electromagnetic waves around and is emitted by the sea surface. The reflected echo we receive is divided according to the distance unit. Each distance unit contains N pulses, which we denote as $\mathbf{x} = [x(1), x(2), \ldots, x(N)]$. The received echo signal can be divided into two types, one is the clutter unit including sea surface scattering clutter and noise, and the other is the target unit including sea surface scattering clutter, noise and target. Therefore, the detection of weak targets at sea can be regarded as a binary classification problem:

$$\begin{cases} H_0 : \begin{cases} \mathbf{x}_k = \mathbf{c}_k + \mathbf{n}_k, & k = 1, 2, \ldots, K \\ \mathbf{x}_k^p = \mathbf{c}_k^p + \mathbf{n}_k, & p = 1, 2, \ldots, P \end{cases} \\ H_1 : \begin{cases} \mathbf{x}_k = \mathbf{s}_k + \mathbf{c}_k + \mathbf{n}_k, & k = 1, 2, \ldots, K \\ \mathbf{x}_k^p = \mathbf{c}_k^p + \mathbf{n}_k, & p = 1, 2, \ldots, P \end{cases} \end{cases} \tag{1}$$

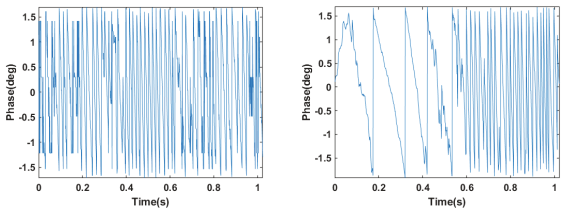


Fig. 2. IP for (a) sea clutter cell and (b) target cell

where the null hypothesis H_0 represents that the echo unit is a clutter unit. The hypothesis H_1 represents that the echo unit is a target unit. $\mathbf{x}_k$, $\mathbf{s}_k$, $\mathbf{c}_k$, and $\mathbf{n}_k$ mean the received echo, target echo, clutter echo and noise echo respectively. K is the number of distance units and P is the number of reference distance cells. Due to the strong energy of sea clutter, the noise component is ignored int the following research.

We use the public IPIX dataset to verify the proposed detector. The classification result matrix can be represented as Table 1.

Table 1. classification result matrix

Actual data	Test results	
	H_0 (Sea Clutter)	H_1 (Target)
H_0 (Sea Clutter)	TN	FP
H_1 (Target)	FN	TP

TN = true negative; FP = false positive; FN = false negative; TP = true positive.

3 Proposed Method

3.1 Instantaneous Phase

The discrepant scattering mechanisms of the clutter and the target will lead to different phase characteristics of the target unit and the sea clutter unit. The Instantaneous Phase of a radar unit $x(n)$ is written as:

$$\phi(n) = \arg\left[x(n)\right] \tag{2}$$

where arg(•) Uses to calculate the phase of target and clutter. Figure 2 shows the difference between the target and the sea clutter for the instantaneous phase feature at 1.024 s observation time. The phase change of the target is slower and more orderly than that of the sea clutter.

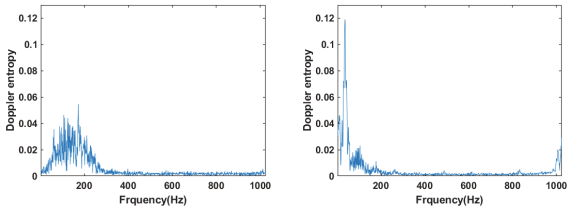


Fig. 3. DSE for (a) sea clutter cell and (b) target cell

3.2 Doppler Spectral Entropy

The Distribution of Doppler spectrum is also an important feature. For the received echo $x(n)$, its Doppler amplitude spectrum can be expressed as

$$X\left(f_d\right) = \frac{1}{\sqrt{N}} \left| \sum_{n=1}^{N} x(n) \exp\left(-2\pi f_d n T_r\right) \right| \\ -\frac{1}{2T_r} \leq f_d \leq \frac{1}{2T_r} \tag{3}$$

where f_d represents the Doppler frequency, and Tr is the pulse repetition frequency. Doppler spectral entropy can be expressed as:

$$\mathrm{DSE}(\mathbf{x}) \equiv -\tilde{X}\left(f_d\right) \log \tilde{X}\left(f_d\right), \tilde{X}\left(f_d\right) = \frac{X\left(f_d\right)}{\sum\limits_{f_d} X\left(f_d\right)} \tag{4}$$

Figure 3 shows the DSE of the sea clutter echo and the target echo. The spectrum of sea clutter unit is more dispersed, while the frequency component of target unit is more concentrated.

3.3 Hilbert Time Marginal Spectrum

The Hilbert-Huang Transform (HHT) performs linear and stationary processing on nonlinear and non-stationary data. Geting intrinsic mode function (IMF) components through empirical mode decomposition (EMD) and retains the characteristics of the data itself during the decomposition process.

HHT consists of two parts, namely EMD and Hilbert transform (HT). Find m IMFs $c_i(n)$ obtained by EMD and perform HT respectively:

$$H\left[c_i(n)\right] = \frac{2}{\pi} \sum_{k=1}^{n} \frac{x\left(n-2k+1\right)}{(2k+1)} \tag{5}$$

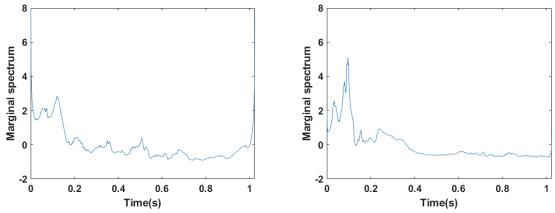


Fig. 4. Hilbert time marginal spectrum for (a) sea clutter cell and (b) target cell

$c_i(n)$ and $H\left[c_i(n)\right]$ are conjugate complex pairs, then the analytical signal $Z_i(n)$ is constructed as:

$$Z_i(n) = H\left[c_i(n)\right] + c_i(n) = a_i(n)e^{j\varphi_i(n)} \tag{6}$$

The amplitude function $a_i(n)$ and phase function $\varphi_i(n)$ are:

$$a_i(n) = \sqrt{c_i^2(n) + H^2\left[c_i(n)\right]} \tag{7}$$

$$\varphi_i(n) = \arctan\left(\frac{H\left[c_i(n)\right]}{c_i(n)}\right) \tag{8}$$

According to Eq. 8, the instantaneous frequency of $c_i(n)$ can be calculated.

$$\omega_i\left(n\right) = \frac{\varphi_i\left(n + \Delta h\right) - \varphi_i\left(n\right)}{\Delta h} \tag{9}$$

$c_i(n)$ can be expressed as:

$$c_i\left(n\right) = \mathrm{Re}\{a_i\left(n\right)e^{j\varphi_i(n)}\} = \mathrm{Re}\{a_i\left(n\right)e^{j\sum\omega_i(n)}\} \tag{10}$$

Ignoring the residual signal, the Hilbert time spectrum of the original signal is:

$$H\left(\omega, n\right) = \mathrm{Re}\{\sum_{i=1}^{m} a_i\left(n\right)e^{j\sum\omega_i(n)}\} \tag{11}$$

$H\left(\omega, n\right)$ describes the variation of signal amplitude with time and frequency. Corresponding the Hilbert time spectrum to the time axis, the time edge distribution of the Hilbert time spectrum can be obtained, which is the Hilbert time marginal spectrum.

$$HTMS\left(n\right) = \sum\nolimits_{\omega} H\left(\omega, n\right)d\omega \tag{12}$$

This feature reflects the energy accumulation of the corresponding frequency bands of the signal at different times. As can be seen from Fig. 4, compared with the HTMS of the clutter unit, the HTMS of the target presence unit has obvious spikes.

This paper uses a bidirectional LSTM model, which has the advantage of considering more sufficient information of time series data.

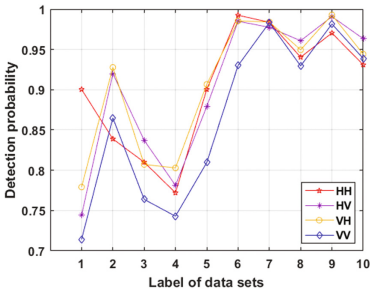


Fig. 5. Detection probabilities for the ten IPIX datasets

4 Experience Results and Performance Evaluation

To analyze the detection performance of the proposed detector under different sea conditions and radar polarizations, we performed target detection on all ten sets of data, and the results are shown in Fig. 5. We can see the detection probability of cross-polarization data is significantly higher than that of the same polarization, because in the case of cross polarization, the amplitude of sea clutter is lower, which means that our detection is easier.

We compare the proposed detector with other feature-based detectors under the HH polarization mode, and the experimental results are shown in Table 2. Comparing the detection performance of different deep learning-based detectors, it is not difficult to find that the detection probability of the proposed detector is much higher than that of other detectors. However, the false alarm rate is not so good, indicating we can further improve the detector in the future.

Table 2. Comparison of detection performance

Feature	Model	P_d	P_{fa}
STFT	VGG	0.7322	0.0006
STFT	LeNet	0.8378	0.01
Amp	LeNet	0.7865	0.1
IP + DSE + HTMS	Bi-LSTM	0.856	0.06

Next, we analyze the receiver operating characteristic (ROC) of the proposed detector in different marine environments. Figure 6 shows the ROC curves of the proposed detector under No. 2 and No. 3 data of the complex marine environment, and No. 8 and No. 9 data of the simple marine environment, and compared with Fractal detector [6], Tri-feature detector [7] and ALTG detector [8]. It can be seen that the detection probability of the proposed detector increases with the increase of the false alarm rate, and the detection performance is obviously better than that of the fractal detector and the three-feature detector. In complex

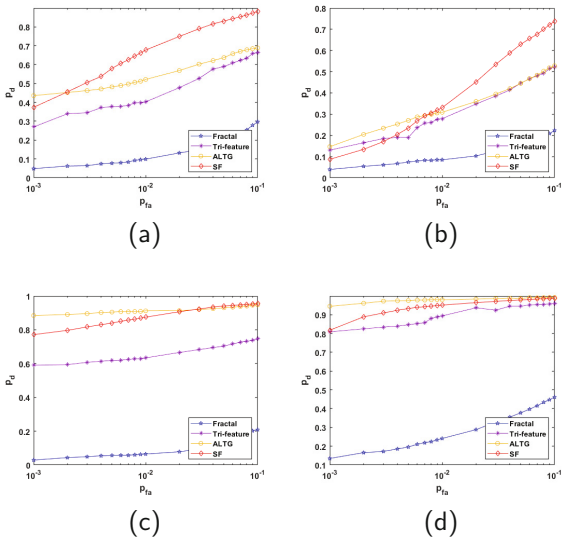


Fig. 6. Receiver operation characteristic (ROC) curves of proposed detector and the other three detector detectors at (a) No. 2, (b) No. 3, (c) No.8, and (d) No. 9 data.

marine environment, the detection probability of proposed detector is higher than that of ALTG detector. In simple marine environment, the characteristics of the sequence detector are not very good. The reason is that the ALTG detector uses the CNN network to learn the time-frequency information of target and clutter. In the case of relatively simple sea conditions, the time-frequency domain characteristics are very obvious, so the detection probability is very high. Correspondingly, when the sea state is complex and the signal-to-clutter ratio is low, the time-frequency information of the target is submerged in the clutter, and the detection effect becomes poor. The proposed detector comprehensively considers the characteristics of IP, DSE and HTMS, and this multi-domain feature joint detector has more advantages in the face of complex sea conditions.

5 Conclusion

From the perspective of deep learning, this paper proposes a sequence feature detector for the limitations of convolutional neural networks. The detector is trained using the measured IPIX data set to have the ability to distinguish sea clutter and targets in complex sea surface environments. Compared with several existing deep learning detectors, it has better detection performance and can further improve the detection performance under low false alarm rate in the future.

References

1. Su, N., Chen, X., Jian, G., Li, Y.: Deep CNN-based radar detection for real maritime target under different sea states and polarizations. In: 3rd International Conference on Cognitive Systems and Information Processing (2018)
2. Zhang, L., You, W., Wu, Q., Qi, S., Ji, Y.: Deep learning-based automatic clutter/interference detection for HFSWR. Remote Sens. **10**(10), 1517 (2018)
3. Su, N., Chen, X., Guan, J., Huang, Y.: Maritime target detection based on radar graph data and graph convolutional network. IEEE Geosci. Remote Sens. Lett. **19**, 1–5 (2022)
4. Zhang, X , Zuo, L., Wang, Y., Li, M.: An efficient method for detecting slow-moving weak targets in sea clutter based on time-frequency iteration decomposition. IEEE Trans. Geosci. Remote Sens. **51**(6), 3659–3672 (2013)
5. Wan, H., Tian, X., Liang, J., Shen, X.: Sequence-feature detection of small targets in sea clutter based on Bi-LSTM. IEEE Trans. Geosci. Remote Sens. **60**, 1–11 (2022)
6. Hu, J., Tung, W.-W., Gao, J.: Detection of low observable targets within sea clutter by structure function based multifractal analysis. IEEE Trans. Antennas Propag. **54**(1), 136–143 (2006)
7. Shui, P.-L., Li, D.-C., Xu, S.-W.: Tri-feature-based detection of floating small targets in sea clutter. IEEE Trans. Aerosp. Electron. Syst. **50**(2), 1416–1430 (2014)
8. Shi, S., Dong, Z., Yang, J., Yang, C.: Sea-surface small target detection based on autonomic learning of time-frequency graph. Syst. Eng. Electron. **43**(1), 33–41 (2021)

Research on Mid-Range Measurement of Space Non-cooperative Target Based on ToF

Zheng Yuan[1,2], Dongjun Yu[1(✉)], and Sen Li[3]

[1] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
njyudj@njust.edu.cn
[2] China Academy of Launch Vehicle Technology, Beijing 100076, China
[3] Tianjin Key Laboratory of Wireless Communication and Energy Transmission, Tianjin Normal University, Tianjin 300387, China

Abstract. This paper mainly analyzes the application of laser Time-of-flight (ToF) measurement in the measurement of medium-range distance of non-cooperative targets in space, in order to handle the problem of different measurement methods in the distance measurement of non-cooperative targets in space. Firstly, the different methods of distance measurement of non-cooperative targets in space are summarized. Then, the principle of laser measurement and the description of different laser measurement systems are introduced. The performance of pulse-ToF measurement system to measure the mid-range distance measurement is also analyzed. As the result, the relative distance measurement accuracy can meet the target approach and principle task requirements from 11 km to 20 m.

Keywords: Time-of-flight (ToF) · Non-cooperative target · Mid-range measurement

1 Instruction

Since modern times, the vigorous development of the aerospace industry in human society and the continuous progress of related technologies have supported the continuous expansion of the scope of human activities in space, from artificial earth satellites performing a single mission to large-scale space stations with expandable and multi-purpose, from near-Earth space exploration around the earth to deep space exploration represented by lunar exploration, fire exploration and exoplanet exploration missions, people have a lot of interest in spacecraft. The mission capability in space has put forward a wide variety of new requirements, among which the on-orbit rendezvous and docking mission for non-cooperative targets has become a research hotspot in the field of aerospace at home and abroad in recent years.

2 Technical Analysis of Distance Measurement for Non-cooperative Targets

The typical mission profile process related to space targets can be briefly described as "discovery → approach → task execution → retreat away". Corresponding to this

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 151–161, 2023.
https://doi.org/10.1007/978-981-99-2653-4_19

typical mission profile, the work process of the space target detection system carried by the spacecraft can also be divided into the following three stages**:**

(1) Long-distance (More than 10 km) target recognition and orientation guidance: The main tasks of the detection system at this stage include identifying specific targets through feature identification and interference factor elimination in the space environment, and completing target relative through specific algorithms. Based on the azimuth calculation of the detection platform, it is used as the input information for the long-distance relative approach/away guidance control;
(2) Relative position/velocity measurement at medium distances (20 m to 10 km): The main task of the detection system at this stage is to keep the target locked and tracked, and to continuously measure the position and velocity of the target relative to the detection platform by means of measurement, and continuously provide the detection platform with Navigation information for further approaching the target;
(3) Relative position/velocity/attitude measurement at close range (within 20 m): The main task of the detection system at this stage is to accurately measure the three-dimensional shape of the target by means of close-range high-precision measurement on the basis of relative position/velocity measurement. Identify and complete the relative attitude calculation between the target and the platform for the detection platform to carry out the task of close rendezvous and docking [1].

For the space target detection system, the spacecraft has strict weight, volume and power consumption resource constraints. It not only needs to have the initial identification of long-distance targets, but also needs to have high-precision relative pose calculation for short-range targets. Therefore, it is necessary to select a suitable measurement method under the existing technical level. According to the basic principles of various measurement techniques and the current level of sensor technology in the aerospace field at home and abroad, the main technical means suitable for the detection of non-cooperative space targets are sorted out, as shown in Table 1.

Table 1. Analysis on the technical means of space non-cooperative target detection

Detection technology	Measurable items	Applicable detection range	Measurement accuracy	Interfering factors
Optical camera	Target shape Target orientation	Medium and long range	Low	Stray light
Infrared measurement	Heat distribution Target orientation	Medium range	Medium	Heat source
Structured light measurement	3D shape	Close range	Extremely high	Stray light
Laser measurement	Distance 3D shape	Close, medium and long range	Extremely high	None

(*continued*)

Table 1. (*continued*)

Detection technology	Measurable items	Applicable detection range	Measurement accuracy	Interfering factors
Radio ehco measurement	Distance	Close, medium and long range	High	RF leakage

The research object of this paper is mainly the medium-distance measurement of space non-cooperative targets. It can be seen from the above table that laser measurement has great advantages in medium-distance measurement [2].

3 Laser Measurement

3.1 Fundamental

Laser measurement is a radar system that emits a laser beam to detect the position, speed and other characteristics of the target, sends a detection signal to the target, and then receives the signal reflected from the target and compares and processes the transmitted signal to obtain the relevant information of the target: target distance, azimuth, speed, etc., so as to realize the detection, tracking and identification of various targets.

A conventional laser measurement system is mainly composed of four parts: a laser transmitter module, a laser receiver module, an optical path control module and an information processing module, as shown in Fig. 1. The laser transmitter module will emit a laser beam of a specific wavelength as the signal source of the entire measurement system; the optical path control module modulates the emitted and received laser beams as required by means of optoelectronic techniques such as reflection, refraction, focus adjustment, and collimation control. The laser beam position, direction and beam width are determined by the laser beam receiving module; the laser receiving module will collect the laser light emitted from the target, and output the electricity that can be used for back-end information processing through the mixing of reflected light and this array light, weak signal amplification, photoelectric conversion, etc. Signal; the information processing and main control module will collect the electrical signal output by the laser receiving module, process and calculate the signal information according to the established algorithm, obtain data such as distance, speed and attitude, and control the laser emission, reception and optical path in real time according to the working status of the system and other functional modules.

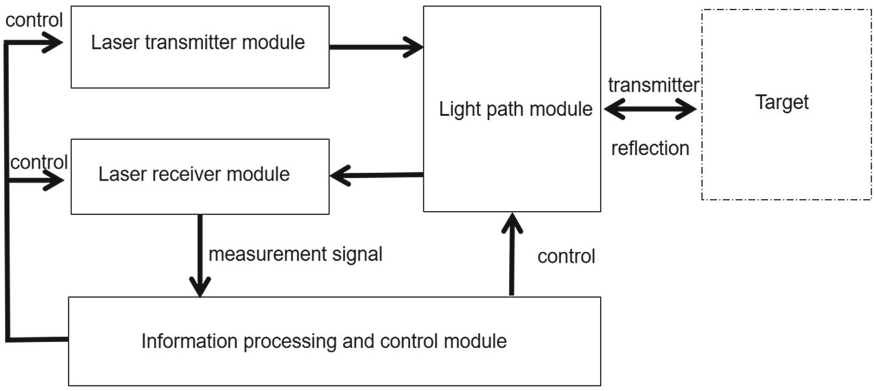


Fig. 1. Conventional Laser Measurement System

According to the specific measurement principle, the laser measurement system can be divided into: pulsed ToF measurement system, continuous wave measurement system, and laser triangulation measurement system.

The pulsed ToF measurement system is the most typical laser measurement method. The timer calculates the time difference between laser emission and reception to obtain the relative distance between the measurement device and the target. The main advantage is that the detection distance is relatively long, and the disadvantage is the accuracy of short-range ranging Slightly worse; the biggest difference between continuous wave ToF measurement and pulsed ToF measurement lies in the way of calculating the time of flight of the laser. The disadvantages are similar to the latter; laser triangulation is mainly suitable for the situation where the relative distance and pointing relationship between the laser transmitter and the detector are fixed and known. Through the geometric relationship between the pixel position on the detector and the relative distance of the target, The relative distance between the measurement system and the target can be obtained by using the proportional relationship of similar triangles. Since the position of the detector pixel can be accurate to the size of a single pixel, and the relative distance and pointing relationship between the laser emitter and the detector can be accurately measured Therefore, the measurement accuracy of this method is extremely high. The disadvantage is that due to the geometric size of the measurement system, the relative distance between the laser emitter and the detector cannot be too far, so the effective measurement distance is relatively short, and it is only suitable for short distances. Measurement.

This paper is mainly aimed at target detection at medium distances, so the pulsed ToF measurement system is selected for more in-depth research.

3.2 Performance Analysis of Pulsed ToF Measurement System

The principle of pulsed laser measurement based on ToF is shown in Fig. 2.

The specific process of the measurement method is as follows: a laser pulse (detection light source) is emitted by the laser, and after the laser irradiates the surface of the detected target, the laser beam is emitted back into space according to the reflection and scattering

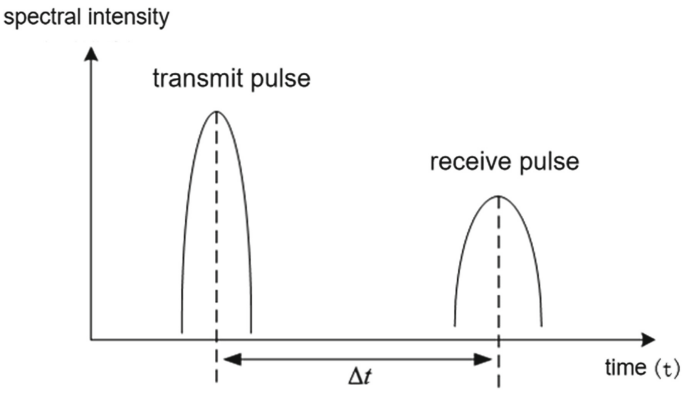


Fig. 2. The Principle of Pulsed Laser Measurement Based on ToF

characteristics of the surface, and part of the reflected light energy enter the detection light path and be captured by the laser detector. In this process, the time module of the detection system records the emission time t1 of the detection laser pulse and the time t2 when the echo pulse is detected, respectively. Through the time calculation of the ranging timer, the flight time of the optical pulse is obtained, then the relative distance is obtained by data calculation:

$$l = \frac{c \cdot \Delta t}{2} = \frac{c \cdot N \cdot T}{2} \tag{1}$$

In formula 1, c = 3 × 10^8 m/s (the speed of light), Δt is the flight time of the light pulse; N is the standard pulse number of the counter clock; T is the clock pulse width; l is the calculated relative distance between the target and the detector [3].

Considering the handover between different measurement systems and retaining an appropriate margin, the performance of the laser ToF measurement system should meet the following requirements: the detection distance is between 15 m and 11 km, and the ranging accuracy is better than ±0.1 m at 20 m, and better than ±2 m at 10 km.

Regarding the detection distance, for the laser ToF echo measurement system, it is equivalent to that the laser detector can effectively detect the echo photons at the required relative distance, and the effective detection distance can be obtained according to the laser ranging radar formula:

$$R = \left(\frac{16 \cdot \eta_q \cdot E_T \cdot S \cdot A_n \cdot A_r \cdot \eta_t \cdot \eta_r \cdot T_a^2 \cdot \alpha}{\pi^2 \cdot \theta^2 \cdot \theta_s^2 \cdot N_s}\right)^{\frac{1}{4}} \tag{2}$$

In formula 2, η_q is the quantum efficiency of the detector, η_t is the quantum efficiency of the transmitting light path, η_r is the quantum efficiency of the receiving light path, N_s is the average number of photons received by the detector, A_n is the effective reflection area of the target, A_r is the effective receiving area of the receiving light path, and T_α is the transmittance of the transmission medium. Pass rate (approximately 1 in vacuum), θ is the divergence angle of the laser beam, θ_s is the target reflection divergence angle, α

is the attenuation factor, E_T is the emitted laser pulse energy, S is the number of photons contained in the laser per joule of energy, and R is the relative distance.

The detector can effectively detect the echo reflected by the target. From the perspective of quantum theory, it is equivalent to the probability that the detector can successfully detect the number of effective photons reflected by the target. According to engineering statistical experience, if the detection probability is not less than 99.99%, Considering effective detection, the calculation method of detection probability is as follows:

$$P_d = 1 - \sum_{r=0}^{N_t-1} \frac{(N_s)^r e^{-N_s}}{r!} \tag{3}$$

In formula 3, P_d is the detection probability, N_t is the equivalent electron number of the detection threshold level of the detection device which is defined as in Eq. 2. According to Eq. 2 and Eq. 3, the relationship between the ToF effective detection distance, the laser emission power, and the detection threshold of the detection device can be calculated, and the relevant technical indicators can be determined accordingly [4].

4 Key Technology

4.1 High Repetition Frequency ToF Laser Pulse Modulation

When the ToF method is used for distance measurement, the main calculation variable is the flight time of the laser beam, and the high repetition frequency pulse modulation method can not only effectively improve the signal-to-noise ratio of the laser signal, but also shorten the interval between adjacent periodic echoes, effectively to improve the real-time measurement of the system, in the high repetition frequency modulation method, the relative time relationship between the transmitted pulse sequence and the received pulse sequence is shown in Fig. 3.

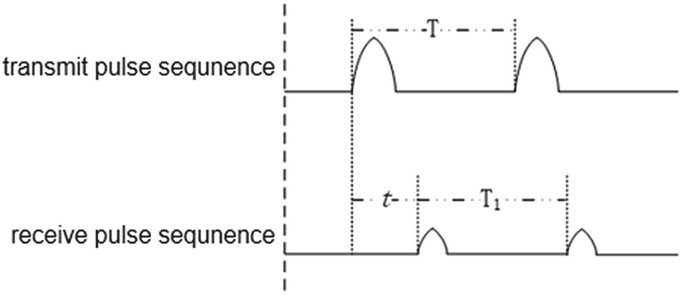


Fig. 3. High Repetotopm Frequency ToF Pulse Modulation

In the high-repetition ToF pulse modulation process shown in Fig. 3, the laser transmits the detection pulse sequence according to the period T, and the pulse period received by the detector is T1 (although the relative distance between the measurement system

and the target is not fixed, due to the mission scenarios targeted by this study, the relative velocities are not large, all in the order of 10 m/s, so it can be considered that T1 = T) Through the accumulation of multiple cycles of receiving-transmitting pulses, the measurement system completes the relative distance capture of the target. In order to ensure that the above accumulation can be carried out correctly, it should be ensured that the detector has successfully received the echo signal of the previous cycle before the next cycle of pulse transmission during pulse modulation, that is, in ToF pulse modulation, it should be ensured that:

$$t < T \tag{4}$$

In addition to the constraints described in Formula 4, in order to ensure that the ToF ranging range L meets the requirements of 15 m–11 km (considering the handover margin), the following relationship needs to be satisfied:

$$L_{\min} < L = \frac{c \cdot t}{2} < L_{\max} \tag{5}$$

According to the above two formulas, the pulse modulation period T should be no less than 73.3 μs. For the convenience of calculation, the pulse modulation period is taken as 100 μs here, that is, the pulse signal transmission frequency is 10 kHz.

From the above-mentioned high repetition frequency ToF pulse modulation method, it can be seen that the power of the laser can be effectively reduced without affecting the accuracy of the laser signal received by the detector, and the signal-to-noise of the signal can be effectively improved by self-calibration of the peak-to-valley signal strength. Ratio, improving the ToF measurement quality.

5 Laser Pulse Peak-Seeking Algorithm

The laser signal has the characteristics of narrow pulse width, so when the calculation accuracy of the moment position of the pulse peak is not enough, it will directly affect the accuracy of ranging. At present, the commonly used peak-finding algorithm is the time-expanded polynomial fitting curve method. This algorithm is not highly dependent on the positioning accuracy, but has weak resistance to external disturbances, and the peak-finding accuracy will be improved after being disturbed. Therefore, this section proposes an optimized laser pulse peak-finding algorithm. By means of multi-pulse waveform fitting and reconstruction, the position and positioning part of peak-finding is optimized to improve the laser pulse peak-finding accuracy and increase the adaptability of the algorithm.

The laser pulse is affected by a series of factors such as channel characteristics and transmission distance changes during the transmission process, as well as the impedance generated by the lidar detector during photoelectric conversion. When the echo pulse is received, the width of the left peak is usually narrower than that of the right peak., the function waveform of the echo signal is not symmetrical, the specific shape is shown in Fig. 4.

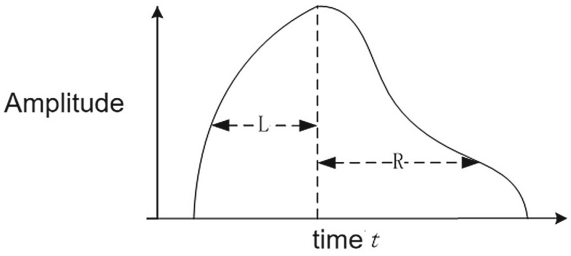


Fig. 4. Schematic Diagram of Asymmetric Pulse Shape

According to domestic and foreign engineering practice experience and theoretical research, for the laser asymmetric pulse shape, the initial error theoretical value can be estimated based on the Sinc function, and on this basis, the convolution correction method is used to correct the peak finding accuracy:

$$F(\mathrm{w}) = \int_0^w f(t-w)e(t) \tag{6}$$

In Formula 6, e(t) is the uncorrected error theoretical function, w is the function integral variable, f(t) is the error decay function, and its expression is:

$$f(t) = \begin{cases} \frac{1}{\tau}, & t \geq 0 \\ 0, & t < 0 \end{cases} \tag{7}$$

In formula 7, τ is a constant for correction, which is determined by the initial deviation positioning of the peak value and the correction compensation value of the peak value. Specifically, in the high repetition frequency ToF pulse modulation system, the new pulse signal is recorded and accumulated by the laser emission time of the main wave corresponding to the echo of each frame, and the noise in the emission signal is filtered out. The false peak points of the noise are not deleted, and are saved together with the real peak value image. The peak point range is obtained by calculating the differential signal, and the peak points in the range where the differential signal is equal to zero are all within the range of the solution. The differential signal is calculated by taking the difference of the pulse signal. The noise peaks generated under the interference may be in the range of "differential signal equals zero"; because the interference deviation is small. Therefore, in addition to the method of setting the slope threshold, the differential signal can also be filtered. The Gaussian filtering method is commonly used to filter noise. The filtering rate of this method can reach more than 85%. After Gaussian filtering, the false peak points outside the amplitude threshold are eliminated. The false peak points are basically eliminated, and finally the position of the signal waveform is reconstructed by peak point correction, and the function of generating the correction amount τ is calculated.

6 Verification

In order to verify the performance of the laser pulse modulation and peak-seeking algorithm proposed for the non-cooperative target detection system in ToF measurement, a simulation model was built under the MATLAB/Simulink platform to verify the measurement accuracy at different distances.

The main simulation models and parameters are set as follows:

(1) Target model: set the target surface as a geometric plane, high-order cosine scattering characteristics, and the average reflectivity is 0.5;
(2) Relative distance model: Set a one-dimensional linear function to represent the distance between the target and the detection system, and its slope is the relative motion speed between the two. Aspects as measured truth values;
(3) Laser model: set its emission pulse width to 100 ns, pulse period to 10 kHz, and output wavelength to 850 nm;
(4) Emission light path model: transmittance 85%;
(5) Space transmission model: Considering the transmission characteristics of laser in vacuum, the space transmission model is simplified to a one-dimensional linear function, and its output is proportional to the relative distance;
(6) Detector model: Set the sensitivity, response time and other parameters of the detector according to the design results of the ToF measurement index above;
(7) Error setting: According to the Monte Carlo method, a random number-based error factor is added to the detector to simulate the background noise of the entire measurement system;
(8) Clock setting: In order to improve the simulation accuracy of pulse rising/falling edge and peak finding algorithm, set the clock frequency to 100 MHz.

Different sections between 11 km and 15 m are selected as simulation examples, and the relative movement speed is considered as a uniform speed, as shown in Table 2.

Table 2. Design of Laser ToF Simulation Example

Working condition name	Relative distance range	Relative velocity	Remark
Condition 1	9 km~11 km	10 m/s	relatively close
Condition 2	2 km~1 km	5 m/s	relatively close
Condition 3	100 m~80 m	0.2 m/s	relatively close
Condition 4	25 m~15 m	0.05 m/s	relatively close

Carry out the simulation according to the calculation example planned in the above table, and the simulation results are shown in Fig. 5.

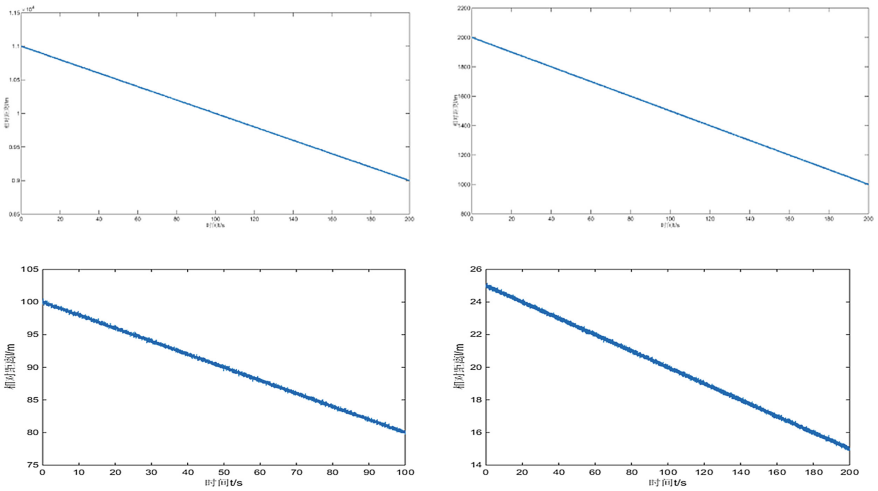


Fig. 5. Condition 1–4 Relative Distance Simulation Results (From left to right, top to bottom)

Error statistics are performed on the simulation results of the above four typical working conditions, and the results are shown in Table 3.

Table 3. Laser ToF Simulation Error Statistics

Working condition name	Relative distance range	Standard deviation	Maximum deviation
Condition 1	11 km~9 km	0.934 m	1.833 m
Condition 2	2 km~1 km	0.367 m	1.065 m
Condition 3	100 m~80 m	0.188 m	0.503 m
Condition 4	25 m~15 m	0.087 m	0.143 m

For laser ToF ranging at medium-range relative distances, after using high-frequency laser pulse modulation and laser pulse peak-finding algorithms, the relative distance measurement accuracy can meet the target approach and principle task requirements from 11 km to 20 m. Compared with existing projects technology, the accuracy has been improved to a certain extent.

References

1. Zhou, J.: Rendezvous and Docking Technology for Space Flight. National Defense Industry Press, Beijing (2013)
2. Alessia, N., Roberto, O., Giancarmine, F., et al.: Analysis of LIDar-based Relative Navigation Performance During Close-range Rendezvous Toward an Uncooperative Spacecraft[C]. IEEE. Pisa(IT): International Workshop on Metrology for AeroSpace, 2020, pp. 446–451 (2020)

3. Ventura, J., Fleischner, A., Walter, U.: Pose Tracking of a Noncooperative Spacecraft During Docking Maneuvers Using a Time-of-Flight Sensor. AIAA Scitech. Kissimmee: AIAA Guidance, Navigation and Control Conference (2015)
4. Remondino, F., Stoppa, D.: Time-Of-Flight Ranging Imaging Camera. National Defense Industry Press, Beijing (2013)

Research on Quadrotor UAV Path Planning Optimization Based on Multi-source Information Fusion Technology of Ant Colony Optimization Algorithm

Mengyu Wang(✉)

School of Information Engineering, Suzhou University, Suzhou 234000, Anhui, China
wangmengyu009@126.com

Abstract. In order to solve the problem of the lack of perceived environmental capability of the traditional quadrotor UAV(Unmanned Aerial Vehicle), a multi-sensor information fusion quadrotor UAV path planning technique based on Ant Colony Optimization Algorithm (ACO) algorithm is proposed. The path planning problem is transformed into a penalty function optimization, and global optimization is sought through the ACO algorithm. For the adaptive function including the obstacle limit and the length of the path information, the ACO algorithm finally obtains the shortest path. Through the integration of design ideas, sensor data processing data fusion and quadrotor UAV control and other key technologies, the quadrotor UAV responded in real time according to environmental changes, realizing the self-control of the quadrotor UAV. The experimental analysis shows that the multi-sensor information fusion control system used in this paper can sense the changes of the surrounding environment in real time, adjust the path planning of the quadrotor UAV in real time, and effectively improve the reliability and work efficiency of the quadrotor UAV.

Keywords: Multi-sensor · Data acquisition · Data fusion · Ant colony optimization

1 Introduction

The intelligent mobile quadrotor UAV is an important part of the UAV field. In the control of the mobile quadrotor UAV, the sensor is the medium for obtaining external information and is the core of comprehensive sensing of external information [1]. It is responsible for real-time acquisition of quadrotor UAV.

Path planning is a very important topic in the research field of intelligent quadrotor UAV. The content of this paper is the multi-sensor information fusion quadrotor UAV intelligent path planning [2, 3]. The quadrotor UAV uses infrared sensors, ultrasonic sensors and visual sensors to collect environmental information, and plans optimal or suboptimal paths based on various environmental information collected by the sensors. Performing outlier judgment and culling on the collected environmental data information

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 162–170, 2023.
https://doi.org/10.1007/978-981-99-2653-4_20

to reduce the impact of sudden erroneous data on the path planning of the quadrotor UAV, and then merging the processed data, and finally based on the fused data information to the quadrotor The machine path is intelligently planned [4, 5].

2 Theory of Multisensor Information Fusion

Multi-sensor fusion technology originated in the 1970s and began to become a systematic research field from sporadic dispersion research in the late 1980s [6]. The International Society of Information Fusion holds a conference each year to summarize the results of the field, and now has formed a relatively complet system [7].

The weighted average method is the simplest method of multi-sensor fusion. The weighted average method weights the original information of a set of redundant sensors and uses the result as a fusion value [8]. This method can process the original dynamic data of the sensor in real time.

Bayesian estimation method expresses information as a form of probability distribution. When merging multi-sensor information, it expresses the environmental information collected by each sensor as a probabilistic form, and treats independent decisions as a partition of a sample space, and then uses Bayes conditions [9]. The probabilistic algorithm processes them and performs a consistency check on the sensor measurements to ensure that the measurements represent the same entity.In this paper, the improved Ant Colony Optimization (ACO) algorithm (ACO) is used to process the data.

3 System Design

3.1 The SYstem's Overall Structure

The quadrotor UAV mainly uses multi-sensors to collect environmental information, processes the obtained data information, and uses the obtained data information to plan the moving path. Firstly, ultrasonic sensors (SRS), infrared sensors (IRS) and visual sensors are used to collect environmental information, and the collected environmental data information is separately preprocessed, and the preprocessed data is transmitted to the processor for information fusion. Finally, using the fused data, the ant colony algorithm is used to plan the quadrotor UAV path. Below, we will study and design the various modules of the system according to the structure of the system [10].

3.2 The SYstem's Functional Module Structure

The main controller unit in this article uses preprocesses the obtained data, and sends the preprocessed data to the central processing unit through the wireless communication or the serial cable, and the central processor The data is fused with the visual sensor [11].

3.3 Ultrasonic Sensor Circuit Diagram Design

Ultrasonic sensors convert ultrasonic signals into other energy signal sensors and are one of the most commonly used sensors in mobile quadrotor drones. The ultrasonic sensor module is the HC-SR04, which provides a non-contact distance measurement of 2 to 400 cm. The accuracy of the distance measurement is up to 0.3 cm. The ultrasonic sensor module mainly includes three parts: an ultrasonic transmitter, a receiver and a control circuit [12].

3.4 Vision Sensor

The vision sensor uses a CCD camera for image capture. An obstacle is an object on the ground that is higher than the threshold we set. The quadrotor drone is difficult to cross by its own motion, and the stereo vision-based method is used to detect the obstacle [10]. The main problem of detecting obstacles based on stereo vision is to compare the two images to calculate the disparity map. The key point and difficulty lies in the stereo matching of each pixel. The goal is to realize the comprehensive accuracy, anti-interference, application range and real-time of obstacle detection.

The information sampling module of visual sensor consists of CDD camera, WIFI module, steering gear and controller, as shown in the figure. CDD camera consists of image sensor, lens, automatic gain control, A/D, CDD driver, image signal processing circuit, power supply and D/A conversion circuit. The principle of visual sampling is that the light reflected by the camera object is transmitted to the lens, and the picture signal is generated on the lens to the chip. Then, after AGC automatic gain, amplification circuit and analog-to-digital conversion, a standard video signal is processed and output by the image processing chip.The main function of WIFI module is to complete the communication with the host computer and Aduino for different control and video transmission. In this way, the WIFI module is used to realize the video transmission to the host computer. First, the video encoding software encodes the video of the USB camera, and then transmits it to the PC through the WIFI module, so the collected image can be seen on the PC side. At the same time, WIFI can communicate with Arduino. It can send instructions to Arduino serial port on PC, and then control the level of IO port through Arduino serial port to drive the motor.

The control of the steering gear can make the field of vision of the CDD camera more comprehensive. Its working principle is that the controller sends control pulses to the steering gear to drive the motor to rotate. The rotation angle of the steering gear is measured by the potentiometer which rotates together at the last stage of the gear set, and then the relative rotation angle of the steering gear is controlled. The information needed to control the steering gear is the pulse width. The steering angle is 0 to 180 degrees, and the pulse width is 0.5 to 2.5 Ms. The steering gear structure and rotation principle are shown in the figure. In the visual information sampling module, the platform control of CDD camera is controlled by steering gear.

4 Odor Source Location Strategy and Algorithmic Design

In the multi-source information robot odor source localization strategy, the mobile robot odor source localization and tracking process is divided into five modules: odor source

localization search module, visual search module, visual approach module, obstacle avoidance module, odor source localization confirmation module. The input information of the sensor is the data of the chemical concentration sensor, the data of the visual sensor and the data of the infrared obstacle avoidance sensor.

(1) Search module of odor source location: Search, identify and track gas according to gas concentration.
(2) Visual search module: call the prior knowledge base to get the feature information of odor source location and search the target color.
(3) Visual approach module: If the color of odor source location is found to be a suspicious target, the mobile robot immediately approaches the target and stops at a certain distance to judge whether the suspicious target is a real odor source location source.
(4) Obstacle avoidance module: Discover suspicious target, stop gas identification; If it is odor location source, complete the task; No, obstacle avoidance.
(5) odor source localization module: when there is a suspicious target, the mobile robot stops moving, uses odor source localization model base to judge odor source localization based on fusion data; if so, stops; if not, the mobile robot continues its task.
(6) Alarm module: When the odor source is located, the alarm responds.

Rule-based odor source localization algorithm is to compare the concentration change rate and concentration measured by left and right chemical concentration sensors, and then judge the direction of motion. Finally, according to the rules, the corresponding behavior module is carried out to locate and determine odor source localization and tracking.

The odor information sensing module mainly uses MQ-3 gas sensor, and the MQ-3 type chemical concentration sensor is composed of a gas sensitive material, tin dioxide (SnO2). The odor information alarm module consists of an LED, a gas concentration sensor, a buzzer and a controller. The working principle is that when the odor concentration sensor detects that the gas concentration value is greater than or equal to 318ppm, the LED starts to flash and the buzzer sounds an alarm [12]. When the gas concentration value is less than 318ppm, the LED is off and the buzzer does not sound.

The principle of visual sampling is to pass the light reflected by the camera object to the lens, and the picture signal is generated by the CDD on the lens to the CDD chip, and then processed by the image processing chip through the AGC automatic gain, amplification circuit and analog-to-digital conversion. Video signal.

Through the analysis of the odor source model, the target features that act on the odor source localization are extracted. In this paper, the odor plume concentration and color are extracted as the target features, and the color feature extraction is extracted based on the *HSV* model. The image of the odor source localization model is calculated by *MATIAB* to obtain the feature vector F_c.The odor concentration feature extraction was performed by the $MQ-3$ odor concentration sensor and subjected to *MATIAB* pretreatment, and the concentration characteristic vector was 318 ppm.

$$F_c = [0.6806,\ 0.1158,\ -2.4865,\ 0.0564,\ 0.0247, 0.9264, 0.3847, 0.2247, 0.4531].$$

4) Setting of fuzzy density and fuzzy membership

When the characteristics of the scent source model are fuzzified, if only the affiliation of the feature is used to determine the scent source, a certain error will occur. Therefore, we will also refer to the fuzzy density for the auxiliary decision. The fuzzy density of the color moment and odor concentration in this paper is:

$$g(F_c) = 0.8$$

$$g(P) = 1$$

The characteristic membership degree of the odor source is trapezoidal distribution, as shown in Fig. 2, in which the membership of the characteristic quantity of the odor source is set to 1, and the feature quantity membership degree of the non-odor locating source is set to 0. Where a is the average value of the feature quantity, a_1 is the range in which the feature quantity membership degree is set to 1, and a_2 is the range in which the feature membership degree is set.

to 0.

5) Multi-source information fusion rules]

The fuzzy density and the membership function of the feature are fused by the formula (2–9) to obtain the fuzzy confidence of the scent source:

$\phi_i = \mu_i(x_1) \times g_i(x_1) + \mu_i(x_2) \times g_i(x_2) + \ldots\ldots \mu_i(x_n) \times g_i(x_n)$

When ϕ_i is greater than the set threshold, it can be determined as the odor localization source, where ϕ_i is the confidence of the i-th target, and $u_i(x_n)$, $g_i(x_n)$ represents the membership degree and blur density of the i-th target feature, respectively.

5 Experiment Results and Analysis

According to the search strategy requirements, the experimental platform consists of a concentration search module, a vision module, a motor control module, an infrared obstacle avoidance ranging module, and an Arduino Mega2560 controller. Among them, the motor control module is composed of double H bridge 293D motor drive expansion board, the vision module is composed of camera and WIFI module, the concentration search module is composed of MQ-3 alcohol concentration sensor, and the infrared obstacle avoidance distance measurement module is two pieces. IR2 infrared obstacle avoidance sensor. The construction of the CPT experimental platform was completed using the above hardware modules.

In order to verify the feasibility of the multi-information robot odor source location tracking strategy, two odor sources with the same color characteristics were placed in the experimental site, one with a leaking substance (odor) and the other with a smoke color-like odor source. Leaking material smell. The odor source is composed of a humidifier containing odor substances, and the experimental site is in a windless natural state. The experimental site is shown in Fig. 6. According to the experimental design requirements,

the mobile robot performs 30 experiments from the starting point to the odor source position to verify the effect of the multi-information robot odor source location tracking strategy.

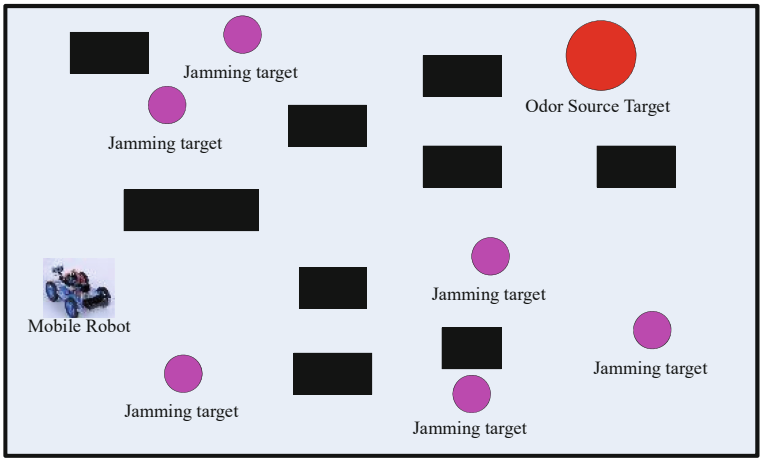


Fig. 6. Experimental site

5.1 Experiment Analysis

By using MATLAB software to draw the mobile robot odor source location road map, the coordinate origin is in the lower left (0,0) of the experimental area, the unit m, there are two red circles in the figure, one red circle position (1.4, 4) is expressed as The source of the scent, another red circle position (2.5, 2.4) is expressed as a suspected scent source, and a blue rectangle position (2.1, 0) is represented as a mobile robot, and the coordinates are also the starting point of the mobile robot, and the black rectangle is expressed as The initial distance between the obstacle, the mobile robot and the scent source is 3.92 m, and the coordinate value is compared with the actual scale. The MATLAB software is used to draw the concentration signal voltage difference curve during the actual odor source localization process. There are two curves indicating the difference of the concentration signal voltage between the current time and the sampling time, wherein the blue curve represents the change of the concentration signal voltage difference of the left sensor, the red curve Indicates the change in the concentration signal voltage difference of the right sensor.

It can be seen from Fig. 7 that when the mobile robot starts the odor source localization search, it first avoids the obstacle object, then searches for the suspected odor source, analyzes and determines the direction, and finally searches for the odor source.

It can also be seen from Fig. 8 that the concentration signal of the concentration sensor in the initial stage is almost the same, but the mobile robot still moves to the right and quickly turns to the suspected odor source. Finally, the mobile robot reaches

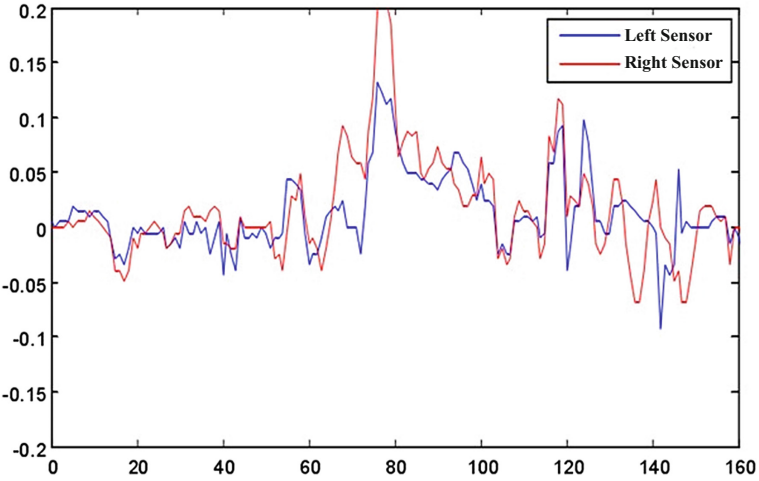


Fig. 8. Concentration signal voltage difference curve

the odor source, and the concentration of the left and right concentration sensors. Signal changes tend to be the same from huge fluctuations.

The MATLAB software is used to draw the voltage curve of the gas concentration signal collected during the search process of the mobile robot. As shown in Fig. 9, there are two undulating curves in the figure. The amplitude of the concentration signal indicated by the red curve is collected by the right sensor. The amplitude signal amplitude represented by the blue curve is acquired by the left sensor, where the time unit is 0.3 s.

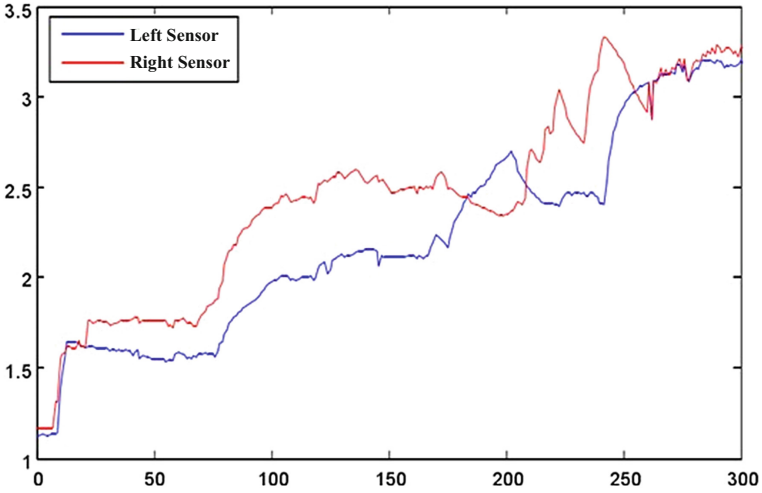


Fig. 9. Concentration signal voltage curve

Initially, due to the distance from the odor source, the gas concentration signal voltage amplitude is lower; when the 6s is, the mobile robot moves to the right and avoids the obstacle approaching the odor source, at which time the gas concentration signal voltage amplitude rapidly increases. After the 24th, as the mobile robot approaches the odor source quickly, the amplitude of the gas concentration signal voltage increases continuously; at the 60s, the mobile robot reaches the suspected odor source closer to the odor source, and the gas concentration signal voltage amplitude A large fluctuation occurred. When the mobile robot approached the odor source after 81 s, the voltage difference between the left and right concentration signals approached zero.

6 Conclusion

In this paper, we studied the control of quadrotor UAV based on multi-sensor information fusion, especially the pre-processing of the environmental information data collected by the sensors of the quadrotor UAV. The data collected by the sensor is processed and integrated, and we planed the quadrotor UAV movement path. Simulation using ant colony algorithm. The results show that preprocessing the data collected by the sensor can greatly improve the accuracy and effectiveness of the data. The improved ant colony optimization (ACO) algorithm is used for data fusion and path planning. Under the premise of accuracy, the convergence speed is accelerated.

Acknowledgment. School level Industrial College of Suzhou University (szxy2021cxxy04) top talent project of Anhui Universities (gxbjzd43), natural science research project of Anhui Provincial Department of Education (kj2018a0453), natural science research project of Anhui Provincial Department of Education (kj2018a0453), natural science research key project of Anhui Provincial Department of Education (kj2021a1110)School-Level Scientific Research Platform of Suzhou University (2021xjpt51).

References

1. Thomas, J., Ghose, D.: Strategies for locating multiple odor sources using glowworm swarm optimization. In: Proceedings of the 4th Indian International Conference on Artificial Intelligence, India: Karnataka, pp. 842–861 (2009)
2. Ishida, H., Nakayama, G., Nakamoto, T., et al.: Controlling a gas/odor plume-tracking robot based on transient responses of gas sensors. IEEE Sens. J. **5**(3), 537–545 (2005)
3. Han, G., Li, Y.: Multi-sensor information fusion algorithm research and application. In: Assembly and Symposium on, pp. 41–44 (2013)
4. Wang, Z., Zhen, Z., Hu, Y.: Information fusion based filtering for muti-sensor system. In: Control and Decision Conference, pp. 427–430 (2008)
5. Wen, Yu., Rosen, J.: Neural PID control of robot manipulators with application to an upper limb exoskeleton. IEEE Trans. Cybernetics **43**(2), 673–684 (2013)
6. Zhang, M., Cui, P.: Multi-sensor optimal robust information fusion for uncertain systems with colored noises. In: 4th IEEE Conference on Industrial Electronics and Applications, ICIEA 2009, pp. 1220–1224 (2009)
7. Cucchiara, R., Grana, C., Piccardi, M., et al.: Improving shadow suppression in moving object detection with HSV color information. In: Proceedings of the IEEE Intelligent Transportation Systems, pp. 334–339 (2011)

8. Marques, L., Nunes, U., Almeida, A.T.: Odour searching with autonomous mobile robots: an evolutionary-based approach. In: Proceedings of 11th International Conference on Advanced Robotics, Portugal: Coimbra, pp. 494–500 (2003)
9. Russell, R.A.: Locating underground chemical sources by tracking chemical gradients in 3 dimensions. In: Proceedings of the 2004. IEEE/RSJ International Conference on Intelligent Robot and Systems. USA,004, pp. 325–330. IEEE (2004)
10. Raol, J.R., Girija, G.: Sensor data fusion algorithms using square-root information filtering. Radar, Sonar and Navigation, IEE Proceedings. **149**(2), 89–96 (2002)
11. Chen, B., Li, Y., Zhang, W.-A., Liu, A.: Robust information fusion estimator for multiple delay-tolerant sensors with different failure rates. IEEE Trans. Circuits Syst. I: Regular Papers **60**(6), 401–414 (2013)
12. Zhang, Y.L., Ma, X.P., Miao, Y.Z.: Localization of multiple odor sources using modified glowworm swarm optimization with collective robots. In: Proceedings of the 30th Chinese Control Conference, China: Yantai, pp. 1899–1904 (2011)

Research on Odor Source Location Tracking Method Based on Multi-sensor Information Fusion Technology of Fuzzy Integral Fusion Algorithm

Mengyu Wang(✉)

School of Information Engineering, Suzhou University, Suzhou 234000, Anhui, China
wangmengyu009@126.com

Abstract. In order to improve the effect of odor source location tracking, multi-sensor information fusion is used to extract the odor source location tracking environment feature information, and these feature information are fused by multi-information fusion technology to improve the actual effect of odor source location tracking. Firstly, the fuzzy-integrated information fusion algorithm is used to extract the features of the odor source model from the concentration, visual detection and infrared obstacle avoidance sensors. The fuzzy membership degree and the fuzzy density are determined. The fusion rule is used to determine the target is true. Confidence; Then, with the support of the prior knowledge base and the scent source model library, the odor source location tracking strategy of multi-source information fusion is designed. Finally, the odor source location tracking experiment platform was built according to the experimental requirements, and the validity and accuracy of the odor source location tracking strategy based on multi-source information fusion were verified.

Keywords: odor source · Multi-sensor information fusion · Fuzzy integral · Fusion Algorithm

1 Introduction

In nature, many animals can use the odor source to sense the air or underwater chemicals to find food, accompaniment and defense against natural enemies [1]. Humans can also use the sensor with the detection chemical to match the robot to the "scent source positioning" of the bionic animal for odor source localization search, and finally confirm the location of the odor source.

Due to the limited information of the odor source obtained by a single sensor, the odor source information cannot be fully reflected, resulting in the robot not being able to make an efficient and accurate judgment to find the odor source [2]. Information fusion refers to the process of filtering and correlating information collected by a single sensor and multiple sensors to obtain fusion information.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 171–180, 2023.
https://doi.org/10.1007/978-981-99-2653-4_21

In this paper, according to the requirements of multi-sensor odor source location tracking, the odor source location tracking system is designed: hardware circuit including odor concentration sensor, CDD camera module, infrared obstacle avoidance sensor, wireless WIFI module, motor drive control and serial communication [3]. Then, an odor source location tracking strategy based on multi-information fusion is proposed. By using various sensors to obtain the feature information of the scent source, the scent source location tracking system has the ability to accurately locate the odor source in a complex environment.

2 Theory of Multi-source Information Fusion Based on Fuzzy Integral

Multi-source information fusion is based on the acquisition of target feature data for multiple sensors, and then the target feature vector is fused into a new feature vector. Finally, the target recognition is estimated based on the joint feature vector [4]. By analyzing the characteristics of the odor source, the definition of fuzzy density and fuzzy membership is completed. Then, the fuzzy source and fuzzy fusion method are used to obtain the odor source characteristic information [5]. The confidence of the odor source is calculated based on different feature fusions to realize the odor source localization. Self-tracking.

3 Theory of Multisensor Information Fusion

1) Fuzzy density

Let $X = \{x_1, x_2 \ldots\ldots x_n\}$, the definition function g is the density of the fuzzy measure:

$x_i \in X \rightarrow [0, 1]$, $g_i = s_\lambda(\{x_i\}), i = 1, 2 \ldots\ldots n$ then the fuzzy metric of the set g_λ is [6]:

$$g_\lambda(X) = \sum_{i=1}^{n} g_i + \lambda \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^{n} g_{i_1} g_{i_2} + \ldots + \lambda^{n-1} g_1 g_2 \ldots g_n$$

$g_\lambda(X)$ is the λ metric of the importance of any feature set X, when $\lambda \neq 0$,

$$g_\lambda(X) = \frac{1}{\lambda}\left[\prod_{i=1}^{n}(1 + \lambda g_i) - 1\right]$$

Using g(X) = 1, the constant λ can be obtained by the following formula

$$\lambda + 1 = \prod_{i=1}^{n}(1 + \lambda g_i)$$

2) Fuzzy integral
Let (X, B, g) be the space of the fuzzy metric, and $f : X \rightarrow [0, 1]$ is the measurement function of B, then the Choquet integral of the fuzzy metric is defined as follows [7]:

$$\int_A f(x)dg(\cdot) = \sum_{i=1}^{n} \left[f(x_i) - f(x_{i-1})\right] g(A_i)$$

$$f(x_0) = 0, \quad 0 \leq f(x_1) \leq f(x_2) \leq \ldots f(x_n) \leq 1 \quad \text{And } A_i = \{x_1, x_2 \ldots\ldots x_n\}$$

Let X be the set, the function of $h : x_i \in X \rightarrow [0, 1]$, and the *Sugeno*-fuzzy integral of the fuzzy metric on $A \subseteq X$ and h:

$$\int_A h(x)dg(\cdot) = \sup_{E \subseteq X} \left[\min\left(\min_{x \in E} h(x), g(A \cap E)\right)\right] = \sup_{a \in [0,1]} \left[(\alpha, g(A \cap F_a))\right]$$

where $F_a = \{x|h(x) \geq a\}$, when X is the determined set, can be calculated by *Sugeno* integral, assuming $h(x_1) \geq h(x_2) \geq h(x_3) \geq \ldots\ldots h(x_n)$

4 System Design

4.1 The System's Overall Structure

Through the analysis of multi-source information systems, the structure of the odor source location tracking system should be divided into the following modules: odor concentration sensor, CDD camera module, pattern recognition system, signal preprocessing system, multi-sensor information fusion system. In this paper, Arduino MCU is used as the system MCU [8]. Among them, there are chemical information acquisition module, infrared obstacle avoidance distance measurement module, visual information acquisition module, WiFi module, chemical information alarm module, motor drive circuit and servo drive circuit.

4.2 Infrared Obstacle Avoidance Ranging Module

The operating principle of the sensor is to detect whether there is an obstacle in front by the reflected infrared signal, where in the sensor has an infrared emitting end and an infrared receiving end [9]. The infrared obstacle avoidance distance measuring module adopts an infrared obstacle avoidance sensor for obstacle avoidance and ranging, and the program flow chart is shown in Fig. 1.

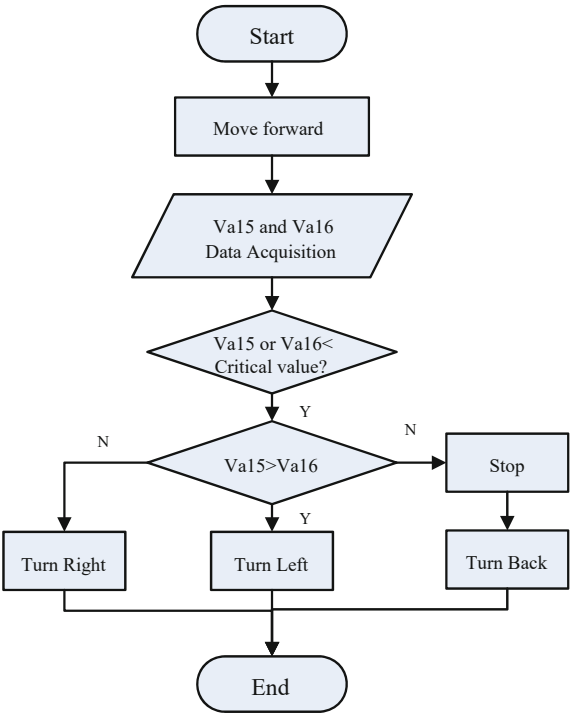


Fig. 1. Infrared obstacle avoidance ranging module program flow chart

5 Multi-source Information Fusion Odor Source Location Strategy

The odor source location tracking success is the correct identification of the odor source. In order to improve the efficiency and accuracy of tracking, a color model based odor source knowledge base and odor source model library are established, and the odor source localization tracking of multi-source information fusion is proposed. Strategy [10].

A color-based odor source knowledge base is established to retrieve the color information of the scent source from the scent source knowledge base when the scent source location tracking system detects the odor source.

1) Feature extraction based on HSV color model:

Use HSV color space. The HSV color space is represented by chromaticity, saturation, and brightness, respectively. The conversion of images from RGB space to HSV space is as follows: first set the red, green and blue coordinates of the color to (*r, g, b*,), and let

min be the smallest of r, g and b, and let max be the largest of r, g and b [11].

$$h = \begin{cases} undefined\ , \max = \min \\ 60^{\circ} \times \frac{g-b}{\max-\min} + 0^{\circ}, \max = r \quad g \geq b \\ 60^{\circ} \times \frac{g-b}{\max-\min} + 0^{\circ}, \max = r \quad g < b \\ 60^{\circ} \times \frac{g-b}{\max-\min} + 0^{\circ}, \max = g \\ 60^{\circ} \times \frac{g-b}{\max-\min} + 0^{\circ}, \max = b \end{cases}$$

$$s = \begin{cases} 0, \max = 0 \\ \frac{\max-\min}{\max} = 1 - \frac{\min}{\max}, \end{cases}$$

where: h, s, and v are the chromaticity, saturation, and brightness of the image, respectively.

Color feature matching: image matching is performed by visual tracking. When the color moment vectors of the two images are generated, the Euclidean distance algorithm is used to obtain the similarity between the color vectors.

$$D(I, J) = \left\{ \sum_{k=0}^{8} [C_I(k) - C_J(k)]^2 \right\}^{\frac{\dot{2}}{2}}$$

where: $C_I(k)$ and $C_J(k)$ represent the image I of the gallery and the color moment vector of the current image J.

2) Smell source knowledge base based on HSV color model

Alcohol as a source of odor for color feature extraction, alcohol color feature vectors and rules to establish a color model based odor source knowledge base, as long as the search for objects that match the color template to find a specific target [12].

The establishment of a knowledge base of odor source based on color model:

If the scent source tracking system detects the CL2 gas then searches for the yellow-green gas AND triggers the scent source model library 1.

If the scent source tracking system detects NO2 gas then search reddish brown gas AND triggers scent source model library 2.

If the scent source tracking system detects the gas then searches for the pale yellow gas AND triggers the odor model library n.

$$F_c = [0.690726, 0.168038, -2.60599, 0.044384, 0.0270044, 0.916986, 0.691569, 0.129749, 0.242233]$$

$$D(I, J) = \left\{ \sum_{K=0}^{8} [C_I(k) - C_J(k)]^2 \right\}^{\frac{1}{2}}$$

5.1 Establishment of Odor Source Model Library

The scent source positioning tracking system is composed of sensors such as an odor concentration sensor, a CCD camera, and an infrared sensor. The sensor system provides characteristic information for the odor source knowledge base. Because the information output by the sensor has errors, the information needs to be filtered, and the feature data is extracted to provide support for the knowledge base and decision function in the scent source model. Among them, the odor source knowledge base adopts color feature vector, distance and concentration characteristics.

1) Extraction of characteristic values of odor source concentration

The data collected by the sensor is optimized by the sliding average filtering method on MATLAB. Finally, by analyzing the optimal data distribution map, it is determined that the odor source concentration characteristic value is P = 318 ppm.

2) Odor source model library

The gas chemical information characteristic value in the knowledge base is 318 ppm, and the distance characteristic value is set to 5 cm. The color model refers to the color feature vector information and the rule-based prior knowledge base of the target object collected by the mobile robot, and the odor is set. The source color template is:

$$F_c = [0.690726, 0.168038, -2.60599, 0.044384, 0.0270044, 0.916986, 0.691569\ 0.129749, 0.242233]$$

$$D(I, J) = \{\sum_{k=0}^{8} [C_I(k) - C_J(k)]^2\}^{\frac{1}{2}} \leq 0.7096$$

The decision module in the scent source model library is mainly responsible for the fusion of the feature information of the odor source in the knowledge base and the use of the decision function to determine whether the suspect scent source is a true scent source. The knowledge base uses fuzzy integral fusion algorithm to fuse feature information, calculate the fuzzy confidence of the scent source, and use the weighted fusion algorithm to assist decision-making.

Feature layer fusion rules based on fuzzy integrals:

$$\phi_i = \mu_i(x_1) \times g_i(x_1) + \mu_i(x_2) \times g_i(x_2) + \ldots\ldots\ldots \mu_i(x_n) \times g_i(x_n)$$

Weighted fusion algorithm, expressed as:

$$g = \sum_{i=1}^{n} w_i d_i = \frac{1}{n} \sum_{i=1}^{n} d_i$$

In the formula, $n = 11$ and W_i are the weighting coefficients of the i-th sensor measurement data, and respectively represent the distance value and the odor concentration

value measured by the robot. The decision function determines whether the scent source is true or not based on the result of the feature fusion:

$$Z = \begin{cases} g \geq h \cap \phi_1 \geq q \\ g < h \cap \phi_1 = q \end{cases}$$

where h, q are the decision thresholds, h = 30.028 and q = [0, 1.8], which are important parameters of the decision function. Z = 1 is the source of the odor and Z = 0 is the source of the non-odor.

6 Experiment Results and Analysis

In the obstacle environment set above, we set the simulation parameters as: N = 200; β = 1; ρ = 0.7; C = 0.5; ε = 0.001. The simulation results are shown in the figure. As the iteration increases, it can obtain better global path planning results.When the number of iterations is 25, the path planning result and the path length change with the number of iterations are shown in Fig. 2.

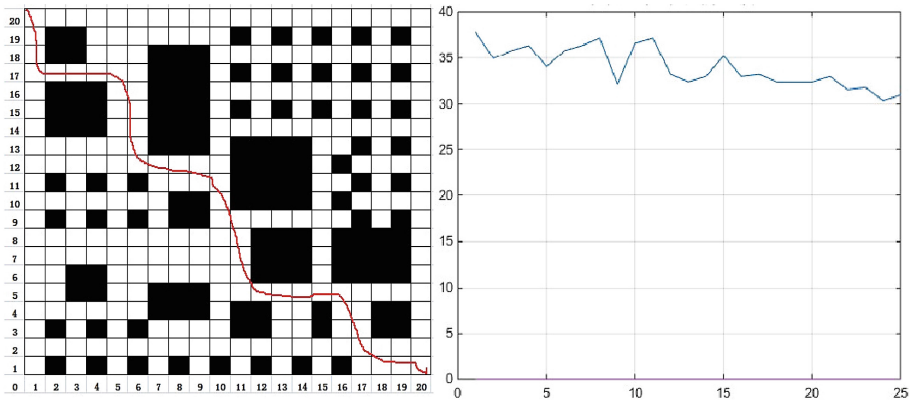


Fig. 2. The number of iterations is 25 when the path planning and path length

When the number of iterations is 50, the path plan graph and path length change with the number of iterations as shown in Fig. 3.

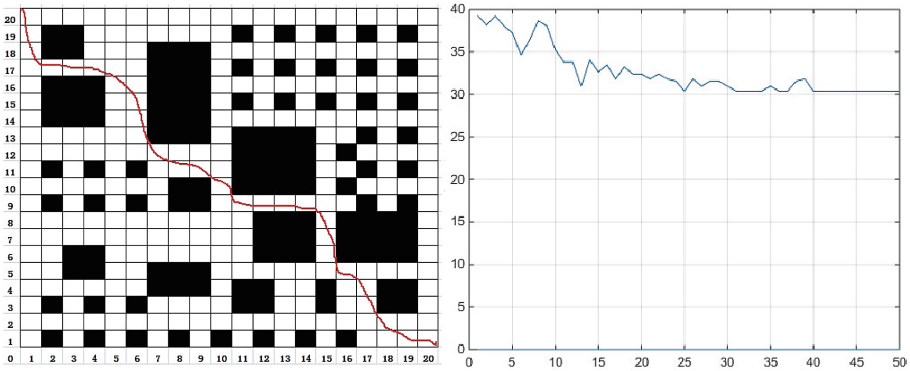


Fig. 3. Path planning and path length when the number of iterations is 50

When the number of iterations is 100, the path plan graph and path length change with the number of iterations as shown in Fig. 4:

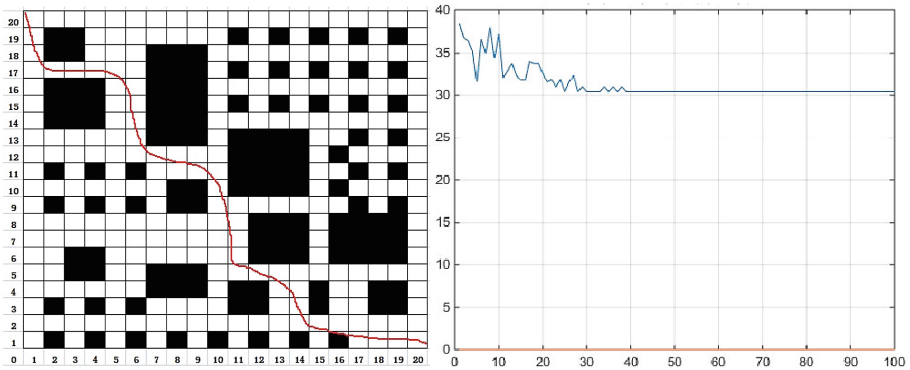


Fig. 4. Path planning and path length when the number of iterations is 100

From the simulation results, we can see that in the obstacle environment, as the number of iterations increases, the path gradually decreases and the length converges to a minimum. In this experiment, when the number of iterations is 38, the path begins to converge to Minimum value.

The convergence speed is faster.

Based on the basic principle of multi-sensor information fusion, the ant colony optimization algorithm is used to fuse the information acquired by multiple ultrasonic sensors and infrared sensors and visual sensors. The control decision of the quadrotor UAV movement is obtained, and the autonomous obstacle avoidance and navigation are realized. The simulation results show that the optimal path and faster convergence speed can be obtained by the improved ant colony (ACO) optimization algorithm in the obstacle avoidance process of the quadrotor.

7 Conclusion

In this paper, the design of the odor source location tracking system is designed, and the hardware circuit and software program of the odor source location tracking system are designed. A multi-source information fusion model for odor source localization was established, and a rule-based odor source location intelligent decision-making system was designed. By designing the rules, knowledge and reasoning mechanism, the odor source location tracking system can be autonomous in different environments. The decision performs the corresponding tracking behavior. The odor source location tracking experiment was carried out, and the experimental scheme was designed according to the requirements of the odor source location tracking strategy of multi-source information fusion. Through careful analysis of the odor source location tracking experiment results, the validity and accuracy of the multi-source information fusion odor source location tracking strategy are proved.

Acknowledgements. School level Industrial College of Suzhou University (szxy2021cxxy04) top talent project of Anhui Universities (gxbjzd43), natural science research project of Anhui Provincial Department of Education (kj2018a0453), natural science research key project of Anhui Provincial Department of Education (kj2021a1110)School-Level Scientific Research Platform of Suzhou University (2021xjpt51).

References

1. Cyrill, S.: Robotic mapping and exploration. Springer Tracts in Advanced Robotics, vol. 55, ISBN: 978-3-642-01096-5 (2009)
2. Fang, Z., Liyan, H.: A survey of multi-sensor information fusion technology. J. Telemetry Tracking Command **3** (2006)
3. D.A. Pomerleau Neural network perception for mobile robot guidance. Springer Science & Business Media (2012)
4. Li, M.A., Park, H.A., Lee, C.S.G.: Closed-form inverse kinematic joint solution for humanoid robots. In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, Taipei, Taiwan, 010, pp. 704–709
5. Di Giampaolo, E., Martinelli, F.: A passive UHF- RFID system for the localization of an indoor autonomous vehicle. IEEE Trans. Ind. Electron. **59**(10), 3961–3970 (2012)
6. Xiao, P., Luan, Y.Q., Guo, R., et al.: Research of the laser navigation system for the intelligent patrol robot. Autom. Instrum. **27**(5), 5–9 (2012)
7. Hess, W., Kohler, D., Rapp, H., et al.: Real-time loop closure in 2D LIDAR SLAM. In: 2016IEEE International Conference on Robotics and Automation (ICRA), pp. 1271–1278. IEEE (2016)
8. Qin, H., Cong, L., Sun, X.: Accuracy improvement of GPS/MEMS-INS integrated navigation system during GPS signal outage for land vehicle navigation. J. Syst. Eng. Electron. **23**(2), 256–264 (2012)
9. Gao, J.B., Harris, C.J.: Some remarks on Kalman filters for the multisensory fusion. Inf. Fusion **3**, 191–201 (2002)
10. Weckenmann, A., Jiang, X.: Multi sensor data fusion in dimensional metrology. CIRP Ann. Manuf. Technol. **58**, 701–721 (2009)

11. Rigatos, G.G.: Extended kalman and particle filtering for sensor fusion in motion control of mobile robots. Math. Comput. Simul. **81**, 590–607 (2010)
12. Besada-Portasa, E., Lopez-Orozcoa, J.A., Besada, J.: Multi sensor fusion for linear control systems with asynchronous, Out-Of-Sequence and erroneous data. Automatica **47,** 1399–1408 (2011)

Dynamic Load Forecasting and Failure Analysis of the Integrated Intelligent System

Biao Lu, Wansu Liu(✉), and Wenping Wu

Information Engineering Department, Suzhou University, Suzhou 234000, China
szxylws@ahszu.edu.cn

Abstract. The Internet of Things technology is introduced into the fault diagnosis, using its advantages in information perception, collection, and communication to ensure faults. First, a new generation of smart grid sensors and cloud platforms are used to combine a distribution line monitoring system. Second, a unified method for identifying and locating phase-to-phase short-circuits and single-phase grounding faults based on the principle of current cross-correlation is proposed. Finally, the fuzzy theory is introduced into the fault diagnosis of power transmission and transformation equipment to realize the preprocessing and local fusion diagnosis of fault information.

Keywords: Internet of Things · integrated intelligent system · dynamic load · failure analysis

1 Introduction

The Internet of Things technology has begun to be applied to smart grids, making smart grids more information-based than traditional grid operations. The existing distribution automation system emphasizes the realization of relay protection and switch opening and closing control. Among them, the remote real-time measurement data is insufficient, resulting in the lack of real-time monitoring data and visualization of the normal operation of the distribution network, and rapid fault location and isolation recovery after a fault occurs. The treatment methods are limited [1–5]. Therefore, in combination with the latest developments in modern sensing and communication technology, research the application of Internet of Things technology in distribution line monitoring and fault analysis and positioning, enrich the functions of the existing distribution automation system, and improve the economic efficiency of distribution network operation and eliminate harmonics. Reduction of the average outage time and average number of outages of the distribution network system, and enhancement of other reliability indicators are all of great significance and value [6].

Traditional box-type measuring devices, including power distribution terminals such as FTU, DTU and TTU, are mainly connected to voltage transformers and current transformers for secondary connection, requiring a large installation space, independent power supply and remote communication equipment are difficult and the investment is relatively large. Based on the high-speed synchronous sampling function of the front-end

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 181–187, 2023.
https://doi.org/10.1007/978-981-99-2653-4_22

sensors of the integrated intelligent system of the Internet of Things, this paper proposed a unified cross-correlation algorithm to realize the short-circuit and single-phase ground fault line selection and location of the distribution line. The simulation results showed that the algorithm can be better. It overcome the difficulty in setting the differential current of the main line when there are many branch load lines, and improves the accuracy and effectiveness of fault location. Finally, based on the large disturbance events recorded in the on-site network trial operation of the system, this article details the entire process of rapid identification and recovery of two 10 kV line faults and post-accident analysis, which verifies the correctness and practicability of the algorithm and system.

2 Dynamic Load Forecasting and Failure Analysis of the Integrated Intelligent System of the Internet of Things

2.1 Hierarchical Architecture of the Internet of Things

The sensor level of the Internet of Things uses induction to take power, with a 32-bit low-power processor as the core, and a global positioning system (GPS) receiver, 4G communication module, high-precision analog-to-digital converter (ADC), temperature sensor, large-capacity data storage, etc.

The central processing unit adopts a 32-bit low-power processor, which is powered by the power output from the inductive power-taking circuit. Real-time sampling of the power line current is realized by controlling the high-precision analog-to-digital converter, and the power line temperature acquisition is realized by controlling the temperature sensor. The sampled data will be marked with a unified GPS time stamp, and then cached in a large-capacity data storage. The sensor realizes the remote communication interaction with the cloud platform system through 4G wireless communication.

$$Y(x) = \sum_{x=1,y=1}^{n} p(y|x) * f(y|x) \tag{1}$$

$$g(x) - \sum_{i,j=1}^{n} (s(1,i) + s(2,i) + ... + s(j,i))/s(i,j) = 0 \tag{2}$$

The radio frequency identification system consists of three parts: an electronic tag, a reader, and a data management system. The electronic tag is composed of an antenna and a chip, and each chip contains a unique identification code and can store certain data. The reader uses the corresponding protocol to read and write the electronic tag information as needed, and communicates with the tag through electromagnetic waves to complete the reading, analysis and identification of the electronic tag information. The energy transfer between the electronic tag and the reader is realized through coupling elements. The data management system completes the storage and management of data information, and can control the reading and writing of tags.

$$|x(1) - f(x)| + |x(2) - f(x)| + ... + |x(n) - f(x)| = n * f(x) \tag{3}$$

$$\sigma(x,y) = 1/E * (\varepsilon(x,x) - t(\varepsilon(y,y))) \tag{4}$$

It is usually a self-organizing wireless network composed of a group of nodes with perception, computing and processing capabilities, and wireless communication capabilities. Through the cooperation between nodes, it collects, monitors, and processes target information in the network coverage area. WSN nodes are usually divided into ordinary nodes and sink nodes.

$$\begin{cases} \varepsilon(x,x) = \dfrac{\partial f(x,y,z)}{\partial x} \\ \varepsilon(y,y) = \dfrac{\partial f(x,y,z)}{\partial y} \\ \varepsilon(z,z) = \dfrac{\partial f(x,y,z)}{\partial z} \end{cases} \tag{5}$$

$$\frac{p'(x) - p(x)}{p(x)} - \int \frac{\partial f(x)}{\partial x} dx = 0 \tag{6}$$

By deploying multiple types of sensors, it can sense and collect comprehensive and rich online monitoring status information of power transmission and transformation equipment, and transmit it to the sink node in a multi-hop manner. Finally, the sink node transmits this information through wireless communication networks such as 2G/3G networks.

2.2 Dynamic Load Node Distribution

The phase of the zero sequence current of the faulty line should lag the zero sequence voltage by 90 degrees and be opposite to the zero sequence current of the normal line. If the zero sequence current of all lines is in the same phase, the bus is grounded. The traditional amplitude-comparison method has certain shortcomings in signal processing and anti-interference, and its use effect is not good. The modern microcomputer-based intelligent amplitude-phase method uses Butterworth digital filters with good passband characteristics to perform effective digital filtering on the signal, extracting more reliable signal components, and improving the accuracy of line selection. When the voltage is close to the maximum value, if a ground fault occurs and the capacitor voltage has a sudden change, the capacitor charge of the fault phase will be discharged to the fault point through the fault line. The distributed inductance and distributed capacitance of the fault line make the current have attenuation and oscillation characteristics, and the transient current does not pass through. The arc suppression coil is not affected by the arc suppression coil. The line selection device of this principle cannot reflect the ground fault when the phase voltage is low, and is easily affected by the operation mode, the time of the fault occurrence and the ground resistance, and there is a dead zone in operation. Figure 1 is a dynamic load node distribution network.

As an intelligent information processing technology, information fusion can intelligently synthesize multiple sources of information from the same target, and produce more accurate and reliable estimates and judgments than a single source of information. Therefore, it is necessary to introduce information fusion into the fault diagnosis of power transmission and transformation equipment, fully excavate the redundant and complementary information between various information, comprehensively describe the

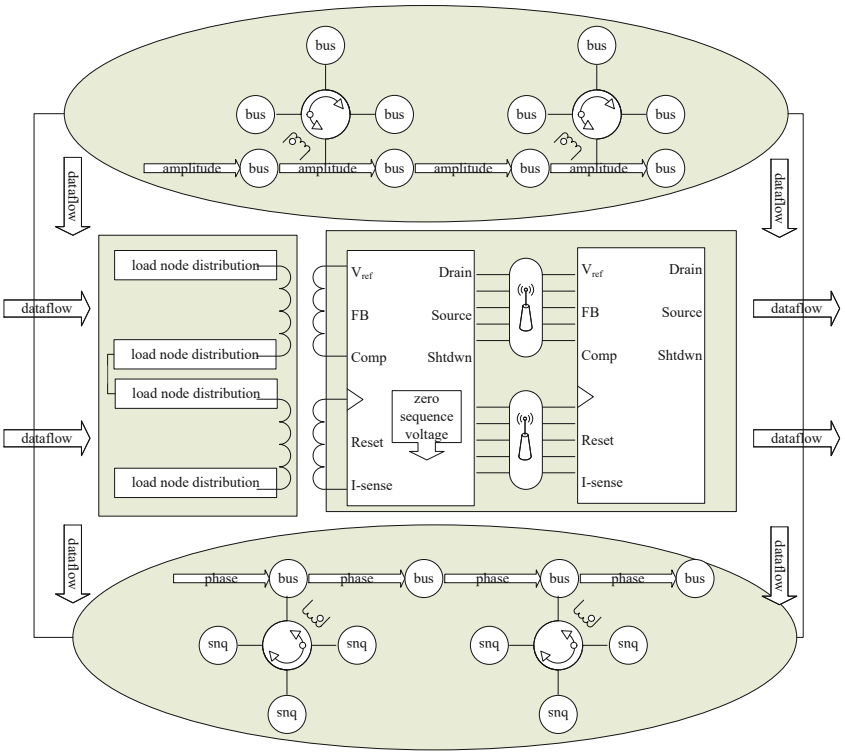


Fig. 1. Dynamic load node distribution network

diagnosis object, and effectively merge the multi-source information to improve the fault. The basic idea of fuzzy theory is to make the absolute membership in the ordinary set flexible, so that the membership degree of the element to the set can be expanded from 0 and l to any value in the interval [0, 1]. Therefore, it is suitable for the uncertainty of fault symptom information. In the fault diagnosis, there is a large amount of uncertainty between the fault mode and the fault symptom information, that is, a fuzzy relationship, and the boundary of the symptom information is also uncertain. Therefore, it is the most important thing to introduce fuzzy theory into the fault diagnosis of power transmission and transformation.

2.3 IoT Integrated Intelligent Data Preprocessing

The on-site monitoring platform of the Internet of Things system for power transmission and transformation equipment is based on the excellent monitoring software Nagios for in-depth customization, integration and secondary development. It obtains server-side monitoring data for modeling, analysis and processing, and then forms a display map. The entire system is composed of the monitored object, the monitored layer, the network layer, the monitoring layer, and the display layer. The electric load monitoring system mainly samples the working current of the electric load and calculates the power value

of the electric load. Since the power metering circuit outputs a rectangular wave, the frequency value of the rectangular wave is proportional to the power value of the electric load. Therefore, before calculating the power value of the electric load, the frequency of the rectangular wave output by the power metering circuit must be measured. Therefore, it is necessary to study the related methods of frequency measurement. In the process of feature extraction, in order to increase the richness of feature quantities, it is not enough to extract physical features such as current effective value, harmonic current and power value. For example, we calculate the mean value, variance, standard deviation, maximum value, extreme value, kurtosis, skewness, etc. of the electric load power value for a continuous period of time. The electrical load identification technology based on steady-state features is characterized by easy collection of feature quantities and simpler state analysis. This identification method can not only identify the type of electrical load, but also the operating state of a certain electrical load.

The specific functions that Nagios can achieve are manifested through configuration files. To monitor an indicator of one type, you first need to determine the plug-in that monitors the indicator. Then we add the corresponding host definition for this type in the configuration file, add the corresponding service definition for the indicator, and add the corresponding command definition for using the plug-in. Finally, after checking that the configuration file is correct, restart Nagios to realize the monitoring data collection of this indicator. The current cross-correlation value of adjacent nodes close to 1 indicates a strong correlation, indicating that the current waveforms flowing through the two are similar; a small or negative cross-correlation value indicates a weak correlation, indicating that the current waveforms flowing through the two are not similar. The cross-correlation values of the current waveforms of the A and B phases of the sensors 0205 and 0206 are very small, even negative, and the correlation coefficients of other adjacent points are close to 1. Based on this, it can be clearly judged that the fault type is A and B. The phase is short-circuited, and the fault occurs between the sensors 0205 and 0206. Since the cross-correlation analysis algorithm focuses on the similarity of the waveform, the amplitude of the current has no effect on the effectiveness of the algorithm. Compared with the pure differential algorithm, it can greatly reduce the measurement error, even the influence of the branch line, and improve the fault identification.

2.4 Simulation of Dynamic Load Forecasting System

In the dynamic load supply entrance, the hardware of the monitoring system is installed and fixed at the power entrance to collect the power information of different electric loads and upload the power information to the cloud server through the Wi-Fi. The system is composed of 4 groups (12 in total) smart grid sensors and cloud platform. It is mainly used to remotely monitor the load current, temperature and harmonic magnitude at each point of the line, and to locate the fault after a fault occurs. There are four measurement nodes along the line, each node contains a group of smart grid sensors, and the instantaneous current range of each sensor is set to 2500A. The node identification numbers (IDs) from the power source to the load side are 8000, 0205, 0206, 0201, in sequence, and between 0205 and 0206 is a section of overhead bare conductor and no branch load line. The system has been in operation for six months and has been working normally. It has recorded and warned the instantaneous overcurrent disturbance caused by the load start.

It accurately captured and automatically located a 10 kV line short-circuit fault on a heavy snowy night, and recorded the fault from its occurrence to its recovery. Figure 2 is the harmonic distribution curve of the dynamic load forecasting node.

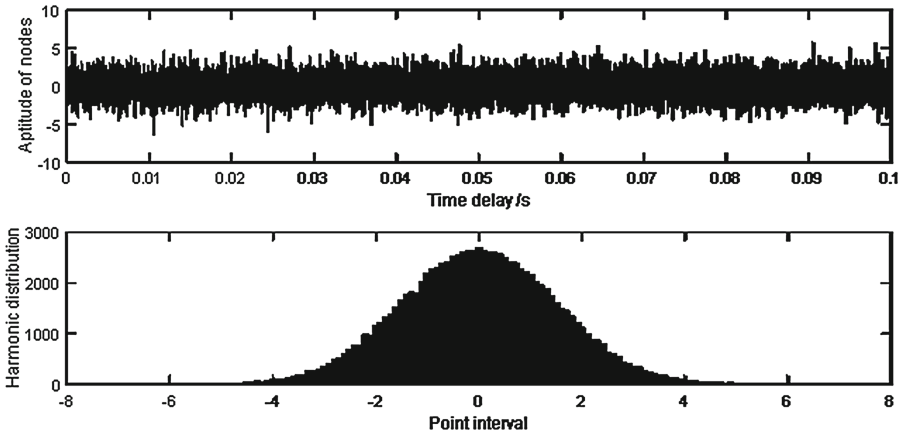


Fig. 2. Harmonic distribution curve of dynamic load forecasting nodes

From the above test results, it can be seen that the functions of monitoring local private information, adding or deleting equipment, displaying equipment operating status, and remotely controlling equipment can be realized by the on-site system monitoring platform of the Internet of Things for power transmission and transformation equipment. At the same time, during the test, it can be seen that the on-site monitoring platform interface is simple and easy to operate, and the response time is fast. Therefore, the lightweight on-site monitoring platform designed in this topic can be applied to the power transmission and transformation Internet of Things site with many information collection points and wide communication area.

3 Conclusion

The Internet of Things technology is introduced into the fault diagnosis, using its advantages in information perception, collection, and communication to ensure faults. First, a new generation of smart grid sensors and cloud platforms are used to combine a distribution line monitoring system. Second, a unified method for identifying and locating phase-to-phase short-circuits and single-phase grounding faults based on the principle of current cross-correlation is proposed. Finally, the fuzzy theory is introduced into the fault diagnosis of power transmission and transformation equipment to realize the preprocessing and local fusion diagnosis of fault information. The simulation results demonstrated the effectiveness of the proposed method.

Funding Statement. This work was supported in part by (Excellent academic and technical backbone University (No. 2020XJGG01) (Key scientific research project of Suzhou University

in 2020 (No. 2020yzd01); New engineering pilot project (szxy2018xgk05); Key disciplines of computer science and technology (2019xjzdxk1); Teaching research project (2018jyxm0960); Collaborative Innovation Center - cloud computing industry (4199106).

References

1. Chen, S., Wen, H., Wu, J., et al.: Internet of things based smart grids supported by intelligent edge computing. IEEE Access **7**, 74089–74102 (2019)
2. Zikria, Y.B., Afzal, M.K., Ishmanov, F., et al.: A survey on routing protocols supported by the Contiki Internet of things operating system. Futur. Gener. Comput. Syst. **82**, 200–219 (2018)
3. Heidari, A., Jabraeil Jamali, M.A., Jafari Navimipour, N., et al.: Internet of things offloading: ongoing issues, opportunities, and future challenges. Int. J. Commun. Syst. **33**(14), e4474 (2020)
4. Iqbal, J., Khan, M., Talha, M., et al.: A generic internet of things architecture for controlling electrical energy consumption in smart homes. Sustain. Cities Soc. **43**, 443–450 (2018)
5. Srinidhi, N.N., Kumar, S.M.D., Venugopal, K.R.: Network optimizations in the Internet of Things: a review. Eng. Sci. Technol. Int. J. **22**(1), 18–21 (2019)
6. Sahni, Y., Cao, J., Zhang, S., et al.: Edge mesh: a new paradigm to enable distributed intelligence in internet of things. IEEE access **5**, 16441–16458 (2017)

A Big Data Analysis Based Method for Sensor Array Reconstruction of Lidar 3D Images

Biao Lu, Wansu Liu(✉), and Wenping Wu

Information Engineering Department, Suzhou University, Suzhou 234000, China
szxylws@ahszu.edu.cn

Abstract. It is important to find an effective data processing method for sensor array reconstruction of lidar 3D images. Lidar point cloud is the smallest unit that describes the outer surface of targeted objects, and it is also an important factor for the accurate sensor array reconstruction of 3D target object images. Hence, how to effectively conduct lidar point cloud data classification becomes important for data processing of sensor array reconstruction. This paper suggested a sensor array reconstruction model for lidar 3D images based on big data analysis, performed the surface reconstruction of triangular meshes. The results showed the effectiveness of the proposed method.

Keywords: Big data · Lidar 3D image · Sensor array · Triangular mesh

1 Introduction

The lidar point cloud data classification technology is an important subject for airborne lidar [1]. So, how to effectively classify lidar point cloud data has become a key issue in data processing. Lidar point cloud data classification method based on big data analysis uses the law of gravitation, which converts the quality of different lidar point cloud data into the horizontal and longitudinal curvature description parameters of the laser point, and calculates the difference between different points in the lidar point cloud data [2]. Gravitational distribution, divide the data types reconstructed by the 3D image sensor array, complete the type point classification, and then perform the attribute type clustering of the classified types to complete the lidar point cloud data classification [3]. Although the 3D image sensor array is always reconstructed along the lidar in most cases, sometimes there is a smaller angle [4]. Therefore, the algorithm obtains the template by rotating the template several sampling angles, and finally calculates the largest correlation coefficient to determine the matching the position and angle of the best sub-images [5]. The difficulty of sensor array tracking lies in how to avoid mist-racking and missing tracking in the case of multiple sensor arrays, and to ensure the stability of the output data. Big data analysis adopts multi-hypothesis tracking algorithm and linear filter for data association and prediction update respectively.

Based on the analysis of previous research results, this paper proposed a sensor array reconstruction model of lidar 3D images based on big data analysis, performs the surface reconstruction of triangular meshes, explores the surface integration of triangular

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 188–194, 2023.
https://doi.org/10.1007/978-981-99-2653-4_23

meshes, discusses the system design of sensor array reconstruction for lidar 3D images based on big data analysis, and finally carries out simulation experiment and it result analysis.

2 Median Filtering Algorithm

Sensor array reconstruction of lidar 3D images generally determines the distance by emitting light pulses to the target and measuring the round-trip time x of the flight:

$$A(x) = \frac{(1 - a_x)}{(b_x - 1)} \sum_{x=1}^{n} \Delta c_x \tag{1}$$

where $A(x)$ is the distance to be measured; a_x is the propagation speed of light in vacuum; b_x is the round-trip time of the lidar signal; c_x is the average refractive index of the transmission medium; n is the stored value corresponding to the object high.

The backscattered echo received by the laser is the convolution of the transmitted pulse and the shape parameters of the obstacle. When the transmitted pulse is a wave, the echo is the superposition of wave components. The reflection time point $B(x)$ of the obstacle can be obtained by decomposition:

$$B(x) = \frac{(1 - d_x)e(x)}{(1 - f_x)g(x)} + \frac{h(x)}{k(x) - l(x)} \tag{2}$$

where d_x is the transmitted waveform; $e(x)$ is the cross-sectional equation of the obstacle; f_x is the distance between the lidar and the obstacle; $g(x)$ the backscatter waveform; $h(i)$ is the received signal; $k(x)$ is the amplification factor, and $l(x)$ is the response function.

The original cluster center is arbitrarily selected n data points. According to the nearby allocation, the data points are grouped into nearby categories, that is, the minimum solution $C(x)$ of idle nodes is calculated:

$$C(x) = \frac{m(x) + o(x)}{p(x) + q(x) - s(x)} \tag{3}$$

where $m(x)$ is the point cloud big data sample; $o(x)$ is the cluster center of all samples; $p(x)$ is the total number of all samples; $q(x)$ is the total sample data of the point cloud big data; $R(x)$ is a sample data of point cloud big data.

If the reconstruction workload of the lidar 3D image sensor array is large, or the reconstruction object is cumbersome and complicated, and the corresponding large data volume is also very large, it is very important to find an effective large data processing method. In the process of using big data analysis technology to process the sensor array reconstruction of lidar 3D images, the data processing delay is a key factor affecting the image reconstruction performance; the shorter the delay, the stronger the reconstruction performance of this method. For this reason, the sensor array reconstruction needs to use an image reconstruction method based on a multi-depth fusion model and an image reconstruction method based on quasi-phase closure.

3 Proposed Method

3.1 Data Collection and Processing

When the sensor array reconstruction of lidar 3D images based on big data analysis deals with some local minimum problems, its model shows that the average distance between two point sets will fluctuate in an area, and iteration may stop at any time. Therefore, the sensor array reconstruction changes the iteration termination condition with a sufficiently small change in the distance error between the two point sets to a sufficiently small translation distance. This not only solves the local minimum problem to a certain extent, but also avoids the sudden stop of iteration due to the randomness of the point set distance fluctuation caused by sampling. The percentage judgment condition of big data analysis can solve the local minimum problem together with the dynamic percentage calculation algorithm, so the overall algorithm effect and robustness are greatly improved. The sensor array reconstruction of lidar 3D images based on big data analysis can calculate the distance from the point in the clump except the starting point and the end point to the straight line with the starting point and the end point as the end point; if the distance is greater than a set threshold, the distance is the point before the current point is a clump, otherwise all the points up to the current point belong to the same clump. The other way is to extract the data from straight lines and planes to obtain the plane model of the current frame. Finally, the plane model of the current frame and the existing plane model are merged through the obtained position change to enlarge the entire scene.

3.2 Deep Data Matching

If all the distances between all the point cloud data in the 3D image sensing array are calculated and analyzed, it will not be realized due to excessive calculation workload. For this kind of problem, big data can be used to analyze the representation method of the neighbor relationship, establish the index structure between the point clouds, calculate the distance of the same type of 3D image sensor array, and obtain the difference between other point clouds by the index method. To analyze the degree of influence of unilateral error on system error, it is necessary to transform the coordinates of the rigid body into the geodetic coordinate system for analysis. It can ensure that under the condition that other conditional parameters remain unchanged; a single parameter is changed to determine the degree of its influence on the whole (Fig. 1). This denoising method based on the combination of statistical filtering and radius filtering can effectively remove the point cloud noise points with larger elevation values, and also effectively remove and smooth the noise point clouds with smaller elevation values.

Big data analysis can define the following sub-structures and their relationships and each 3D image contains rows and columns; rows are defined by the minimum and maximum sizes, and must be located in the columns. The lidar 3D image is composed of a series of columns, and the course is composed of a single horizontal plane. The horizontal plane is closed by the horizontal plane boundary, and the intersection of the two horizontal planes is a common horizontal plane contour. All the line segments that belong to only one horizontal plane are called the single horizontal contour, which

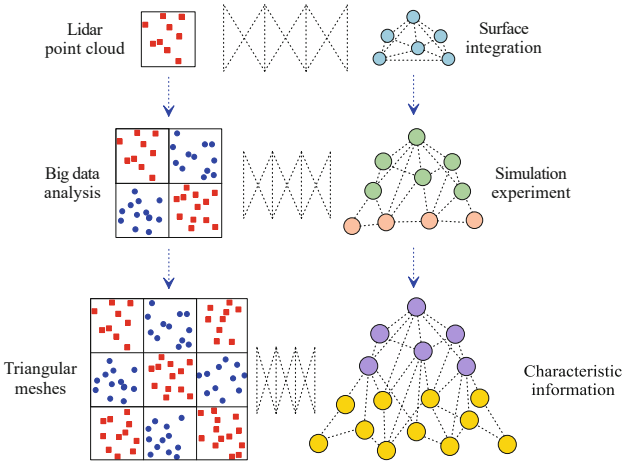


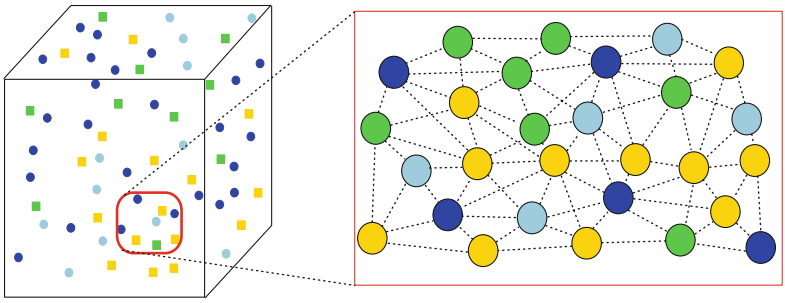
Fig. 1. Sensor array reconstruction model

is a part of the closed horizontal boundary. Through the above description of spatial topological relationship, the spatial topological relationship of 3D image components can be strictly defined, which provides model support for 3D image extraction.

3.3 Surface Integration of Triangular Mesh

Since the original 3D lidar image is obtained by triangulation between multiple frames of a monocular camera, the solution of this process is a numerical optimization process, and there are numerical solution errors. In addition, for feature points with greater depth, due to their smaller parallax, the matrix for optimally solving the displacement is close to ill-conditioned, so the result obtained by the optimal solution is noisy. In order to simulate the planar lidar, the spatial points need to be projected and filtered to obtain two-dimensional discrete data points. In view of the high degree of dispersion and high noise of the sensor array reconstruction, the distribution expression of the projection data of the frame is obtained through the statistics of the set of data points and distribution. In big data analysis, for the situation of missing data texture information and high degree of dispersion, the neighborhood distribution statistics of the filtered data null points are performed, and the expression is used to perform linear interpolation, and at the same time, the interpolation feasibility criterion is established for judgment [6]. The reconstruction of the 3D image sensor array performs planar projection on the spatial point cloud obtained by big data analysis, and performs single-laser filtering according to the light propagation principle of a single laser. It uses the processed global data points to express the difference distribution between statistical data points, and uses this distribution to filter the global data so as to retain only the data with obvious geometric distribution characteristics as Fig. 2 shows.

The core is to establish the Lidar capture image probability calculation model, and establish the lidar sensor model parameter description accordingly. The 3D image sensor array requires a calculation model and an image pixel calculation model. According to



(a) Data collection and Processing (b) Deep data matching

Fig. 2. Surface reconstruction and integration of triangular meshes

the simulation results, the general rules for the application of the lidar simulator model are given and applied to the distributed interactive simulation system.

4 Simulation Results

In order to test the application performance of the sensor array reconstruction of lidar 3D images method based on big data analysis in the realization of image reconstruction and processing, this research conducted simulation experiment analysis. The width of the dynamic scene is 12; the sampling pixel value was 120 × 120; the original image input of the dynamic scene used the lidar 3D image with the size of 180 × 120 pixels; the image imaging used the 3D laser Scanning technology is carried out; the scale segmentation coefficient is 0.18; the maximum number of iteration steps is 1200; the block threshold of image fusion is 0.22. According to the above simulation environment and parameter settings, big data analysis obtains the experimental environment for laser image reconstruction of the scanned original dynamic scene. Big data analysis performs sensor array reconstruction and dynamic simulation on the lidar 3D image, performs grid area segmentation on the collected dynamic scene image in the laser transmission space, and extracts the brightness and color difference characteristic information [7, 8]. Global splicing mainly relies on establishing a global coordinate system to unify all point cloud information from different perspectives into the global coordinate system. This splicing method can optimize the distribution of errors to each local area, so higher splicing accuracy can be obtained.

When the array image correlation is used, the photographic baseline of the lidar 3D image sensor array is not unique, so the array with the same name cannot be determined by the conventional method of defining the array. Big data analysis takes the left image as the reference image, and uses each scan line on it as the left array; a set of images with the same name on the right image of all pixels on a certain left array. The left array is a straight scan line, but the right array is a curved line, and the degree of curvature varies with the attitude of the sensor and the topography. The lidar 3D image is for each pixel on a left array, first obtain its corresponding ground point position according to the collinear equation, and then use the collinear equation to calculate the constellation of the ground tile point on the right image coordinates. The 3D image sensor array

can be approximately considered that the ordinate of the pixel on the right array is a polynomial function of its abscissa, and the polynomial coefficients are obtained from a certain number of known image points with the same name by polynomial fitting. Obviously, if the array with the same name determined by this method is used for array correlation, after obtaining the image points with the same name, as the initial value of the correlation, the big data analysis can receive satisfactory results (Fig. 3).

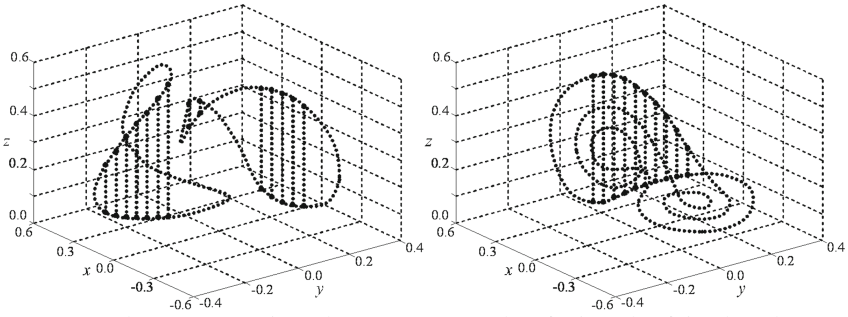


Fig. 3. Surface reconstruction (a) and integration (b) results of triangular meshes in system design of proposed method

5 Conclusion

This paper proposed a big data analysis based method for sensor array reconstruction of lidar 3D images. It analyzed data collection and processing, conducts deep data matching, proposes a sensor array reconstruction model of lidar 3D images based on big data analysis, and finally carries out simulation experiment and it result analysis. The principle of big data analysis is to set different parameter levels for each structural factor on the basis of the set reference model, and reconstruct them separately to obtain the optimal value under different parameter settings. Big data analysis shows the discrete echo and full waveform results obtained by the small spot airborne lidar when scanning different objects, which can also be seen as a process of sorting the sensor array from disorder to order. The study results of this paper provide a reference for further researches on the sensor array reconstruction of lidar 3D images based on big data analysis.

Funding Statement. New engineering pilot project (szxy2018xgk05); Teaching research project (2018jyxm0960); Collaborative Innovation Center - cloud computing industry (4199106); Key disciplines of computer science and technology (2019xjzdxk1).

References

1. Tachella, J., et al.: Real-time 3D reconstruction from single-photon lidar data using plug-and-play point cloud denoisers. Nat. Commun. **10**(1), 1–6 (2019)
2. Padmanabhan, P., Zhang, C., Charbon, E.: Modeling and analysis of a direct time-of-flight sensor architecture for LiDAR applications. Sensors **19**(24), 5464 (2019)
3. Song, Y., Zhu, Y., Nan, T., Hou, J., Du, S., Song, S.: Accelerating faceting wide-field imaging algorithm with FPGA for SKA radio telescope as a vast sensor array. Sensors **20**(15), 4070 (2020)
4. Tobin, R., Halimi, A., McCarthy, A., Soan, P.J., Buller, G.S.: Robust real-time 3D imaging of moving scenes through atmospheric obscurant using single-photon LiDAR. Sci. Rep. **11**(1), 1–13 (2021)
5. Jiang, Y., Karpf, S., Jalali, B.: Time-stretch LiDAR as a spectrally scanned time-of-flight ranging camera. Nat. Photonics **14**(1), 14–18 (2020)
6. Hansman, R.J., Ring, U.: Workflow: From photo-based 3-D reconstruction of remotely piloted aircraft images to a 3-D geological model. Geosphere **15**(4), 1393–1408 (2019)
7. Morrison, C., Brown, B., Lin, D.Y., Jaarsma, R., Kroon, H.: Analgesia and anesthesia using the pericapsular nerve group block in hip surgery and hip fracture: a scoping review. Reg. Anesth. Pain Med. **46**(2), 169–175 (2021)
8. Abdelghany, M.S., Ahmed, S.A., Afandy, M.E.: Superficial cervical plexus block alone or combined with interscalene brachial plexus block in surgery for clavicle fractures: a randomized clinical trial. Minerva Anestesiol. **87**(5), 523–532 (2021)

Research on Similarity Recognition of Face Images Based on Data Dimensionality Reduction

Shanshan Li, Wenquan Tian(✉), Chengfang Tan, and Biao Lu

Suzhou University, Suzhou 234000, Anhui, China
1106819868@qq.com

Abstract. Feature dimensionality reduction can remove redundant and useless information in data and improve the accuracy of pattern recognition. Based on this, a face image recognition algorithm based on singular value decomposition is proposed. Firstly, the feature extracted data is decomposed by SVD, and the top N values with the largest feature are retained, and the sample data after feature reduction is obtained. Then, through calculation, the category corresponding to the maximum similarity is determined, and the similarity recognition of face images based on data dimensionality reduction is realized. Through experiments on the face image data set, the experimental results show that the algorithm has a good effect.

Keywords: singular value decomposition · similarity measure · face image recognition · dimension reduction

1 Introduction

Face recognition [1] is a biometric identification technology [2] based on human facial feature information, which uses computer technology for analysis and comparison to identify faces. In the application of face recognition, the required facial feature information can be extracted by the computer. The quality of the image features determines the expression of the image information and the accuracy of the final recognition. In order to improve the effectiveness of feature expression, fine-grained description of images is required, but with the deepening of image analysis, the problem of increased feature dimension will arise, resulting in a dimensional disaster. The increase of the image dimension will also face the increase of invalid features. In order to solve this problem, it is necessary to perform feature dimension reduction [3, 4] on the collected features. In order to improve the features in the image and improve the stability of the feature information expression. Singular Value Decomposition (SVD) [5–7] is used to reduce the dimension of face image features, and transform the original features by transforming the original features.

In terms of image classification, ORL face image dataset is used, and each dataset corresponds to a label category. A face image recognition algorithm based on singular value decomposition is proposed. Firstly, SVD is used to calculate the singular value of the original image feature data, and the first k singular values are selected to realize

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 195–201, 2023.
https://doi.org/10.1007/978-981-99-2653-4_24

feature dimension reduction. Then, the data after dimensionality reduction is used as input, and the calculation is carried out through the cosine similarity model. The final maximum similarity is the category corresponding to the predicted image. The experimental results show that the algorithm can effectively predict the image category and realize the calculation of the similarity measure of the image.

2 Related Works

2.1 SVD Algorithm

Assuming that there is a training matrix $A \in R^{m \times n}$ containing m samples and n features, there is a decomposition that satisfies:

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^{T} \tag{1}$$

So the original matrix A is decomposed into the product of U, Σ, V^T matrices, we define this as singular value decomposition. This decomposition process is called singular value decomposition. According to the principle of principal components, the larger the value of SVD, the more information it contains. Based on this idea, only the first k features are important, and the rest are unimportant features. Therefore, singular value decomposition can be used for dimensionality reduction or denoising processing of data to achieve the extraction of important features in the data.

2.2 Cosine Similarity Metrics

Two vectors are determined by calculating the cosine value of the cosine similarity [8] to determine whether they are similar. For two vectors A and B, A_i and B_i are the respective components of A and B, and the calculation formula is:

$$\text{similarity} = cos\theta = \frac{\sum_{i=1}^{c} \boldsymbol{A_i B_i}}{\sqrt{\sum_{i=1}^{c} \boldsymbol{A_i^2}} \sqrt{\sum_{i=1}^{c} \boldsymbol{B_i}^2}} \tag{2}$$

Since the similarity between images is only positive correlation, the cosine value of cosine similarity is [0,1]. The higher the similarity of two vectors, the closer the cosine value is to 1. The more the cosine value tends to 0, the lower the similarity of the vector.

3 Face Image Recognition Based on SVD and Cosine Similarity

3.1 Face Image Data Dimensionality Reduction Processing Based on SVD

For the image features matrix R, R is an m × n face feature matrix, and the SVD decomposition method is used to decompose the matrix R into the product of three matrices, as shown in formula (3):

$$R_{m \times n} = U_{m \times r} \cdot S_{r \times r} \cdot V_{r \times n} \tag{3}$$

Among them, U is regarded as an $m \times r$ orthogonal matrix, V is regarded as an $r \times n$ orthogonal matrix, S is regarded as a diagonal matrix, and the elements on the diagonal are singular values. The first singular value is the largest, which represents the direction that contains the most information. With the change of the singular value, the content of useful information in the data also changes. Therefore, using the first k values can effectively describe the sample data set and realize the dimensionality reduction of the data set.

In the data dimensionality reduction method, the matrix singular value feature can realize local feature extraction and dimensionality reduction of high-dimensional features. At the same time, singular value decomposition has the data processing ability under the global optimum, because the top 20% to 30% of the largest eigenvalues can effectively retain most of the information of the data set, so this paper uses SVD for the face [9, 10] image feature data set to realize the data analysis. The dimensionality reduction process of SVD calculates the eigenvalues of the multi-label feature data, and retains the top 20% of the largest eigenvalues to realize the dimensionality reduction of the feature data by SVD.

The original data set R is an $m \times n$ matrix, which becomes $\boldsymbol{R_D}$ after feature dimension reduction, and $\boldsymbol{R_D}$ is an $m \times k$ matrix, where $k \ll n$.

3.2 Image Classification Algorithm Based on Cosine Similarity

The algorithm selects the most similar class by measuring the similarity between the test images and target images. The basic idea is: for a given set of test images, through the cosine similarity measure [11], the features of the test images are compared with the features of the target images, and then the nearest k data are found as judgments, and the categories of these neighboring data are judged by probability accumulation. Finally, the label with the highest probability is given.

The matching experiments on images are designed to evaluate the similarity of images by optimizing the cosine similarity. The experiment is to calculate the cosine similarity of a test label corresponding to other labels, calculate all cosine similarities according to the label, and sum up the cosine similarity values of each label. Then, all labels are reordered to establish the horizontal axis, and the sum of cosine similarity is used to establish the vertical axis to generate a probability map similar to the label distribution. When inputting image data, the output is the image corresponding to the maximum value of the sum of cosine similarity.

4 Experiment and Result Analysis

4.1 Experimental Dataset

In order to analyze the experimental performance of the algorithm, the ORL dataset (http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html) of Cambridge University is selected for experiments in this paper. The dataset contains 40 face images of different ages and genders. Each group represents 10 different pose photos of the same person. These 10 face images are collected from the same person with different facial expressions and from different angles.

4.2 Experimental Results and Analysis

Experimental Program 1:

Experiments are performed on real datasets. Choose one image from each category as the test set, and all the remaining images as the comparison data. In order to show the difference between the original data and the dimensionality reduction data, the prediction results of each category are displayed. The experimental results are shown in Figs. 1, 2.

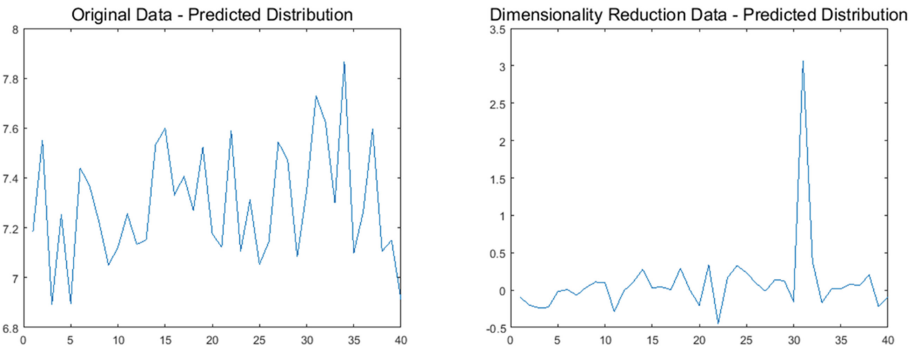


Fig. 1. Experimental comparison of different data on test image 31

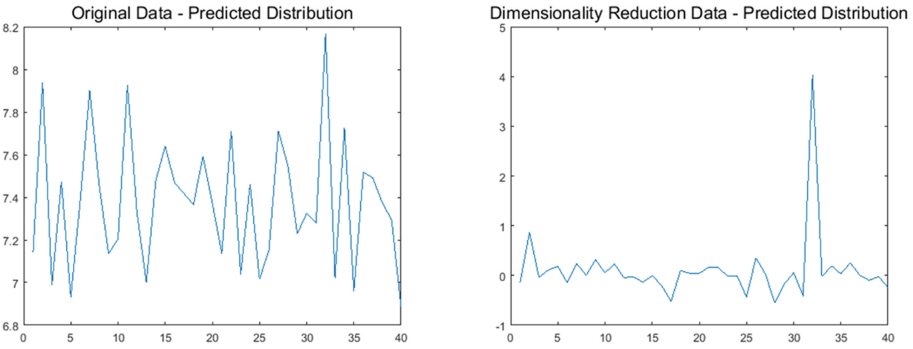


Fig. 2. Experimental comparison of different data on test image 32

Figures 1, 2 show the experimental results of the original data and the dimensionality reduction data on the test images. For intuitive display, the pictures of the same group of prediction categories are displayed, which can more intuitively explain the accuracy of the prediction results. The experimental results in Fig. 1 show that the use of SVD algorithm for data dimensionality reduction can accurately predict the category of the image, which shows that face image recognition based on SVD algorithm can effectively improve the model prediction accuracy. The experimental results in Fig. 2 show that both the original data and the dimensionality reduction data can be used to accurately predict the experimental results, but through the analysis of the predicted numerical distribution [12, 13], it can be seen that the data processed by dimensionality reduction

has better distinguishability and can be accurately segment the predicted category with other categories to better complete data prediction.

Experimental Program 2:

Experiments were compared on the ORL dataset using 10-fold cross-validation. The dataset was randomly divided into 10 parts for experiments. Record the mean and standard deviation of the original data and the dimensionality-reduced data in terms of prediction accuracy, and the results are as follows.

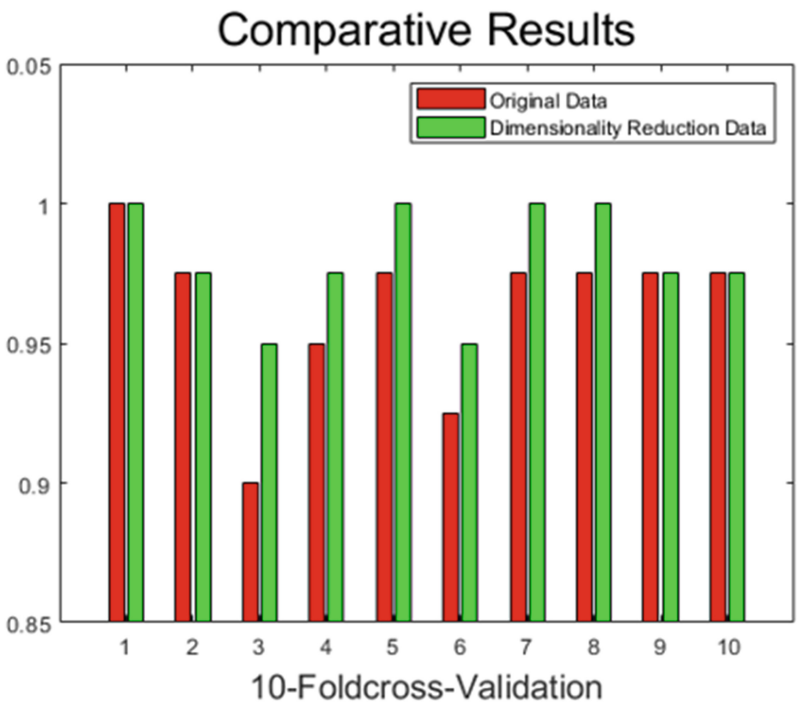


Fig. 3. 10-fold cross-experiment comparison between different data (plot)

Figures 3, 4 show the experimental results of ten-fold crossover between the original data and the reduced-dimensional data. Figure 3 shows the result display of the original data and the dimensionality reduction data by means of ten-fold crossover. The experiment shows that in the ten experiments, the SVD feature dimensionality reduction data is better than the original results 5 times, and the other 4 times the results are the same as the original results. Figure 4 shows the experimental comparison between original data and dimensionality-reduced data in different dimensions. Experiments show that in the data dimension reduction experiment, when the dimension is 14, the prediction accuracy of the original data and the dimension-reduced data is the same. When the dimension is greater than 15, the experimental precision of the matrix factorization dimension reduction is better than the original data, and with the dimension increases, the prediction accuracy approaches 1 and tends to stabilize.

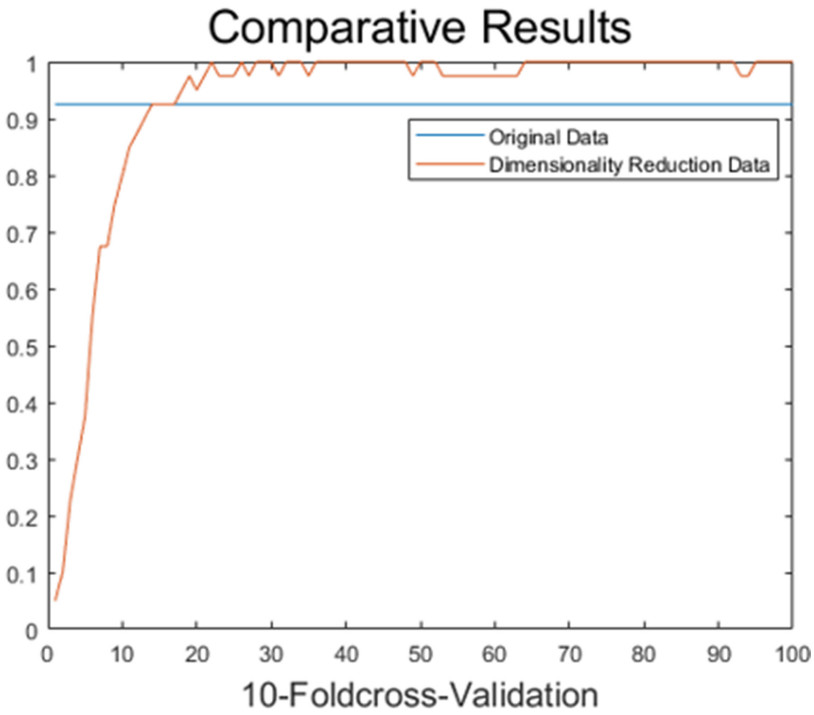


Fig. 4. 10-fold cross-validation experiment comparison with different data dimensions

5 Conclusion

In this paper, a similarity metric recognition of face images based on singular value decomposition is proposed. Attribute reduction is performed on the original data through SVD, and similarity measurement is performed on the processed features. Compared with the original data set, the results show that the algorithm has a good effect, and can accurately predict the label of the image in the real prediction.

This paper focuses on the impact of data dimensionality reduction on the results of image similarity measurement. In the follow-up research, different methods will be considered for dimensionality reduction analysis of image features.

Acknowledgment. This work was supported by Domestic Visiting Program for Outstanding Young Teachers in Colleges and Universities (gxgnfx2021154); Suzhou University Scientific Research Platform Open Project (2020ykf01); Key disciplines of computer science and technology (2019xjzdxk1); Top talent project of colleges and universities in Anhui Province (gxbjZD43); Natural science research project of Anhui Provincial Department of Education (KJ2018A0453); School-Level Scientific Research Platform of Suzhou University (2021xjpt51); Provincial Industrial College (2021cyxy069); Key project of Natural Science Research of Anhui Provincial Department of Education (KJ2021A1110); School level Industrial College of Suzhou University (szxy2021cxxy04); Anhui Province Natural Science Foundation (2022AH051372); Key Project of Natural Science Research in Universities of Anhui Province (2022AH051372).

References

1. Zhi, H., Liu, S.: Face recognition based on genetic algorithm. J. Vis. Commun. Image Represent. **58**, 495–502 (2019)
2. Kute, R., Vyas, V., Anuse, A.: Transfer learning for face recognition using fingerprint biometrics. Journal of King Saud University - Engineering Sciences (2021)
3. Zebari, R., Abdulazeez, A., Zeebaree, D., et al.: A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. J. Appl. Sci. Technol. Trends **1**(2), 56–70 (2020)
4. Espadoto, M., Martins, R.M., Kerren, A., et al.: Toward a quantitative survey of dimension reduction techniques. IEEE Trans. Visual Comput. Graphics **27**(3), 2153–2173 (2019)
5. Akritas, A.G., Malaschonok, G.I.: Applications of singular-value decomposition (SVD). Math. Comput. Simul. **67**(1–2), 15–31 (2004)
6. Henry, E.R., Hofrichter, J.: [8] Singular value decomposition: Application to analysis of experimental data. Methods Enzymol. Academic Press **210**, 129–192 (1992)
7. Chen, R., Pu, D., Tong, Y., et al.: Image-denoising algorithm based on improved K-singular value decomposition and atom optimization. CAAI Trans. Intell. Technol. **7**(1), 117–127 (2022)
8. Gu, F., Lu, J., Xia, G., et al.: Face verification technology based on FaceNet similarity recognition network. In: 2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS), pp. 1362–1367. IEEE (2021)
9. Li, J., Qiu, T., Wen, C., et al.: Robust face recognition using the deep C2D-CNN model based on decision-level fusion. Sensors **18**(7), 2080 (2018)
10. Kortli, Y., Jridi, M., Al Falou, A., et al.: Face recognition systems: a survey. Sensors **20**(2), 342 (2020)
11. Saha, S., Ghosh, M., Ghosh, S., et al.: Feature selection for facial emotion recognition using cosine similarity-based harmony search algorithm. Appl. Sci. **10**(8), 2816 (2020)
12. Wang, M., Feng, T., Shan, Z., et al.: Attribute and label distribution driven multi-label active learning. Appl. Intell. **52**, 1–16 (2022)
13. Wang, Y., Tian, W., Cheng, Y.: Heterogeneous ensemble learning algorithm based on label distribution learning **32**(10), 945–954 (2019)

Research on Computation Offloading Schemes for Vehicular Edge Computing

Wei Zhou(✉)

Information Engineering Department, Suzhou University, Suzhou 234000, China
weizhou@ahszu.edu.cn

Abstract. The development of mobile edge computing technology has provided support for the Internet of Vehicles. Due to the limited computing power of vehicles, the Internet of Vehicles service has higher processing priority and lower latency requirements than the traditional cloud computing service. The computation tasks of the on-board terminal can be offloaded to the edge server for processing. In this paper, we studied the in-vehicle network architecture based on DNS, we integrate the computation offloading problems and mobile edge computing problems to build the vehicle-edge-cloud architecture, and proposes a dynamic computation offloading scheme based on DNS in the Internet of Vehicles. Aiming at the clustering problems and the calculation and offloading problems of the tasks of Internet of Vehicles, offloading tasks is clustered into different clusters, and the tasks in the cluster have similar computing requirements and delay requirements, and the task data parameters are normalized, and the K-means algorithm is improved according to the characteristics of the task data in the Internet of Vehicles. Consider summarizing the status of the edge server through centralized SDN, offloading the computing tasks to the relatively idle MEC servers, and balancing the load between the edge servers, so as to solve the problems of high energy consumption and high latency in the Internet of Vehicles. Simulation results show that our method can reduce the average latency and reduce the energy consumption of on-board equipment in the edge computing network, which meets the needs of real-time and low energy consumption of the Internet of Vehicles.

Keywords: Mobile edge computing · internet of vehicles · computation offloading

1 Introduction

With the development of smart devices and mobile communication technologies, mobile devices such as mobile phones, wearable devices, augmented reality (AR) [1], etc. generate a large number of data and computing requirements, which increase network latency and traffic burden, users have higher and higher requirements for network performance such as network service quality and request delay, and edge devices often face limitations such as computing resources, storage resources, and power resources. We need to offload compute tasks to a relatively resource-rich edge computing server. Although the traditional cloud computing has strong computing power, but the central server and

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 202–211, 2023.
https://doi.org/10.1007/978-981-99-2653-4_25

terminal equipment are too far away to respond to the needs of real-time users, a large number of terminal nodes will transmit data to the cloud, which will cause network congestion. In order to respond to user needs, the energy consumption of cloud computing center is high, the energy consumption of terminal nodes in the process of data transmission to cloud servers is also high. Mobile edge computing (MEC) [2] is deployed in an edge location closer to the terminal, which can offload tasks to a nearby edge server without communicating with a remote server, reducing communication load and response latency, reducing the burden and energy consumption of the central server, and improving the operational efficiency and user experience on user devices [3].

In-vehicle edge computing is an important application scenario of mobile edge computing. The in-vehicle network realizes instant communication between vehicles and vehicles (V2V), vehicles and everything (V2X) through wireless channels, and integrates various sensors, communication, image recognition and other data collected from vehicles for autonomous driving, road traffic control, in-vehicle navigation and augmented reality [4]. However, due to the limited computing resources, storage and communication resources of vehicle nodes, data-intensive applications such as in-vehicle entertainment and augmented reality need to process a large amount of data, the in-vehicle unit does not have the processing capability of a large amount of data, so a better practice is to offload the in-vehicle tasks to the remote cloud server for processing, due to the latency-sensitive nature of the Internet of Vehicles, this kind of computation offloading strategies require a large amount of bandwidth resources and cannot meet the requirements of the high real-time nature of the Internet of Vehicles. Therefore, some of the vehicle's compute-intensive tasks are offloaded to edge nodes for execution to reduce latency while enabling efficient service delivery and user experience.

Computation offloading technology is an important part of the research in the Internet of Vehicles. Compute offloading uploads the computing requirements of resource-constrained endpoints to the edge MEC server, and transmits the computing results back to the terminal with the powerful computing power, bandwidth, and storage of the edge MEC server. [5] studies the task scheduling problem of VEC in compute-intensive scenarios, unloads tasks to different edge RSU servers according to dependency relationships, converts them into NP-difficult optimization problems, and designs a multi-task prioritization scheduling strategy to reduce completion time of in-vehicle tasks [6–8]. Studied the data forwarding, road channel interference and relay routing problems, designed efficient energy-efficient multicast routing protocol and equivalent bandwidth deployment, then designed a 5G cellular cluster model, taking time and energy consumption as the optimization index. [9, 10] considers optimization strategy of computation offloading problems, combines the computing and storage resources of MEC servers and centralized cloud computing centers for computation offloading, and adopts FiWi network architecture to provide support for edge-cloud communication [11, 12]. Takes the mobility of the vehicle as the influencing factor of task unloading, considers the approximate optimal unloading scheme of the occupied vehicle, transforms the global optimization problem into a local optimization problem, and solves it by numerical method [12].

This paper studies the computation offloading problems in the single-lane mode, clusters the uninstall tasks into different clusters, and the tasks in the cluster have similar computing requirements and latency requirements, and the in-vehicle tasks can be transmitted to the idle MEC to balance the network load by summarizing the edge server status through centralized SDN.

2 Network Architecture

This paper proposes an in-vehicle edge computing network architecture based on SDN. As shown in Fig. 1, this network architecture can be divided into four parts: the user layer which combines 5G and wireless access networks, the edge layer consisting of RSU and VEC, SDN which provide centralized management and cloud layer. Using vertical end-edge-cloud collaboration, horizontal user layer collaboration and edge layer collaboration, SDN center can balance the uneven traffic in the network.The user layer collects vehicle information through sensors, cameras, radars, global positioning systems, etc., and uses 5G access to use V2V and V2R communication methods to interact with the terminal node or RSU. VEC servers in edge layer use 5G mobile communication equipment, WIFI and short-distance wireless communication, combined with MEC and network slicing technology for data transmission and computation offloading services.The control layer include SDN controller, enables automation and network control, provides storage and computing facilities for the vehicle through SES. SDN-based Road Side Unit (SRSU) communicates with vehicles in range under the management of SDN controller. The cloud layer include OpenFlow management software and resource

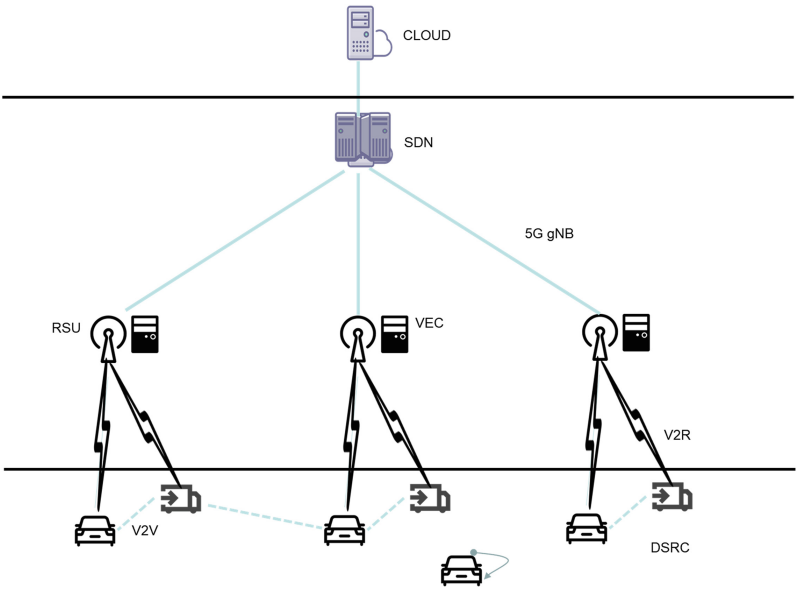


Fig. 1. The on-board edge network architecture based on SDN controller

management software, realizes advanced driving, collaborative sensing, vehicle formation and CCAM applications in the cloud layer. By caching content of edge users, the compute capabilities of the center can be downgraded to the edge layer on the user side.

The SDN part of the in-vehicle edge computing network architecture provides centralized intelligent control management for the entire network architecture, and realizes the decoupling of data forwarding and control by sensing the overall network state. Bandwidth, recovery, security, and policy management in the Internet of Vehicles can be automated and highly intelligent, and the SDN architecture can support collaborative computing, operations support/business support systems, and business-critical network applications. SDN implementation functions mainly include: (i) using OpenDayLight controller, monitoring the network topology discovery of physical devices such as switches, routers, hosts, etc., routing policies and other network status information, adjusting the quasi-sending policy according to the data information received by the flow table policy docking, and integrating network resources. (ii) Use FlowVisor technology to virtualize SDN hardware to achieve network resource sharing. (iii) Network load balancing, using the Markov process to evaluate the communication link weights, select the control switching equipment with low load for traffic migration, and realize traffic monitoring and network load balancing.

3 System Model

3.1 Computation Model

The real-life vehicle networking environment is extremely complex, this article considers the one-dimensional one-way road, RSU equidistant distribution, each RSU is connected to the VEC server through an optical cable, and the transmission time between the RSU and VEC is negligible. V = {1, 2,…, N} represents a collection of vehicles, we represent R = {1,2,…,M} as a set of on-board tasks waiting to be completed, assuming that vehicles have a non-interfering, mutually independent calculation task that needs to be performed within a certain period of time. Because the application scenarios of the Internet of Vehicles are not the same, the types of tasks generated are also different, the tasks are fine grained divided, and each calculation task is represented by a triplet $T = \{ l_i, c_i, t_i^{max} \}$, l_i indicates the amount of data for the task i to be calculated, the data size of the task affects the transmission time and execution time, c_i indicates the CPU cycle required to complete the task, the task calculation time, and the upper limit of the tolerance delay of the task.

The data transmission rate of vehicle v is shown in Eq. 1, R_v represents the transmission power of the vehicle user v, g_v represents the channel gain of the vehicle communicating with the roadside unit, I represents the noise power of the channel, and W represents the bandwidth of the channel.

$$R_v = W \log_2\left(1 + \frac{p_v g_v}{I}\right) \tag{1}$$

When the task offloading strategy decides to put the calculation on the vehicle side, there is no need to calculate the task transmission time, and the total delay is the vehicle calculation time. The tasks which can't be offloaded are performed locally in the vehicle,

and the part of the task that can be offloaded will be performed either by the on-board processor and transmitting to the VEC server on the RSU side. The curbside unit detects the task status of each vehicle, minimizes the overall service delay according to the state of the vehicle and the task, and carries out the planning of the offloading strategy and the sequential scheduling of the task scheduling. When vehicle i decides to perform its computing task locally, the processing delay for that task depends only on the vehicle's computing power.The latency and energy consumption of local execution can be calculated, f_i^{local} is the CPU cycle frequency of vehicle i. l_i^{local} is the size of the amount of data that task i performs locally.

$$\begin{aligned} T^{local} &= \sum_{i=1}^{n} \frac{l_i^{local}}{f_i^{local}} \\ E^{local} &= Pv * \sum_{i=1}^{n} \frac{l_i^{local}}{f_i^{local}} \end{aligned} \tag{2}$$

When a vehicle has too many tasks to deal with, it can offload its computing task to the Edge Server j on the RSU, the total compute task processing delay consists of the waiting time for the task to execute and the data transfer time, the task can be offloaded to the RSU at the edge or to the vehicle with signal coverage, and the delay of the total task can be calculated by a formula, where f_j^{mec} is the CPU cycle frequency of the j-th edge server on the RSU. Vehicle i can choose V2R communication or V2V communication, if vehicle i decides to communicate through V2R, its computing task is offloaded to the edge server on RSU j, and the transmission delay of the calculation task r is the ratio of the data size to the data transmission rate, and the total transmission and energy consumption of the on-board task offloaded to the edge server can be calculated.

$$\begin{aligned} E^{mec} &= Pt * \sum_{j=1}^{n} \frac{l_j^{mec}}{f_j^{mec}} + Pw * \sum_{j=1}^{n} \frac{l_j^{mec}}{W \log_2\left(1 + \frac{p_v g_v}{I}\right)} \\ T^{mec} &= \sum_{j=1}^{n} \frac{l_j^{mec}}{f_j^{mec}} + \sum_{j=1}^{n} \frac{l_j^{mec}}{W \log_2\left(1 + \frac{p_v g_v}{I}\right)} \end{aligned} \tag{3}$$

3.2 Task Clustering Problem

The application scenarios of the Internet of Vehicles are different, the types of tasks generated are also different, and the tolerance of delays and the demand for resources of different types of tasks are different. The task is divided into fine-grained, each computing task is represented by a triplet T = { l_i,c_i,t_i^{max} } , according to the data size of the task, the need for CPU resources and latency requirements can be divided into compute-intensive type, bandwidth-demanding type and latency type. Cluster different tasks to reduce the characteristic dimension of the task. Each cluster after clustering is a collection of tasks. In the same task cluster, the task attribute characteristics waiting for scheduling are more similar, but the task attribute characteristics waiting for scheduling are less similar to those in other clusters. Firstly normalize the CPU resources and delay required by the

task, and the euclidean distance calculation method is used to obtain the normalized task sample distance.

$$\begin{cases} l_{\mathrm{i}}^{d} = \dfrac{l_{\mathrm{i}}}{\sum\limits_{i=1}^{n} l_{\mathrm{i}}} \\ c_{\mathrm{i}}^{d} = \dfrac{c_{\mathrm{i}}}{\sum\limits_{i=1}^{n} c_{\mathrm{i}}} \\ t_{\mathrm{i}}^{\max d} = \dfrac{t_{\mathrm{i}}^{\max}}{\sum\limits_{i=1}^{n} t_{\mathrm{i}}^{\max}} \end{cases} \tag{4}$$

We use a K-Means-based clustering algorithm Cluster different types of vehicle tasks. Tasks within the cluster have similar characteristics of latency, calculation and data volume, and can be selected to offload similar tasks according to the characteristics of the task, such as offloading tasks with large data volume and large computation volume to the VEC server can improve efficiency.

First we randomly select K different types of tasks as the initial clustering points, and then calculate the Euclidean distance between each task and the cluster center, when the distance is less than the specified threshold, the task is absorbed into the cluster, and the task is marked as a clustered task, the mean of all points is calculated for the constituent clusters, and then the task furthest from the initial cluster point is selected as the next unmarked task as the new cluster center, repeat the above steps until the task is fully completed.

$$D(t_i, t_j) = \sqrt{(l_{\mathrm{i}}^{d} - l_j^{d})^2 + (c_{\mathrm{i}}^{d} - c_j^{d})^2 + (t_{\mathrm{i}}^{\max d} - t_j^{\max d})^2} \tag{5}$$

The K-means clustering algorithm is used to evaluate the similarity by clustering on the basis of minimizing error, t_i is the sample data. Based on the error score and the continuous attribute as a cluster objective function to characterize the sample and cluster center density, when the SSE is smaller, the higher the sample similarity in the cluster.

$$SSE = \sum_{i=1}^{k} \sum_{i=1}^{n} D(t_i, \frac{1}{n_i} \sum_{t_i=1}^{m} t_i) \tag{6}$$

3.3 Offloading Strategies Considering Load Balance

Different MEC servers have different loads, and when a MEC server with insufficient computing resources undertakes a large number of computing tasks, it is not only impossible to ensure the completion of the computing tasks, but also causes the MEC server to be overloaded and affects the processing of other services. Edge Server performance degrades as the number of computation offloading increases, and in dense network areas, vehicles offload a large number of tasks to adjacent RSU at the same time, which leads to

increased task processing latency due to the decrease in computing resources available on edge servers. For areas where road traffic is idle, fewer vehicles will offload tasks to the edge VEC servers, and the calculation offload algorithm does not take into account the load of the servers, and some edge servers will be overloaded with computing and some servers idle.

This paper proposes an offloading algorithm that combines load balancing with Edge Server. Assume that all Edge Servers are optional servers for vehicles to offload their computing tasks, and then consider the balance of execute latency and transmission latency to determine which Edge Server to choose to offload tasks to execute. Tasks can be decided to execute locally in edge vehicles, transmit to the edge VEC server of the nearest RSU or transmit to the edge VEC servers of the remote RSU.

The task is calculated first under all possible offload decisions, calculate the locally processing delay of task t_i^{local}, and Offload latency to the edge cloud t_i^{mec}. The task is removed from the unload collection if the task latency is greater than the maximum tolerated delay. Based on the current network channel state calculation, the compute task with the minimum value density is found, the MEC server queue is removed, and the computing resources consumed are recalculated until all compute tasks consume no more than the total computing resources of the MEC server. If the on-premises processing latency is less than the offload to the edge cloud, the task is removed from the offload collection. Tasks that cannot be executed locally on the vehicle CPU are offloaded to the edge VEC server by comparing the execute latency caused by different offload decisions. Determine offload decisions for the remaining tasks by updating the network state.

4 Simulation Results

The simulation scenario in this paper is a distributed network architecture for the deployment of VEC cloud servers for ultra-dense mobile edge computing in urban areas. In our simulation, there were 20 moving vehicles, with RSU and MEC servers deployed along the road section. Set the channel bandwidth to W = 25 MHz, the noise power of the channel to 10–10 W, and the computing power of the VEC server to 5–20 GHz. The computing power of each vehicle-mounted user CPU is randomly distributed at 0.1–5 GHz, the computation offloading strategy is controlled by the SDN controller, and the random parameters are independent for different vehicles (Table 1).

Table 1. Network load node distribution

Symbol	Meaning	Value
N	Number of vehicles	20
M	Number of tasks	20–300
l_i	Task size	0.1–5 MB
c_i	The number of task CPU cycles	0.1–2 GHZ

(*continued*)

Table 1. (*continued*)

Symbol	Meaning	Value
W	Wireless bandwidth	25 MHz
f_j^{mec}	CPU cycle frequency of server	5–20 GHZ
f_i^{local}	CPU cycle frequency of local	0.1–5 GHZ
$t_i^{\max}$	Maximum tolerance delay	0.1–3 s
I	Noise	−100 dBm

Compare the solution of the computation offloading strategy with other offload scenarios to test the performance of the computation offloading algorithm. Compare the performance of tasks performed locally, fully offloaded to the server, and offloaded strategies using improved algorithms, in local calculations each vehicle user's calculation task does not consider the offloading situation, the task is calculated on the local processor, which is represented by LOCAL. In complete offloading the computing tasks that each vehicle user needs to process are not performed locally, the task is transmitted to the VEC server, which is represented by VEC. Compare the performance of the three offloading strategies, Fig. 2 shows the average delay of vehicles. In different offloading schemes, the system delays with the increase of the number of tasks, and it can be seen that these methods increase with the increase of the number of vehicle users, and the amount of task data that needs to be performed increases, and the delay increases. Comparing several schemes, it can be seen from the simulation results that the improved BLCVEC algorithm achieves the smallest overall response delay and the best performance. The local computing strategy has the largest latency, which is because all vehicle computing tasks are completed on their own processing equipment, each vehicle's computing resources are limited, and the speed of processing data is slower, so the time consumption of computing tasks is much higher than that of other algorithms. The overall delay of offloading to VEC is higher than that of the BLCVEC algorithm, which is due to the BLCVEC algorithm's decision to calculate or offload each user's task locally, and also clustered according to the characteristics of the task, so that the overall time consumption is minimal (Fig. 3).

Analyzing the energy consumption of vehicles under different number of on-board users, it can be seen that the trend of the three schemes is to increase the energy consumed with the increase of the amount of tasks. This paper considers the calculation tasks with four parameters, resource-intensive computation tasks with large amount of data and computation, latency-sensitive tasks with a small maximum tolerance time, compute-intensive computation tasks with large amount of computation, and data-intensive computation tasks with a lot of data. Among them, the local computing scheme consumes the most energy, because the computation of the task is intensive, and the computing power of the on-board CPU is low, which will not only cost a large amount of energy, but also has a long delay in the system. The VEC offloading scheme consumes the lowest energy of vehicles, at which time the vehicle user only needs to bear the energy of data transmission and communication, and the calculation consumption is borne by the

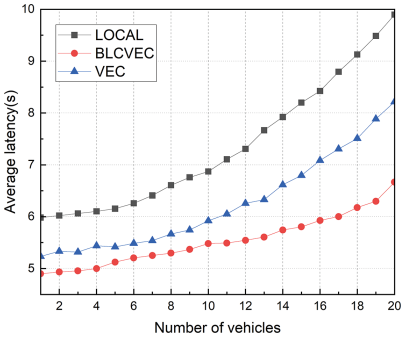


Fig. 2. Average delay

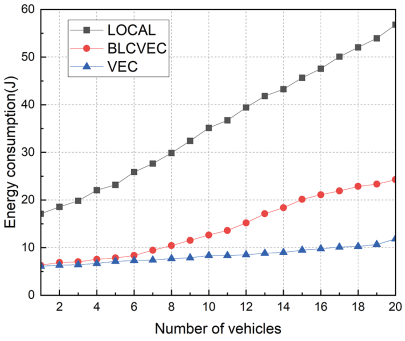


Fig. 3. Energy consumption

VEC server. The energy consumption of the BLCVEC scheme is similar to the energy consumption of the full unloading scheme when the number of tasks is small, because the tasks will be offloaded fully when the amount of computation tasks is small. However, as the amount of computing tasks increases, the BLCVEC scheme consumes more energy than the full offloading method, and some tasks will be executed locally, which is the lowest energy consumption cost to make the overall delay, and by consuming some energy, the response delay is greatly reduced. For automotive users, the advantages of the BLCVEC solution are more prominent, and they can better meet the latency sensitivity and safety needs of automotive applications.

5 Conclusion

This paper studies the in-vehicle network architecture, introduces the computation offloading problem of software-defined network and mobile edge computing integration to build the edge cloud architecture, clusters the offloading tasks into different clusters, and the tasks in the cluster have similar computing requirements and latency requirements, and considers summarizing the edge server status through centralized SDN, and offloading the computing tasks to the more idle MEC servers to balance the network load. We expected to solve the problems of high energy consumption and high latency in the Internet of Vehicles, and the algorithm has been improved according to the relevant problem research. The environment of the simulation experiment is configured, and the performance parameters required for the simulation experiment are determined according to the literature related to the Internet of Vehicles and computation offloading, and the experimental simulation results further prove the effectiveness of computation offloading strategy we proposed.

Acknowledgments. This work was supported by Key disciplines of computer science and technology (2019xjzdxk1), Teaching Research Project (2018jyxm0960), College of modern industry (szxy2021cyxy04), School level Industrial College of Suzhou University (szxy2021cxxy04), top talent project of Anhui Universities (gxbjzd43), natural science research project of Anhui Provincial Department of Education (kj2018a0453), Innovation and entrepreneurship project for college

students (S202110379167), natural science research key project of Anhui Provincial Department of Education (kj2021a1110), School-Level Scientific Research Platform of Suzhou University (2021xjpt51), Provincial Industrial College (2021cyxy069), Teaching and research project of education (2020jyxm2220), Research Project (2022xhx039), Collaborative education project (202102326030), Teaching and research project of education (szxy2019jyxm12), Outstanding talent training innovation project (2020zyrc163).

References

1. Shi, W., Cao, J., Zhang, Q., Li, Y., Lanyu, X.: Edge computing: vision and challenges. IEEE Internet Things J. **5**, 637–646 (2016)
2. Burcu, C.G., Begen Mehmet, A.: Exact optimization and decomposition approaches for shelf space allocation. Eur. J. Oper. Res. (2), 432–447 (2022)
3. Deng, X., Li, J., Liu, E., et al.: Task allocation algorithm and optimization model on edge collaboration. J. Syst. Architect. **9**, 101778 (2020)
4. Liu, L., Chen, C., Pei, Q., et al.: Vehicular edge computing and networking: a survey. Mob. Networks 1145–1168 (2020)
5. Liu, Y., et al.: Dependency-aware task scheduling in vehicular edge computing. IEEE Internet of Things J. 4961–4971 (2020)
6. Eshteiwi, K., Kaddoum, G., Selim, B., Gagnon, F.: Impact of co-channel interference and vehicles as obstacles on Full-Duplex V2V cooperative wireless network. IEEE Trans. Veh. Technol. **99**, 7503–7517 (2020)
7. Li, B., Hou, P., Wang, K., Peng, Z., Jin, S., Niu, L.: Deployment of edge servers in 5G cellular networks. Trans. Emerg. Telecommun. Technol. 3841–3849 (2020)
8. Kadhim, A.J., Seno, S.A.H.: Energy-efficient multicast routing protocol based on SDN and Fog computing for vehicular networks. Ad Hoc Networks (5), 68–81 (2019)
9. Guo, H., Liu, J.: Collaborative computation offloading for multiaccess edge computing over fiber-wireless networks. IEEE Trans. Veh. Technol. **5**, 4514–4526 (2018)
10. Wang, J., Feng, D., Zhang, S., Tang, J., Quek, T.Q.S.: Computation offloading for mobile edge computing enabled vehicular networks. IEEE Access (7), 62624–62632 (2019)
11. Zhan, W., Luo, C., Min, G., Wang, W., Zhu, Q., Duan, H.: Mobility-aware multi-user offloading optimization for mobile edge computing. IEEE Trans. Veh. Technol. **3**, 3341–3356 (2020)
12. Wu, Q., Ge, H., Liu, H., Fan, Q., Li, Z., Wang, Z.: A task offloading scheme in vehicular fog and cloud computing system. IEEE Access **8**, 1173–1184 (2020)

Research and Design of CO_2 Concentration Monitoring System Based on Sensor Information Acquisition Technology

Bingqian Ni and Shiyuan Zhang(✉)

School of Information Engineering, Suzhou University, Suzhou 234000, Anhui, China
shiyuanzhang007@126.com

Abstract. In order to solve the problem of detecting CO_2 concentration in the air, an information acquisition and concentration monitoring system based on CO_2 concentration sensor is designed. Firstly, the overall framework of CO_2 concentration monitoring system is designed, including sensor module design, signal conditioning circuit module design, main control chip circuit design and lithium battery voltage rise and fall circuit design. Then, the hardware circuit connection design is carried out, including the overall connection circuit design of the main control chip stm32f103vet, the sensor circuit connection design, the LCD drive circuit design, the USB communication and upper computer data storage circuit design and the power supply circuit design. Finally, the whole software flow of CO_2 concentration monitoring system is designed. The system realizes the real-time detection and storage of CO_2, and can realize the information exchange and information processing of the upper computer through the data transmission interface. The system plays a very important role in improving production efficiency and energy saving and emission reduction.

Keywords: CO_2 concentration · Information collection · Concentration monitoring · Sensor

1 Overall Design of CO_2 Concentration Monitoring System

With the development of human economy and social progress, more and more industrial fields use coal and other resources for a long time, resulting in higher and higher CO_2 content in the air. The "greenhouse effect" slowly caused by CO_2 is destroying human living environment step by step. Therefore, it is more and more important to monitor and control the concentration of CO_2 in the air regularly and accurately [1]. In the medical field, the concentration and content of CO_2 pose a very serious threat to human respiratory diseases. In the field of agroecology, the concentration of CO_2 also has an important impact on the growth of crops. Therefore, the rapid realization of CO_2 concentration detection technology has become increasingly urgent. It plays an important role in improving people's health, economic level and social development [2]. In order to realize the real-time detection and real-time control of carbon dioxide content in the

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 212–218, 2023.
https://doi.org/10.1007/978-981-99-2653-4_26

air, a CO_2 concentration data acquisition system based on embedded chip stm32f103vet is designed in this paper.

The CO_2 concentration monitoring system uses stm32f103vet chip as the main controller, and its work flow is divided into three steps [3]: (1) first, the information collected by the CO_2 concentration detection sensor is amplified to output analog signals; (2) Then, the concentration of carbon dioxide in the air is calculated by the digital to analog converter of the signal conditioning circuit; (3) After being processed by mb1322 chip, the CO_2 concentration value is transmitted to LCD screen for digital display. The system provides USB interface, which can be used to export and read system data, and realize information docking and information interaction between CO_2 acquisition system and host computer. Among them, the power supply system of the system uses conventional lithium battery for power supply to ensure the normal operation of the system.

2 System Hardware Design

2.1 Stm32f103vet Minimum System

Stm32f103vet main control chip has a large number of registers and i/o interfaces, which can provide fast information processing capacity and fast system response speed for CO_2 concentration acquisition system, and support system C language programming and system debugging [4]. At the same time, the chip has a small volume, which greatly reduces the use of system space in circuit connection design. The system mainly includes high-speed program uploading and downloading, circuit debugging, system power on, system restart, high-speed and low-speed chip oscillation circuit and power interface [5]. The circuit design connection of stm32f103vet main control chip is shown in Fig. 1.

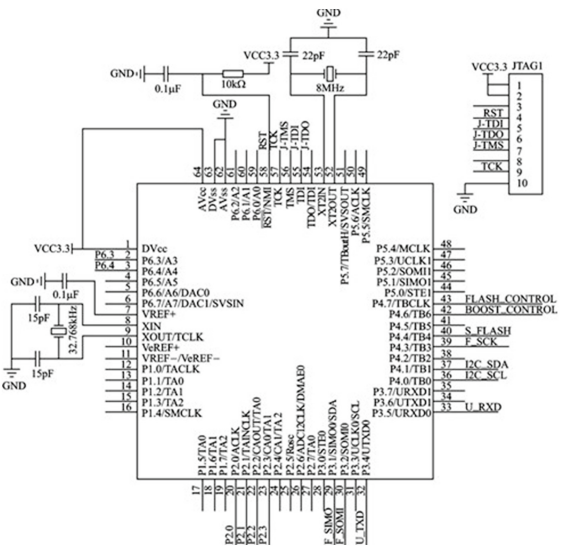


Fig. 1. Circuit connection diagram design of stm32f103vet

2.2 CO_2 Sensor Circuit Design

The CO_2 concentration sensor adopts the lp908 sensor produced by Nanjing Jingwei Company. The sensor completes the functions of voltage signal acquisition, output and data processing input through the sensor output and input pins p6.3 and p6.4 of the main chip [6]. In a specific temperature and humidity environment, the voltage value received by the lp908 is inversely proportional to the CO_2 concentration in the air. The CO_2 sensor voltage output value obtained after digital to analog conversion can be converted into the CO_2 concentration value through the inversely proportional relationship. The circuit design of CO_2 sensor is shown in Fig. 2.

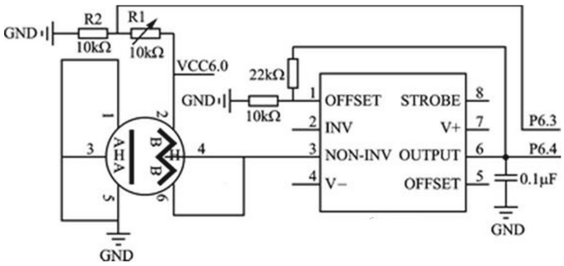


Fig. 2. Circuit design of CO_2 sensor

2.3 LCD Driving Circuit Design

In order to achieve the accuracy and stability of CO_2 concentration value display, mb1322 multi-function LCD module and digital display system are used in the screen display driver circuit. The system can realize command modes such as data reading, storage, reading and writing modification, and can directly display digital information on the LCD screen [7]. The serial ports such as read, write, data and Cs on mb1322 can be

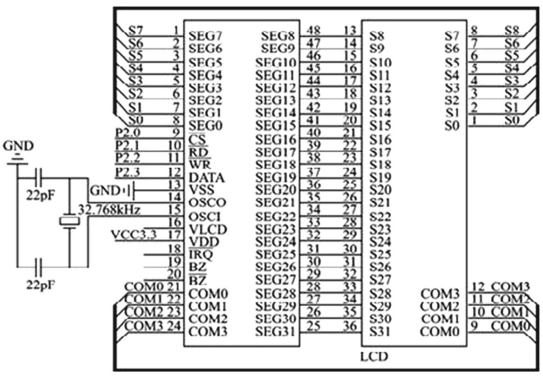


Fig. 3. ht1621 driving segment LCD circuit

connected with P2 on stm32f103vet X and other interfaces directly connect data to complete the functions of system initialization, data reading and writing, data storage, data processing and output. The design of LCD driver circuit is shown in Fig. 3.

2.4 USB Communication and Data Storage Chip Circuit Connection Design

In terms of data storage of the concentration monitoring system, a 16 MB k36p90 memory chip is used, which can read and write data according to the number of pages, such as dynamic, static, slice and overall modification. K36p90 chip uses a special KPI controller to directly control and automatically generate the system automatic control timing. When k36p90 chip completes data storage, lp1108 serial port chip is used to realize the interactive function between data and upper computer. The upper computer interacts with k36p90 by using USB interface and UART to complete data transmission. The request response class of the controller can complete the control function of UART. The circuit connection design of USB communication and data storage chip is shown in Fig. 4.

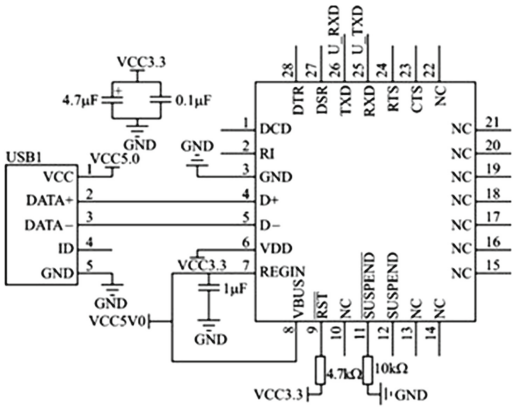


Fig. 4. Circuit connection design of USB communication and data storage chip

2.5 Circuit Design of Power Supply System

The power supply voltage of the carbon dioxide concentration monitoring system mainly includes the 3.6 V voltage circuit of stm32f103vet, mb1322 and lp908 chips and the 6 V supply voltage of the sensor. However, the normal voltage provided by the lithium battery used in the system is the 4 V voltage circuit. In order to ensure the normal operation of the system chip, it is necessary to design the boost circuit design and step-down circuit design based on the lithium battery voltage. The boost chip adopts dsp1708 chip of 3Com company, and the step-down chip adopts nax16993 chip of delivery company. The design of the lithium battery voltage rise and fall circuit is shown in Fig. 5, in which the left half is the design connection of the step-down circuit and the other half is the design connection of the step-up circuit.

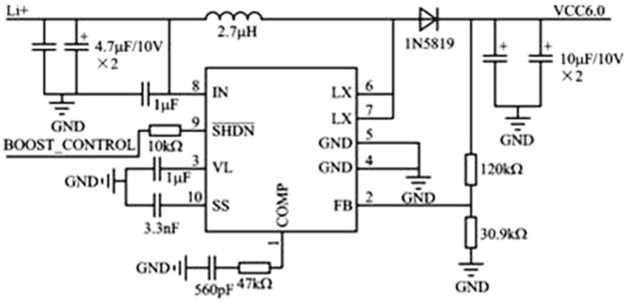


Fig. 5. Connection design diagram of step-down and step-up circuit

3 Software Design of CO_2 Concentration Monitoring System

3.1 System Program Design

The CO_2 concentration monitoring system is flexible and modular program design is adopted for later upgrading and maintenance. The software design driver module of the system corresponds to the hardware circuit driver module one by one. The system mainly includes the main program design, sensor information acquisition program, information processing, LCD digital display, system interrupt and restart, USB communication and data host computer storage interaction, etc.

4 CO_2 Concentration Monitoring System Test

4.1 System Accuracy Test

8 groups of samples were taken for system test, and the test was repeated 10 times, totaling 80 groups of data. The calculation shows that the detection accuracy range of the system is –1.31% –0.87%, which meets the national gas detection error standard (–2.48% –4.10%), and the relative deviation range is 2.172% –4.184%, indicating that the system has good repeatability and practicability.

4.2 System Stability

Use the CO_2 concentration detection system to flush out 8 groups of gas samples for 10 times, and draw the data in Table 1. into the concentration curve. See Fig. 6 for the results. It can be seen from the figure that the concentration detection system has strong system stability within a certain error range.

Table 1. Performance indexes of CO_2 concentration monitoring system

No	Concentration reference	Average value	Absolute error	Relative error	Standard deviation	Relative deviation
1	291	280.1	– 0.18	– 0.32	12.7	2.172
2	352	340.5	0.76	0.65	14.6	2.318
3	467	481.8	0.32	0.53	30.9	4.121
4	589	579.2	0.87	0.72	36.5	2.863
5	765	783.1	0.43	0.58	42.2	3.162
6	997	985.6	– 0.87	– 0.72	50.3	3.926
7	1129	1101.7	– 1.31	– 1.22	68.2	3.426
8	1568	1509.8	– 1.29	– 1.37	80.4	4.184

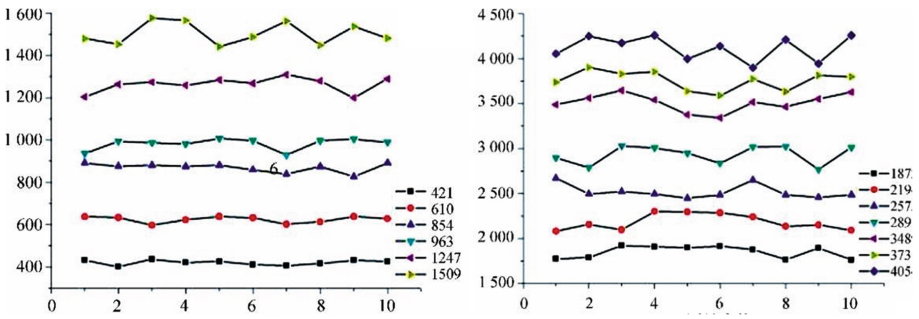


Fig. 6. CO_2 test data curve

5 Conclusion

This paper designs a set of CO_2 concentration information acquisition and concentration monitoring system based on the master chip stm32f103vet, which realizes the real-time monitoring and digital display of CO_2 concentration in the air. The overall design of the system mainly includes sensor module, signal conditioning circuit module, main control chip circuit module and lithium battery voltage rise and fall circuit module. In terms of hardware circuit design, it mainly includes the overall connection circuit of the main control chip stm32f103vet, sensor circuit connection, LCD drive, USB communication and host computer data storage, lithium battery boost and step-down and other functional design. In terms of software flow design, it mainly realizes the software flow of the modular CO_2 concentration monitoring system, ensures the normal operation and coordinated operation of the hardware system, and realizes the real-time detection, real-time storage, information exchange and information processing of CO_2. The CO_2 concentration information collection and concentration monitoring system provides a guarantee for improving energy conservation and emission reduction, economic benefits and living standards.

Acknowledgment. Key disciplines of computer science and technology (2019xjzdxk1); Top talent project of colleges and universities in Anhui Province (gxbjZD43) Natural science research project of Anhui Provincial Department of Education (KJ2018A0453) School-Level Scientific Research Platform of Suzhou University (2021xjpt51) Provincial Industrial College (2021cyxy069) Key project of Natural Science Research of Anhui Provincial Department of Education (KJ2021A1110) School level Industrial College of Suzhou University (szxy2021cxxy04).

References

1. Romain, E.L., Giuliani, D.: Deep-neural network approaches for speech recognition with heterogeneous groups of speakers including children. Nat. Lang. Eng. **23**(3), 325–350 (2016)
2. Chen, J., Liu, Z.: Research on cold chain logistics path optimization based on ant colony algorithm. Transport. Technol. Econ. **23**(5), 38–44 (2021)
3. Jin, J.T., He, W.J., Xu, C.Y., Zhang, J.Y., Dai, D.: Research on logistics distribution path optimization based on improved ant colony algorithm. Mall Modernization **17**(10), 46–48 (2021)
4. Zhu, Z.G., Guang, X.P., Guo, M.: Multi vehicle logistics distribution path optimization under uncertain environment. Transport. Technol. Econ. **23**(2), 6–12 (2021)
5. Emerson, A., Jinn, T., Daniels, B.C.: Locating decisionmaking circuits in a heterogeneous neural network. Front. Appl. Math. Statist. **4**(11), 11–13 (2018)
6. Su, J.H.: Simulation research on optimal selection of automatic logistics distribution route. Comput. Simul. **36**(8), 408–411 (2019)
7. Yang, J.Q., Cao, L.: Urban cold chain logistics path optimization with multiple logistics centers. Logistics Technol. **38**(7), 71–75 (2019)
8. Zou, J.H.: Discussion on logistics path optimization method based on hybrid genetic algorithm. Comput. Products Circul. **17**(12), 128–130 (2018)

Producer Mobility Support Based on Dynamic Forwarding Path Management in ICN

Yanpeng Ji[1(✉)], Jianli Zhao[1], Ze Chen[1], Xiaodong Du[1], and Yachun Mao[2]

[1] Electric Power Research Institute, Hebei Electric Power Corporation, Shijiazhuang 050021, Hebei, China
ymao@tju.edu.cn

[2] School of Microelectronics, Tianjin University, Tianjin 300072, China

Abstract. Information-centric networking (ICN), as a promising future Internet architecture, its user mobility management is one of the important research contents. ICN brings native support for consumer mobility due to its receiver-driven content retrieval model. However, producer mobility is still a challenging issue. Once the producer moves, the routing table of some ICN nodes will be invalid. As a result, the mobile producer cannot be routed, which interrupts the requests of consumers and reduces the quality of network service. To solve the above problems, a mobility support mechanism based on dynamic forwarding path management is proposed to ensure the stability of the network. This mechanism updates the forwarding state of ICN routers on the data plane by dynamically tracking the producer's network location to reduce the unreachability of content during producer movement. The simulation results show that DYFP can fully support producer mobility. In the speed range of 0–30 m/s, the average packet loss rate caused by mobility is 8.1%, and the average retrieval delay is stable at 287.3ms.

Keywords: Information centric networking · Mobility management · Forwarding strategy · Transmission delay · Request response ratio

1 Introduction

Information-centric networking [1] (ICN) is a proposed Internet architecture, which proposes a new retrieval model that only focuses on the content itself but not the location of the content, and realizes flexible data transmission by naming content to cope with the increasing growth of data traffic [2]. Furthermore, the identity-location separation and stateless connection of the ICN paradigm can potentially facilitate user mobility [3]. In ICN, user mobility can be divided into consumer mobility and producer mobility [4]. Because of the consumer-driven nature, ICN can naturally support consumer mobility [5]. However, since there is no separation between locators and content identifiers, how to improve producer mobility is a major challenge [6–8].

Many techniques have been proposed to solve the producer mobility problem [9]. The techniques can be classified into the following three categories: the routing-based approach [10, 11], which utilizes the multi-path forwarding properties of ICN's Interests (interest packets) [12, 13] and the ability of routers to advertise prefixes to facilitate

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 219–227, 2023.
https://doi.org/10.1007/978-981-99-2653-4_27

routing update, but it will cause frequent routing updates and low routing aggregation. The indirection approach [14, 15] maps a producer's permanently valid content name to a routable name, which can identify the producer's temporary location through an indirection node. During the movement of the producer, the mapping relationship will be updated according to the new routable name of the producer. However, this approach needs to re-encapsulate the consumer's Interests [16], affecting the communication efficiency. The tracing-based approach [17] leverages and extends the stateful forwarding plane to retrieve the consumer Interests. Whenever the producer moves, it sends trace command Interests towards the RV (Rendezvous Server), creating or updating a soft-state hop-by-hop reverse path. Due to the uncertainty of the mobile producer's movement, this approach may cause high consumption and delay when maintaining the tracing path [18].

To tackle such problems in a simple and effective way, a mobility management mechanism, named DYFP, which introduces a dynamic update method of forwarding table is proposed. By tracking the producer's network location and updating the router's forwarding state on the data plane, DYFP improves the reachability of the mobile producer content. Firstly, by utilizing the consumer-driven characteristics of ICN, a new control message is designed to track the mobile producers' position dynamically. Secondly, through the producer's moving information, a Temporary FIB (TFIB) is created, and the pending interest table (PIT) is update in ICN router to correctly forward consumer requests. Finally, DYFP designs a fast recovery mechanism for consumer requests, which reduced the packet loss rate and retrieval delay while avoiding possible forwarding loops. This mechanism makes full use of the decoupling of locator and content in ICN, and realizes seamless communication between consumers and the mobile producers. Meanwhile, it reduced the handover delay caused by mobile producers and ensured a sufficiently low packet loss rate.

2 Model Design of DYFP

2.1 System Overview

DYFP creates a temporary forwarding path to ensure normal communication between content consumers and mobile producers. With the help of a rendezvous server (RV) that is reachable through the routing plane, DYFP can track the mobile producer's position by leveraging real-time control messages of the producer, which is designed a fast and lightweight handover. Meanwhile, DYFP keeps the session continuity by dynamically building a temporary FIB (TFIB), which relays current producer's position and the RV. In addition, in order to enable the mobile producers to obtain consumer requests that they did not receive during handover, DYFP designs a PIT update mechanism to achieve fast forwarding mechanism of the producer's data.

According to different handover events, a mobile producer sends different mobility messages to RV during the movement process. After that, the RV records the mobility information of the producer and replies to corresponding Data (data packet). DYFP relies on the communication between RV and the producer to create/update the TFIB and PIT of ICN routers to retrieve mobile producers' content. The dynamic mobility messages of the producer enable RV to know the producer's movement status at any time. Then

the RV sends corresponding Data to establish a forwarding path to the new location of the producer. This mechanism allows consumers to obtain the requested content from the mobile producer in time.

DYFP designs special Interest/Data to trigger ICN routers to create/update the TFIB and PIT. Firstly, the Content name field is used as the unique identifier of the data, which is composed of "/RV prefix" and "/Content prefix". The "/RV prefix" is announced to the routing plane by RV, through which the producer's special Interests can be directed to the RV. The "/Content prefix" is the name prefix of the mobile producer and is the unique name identifier of the consumer requesting producer's data. In addition, DYFP adds an extra field Label based on the original ICN packet format, and the field is filled with different flags to indicate the different movement states of the producers. This field is used to trigger the update operation of RV and ICN routers. The ICN routers will be triggered by the special flag and extract the "/Content prefix" from the Content name field to create/update TFIB and PIT entries. Finally, the Signature Information name component field is also added to the Interest format. This field contains relevant information necessary to verify the producer's content (such as the mobile producer's ID, signature information, etc.), so that RV can verify the origin of the "/Content prefix" and ensure the authenticity and reliability of the producer's data.

DYFP adds an extra field Label to the producer's movement information to trigger special operations of RV and ICN routers. The producer will send different mobile Interests MI or CI according to different mobile states. Among them, the Label field of MI is marked with "MoB", which indicates the disconnection state of the producer. The Label field of CI is marked with "CoN", indicating the connection status of the producer.

2.2 Forwarding Path Establishment

By judging and operating MD/CD, ICN router establishes a temporary forwarding path to correctly retrieve the consumer Interests. Figure 1 shows the process of the router establishing/updating the TFIB and PIT. After the router receives MD/CD, it will first process it as a normal Data, that is, match it with the PIT entry. If a match is found, the router queries its Label field and forwards it downstream. If a matching flag is found in the Label field, the router will record the corresponding interface information of the corresponding MD/CD and the content prefix/Content prefix of the producer. After that, the router will judge the Label field. For different flags, the router will perform different operations. If marked as "MoB", the router extracts the incoming interface information of the MD. If marked as "CoN", the router will extract the incoming interface information of the CI in the PIT table. Next, the router creates/updates the TFIB entry with the extracted /Content prefix and interface information. Finally, for the "CoN" flag, the router accesses the content data field in the CD. If this field contains the consumer's Interests information, the router extracts this information and retrieves the PIT table. If the corresponding PIT entry exists, the router adds the incoming interface information of the CD to the interface information of the PIT entry as the incoming interface of Interests. If the PIT entry does not exist, the router extracts the CD's incoming interface information to create a new PIT entry.

The FIB table created or updated by the ICN router during packet processing is called a temporary forwarding table (Temporary FIB, TFIB). The TFIB coexists with the normal

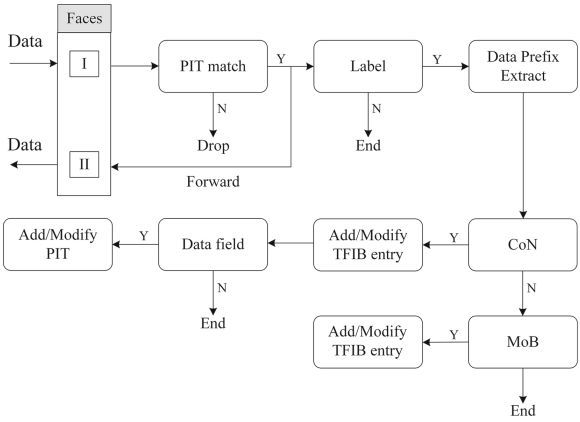


Fig. 1. Forwarding strategy proposed

FIB entry created by the routing plane, but has a higher priority than the normal FIB. The consumer Interests will first retrieve the TFIB during the forwarding process, and then retrieve the FIB after the lookup fails. When creating a TFIB entry, if no "/Content prefix" entry exists in the router, a timer is set when the entry is created. The timer is updated if the entry exists. This timer mechanism prevents a sharp increase in entries in TFIB. In addition, the producer will dynamically send the movement information to update the TFIB entry during the movement process to avoid the entry being invalid because of the disconnection of the producer. Finally, DYFP improves the lifetime of TFIB entries and avoids the expiration of TFIB entries causing consumer Interests to be forwarded to the producer's old location.

3 Performance Evaluation

This section details the evaluation of DYFP. We implemented DYFP through the ndnSIM simulator [19]. Firstly, DYFP is implemented completely in ndnSIM, and then several adopt metrics and simulation experiments are carried out. Finally, we analyze the obtained results to verify the feasibility and effectiveness of DYFP.

3.1 Simulation Setup

Figure 2 shows a grid-5d infrastructural network topology for simulation consists of 25 routers with distance 100m between each two. Among them, each router functions as both PoA (Point of Access) and forwarder. Wired links have a capacity of C = 100Mb/s and 2ms delay. In addition, we set the transmission range of PoA nodes as 50m, which indicates that once the producer moves out of the communication region, it may no longer receive requests from consumers. Finally, the consumer is connected to node 5, the producer's initial location is at node 1, and the RV or map server is connected to node 13.

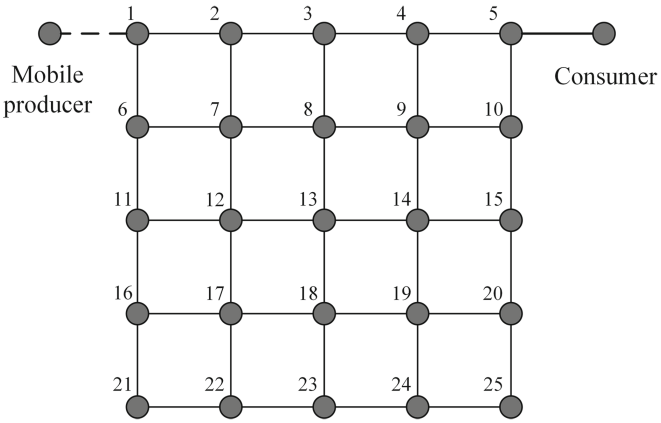


Fig. 2. Simulation network topology

To evaluate the performance of DYFP, a dynamic scenario simulating producer mobility is designed. The Random Walk model is adopted [20], in which the producer moves a fixed distance at a constant speed and then randomly changes direction. Moreover, we vary mobile producer's speed from 0m/s to 30m/s and perform 200 simulation runs for each different speed. In each simulation, the consumer continuously retrieves data from a mobile producer for 100 s. The specific simulation parameter values are shown in Table 1.

Table 1. Simulation parameters

parameter	Value
Wired link bandwidth/(Mbit·s^{-1})	100
Wired link delay/ms	2
Link layer protocol	IEEE 802.11n
Wireless rate	DsssRate1Mbps
Wireless signal transmission range/m	50
Propagation loss model	Rayleigh-fading model
Mobility model	Random Walk model
Fixed moving distance/m	200
Ways to router cache	Nocache
Interest sending rate/(packet·s^{-1})	1
Producer moving speed/(m·s^{-1})	0 ~ 30
Simulation time/s	100

To analyse the pros and cons of the experimental performance, DYFP is compared with two mobility support strategies, the tracing-based scheme KITE [21] and the rendezvous-based scheme (RB) [22]. In addition, the packet loss ratio (Failed Answer Ratio, FAR) and average request delay (Average Request Delay, ARD) indicators were measured to verify the mobility support capability of the solution.

3.2 Results and Discussion

1) Analysis of packet loss rate

Figure 3 shows the simulation results of the average packet loss rate FAR, one of the performance indicators for delay-sensitive traffic. FAR is represented by the number of retransmitted Interests over the total number of Interests sent by all the consumers in the topology. The experienced average packet loss is a consequence of mobile speed (from 0m/s to 30m/s) of different solutions: 8.1% for DYFP, 17.6% for KITE, 31.9% for RB.

The reason is that DYFP utilizes the producer's mobile information to dynamically establish an effective forwarding path. When the producer disconnects, the consumer requests will be recorded by RV. When the producer reconnects to the network, it will send the data requested by the consumer immediately. At the same time, DYFP sets an appropriate lifetime of the TFIB table, so that the consumer requests will not be forwarded to the old location of the producer due to expiration, thereby reducing the FAR. In KITE, the temporary forwarding path established by the mobile producer is invalidated each time it disconnects, so the consumer's requests are forwarded to the producer's old location and discarded due to timeout. New traces are established only after the producer reconnects to the network. In RB, consumers request the producer's new location information from the mapping server before requesting data. However, the mapping information is not real-time, which results in high FAR. In addition, the higher handover frequency of producers increases the hysteresis of the location information obtained by consumers. Therefore, as the speed of producers increases, the FAR

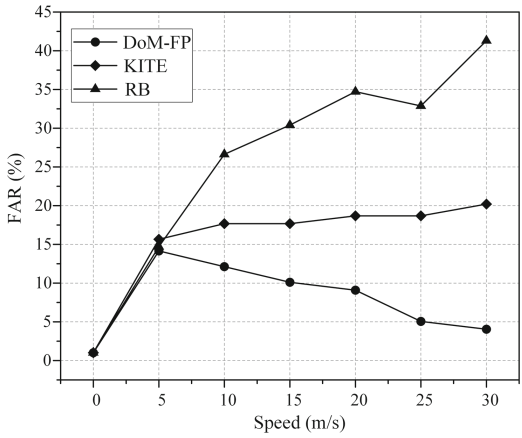


Fig. 3. Evaluation of packet loss rate

gradually increases. In conclusion, the FAR of DYFP is lower than that of KITE and RB.

2) Analysis of Average Request Delay

Figure 4 shows the simulation results of the average request latency ARD for consumer requests, which indicates the efficiency of the solution to support the producer's mobility. ARD specifies the duration time, in second, between consumers send Interests and the arrival of the corresponding contents. In the speed range of 0-30m/s, as the speed increases, the ARD of DYFP gradually decreases, and finally stabilizes at about 230ms. The ARD of KITE gradually stabilized at 600ms, while the ARD of RB continued to increase.

In principle, the fast recovery mechanism of DYFP reduces its ARD. After receiving the CD returned by the RV, the ICN router will extract the consumer requests from the CD's data field to create/update the PIT table. This mechanism enables the mobile producer to return the Data requested by the consumer immediately, without waiting for the consumer to resend Interests, thereby reducing the ARD. For KITE, it maintains a constant transmission delay. The consumer requests may be forwarded to the old location of the producer, and the Interest retransmitted by the consumer will be discarded due to the PIT entry, so its ARD is higher than DYFP. In RB, the producer will handover frequently at high speed, which causes the location information obtained by the consumer expiring frequently. The consumer needs to re-request the latest location information of the producer from the mapping server. Therefore, RB is not suitable for frequent handover of the mobile producer. To sum up, compared with KITE and RB, DYFP can support the communication between consumers and high-speed producers.

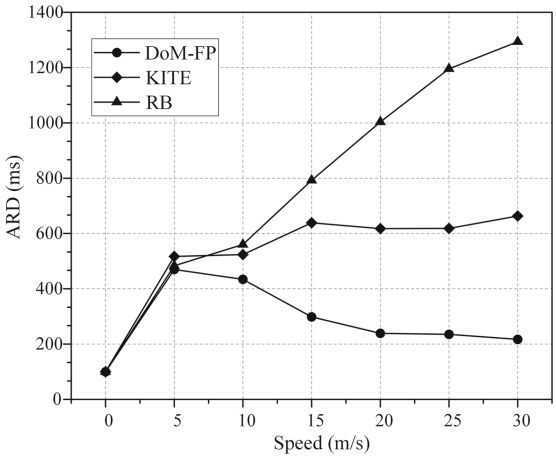


Fig. 4. Evaluation of average request delay

4 Conclusion

Information-centric networking is a proposed Internet architecture, which aims to cope with the increasing traffic demands on the Internet. The research on producer mobility is of great significance for the development and improvement of ICN. Firstly, the limitations of existing schemes are analyzed, and the challenges are summarized for producer mobility. Then, a mobility management mechanism, named DYFP, based on the dynamic update of the forwarding table is proposed. DYFP involves the creation of the TFIB, the update of the PIT, the design of the new control messages, and the response of the routers to different messages. DYFP designs the producer's mobile information to establish a forwarding path between the routable RV and the mobile producer, which effectively reduces the packet loss during handover and ensures the normal communication of users in the mobile environment. In addition, DYFP designs a fast recovery mechanism for consumer requests, which improves the content response rates. Finally, the simulation results show that DYFP can fully support the producer's mobility in the wireless mobile environment, and significantly reduce the packet loss rate and retrieval delay during the movement of the producer. In the future, it is necessary to further evaluate the efficiency of DYFP in practical scenarios and conduct correlation analysis on multiple performance indicators.

Acknowledgements. This work was supported by the Tianjin Science and Technology Plan Project under Grant No. 20JCQNJC01490.

References

1. Ahlgren, B., Dannewitz, et al.: A survey of information-centric networking. In: IEEE Communications Magazine Articles News & Events of Interest to Communications Engineers (2012)
2. Li, Z., Liu, J., Yan, L., Zhang, B., Luo, P., Liu, K.: Smart name lookup for NDN forwarding plane via neural networks. In: IEEE/ACM Transactions on Networking, vol. 30, no. 2, pp. 529–541
3. Tyson, G., Sastry, N., Cuevas, R., et al.: A survey of mobility in information-centric networks. Commun. ACM **56**(12), 90–98 (2013)
4. Araujo, F., Sousa, A.D., Sampaio, L.N.: SCaN-Mob: an opportunistic caching strategy to support producer mobility in named data wireless networking. Comput. Netw. **156**, 62–74 (2019)
5. Li, Z., Liu, K., Zhao, Y., et al.: MaPIT: an enhanced pending interest table for NDN with mapping bloom filter. IEEE Commun. Lett. **18**(11), 1915–1918 (2014)
6. Fang, C., Yao, H., Wang, Z., et al.: A survey of mobile information-centric networking: research issues and challenges. IEEE Commun. Surveys Tutorials **20**(3), 2353–2371 (2018)
7. Gündoğan, C., Kietzmann, P., Schmidt, T.C., et al.: A mobility-compliant publish–subscribe system for an information-centric Internet of Things. Comput. Netw. **203**, 108656 (2022)
8. Zhuo L, Chen Y, Shi H, et al. NDN-GSM-R: a novel high-speed railway communication system via Named Data Networking. EURASIP J. Wirel. Commun. Netw. 2016, 2016(1)
9. Hussaini, M., Naeem, M.A., Kim, B.S., et al.: Efficient producer mobility management model in information-centric networking. IEEE Access **7**, 42032–42051 (2019)

10. Meddeb, M., Dhraief, A., Belghith, A., et al.: AFIRM: Adaptive forwarding based link recovery for mobility support in NDN/IoT networks. Futur. Gener. Comput. Syst. **87**, 351–363 (2018)
11. Kim, D., Kim, J., Kim, Y., et al.: End-to-end mobility support in content centric networks. Int. J. Commun Syst **28**(6), 1151–1167 (2015)
12. Z. Li, Y. Xu, B. Zhang, L. Yan, and K. Liu, Packet forwarding in named data networking requirements and survey of solutions, IEEE Commun. Surveys Tuts. **21**(2), 1950–1987, 2nd Quart (2019)
13. Yan L, Li Z, Liu K. Learning Tree: Neural Network-based Index for NDN Forwarding Plane[C]// the ACM SIGCOMM 2019 Conference Posters and Demos. ACM, 2019
14. Hermans F, Ngai E, Gunningberg P. Global source mobility in the content-centric networking architecture. In: Proceedings of the 1st ACM workshop on emerging name-oriented mobile networking design-architecture, algorithms, and applications. 2012: 13–18
15. Yan Z, Park Y J, Leau Y B, et al. Hybrid network mobility support in named data networking. In: 2020 International Conference on Information Networking (ICOIN). IEEE pp. 16–19 (2020)
16. Augé, J., Carofiglio, G., Grassi, G., et al.: Map-me: Managing anchor-less producer mobility in content-centric networks. IEEE Trans. Netw. Service Manag. **15**(2), 596–610 (2018)
17. Zhang, Y., Afanasyev, A., Burke, J., et al.; A survey of mobility support in named data networking. In: 2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). pp 83–88, IEEE (2016)
18. Alajlan, M., Belghith, A.: Supporting seamless mobility for real-time applications in named data networking. Procedia Comput. Sci.**110**, 62–69 (2017)
19. Mastorakis, S., Afanasyev, A., Zhang, L.: On the evolution of ndnSIM: An open-source simulator for NDN experimentation. ACM SIGCOMM Comput. Commun. Rev.**47**(3), 19–33 (2017)
20. Camp, T., Boleng, J., Davies, V.: A survey of mobility models for ad hoc network research. Wirel. Commun. Mob. Comput. **2**(5), 483–502 (2002)
21. Zhang, Y., Xia, Z., Mastorakis, S., et al.: Kite: producer mobility support in named data networking. In: Proceedings of the 5th ACM Conference on Information-Centric Networking. pp. 125–136 (2018)
22. Han, D., Lee, M., Cho, K., et al.: Publisher mobility support in content centric networks. In: The International Conference on Information Networking 2014 (ICOIN2014). IEEE, pp. 214–219 (2014)

SR-FIB: A NewIP Forwarding Information Base via Neural Networks

Hui Fan[1], Tianxiang Ma[2], Xin Duan[2], Weina Liu[2], Zhuo Li[3(✉)], Jindian Liu[3], and Binzhi Wang[3]

[1] State Grid Hebei Electric Power Corporation, Shijiazhuang 050021, China
[2] State Grid Hebei Electric Power Research Institute, Shijiazhuang 050021, China
[3] CoopNet Co., Ltd. at Shanxi, Shanxi 037000, China
zli@tju.edu.cn

Abstract. Aiming at the problem of fast indexing and efficient storage of addresses in the SRv6 protocol of NewIP network, a forwarding information base (FIB) that supports fast retrieval of IPv6 addresses was designed, called SR-FIB. The neural network was used to learn the distribution characteristics of IPv6 addresses in the memory and build an index mapping model that can achieve uniform address mapping. The storage structure and retrieval algorithm that can effectively support the mechanism of longest prefix matching were designed. The simulation results show that SR-FIB can achieve rapid retrieval of IPv6 addresses under the SRv6 protocol. Under the requirement of 1% network packet loss rate, the memory consumption is only 26.14 MB and the throughput is 11.48 MPPS, which can meet the requirements of network retrieval.

Keywords: NewIP · forwarding information base · route retrieval · neural network · data structure

1 Introduction

With the emergence and application of big data, cloud computing, artificial intelligence, etc., the Internet has ushered in a new round of revolution. Holographic communication, consciousness communication, and integrated communication of space, sky, earth and ocean have come, and an intelligent world where everything perceives and is interconnected with all networks is coming [1]. Numerous new applications and technologies have brought some requirements and challenges to the current IP network. However, compared with the rapid iteration of the Internet applications, the TCP/IP protocol has not made substantial changes in the past 40 years as the basis of the Internet, and the capabilities of the IP network have to be enhanced urgently. As an improvement of IPv4 network, IPv6 network solves the problem of address exhaustion and security, but it has not changed the core of TCP/IP technology, and its inherent defects have not been solved yet [2].

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 228–236, 2023.
https://doi.org/10.1007/978-981-99-2653-4_28

To this end, the Huawei research team focused on the key requirements for data networks in typical application scenarios of network 5.0 [3], and proposed a new network protocol system, NewIP [4]. Through the top-level design, it interconnects a large number of heterogeneous communication entities such as devices, contents, services and people. Meanwhile, it provides a more powerful capacity of connection, transmission, endogenous security and customization. Among them, the new addressing and control technology is one of the most critical enabling technologies in NewIP technology, which can support massive heterogeneous communication entities and interconnection between heterogeneous networks [5]. As a technology to realize the functions of the NewIP network, SRv6 protocol (Segment Routing IPv6, SRv6) is compatible with IPv6 in the network layer protocol, and realize segment routing in the forwarding plane. The SRv6 protocol implements segment routing according to the segment ID (Segment ID, SID) in the segment routing header (Segment Routing Header, SRH) constantly changing the SRv6 destination address [6].

Due to the uncertainty in the scale and distribution of IPv6 addresses, the related algorithms based on the prefix distribution characteristics will be affected significantly. Simultaneously, because of the frequent updates, the related algorithm based on the prefix will aggravate the imbalance of the trie. Thus, it is difficult for the address retrieval in the current network to transplant to the NewIP network. Considering the lack of research on forwarding information base (FIB) that supports fast lookup and forwarding of IPv6 addresses under the SRv6 protocol, a neural network-based FIB called SR-FIB is proposed in this paper to build a FIB under the NewIP network architecture from the bottom up, which can realize fast indexing of IPv6 addresses so that realizing fast packets forwarding.

2 Design Elements of FIB Under SRv6 Protocol

The SID in SRv6 is an IPv6 address with the length of 128-bits, in which 128 bits can be flexibly allocated to the three fields of locator, function and parameter. Among them, the locator represents the unique node identifier in the network, the function represents the function that the network node needs to perform, and the parameter, as a supplement to the function, is optional. Due to the flexibility of SID arrangement, the IPv6 address hierarchy in the SRv6 protocol is more diverse, which brings great challenges to the rapid IPv6 addresses lookup in FIB.

First, IPv6 addresses in NewIP network are highly personalized in different scenarios. When forwarding packets, it is necessary to accurately retrieve the address information in the FIB to match the next-hop interface. Therefore, one of the design elements of the FIB is to build an index with a low false positive probability and fast lookup speed.

Secondly, the future communication is not limited to the hosts, so the number of addresses used will surge due to the addition of various communication subjects and heterogeneous networks [7]. Meanwhile, the number of FIB entries

will reach millions, resulting in the large memory consumption of FIB. The current fast SRAM memory has a small storage space, and the DRAM memory with a large memory space is slow so that it is difficult to meet the requirements of fast forwarding. [8] Therefore, the second design element of FIB is to design a storage structure that is suitable for the current hardware memory storage.

Finally, IPv6 needs to perform the Longest Prefix Matching (LPM) when retrieving the address in FIB [9]. The changes of IPv6 address characteristics make LPM more difficult. Therefore, the third design element of FIB is a fast retrieval algorithm that supports the LPM for IPv6 addresses.

3 SR-FIB

3.1 The Design Overview of SR-FIB

The principle of the proposed index, which can achieve uniform mapping, is as follows. Firstly, a large number of IPv6 addresses are used to construct the training set for the neural network index model, the addresses are sorted according to a certain rule, and the sequence number is used as the label. The neural network model is trained by the above dataset, and the cumulative distribution function (CDF) of IPv6 addresses in memory is learned. After training, the parameters of neural network are extracted to construct the index. When the packet is forwarded, the neural network index model predicts the CDF value of the IPv6 address, and multiplies it by the size of the mapping table to obtain the offset address in the mapping table, then retrieves the actual address in the memory. Due to the uniform distribution characteristics of the CDF, the positions mapped in the table also subject to the uniform distribution.

3.2 The Structure of SR-FIB

The structure of the SR-FIB is shown in Fig. 1, which mainly includes an input unit, an indexing unit, and a mapping unit. The input unit processes 128-bit IPv6 addresses into input vectors of neural network with moderate length. When the dimension of the input vector is too large, the lookup speed will decrease and the memory consumption will increase. In turn, the shorter input vector will cause the concentrated data distribution so that the neural network cannot learn the distribution well. The index unit is used to learn the distributions of IPv6 addresses in the memory, and the neural network in the indexing unit is related to the lookup speed and memory consumption of the index. In this paper, the Back Propagation Neural Network (BPNN) [10] is used to achieve fast calculation, and it is built into a two-layer tower structure to deal with millions of data. The first layer contains a BPNN, which divides millions of data into 1,000 categories. The second level contains 1,000 BPNNs, corresponding to the 1,000-class classification results generated from the first layer, and each BPNN learns the distribution of data retrieved in the memory in each category. In Fig. 1, $\text{BPNN}_{i,j}$ represents the j-th neural network at the i-th. After training,

the neural network at the first layer divides the training set into 1,000 regions, and trains the $\text{BPNN}_{2.j}$ by using the data in corresponding region to learn the CDF in each region. Finally, the second layer of 1,000 BPNNs will fit the CDF of the entire memory space.

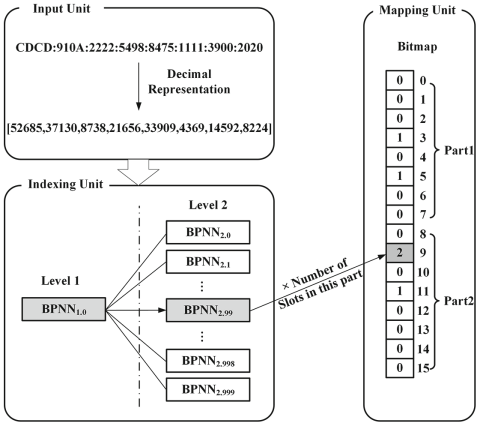


Fig. 1. Basic structure of the model.

The mapping unit deploys a Bitmap to record the offset of the actual index address. The Bitmap is divided into several parts evenly, and each slot records the order where the IPv6 address is inserted into the part to realize the dynamic allocation of the memory. The mapping position can be obtained by multiplying the CDF value predicted by the model unit by the total number of slots in Bitmap. Finally, according to the address offset recorded in the slot, the off-chip actual memory address can be accessed. An example of an IPv6 address lookup through this index model is shown in Fig. 1. For an input IPv6 address CDCD:910A:2222:5498:1111:3900:2020, first convert the hexadecimal to decimal representation to get the input vector [52685, 37130, 8738, 21656, 33909, 4369, 14592, 8224]. The vector is input into the index unit, the $\text{BPNN}_{1.0}$ calculates the class of 99, then the input vector is input the $\text{BPNN}_{2.99}$ to calculate the CDF value. Assuming that the CDF value calculated by $\text{BPNN}_{2.99}$ is 0.6, the position of the IPv6 address in the Bitmap is 0.6×15 (number of slots) $= 9$. Since position 9 is located in the second part of the Bitmap, where the recorded offset is 2. Thus, the actual of the IPv6 address is equal to the base address of the second part plus the offset address of 2.

4 Index Architecture and Operation Algorithm of SR-FIB

4.1 Index Architecture of SR-FIB

As shown in Fig. 2, the on-chip storage unit uses SRAM, which can retrieve data with a faster processing speed. The off-chip memory uses DRAM with larger storage space to store millions of address entries. Four index models are deployed on SRAM for the IPv6 address lengths range of 1-32, 33-64, 65-96 and 97-128 respectively. A series of storage pools corresponding to the on-chip model are deployed on DRAM to store the massive forwarding information. When the conflict happens, that is different data indexing to the same address, the chain address method is used in the data storage pool to deal with the conflict.

4.2 Operation Algorithm of SR-FIB

The operation algorithm in SR-FIB mainly includes the lookup and update algorithm of packet forwarding information. The lookup process of SRv6 packets in FIB is as follows. First, the IPv6 address in packet is obtained, and divided into four parts: 1-32, 33-64, 65-96 and 97-128 according to the prefix length. And then input them into SR-FIB with the prefix length from high to low. If there is a matching result in the SR-FIB, the forwarding information corresponding to the longest address prefix is as the next hop interface. If there is no matching result, it indicates that the next hop of the packet does not exist in the router.

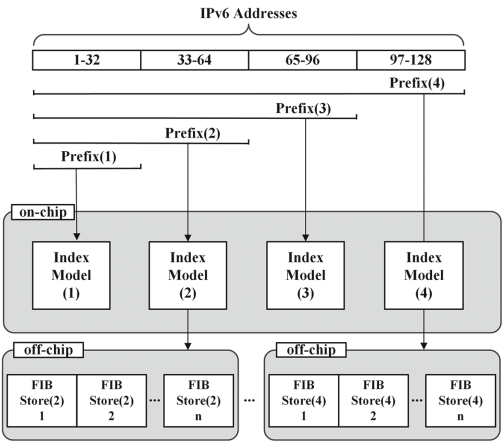


Fig. 2. Index architecture of SR-FIB.

Each node that supports SRv6 will maintain a local SID table, and the FIB will perform update operations such as inserting, deleting, and modifying routing data according to the local SID table. When the update operation is performed,

the longest prefix matching operation needs to be performed on the IPv6 address, and the IPv6 address is inserted into the corresponding SR-FIB. Afterwards the off-chip memory address is indexed according to the output generated from the index model to store the forwarding information. When performing the modification operation, input the IP address to be modified into the index model to calculate the memory address, and modify the forwarding information in the corresponding off-chip storage pool. When the delete operation is performed, the forwarding information stored in off-chip memory and the corresponding storage information in the index model need to be deleted at the same time.

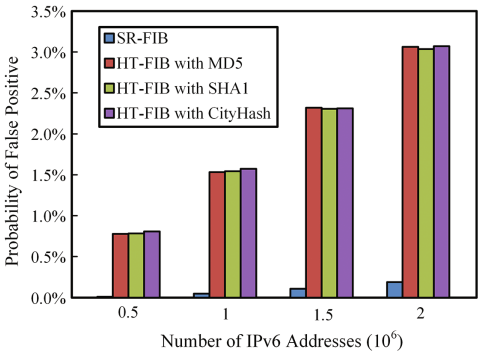


Fig. 3. False positive probability of SR-FIB.

5 Performance Evaluation of SR-FIB

5.1 Experimental Setup

In the section, the performance of SR-FIB is evaluated in terms of false positive probability, memory consumption and throughput. These results are compared with Patricia Trie-FIB [11] and HT-FIB [12]. The performance is tested on a workstation with Intel Xeon E5-1650 v2 3.50 GHz, DDR3 24 GB SDRAM. The SR-FIB index model is trained with Neural Network Toolbox in MATLAB. All the index schemes are implemented in C++ for a fair comparison.

In SR-FIB, each BPNN has 8 input neurons, 20 hidden neurons and 1 output neuron. The size of each slot in the Bitmap is set to 2 bytes. In Patricia Trie-FIB, each entry has 4 bytes, which is a pointer to access the forwarding information. For HT-FIB, it is implemented with MD5, SHA1 and CityHash256 respectively, and the size of each entry also has 4 bytes.

In order to train a relatively perfect neural network model, 100 million IPv6 addresses is used as the training set, and 0.5 million, 1 million, 1.5 million, and 2 million IPv6 addresses are used as the testing sets.

5.2 False Positive Probability

In this subsection, the false positive probability of SR-FIB and the HT-FIB is compared. As Patricia Trie-FIB does not have false positive, it is not tested in this experiment. For fair comparison, the number of slots in Bitmap was set to 32 million. As illustrated in Fig. 3, for 0.5-, 1-, 1.5-, and 2- million IPv6 addresses, the false positive probability of SR-FIB is 0.01%, 0.05%, 0.11%, and 0.19% respectively, which is much lower than the tolerance of the Internet's 1% packet loss rate. On the contrary, the false positive probability of HT-FIB is much higher than that of SR-FIB. Considering the comprehensive performance of hash functions, CityHash256 is used as the hash function representative for other tests.

5.3 Memory Consumption

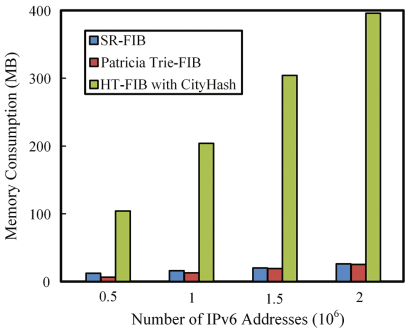


Fig. 4. Memory consumption of SR-FIB.

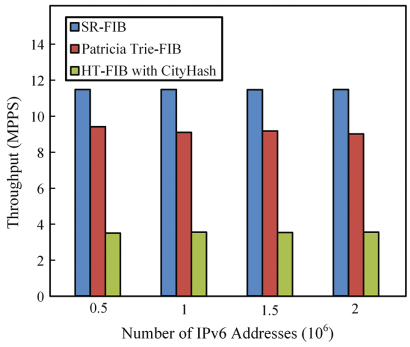


Fig. 5. Throughput of SR-FIB.

As shown in Fig. 4, the memory consumption of Bitmap in SR-FIB is 6 MB, 10 MB, 14 MB and 20 MB for different testing sets respectively. Additionally, the memory consumption of neural network models is $(20 \times 8 + 20 \times 1 + 1 \times 20 + 1 \times 1) \times 8\text{B} \times 1{,}001 \times 4 \approx 6.14\,\text{MB}$. Therefore, the total memory consumption of SR-FIB is 12.14 MB, 16.14 MB, 20.14 MB and 26.14 MB for 0.5-, 1-, 1.5- and 2 million data respectively. The memory consumption of the SR-FIB and Patricia Trie-FIB is much lower than that of HT-FIB scheme, as SR-FIB learns the distributions of IPv6 addresses in memory by neural network to achieve uniform mapping, which greatly reduces memory consumption. The Patricia Trie-FIB reduces memory consumption by aggregating the same prefix. With the increase of the number of data retrieved, however, the memory consumption of SR-FIB tends to be better than that of Patricia Trie-FIB.

As illustrated in Fig. 5, the throughput of SR-FIB and Patricia Trie-FIB is much higher than that of HT-FIB, as SR-FIB and Patricia Trie-FIB can be deployed on SRAM. The throughput of SR-FIB has reached 11.49 MPPS (Million Packets Per Second). The reason is that even if the training process of SR-FIB is extremely complicated, the it is not considered in lookup process, and the trained parameters are only extracted for calculation. In addition, the BPNN used in SR-FIB has a compact structure, fast running speed, and can be deployed on SRAM, so the lookup speed is only determined by the on-chip model indexing speed. The throughput of Patricia Trie-FIB is lower than that of the SR-FIB scheme due to its higher depth which affects the lookup speed.

6 Conclusion

This paper proposed an SR-FIB that supports IPv6 address retrieval and forwarding in NewIP network SRv6 protocol. By using a neural network to learn the distributions of IPv6 addresses in memory, an SR-FIB can achieve uniform data mapping. Moreover, SR-FIB adopts on-chip and off-chip two-level storage unit structure, which realizes fast lookup speed and efficient storage of forwarding information. With the designed operation algorithm, efficient indexing of data is realized under the premise of satisfying the longest prefix matching. The experimental results show that under the condition of 1% false positive probability, the memory consumption of SR-FIB is only 26.14 MB and the throughput is as high as 11.49 MPPS for millions of data. The future work is to study its deployment strategy and test the comprehensive performance of SR-FIB in the FPGA.

Acknowledgments. This work was supported by the Key Research and Development Projects of Hebei Province under Grant 20314301D and the Science and Technology Project of SGCC under Grant kj2022-023 and the Tianjin Science and Technology Plan Project under Grant 20JCQNJC01490.

References

1. Zhuo, L., Yaping, X., Baichuan, Z., Liu, Y., Kaihua, L.: Packet forwarding in named data networking requirements and survey of solutions. IEEE Commun. Surv. Tutor. **21**, 1950–1987 (2019)
2. Zhuo, L., Jindian, L., Liu, Y., Baichuan, Z., Peng, L., Kaihua, L: Smart name lookup for NDN Forwarding plane via neural networks. IEEE/ACM Trans. Networking **30**, 529–541 (2022)
3. Xiuli, Z., Sheng, J., Chuang, W.: A potential direction of next generation data communication network-network 5.0. Inf. Commun. Technol. **11**, 37–44 (2020)
4. Zhe, C., Chuang, W., Guanwen, L., Zhe, L., Sheng, J., Alex, G.: New IP framework and protocol for future applications. In: NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium, pp. 1–5. IEEE, Budapest (2020)
5. Xiuli, Z., Sheng, J., Chuang, W.: NewIP: new connectivity and capabilities of upgrading future data network. Telecommun. Sci. **35**, 2–11 (2019)

6. Filsfils, C., Dukes, D., Previdi, S.: IPv6 segment routing header (SRH): RFC8754. https://www.rfc-editor.org/rfc/rfc8754.html
7. Zhuo, L., Yutong, C., Heping, S., Kaihua, L.: NDN-GSM-R: a novel high-speed railway communication system via named data networking. EURASIP J. Wirel. Commun. Netw. **48**, 1–5 (2016)
8. Zhuo, L., Kaihua, L., Yang, Z., Yongtao, M.: MaPIT: an enhanced pending interest table for NDN with mapping bloom filter. EURASIP IEEE Commun. Lett. **18**, 1915–1918 (2014)
9. Dharmapurikar, S., Krishnamurthy, P., Taylor, D,E.: Longest prefix matching using bloom filters. IEEE/ACM Trans. Netw. **14**, 397–409 (2006)
10. Liu, Y., Zhuo, L., Kaihua, L: Learning tree: neural network-based index for NDN forwarding plane. In: Proceedings of the ACM SIGCOMM 2019 Conference Posters and Demos, pp. 62–65. ACM, Beijing (2019)
11. Morrison, D,R.: PATRICIA-practical algorithm to retrieve information coded in alphanumeric. J. ACM (JACM) **15**, 514–534 (1968)
12. Kirsch, A., Mitzenmacher, M., Varghese, G.: Hash-based techniques for high-speed packet processing. In: Cormode, G., Thottan, M. (eds.) Algorithms for Next Generation Networks. Computer Communications and Networks, pp. 181-218. Springer, London (2010). https://doi.org/10.1007/978-1-84882-765-3_9

A Learned Classification Index for PIT in Named Data Networking

Xin Duan[1(✉)], Tianxiang Ma[1], Zhiyuan Zhang[2], Gengsen Wang[1], Weina Liu[1], Jindian Liu[3], Zhengzhou Han[3], and Zhuo Li[3]

[1] State Grid Hebei Electric Power Research Institute, Shijiazhuang 050021, China
m18931223557@163.com
[2] State Grid Hebei Electric Power Corporation, Shijiazhuang 050021, China
[3] CoopNet Co., Ltd. at Shanxi, Shanxi 037000, China

Abstract. In order to tackle the issues of rapid name lookup, efficient storage and the name lookup algorithm in Pending Interest Table(PIT) of Named Data Networking. An index for PIT called LCI-PIT was proposed, which can learn the distribution regularity of names in each category in the static memory to achieve more uniform mapping and improve memory utilization after dividing the name space into many categories. The on-chip memory using SRAMs deployed classification mapping model, while the off-chip memory using DRAMs deployed multiple PIT stores. It not only can retrieve the forwarding information of the Interest packets and the Data packets by using the corresponding name lookup algorithm, but also realize fast name lookup. The experimental results show that LCI-PIT can achieve the efficient performance in terms of memory consumption, false positive probability and throughput. LCI-PIT significantly reduced memory consumption to 1.08MB with the probability of false positive less than 1%, which means it can deploy on SRAMs in commercial line cards. The throughput of LCI-PIT was about 290.36 million search per second, which met current network requirements of fast packet processing.

Keywords: named data networking · pending interest table · neural network · name lookup

1 Introduction

With the continuous expansion of the Internet scale and the emergence of innovative technologies, such as big data and high-definition video transmission, Internet services are gradually evolving towards content and personalization. However, the traditional Internet architecture has been unable to meet the business requirements of the future network [1]. Therefore, many research teams are actively exploring and studying the innovative Internet architecture. As one of the typical paradigms of new Internet architectures, named data networking (NDN) was proposed in 2010 [2].

In NDN, all communication is driven by consumers [2], through the exchange of two distinct types of packets carrying name identifiers: Interest and Data.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 237–244, 2023.
https://doi.org/10.1007/978-981-99-2653-4_29

In order to fetch the desired data, a consumer sends an Interest packet with a name identifying name to the network. When an Interest packet reaches at a node, the pending interest table (PIT) will aggregate forwarding information such as the incoming interface according to the name of Interest packet, and then wait for the return of the required data. When the Interest arrives at a node that has the desired data, the data will retrieve the Interest matching its name in the PIT, and then transmit it back to the consumer according to the incoming interface of the Interest recorded in the PIT [3]. Since the forwarding process of both Interest and Data requires needed to be retrieved in PIT, the design of the PIT will directly affect the performance of the forwarding plane of NDN. However, the research on PIT still faces a series of issues and challenges. Firstly, the NDN names that are variable-length with no upper bound are more complex than fixed-length IP addresses [2], which brings higher requirements for fast lookup speed of PIT. Secondly, the number of PIT entries in the millions makes efficient storage of PIT a challenge. Finally, the Exact Name Matching (ENM) and the All Name Prefix Matching (ANPM) in PIT are still problems to be solved [4].

Since NDN was proposed in 2010, many solutions have been proposed for the fast name lookup and efficient storage of PIT. The current scheme for PIT is mainly based on three data structures. Trie, hash table and Bloom filter. For example, Dai et al. proposed the CONSTER, which uses ASCII code to store forwarding information efficiently [4]. Song et al. proposed the binary Patricia Trie, which uses the binary as the granularity to minimize the impact of redundant information at memory [5]. Yuan proposed the FHT to reduce memory consumption by storing fixed-length fingerprints [6]. Li et al. proposed an enhanced data structure called mapping Bloom filter, which utilizes Bloom filter and bitmap to reduce memory consumption [7]. However, the currently proposed schemes are difficult to take into account the requirements of PIT for name lookup speed and memory consumption, and do not sufficiently consider how to support the ENM and ANPM for name lookup. To this end, an efficient learned index for PIT called LCI-PIT (Learned Classification Index-PIT) is proposed in this paper.

2 Overview and Architecture of LCI-PIT

2.1 Overview of LCI-PIT

In order to meet the requirements of PIT, a learned classification index for PIT called LCI-PIT is proposed to efficient name lookup of PIT. LCI-PIT divides the huge name space into lots of name subspace and uses neural network to learn the distribution of names in each subspace to fits the cumulative distribution function (CDF) to build a classification mapping model. The namespaces classification allows the neural network to learn more precisely and fit a more accurate CDF so as to achieve uniform mapping, efficient storage, and fast name lookup.

The structure of the classification mapping model includes three parts: input unit, classification unit and mapping unit, as shown in Fig. 1.

The input unit converts the variable-length name into a 5-dimensional fixed-length vector to input into the neural network. Suppose the input to the neural

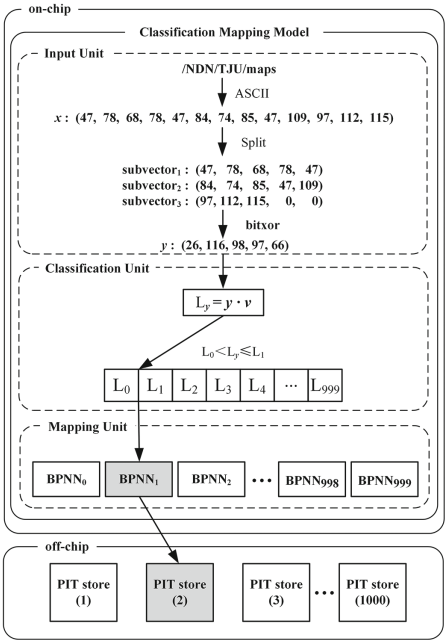


Fig. 1. Index architecture of LCI-PIT.

network is $\boldsymbol{y} = \left(y_0, y_1, \ldots, y_{N-1}\right)$, in order to convert a name with length n into a 5-dimensional vector, $\boldsymbol{y}$ can be treated as a n-dimensional vector $\boldsymbol{x} = (x_0, x_1, \ldots, x_{n-1})$. For $\boldsymbol{x}$, each dimension is the ASCII of the characters. The generation of the input vector is as follows:
if $n \leq N$,

$$y_i = \begin{cases} x_i, & 0 \leq i < n, \\ 0, & n \leq i < N, \end{cases} \tag{1}$$

if $n > N$,

$$\boldsymbol{y} = f\left(subvector_1, subvector_1, \ldots, subvector_N\right), \tag{2}$$

where *subvector* is the subvector obtained by splitting the vector by each 5 elements. Thus, the name can be converted into a 5-dimensional vector. Due to the use of ASCII and the bitwise XOR operation, $\boldsymbol{y}$ satisfies $y_i \in [0, 127] \cap \mathbb{Z} (i = 0, 1, 2, 3, 4)$, where $\mathbb{Z}$ is the set of integers.

Secondly, the future communication is not limited to the hosts, so the number of addresses used will surge due to the addition of various communication subjects and heterogeneous networks. Meanwhile, the number of FIB entries will reach millions, resulting in the large memory consumption of FIB. The current fast SRAM memory has a small storage space, and the DRAM memory with a large memory space is slow so that it is difficult to meet the requirements of fast

forwarding. Therefore, the second design element of FIB is to design a storage structure that is suitable for the current hardware memory storage.

The classification unit utilizes an array called classified array to find the corresponding category for each name. Classified array is constructed as follows. Firstly, collect a large number of names, and convert them into a 5-dimensional vector by input unit. Then, sort the vector with its values from small to large. Subsequently, the vector set is equally divided into 1000 categories, each of which serves as a training set for the neural networks in mapping unit. Since the elements in the input vector are all integers from 0 to 127, it can be regarded as being obtained by adding L base vectors $\boldsymbol{b}^N = (0, 0, \ldots, 0, 1)$ according to 128 system, that is $\boldsymbol{y} = L\boldsymbol{b}$. For example, $(0, 0, \ldots, 1, 127) = (0, 0, \ldots, 1, 0) + (0, 0, \ldots, 0, 127) = (128 + 127)\,(0, 0, \ldots, 0, 1)$. At this point, the L corresponding to the last vector $\boldsymbol{y}^{(i)} = \left(y_0^{(i)}, y_1^{(i)}, \ldots, y_{N-1}^{(i)}\right)\,(i = 0, 1, \ldots, m-1)\,(i = 0, 1, \ldots, 999)$ in all classes is calculated. L is calculated as follows,

$$L_i = \boldsymbol{y}^{(i)} \cdot \boldsymbol{v}, \tag{3}$$

where $\boldsymbol{v} = \left(128^{N-1}, 128^{N-2}, \ldots, 128, 1\right)$. Finally, L_i will be stored in the array to build the classified array. For an input vector $\boldsymbol{y}$, is calculated by formula (3), and then the class number of $\boldsymbol{y}$ can be determined by comparing with the elements in the categorical array. If $L_i < L_{\boldsymbol{y}} \leq L_{i+1}$, it indicates that the class number of the name is $i+1$. In order to improve the efficiency of the comparison, the classification unit adopts a binary search algorithm with a time complexity of $\boldsymbol{O}\left(\log_2 m\right)$.

The mapping unit consist of 1000 small back propagation neural networks (BPNN), each BPNN trains the name in a category, and fits the CDF of names. After training, the mapping unit maps names to the static memory.

To access the forwarding information of the name, each slot in the classified array sets a pointer to point to a static memory that stores the forwarding information of a class of name. For a NDN name, the CDF value of the name will be calculated by the corresponding BPNN, and then multiply the CDF value by the total number of slots in the corresponding part to obtain the offset address of the name in the memory. Finally, the address of the name in the memory can be obtained by adding the base address pointed to in the classified array slot and the calculated offset address.

2.2 Architecture of LCI-PIT

In order to realize the rapid name lookup of PIT under the current memory devices, the architecture of LCI-PIT adopts a hierarchical structure, which is composed of two parts: an on-chip storage unit and an off-chip storage unit. The architecture is shown in Fig. 1, the on-chip storage unit utilizes SRAM to realize fast data packet processing, and the classification mapping model is deployed on it. The off-chip storage unit utilizes DRAM with a large memory space, and 1000 PIT stores are deployed on it to store the forwarding information of names. In addition, when multiple different names are mapped to the same off-chip address, the chain address method will be used to handle conflicts in the PIT stores.

2.3 Lookup Algorithm of LCI-PIT

To ensure the normal forwarding of packets in NDN, LCI-PIT will perform different algorithms for Interest and Data. When an Interest arrives at LCI-PIT, the on-chip storage unit retrieves the name of Interest to obtain the memory address, and then the ENM is performed in the corresponding PIT store. If there is a matching entry, its forwarding information is updated, and the Interest is discarded. Otherwise, a new PIT entry is created to store the forwarding information of the Interest. When a Data arrives at the LCI-PIT, the on-chip storage unit retrieves the name of Data to obtain the memory address, then all name prefixes are matched in the PIT store. That is all names with the same name prefix as the Data packet are retrieved. If the matching entry exists, the Data is sent back according to the Interest forwarding information stored in the memory. Otherwise, the Data packet is discarded.

Figure 1 shows an example of name lookup in LCI-PIT. For a packet named /NDN/TJU/maps, input unit converts the name to a 5-dimensional vector (26, 116, 98, 97, 66). Then, calculate L of the vector in classification unit, and determine the category in the classified array. If $L_0 < L \leq L_1$, he category is 1. Then the CDF value is calculated by BPNN_1 in mapping unit. Assuming that the CDF value is equal to 0.5 and the total number of slots of PIT store is 10, the CDF value is multiplied by the total number of slots, and the offset address of the name in PIT store is $0.5 \times 10 = 5$. Therefore, the actual address of the name in the PIT store is the base address pointed by the slot L_1 of classified array plus the offset address 5. In this way, the forwarding information can be accessed in the off-chip PIT store.

3 Performance Evaluation of LCI-PIT

In this section, the performance of LCI-PIT in terms of memory consumption, false positive probability, and throughput is evaluated and compared with Binary Patricia Trie-PIT [6] and Hash Table-PIT [?], which are the typical indexes for NDN forwarding plane.

3.1 Experimental Setup

All experiments are performed on a small workstation configured with Intel i9-10920X 3.5 GHz, DDR4 64 GB SDRAM. The neural network is trained using MATLAB neural network toolbox and implemented by C++. For fair comparison, all contrast schemes are also implemented in C++. In LCI-PIT, each BPNN in mapping unit has 5 input neurons, 20 hidden neurons, and 1 output neuron, and all weight and bias parameters are 8-byte double-precision floating point number [9]. As both the integer numbers and pointers stored in each slot of classified array are 4 bytes, the size of each slot is 8 bytes. For Binary Patricia Trie-PIT, each node needs to record the address pointer with 4 bytes. For Hash Table-PIT, MD5 [?], CityHash256 [?] and xxHash [?] are used as hash functions, and each entry is set to 4 bytes.

Table 1. Memory consumption of LCI-PIT

# of Names	LCI-PIT (MB)	BPT-PIT (MB)	MD5-PIT (MB)	CityHash256-PIT (MB)	xxHash-PIT (MB)
500,000	58.258	16.265	400.000	408.000	408.000
1,000,000		32.507			
1,500,000		48.762			
2,000,000		65.046			

3.2 Memory Consumption

In order to deploy the index on SRAM to realize fast data processing, the memory footprint of the index needs to be compressed as much as possible. For LCI-PIT, Binary Patricia Trie-PIT and Hash Table-PIT, the memory consumption for storing forwarding information is the same, so the memory consumption of this part is not considered. In addition, to meet the requirement that the current Internet packet loss rate should be less than 1% [10], the tests of each index are carried out under the condition of false positive probability is below 1%. For LCI-PIT, its memory consumption is fixed (Table 1).

For Hash Table-PIT, since its memory space is static, the memory consumption of Hash Table-PIT will be evaluated in the case of 2 million names to ensure that the false positive probability for all testing sets is less than 1%. For the Binary Patricia Trie-PIT whose number of nodes varies with the number of names, its memory consumption will be evaluated for 0.5-, 1-, 1.5- and 2 million names, respectively. In LCI-PIT, memory consumption consists of two parts, which are classified arrays and neural network parameters. The memory consumption for the classified array is $1000 \times 8\text{B} = 8000\text{B}$, and the memory consumption caused by neural network parameters is $(20 \times 5 + 20 \times 1 + 1 \times 20 + 1 \times 1) \times 8\text{B} \approx 1.07\,\text{MB}$. Therefore, the memory consumption for LCI-PIT is $8000\text{B} + 1.07\,\text{MB} \approx 1.08\,\text{MB}$. The experimental results are listed in Table 2. The memory consumption of LCI-PIT and BPT-PIT is much lower than that of Hash Table-PIT.

3.3 Probability of False Positive

Aiming at the requirement of the Internet that the packet loss rate should be lower than 1%, the experiment tested the false positive probability of LCI-PIT and Hash Table-PIT when the total number of mapping slots is 32 million. The experimental results are shown in Table 4. For 0.5-, 1-, 1.5- and 2 million names, the false positive probability of LCI-PIT are 0.09%, 0.26%, 0.50%, and 0.82%, respectively. However, the false positive probability of Hash Table-PIT is much higher than that of LCI-PIT, which cannot meet the current Internet requirement mentioned above. The reason is that LCI-PIT can learn the distribution

Table 2. Probability of false positive of LCI-PIT

# of Names	LCI-PIT	MD5-PIT	CityHash256-PIT	xxHash-PIT
500,000	0.09%	0.77%	0.78%	0.75%
1,000,000	0.26%	1.55%	1.56%	1.56%
1,500,000	0.49%	2.32%	2.32%	2.31%
2,000,000	0.82%	3.06%	3.07%	3.08%

of names, hence the mapping of LCI-PIT is more uniform and the false positive probability is lower. On the contrary, there are a lot of conflicts in Hash Table-PIT due to the uneven mapping of hash funtion.

3.4 Throughput

The processing speed of packet affects the performance of the forwarding plane of NDN to a large extent. For this reason, the throughput of LCI-PIT, BPT-PIT and Hash Table-PIT is evaluated for different number of names. The experimental results are shown in Table 5, since LCI-PIT and Binary Patricia Trie-PIT can be deployed on SRAM for fast packet processing, the throughput of them is higher than that of Hash Table-PIT.

For LCI-PIT, its throughput can reach up to 290.36 million searches per second (MSPS). The reason is that a fast binary search algorithm is used to search the category of the name and the data mapping uses a BPNNs with simple structure. Although Binary Patricia Trie-PIT can also be deployed in SRAM, the name lookup process needs to traverse in a tree structure, and the larger depth affects the lookup speed. Since the memory consumption of Hash Table-PIT does not allow it to be deployed on SRAM, it can only be deployed on DRAM with larger capacity but lower speed. Therefore, the throughput of them is much lower than that of LCI-PIT and Binary Patricia Trie-PIT, which is difficult to satisfy current network requirements for packet processing speed (Table 3).

Table 3. Throughput of LCI-PIT

# of Names	LCI-PIT (MSPS)	BPT-PIT (MSPS)	MD5-PIT (MSPS)	CityHash256-PIT (MSPS)	xxHash-PIT (MSPS)
500,000	277.20	42.80	2.34	4.21	16.38
1,000,000	287.67	37.31	2.40	4.65	19.37
1,500,000	290.36	34.91	2.43	4.80	20.72
2,000,000	287.86	33.23	2.47	4.84	21.54

4 Conclusion

This paper proposes a learned classification index for PIT called LCI-PIT. LCI-PIT divides the huge namespace into lots of subnamespace, and utilizes neural networks to learn the distribution of names, which achieves a more uniform mapping and efficient storage and lookup. The experimental results show that under the condition that the false positive probability is less than 1%, the memory consumption of LCI-PIT is 1.08 MB for millions of names, and the throughput can reach up to 290.36 MSPS. Therefore, LCI-PIT can well meet the requirements of the current network. In the future work, the performance of LCI-PIT will be tested in the actual communication environment. Meanwhile, its deployment scheme will be studied.

Acknowledgments. This work was supported by the Key Research and Development Projects of Hebei Province under Grant 20314301D and the Science and Technology Project of SGCC under Grant kj2022-023 and the Tianjin Science and Technology Plan Project under Grant 20JCQNJC01490.

References

1. Zhuo, L., Yaping, X., Baichuan, Z., Liu, Y., Kaihua, L: Packet forwarding in named data networking requirements and survey of solutions. IEEE Commun. Surv. Tutor. **21**, 1950–1987 (2019)
2. Zhuo, L., Jindian, L., Liu, Y., Baichuan, Z., Peng, L., Kaihua, L.: Smart name lookup for NDN forwarding plane via neural networks. IEEE/ACM Trans. Networking **30**, 529–541 (2022)
3. Lixia, Z., Afanasyev, A., Burke, J.: Named data networking. ACM SIGCOMM Comput. Commun. Rev. **44**, 66–73 (2014)
4. Huichen, D., Bin, L.: CONSERT: constructing optimal name-based routing tables. Comput. Netw. **94**, 62–79 (2016)
5. Tian, S., Haowei, Y., Crowley, P: Scalable name-based packet forwarding. In: Proceedings of the 2nd ACM Conference on Information-Centric Networking, pp. 19–28. ACM, San Franciscot (2015)
6. Haowei, Y.: Data structures and algorithms for scalable NDN forwarding. School of Engineering & Applied Science, Washington University in St. Louis, St Louis (2015)
7. Zhuo, L., Kaihua, L., Yang, Z., Yongtao, Ma.: MaPIT: an enhanced pending interest table for NDN with mapping bloom filter. IEEE Commun. Lett. **18**, 1915-1918 (2014)
8. Varvello, M., Perino, D., Linguaglossa, L.:On the design and implementation of a wire-speed pending interest table. In 2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 369-374. IEEE, Turin (2013)
9. Liu, Y., Zhuo, L,. Kaihua, Liu.: Learning tree: neural network-based index for NDN forwarding plane. In: ACM SIGCOMM 2019 Conference Posters and Demos, pp. 63-65. ACM, Beijing (2019)
10. Zhuo, L., Yutong, C., Heping, S., Kaihua, L.: NDN-GSM-R: a novel high-speed railway communication system via named data networking. EURASIP J. Wirel. Commun. Netw. **48**, 1–5 (2016)

Router Forwarding Congestion Control Scheme in Information Center Networking

Xin Duan[1(✉)], Zizi Zhang[1], Hongliang Shen[2], Ersong Chen[1], and Xingdi Shen[3]

[1] State Grid Hebei Electric Power Research Institute, Shijiazhuang 050021, China
123100784@qq.com
[2] State Grid Hebei Electric Power Corporation, Shijiazhuang 050021, China
[3] CoopNet Co., Ltd. at Shanxi, Shanxi 037000, China

Abstract. In order to solve the problem that the original path forwarding traffic caused by congestion that cannot be recovered in time in Information Center Networking (ICN), a congestion control scheme RFCC (Router Forwarding Congestion Control) based on path traffic recovery is proposed. This scheme combines receiver rate adjustment and router forwarding and diversion into a hybrid congestion control scheme. Active queue management technology is used to detect congestion and notify downflow nodes (consumers and routers) in time, which trigger downflow consumers to adjust the Interest sending window and intermediate routers to transfer traffic. In addition, the original path traffic is restored in time through the path traffic recovery policy. This paper implements the scheme in ndnSIM, and extensive experiments show that the proposed scheme achieves good network performance, which can not only obtain high and stable throughput, but also obtain shorter and more stable transmission delay. In addition, the optimal path transmission volume and link utilization are improved.

Keywords: Information Center Networking · Path Traffic Recovery Strategy · Active Queue Management · Hybrid Congestion Control · Future Network Architecture

1 Introduction

Information Center Networking (ICN) [1] is an emerging network architecture, changing the traditional network communication model. Information access in ICN is driven by the receiver, which requests data through the content name, and the routing node forwards the data according to the content name. This new model supports multicast data delivery, intra-network, multi-source caching, and multi-path forwarding [2, 9, 10]. However, this new network architecture also complicates network congestion control because existing traditional congestion control solutions cannot be directly applied[11, 12, 13]. In ICN, the concept of end-to-end connection is not applicable, because data blocks of the same content can be obtained from different producers or different node caches on the path to these producers. These different content sources will lead to different retrieval delays, and the requester cannot distinguish them. Therefore, the traditional detection based on RTT timeout has become an unreliable congestion detection index [3].

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 245–251, 2023.
https://doi.org/10.1007/978-981-99-2653-4_30

This paper proposes a hybrid congestion control scheme RFCC based on Active Queue Management (AQM), which can solve the problem of path recovery after congestion is relieved through ICN intermediate node forwarding. The scheme detects congestion and generates marking information by deploying the improved AQM scheme PIE [4] in the intermediate routers, and then adds the marking information in the packets (Data) and forwards them to the consumer and downflow routers. Through adjusting interest sending rate and diverts traffic to other alternative paths to control congestion. Moreover, for the transmission recovery problem of subsequent congested paths, this paper designs a double recovery mechanism. When the congested path is detected to work normally, all traffic is transferred from the sub-optimal link back to the original path in time, and the transmission continues on the optimal path.

2 Router Forwarding Congestion Control Scheme

2.1 Congestion Control Strategy

The hybrid congestion control strategy includes two control methods: mild congestion control and heavy congestion control to adapt to different congestion levels. In mild congestion control, when the shunt node receives the mild congestion mark "01", it relieves upflow congestion by diverting traffic [5]. At the same time, when the consumer side receives the mild congestion mark "01", it will multiply the window by a smaller coefficient that can reduce the size of the sending window. However, if the shunt node cannot handle the current congestion through forwarding control, that is, there is no available interface or the data queue exceeds 90%. The shunt node will update the mark of a Data packet to the heavy congestion mark "11", and then enter the heavy congestion control. For the heavy congestion mark "11", the shunt node and the consumer will take different control actions. On the one hand, the shunt node will stop the shunting and hand over all control to the consumer side. On the other hand, in order to alleviate network congestion, the consumer multiplies the window size by a larger coefficient to multiply the window size.

2.2 Path Traffic Recovery Strategy

The path traffic recovery strategy judges whether to restore the optimal path traffic according to the congestion mark information. When a congested interface is disabled, the congested interface can resume normal operation only when the interface receives the normal mark "00" again (indicating upflow congestion relief). However, if the congested interface continues to receive the Congestion Mark "01" or the Invalid mark "10", it indicates that the congestion status of the optimal path is not relieved at this time, and the interface will continue to be disabled.

After the shunt node receives the first Congestion Mark from an interface (the current optimal interface), the interface will be temporarily disabled, and it will not continue to forward Interest requests from this interface. However, it will forward traffic to other standby interfaces. Usually, when a congested interface is temporarily disabled, the interface still receives some data requested before the disablement, which is called subsequent

data packets. Therefore, the scheme in this paper firstly judges whether to re-enable the congested interface according to the marking information carried by the subsequent data packets. When an interface receives Congestion Mark "01", the shunt node will start a timer of Timer_1 for this interface. After that, whenever a new data arrives, the shunt node will use the sample RTT to update the moving weighted average round-trip delay of RTT_{average}.

$$RTT_{average} = \omega RTT + (1 - \omega) RTT_{last} \tag{1}$$

RTT_{last} is defined as the latest moving average round-trip delay, and RTT is the sample RTT of the current subsequent data packets. ω is a weight factor used for adjustment. Whenever the congested interface receives a subsequent data packet, it will update the value of the RTT_{average}, and then record the value of the timer of the Timer_1 at this time and compare it with RTT_{last}. If $RTT_{\text{last}} \geq Timer\_1$, it means that the interface received subsequent packets within the latest RTT time, then reset Timer_1 to 0. If RTT_{last}. If $RTT_{\text{last}} < Timer\_1$, it means that the congested interface has not received subsequent data packets within the latest RTT. It is considered that all Interest requests sent before the congested interface is disabled have been responded to, because the interface is temporarily disabled and no subsequent data packets will be returned after that.

Figure 1 is an example of ICN recovering a congested path. For case A, the current optimal transmission path is R1-R2. After receiving Congestion Mark "01" from R2, R1 adjusts the forwarding interface to R3, and the transmission path becomes R1-R3. The congested interface connected to R3 starts Timer_1. If R1 receives a subsequent packet carrying "00" from R2, it means that the R2 congestion has disappeared, so the path R1-R2 is re-enabled, and R1 turns off Timer_1 (case A to case B in Fig. 3). If R1 receives a data packet carrying "01" or "10" from R2, it indicates that the congestion of R2 has not been relieved at this time, and it needs to continue to be disabled, such as case C. In case D of Fig. 3. Since subsequent packets from R2 are exhausted and the interface is temporarily disabled without forwarding new Interest requests, R1 will no longer receive packets from R2 after that. Considering this situation, R1 will start another

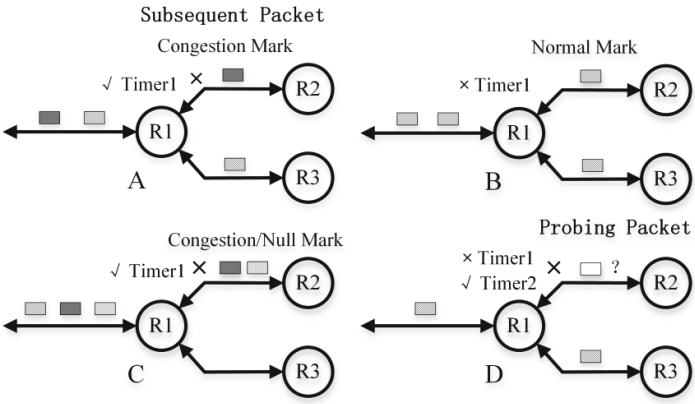


Fig. 1. Example of Path Traffic Recovery

timer Timer_2, and periodically and actively send probing Interest to R2 according to the time of Timer_2 to obtain the congestion status of the congested path.

In the above example, it can be seen that it is not enough to recover traffic only by subsequent data packets in the case of severe congestion. Therefore, actively sending probing Interest requests to the congested interface is a necessary design for traffic recovery of congested paths. In this solution, when the subsequent data packets are exhausted, enough probe Interest requests are actively sent to help the congested path recover traffic without increasing the congestion status of the congested path. In addition, in reference [6], the authors show that Interests occupy a non-negligible part of the link bandwidth, prompting this paper to consider the influence of Interest flow when sending probing Interests. In order to make the probing Interest request not aggravate the congestion state, the remaining available bandwidth of the congested link needs to be evenly distributed to each Interest flow and Data flow.

3 Experiment and Performance Evaluation

This paper uses the NDN open source simulation platform ndnSIM [7, 8] based on NS-3 to implement the proposed congestion control mechanism, and compares it with other schemes.

3.1 Experimental Setup

This paper uses ndnSIM version 2.1 to implement the proposed ICN congestion control methods RFCC based on path recovery strategy, and uses representative network typologies to evaluate and compare the two schemes. The following four indicators are used, which are network throughput, packet loss rate, queue length, and transmission delay time.

In this paper, the multi-source topology shown in Fig. 3. All four consumers access the network through high-speed access links (100Mbps). Two consumers C1 and C2 share the same bottleneck link R1-R2. Both C1 (starting at 0s) and C2 (starting at 1s) go to the network. Producer P1 requests 50M content with a prefix of /Prefix/A. The specific link bandwidth and delay settings are shown in the figure. The other two consumers, C3 and C4, access the network through R2. C3 (starting at 5s and stopping at 6s) and C4 (starting at 8s and stopping at 9th) send the prefix /Prefix/B to P1 with ConsumerCbr at the specified time. The frequency is 5000 Interest packets per second.

In addition, the size of the Interest packet is set to 40 Bytes, the size of the Data packet is set to 1024Bytes, and the queue length capacity of each router node in the topology is set to 200Pkts.

3.2 Analysis of Simulation Results

In this paper, In the multi-source topology in Fig. 2, C1 starts to request data from P1 at 0s, then another user C2 is opened 1s later, and all consumers first request data from P2. Under this topology, this paper compares and analyzes the performance indicators of RFCC, such as throughput and average delay.

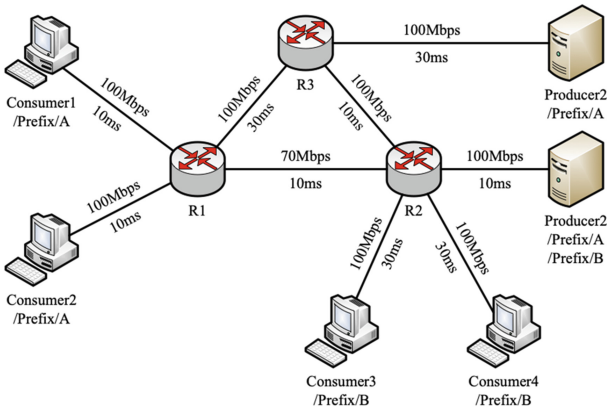


Fig. 2. Multi-Source Network Topology

Figure 3 compares the C1, C2 and overall throughput of RFCC under the multi-source topology. It can be seen from the results that the overall throughput of RFCC is high (7867.8 KB/s and standard deviation of 2256.9). Due to the larger RTO time of RFCC, multiple invalid retransmissions are reduced, and more accurate congestion detection and adaptive path recovery control strategy are used. Therefore, the network utilization is higher and the overall throughput performance is also better. The RFCC uses adaptive node forwarding and the receiver. The hybrid congestion scheme of window adjustment, so the overall control effect is better. Even if the constant burst volume is increased in the 5s and 8s, good control can be achieved. Due to the path recovery strategy, it can restore traffic in a very short time and ensure network performance.

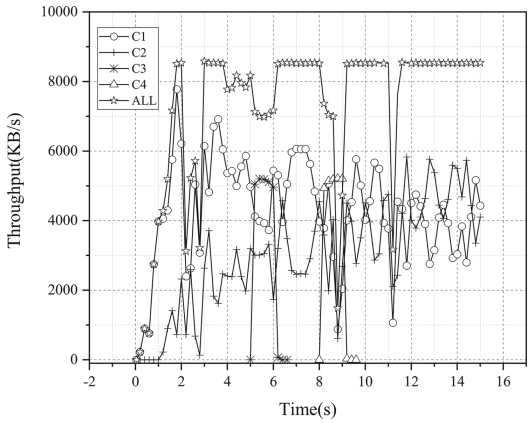


Fig. 3. Throughput of RFCC

Figure 4 compares the average delay and standard deviation of C1 and C2 in the multi-source topology. It can be seen from the results that both RFCC C1 and C2 have

lower transmission delay (68.69 ms and 67.08 ms). The transmission delay of each consumer (17.03 and 14.04) is also more stable. The average delay of RFCC is 67.8 ms.

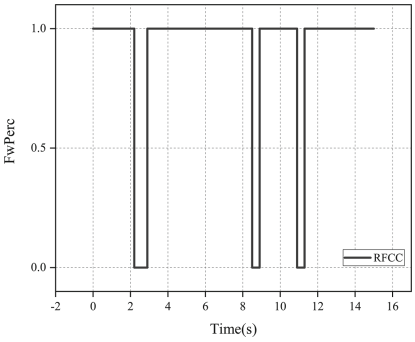


Fig. 4. Comparison of Average Delay and Standard Deviation

Figures 5 and 6 compare the forwarding ratio of the interest in R1's every interface (258, 259). Since RFCC has router forwarding control, congestion is detected at about 2s, 8s and 11s, then the traffic is transferred from 258 interface to 259 interface, and then the path recovery strategy is implemented. The congestion in 8s and 11s quickly recovers the traffic after the congestion is relieved according to the subsequent packet recovery mechanism, while the congestion in 2s sends a probing interest packet, so the time to recover the traffic is slightly longer.

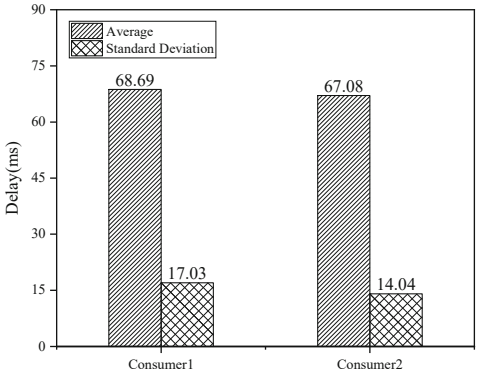


Fig. 5. Forwarding ratio of R1 node 258 interface

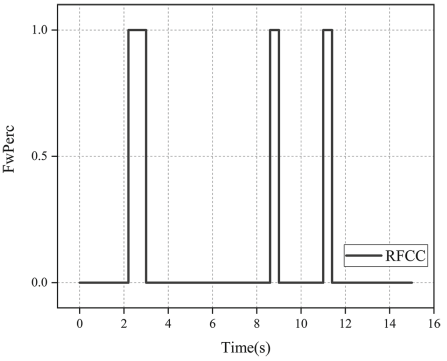


Fig. 6. Forwarding ratio of R1 node 259 interface

4 Conclusion

This paper proposes a congestion control solution RFCC based on path traffic recovery for the forwarding of nodes in the information center networking that cannot restore the

original path transfer traffic caused by congestion in time. RFCC combines consumer rate adjustment and intermediate node forwarding and offloading into a hybrid congestion control scheme, which can not only coordinate the receiver and the router to dynamically relieve congestion, but also restore the optimal path traffic in time to ensure its transmission performance. The scheme in this paper is implemented in ndnSIM and tested under a representative multi-source forwarding network topology. Extensive experiments show that the throughput of RFCC is better than other comparison schemes, and more stable transmission delay and throughput can be obtained. In addition, the optimal path transmission volume and link utilization are improved.

Acknowledgements. This work is supported by the Key Research and Development Projects of Hebei Province under Grant 20314301D, the Science and Technology Project of SGCC No.kj2022–023, Tianjin Science and technology planning project under Grant No. 20JCQNJC01490.

References

1. Zhang, L., Estrin, D., Burke, J., et al.: Named Data Networking (NDN) Project (2012)
2. Schneider, K., et al.: A practical congestion control scheme for named data net-working. ACM Conference on Information-centric Networking. ACM (2016)
3. Ren, Y., et al.: Congestion control in named data networking – a survey. Comput. Commun. **86**(15), 1–11 (2016)
4. Pan, R., Natarajan, P., Baker, F., et al.: Proportional integral controller enhanced (pie): a lightweight control scheme to address the bufferbloat problem (2017)
5. Ren, Y., et al.: An explicit congestion control algorithm for Named Data Networking. Computer Communications Workshops IEEE (2016)
6. Wang, Y., et al.: An improved hop-by-hop interest shaper for congestion control in named data networking. ACM SIGCOMM Comput. Commun. Rev. **43**, 4 (2013)
7. Afanasyev, A., Moiseenko, I., Zhang, L.: ndnSIM: NDN simulator for NS-3 (2012)
8. Mastorakis, S., Afanasyev, A., Moiseenko, I., Zhang, L.: ndnsim 2.0: a new ver-sion of the ndn simulator for ns-3. NDN, Technical Report NDN-0028 (2015)
9. Li, Z., Xu, Y., Zhang, B., et al.: Packet forwarding in named data networking requirements and survey of solutions. IEEE Commun. Surv. Tutor. **PP**, 1 (2018)
10. Li, Z., Liu, J., Yan, L., et al.: Smart Name lookup for NDN forwarding plane via neural networks. IEEE/ACM Trans. Netw. **29**, 1–13 (2021)
11. Yan, L., Li, Z., Liu, K.: Learning tree: neural network-based index for NDN forwarding plane. In: The ACM SIGCOMM 2019 Conference Posters and Demos. ACM (2019)
12. Li, Z., Liu, K., Zhao, Y., et al.: MaPIT: an enhanced pending interest table for NDN with mapping bloom filter. IEEE Commun. Lett. **18**(11), 1915–1918 (2014)
13. Li, Z., Chen, Y., Shi, H., et al. NDN-GSM-R: a novel high-speed railway communication system via named data networking. EURASIP J. Wirel. Commun. Netw. **2016**(48), 1–5 (2016)

Memory Bank for Unsupervised Domain Adaptation Person Retrieval

Di He[1], Yu Lian[1], Peng Xia[2], Suoku Wei[3], and Zhong Zhang[1](✉)

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
zhong.zhang8848@gmail.com
[2] Tianjin Dolphin AI Technology Co., Ltd., Tianjin, China
[3] Tianjin Shengna Science and Technology Co., Ltd., Tianjin, China

Abstract. With the development and progress of deep learning, unsupervised domain adaptation (UDA) person retrieval task becomes more and more popular. But there is a huge challenge in pseudo-label UDA person retrieval method, and therefore it is necessary to maintain consistency between similar features in metric learning. In this paper, we introduce the memory bank based methods for UDA person retrieval, where memory bank can store a large number of pedestrian features and maintain feature consistency by dynamically updating. Meanwhile, we compare the performance of memory bank based UDA person retrieval methods on benchmark datasets to demonstrate the effectiveness of memory bank.

Keywords: person retrieval · unsupervised domain adaptation · memory bank

1 Introduction

Person retrieval aims to search the same pedestrian from cross camera, which is also known as person re-identification (Re-ID), and it is a popular task in computer vision [1–4]. Recently, person retrieval on annotated datasets has achieved impressive performance, but the performance degrades significantly when a person retrieval model trained on annotated datasets is used to test an unknown dataset. Furthermore, there are currently limited pedestrian datasets with annotations and manual annotation of pedestrian datasets is expensive. Hence, many researchers begin to concern about and explore unsupervised domain adaptation (UDA) person retrieval [5–8], which is closer to practical applications.

In particular, memory bank has recently become popular in UDA person retrieval task, where it is mainly used to help access pedestrian features so as to better learn pseudo labels and efficiently calculate the loss [9–12]. More importantly, pseudo-label-based UDA person retrieval methods achieve state-of-the-art performance thanks to the introduction of memory bank [10,11].

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 252–258, 2023.
https://doi.org/10.1007/978-981-99-2653-4_31

In this paper, we review and introduce some existing memory bank based UDA person retrieval methods, and these methods are divided into two categories according to the update manner of the features in the memory bank: Instance-based Memory Bank Methods and Queue-based Memory Bank Methods. Moreover, there is another category: Hybrid Memory Bank Methods. Subsequently, we compare the performance of memory bank based methods on benchmark datasets to demonstrate the effectiveness of memory bank.

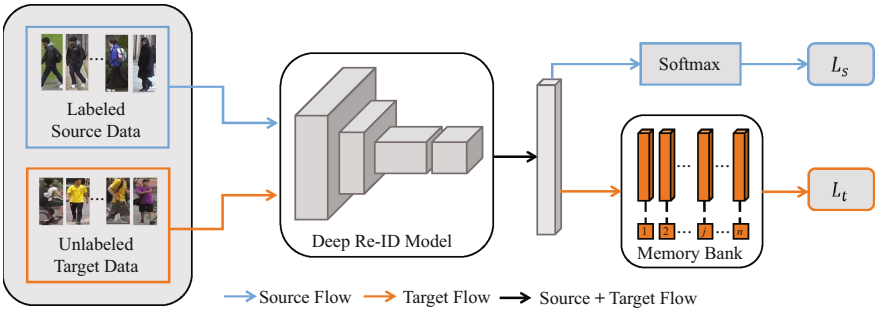


Fig. 1. The network framework of memory bank based UDA person retrieval methods, in which memory bank is used to store the up-to-date features of target data and help to compute the loss of unlabeled target data.

2 Existing Methods

The UDA person retrieval model is generally trained by the labeled source domain and the unlabeled target domain, and the general pipeline of memory bank based method for UDA person retrieval is shown in Fig. 1. With the help of memory bank, researchers have made great progress in UDA person retrieval task. Here, we introduce three kinds of UDA person retrieval methods with memory bank, i.e., Instance-based Memory Bank Methods, Queue-based Memory Bank Methods and Hybrid Memory Bank Methods.

2.1 Instance-Based Memory Bank Methods

Chen *et al.* [13] regard the person retrieval task as a classification task and a rank task, and metric learning plays an important role in the rank issue. However, there is still a challenging problem in metric learning. Since a large number of negative samples in metric learning can enhance the discrimination of feature representation learning and the batch size of model training is limited, the problem is how to sample a large number of negative samples for feature learning.

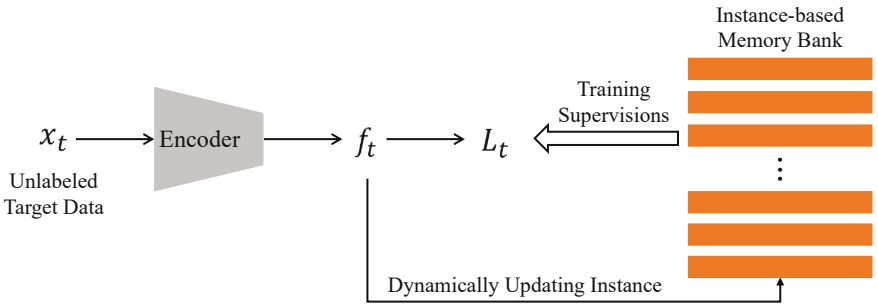


Fig. 2. The flowchart of the instance-based memory bank. The orange rectangular bar represents the target feature.

Furthermore, Wu *et al.* [14] propose the concept of memory bank and describe a method based on memory bank to sample negative samples for contrastive learning, where they treat each sample as a separate class, that is instance level classification task, so as to efficiently compute non-parametric softmax loss.

Inspired by [14], in the field of UDA person retrieval, Zhong *et al.* [9] construct an exemplar memory module for storing the up-to-date features of all target images, where they treat each target sample as an individual identity and perform features update in the form of key-value pairs so as to implement three types of invariance learning. The flowchart of the instance-based memory bank is shown in Fig. 2, where the form of dynamically update instance is as follows:

$$M^i \leftarrow \alpha_t M^i + (1 - \alpha_t) f_t^i, \tag{1}$$

where M^i is the feature of the i-th slot in the memory bank, and α_t is a parameter that controls the update scale.

2.2 Queue-Based Memory Bank Methods

In order to keep key encoder and query encoder in a consistent state, He *et al.* [15] point out that the unsupervised learning of computer vision needs to establish a dictionary, and propose that the dictionary depends on two necessary conditions: 1) The size of the dictionary is large enough; 2) The features in the dictionary need to be consistent. Meanwhile, they use a queue to achieve the first condition and an encoder updated by momentum to implement the second condition. The update rule of the momentum encoder is as follows:

$$\theta_k \leftarrow m\theta_k + (1 - m)\theta_q, \tag{2}$$

where θ_q and θ_k indicate the network parameters of two different encoders, and m is a parameter that controls the update scale.

Inspired by [14], Zheng *et al.* [10,11] design the mean teacher network [16] with queue-based memory bank for UDA person retrieval, where they combine

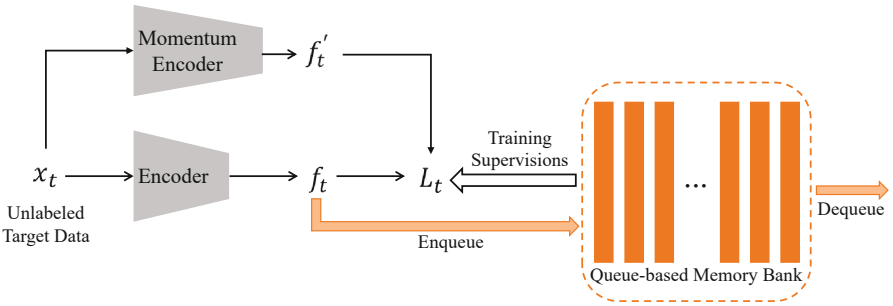


Fig. 3. The flowchart of the queue-based memory bank. The orange rectangular bar represents the target feature.

the contrastive loss with the target instance memory bank in order to jointly optimize positive and negative pairs for a query/anchor sample in the memory bank. The flowchart of the queue-based memory bank is shown in Fig. 3.

2.3 Hybrid Memory Bank Methods

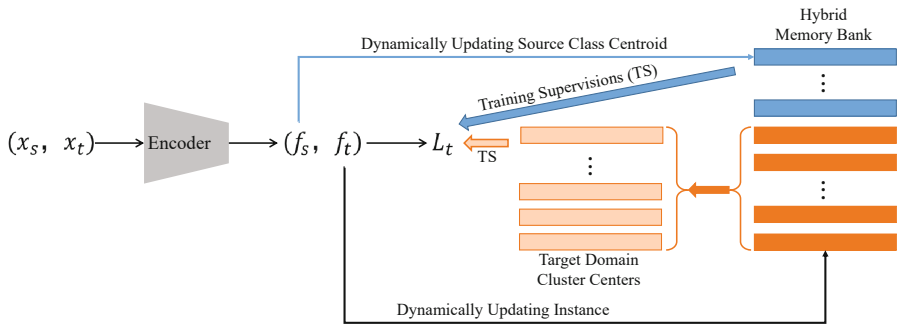


Fig. 4. The flowchart of the hybrid memory bank. The rectangular bar represents the feature.

Recently, Ge *et al.* [12] propose a novel self-paced contrastive learning framework with the hybrid memory bank for UDA person retrieval, which includes source domain class centroids and target domain instance features. The flowchart of the hybrid memory bank is shown in Fig. 4, where the source domain class centroids are added to enhance features learning, so that improving the domain adaptability of the model. The update criteria for the source domain class centroids are as follows:

$$w^k \leftarrow \alpha_s w^k + (1-\alpha_s)\frac{1}{|B_k|}\sum_{f_s^i \in B_k} f_s^i, \tag{3}$$

where B_k indicates the features belonging to source domain class k under the current batch, and α_s is a parameter that controls the update scale. Furthermore, Zheng *et al.* [10,11] also add source domain class centroids in memory bank, and achieve the state-of-the-art results on benchmark datasets.

3 Experiments

Table 1. The evaluation results (%) of UDA person retrieval method on Market→Duke and Duke→Market. Here, MB and HMB indicate memory bank and hybrid memory bank, respectively.

Methods	Market→Duke				Duke→Market			
	mAP	R1	R5	R10	mAP	R1	R5	R10
Baseline	53.9	71.8	81.9	85.5	64.7	83.0	93.0	95.1
Baseline+MB	57.6	74.8	84.8	88.2	68.6	85.3	93.9	96.2
Baseline+HMB	59.5	76.0	85.3	89.3	69.3	86.0	94.3	96.7

We make some evaluations on two benchmark pedestrian datasets, i.e., Market-1501 (Market) [17], DukeMTMC-reID (Duke) [18]. Market consists of 32668 annotated images of 1501 identities captured by 6 cameras, where it contains 12936 training images of 751 identities and 19732 test images of 750 identities. Duke has 36411 annotated images taken by 8 camera, in which it contains 16522 training images of 702 identities and 19889 test images of another 702 identities.

We adopt ResNet50 [19] as the backbone of the deep person retrieval model where it is initialized by the pre-trained parameters of ImageNet [20], and the target ID loss and the triplet loss as our objective optimization function.

In the stage of training, all pedestrian images are resized to 256×128 before feeding into the networks. The training images are augmented by random horizontal flipping, random erasing, random cropping and padding [21]. The batch size is set to 64 which includes 16 target domain identities, and each identity contains 4 images. Note that the target domain samples are utilized to optimize the whole network, while the source domain samples are applied to construct the source domain class centroids in hybrid memory bank. The network is optimized by the Adam optimizer [22] with the weight decay of 0.0005. The number of epochs is set to 80, and the learning rate is fixed at 0.00035. All update parameters are empirically set as 0.2. In the clustering stage, we adopt DBSCAN to cluster and obtain the pseudo labels, where the $eps = 0.6$ for all datasets about DBSCAN settings. In the testing stage, the output of the model is treated as the representation of pedestrian image to calculate the cosine similarity.

Because our baseline adopts the structure shown in Fig. 2, we only evaluate the instance-based memory bank UDA person retrieval method and the hybrid memory bank UDA person retrieval method, as shown in Table 1, where the update rule of memory bank is Eq. 1. From the table, we can draw two conclusions: 1) By adding the memory bank, the performance of $Baseline + MB$ is higher than $Baseline$, which proves the effectiveness of memory bank. 2) By introducing the hybrid memory bank, the performance of $Baseline + HMB$ is higher than $Baseline + MB$, which proves that introducing the source domain class centroids can enhance domain adaptability of the model.

4 Conclusion

In this paper, we have introduced representative memory bank based UDA person retrieval methods, including instance-based memory bank, queue-based memory bank and hybrid memory bank methods. Moreover, we have demonstrated the effectiveness of memory bank in UDA person retrieval methods by evaluation experiments. Hence, it is necessary to consider memory bank in the pseudo-label based UDA person retrieval methods, which can enhance the discrimination of pedestrian feature representation, so as to improve the robustness and domain adaptability of the model.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant No. 62171321, 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 the Key Project of Graduate Scientific Research Innovation Project of Tianjin Normal University under Grant No. 2022KYCX032Z.

References

1. Liao, S., Hu, Y., Zhu, X., Li, S.Z.: Person re-identification by local maximal occurrence representation and metric learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2197–2206 (2015)
2. Ge, Y., Chen, D., Li, H.: Mutual mean-teaching: pseudo label refinery for unsupervised domain adaptation on person re-identification. In: Proceedings of the International Conference on Learning Representations (2020)
3. Ye, M., Shen, J., Lin, G., Xiang, T., Shao, L., Hoi, S.C.: Deep learning or person re-identification: a survey and outlook. IEEE Trans. Pattern Anal. Mach. Intell. **44**(6), 2872–2893 (2022)
4. Zhang, Z., Wang, Y., Liu, S., Xiao, B., Durrani, T.S.: Cross-domain person reidentification using heterogeneous convolutional network. IEEE Trans. Circuits Syst. Video Technol. **32**(3), 1160–1171 (2022)
5. Fan, H., Zheng, L., Yan, C., Yang, Y.: Unsupervised person reidentification: Clustering and fine-tuning. ACM Trans. Multimed. Comput. Commun. Appl. **14**(4), 1–18 (2018)

6. Li, Y.J., Lin, C.S., Lin, Y.B., Wang, Y.C.F.: Cross-dataset person re-identification via unsupervised pose disentanglement and adaptation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 7919–7929 (2019)
7. Huang, Y., Wu, Q., Xu, J., Zhong, Y., Zhang, Z.: Unsupervised domain adaptation with background shift mitigating for person re-identification. Int. J. Comput. Vision **129**(7), 2244–2263 (2021)
8. Zheng, Y., et al.: Online pseudo label generation by hierarchical cluster dynamics for adaptive person re-identification. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 8371–8381 (2021)
9. Zhong, Z., Zheng, L., Luo, Z., Li, S., Yang, Y.: Invariance matters: exemplar memory for domain adaptive person re-identification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 598–607 (2019)
10. Zheng, K., Lan, C., Zeng, W., Zhang, Z., Zha, Z.: Exploiting sample uncertainty for domain adaptive person re-identification. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 3538–3546 (2021)
11. Zheng, K., Liu, W., He, L., Mei, T., Luo, J., Zha, Z.: Group-aware label transfer for domain adaptive person re-identification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5310–5319 (2021)
12. Ge, Y., Zhu, F., Chen, D., Zhao, R., Li, H.: Self-paced contrastive learning with hybrid memory for domain adaptive object Re-ID. In: Proceedings of the Advances in Neural Information Processing Systems, pp. 11309–11321 (2020)
13. Chen, W.H., Chen, X.T., Zhang, J.G., Huang, K.Q.: A multi-task deep network for person re-identification. In: Proceedings of the AAAI Conference on Artificial Intelligence (2017)
14. Wu, Z.R., Xiong, Y.J., Yu, S.X., Lin, D.H.: Unsupervised feature learning via non-parametric instance discrimination. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3733–3742 (2018)
15. He, K., Fan, H., Wu, Y., Xie, S., Girshick, R.: Momentum contrast for unsupervised visual representation learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9729–9738 (2020)
16. Tarvainen, A., Valpola, H.: Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In: Advances in Neural Information Processing Systems, pp. 1195–1204 (2017)
17. Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., Tian, Q.: Scalable person re-identification: a benchmark. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1116–1124 (2015)
18. Ristani, E., Solera, F., Zou, R.S., Cucchiara, R., Tomasi, C.: Performance measures and a data set for multi-target, multi-camera tracking. In: Proceedings of the European Conference on Computer Vision, pp. 17–35 (2016)
19. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
20. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: a large-scale hierarchical image database. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255 (2009)
21. Zhong, Z., Zheng, L., Kang, G., Li, S., Yang, Y.: Random erasing data augmentation. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 13001–13008 (2020)
22. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

Sink Trajectory Planning for Data Collection in Mobile Wireless Sensor Network

Xiaoqing Tian(✉), Siwei Li, and Jing Liang

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China
xiaoqing_tian@163.com

Abstract. In mobile wireless sensor networks, the efficient collection of sensors' data is the key to ensuring the quality of space monitoring. In this paper, considering the trajectory length and energy consumption of network data collection, we propose the virtual force and improved LEACH (Low Energy Adaptive Clustering Hierarchy) based algorithm (V-ILEACH) to realize the trajectory planning of sink. Compared with the algorithm based on traveling salesman problem (TSP), V-ILEACH achieves significant reduction of trajectory length on the basis of ensuring complete data collection.

Keywords: ILEACH · Virtual force · Trajectory planning · MWSN · Data collection

1 Introduction

Mobile Wireless Sensor Network (MWSN) consists of a series of movable sensor nodes. Due to the characteristics of scalable network scale, low node energy consumption, and flexible mobility [1], MWSN has been widely used in intelligent transportation, disaster warning, abnormal monitoring, and other fields. In the process of MWSN task execution, a reasonable node movement trajectory can shorten the task execution time and avoid the threat sources, which is the key to ensuring the quality of the task completion.

In recent years, the focus of research is to introduce a mobile sink in MWSN [2,3] to move along the planned trajectory to collect data to avoid excessive energy consumption of relay nodes due to the transmission of large amounts of data. In the traditional TSP-based method [3], the sink needs to visit each node. When the network scale increases, the trajectory length will increase sharply. It is worth mentioning that as long as the sink is within the communication radius of the node, the node can send information to the sink remotely [4], so the sink does not have to traverse all nodes. In [5], the traveling salesman problem with neighborhoods (TSPN) is proposed, which significantly reduces the sink's trajectory length and resource overhead. Some scholars use clustering to

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 259–266, 2023.
https://doi.org/10.1007/978-981-99-2653-4_32

complete data collection. The clustering algorithm clusters the nodes in MWSN. Each cluster selects one cluster head, and the nodes in the cluster send the data to the cluster head [6]. The sink traverses each cluster head to collect data. The fewer clusters, the fewer places the sink needs to stay, and the shorter its trajectory. However, the corresponding number of nodes in the cluster increases, and the cluster heads will quickly run out of energy due to too much data transmission [7]. Therefore, it is essential to design an appropriate sink trajectory under the premise of balancing the trajectory length and the energy consumption of the nodes. As a clustering algorithm considering the energy consumption, LEACH [8] is widely used in wireless sensor networks. Most of the existing algorithms are based on the ideal assumption that the search environment is free of obstacles and that movement of nodes is not restricted. However, there are various threat sources in the actual environment, so it is urgent to design the trajectory planning of restricted nodes in complex environments.

This work proposes a virtual force [9] and improved LEACH based trajectory planning algorithm (V-ILEACH) in three-dimensional space. The contributions of this work are presented as follows:

1) Considering the nodes' energy and environmental threat, an improved LEACH is proposed to avoid the nodes in high-threat regions becoming cluster heads.
2) After the clustering is completed, the virtual force is used to comprehensively consider the energy of cluster head, the number of packets in the cluster and the distance between cluster heads to determine the sink's stay positions.

2 Model Construction

Assuming that the size of the monitoring space is $a \times b \times c$, there are threat sources such as radar and anti-aircraft artillery in this space. Nodes are randomly distributed, periodically perceive information, and their positions remain unchanged after deployment. The sink collects the data of all nodes once is called a "round".

2.1 Threat Source Model

As a common threat source, the threat degree of enemy radar is defined as:

$$TD_r = \begin{cases} 1 - \left(\frac{\|(x,y,z) - (x_{r_0}, y_{r_0}, z_{r_0})\|}{d_{r_{\max}}} \right)^4, & \|(x,y,z) - (x_{r_0}, y_{r_0}, z_{r_0})\| \le d_{r_{\max}} \\ 0, & else \end{cases} \tag{1}$$

where (x, y, z) represents the coordinates of one point in this space, $(x_{r_0}, y_{r_0}, z_{r_0})$ is the coordinates of enemy radar, $d_{r_{\max}}$ is the longest distance that radar can detect, and $\|\bullet\|$ is the Euclidean distance of two points.

The threat degree of anti-aircraft artillery is defined as follows:

$$TD_g = \exp\left(-\frac{\|(x,y,z)-(x_{g_0},y_{g_0},z_{g_0})\|^2}{d_{g_{\max}}^2/9}\right) \tag{2}$$

where $(x_{g_0}, y_{g_0}, z_{g_0})$ is the coordinates of antiaircraft artillery and $d_{g_{\max}}$ is the longest range of artillery attackable.

The threat degree of node caused by flight altitude is:

$$TD_h = \begin{cases} v_1, h_{\min} \le h \\ v_2, h_{\min} < h < h_{\max} \\ v_3, h_{\max} \le h \end{cases} \tag{3}$$

$h_{\min}$ and $h_{\max}$ represent the lowest and highest heights, respectively.

For the two independent threat sources with threat degrees P_1 and P_2, when the threat source belongs to the same party and the enemy, the total threat degree are:

$$TD_{\texttt{total}} = TD_1 \oplus TD_2 = 1-(1-P_1)(1-P_2) \tag{4}$$

$$TD_{\texttt{total}} = TD_1 \odot TD_2 = P_1(1-P_2) \tag{5}$$

2.2 Energy Consumption Model

Energy consumption of node mainly comes from sending, receiving data, and overcome the threat source. The energy consumption of transmitting, receiving, and merging u bit data is E_{Tx}, E_{Rx}, and E_{Mx}, respectively [10].

$$E_{Tx}(u,d) = \begin{cases} u \times E_{\text{elec}} + u \times \gamma_1 \times d^2, d < d_0 \\ u \times E_{\text{elec}} + u \times \gamma_2 \times d^2, d \ge d_0 \end{cases} \tag{6}$$

$$E_{Rx}(u,d) = u \times E_{elec} \tag{7}$$

$$E_{Mx}(u,d) = u \times E_{da}n \tag{8}$$

The energy consumption of sending u bit data to overcome environmental threat is:

$$E_{Dx} = u \times E_{ob} \times Th \tag{9}$$

γ_1 is the amplification factor of free space model, d is the distance between transmitter and receiver, d_0 is the distance threshold, γ_2 is the amplification factor of two-path fading model, and Th is the threat degree at the target location. E_{elec}, E_{da} and E_{ob} are the energy consumption of nodes receiving or transmitting, fusing, and sending 1 bit data to the threat source region, respectively.

Energy consumption is divided into two stages. In the initial stage of cluster head selection, the energy consumption of nodes mainly comes from transmitting data to the threat source region [11], which is calculated by Formulas 6 and 9. The energy consumption of the cluster head comes from receiving and merging the data sent by nodes in the cluster, which is calculated by Formulas 7 and 8. In the second stage, we need to adjust the sink's stay point, and divide the cluster heads obtained in the first stage into each sink's stay point. The energy consumption comes from the cluster heads transmitting all the data obtained in the first stage to the sink, which is calculated by Formulas 6 and 9.

3 Virtual Force and Improved LEACH Based Trajectory Planning Algorithm

3.1 First-generation Cluster Head Selection Based on ILEACH

The traditional LEACH method calculates the cluster head measurement factor $V(n)$ of each node in each round, and $V(n)$ determines whether the node can become the cluster head.

$$V(n) = \begin{cases} \frac{p_1}{1-p_1[\ \mathrm{mod}\ (r,1/p_1)]} \times \left(\frac{E_i(r)}{E_{\mathrm{avg}}(r)}\right)^2, & n \in G \\ 0, & n \notin G \end{cases} \tag{10}$$

where r is the current network running round, p_1 is the pre-set proportion of cluster heads. $E_i(r)$ and E_{avg} are the energy of node i under the current round and the average energy of all nodes in the network, respectively. G is the set of nodes that have not been cluster heads in the previous $1/p_1$ rounds. If $V(n)$ of node i is greater than Th_i, then node i becomes the cluster head of this round. Th_i is a number in $[0, 1]$, which is randomly generated by each node in each round.

In actual information collection, we expect nodes with the high regional threat to avoid becoming the cluster head, because such nodes are easily damaged, and the information they hold cannot be transmitted to the monitoring center. Once the cluster head is damaged, the information of all nodes in the cluster becomes invalid. Considering the node energy and environmental threat degree comprehensively, an improved LEACH algorithm is proposed. The improved cluster head measurement factor $V(n)$ is as follows:

$$V(n) = \begin{cases} \frac{p_1}{1-p_1[\ \mathrm{mod}\ (r,1/p_1)]} \times \left(\frac{E_i(r)}{E_{\mathrm{avg}}(r)}\right)^2 \times \left(1 - \left(\frac{\mathrm{Threat}_i(r)}{\mathrm{Threat}_{\mathrm{avg}}(r)}\right)^2\right), & n \in G \\ 0, & n \notin G \end{cases} \tag{11}$$

where Threat_i is the threat degree of the location of node i, and $\mathrm{Threat}_{\mathrm{avg}}$ is the average threat degree of the current round of network nodes. Calculate $V(n)$ of each node and compare it with Th_i to determine whether this node becomes the cluster head.

3.2 Adjustment of Sink's Stay Point Based on Virtual Force

After obtaining the first-generation cluster head of the network, in order to further reduce the trajectory length, we use the idea of TSPN to find the sink's stay points set S_{stop}, and complete the network information collection by visiting the points in S_{stop}. In this Section, we use virtual force [9] to determine the stay points of sink. Assuming that *Count* cluster heads determine a sink stay point *Stop*, the position of the initial time *Stop* is the coordinates of the randomly selected cluster head p, and then select the $Count - 1$ cluster heads closest to p to complete the determination of a stay point together. The stay point is affected

by three forces F_{dis}, F_{e}, F_{num}, which are caused by the distance dis , energy e and the number of nodes in the cluster num respectively. The stay point is subjected to the force of each cluster head as:

$$F_{\text{dis}} = \begin{cases} w_a \left(D - D_{\text{cavg}}\right) \times \boldsymbol{d}, & D > D_{\text{cavg}} \\ 0, & D \leq D_{\text{cavg}} \end{cases} \tag{12}$$

$$F_{\text{e}} = \begin{cases} w_b \left(E - E_{\text{cavg}}\right) \times \boldsymbol{d}, & E < E_{\text{cavg}} \\ 0, & E \geq E_{\text{cavgc}} \end{cases} \tag{13}$$

$$F_{\text{num}} = \begin{cases} w_c \left(N - N_{\text{cavg}}\right) \times \boldsymbol{d}, & N > N_{\text{cavg}} \\ 0, & N \leq N_{\text{cavgc}} \end{cases} \tag{14}$$

where the average energy of $Count$ cluster heads is E_{cavg}, the average distance from the initial stay point is D_{cavg}, and the average node number of each cluster group is N_{cavg}. D, E and N are the distance from the cluster head to the initial stay point, the remaining energy of the cluster head, and the number of nodes in this cluster, respectively. The direction of the force is directed from each cluster head to S_{stop}. w_a, w_b, w_c are coefficients. $\boldsymbol{d}$ is the direction of force. The resultant force on the sink stay point is F_{total}.

$$F_{total} = F_{dis} + F_e + F_{num} \tag{15}$$

The sink moves under the action of resultant force F_{total}. Assuming that the single moving step is $Step$, and the initial stay position is p, the next point p' is:

$$p' = p + F_{\texttt{total}} \times Step \tag{16}$$

3.3 Solving the Shortest Trajectory

After getting the sink's stay point set S_{stop}, the genetic algorithm is used to find the shortest trajectory. The fitness function of the algorithm is the reciprocal $f = 1/l$ of the trajectory length l of each stay point. Each individual in this problem is the order of access to points in S_{stop}. After crossover and mutation, repeat judgment is made to avoid multiple access to one point and not to others.

4 Simulation and Result Analysis

In the three-dimensional space with the simulation area of $200 \times 200 \times 30$, the initial energy of all nodes is $E_0 = 1$J, the communication radius is R_c, the total number of network nodes is N, the total running round is $round$, the percentage of cluster head is $P = 0.4$, and the step length is $Step = 1$. $w_a = 1/20$, $w_b = 1/2$, $w_c = 1/10$, $E_{elec} = E_{Tx} = E_{Rx} = 5 \times 10^{-8}$ J, $E_{DA} = 5 \times 10^{-9}$ J, $E_{ob} = 5 \times 10^{-9}$ J, $\gamma_1 = 10 \times 10^{-12}$ J, $\gamma_2 = 13 \times 10^{-16}$ J. Actual energy consumption against threat sources needs to be multiplied by the threat level of target location. There is a anti-aircraft artillery in space $[30, 50, 1]$ with a maximum threat distance of 80. $H_{min} = 20$, $H_{max} = 60$, $v_1 = 0.01$, $v_2 = 0.02$, $v_3 = 0.03$. Nodes are randomly distributed in the MWSN.

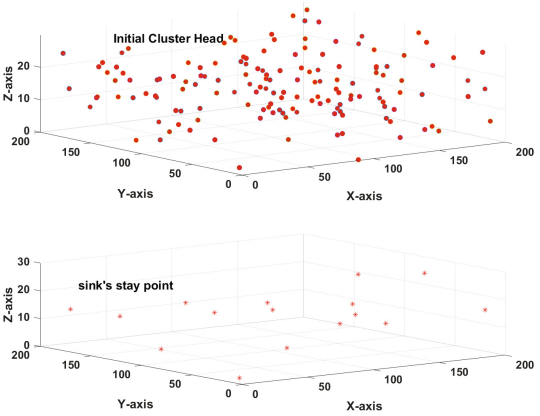


Fig. 1. The initial cluster heads and sink's stay points

Firstly, $N = 500$, $R_c = 40$, $Round = 1000$, The total number of packet bags $Packetbag$ is 4×10^4. Figure 1 shows the initial cluster heads after ILEACH clustering and the sink's stay points obtained by V-ILEACH. It can be seen that after V-ILEACH, the number of sink stay points is far less than the number of initial nodes, which greatly shortens the trajectory length of data collection.

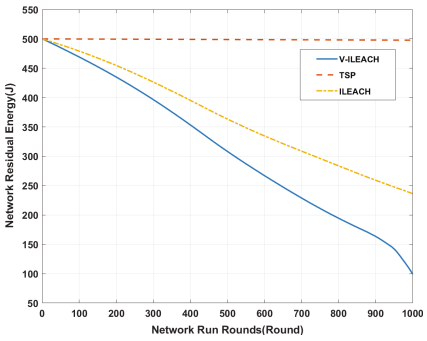


Fig. 2. Network residual energy

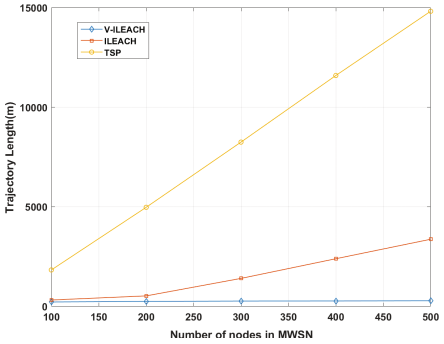


Fig. 3. Trajectory length under different number of nodes

Figure 2 shows the MWSN residual energy. In this experiment, based on the traditional TSP algorithm that the sink traverses each node to collect information, the node does not have the energy consumption of data transmitting, and only the energy consumption of overcoming the threat sources, so the network energy is almost unchanged. In the ILEACH algorithm, the energy consumption of MWSN mainly comes from the nodes sending information to its cluster head and overcoming the threat sources. In the V-ILEACH algorithm, in addition to

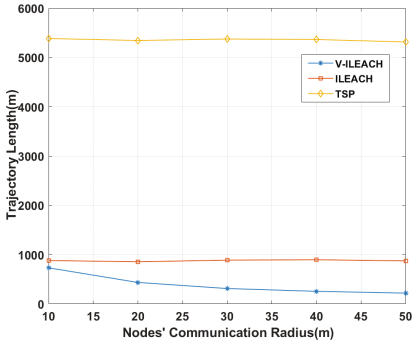


Fig. 4. Trajectory length under different nodes' communication radius

the energy loss in the ILEACH clustering, the cluster heads also need to send data to the sink, so the energy consumption is the largest. It can be seen that under the 1000 rounds of operation, the V-ILEACH consumes 80% of the network energy. Although the network energy decreases faster under the V-ILEACH method, it still ensures that the network can collect enough rounds of data.

Figure 3 demonstrates the trajectory length under different number of nodes. The sink needs to traverse all nodes based on TSP, so the trajectory length increases linearly, and is always the longest. Based on ILEACH, the number of stay points of the sink is less than that of TSP, and the trajectory length is shorter. The V-ILEACH algorithm uses the virtual force to adjust the position of the stay points of the ILEACH algorithm, and determines a sink stay point in *Count* cluster heads. The cluster heads whose distance from the stay point are less than the communication radius will be assigned to this stay point, which further reduces the trajectory length, and the increase in the network size will hardly increase the number of stay points. When the number of sensor nodes in MWSN is N, there are N points for TSP, $N \times p_1$ points for ILEACH and $N \times p_1/$ `Count` points for V-ILEACH. According to Fig. 2 and Fig. 3, the reduction of V-ILEACH trajectory length is at the expense of energy.

It can be seen from Fig. 4 that the trajectory length of V-ILEACH algorithm is the shortest, and with the increase of R_C, more cluster heads are divided into the same stop point, and the trajectory length decreases. The trajectory length obtained by TSP algorithm and ILEACH algorithm is almost unchanged. This is because TSP-based algorithm needs to traverse all nodes, and ILEACH clustering algorithm is independent of communication radius.

5 Conclusion

In this paper, taking into account the network nodes' characteristics, distribution characteristics and environmental threat, a virtual force and improved LEACH (V-ILEACH) based trajectory planning algorithm is proposed, which shortens the trajectory length of the sink for data collection and reduces the network

energy consumption. Compared with TSP and ILEACH, the trajectory length of V-ILEACH is reduced by 98% and 91%, respectively, but the energy consumption is increased by 12% and 2.2%, respectively, indicating that the trajectory length is greatly reduced at the expense of a small amount of energy. This method is particularly useful in the case of data collection without multi-round execution.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (61731006, 61671138), and was partly supported by the 111 Project No. B17008.

References

1. Perillo, M., Cheng, Z., Heinzelman, W.: An analysis of strategies for mitigating the sensor network hot spot problem. In: International Conference on Mobile & Ubiquitous Systems: Networking & Services (2005)
2. Zhao, M., Yang, Y.: Bounded relay hop mobile data gathering in wireless sensor networks. IEEE Trans. Comput. **61**(2), 265–277 (2010)
3. Gao, S., Zhang, H., Das, S.K.: Efficient data collection in wireless sensor networks with path-constrained mobile sinks. IEEE Trans. Mob. Comput. **10**(4), 592–608 (2011)
4. Cheng, C.F., Yu, C.F.: Data gathering in wireless sensor networks: a combine-tsp-reduce approach. IEEE Trans. Veh. Technol. **65**(4), 2309–2324 (2016)
5. Ma, M., Yang, Y., Zhao, M.: Tour planning for mobile data-gathering mechanisms in wireless sensor networks. IEEE Trans. Veh. Technol. **62**(4), 1472–1483 (2013)
6. Olariu, S., Stojmenovic, I.: Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting. In: Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications, pp. 1–12 (2006). Citeseer
7. El-Moukaddem, F., Torng, E., Xing, G.: Mobile relay configuration in data-intensive wireless sensor networks. IEEE Trans. Mob. Comput. **12**(2), 261–273 (2011)
8. Heinzelman, W.R.,Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS 2000) (2002)
9. Zou, Y., Chakrabarty, K.: Sensor deployment and target localization based on virtual forces. In: Joint Conference of the IEEE Computer & Communications IEEE Societies (2003)
10. Younis, O., Fahmy, S.: Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans. Mob. Comput. **3**(4), 366–379 (2004)
11. Olariu, S., Stojmenovic, I.: Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting. In: Infocom IEEE International Conference on Computer Communications (2006)

Control System Modeling and Solving for the Multi-objective Level Balance

Jinlin Huang(✉) and Guanghua Cao

School of Electrical Engineering, Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241000, Anhui, China
0125000562@ahcme.edu.cn

Abstract. This paper is devoted to the problem of the material level balance in six towers during the production process of Hydrometallurgy, introduced the process. According to the demand of the maximum material emissions and the less cycles to achieve balance, established the mathematical model. Used the basic decision-making theories of Voting Analytic Hierarchy Process (VAHP) to solve the model. In the Siemens S7-300PLC, designed the program, obtained the material emissions of six towers. Through the simulation software of PLCSIM, results show that after the end of each production cycle, this program can automatically determine the material emissions which can make the level trends to balance.

Keywords: Multi-objective · Mathematical Model · Level · Balance

1 Introduction

In the process of hydrometallurgy production, the level height of the six-way tower is usually manually adjusted, which often leads to the imbalance of the level of the six-way tower and directly affects the production efficiency [1]. To ensure the high quality of products, high economic benefits, with the least input to achieve the maximum output, it is required that the material level of each tower after the discharge of materials tends to be consistent, which belongs to the category of multi-objective control decision-making.

At present, the methods of multi-objective control include linear weighted sum method, multi-attribute utility method, analytic hierarchy process, multi-objective fuzzy decision making, hierarchical sequence method and so on [2]. The Voting Analytic Hierarchy Process (VAHP) refers to ranking all targets in order of importance. Firstly, the optimal solution of the first and most important objective is obtained, and then the optimal solution of the next objective is obtained successively on the premise of ensuring the optimal solution of the previous objective, until the last objective is obtained, which provides a basis for the multi-objective level balance control of metal hydrometallurgy [3]. PLC is widely used in various industries in the field of automatic control, it can through the solution process of VAHP, the water metallurgy multi-objective level control programming to achieve automatic level balance [4].

In this paper, the six tower is taken as the controlled object, the material balance principle is applied to build a mathematical model, and the VAHP is used to analyze the

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 267–273, 2023.
https://doi.org/10.1007/978-981-99-2653-4_33

model. In Siemens S7-300PLC, the level balance control decision procedure is designed, and the level height of the six tower tends to be the same quickly by controlling the material discharge.

2 Hydrometallurgical Process

Resin ion exchange is used in water metallurgy process [5], as shown in Fig. 1, first of all, the ore is mined underground, crushed and leached to the storage yard for heap leaching, and the leached stock solution is pumped to the stock solution elevation tank. When the feed valve is opened, the stock solution flows into the adsorption tower for adsorption, and the resin selectively absorbs the metal ions in the stock solution. When the adsorption is saturated, the feed valve is closed, and after a delay of a period of time, the discharge is opened to discharge the resin into the press-feeding tank, and then the resin is lifted to Tower 2 for cleaning. The resin is washed in tower 3, Tower 4 and Tower 5, and the resin is pressed into tower 6. The function of Tower 6 is to restore the readsorption capacity of the resin, and then lift the resin to Tower 1 for readsorption to realize the cycle process of the whole process, and the qualified metal liquid flows out of tower 3. To improve the adsorption efficiency and yield, it is necessary to control the material level in the six towers to reach the same height after resin discharge in each tower.

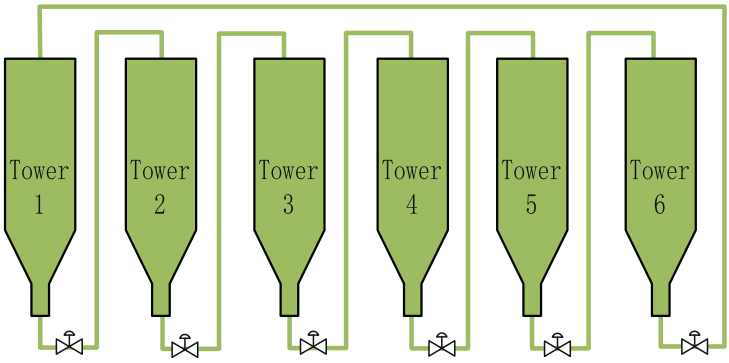


Fig. 1. The process of the Multi-objective Level Balance Control System

3 Mathematical Model of Level Balance

In production, the shape and specifications of each tower are the same. The diameter of the tower is R_1, the height is H_1, and the cross-sectional area of the tower can be calculated as s. What this paper wants to control is the corresponding height of each tower resin discharge, so that the final height of each tower is consistent. The ideal level

height h_v of each tower is the average value of the initial level height of each tower, expressed as:

$$h_v = \frac{\sum_{i=1}^{6} h_i}{6} \tag{1}$$

where, h_1– h_6 represents the initial object height of each tower.

Take tower 1 as an example, the initial object level height is h_1, assume that the height increase corresponding to flowing from tower 6 to tower 1 is x_6, and the height reduction corresponding to flowing from tower 1 to tower 2 is x_1, then the object level height of Tower 1 after adjustment satisfies the following formula:

$$h_1\_new = h_1 - x_1 + x_6 \tag{2}$$

To balance the level in the hydraulic system, the square of the difference between the average level height and the adjusted level height should be minimized:

$\min f_{11}(x)$, $\min f_{21}(x)$, $\min f_{31}(x)$, $\min f_{41}(x)$, $\min f_{51}(x)$, $\min f_{61}(x)$.

Where, $f_{11}(x) = [h_v - (h_1 - x_1 + x_6)]^2$, $f_{21}(x) = [h_v - (h_2 - x_2 + x_1)]^2$,

$f_{31}(x) = [h_v - (h_3 - x_3 + x_2)]^2$, $f_{41}(x) = [h_v - (h_4 - x_4 + x_3)]^2$,

$f_{51}(x) = [h_v - (h_5 - x_5 + x_4)]^2$, $f_{61}(x) = [h_v - (h_6 - x_6 + x_5)]^2$.

Under the condition that the level of each tower tends to balance, make the resin discharge x_1, x_2, x_3, x_4, x_5, x_6 of each tower as large as possible and make the level quickly tend to balance.

Since tower 1 is an adsorption tower, in order to discharge as much adsorbed saturated resin as possible, it is assumed that the upper limit of material discharge height x_1 of Tower 1 is $Q1\_up$ and the lower limit is $Q1\_down$, that is, $x_1 \in [Q1\_down, Q1\_up]$. Assume that the upper limit of material discharge height of other paths is $Q2\_up$ and the lower limit is $Q2\_down$, i.e. $xi \in [Q2\_down, Q2\_up]$.

So the constraint is:

$$\begin{cases} Q1\_down \le x_1 \le Q1\_up \\ Q2\_down \le x_i \le Q2\_up \ \ (i = 2, \cdots, 6) \end{cases} \tag{3}$$

In summary, the objective function of the mathematical model of the multi-objective level system of hydrometallurgy is as follows:

$\min f_{11}(x)$, $\min f_{21}(x)$, $\min f_{31}(x)$, $\min f_{41}(x)$, $\min f_{51}(x) \min f_{61}(x)$, $\max x_1$, $\max x_2$, $\max x_3$, $\max x_4$, $\max x_5$, $\max x_6$.

$$s.t \begin{cases} Q1\_down \le x_1 \le Q1\_up \\ Q2\_down \le x_i \le Q2\_up \ \ (i = 2, \cdots, 6) \end{cases}$$

4 VHAP for Multi-objective Level Balance Control

According to the principle of hierarchical sequence method [6], all the objects (x_1, x_2, x_3, x_4, x_5, x_6) are sorted in order of their importance on the premise that the deviation

between the object level height and the average object level height is minimum. Since tower 1 is an adsorption tower, in order to improve the adsorption efficiency, should approach to 0 as soon as possible, and x_1 should be as large as possible. Firstly, the value of x_1 is determined according to the deviation between the initial object level height of tower 1 and the average object level height, namely $x_1 = h_1\text{-}h_v + Q1\_up$, if:

$$\begin{cases} x_1 \geq Q1\_up, x_1 = Q1\_up \\ x_1 \leq Q1\_down, x_1 = Q1\_down \\ Q1\_down < x_1 < Q1\_up, x_1 = x_1 \end{cases} \tag{4}$$

After determining x_1, $x_6 = h_v + x_1\text{-}h_1$ can be obtained according to the requirement $min\, f_{11}(x) = [h_v\text{-}(h_1\text{-}x_1 + x_6)]^2$, Similarly, if:

$$\begin{cases} x_6 \geq Q2\_up, x_6 = Q2\_up \\ x_6 \leq Q2\_down, x_6 = Q2\_down \\ Q2\_down < x_6 < Q2\_up, x_6 = x_6 \end{cases} \tag{5}$$

And the same goes for $x2$, $x3$, $x4$, $x5$.

After the resin discharge is carried out according to the value of x1– x6, the theoretical calculated value of the tower level height is:

$$h_{ik} = h_{ik-1} - x_i + x_{i-1} \tag{6}$$

In the next production cycle, the values of x1– x6 are recalculated until the height of each tower level tends to the mean value.

5 Solving for Multi-objective Level Balance Control

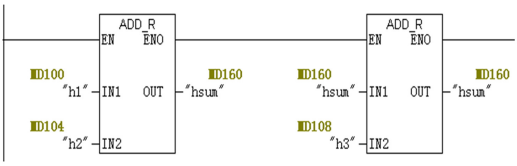
5.1 PLC Program Design of Level Balance Control

According to the analysis process of VHAP, the multi-objective level balance control program of hydrometallurgy is programmed by STEP V5.4 version of PLCS7–300. STEP7 adopts the structured programming method, which places the program written by users and the data required by the program in the "block", standardizes a single program component, and structured the user program through similar subroutine calls within or between blocks [7].

In the main program OB1, the function of level balance FC1 is called. In FC1, the average value of the initial level height is firstly solved according to the detected level value of the six-way tower, and then the corresponding height of the material discharge of each tower is obtained by calling FB1. Part of the program of FC1 is shown in Fig. 2. In the figure, program segments 1, 2 and 3 are to calculate the average object level height, and 4 and 5 are to calculate the material discharge height of tower 1 and tower 6 respectively. The material discharge of Towers 2 to 5 is calculated similarly.

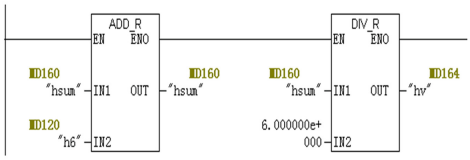
In the functional block FB1, the material discharge is first calculated according to the analytic hierarchy process, and then the obtained material discharge height is compared with the constraint conditions to finally determine the value of the material discharge.

Segment 1: Calculation hv=(h1+h2+h3+h4+h5+h6)/6

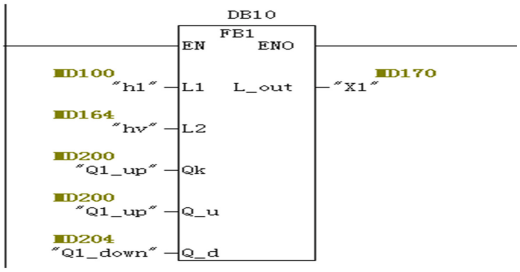


Segment 2:

Segment 3:



Segment 4:



Segment 5:

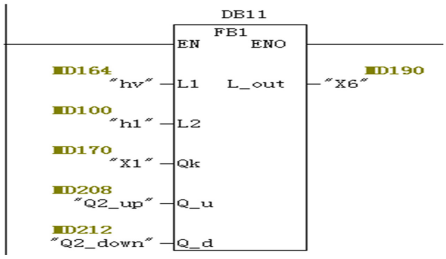


Fig. 2. FC1 program

5.2 Operation Results and Analysis

In S7-PLCSIM simulation software, the known object level height value is realized through the PLC program of analytic hierarchy process, and the material discharge is obtained for simulation. It is assumed that the initial object level height of each tower is 11.0, 9.0, 9.5, 10.5, 10.2, 9.8, and the constraint condition of X1 is [2,2.3]. The constraints of X2– X6 are [1.1, 2.3]. In the next cycle, the material level height of the obtained six-way tower is taken as the initial material level height to solve X1– X6, so that the material level automatically tends to balance. After reaching the control requirements, the material discharge height of each tower is discharged according to the maximum quantity.

To explain the changes of the material level of each tower, the simulation calculation of the material level height after resin discharge in each cycle is carried out, and the program is written in the function block FB2, as shown in Fig. 3. Where, *hnew* represents the material level height after discharge, hi represents the initial height, *Qout* represents the material level value discharged from the tower, and *Qin* represents the material level value flowing into the tower. Find the height according to *hnew* = *hi-Qout* + *Qin*.

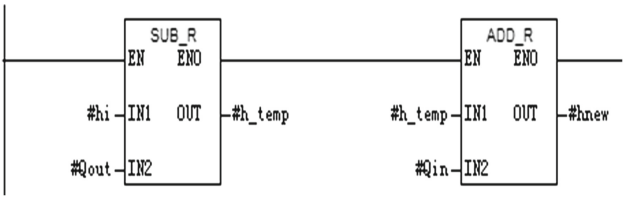


Fig. 3. Post-discharge level height program

The simulation results of material discharge and level height after resin discharge are shown in Table 1.

Table 1. The initial height of each tower and the height of the discharge

	Tower 1	Tower 2	Tower 3	Tower 4	Tower 5	Tower 6
Initial height	11	9	9.5	10.5	10.2	9.8
First discharge	2.3	1.3	1.1	1.6	1.8	1.3
Height after first discharge	10	10	9.7	10	10	10.3
Secondary discharge	2.3	2.3	2	2	2	2.3
Height after the second discharge	10	10	10	10	10	10

As shown in Table 1, at the end of the second cycle, the height of each tower is 10 m, which has reached equilibrium. In the later production cycle, the discharge height of each tower is discharged according to the maximum quantity, and the material discharge height corresponding to the tower is 2.3 m.

In this system, the level sensor is used to obtain the level height of each tower, and the resin emission of each production cycle can be obtained through the above procedure, and the level automatically tends to balance.

6 Conclusion

In this paper, a mathematical model is established for the level balance problem of the six-way tower in the production process of a metal hydraulic multi-objective level system, and the model is analyzed by using VHAP. The resin emission in each tower is obtained by programming in PLC S7–300, and the simulation in PLCSIM shows that the method can automatically realize the level balance of the six-way tower. It replaces the traditional production mode of manually adjusting material level, greatly saves human resources and improves adsorption efficiency.

Acknowledgements. This work was supported by the Excellent Young Talents support Program of Anhui Province, China (NO.gxyq2021259), Natural Science Research Project of Anhui Province of P. R. China (2022AH052368), Quality Project of Anhui Province of P. R. China (2020SJJD029).

References

1. Mudd, G.M.: Acid in Situ Leach Uranium Mining-USA an Australia. Tail. Mine Waste. **2020**, 23–26 (2020)
2. Uranium Energy Corporation: Palangana Hobson Operation Tour. Uranium Energy Corporation, Vancouver (2017)
3. Mooiman, M.B.: Challenging the traditional hydrometallurgy curriculum—an industry perspective. Hydrometallurgy **79**, 80–88 (2019)
4. The extraction of adsorption material in the metal and other useful components. Beijing: Metallurgical Industry Press, pp. 47–72 (2015)
5. Tinggang, Z.: Mathematical modeling methods and mathematical models. Science Press, Beijing (2011)
6. Xiaopeng, W.: Three-Tank water tank level control system dynamic simulation and multi-line monitoring of process variables. Shandong University master's degree thesis (2018)
7. Hecker, S.: Enhanced LFR-toolbox for Matlab. Aerosp. Sci. Technol. **9**, 173–180 (2005)

Soil Moisture Extraction Based on Time-Frequency Analysis

Peng Chen, Yuefei Duan, and Jing Liang(✉)

School of Information and Communication Engineering,
University of Electronic Science and Technology of China, Chengdu, China
liangjing@uestc.edu.cn

Abstract. Ultra-wideband sensor has been shown to capture soil moisture well without contact, but the processing of soil echo signals mostly stays in the time domain. This paper introduces frequency domain features into the analysis of soil echo signals, and a soil moisture acquisition method is proposed based on time-frequency analysis and deep learning. We explore the influence of the size of the window on the network performance based on the short-time Fourier transform(STFT). Then we compare the impact of the three time-frequency analysis methods of STFT, Continue wavelet transform(CWT), and Wigner-Ville Time-Frequency Distribution(WVD) on the ability of the network to extract soil moisture. The results show that the network based on STFT has better soil moisture identification when the window size is suitable.

Keywords: UWB · soil moisture · time-frequency analysis · short-time Fourier transform

1 Introduction

The arbitrary destruction of soil resources has gradually attracted researchers' attention in recent years. Soil moisture, as one of the important factors determining whether the soil is suitable for planting has become the focus of research. The extraction of soil moisture can be mainly divided into three categories: (a). Drying method. (b). Intrusive electromagnetic signal measurement method [1–3]. (c). Non-invasive electromagnetic signal measurement method [4,5].

The drying method obtains soil moisture by drying it. The intrusive electromagnetic signal measurement method is mainly to measure the transmission time of electromagnetic wave in soil and then calculate its dielectric constant. Then calculate the soil moisture according to the model formulas [6,7]. But both drying methods and intrusive electromagnetic signal measurements will destroy soil structure and can not monitor soil moisture in real-time. Non-invasive measurement should be the future research direction of soil moisture. Wang et al. [8,9]. Used ultra-wideband sensors and provided a new idea for non-contact soil moisture extraction in the time domain. In addition, in their studies, deep learning and machine learning [10,11] have proved to be highly effective means to extract soil characteristics.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 274–281, 2023.
https://doi.org/10.1007/978-981-99-2653-4_34

The frequency domain must also contain soil moisture information as another signal feature. This paper investigates time-frequency analysis to extract soil moisture. It compares the three time-frequency analysis methods of short-time Fourier transform(STFT), wavelet transform(CWT), and Wigner-Ville Time-Frequency Distribution(WVD) on the ability to extract soil moisture.

The arrangement of this paper is as follows: Section 2 Introduce the acquisition of soil signals by ultra-wideband sensors and the time-frequency analysis of signals. Section 3 Introduce the soil moisture extraction system based on time-frequency analysis and CNN. Section 4 Compare the performance of the proposed method with the existing methods. Section 5 Conclude this paper.

2 Signal Preprocessing and Time-Frequency Analysis Theorys

The data used in this paper are collected in Chengdu using p440 ultra-wideband sensor(Figs. 1 and 2) and soil moisture meter TDR300. P440 has high resolution and small volume, which makes it easier to extract soil water from UAVs and other vehicles. TDR300 is an accurate instrument for measuring soil water content, and its measurement error is ±3%.

Fig. 1. Ultra-wideband chip module-P440

Fig. 2. Self-made antennas

The signal's time-frequency analysis can combine the signal's time domain and frequency domain well. This section will briefly introduce the three time-frequency analysis theories (Short-Time Fourier Transform, Wigner-Ville Time-Frequency Distribution, and Wavelet Transform).

2.1 Data Collection and Processing

In the experiment, the drone was equipped with a P440 ultra-wideband sensor to collect data at a distance of 85 cm from the ground, and the probe length of tdr300 is 12 cm.

Through the experiment, we obtained 130 sets of data. The echoes are shown in Fig. 3.

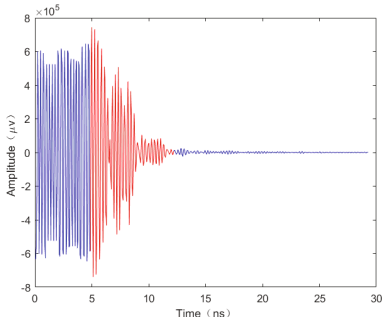


Fig. 3. The soil echo signal

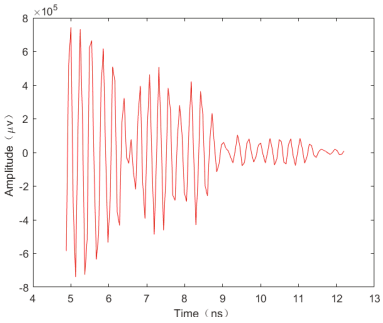


Fig. 4. Effective soil echo signal

However, a large part of the signals we collected did not propagate in the soil, so they do not carry soil moisture information [12,13]. Through the Eq. 1 as follow:

$$t = s/v = s/\frac{c}{\sqrt{\xi\mu}} \tag{1}$$

where ξ is the Dielectric constant, t is the propagation time of electromagnetic wave in soil, v is the velocity of the electromagnetic wave in soil, s is the distance between two probes, and c is the speed of light. We usually take μ to be 1.

we calculate that the signal carrying soil moisture is between $4.94\,ns$ and $12.2\,ns$ as shown in Fig. 4.

2.2 Time-Frequency Analysis of Signals

The stationary signal has been constant in time. However, a particular statistic of many artificial signals and natural signals is a function of time, and we call these signals non-stationary signals. For these signals, the Fourier transform is no longer applicable, but local transforms can solve the limitations of Fourier transforms well.

The following equations are the local transformations used in this paper:

Short-Time Fourier Transform(STFT)

$$STFT(m,n) = \sum_{k=-\infty}^{\infty} z(k)g^*(kT - mT)e^{-j2\pi(nF)k} \tag{2}$$

where $z(k)$ is the discrete signal.$g(t)$ is the window function.

Its spectrum is defined as the square of $STFT(m,n)$, namely:

$$SPEC(t,\omega) = |STFT(t,\omega)|^2 \tag{3}$$

Wigner-Ville Time-Frequency Distribution(WVD): The time-frequency distribution is defined as:

$$P(t,\omega) = \int_{-\infty}^{\infty} R(t,\tau) e^{-j\omega\tau} d\tau \tag{4}$$

where $R(t,\tau)$ is the local correlation function.

Choose the time impulse function$\phi(u-t,\tau) = \delta(u-t)$ as the window function of $R(t,\tau)$, and get wigner-ville time-trequency distribution as:

$$W_z(t,\omega) = \int_{-\infty}^{\infty} z\left(t+\frac{\tau}{2}\right) z^*\left(t-\frac{\tau}{2}\right) e^{-j\omega\tau} d\tau \tag{5}$$

Wigner-ville time-trequency distribution can describe the local energy distribution very well.

Wavelet Transform(CWT)

$$WT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) h^*\left(\frac{t-b}{a}\right) dt = \int_{-\infty}^{\infty} x(t) h_{ab}^*(t) dt = \langle x(t), h_{ab}(t) \rangle \tag{6}$$

where the kernel function [15] $h_{ab}(t)$ is:

$$h_{ab}(t) = \frac{1}{\sqrt{a}} h\left(\frac{t-b}{a}\right) \tag{7}$$

The scale map of the wavelet transform is defined as:

$$SCAL(a,b) = |WT(a,b)|^2 \tag{8}$$

3 The Soil Moisture Extraction System

In this section, we will group the experimental data, use time-frequency analysis to obtain the spectrum or scale map of the signal, and use the powerful image processing capability of deep learning to complete the extraction of soil moisture.

3.1 Data Selection

Some scholars [5,9,13] have treated each echo in previous studies as a separate category, mainly because the sampling location is fixed. In this case, the echo will not be disturbed by the soil structure itself or buried objects. To make the research more realistic, we collect data from different locations, and we should group the experimental data to avoid soil structure and submerged objects affecting the experimental results.

Since the experiment is conducted in the rainy season, our soil moisture is concentrated between 35% and 45%. Therefore, we cannot use all the data (130 sets). We divided the collected data into seven groups at 5% intervals. Each group contains 9 data, as shown in the Table 1:

Table 1. Soil Moisture Grouping.

Group1	Group2	Group3	Group4	Group5	Group6	Group7
19.62%	25.9%	31.02%	36.14%	40.92%	45.68%	51.26%
20.42%	27.4%	31.34%	36.34%	41.30%	45.88%	51.54%
21.80%	27.4%	31.50%	36.44%	41.40%	46.04%	51.66%
21.90%	27.8%	32.00%	36.64%	41.70%	46.06%	51.76%
22.88%	29.7%	32.10%	39.58%	45.06%	49.00%	52.14%
22.88%	30.0%	32.38%	39.78%	45.36%	49.38%	52.44%
23.18%	30.2%	34.50%	39.88%	45.46%	49.38%	53.32%
25.24%	30.2%	34.88%	40.24%	45.56%	50.08%	55.28%
25.26%	30.3%	35.72%	40.36%	45.64%	50.86%	58.14%

3.2 Time-Frequency Analysis of Soil Echo Signal

Each data in the group has 100 ultra-wideband echo signals corresponding to it(For example, we have 100 echo signals corresponding to soil water content of 19.62%). But to reduce the number of training, we only select 20 echoes for processing, which does not lead to a decrease in the recognition rate. We choose a signal whose soil moisture is 19.62% to display the spectrum or scale image obtained by the three time-frequency analysis theories.

First of all, for the short-time Fourier transform, the window function we choose is the Hanning window, and the size of the window is 5 and 100, respectively.

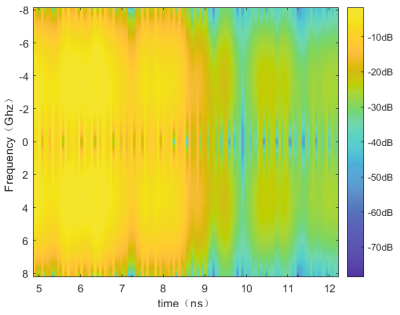


Fig. 5. STFT,window size is 5

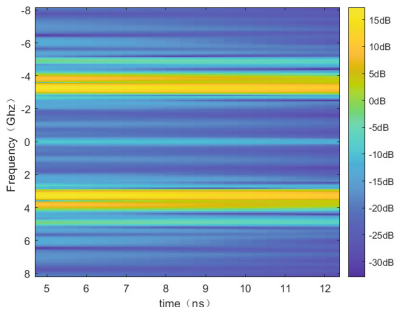


Fig. 6. STFT,window size is 100

From Figs. 5 and 6, we can see that as the value of the window becomes larger, the time domain resolution of the general map is getting lower and lower, and the frequency domain resolution is getting higher and higher.

Secondly, the spectrograms generated by WVD and Wavelet Transform is as shown as Figs. 7, 8.

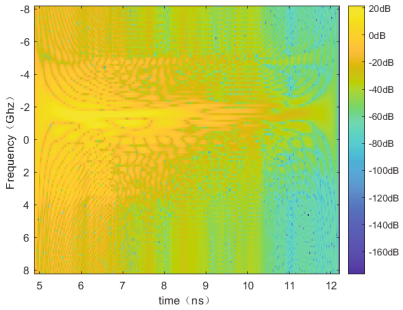


Fig. 7. WVD spectrogram

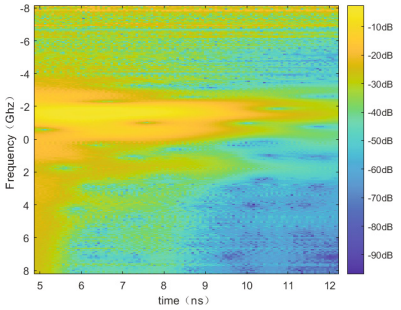


Fig. 8. Wavelet Transform Scale Map

3.3 Soil Moisture Extraction System

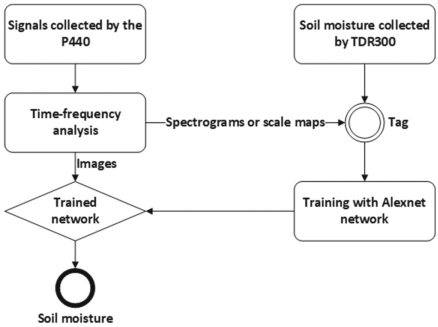


Fig. 9. The soil moisture extraction system

The soil moisture extraction system proposed in this paper is shown in Fig. 9. We perform time-frequency analysis on the signals(after interception) collected by the p440 to obtain a spectrogram or a scale map. After that, use the soil moisture collected by the TDR300 to label the pictures. And then we feed the labeled images into the ALexnet network to train the network. Finally, we input the untrained spectral map or scale map into the trained network, and then we can wait to get the soil moisture.

4 Results Analysis and Performance Comparison

We mentioned that the size of the window function for the STFT would affect the resolution of the time and frequency domain. Therefore, the spectrums generated by STFT with different window sizes also affect the network's Performance.

As is shown in Fig. 10, the best possible state for STFT is when the size of the window function is 10. When the window selection is smaller than 10, the network's performance also decreases. So, the window size should be flexibly adjusted according to the signal type and specific application. We can also see that the signal's time-domain mainly determines the information on soil moisture. Still, when the window function size is less than 10, the performance of the network will also decrease as the frequency resolution of the signal decreases, so we can get the frequency domain of the signal containing soil moisture information.

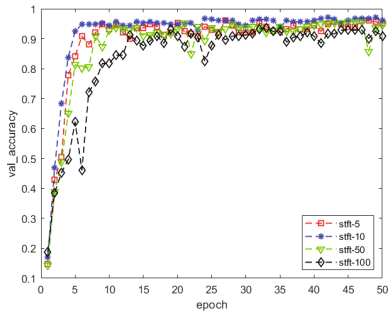


Fig. 10. The accuracy of STFT

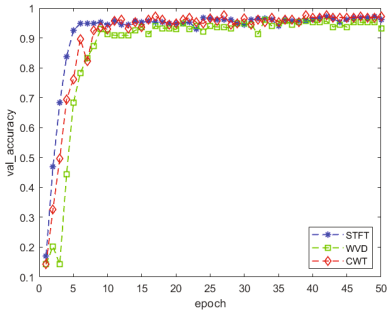


Fig. 11. The accuracy of three methods

For the performance comparison of the three methods, as shown in Fig. 11, we can see that the performance of STFT is the best, followed by CWT, and WVD is the worst. The main reason for this result is that WVD can represent the signal's energy distribution over time, but it ignores some information in the frequency and time domains. Studies have shown that the frequency spectrum will shift with the transformation of soil moisture for soil echo signals, so the resolution at each frequency should be consistent. For CWT, it has multi-resolution characteristics and can use different keys for signals of different frequencies, which leads to its performance is not as good as STFT.

5 Conclusion

This paper provided a method based on time-frequency analysis and machine learning to solve the non-destructive soil moisture extraction problem. We compared the average correct identification rate of soil moisture based on STFT with different window sizes. We also analyzed the average correct identification rate between the STFT and the other two methods(WVD and CWT).

The average correct identification rate of soil moisture of the STFT-based network will gradually improve when the window is reduced from 100 to 5. The average correct identification rate of soil moisture of the network based on STFT

is the highest when the window is 10. For the networks under the three time-frequency analysis methods, through performance comparison experiments, we found that the average correct identification rate of soil moisture based on STFT is the best, and that of CWD is the worst. In general, the soil moisture extraction system can better apply soil monitoring and contribute to reducing soil resource loss.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (61731006, 61671138), and was partly supported by the 111 Project No. B17008.

References

1. Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., Wagner, W.: An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations. IEEE Trans. Geosci. Remote Sens. **47**(7), 1999–2013 (2009)
2. Will, B., Rolfes, I.: Sensors. In: 2013 IEEE, pp. 1–4. IEEE (2013)
3. Piuzzi, E., Cataldo, A., Cannazza, G., De Benedetto, E.: An improved reflectometric method for soil moisture measurement exploiting an innovative triple-short calibration. IEEE Trans. Instrum. Meas. **59**(10), 2747–2754 (2010)
4. Taylor, R.A., et al.: In: IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, pp. 3974–3977. IEEE (2019)
5. Yang, C., Liang, J.: Soil PH value forecasting using UWB echoes based on ensemble methods. IEEE Access **7**, 173,249–173,256 (2019)
6. Alharthi, A., Lange, J.: Soil water saturation: dielectric determination. Water Resour. Res. **23**(4), 591–595 (1987)
7. Topp, G.C., Davis, J., Annan, A.P.: Electromagnetic determination of soil water content: measurements in coaxial transmission lines. Water Resour. Res. **16**(3), 574–582 (1980)
8. Malajner, M., Gleich, D.: In: 2016 IEEE/ACES International Conference on Wireless Information Technology and Systems (ICWITS) and Applied Computational Electromagnetics (ACES), pp. 1–2. IEEE (2016)
9. Wang, T., Liang, J., Liu, X.: Soil moisture retrieval algorithm based on TFA and CNN. IEEE Access **7**, 597–604 (2018)
10. Liu, H., Yang, Y., Wan, X., Cui, J., Zhang, F., Cai, T.: In: 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), pp. 46–51. IEEE (2021)
11. Babaeian, E., Paheding, S., Siddique, N., Devabhaktuni, V.K., Tuller, M.: Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning. Remote Sens. Environ. **260**, 112,434 (2021)
12. Liang, J., Liu, X., Liao, K.: Soil moisture retrieval using UWB echoes via fuzzy logic and machine learning. IEEE Internet Things J. **5**(5), 3344–3352 (2017)
13. Liu, X., Yu, X., Ren, J., Liang, J.: Soil moisture retrieval using UWB echoes via ANFIS and ANN. In: Liang, Q., Mu, J., Jia, M., Wang, W., Feng, X., Zhang, B. (eds.) CSPS 2017. LNEE, vol. 463, pp. 1261–1268. Springer, Singapore (2019). https://doi.org/10.1007/978-981-10-6571-2_151
14. Stoica, P., Moses, R.L., et al.: Spectral analysis of signals (2005)
15. Valens, C.: A really friendly guide to wavelets. ed. Clemens Valens (1999)

Cross-Modality Person Retrieval with Cross-Modality Loss Functions

Qing Dong[1], Jianglin Zhou[1], Jian Li[1], Song Gao[2], Shaoyan Gong[3], and Zhong Zhang[1(✉)]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
zhong.zhang8848@gmail.com

[2] The National Innovation Center of Intelligent and Connected Vehicles, Beijing 100176, China
gaosong@china-icv.cn

[3] Global Energy Interconnection Research Institute Co., Ltd., Beijing, China

Abstract. Cross-modality person retrieval aims to search a pedestrian image of RGB modality from infrared (IR) pedestrian images and vice versa. Many approaches utilize dual-stream models to extract the features of images from different modalities. Then, they use different types of loss functions to overcome the intra-class modality variations and the large cross-modality discrepancy. In this review, we introduce three types of loss functions used in cross-modality person retrieval including probability based, distance-based, and center-based loss functions. Afterwards, we combine the baseline network with different loss functions, and perform the experiments on a publicly available dataset, i.e., SYSU-MM01 to evaluate their impact on cross-modality person retrieval.

Keywords: Person retrieval · Cross-modality · Loss functions

1 Introduction

Given an image of pedestrian, person retrieval aims to search the same identity under non-overlapping cameras. However, person retrieval suffers from the challenges of large intra-class differences and small inter-class differences caused by view change, pose variation and occlusion. In order to overcome these challenges, some works [1–3] learn the robust features of pedestrian images. Meanwhile, many existing methods [4–6] train the networks with various loss functions to overcome these challenges. Nevertheless, conventional person retrieval is performed in a single modality and the methods for single modality person retrieval do not match people in the dark environment. Therefore, cross-modality person retrieval receives more interest.

Cross-modality person retrieval [7–10] which suffers from intra-modality variations and cross-modality discrepancy between RGB images and infrared (IR) images. To narrow the gap of different modalities, most methods [11–13] design

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 282–289, 2023.
https://doi.org/10.1007/978-981-99-2653-4_35

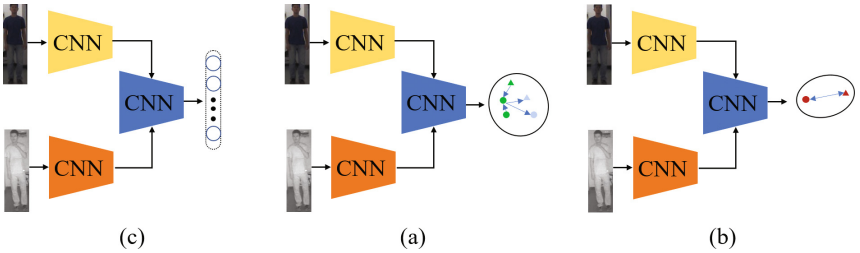


Fig. 1. The framework using (a) probability-based, (b) distance-based and (c) center-based losses for cross-modality person retrieval.

dual-stream networks to extract the features of images from different modalities, including the modality-specific information and the modality-shared information. Then, they use cross-modality loss functions to supervise the networks as shown in Fig. 1. Most methods [16,21,25,27] often design loss functions which are adopted to measure the similarity or distance between data. There are three commen types of loss functions that are probability-based, distance-based, and center-based. As for cross-modality person retrieval, loss functions allow deep model to extract more discriminable features of pedestrian images. For example, Wu et al. [14] consider cross-modality person retrieval as a classification task, which utilizes Cross-Entropy (CE) function to estimate the difference between the predicted probability distribution of each image belonging to each identity and the true probability distribution. Then, Ye et al. [15,21] desgin some loss functions based on the distance between images of different modalities, and they aim to learn an embedding space where images from different modalities belonging to the same identity are close to each other. Moreover, some approaches [16,18,22,28] propose to optimize the distance between centers of heterogenous modalities. That improves the intra-class similarity of images from different modalities. The distance-based loss and center-based loss always supervise network learning jointly with the CE function.

In this review, we introduce three common cross-modality loss functions for cross-modality person retrieval, such as probability-based CE loss [14], distance-based triplet loss [15,23] and center-based Hetero-Center (HC) loss [17]. The CE loss function is commenly used in cross-modality person retrieval, and some works [11,14,26] calculate the loss value between the predicted scores of model and the true label distribution for each image by using the CE function. The triplet loss randomly picks a postive image and a negative image accroding to the identity of anchor. Then it is designed to reduce the distance between the anchor and the positive image and push the negative image far from the anchor in the training process. As for cross-modality person retrieval, the triplet loss not only constrains the distance between anchor and positive images with different modalities, but also pushes the negative images from other modality away in an embedding space. Moreover, the HC loss function aims to pull the hetrogeneous centers with the same identity close, which could improve the cross-modality similarity and learn the modality invariant of features. We evluate their

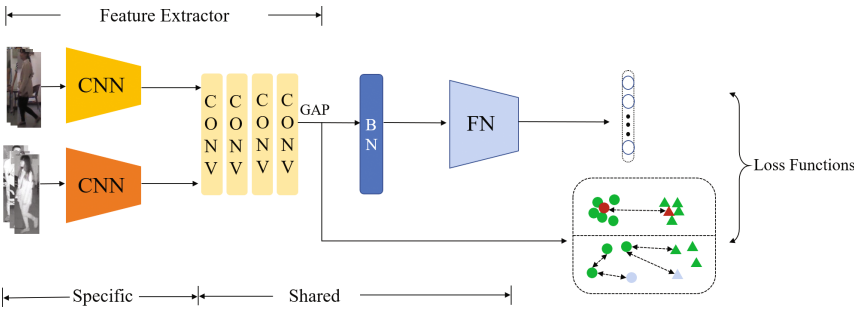


Fig. 2. The architecture of cross-modality person retrieval.

effectiveness on a publicly visible-infrared person images dataset, i.e., SYSU-MM01 [14] for cross-modality person retrieval. Furthermore, we combine three loss functions with the baseline model to obtain better performance.

2 Methods

In this section, we first present the feature extractor of cross-modality person retrieval. Then we introduce three loss functions in detail. Finally, we show our optimization scheme.

2.1 Feature Extractor

The feature extractor [24] consists of specific module and shared module derived from ResNet-50 [19] as shown in Fig. 2. The two specific modules are the first convolutional layers of ResNet-50 with different parameters. Meanwhile, the remaining layers of ResNet-50 are the shared module with the same parameters. We feed the RGB images and IR images into the dual-stream model, and we obtain the feature maps with the size of $W \times H \times C$, where W and H denote the width and the height of feature maps, C is the number of channels. Afterwards, we adopt the Global Average Pooling (GAP) operation on the feature maps to obtian the feature vetors of different modalities $F = \{g_i, r_j\}$. The features of j-th IR image and i-th RGB image are denoted as $r_j \in R^{C\times 1}$ and $g_i \in R^{C\times 1}$ respectively. Then, we leverage the Batch Normalization (BN) to stabilize the feature learning.

2.2 Cross-Entropy Loss

Many methods [10,24,26] employ a shared Fully-Connected (FC) layer as a classifier to obtain the predicted probability distribution. Then, the output of i-th image from the shared FC layer is represented as $q_i \in R^{D\times 1}$, where D is the number of person identities. The supervised CE loss function is designed to measure the similarity between the prediced probability distribution of network and

the real data distribution based on the labeled training images of two modalities, and it is defined as [14]:

$$L_{CE} = -\sum_{i=1}^{N} p_i \log q_i \tag{1}$$

where p_i is the real identity probability of i-th image and N is the total number of RGB images and IR images.

The CE loss narrows the gap between images of RGB modality and IR modality that is identity probability distributions of images from two different modalities with the same identity are similar under the guidance of the CE loss.

2.3 Triplet Loss

Learning distance metric is beneficial to mitigate the intra-class modality variations and the cross-modality discrepancy. Suppose that we obtain an anchor feature $f_i^a \in F$ whether it from RGB modality or IR modality, and we mine its hardest triplets (f_i^a, f_i^p, f_i^n) accroding to the Euclidean distance, where f_i^p and f_i^n are the farthest positive feature and closest negative feature of f_i^a respectively in the embedding space. The triplet loss [5] is defined as:

$$L_{tri} = \sum_{i=1}^{N} \max\left[m + D(f_i^a, f_i^p) - D(f_i^a, f_i^n), 0\right] \tag{2}$$

where $D(,)$ represent Euclidean distance function and m denotes a margin.

The triplet loss function makes sure that the person images with different identities are farther than person images belonging to the same identity by a margin m in an embedding space. Therefore, many methods adopt the triplet loss to improve the discriminability of features.

2.4 Hetro-Center Loss

Intuitively, improving the similarity between feature distributions with the same identity belonging to different modalities is helpful to address the cross-modality discrepancy. However, it is hard to directly constrain the distance between RGB feature distribution and IR feature distribution. So the HC loss [17] is proposed to reduce the distance between the two heterogeneous centers:

$$L_{hc} = \sum_{i=1}^{K} \left[\|c_i^R - c_i^I\|_2^2\right] \tag{3}$$

where $c_i^R = \frac{1}{m_1}\sum_{v=1}^{m_1} g_{i,v}$, $c_i^I = \frac{1}{m_2}\sum_{v=1}^{m_2} r_{i,v}$ denote the centers of two modalities in i-th identity, and K is the number of identities. $g_{i,v}$ and $r_{i,v}$ represents the v-th RGB feature and IR feature of i-th identity, respectively. m_1 and m_2 is the number of RGB images and IR images belonging to the same identity respectively.

The HC loss brigdes the gap between two modalities by constraining the distance between heterogenerous centers, and it promotes network to learn an embedding space where heterogenerous centers with the same identity are close.

2.5 Optimization Scheme

In order to validate the effectiveness of various loss functions in cross-modality person retrieval, we employ the CE loss, the triplet loss and the HC loss to optimize the network jointly. So the total loss function is:

$$L = \lambda_1 L_{id} + \lambda_2 L_{tri} + \lambda_3 L_{hc} \tag{4}$$

where λ_1, λ_2 and λ_3 are the weights to balance the importance between different losses.

3 Experiments

3.1 Dataset

We implement the experiments on a publicy cross-modality person retrieval dataset i.e., SYSU-MM01 [14].

SYSU-MM01 is a large-scale dataset collected by four RGB cameras and two infrared cameras. The training set consist of 22,258 RGB images and 11,909 IR images of 395 identities. In the testing stage, there are 96 identities of images from different modalitites, and IR images are query images while RGB images are gallery set. We employ two test modes to evaluate the effectiveness of three loss functions, i.e., indoor-search and all-search. The gallery set of all-search mode are collected by all RGB cameras. Meanwhile, only two RGB cameras that put in the indoor environment are used to build the gallery set for indoor-search mode.

Table 1. The performance of different loss functions for cross-modality person retrieval. Ri indicates Rank-i.

Methods	*All-search*					*Indoor-search*				
–	R1	R5	R10	R20	mAP	R1	R5	R10	R20	mAP
CE	44.58	72.95	83.28	92.05	43.92	46.69	78.73	89.35	96.33	56.19
tri	28.14	61.84	78.01	90.99	30.75	32.83	74.09	88.39	97.81	46.00
HC	1.18	5.56	10.87	21.55	3.02	1.76	8.75	17.74	36.25	6.87
CE + tri	47.66	74.31	83.27	90.59	43.74	51.63	78.66	88.06	95.05	58.09
CE + HC	43.83	71.30	81.73	90.34	39.87	45.26	74.71	85.98	94.30	52.58
CE + tri + HC	50.53	77.96	87.43	94.67	49.01	54.45	83.27	91.79	97.60	62.51

3.2 Implementation Details

The pedestrian images are resized into 288×144, then we apply the random cropping and the random horizontal flip to augment the data. The batch size is set to 64 where each batch consists of 4 identities, and each identity includes 8 IR pedestrian images and 8 RGB pedestrian images. During the training process, we set the trade-off weights λ_1, λ_2 and λ_3 to 1.0, 0.1, and 0.5 in Eq. 4. We optimize the network using the stochastic gradient descent (SGD) optimizer [29], and we set the epoch number to 80. The learning rate decreases to 0.01 after 30 epochs and decreases to 0.001 after 50 epochs, and we set the initial learning to 0.1 with a warm-up strategy [20].

3.3 Experimental Results

We perform the experiments on SYSU-MM01 in the indoor-search and all-search modes to validate the impact of each loss function. The experimental results are shown in Table 1. *CE*, *tri*, and *HC* represent the dual-stream network with the CE loss, the triplet loss and the HC loss respectively.

From Table 1, we come to the following conclusions. Firstly, the performance of *CE* are 46.69% and 44.58% in Rank-1 on SYSU-MM01 under the indoor-search mode and all-search mode respectively, which are better than *tri* and *HC*. It demonstrates that the model trained by the loss function based on probability is more effective than the metric learning based on distance for cross-modality person retrieval. Secondly, the combination of *CE* and *tri* achieves 51.63% in Rank-1 and 43.74% in mAP under indoor-search mode and all-search mode respectively, which proves that the cooperation between probability-based CE loss and distance-based metric function could guide the model to learn more discriminative features. Finally, the integration of three cross-modality loss functions fully improve the performance. It proves that three different loss functions could optimize the features from different aspects.

4 Conclusion

In this review, we have introduced the CE loss, the triplet loss, and the HC loss that are typical loss functions used in cross-modality person retrieval. Meanwhile, we perform the experiments to evaluate the performance of them repectively. The experimental results on a common-used dataset have demonstrated that the cross-modality loss functions are useful to alleviate the intra-class modality variations and the cross-modality discrepancy for cross-modality person retrieval.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant No. 62171321, in part by the Natural Science Foundation of Tianjin under Grant No. 22JCQNJC00010, and in part by the Scientific Research Project of Tianjin Educational Committee under Grant No. 2022KJ011.

References

1. Zheng, Z., Zheng, L., Yang, Y.: A discriminatively learned CNN embedding for person re-identification. ACM Trans. Multimedia Comput. Commun. Appl. **14**(13), 1–20 (2018)
2. Sun, Y., Zheng, L., Yang, Y., Tian, Q., Wang, S.: Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline). In: Proceedings of the European Conference on Computer Vision, pp. 480-496 (2018)
3. He, T., Shen, X., Huang, J., Chen, Z., Hua, X.S.: Partial person re-identification with part-part correspondence learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9105-9115 (2021)
4. Cheng, D., Gong, Y., Zhou, S., Wang, J., Zheng, N.: Person re-identification by multi-channel parts-based CNN with improved triplet loss function. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1335-1344 (2016)
5. Hermans, A., Beyer, L., Leibe, B.: In defense of the triplet loss for person re-identification. arXiv preprint arXiv:1703.07737 (2017)
6. Guo, Y., Cheung, N.M.: Efficient and deep person re-identification using multi-level similarity. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2335-2344 (2018)
7. Tian, X., Zhang, Z., Lin, S., Qu, Y., Xie, Y., Ma, L.: Farewell to mutual information: variational distillation for cross-modal person re-identification. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1522-1531 (2021)
8. Chen, Y., Wan, L., Li, Z., Jing, Q., Sun, Z.: Neural feature search for RGB-infrared person re-identification. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 587-597 (2021)
9. Wu, Q., Dai, P., Chen, J., Lin, C.W., Wu, Y., Huang, Y.F.: Discover cross-modality nuances for visible-infrared person re-identification. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4330-4339 (2021)
10. Park, H., Lee, S., Lee, J., Ham, B.: Learning by aligning: visible-infrared person re-identification using cross-modal correspondences. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 12046-12055 (2021)
11. Ye, M., Lan, X., Leng, Q., Shen, J.: Cross-modality person re-identification via modality-aware collaborative ensemble learning. IEEE Trans. Image Process. **29**, 9387–9399 (2020)
12. Cheng, D., Li, X., Qi, M., Liu, X., Chen, C., Niu, D.: Exploring cross-modality commonalities via dual-stream multi-branch network for infrared-visible person re-identification. IEEE Access **8**, 12824–12834 (2020)
13. Cheng, Y., Li, X., Xiao, G., Ma, W., Gou, X.: Dual-path deep supervision network with self-attention for visible-infrared person re-identification. In: Proceedings of the IEEE International Symposium on Circuits and Systems, pp. 1-5 (2021)
14. Wu, A., Zheng, W.S., Yu, H.X., Gong, S., Lai: RGB-infrared cross-modality person re-identification. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 5380-5389 (2017)
15. Ye, M., Wang, Z., Lan, X., Yuen, P.C.: Visible thermal person re-identification via dual-constrained top-ranking. In: Proceedings of the International Joint Conference on Artificial Intelligence, pp. 1092-1099 (2018)

16. Ling, Y., Luo, Z., Lin, Y., Li, S.: A multi-constraint similarity learning with adaptive weighting for visible-thermal person re-identification. In: Proceedings of the International Joint Conference on Artificial Intelligence, pp. 845-851 (2021)
17. Zhu, Y., Yang, Z., Wang, L., Zhao, S., Hu, X., Tao, D.: Hetero-center loss for cross-modality person re-identification. Neurocomputing **386**, 97–109 (2020)
18. Sun, J., Li, Y., Chen, H., Peng, Y., Zhu, X., Zhu, J.: Visible-infrared cross-modality person re-identification based on whole-individual training. Neurocomputing **440**, 1–11 (2021)
19. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778 (2016)
20. Luo, H., et al.: A strong baseline and batch normalization neck for deep person re-identification. arXiv preprint arXiv:1906.08332 (2019)
21. Ye, M., Lan, X., Li, J., Yuen, P.C.: Hierarchical discriminative learning for visible thermal person re-identification. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 7501-7508 (2018)
22. Ye, M., Lan, X., Wang, Z., Yuen, P.C.: Bi-directional center-constrained top-ranking for visible thermal person re-identification. IEEE Trans. Inf. Forensics Secur. **15**, 407–419 (2020)
23. Liu, H., Chai, Y., Tan, X., Li, D., Zhou, X.: Strong but simple baseline with dual-granularity triplet loss for visible-thermal person re-identification. IEEE Signal Process. Lett. **28**, 653–657 (2021)
24. Ye, M., Shen, J., Lin, G., Xiang, T., Shao, L., Hoi, S.C.: Deep learning for person re-identification: a survey and outlook. IEEE Trans. Pattern Anal. Mach. Intell. **44**(6), 2872–2893 (2021)
25. Zhang, L., Du, G., Liu, F., Tu, H., Shu, X.: Global-local multiple granularity learning for cross-modality visible-infrared person reidentification. IEEE Trans. Neural Netw. Learn. Syst. 1-11 (2021)
26. Ye, M., Shen, J., Crandall, D.J., Shao, L., Luo, J.: Dynamic dual-attentive aggregation learning for visible-infrared person re-identification. In: Proceedings of the European Conference on Computer Vision, pp. 229-247 (2020)
27. Hu, X., Zhou, Y.: Cross-modality person ReID with maximum intra-class triplet loss. In: Proceedings of the Chinese Conference on Pattern Recognition and Computer Vision, pp. 557-568 (2020)
28. Han, C., Pan, P., Zheng, A., Tang, J.: Cross-modality person re-identification based on heterogeneous center loss and non-local features. Entropy **23**(7), 919 (2021)
29. Ketkar, Nikhil: Stochastic gradient descent. In: Deep Learning with Python, pp. 111–130. Apress, Berkeley, CA (2017). https://doi.org/10.1007/978-1-4842-2766-4_8

Channel Attention Cloud Detection Network for Ground-Based Cloud Detection

Jiafeng Zhang[1,2], Longxu Zhang[1,2], Jian Li[1,2], Zhong Zhang[1,2], and Shuang Liu[1,2(✉)]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin, China
shuangliu.tjnu@gmail.com

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

Abstract. Cloud detection is indispensable in ground-based cloud observation, and it can implement automatic cloud cover estimation. Cloud detection is quite challenging because of blurred cloud boundaries and variable shapes. In this paper, we propose a new network Channel Attention Cloud Detection Network (CACDN) for ground-based cloud detection. The proposed CACDN is an encoder-decoder architecture, and we design the cloud channel attention (CCA) module to filter information for accurate cloud detection. We conduct the experiments on TLCDD, and the experimental results show that our method achieves better results than other methods, thus proving the effectiveness of the proposed CACDN.

Keywords: ground-based cloud detection · CNN · encoder · decoder

1 Introduction

As a product of nature, clouds largely reflect the hydrological cycle and heat balance of the local area. Hence, accurate cloud observation is necessary for environmental monitoring, etc., but it is challenging due to the numerous weather and variable clouds [1,2]. Cloud observation is generally categorized into two types: ground-based cloud observation and satellite-based cloud observation [3]. Compared with satellite-based cloud observation, ground-based cloud observation can flexibly acquire characteristics of the local cloud bottom, i.e., cloud cover estimation, cloud type, and cloud bottom height [4].

Recently, ground-based cloud observation mainly relies on meteorologists, which may result in inconsistent when facing the same observation. To deal with such situations, many researchers [5,6] work on developing automatic ground-based cloud observation methods for cloud base height measurement, cloud type classification and cloud cover estimation. In this paper, we focus on automatic cloud cover estimation for ground-based cloud observation.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 290–296, 2023.
https://doi.org/10.1007/978-981-99-2653-4_36

Automatic ground-based cloud cover estimation can be accomplished by employing cloud detection technology. Cloud detection predicts each pixel of the input ground-based cloud image as a cloud or sky label. However, the blurred boundaries and the variety of clouds make cloud detection quite challenging. Hence, a lot of methods [7,8] are proposed for ground-based cloud detection. These methods generally are categorized into two types: traditional methods and deep learning methods.

Since the light scattering leads to different colors of clouds and sky, most traditional methods are based on thresholding of color features for cloud detection. For example, Kreuter *et al.* [9] utilized a red to blue ratio (R/B) threshold of 0.77 for cloud detection. Souzaecher *et al.* [10] chose B-R as the threshold for accurate cloud detection. Shi *et al.* [11] adopted the superpixel segmentation method that incorporates texture features as complementary features for cloud detection. Despite the progressive improvement in performance, the results are not satisfactory in real applications.

Since convolutional neural networks (CNNs) [12–14] have good characterization capability, CNN-based methods become mainstream for cloud detection. Dev *et al.* [15] applied convolutional layers, pooling layers and upsampling to construct a typical CNN-based encoder-decoder architecture for cloud detection called CloudSegNet. Afterwards, Xie *et al.* [16] proposed SegCloud network that applies pooling indices to increase the interaction between the encoder and decoder. Shi *et al.* [17] proposed a deep learning method named EFCN that employs histogram equalization to improve the detection performance.

In the paper, we propose a new network called Channel Attention Cloud Detection Network (CACDN). The CACDN is composed of encoder, decoder and the cloud channel attention (CCA) module. The proposed CACDN utilizes the encoder to learn the representation features. The CCA module is designed to model the interaction dependencies between channels of representation features. In this way, we can learn the importance on each channel of the feature maps, which emphasizes the useful information and suppresses the irrelevant information. Finally, the decoder achieves pixel-level classification by using the representation features from the CCA module output for implementing cloud detection.

The CACDN applies depth wise convolutional layer (DWConv) instead of convolutional layer for feature extraction. DWConv reduces the number of parameters in the network, thus improving the efficiency of inference. DWConv also helps the CCA module to conveniently capture channel information of feature maps. We use GeLU [18] as the activation function, so that better performance can achieve by introducing random regularization. We conduct the experiments to verify the effectiveness of CACDN on TLCDD [19], and the experimental results demonstrate that CACDN outperforms other state-of-the-art methods.

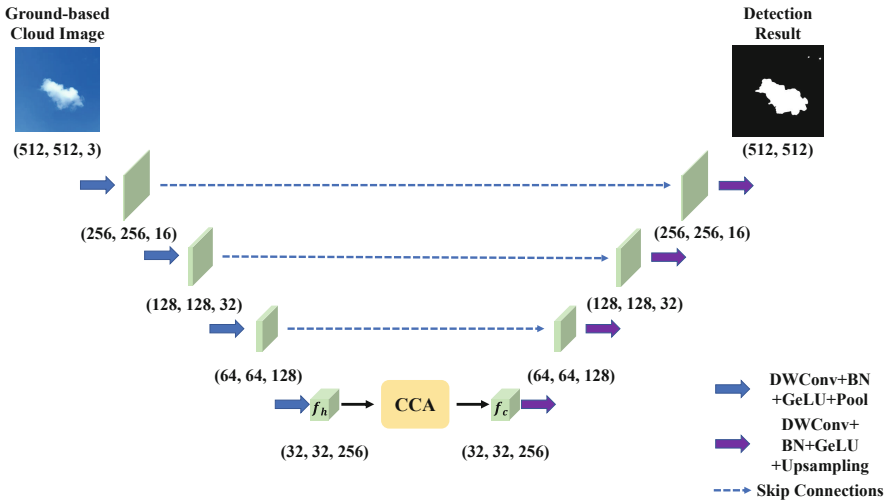


Fig. 1. The framework of the proposed CACDN.

2 Approach

We propose CACDN to effectively interact information between channels of feature maps for cloud detection. The architecture of CACDN is shown in Fig. 1. In this section, we describe the compositions of the proposed CACDN in detail, including the encoder, the CCA module and the decoder.

2.1 Encoder

The encoder is designed to extract abstract information from the input ground-based cloud image for learning representation features. The encoder of CACDN consists of four depth wise convolutional layers (DWConv), batch normalization layer (BN), Gaussian error linear unit activation function (GeLU) and maximum pooling layer (Maxpooling).

We first apply the DWConv, which reduces the number of parameters of the network. It can accelerate the speed of training and learn contextual information. The filter size of DWConv is 3×3 with the stride of 1 and padding of 1. Then, we apply BN to accelerate the convergence speed of the network and prevent over-fitting. Furthermore, we add GeLU after BN to enhance the network representation capability. Finally, we use the Maxpooling filter with the size of 2×2 and stride of 2 to extract local significant features. In addition, this operation reduces the resolution of feature maps $f_h \in \mathbb{R}^{32\times32\times256}$ to increase the receptive field in the encoder.

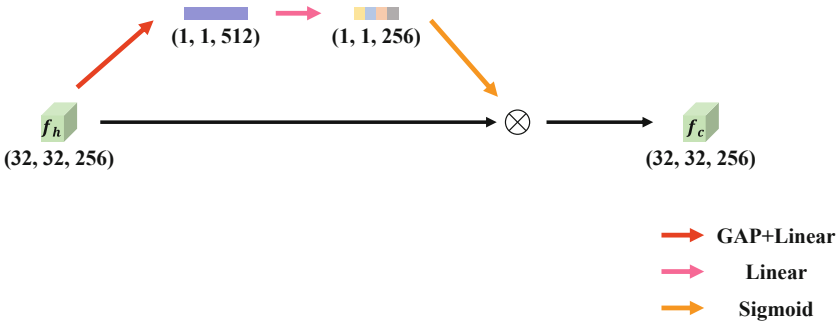


Fig. 2. The flowchart of the cloud channel attention module.

2.2 Cloud Channel Attention Module

CNNs extract different semantic information from cloud images and distribute them to different channels of feature maps. However, not all channel information is beneficial for cloud detection, so we utilize the channel attention [20] for filtering channel information, called the cloud channel attention (CCA) module. The CCA module can weight different channels, so it enhances useful information and reduces noisy information.

Figure 2 shows the framework of the CCA module. Specifically, we first use global average pooling (GAP) and linear layers to change the size of the features to $1 \times 1 \times 512$. Then, we apply the linear layer and the Sigmoid to acquire the weight matrix $1 \times 1 \times 256$. Finally, we perform channel-wise multiplication between the weight matrix and the input f_h to obtain the CCA module output $f_c \in \mathbb{R}^{32\times32\times256}$.

2.3 Decoder

The decoder implements the pixel-level classification with representation features f_c. The decoder of CACDN consists of four DWConv, BN, GeLU, upsampling and three skip connections. For representation features f_c, we apply upsampling by a factor of 2, which can enlarge the size of the feature map. Here, we use the bilinear interpolation for smoother results. We also employ skip connections to combine shallow and deep features to learn contextual information.

Specifically, we take f_c as the input of the decoder. We apply DWConv and upsampling to reduce the number of channels and expand the size of feature maps, and then the feature maps are combined with shallow feature maps from the encoder to learn contextual information by skipping connections.

After four such operations, we get the features with the same size as the input image. This feature after the Softmax operation yields a 512×512 detection result which only contains the pixel labels of the cloud or the sky.

3 Experimental Results and Analysis

In this section, we evaluate the performance of CACDN on a publicly available ground-based cloud detection database. We first describe the database. Then, we illustrate experimental implementation details. Finally, we compare our proposed method with other state-of-the-art methods to demonstrate the effectiveness of our proposed method.

3.1 Database

The TLCDD [19] is a large-scale ground-based cloud detection database. This database is collected over two years in nine provinces. Therefore, it has the diversity of cloud samples, which makes the experimental results more convincing. The TLCDD contains 5000 cloud samples and the corresponding manually labeled ground-truth cloud masks, where we divide into 4208 training samples and 792 test samples. The cloud image of the database is 512×512 and is stored in the PNG format.

Table 1. Comparison with state-of-the-art methods.

Methods	Pre	Rec	F_s	Acc	IoU
B/R (Otsu) [22]	55.98	77.48	57.26	67.72	45.39
B-R (Otsu) [22]	57.91	61.47	50.80	66.92	38.34
FCN [14]	63.20	73.77	57.00	66.49	46.75
CloudSegNet [15]	64.46	77.61	57.79	64.59	47.78
U-Net [13]	68.80	80.43	67.32	74.13	58.16
SegCloud [16]	68.35	81.50	66.95	73.06	57.76
CACDN	**68.39**	**81.67**	**69.18**	**82.55**	**59.79**

3.2 Implementation Details

We set the total training epoch to 50 in all experiments. For the optimizer, we adopt Adam [21] with an initial learning rate of 0.001 and a weight decay of 0.0001. Furthermore, we employ binary cross-entropy (BCE) loss as the objective function.

To quantitatively evaluate different methods, we use five evaluation criteria, namely Precision (Pre), Recall (Rec), F-score (F_s), Accuracy (Acc) and Intersectionality (IoU).

3.3 Comparisons with State-of-the-Art Methods

We compare the proposed CACDN with other state-of-the-art ground-based cloud detection methods, i.e. traditional methods and deep learning methods.

The comparison results of different methods on TLCDD are shown in Table 1. From Table 1, we can draw several conclusions. First, our proposed CACDN achieves the best results on all five evaluation criteria. Second, deep learning methods generally outperform traditional methods for ground-based cloud detection. Traditional thresholding methods are easily influenced by illumination, while deep learning methods can learn more robust representational features.

4 Conclusion

In this paper, we have proposed CACDN for ground-based cloud detection. Specifically, we employ the encoder to mine semantic information from the input ground-based cloud image. Then, we utilize the CCA module to filter the features and get more discriminative representation features. Finally, we design the decoder to upsample the representation features for pixel-level classification so as to conduct ground-based cloud detection. We have conducted a number of experiments on TLCD, and the experiment results demonstrate the effectiveness of the proposed CACDN.

Acknowledgement. This work was supported in part by the National Natural Science Foundation of China under Grant No. 62171321, in part by the Natural Science Foundation of Tianjin under Grant No. 22JCQNJC00010, and in part by the Scientific Research Project of Tianjin Educational Committee under Grant No. 2022KJ011.

References

1. Zhang, Z., Yang, S., Liu, S., Xiao, B., Cao, X.: Ground-based cloud detection using multiscale attention convolutional neural network. IEEE Geosci. Remote Sens. Lett. **19**, 1–5 (2021)
2. Liu, S., Duan, L., Zhang, Z., Cao, X., Durrani, T.S.: Multimodal ground-based remote sensing cloud classification via learning heterogeneous deep features. IEEE Trans. Geosci. Remote Sens. **58**(11), 7790–7800 (2020)
3. Shi, C., Zhou, Y., Qiu, B., Guo, D., Li, M.: Cloudu-Net: a deep convolutional neural network architecture for daytime and nighttime cloud images' segmentation. IEEE Geosci. Remote Sens. Lett. **18**(10), 1688–1692 (2020)
4. Ye, L., Cao, Z., Xiao, Y.: Deepcloud: ground-based cloud image categorization using deep convolutional features. IEEE Trans. Geosci. Remote Sens. **55**(10), 5729–5740 (2017)
5. Taravat, A., Del Frate, F., Cornaro, C., Vergari, S.: Neural networks and support vector machine algorithms for automatic cloud classification of whole-sky ground-based images. IEEE Geosci. Remote Sens. Lett. **12**(3), 666–670 (2014)
6. Allmen, M.C., Kegelmeyer, W.P., Jr.: The computation of cloud-base height from paired whole-sky imaging cameras. J. Atmos. Oceanic Technol. **13**(1), 97–113 (1996)
7. Dev, S., Lee, Y.H., Winkler, S.: Color-based segmentation of sky/cloud images from ground-based cameras. IEEE J. Selected Top. Appl. Earth Observ. Remote Sens. **10**(1), 231–242 (2016)

8. Long, C.N., Sabburg, J.M., Calbó, J., Pagès, D.: Retrieving cloud characteristics from ground-based daytime color all-sky images. J. Atmos. Oceanic Technol. **23**(5), 633–652 (2006)
9. Kreuter, A., Zangerl, M., Schwarzmann, M., Blumthaler, M.: All-sky imaging: a simple, versatile system for atmospheric research. Appl. Optics **48**(6), 1091–1097 (2009)
10. Souza-Echer, M.P., Pereira, E.B., Bins, L., Andrade, M.: A simple method for the assessment of the cloud cover state in high-latitude regions by a ground-based digital camera. J. Atmos. Oceanic Technol. **23**(3), 437–447 (2006)
11. Shi, C., Wang, Y., Wang, C., Xiao, B.: Ground-based cloud detection using graph model built upon super pixels. IEEE Geosci. Remote Sens. Lett. **14**(5), 719–723 (2017)
12. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
13. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
14. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
15. Dev, S., Nautiyal, A., Lee, Y.H., Winkler, S.: Cloudsegnet: a deep network for nychthemeron cloud image segmentation. IEEE Geosci. Remote Sens. Lett. **16**(12), 1814–1818 (2019)
16. Xie, W., et al.: Segcloud: a novel cloud image segmentation model using a deep convolutional neural network for ground-based all-sky-view camera observation. Atmos. Measurement Tech. **13**(4), 1953–1961 (2020)
17. Shi, C., Zhou, Y., Qiu, B., He, J., Ding, M., Wei, S.: Diurnal and nocturnal cloud segmentation of all-sky imager (ASI) images using enhancement fully convolutional networks. Atmos. Measurement Tech. **12**(9), 4713–4724 (2019)
18. Hendrycks, D., Gimpel, K.: Gaussian error linear units (gelus). arXiv preprint arXiv:1606.08415 (2016)
19. https://github.com/zhongzhang8848/TJNU-Large-Scale-Cloud-Detection-Database
20. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7132–7141 (2018)
21. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
22. Yang, J., Lu, W., Ma, Y., Yao, W.: An automated cirrus cloud detection method for a ground-based cloud image. J. Atmos. Oceanic Technol. **29**(4), 527–537 (2012)

Design and Application of Fault-Tolerant On-Board Computer System with High Reliability

Yukun Chen(✉), Dezhi Zhang, Gang Rong, Xu Wang, and Feng Qiu

China Academy of Launch Vehicle Technology, Beijing 100076, China

Abstract. On-board computer system has a significant role in spacecraft electronic system, and its reliability is especially essential to achieve final mission. In order to still work normally when on-board computer has failure, system architecture, switch method and estimation rule were introduced. On the basis of preserving the state signal between main computer and switch circuit, the state signal between backup computer and switch circuit, and the additional state signal between main computer and backup computer was adopted, then presented a modified independence switch method. By using modified independence switch method, independence switch function still worked normally when independence switch module had failure. Comparator was implemented by adopting software vote and software switch approach, and it could eliminate hardware comparator's key failure. Results indicated that the redundant technology could effectively improve the reliability of space on-board computer system. The scheme has engineering application value for design and application of space on-board computer system with high reliability.

Keywords: Dual Redundancy · Fault-tolerance · On-board Reliability

1 Introduction

With the development of aerospace technique, the spacecraft reliability and security must meet the needs of task. On-board computer system under space atmosphere is influenced by plasma, energetic charged particles, earth magnetic field, solar electromagnetic radiation, meteoroid, and so on, which will degrade performance. On-board computer system in space orbit has the feature of unmaintainability except for space station [1]. Aircraft mission may fail when On-board computer system has failure. Fault tolerance technology has become a urgent topic for on-board computer system to increase reliability.

2 The Dual Redundant Hardware Architecture and Switching Strategy

Backup fault tolerance architecture is common means for fault tolerance system architecture. It has four types, includes cold standby, warm standby, hot standby and duplex mode [2]. Redundancy will increase additional costs. According to performance and reliability, on-board computer system adopts multiple scheme based on flight sequence.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 297–304, 2023.
https://doi.org/10.1007/978-981-99-2653-4_37

2.1 The Fault Tolerance Architecture of Dual Cold Standby

When aircraft is under station control or stable self-control, the scheme of one hot standby and one cold standby is appropriate to insure working life and decrease power consumption [3]. The typical fault tolerance architecture of dual cold standby includes two same sets of processor and multiple I/O, as depicted in Fig. 1. Fault tolerance module manages the switch between dual on-board computer. When host has failure, control right is transferred from host to backup. When backup also has failure while host don't recover, control right is transferred from backup to emergency module.

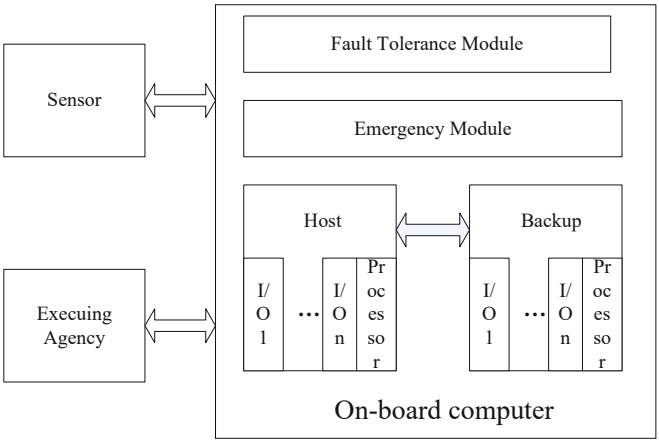


Fig. 1. The typical fault tolerance architecture of dual cold standby

Three factor should be taken into account, one is failure test, two is switch from host to backup, and finally is state recovery. Failure test is indispensable base for fault tolerance architecture of dual cold standby [4]. Failure test includes many methods, such as system self-test, program repeating, choosing two from three data sector, and watchdog technology. Majority fault can be discovered through system self-test, and the key point is that system self-test must work normally.

2.2 The Fault Tolerance Architecture of Dual Hot Standby

To insure aircraft wok normally and deal with fault quickly, on-board computer adopts the architecture of dual hot standby in initial attitude setting phase. When host has failure, control right is transferred from host to backup by commands or autonomous discrimination [5].

Host and backup both have power supply module separately. Control right can be achieved only by either host or backup at the same time. In the hot standby mode, both host and backup can accept system input signal. Processing results of the computer that has control right are chosen as system output through switching circuit. When command centre discovers a computer has severe failure and cannot work, the power supply can be shut down through remote control or autonomous switching circuit. Fault isolation

circuit deletes the computer that has failure. Closing failure computer cannot influence the other computer. Figure 2 shows the typical fault tolerance architecture of dual hot standby.

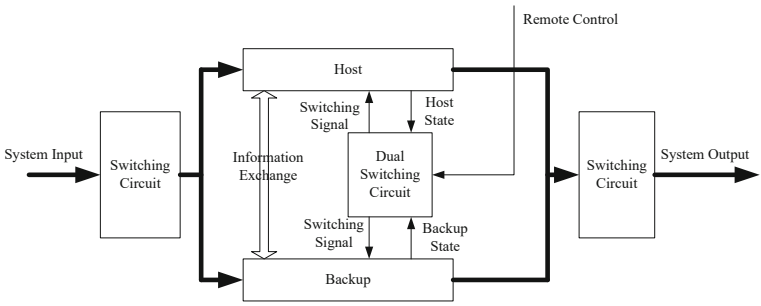


Fig. 2. The typical fault tolerance architecture of dual hot standby

Dual switching circuit is monitored by host and backup. It has timing monitor and corresponding logic circuit, designed by watchdog mechanism [6]. Dual switching circuit has two triggers. Host and backup reset corresponding trigger, but timing signal inside the dual switching circuit set two triggers. If reset signal don't appear before timing signal coming, switching signal will be generated automatically, then the computer working normally will be on duty. If host and backup are both working normally, switching signal is not generated, and host will be on duty. If host and backup are both out of work, host will be on duty, then at least there is still a guarding computer to avoid switching frequently between host and backup.

2.3 The Switching Implementation Mode Between Dual Computer

When aircraft is in the orbit, control right can be switched between dual on-board computer through remote control and switching autonomously [7]. 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.

2.4 The Modified Autonomous Switching Strategy Between Dual Computer

To avoid logic estimation failure between normal computer and faulty computer when autonomous switching module has hardware malfunction. The typical fault tolerance

architecture of dual hot standby is optimized. On the base of host state signal and backup state signal, the working state signal was presented. Autonomous switching could still be achieved if autonomous switching module has hardware failure. It can improve system redundancy. The principle was depicted as follows: host sent regularly its normal state signal to backup under regular condition, but backup could not receive normal state signal when host had failure. Backup could estimate that whether host worked normally through dual computer communication port. If backup worked normally and found failure on host, it would transmit on duty pulse to get control right. System architecture would be reconfigurable. Autonomous switching could be achieved when autonomous switching module inside dual computer switching circuit had failure, and it could tolerance a fault on autonomous switching module. Figure 3 illustrates the modified fault tolerance architecture of dual hot standby.

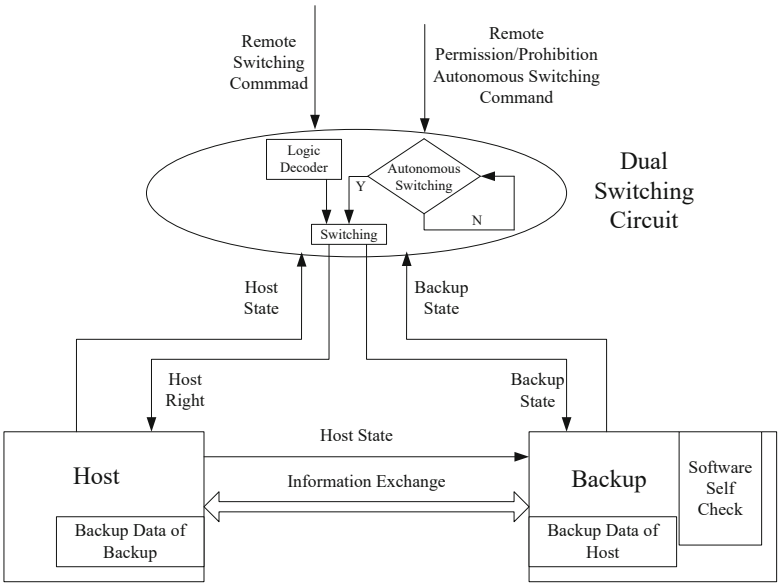


Fig. 3. The modified fault tolerance architecture of dual hot standby

There were separate cache in host and backup computer for exchanging data each other. Host sent its data to cache of backup, while backup sent its data to cache of host. Host and backup had the same component. Figure 4 shows the principle block diagram for dual computer communication.

Figure 5 displays the data flow diagram for dual computer communication. If M represents host, then N represents backup. Whereas if M represents backup, then N represents host.

3 The Design of Comparator in Dual Duplex Mode Architecture

When taking into account influence on system availability caused by success ratio and time of failure judgement, duplex system has greater availability than hot standby and

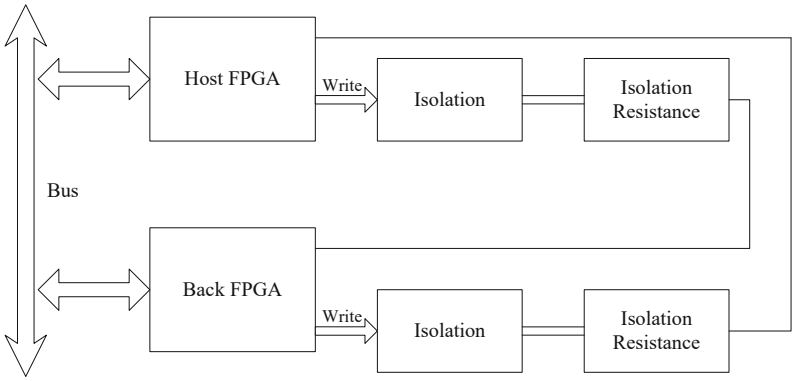


Fig. 4. The principle block diagram for dual computer communication

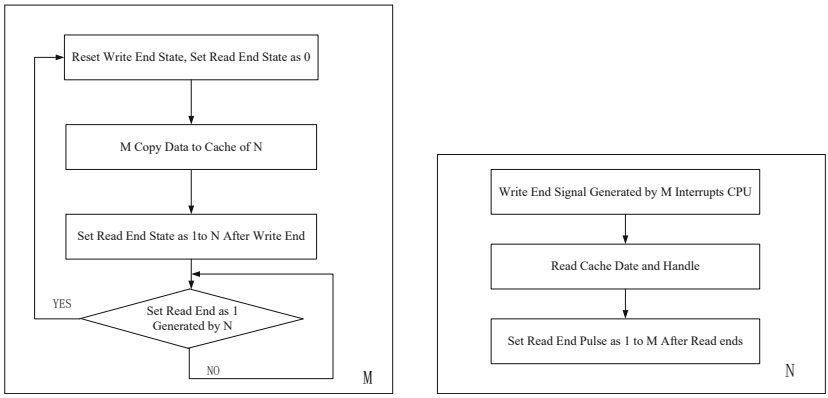


Fig. 5. The data flow diagram for dual computer communication

warm standby system under certain conditions. Besides duplex system has no switch problem, it is suitable to run under real time condition for crucial task, but it increases power consumption and has additional comparison circuit [8]. Dual duplex mode will output the comparison result of host and backup, so comparator is the crucial component for redundant system of dual duplex mode.

3.1 Hardware Design

Comparator is achieved by hardware in common redundant system of dual duplex mode. Hardware comparison module consists of comparison circuit and detection circuit executing agency [9]. Comparator adopts logic circuit for low redundancy level, and adopts independent processor system for high redundancy level. Figure 6 illustrates architecture for single comparator. If comparator has failure, it cannot generate detection signal or indicate faulty output, so the reliability of comparator becomes the new key single point for redundant system. To solve comparator's single point of failure, the problem can be relieved by increasing comparator's redundancy. Figure 7 depicts the architecture

for dual comparators. However when detection and switching circuit for multiple comparators should be introduced, hardware will become further complex. More and more redundancy will decrease the whole system reliability.

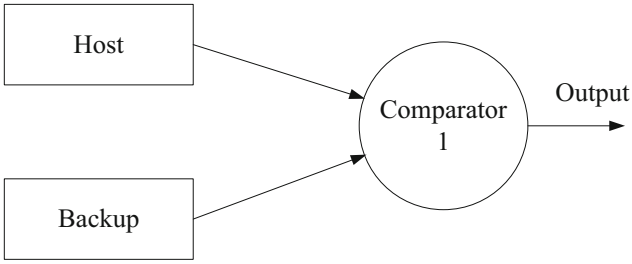


Fig. 6. The architecture for single comparator

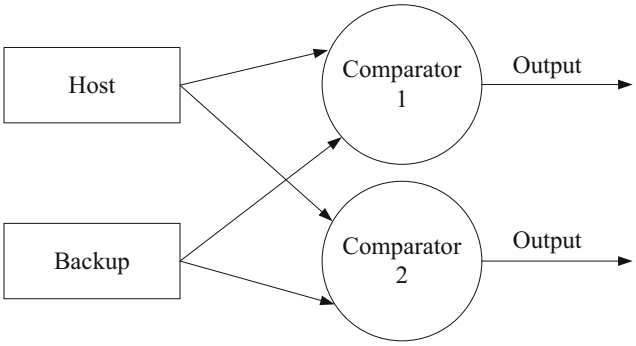


Fig. 7. The architecture for dual comparators

3.2 Software Design

To avoid comparator's disadvantage and increase reliability for real time and embedded on-board computer system, the design for solving comparator's reliability was presented. Based on the scheme of software voting and software switching, hardware unit for comparator was abandoned and comparator was implemented as software. The design was analyzed and detected as a part of system resource, which could solve the reliability problem caused by alone comparator detection, therefore the system could be optimized. From principle analysis, it was available that comparator belonged to system resource. First, comparator had less process load and simpler category, and software reliability could be guaranteed after testing, so it had low probability of leading to failure in system. On the other hand, hardware lifetime is limited according to reliability theory. Software reliability is almost invariable once put into use, so reliability of comparator can be guaranteed. Figure 8 illustrates the determination flow diagram of software for comparator.

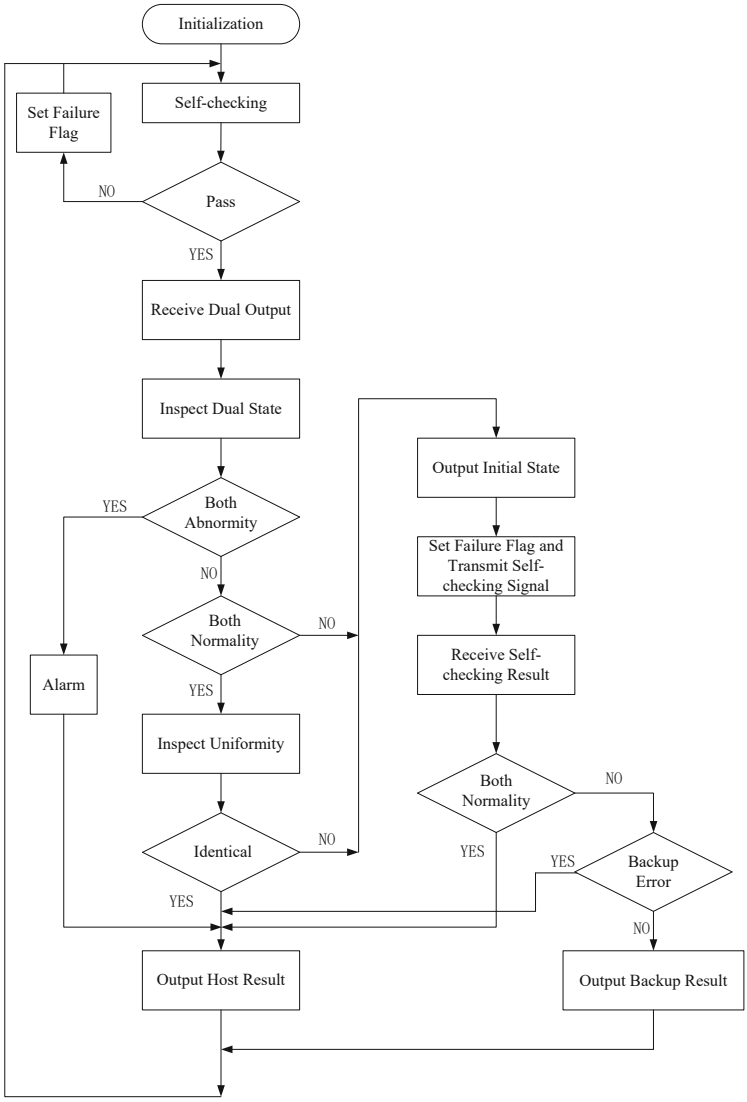


Fig. 8. The determination flow diagram of software for comparator

4 Conclusions

The reliability and security of on-board computer is the crucial topic for aircraft. The paper introduced architecture of fault-tolerant on-board computer system with redundancy function, and analyzed the dual switching strategy and determination criterion, finally presented a modified autonomous switching strategy and software suitable for comparator. Practice indicates that the redundant design improvement can effectively

enhance the performance 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.

References

1. Yang, M., Hua, G., Feng, Y.: Fault Tolerance Techniques for Spacecraft Control Computer. National Defense Industry Press, Beijing (2014)
2. Sun, X., Chen, Z., Gu, Y.: Research on fault-tolerant flight control computer system based on dynamic reconfiguration. J. Syst. Simul. **30**(10), 3957–3963 (2018)
3. Yu, Y., Wang, H.: Deep Learning-based Reentry Predictor-corrector Fault-tolerant Guidance for Hypersonic Vehicles. ACTA ARMAMENTARII **41**(4), 659–665 (2020)
4. Jiang, B., Zhang, K., Yang, H.: Fault-tolerant control of satellite attitude control systems. Acta Aeronautical et Astronautica Sinica **42**(11), 524662 (2021)
5. Wang, Y., Wen, X.: Research status and progress of fault diagnosis technology for spacecraft. Aero Weaponry **23**(5), 71–76 (2016)
6. Xu, A., Xia, D., Zheng, J.: The study of fault tolerance technical in civil aircrafts slat flap control computer. Microelectron. Comput. **32**(6), 36–40 (2015)
7. Xiao, A., Hu, M.: Reliability analysis of the computer with quad-modular redundancy byzantine fault tolerant. Aerosp. Control Appl. **40**(3), 41–46 (2014)
8. Lv, Y.: A fault-tolerant method for space computer memory with low-cost and high-reliability. Aerosp. Control Appl. **46**(3), 66–70 (2020)
9. Wang, Z., Cheng, S.F., Ma, X.B.: Design and implementation of highly reliable fault-tolerant computer with integrated multi-task. Aeronaut. Comput. Tech. **50**(4), 111–112 (2020)

Query Algorithm for Close Contacts of Epidemic Prevention and Control Based on Big Data Analysis

He Dong(✉), Baoguo Yu(✉), Wanqing Wu, Qingwu Yi, Chenglong He, and Chong Chen

State Key Laboratory of Satellite Navigation System and Equipment Technology, The 54th Research Institute of CETC, Shijiazhuang 050081, China
18845144376@163.com

Abstract. Aiming at the phenomenon that the novel coronavirus pneumonia virus is easy to spread in large-scale indoor venues with dense people in recent years, we proposed a method of querying close contacts for epidemic prevention and control in large-scale indoor venues. Firstly, we obtain the location information of all personnel in the indoor venue at all times and store them in a time sequence database. Secondly, determine the personnel and the time period that need to be inquired, and divide the time period into several time slices according to the situation of this personnel entering and leaving different areas in the indoor venue. Finally, settle and count the relevant information of other personnel who have close contact with the personnel need to be required in each time slice, summarize the close connection situation of all time slices, and realize the analysis of the close contact situation of the personnel need to be required. Through experimental verification, this invention can quickly and accurately analyze the close connection situation in large-scale indoor venue with dense personnel, and provide strong guarantee for epidemic prevention and control of indoor venues.

Keywords: Large-scale indoor venue · Epidemic prevention and control · close contact inquiry

1 Introduction

With the rampant spread of COVID-19 in recent years, it has become increasingly important to ensure the prevention and control of COVID-19 in large indoor venues with crowded people [1]. Through the analysis of big data, close contacts can be mined, thus providing a strong guarantee for epidemic prevention and control [2, 3]. The analysis of close contact among people in indoor venues is the core guarantee of epidemic prevention and control in indoor venues. When an infected person with COVID-19 or a high-risk person who has had close contact with a confirmed case is found in an indoor venue, other people who have close contact with this person should be screened and quarantined in a timely manner.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 305–312, 2023.
https://doi.org/10.1007/978-981-99-2653-4_38

When people infected with COVID-19 virus are found in indoor venues, the traditional manual flow adjustment method will screen other people who may have close contact with the infected one by one, which not only requires a lot of human and material resources, but also in the case of untimely flow adjustment, the people at risk of infection will not be isolated and controlled quickly, resulting in the further expansion of the spread of the epidemic.

At present, indoor positioning technologies based on wireless signal propagation (such as ultra-wideband [4], Bluetooth [5], Wi-Fi [6], etc.) have become more and more mature. These technologies can be used to accurately locate the people in the indoor venues, obtain the location information of each individual, and store the information in the time-series database. Based on the location information, we proposed a close contact query method for epidemic prevention and control applied to large indoor venues. This method has the characteristics of short calculation time and accurate calculation results. By processing and analyzing the personnel location information stored in the time-series database, it can quickly and accurately analyze and query the close contacts of infected people in large indoor venues.

2 Proposed Method

2.1 Overall Process Framework

The main calculation and analysis process of the close contact query method for epidemic prevention and control applied to large indoor venues we proposed is shown in Fig. 1.

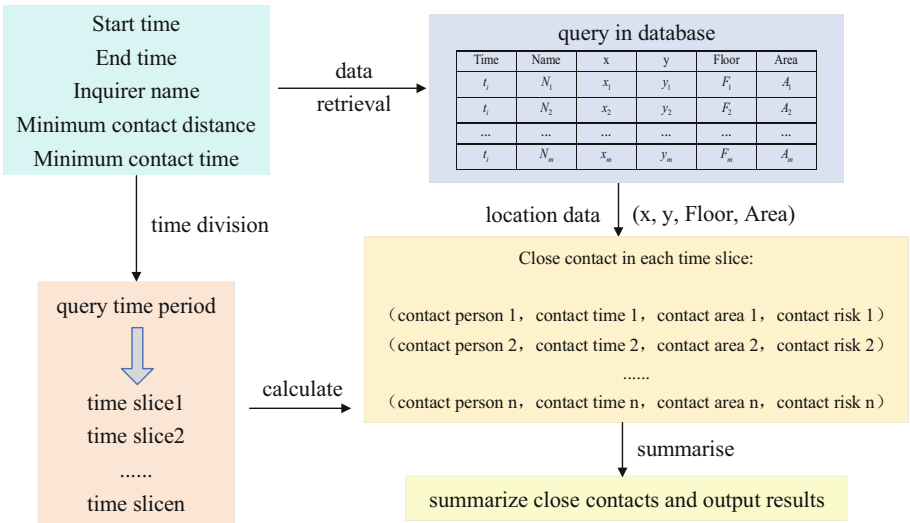


Fig. 1. Overall process framework of the proposed method

In order to accurately determine the contact relationship among people in indoor venues, it is necessary to monitor the position of all people in the venues at all times,

and store theses information in the time-series database. When it is necessary to analyze the people close connection in the venue, we can query the position data stored in the time-series database and perform corresponding calculation and processing.

In our paper, we need to obtain the name of the inquirer and the time period to be inquired when analyzing the close contact of people in the venue. In order to speed up the time of contact analysis, the query time should be divided into several time slices according to the situation of the query person entering and exiting different areas in the venue. Then, for each time slice, it is necessary to calculate and calculate the relevant information of other people who are closely connected with the query person. Finally, we summarize the contact information of all time slices and output the summarized information to realize the personnel contact analysis of indoor venues.

2.2 Establishment of Time-Series Database

People contact analysis of large indoor venues needs to be based on the position data of each person, so the establishment of time-series database is very important. The purpose of the establishment of time-series database is to store the position data of all people in the venue at all times, which can be used to calculate and analyze the personnel contact situation in the venue. The time-series database established in our paper needs to store the position information of all people in the venue at each moment. For the same moment, the data format stored in the temporal sequence database is shown in Table 1.

Table 1. The data format stored in a sequential database

Time	Name	x	y	Floor	Area
t_i	N_1	x_1	y_1	F_1	A_1
t_i	N_2	x_2	y_2	F_2	A_2
...	...	...	...	...	...
t_i	N_m	x_m	y_m	F_m	A_m

Table 1 shows the position data of all people in the venue stored in the time-series database. Assuming there are m persons in the current indoor venue, the time-series database can store the name, current location coordinates, floor and area of each person in the venue at time t_i. At the next moment, the position information of each person will also continue to be stored according to the data format, and these data will also serve as the basis for realizing the close connection analysis of indoor venue personnel.

2.3 Division of Query Time Period

When analyzing the close contact of indoor venues, we need to obtain the name of the person who inquired and the query period. Since people are always on the move in the indoor venues, the inquiring person come in and out of different areas during this period. In order to judge whether there is close contact among people in indoor

venues, it is necessary to calculate the distance between any two people at the same time, which requires a lot of calculation cost and time. Therefore, we proposed the concept of dividing the query time period, which needs to retrieve how many areas the query person have entered and left in the specified time period in the time series database. Then, the query time is divided into several time slices according to the situation of the query person entering and leaving the area, and each time slice corresponds to a region ID. The schematic diagram of query time division is shown in Fig. 2.

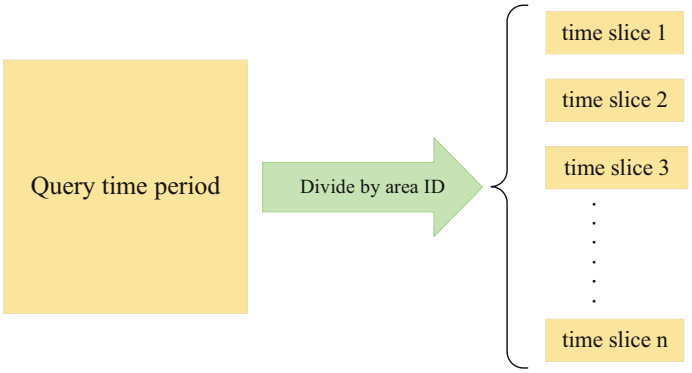


Fig. 2. Divide query time period

After we divide the query time period, other people who are in the same area with the query person in the same time slice can be retrieved in the time-series database. When we calculate the location relationship among people, we only need to calculate with these people one by one, and don't need to calculate other people who are not in the same area at the same time as the query person, which can greatly shorten the calculation time.

2.4 Computing Close Contact Information

When dividing the query time period into several time slices, for each time slice, we first need to obtain the region of the query personnel in this time slice. Then, we filter the name, location coordinates, floor and other location data of other person in the same area as the query person in each time slice in the time-series database. Finally, for each time of the time slice, we need to calculate the distance D between the query person and the filtered person in turn.

$$D = \sqrt{(x - x')^2 + (y - y')^2} \tag{1}$$

where, (x, y) is the coordinates of the inquirer at a certain time in the time slice, and (x', y') is the coordinates of other people in the same area as the inquirer at that time. If the distance between them at this time is less than the distance threshold $D_{\min}$:

$$D <= D_{\min} \tag{2}$$

It indicates that there is a close contact between them. In our paper, the distance threshold $D_{\min}$ is set to 1m, which can be adjusted according to the prevention and control requirements of indoor venues. The smaller the value is, the more rigorous the query of person close contact is.

At this point, we define a list to store the information of other person who have close contact with the query person. We traverse all the times of each time slice. If there is a close contact relationship between another person and the query person at a certain time, we will retrieve whether the information of that person has been stored in the list. If the information is not retrieved, it means that they are in contact for the first time. At this time, we record their current contact time as the initial contact time and their current distance, and set the contact times between them to 1. If the information is retrieved, it indicates that they have been in contact before. At this time, we increase their contact times, and update the current time as their end contact time. In addition, it is also necessary to judge whether their current distance is less than the distance they have previously contacted. If less, we update the distance data.

After traversing each time slice, we need to calculate the names, initial contact time, end contact time, minimum distance, contact times and other information of other people who have close contact with the query person in the time slice, and store the information accordingly.

Finally, we need to summarize the close contacts calculated from all time slices. In our paper, it is set that the distance between the two in the same area is less than the distance threshold $D_{\min}$ (set as 1m in our method) for more than 15 min (900 s), which indicates that there is a close contact relationship between them. Therefore, after summarizing the closeness of all time slices, it is also necessary to screen out people with a total number of closeness more than 900 times, and convert the number of close contacts between the two people into their closeness risk value information. Finally, we output the names of all contact person, initial contact time, end contact time, initial contact area, end contact area, minimum contact distance, contact risk value as the result of contact, so as to realize the analysis of close contact in indoor venues.

3 Verification of Experimental Results

We next experimentally verify the performance of the close contact query method for epidemic prevention and control proposed. When conducting close contact analysis, we first need to input query information, the format of which is shown in Table 2.

The query information in Table 2 includes five parts, namely start time, end time, name of the inquiring person, distance threshold and total contact time. In our paper, the definition of the close relationship between two people is that they are within 1m of each other for more than 15 min. After inputting the query information, the close contact situation of the current venue will be analyzed, and the analysis results will be output in the form of JSON string. The output results are shown in Fig. 3.

The query method for close contacts proposed in our paper can output the information of other people who are closely connected with the query person during the query period after obtaining the information of the query person, query period and search conditions. As can be seen from Fig. 3, when outputting the information of the close contact people,

Table 2. Format of query information

Key	Value
Start_at	2022-02-16 00:00:00
End_at	2022-02-18 00:00:00
User_name	Zhang San
Distance	1
Contact_time	15


```
[
    {
        "name": "ZhangSan",
        "First_time": "2022-07-01 10:56:28",
        "End_time": "2022-07-04 12:56:28",
        "First_area_ID": 1,
        "End_area_ID": 6,
        "Distance": 0.01,
        "risk": 17.71%
    }
    {
        "name": "Wangwu",
        "First_time": "2022-07-02 14:30:00",
        "End_time": "2022-07-03 19:30:00",
        "First_area_ID": 12,
        "End_area_ID": 24,
        "Distance": 0.13,
        "risk": 5.69%
    }
]
```


Fig. 3. Close contact results of calculation output

our method outputs a total of 7 pieces of data, including the name of the close contact people, initial contact time, end contact time, initial contact area, end contact area, shortest contact distance within this time period and the close contact risk value. Using the information, we can accurately obtain information about close contacts and isolate them in a timely manner to prevent the spread of COVID-19.

Finally, it is necessary to verify the real-time and accuracy of the close contact query method proposed in our paper. In our paper, the indoor test environment for the analysis of close contact people is the Winter Olympics Snow Ruyi Stadium. The total number of indoor personnel is about 1,500, and the query time is 2 days. In our paper, the proposed close query method is compared with the traditional artificial flow adjustment method, and the comparison results are shown in Table 3.

Table 3. The performance comparison of the two methods

Method	Retrieval time	Retrieval accuracy
Artificial flow control method	2.4 h	95.3%
Close contact query method	5.6 min	97.6%

From the comparison results in Table 3, it can be seen that the query method for close contacts of epidemic prevention and control proposed in our paper is significantly better than the traditional artificial flow control method in terms of retrieval time and retrieval accuracy, indicating that this method not only can quickly analyze the close connection of people in large indoor venues, but also ensure high calculation accuracy.

4 Conclusion

We proposed a close contact query method for epidemic prevention and control applied to large indoor venues. After experimental verification, our method can quickly and accurately solve and analyze the close contact situation of large indoor venues with dense people, and accurately output the information of close contacts. Using the information of close contacts, high-risk groups can be quickly and accurately identified and isolated, providing a strong guarantee for epidemic prevention and control in indoor venues.

References

1. Wang, J., Zhang, J., Zhou, K., Sun, X.: Analysis and design of epidemic disease monitoring cloud platform. In: 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC), pp. 1989–1992 (2022)
2. Lixin, L., Lin, L.: The big data analysis and mining of people's livelihood appeal based on time series modeling and algorithm. In: 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), pp. 1–5 (2020)
3. Li, H.: Research on big data analysis data acquisition and data analysis. In: 2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA), pp. 162–165 (2021)
4. Zhang, H., Zhang, Z., Zhao, R., Lu, J., Wang, Y., Jia, P.: Review on UWB-based and multi-sensor fusion positioning algorithms in indoor environment. In: 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), pp. 1594–1598 (2021)
5. Sophia, S., Shankar, B.M., Akshya, K.: Bluetooth low energy based indoor positioning system using ESP32. In: 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), pp. 1698–1702 (2021)
6. Wu, W., Yang, B., Yu, H., Wang, H.: High-accuracy WiFi-based 3D indoor positioning using non-uniform planar array. In: 2021 IEEE MTT-S International Wireless Symposium (IWS), pp. 1–3 (2021)
7. Tayeb, H.F., Karabatak, M., Varol, C.: Time series database preprocessing for data mining using Python. In: 2020 8th International Symposium on Digital Forensics and Security (ISDFS), pp. 1–4 (2020)

8. He, P.: Study on epidemic prevention and control strategy of COVID-19 based on personnel flow prediction. In: 2020 International Conference on Urban Engineering and Management Science (ICUEMS), pp. 688–691 (2020)
9. Chen, C., Li, R., Yao, X.: Based on the application of improved Delphi method and analytic hierarchy process in epidemic prevention and control. In: 2021 40th Chinese Control Conference (CCC), pp. 6589–6594 (2021)
10. Ming, J., Zhang, L., Sun, J., Zhang, Y.: Analysis models of technical and economic data of mining enterprises based on big data analysis. In: 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), pp. 224–227 (2018)

Review of Reconstruction Methods for Spectral Snapshot Compressive Imaging

Hao Yuan[1], Xiaoming Ding[1,2(✉)], Qiangqiang Yan[2], Xiaocheng Wang[1], Yupeng Li[1], and Tingting Han[1]

[1] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
xmding@tjnu.edu.cn

[2] CAS Key Laboratory of Spectral Imaging Technology, Xi'an Institute of Optics and Precision Mechanics, Xi'an 710119, China

Abstract. Snapshot compressive imaging (SCI) uses a 2D sensor to obtain higher dimensional data and then reconstructs the underlying high-dimension data by elaborate algorithms. Applying SCI to capture hyperspectral images is known as spectral SCI. Although this technique has been proposed for more than a decade, it has not been widely used, mainly because its reconstruction accuracy and reconstruction speed are not yet satisfactory, which is the research focus on spectral SCI. This paper investigates the literatures on reconstruction methods of spectral SCI, mainly involving coded aperture optimization, model-based reconstruction algorithms and deep learning-based reconstruction algorithms. In this paper, we also provide a summary of studies on noise modeling and denoising for reconstructed spectral SCI data.

1 Introduction

A real-world scene can be modeled as a three-dimensional (3D) data cube (x, y, λ), where (x, y) denotes the spatial dimension and λ denotes the spectral dimension. Spectral imaging technology can capture multiple images of a target scene at different wavelengths in a certain wavelength range to obtain a complete data cube. As the spectral information has a "fingerprint effect", it is used in many fields, such as remote sensing [1], medical diagnosis [2], and mineral detection [3]. However, to acquire a three-dimensional data cube with a two-dimensional (2D) detector, traditional spectral imaging techniques, whether whiskbroom, pushbroom, or staring imagers, require a time-consuming scanning, which limits the application of spectral imaging techniques in dynamic scenes and may introduce motion artifacts that impair the imaging quality. Compressed sensing theory [4] states that if the signal is sparse or compressible (i.e. signal represented in a transform domain is sparse), then it can be reconstructed using compressed measurements obtained with a sampling rate much lower than the Nyquist sampling rate.

In 2007, Gehm M.E. et al. proposed Dual Disperser Coded Aperture Spectral Imaging (DD-CASSI) [5], which applied compressed sensing theory to spectral imaging and achieved the acquisition of a complete data cube in a single exposure time, i.e. spectral snapshot compressive imaging, called spectral SCI here. The following year, Ashwin

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 313–322, 2023.
https://doi.org/10.1007/978-981-99-2653-4_39

Wagadarikar et al. proposed Single Disperser Coded Aperture Spectral Imaging (SD-CASSI) [6], which used only one dispersive element and had a more compact system structure. Currently, spectral SCI systems mostly use CASSI platform or those modified from them, such as the Dual Camera Compressive Hyperspectral Imager (DCCHI) [7]. The DCCHI incorporates a beam splitter in front of the objective lense of CASSI, and one beam enters the CASSI system to obtain scene compressed measurement on the detector, while the other beam is captured by a grayscale camera. Using the grayscale image as supplementary information to the compressed measurement improves the quality of the final reconstructed image. In addition, the captured grayscale image can have other applications, such as aiding coded aperture design [8].

Compressed sensing theory consists of three elements, including the sparse representation of the signal, the measurement matrix, and the reconstruction algorithm. For most signals, they are not sparse by themselves, so it is necessary to find suitable transform basis for them, such as wavelet transform basis, discrete cosine transform basis, Fourier transform basis, etc. Besides, the dictionary learning approach has been used for sparse representation of signals. Restricted Isometry Principle (RIP), proposed and proved by Candes E J, Tao T, Baraniuk R et al. [9–11], gives the criteria to be satisfied by the sensing matrix in order to achieve an accurate reconstruction (see Coded Aperture Optimization). The sparse transform basis is usually fixed, and a good measurement matrix is crucial for the whole compressive measurement and reconstruction process.

Along with compressed sensing theory, many classical reconstruction algorithms have been proposed and used, such as matching tracing algorithm (MP) [12], orthogonal matching tracing algorithm (OMP) [13], and two-step iterative thresholding/shrinking algorithm (TwIST) [14], etc., but these algorithms suffer from the issues of insufficient reconstruction accuracy and high computational complexity. In the study of reconstruction algorithms, model-based reconstruction algorithms usually use prior knowledge or introduce additional information to enhance the quality of reconstructed images. The use of image blocking can reduce the complexity of algorithms and improve the reconstruction speed. In addition, deep learning has powerful feature extraction capability, and numerous deep network models are used for reconstruction of spectral SCI, greatly increasing the speed of reconstruction and enabling real-time reconstruction. Noise is an essential factor affecting imaging quality, and taking denoising in the system into full consideration can improve reconstruction accuracy.

This paper is organized as follows. Section I is the introduction section, which briefly introduces background knowledge of spectral SCI. Section II gives some achievements on improving reconstruction quality by optimizing the coded aperture. Section III introduces model-based reconstruction algorithms. In Section IV, deep learning-based reconstruction algorithms are introduced. Some research results about noise of images are presented in Section V, and finally, the full paper is summarized in Section VI.

2 Coded Aperture Optimization

For spectral SCI, taking SD-CASSI as an example (Fig. 1), the 3D data cube of the input scenes is modulated in the spatial dimension by the coded aperture, then passes through the collimating lense, arrives at the dispersive element (a prism is used here), and is

finally sent to the detector by the reimaging lense, where the coded and compressed measurements are obtained (Fig. 2).

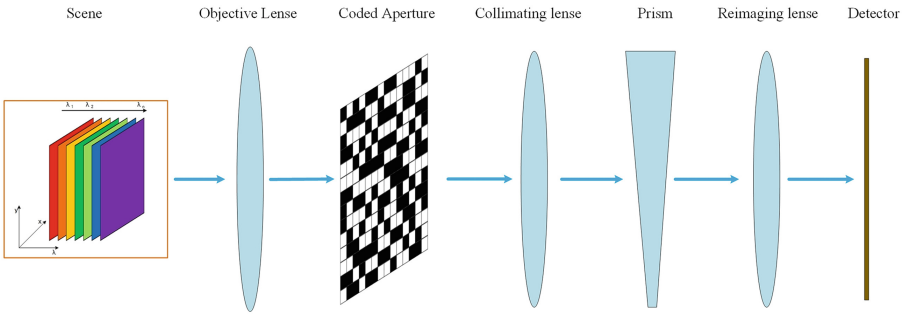


Fig. 1. Sketch of SD-CASSI system

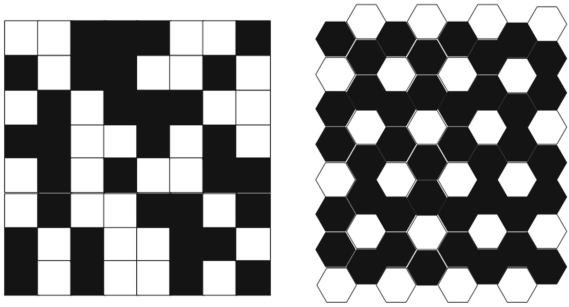


Fig. 2. (left) square coded aperture (right) hexagon coded aperture

The system can be modeled as follows,

$$y = \mathbf{\Phi} x = \mathbf{\Phi}\mathbf{\Psi}\theta = \mathbf{\Theta}\theta \tag{1}$$

where $x = \mathbf{\Psi}\theta$, x denotes the target scene, $\mathbf{\Psi}$ represents the sparse transform basis, and θ is the sparse representation of x on $\mathbf{\Psi}$. $\mathbf{\Phi}$ is the measurement matrix, i.e., the mathematical representation of the coded aperture. $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi}$ is called the sensing matrix. And y is the compressed measurement obtained on the 2D detector.

To exactly solve Eq. (1), the sensing matrix requires to satisfy restricted isometry property (RIP). RIP is formulated as follows, if $\sigma_k \in (0, 1)$, satisfying

$$(1 - \sigma_k)\|\theta\|_2^2 \leq \|\Phi\Psi\theta\|_2^2 \leq (1 + \sigma_k)\|\theta\|_2^2 \tag{2}$$

Then, we say that such a sensing matrix satisfies the RIP condition. Most random measurement matrices can satisfy Eq. (2), such as Gaussian distribution, Bernoulli distribution, uniform distribution, etc.

Compared to the traditional approach of using a series of binary block elements to form an encoding pattern, Henry Arguello et.al extended the compressive capability of

CASSI by using filters as through-light elements of the coded aperture, called colored coded aperture [15]; later, Laura Galvis et al. proposed a moving colored coded aperture that can be moved vertically to obtain two different snapshots, which has the advantage of reducing the number of filters that are required [16]; Considering the serious signal distortion in highly compressed regions when modulating the scene using a randomly distributed coded aperture, Xu Yan et al. proposed a uniformly distributed slit array coding pattern [17]; Zhang Hao et al. abolished traditional squared coded elements and proposed a coded aperture with hexagonal elements, because the shape difference between the coded aperture hexagonal elements and the detector square pixels can introduce equivalent gray-scale spatial modulation, which increases the degree of freedom of the sensing matrix [18]; As mentioned earlier, the grayscale images obtained in a dual camera system can be used to aid the design of the coded aperture pattern, and Ma Xu et al. used the nonlinear thresholding of the grayscale images of the scene to construct an adaptive coded aperture and achieved a purpose of exploiting prior knowledge of the scene [8].

Apart from the above-mentioned independent designs that just focused on the coded aperture, Tao Chenning et al. proposed a simultaneous optimization on the coded aperture and the sparse representation basis [19], specifically, the coded aperture and the sparse dictionary are optimized by using a genetic algorithm and a gradient descent method, respectively, and such simultaneous optimization is performed iteratively to achieve coherence minimization.

The improvement of coded aperture elements, the pattern design of coded masks, the introduction of adaptive methods and the proposal of joint optimization approaches have enriched the forms of coded apertures, but the constant is that the optimization of coded apertures is always guided by RIP.

3 Model-Based Reconstruction Algorithms

The reconstruction for spectral SCI is an underdetermined inverse problem, which can be modeled as follows,

$$\hat{x} = \mathbf{\Psi} \underset{\theta}{\operatorname{argmin}} \left(\|\mathbf{\Phi}\mathbf{\Psi}\theta - y\|_2^2 + \tau \|\theta\|_1 \right) \tag{3}$$

where τ is the parameter that adjusts the weight of the sparsity of the solution in the optimization problem, and the following term $\|\theta\|_1$ restricts the sparsity of the solution [20].

To improve the reconstruction quality of hyperspectral images, an intuitive approach is to take multiple shots, but this gives up the advantage of snapshot. Using priors or additional information can improve the quality of the reconstructed images while preserving the advantages of snapshot.

Priors such as spatial correlation, spectral correlation, and non-local self-similarity are widely used for the reconstruction process of spectral SCI. Zhang Xinyue et al. exploited both spatial and spectral priors of hyperspectral images for reconstruction [21]; Chen Yong et al. proposed a subspace-based nonlocal tensor ring decomposition method that simultaneously captures the global spectral correlation, spatial structure, and

nonlocal self-similarity priors of hyperspectral images [22]; Liu Yang et al. exploited the high-dimensional structure of the signal to improve the reconstruction quality [23]. Zhang Shipeng et al. proposed a dimension-discriminative low-rank tensor recovery (DLTR) model which adaptively characterizes the structure prior in each dimension [24]; Since hyperspectral images have piecewise smooth structure in both spatial and spectral domains, Wu Zongrui et al. proposed to use total variation constraints in all three dimensions, and then used weights to weigh the total variation contribution in different dimensions [25]; Cheng Niankai et al. used the spatial-spectral correlation in hyperspectral images to build high-order tensors and proposed a weight high-order singular value regularization based low-rank tensor recovery model to characterize the structure prior of hyperspectral images [26]; Fu Wei et al. introduced the contextual prior into dictionary learning and the sparse reconstruction [27]. First, by introducing a joint sparsity constraint, they proposed a patch-based online dictionary learning algorithm, then, by introducing a weighted smoothing regularization, they built a reconstruction model. By their method, they successfully preserved the spectral-spatial structure of hyperspectral images and reduced the possible artifacts.

RGB images or various types of compressed measurements are used in the fusion model. For example, Yuan Xin et al. took RGB images as side information of a compressed sensing system and used the combined dataset of compressed measurements and their corresponding RGB images to learn a coupled dictionary [28]. In addition, Chen Limin [29], Tao Chenning [30], and He Wei et al. [31], all adopted the fusion of RGB image information and compressed measurements, aiming to use RGB images to compensate for the lack of spatial accuracy of the CASSI system. It is also possible to use various compressed measurements as fusion information, e.g., Jorge Bacca considered the fusion of compressed measurements from single-pixel cameras and coded aperture snapshot spectral imaging systems [32]; while Edwin Vargas proposed a model for the fusion of compressed sensor data with high spatial/low spectral and low spatial/high spectral resolution [33].

An example of using a blocking strategy to improve the speed of reconstruction is a mathematical model proposed by Henry Arguello et al. for lapped block reconstructions in CASSI [34] (Figs. 3 and 4).

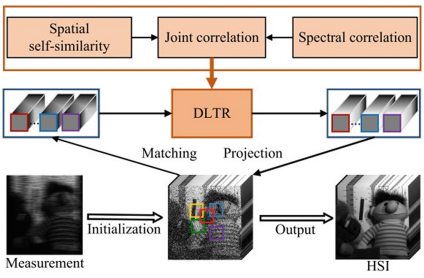


Fig. 3. Reconstruction with DLTR [24]

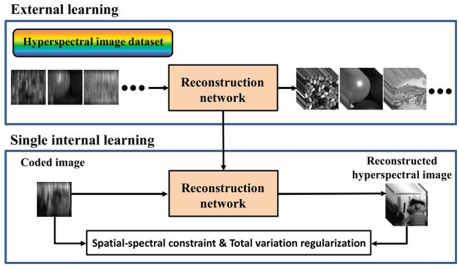


Fig. 4. A CNN-based method [35]

4 Deep Learning-Based Reconstruction Algorithms

Due to the powerful feature extraction capability of deep learning, numerous models in the field of deep learning have been successfully applied in the field of spectral SCI, such as convolutional neural network, generative model, transformer, etc.

Fu Ying et al. proposed a hyperspectral images reconstruction method based on convolutional neural network, which leverages hyperspectral images' spatial-spectral correlation through external learning and ensures its capability to generalize through internal learning [35]; In the method proposed by Wang Lizhi et al., a data-driven prior is used to regularize the optimization problem instead of a hand-crafted prior, and the regularizer is combined with an optimization-inspired network where the complete parameters are learned through end-to-end training, resulting in an algorithm with good robust performance [36]; In the method proposed by Huang Wei et al., two networks are included: a residual-dense network for hyperspectral images reconstruction and a spectral difference reconstruction network. In the first network, local features and global hierarchical features are extracted from all residual dense blocks by cascading them, while in the second network, spectral features of hyperspectral images are better preserved by exploiting the correlation between spectrally neighboring bands [37]; Sun Yubo proposed an unsupervised spatial-spectral network, which is a conditional generative model conditioned on snapshot measurements, using a spatial-spectral attention module to capture the joint spatial-spectral correlation of hyperspectral images [38]; Zheng Siming et al. proposed a plug-and-play approach using a deep learning-based denoiser to serve as a regularization prior for spectral SCI [39]; Lin Jing et al. proposed a transformer-based approach called coarse-to-fine sparse transformer, which embeds hyperspectral images sparsity into deep learning and applies it to hyperspectral images reconstruction, then performes coarse patch selection using a spectral-aware screening mechanism and feeds the selected patches into a customized spectral-aggregation hashing multi-head self-attentive for fine pixel clustering and self-similarity capturing [40].

As mentioned earlier, to improve the reconstruction accuracy of spectral SCI, most of the approaches are either through designing optimized coded apertures or through developing advanced reconstruction algorithms. Using deep learning, Wang Lizhi et al. tied the coded aperture optimization and reconstruction algorithms in a unified framework [41]. In their method, the entities of the coded apertures are obtained by learning as network weights, and the extensive correlation of hyperspectral images across spatial and spectral dimensions is exploited for reconstruction, finally a high-quality reconstruction outcome is got.

Models using deep learning for hyperspectral images reconstruction possess the advantage of high speed, but most networks generalize poorly, while there is wide variability among various spectrometers, a trained network can only achieve good results on a specific dataset, reconstruction accuracy decreases dramatically when it is used on other datasets, and retraining a network takes a long time, usually a week or two weeks. In addition, network models use end-to-end approaches and suffer from a lack of interpretability.

5 Noises and Denoising

Unlike a direct scene-to-image mapping style, spectral SCI first obtains compressed measurements of the scene, and later relies on the information processing to obtain the target images through a carefully designed reconstruction algorithm. Noise is inevitable due to both the impact from imaging equipment and imaging environment. How to effectively model the noise as well as denoising is a key aspect to improve the imaging quality.

Some experiments use Gaussian noise, shot noise, etc. to simply model the noise in spectral SCI. To select the specific noise, an intuitive noise selection method is to add the corresponding noise to the simulation data respectively, and then choose the type of noise according to the comparison of the similarity between the ground truth images and the reconstructed results. However, in practice, the noise in hyperspectral images is complex, and this simple noise modeling approach cannot achieve satisfactory results.

The measurement noise in the actual imaging process has a non-zero mean value, moreover, the non-zero mean value will be amplified in the reconstruction process, which seriously affects the reconstruction quality, so there is a need to model the non-zero mean noise. Zhang Maoqing et al. pointed out that there were two main reasons for the generation of non-zero mean noise, on the one hand, the dark current is generated by the electrical components in the camera, and on the other hand, due to imperfections in the transmittance, reflectance and smoothness of the optical elements, the light in the imaging system is reflected and scattered many times, based on which they established a reliable non-zero mean noise model [42].

In terms of denoising algorithms, considering the independent irrelevance between noise and clean data as well as a certain correlation between adjacent image elements of the image, one idea is to calculate the value of image elements by filtering in a band-by-band manner, such as median filter, mean filter, bilateral filter, etc.; Another idea is to use the difference between noise and clean signal in some transform domains, specifically, first project the image to a certain transform domain, then set a threshold to separate noise and clean signal, and finally complete denoising by an inverse conversion.

Based on the low-rank tensor property of noiseless hyperspectral images data and the sparsity property of outliers and non-Gaussian noise, Li Chang et al. proposed a model based on robust low-rank tensor recovery, which can preserve the global structure of hyperspectral images while removing outliers and different types of noise such as Gaussian noise, impulse noise, deadlines, etc. [43].

6 Conclusion

Spectral SCI greatly reduces the amount of the desired data to be sampled, transmitted and stored and does not need any scanning in the spatial or spectral domain, at the cost of requiring a reconstruction process. Suitable coded apertures and good reconstruction algorithms are crucial for imaging quality.

According to RIP, many optimized coded apertures have been proposed, such as colored coded apertures which replace blocked-unblocked elements with filter arrays, coded apertures using hexagonal elements, and some adaptive schemes, all of which

bring an improvement in PSNR of reconstructed images. The reconstruction of spectral SCI is an ill-posed inverse problem that demands the use of some constraint information to get highly accurate reconstruction outputs. Using image fusion methods, and using rich prior knowledge - such as spatial similarity, non-local self-similarity, and high-dimensional structure of the images et al. - are two strategies frequently used in hyperspectral images reconstruction algorithms. Due to the power of deep learning, the reconstruction speed of spectral SCI is greatly boosted. Besides, as mentioned before, schemes that jointly consider coded aperture optimization and reconstruction algorithms have been proposed.

To summarize, the reconstruction of spectral SCI not only requires high enough accuracy, but also expects real-time reconstruction, which is an ongoing challenge. In addition, the spectral range and noise amplitude of various spectral imagers vary, and the reconstruction algorithm should also have good generalization and robustness for noise.

Acknowledgement. This work was supported by the National Natural Science Foundation of China (NSFC) (62001328, 62001327, 61901301), Scientific Research Project of Tianjin Educational Committee (2021KJ182), the Natural Science Foundation of Tianjin Municipality (20JCYBJC00300), Open Fund of CAS Key Laboratory of Spectral Imaging Technology (No. LSIT202004W) and Natural Science Foundation of Shanxi Province (2021JQ-323).

References

1. Bioucas-Dias, J.M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N.M., Chanussot, J.: Hyperspectral remote sensing data analysis and future challenges. IEEE Geosci. Remote Sens. Mag. Institute of Electrical and Electronics Engineers Inc. **1**, 6–36 (2013)
2. Lu, G., Fei, B.: Medical hyperspectral imaging: a review. J Biomed Opt. **19**(1), 010901 (2014)
3. Neville, R.A., Lévesque, J., Staenz, K., Nadeau, C., Hauff, P., Borstad, G.A.: Spectral unmixing of hyperspectral imagery for mineral exploration: comparison of results from SFSI and AVIRIS (2003)
4. Donoho, D.L.: Compressed Sensing (2004)
5. Gehm, M.E., et al.: Single-shot compressive spectral imaging with a dual-disperser architecture [Internet]. IEEE Trans. Geosci. Remote Sens. SPIE-IS and Electronic Imaging, **41** (2003). http://www.acm.caltech.edu/
6. Wagadarikar, A., John, R., Willett, R., Brady, D.: Single disperser design for coded aperture snapshot spectral imaging (2008)
7. Wang, L., Xiong, Z., Gao, D., Shi, G., Wu, F.: Dual-camera design for coded aperture snapshot spectral imaging. Appl Opt. **54**(4), 848 (2015)
8. Ma, X., Zhang, H., Ma, X., Arce, G.R., Xu, T., Mao, T.: Snapshot compressive spectral imaging based on adaptive coded apertures. SPIE-Int. Soc. Opt. Eng. 2 (2018)
9. Candes, E., Tao, T.: Decoding by Linear Programming (2005)
10. Candès, E.J.: The Restricted Isometry Property and Its Implications for Compressed Sensing [Internet]. www.acm.caltech.edu/
11. Baraniuk, R., Davenport, M., DeVore, R., Wakin, M.: A simple proof of the restricted isometry property for random matrices. Constr. Approx. **28**(3), 253–263 (2008)
12. Mallat, S.G., Zhang, Z.: Matching pursuits with time-frequency dictionaries. IEEE Trans. Sig. Process. **41** (1993)

13. Tropp, J.A., Gilbert, A.C.: Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans. Inf. Theory **53**(12), 4655–4666 (2007)
14. Bioucas-Dias, J.M., Figueiredo, M.A.T.: A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration. IEEE Trans. Image Process. **16**(12), 2992–3004 (2007)
15. Arguello, H., Arce, G.R.: Colored coded aperture design by concentration of measure in compressive spectral imaging. IEEE Trans. Image Process. **23**(4), 1896–1908 (2014)
16. IEEE Signal Processing Society: Optimization of a Moving Colored Coded Aperture in Compressive Spectral Imaging, 7020 p. (2018)
17. Xu, Y., Li, C., Liu, S., Tang, G., Xie, J., Wang, J.: Simulation and verification of uniformly distributed-slit array coded spectral imaging system. Bandaoti Guangdian/Semiconductor Optoelectron. **42**(4), 562–567 (2021)
18. Zhang, H., Ma, X., Arce, G.R.: Compressive spectral imaging based on hexagonal blue noise coded apertures, 21 May 2019. http://arxiv.org/abs/1905.09387
19. Tao, C., Zhu, H., Sun, P., Wu, R., Zheng, Z.: Simultaneous coded aperture and dictionary optimization in compressive spectral imaging via coherence minimization. Opt. Express. **28**(18), 26587 (2020)
20. Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**(2), 489–509 (2006)
21. Zhang, X., Zhang, X.: Compressive hyperspectral imaging with spatial and spectral priors. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **11**(11), 4156–4169 (2018)
22. Chen, Y., Huang, T.Z., He, W., Yokoya, N., Le, Z.X.: Hyperspectral image compressive sensing reconstruction using subspace-based nonlocal tensor ring decomposition. IEEE Trans. Image Process. **29**, 6813–6828 (2020)
23. Liu, Y., Yuan, X., Suo, J., Brady, D.J., Dai, Q.: Rank minimization for snapshot compressive imaging. IEEE Trans. Pattern Anal. Mach. Intell. **41**(12), 2990–3006 (2019)
24. Zhang, S., Wang, L., Fu, Y., Zhong, X., Huang, H.: Computational Hyperspectral Imaging Based on Dimension-discriminative Low-rank Tensor Recovery
25. Wu, Z., Chen, X., Shi, W., Chen, L., Hu, S.: Hyperspectral image reconstruction via block low-rank and three-dimension weighted total variation constraint. IEEE Access **7**, 47698–47713 (2019)
26. Cheng, N., Huang, H., Zhang, L., Wang, L.: Snapshot hyperspectral imaging based on weighted high-order singular value regularization. In: Proceedings - International Conference on Pattern Recognition [Internet], pp. 1267–1274 (2020). http://arxiv.org/abs/2101.08923
27. Fu, W., Lu, T., Li, S.: Context-aware compressed sensing of hyperspectral image. IEEE Trans. Geosci. Remote Sens. **58**(1), 268–280 (2020)
28. Yuan, X., Tsai, T.H., Zhu, R., Llull, P., Brady, D., Carin, L.: Compressive hyperspectral imaging with side information. IEEE J. Sel. Top. Sig. Process. **9**(6), 964–976 (2015)
29. Chen, L., Xu, J., Liu, P.X., Yu, H.: A RGB-guided low-rank method for compressive hyperspectral image reconstruction. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **E101A**(2), 481–487 (2018)
30. Tao, C., Zhu, H., Sun, P., Wu, R., Zheng, Z.: Hyperspectral image recovery based on fusion of coded aperture snapshot spectral imaging and RGB images by guided filtering. Opt. Commun. **1**, 458 (2020)
31. He, W., Yokoya, N., Yuan, X.: Fast hyperspectral image recovery via non-iterative fusion of dual-camera compressive hyperspectral imaging, 30 December 2020. http://arxiv.org/abs/2012.15104
32. Bacca, J., Correa, C.V., Arguello, H.: Noniterative hyperspectral image reconstruction from compressive fused measurements. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **12**(4), 1231–1239 (2019)

33. Vargas, E., Espitia, O., Arguello, H., Tourneret, J.Y.: Spectral image fusion from compressive measurements. IEEE Trans. Image Process. **28**(5), 2271–2282 (2019)
34. Arguello, H., Correa, C.V., Arce, G.R.: Fast lapped block reconstructions in compressive spectral imaging (2013)
35. Fu, Y., Zhang, T., Wang, L., Huang, H.: Coded hyperspectral image reconstruction using deep external and internal learning. IEEE Trans. Pattern Anal. Mach. Intell. 1 July 2021
36. Wang, L., Sun, C., Fu, Y., Kim, M.H., Huang, H.: Hyperspectral image reconstruction using a deep spatial-spectral prior
37. Huang, W., Xu, Y., Hu, X., Wei, Z.: Compressive hyperspectral image reconstruction based on spatial-spectral residual dense network. IEEE Geosci. Remote Sens. Lett. **17**(5), 884–888 (2020)
38. Sun, Y., Yang, Y., Liu, Q.,Kankanhalli, M.: Unsupervised spatial-spectral network learning for hyperspectral compressive snapshot reconstruction, 18 December 2020. http://arxiv.org/abs/2012.12086
39. Zheng, S., Liu, Y., Meng, Z., Qiao, M., Tong, Z., Yang, X., et al.: Deep plug-and-play priors for spectral snapshot compressive imaging. Photonics Res. **9**(2), B18 (2021)
40. Lin, J., Cai, Y., Hu, X., Wang, H., Yuan, X., Zhang, Y., et al.: Coarse-to-fine sparse transformer for hyperspectral image reconstruction
41. Wang, L., Zhang, T., Fu, Y., Huang, H.: HyperReconNet: joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging. IEEE Trans. Image Process. **28**(5), 2257–2270 (2019)
42. Zhang, M., Wang, L., Zhang, L., Huang, H.: Compressive hyperspectral imaging with non-zero mean noise. Opt. Express. **27**(13), 17449 (2019)
43. Li, C., Ma, Y., Huang, J., Mei, X., Ma, J.: Hyperspectral image denoising using the robust low-rank tensor recovery. J. Opt. Soc. Am. A. **32**(9), 1604 (2015)

Fire Risk Prediction Based on the MultiBoost-Replaceable Model

Yue Han[1], Yuhua Song[2], Hengpeng Xu[3(✉)], and Jun Wang[2(✉)]

[1] College of Mathematics and Statistics Science, Ludong University, Yantai 264025, China
[2] Yantai XF Detachment, Yantai 264025, China
junwang@ldu.edu.com
[3] Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China
xuhp@tjnu.edu.com

Abstract. To address the problem that current fire supervision frequency is determined according to the experiences of supervisors, this paper proposes a novel building fire prediction framework based on the MultiBoost-Replaceable model. Specifically, a replaceable learning strategy is fused with the MultiBoost model to analyze the fire supervision data of 5683 buildings in an administrative region. The experimental results show that the proposed model is generally superior in predicting fire risk especially for some special cases such as high-rise buildings and crowded places.

Keywords: Fire risk prediction · MultiBoost · Ensemble learning

1 Introduction

With the rapid development of urban construction, numerous buildings have been and are being put up, bringing increasingly prominent problem of fire safety. As pursuing economic growth, ignoring fire safety is pervasive which has resulted in a huge number of fire accidents each year. Effectively predicting fire risks gives the way to avoid fire accidents that may happen in different types of building units. Daily fire supervisions in China usually adopt a strategy of "double randomness", that is, the randomly selected supervisors inspect target units randomly. Apparently, this strategy is stochastic and heavily relies on the experiences and expertise of supervisors.

In recent years, some researchers have paid close attentions on fire safety inspections via machine learning methods [1]. Chen et al. used a building information model to construct fire safety equipment components for safety information maintenance [2]. Madaio et al. developed a Firebird framework that could prioritise fire inspections or identify new inspection features [3]. Zhang et al. designed a

Y. Han and Y. Song—These authors are equally contributed to this work.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 323–331, 2023.
https://doi.org/10.1007/978-981-99-2653-4_40

crowd sourcing-based method that harvested collective intelligence for large-scale inspection tasks [4]. While most of these approaches involve manual efforts such as expert experiences or collective intelligence, the fire inspections are still subjective. The MultiBoost model [5] arises from the well-known AdaBoost model [6]. Similar to the general Boosting algorithm, MultiBoost can integrate several weak classifiers to obtain a strong one which can significantly improve the overall learning performance. In contrast to AdaBoost, MultiBoost incorporates the Wagging strategy [7], thereby inheriting both the excellent bias reduction from AdaBoost and the excellent variance reduction from Wagging.

In this paper, we design a MultiBoost-Replaceable model to predict the fire risk in various types of buildings, which correspond to different prediction scenarios. We collect the fire inspection data from 5,683 buildings in a city in China, including low-rise and high-rise buildings of different heights, as well as entertainment, schools, offices, and multi-purpose public places, and employ the MultiBoost-Replaceable model for prediction and give practical suggestions according to the evaluation results.

2 The MultiBoost - Replaceable Model

We improve the MultiBoost framework by replacing a standard base classifier (e.g., C4.5 decision tree) with a Replaceable strategy. Specifically, given a set of training data as $S = \left\{ (x_1, y_1) ; (x_2, y_2) ; ...; (x_n, y_n) \right\}$ and the weight of each sample as 1, a specific classifier $\mathcal{F}$ is involved for prediction and then, the weight is adjusted according to the continuous Poisson distribution and normalized in the next iteration. At each iteration t, the classification error is calculated as

$$\epsilon^{[t]} = \frac{1}{n} \sum_{\hat{y}_i \neq y_i} weight\,(x_i)\,, \tag{1}$$

where $\hat{y}_i$ is the prediction value of x_i and $weight(x_i)$ represents its sample weight. The weight of $\mathcal{F}$ in the final combination model is set as

$$\beta^{[t]} = \frac{\epsilon^{[t]}}{1 - \epsilon^{[t]}}. \tag{2}$$

If the classification error is still too high in the last sub-decision group iteration, the algorithm will append one or more sub-decision groups until the convergence requirements are satisfied. The algorithm flow is listed as follows.

Algorithm 1. MultiBoost-Replaceable algorithm.

Input: dataset S, number of iretations T, base classifier $\mathcal{F}$, and sub-decision group I

Output: predictions $\hat{Y}^*$

1: **begin**
2: initialize the weight of each sample in S as 1;
3: **while** "not converged" **do**

reassign sample weights based on a continuous Poisson distribution;
construct the dataset $S^{[t]}$ with sample weights;
$\hat{Y}^{[t]} = \mathcal{F}\left(S^{[t]}\right)$;
calculate $\epsilon^{[t]}$ according to Eq. (1);
calculate $\beta^{[t]}$ according to Eq. (2);
for each $x_i \in S^{[t]}$ **do**
  if $\hat{y}_i^{[t]} \neq y_i$ **then**
    $weight(x_i)^{[t]} = \frac{weight(x_i)^{[t-1]}}{2\epsilon^{[t]}}$;
  else
    $weight(x_i)^{[t]} = \frac{weight(x_i)^{[t-1]}}{2(1-\epsilon^{[t]})}$;
  end if;
end for;
t=t+1;
end while;
$\hat{Y} = \sum_t \hat{Y}^{[t]} \log \frac{1}{\beta^{[t]}}$;
return $\hat{Y}^*$;
end;

3 Data Collection

Fire supervision data from 5,683 buildings were collected for analysis in the following experiments. Thirteen features were extracted as shown in Table 1:

Table 1. Supervision features for fire risk prediction

Fea	Institution	Height	Area	Floors	Rating	Usage	Out.Hydrant
Type	char	num	num	num	nom	nom	num
Fea	In.Hydrant	SpraySys	AlarmSys	SmokeExt	Mana.Level	Num.People	
Type	num	nom	nom	nom	nom	num	

For all of the samples, the buildings whose heights are below 24 m are regarded as the low-rise buildings according to the Ministry of Construction's General Rules [8], while those over 24 m are considered as the high-rise ones. The statistics of fire supervision frequencies for both types of buildings are shown in Fig. 1. It is revealed that the proportion of normal inspection is the highest for both low-rise and high-rise buildings. The percentage of the high inspection frequency for high-rise buildings is nearly 15% larger than that for low-rise ones. This is in line with the urban development situation that high-rise and super high-rise buildings have sprung up with high fire risks.

In addition, we further categorize the buildings into 16 types according to their architectural functions, as demonstrated in Table 2, where the Nos. 1, 2, 3,

Table 2. Different categories of buildings

No	Cat	No	Cat	No	Cat	No	Cat
1	Flammable Sites	2	Hotels	3	Schools	4	Entertainments
5	Malls	6	Tourist Attractions	7	Warehouses	8	Enterprises
9	Museums	10	Nursing Homes	11	Offices	12	Hospitals
13	Administrations	14	Stations	15	Communications	16	Posts

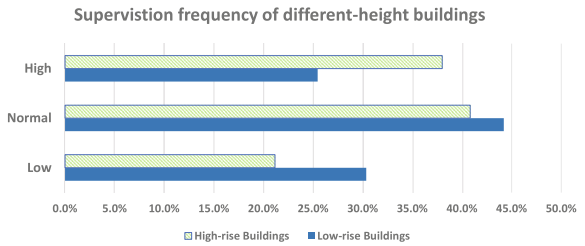


Fig. 1. Statistics of the supervision frequency for different types of buildings

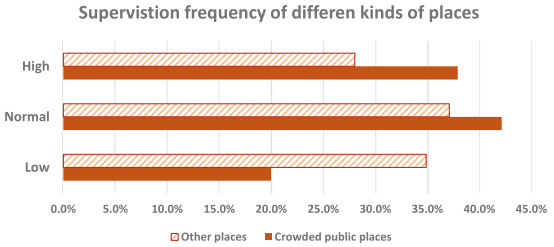


Fig. 2. Statistics of the supervision frequency for different places

4, 5, 6, 8, 9, 10, 12 and 14 are considered as the crowded places. Their statistical analysis of fire supervision frequencies are shown in Fig. 2.

As can be seen from Fig. 2, the percentage of high inspection frequency for crowded places is significantly higher than that of low frequency. This also indicates that fire rescue agencies have strengthened their inspections for the crowded places in recent years, which largely helps to prevent catastrophic fires and serious injuries or deaths.

4 Experimental Analysis

4.1 Replaceable Settings

The following base classifiers were fused into the Replaceable strategy: Random Forest [9], Elastic-Net [10], SVM (Support Vector Machine) [11], Bayesian Network [12], BP (Back Propagation) Network [13], and KNN (K-Nearest Neighbor) [14] classifiers. The number of decision trees in Random Forest was set to

10, linear kernel function was adopted in SVM, BP Network learning rate was fixed to 0.01 and the number of nodes in hidden layers was fixed to 10, and the number of nearest neighbors in KNN was fixed to 5.

4.2 Fire Risk Prediction Results

MultiBoost-Replaceable was compared with the six base classifiers introduced in Sect. 4.1 and AdaBoost combined with them. Each method was implemented 10 times on the dataset built in Sect. 3 and the five-fold cross validation results under the metrics of Accuracy, Precision, Recall, F-measure, and AUC were recorded in Table 3.

According to the table, we can observe that: (i) the overall prediction results of the MultiBoost-Replaceable strategy were comparably excellent, especially when comparing with the base classifiers without any boosting strategies; (ii) boosting strategies sometimes might not provide enough benefits, e.g., Bayesian

Table 3. Prediction results (%): MB and AB represent MultiBoost and AdaBoost, respectively.

	Metrics				
Methods	Accuracy	Precision	Recall	F-measure	AUC
Random Forest	74.94	81.21	59.68	60.71	88.20
Elastic-Net	61.95	54.23	61.40	56.01	71.63
SVM	70.39	70.25	72.86	69.89	82.08
Bayesian Network	72.99	65.62	66.54	64.68	97.23
BP Network	71.04	64.75	73.40	67.05	93.68
KNN	53.51	63.49	48.67	50.43	75.43
AVG.	*67.47*	*66.59*	*63.76*	*61.46*	*84.71*
AB-RandomForest	71.69	89.73	68.73	68.41	94.09
AB-ElasticNet	61.95	54.23	61.40	56.01	71.63
AB-SVM	73.64	72.45	76.34	69.99	90.68
AB-BayesianNetwork	72.99	65.62	66.54	64.68	97.23
AB-BPNetwork	76.04	69.75	78.40	72.05	98.68
AB-KNN	60.65	70.34	53.31	55.44	75.91
AVG.	*69.49*	*70.35*	*67.45*	*64.43*	***88.04***
MB-RandomForest	**78.39**	**90.56**	**81.80**	**77.72**	87.50
MB-ElasticNet	61.95	54.23	61.40	56.01	71.63
MB-SVM	76.88	83.10	74.20	73.31	89.14
MB-BayesianNetwork	72.99	65.62	66.54	64.68	97.23
MB-BPNetwork	76.04	69.75	78.40	72.05	**98.68**
MB-KNN	67.79	72.70	57.90	58.29	76.56
AVG.	***72.34***	***72.66***	***70.04***	***67.01***	86.79

Network classifier made no significant improvements when combining with MultiBoost or AdaBoost. However, it should be noted that this is not a general phenomenon while merely works on this dataset; (iii) MultiBoost-BP Network and AdaBoost-BP Network performed identically, indicating that different boosting strategies might make nearly similar effects in some situations; (iv) if a base classifier performed outstanding, it might be significantly improved after combining with boosting strategies the performance, such as SVM; and (v) MultiBoost-Random Forest performed superior under different metrics.

4.3 Evaluations for High-Rise Buildings

To further test the predictive performance of MultiBoost-Replaceable on the high-rise buildings (taller than 24 m as mentioned before), these buildings were further divided into four data sets according to their building height: $[24, 70), [70, 120), [120, 170), [170, 214]$, and the prediction results for different heights of buildings are shown in Fig. 3.

According to the figure, we can find that: (i) a strategy might provide different prediction effects on the high-rise buildings of different heights. For example, MultiBoost-Bayesian Network performed outstanding in general cases while comparably inferior when building heights were taller than 170; and (ii) generally, the MultiBoost-Replaceable strategies performed less effectively on much taller buildings, such as those taller than 170 m, which might be due to the fewer number of this kind of buildings as well as its higher fire safety requirements.

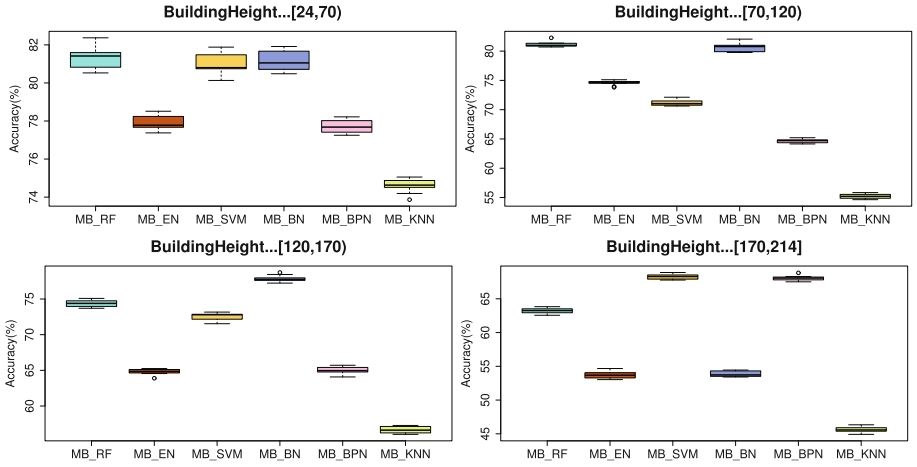


Fig. 3. Prediction performance for different types of high-rise buildings

4.4 Evaluations for Crowded Places

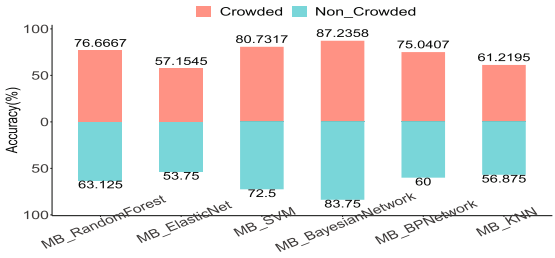


Fig. 4. Prediction performance for crowded and non-crowded places

Based on Table 2, the original dataset is further divided into two datasets respectively with crowded and non-crowded places, and the prediction results are demonstrated in Fig. 4. We can observe that: (i) the prediction performance of the MultiBoost-Replaceable strategies for crowded places was significantly better than that for non-crowded places. This is because both fire protection facilities and fire managements in crowded places have been largely improved in China and these conditional attributes may be conducive to fire risk prevention; (ii) MultiBoost-Bayesian Network achieved the best results for both crowded and non-crowded places, indicating its superior prediction performance; and (iii) the same strategy might show different prediction effects on different places. For instance, the accuracy of MultiBoost-RandomForest and MultiBoost-BPNetwork differed by more than 10% when predicting crowded and non-crowded places.

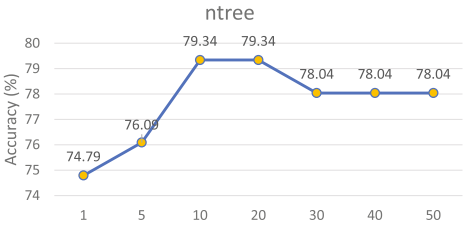
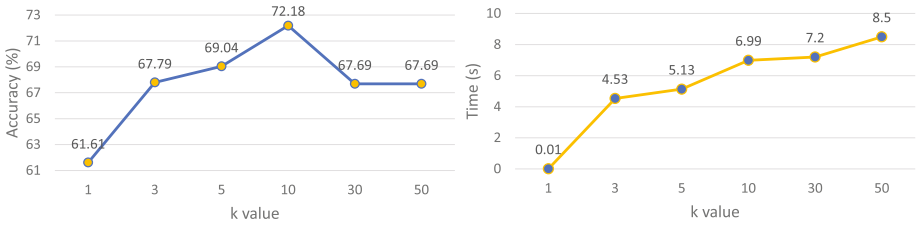


Fig. 5. Accuracy of MultiBoost-Random Forest with increasing the number of trees

4.5 Parameter Sensitivity Evaluations

- **Effects of Number of Trees.** The prediction performance of MultiBoost-Random Forest with different number of trees is illustrated in Fig. 5. The top score was obtained when *ntree* was increased to 10, and subsequently begun to drop when continuing increasing this parameter, which suggested that simply augmenting decision trees might not contribute to performance. Considering effectiveness and efficiency, *ntree* was fixed to 10 throughout the experiments.

- **Effects of Number of Nearest Neighbours.** Given the number of nearest neighbors $k \in [1, 3, 5, 10, 30, 50]$, the prediction results and running costs of MultiBoost-KNN are shown in Fig. 6 (Intel Core i7-4790 CPU, 64GB Memory). As can be seen from the figure, the prediction performance showed a significant rise at first and begun to decline after reaching its highest score when $k = 10$. In terms of running costs, it would significantly scale up with increasing k. Accordingly, the k value of MultiBoost-KNN was fixed to 5 in the experiments.



(a) Accuracy variations (b) Time variations

Fig. 6. Performance of MultiBoost-KNN with increasing the number of nearest neighbors

5 Conclusions

At present, there exist few studies on fire safety supervision via machine learning technologies. This paper constructs a novel fire risk prediction model namely MultiBoost-Replaceable, which shows high prediction performance of 99.33% accuracy on the newly introduced dataset. Applications of machine learning technologies, such as the Boosting strategies, in the field of fire safety protection can greatly approach the project of "smart fire fighting", which brings benefits for personnel safety and urban constructions.

Acknowledgements. We are very grateful to the anonymous reviewers and editor for their helpful and constructive comments. This work is partially supported by the National Natural Science Foundation of China under Grant No. 62106091, the Doctoral Foundation of Tianjin Normal University under Grant No. 52XB2104, and the Shandong Provincial Natural Science Foundations under Grant Nos. ZR2021MF054 and ZR2022MF247.

References

1. Zhang, X., Sun, Y., Liu, H., Hou, Z., Zhao, F., Zhang, C.: Improved clustering algorithms for image segmentation based on non-local information and back projection. Inf. Sci. **550**, 129–144 (2021)
2. Chen, Y.J., Lai, Y.S., Lin, Y.H.: BIM-based augmented reality inspection and maintenance of fire safety equipment. Autom. Constr. **110**, 103041 (2020)

3. Madaio, M., Chen, S.T., Haimso, O.L.: Firebird: predicting fire risk and prioritizing fire inspections in Atlanta. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 185-194 (2016)
4. Zhang, D., Zhang, J., Xiong, H.: Taking advantage of collective intelligence and BIM-based virtual reality in fire safety inspection for commercial and public buildings. Appl. Sci. **9**(23), 5068 (2019)
5. Webb, G.I.: Multiboosting: a technique for combining boosting and wagging. Mach. Learn. **40**(2), 159–196 (2000)
6. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. **55**, 119–139 (1995)
7. Bauer, E., Kohavi, R.: An empirical comparison of voting classification algorithms: bagging, boosting, and variants. Mach. Learn. **36**, 105–139 (1999)
8. GB 50352-2005, General rules for the design of civil buildings
9. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
10. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. J. Roy. Stat. Soc. Ser. B (Stat. Methodol.) **67**(2), 301–320 (2005)
11. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
12. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, Massachusetts (1988)
13. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Nature **323**(6088), 533–536 (1986)
14. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE Trans. Inf. Theory **13**(1), 21–27 (1967)

Design of Intelligent Sorting and Storage System Based on Virtual Simulation Technology

Qingfeng Sun(✉) and Yue Liu

Department of Electronic Engineering, Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241002, Anhui, China
Sunqife@126.com

Abstract. Based on the virtual simulation software factory IO, build a virtual simulation model that integrates the functions of detection, sorting and warehousing. Through communication with PLCSIM, develop and design a three-dimensional warehousing system controlled by PLC. The system can realize the sorting and warehousing of different types of workpieces, which can be applied to the actual storage, and can also be used as a virtual simulation experiment teaching project to improve the quality of PLC practical teaching and improve students' professional ability.

Keywords: Factory IO · warehousing system · professional ability

1 Introduction

Automation technology has made continuous breakthroughs in the industrial, logistics, medical, military and commercial fields, making the storage and transportation of materials change from traditional warehousing to automated warehousing. Intelligent three-dimensional storage system can not only greatly save human resources and reduce human costs, but also transport goods to the designated warehouse location to achieve better storage management, which can improve the storage turnover efficiency of the warehouse and reduce storage costs.

PLC (programmable logic controller) is widely used in industrial control fields such as automatic production line, temperature and liquid level [1]. With the continuous development of industrial network technology, PLC is no longer the application of simple digital quantity and analog quantity, but an industrialized comprehensive application with network communication, digital interaction and advanced controller deployment. However, the traditional production line commissioning based on physical construction has some shortcomings, such as complex construction, a lot of human and material resources are required for commissioning and adjustment design, and commissioning needs to be carried out on site. Under this background, virtual simulation and digital twin technology have developed rapidly, Unity3D [2]、Factory IO [3, 4]and other platforms have become effective tools for industrial production line design and simulation.

Based on virtual simulation software factory IO and Siemens PLC programming software, this paper designs and implements an automatic goods sorting and storage system.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 332–337, 2023.
https://doi.org/10.1007/978-981-99-2653-4_41

2 Design Scheme of Intelligent Storage System

Automatic storage control system, also known as automatic access system, high-rise shelf warehouse, automatic three-dimensional warehouse, etc., is an important part of modern product storage and logistics transportation. It is mainly composed of goods access system, goods transportation system, real-time monitoring and management system, etc. Its working principle is to adopt multi-layer shelves in the vertical direction, make full use of space, and use automation technology to control handling and conveying equipment and automatic sorting equipment to complete the sorting, warehousing and outbound operations of goods [5].

2.1 Overall Function Design

The automatic storage system controls the operation of the whole system through PLC, and finally realizes the sorting, warehousing, outbound and other operations of the objects that need to be stored without manual processing. The mutual cooperation between conveyor belts, automatic guided vehicles (AGVS), stackers and other equipment makes the movement operations in the warehouse more simple and reasonable [6]. Specifically, this design mainly realizes the sorting and warehousing of two types of goods (A and b). The screening of cargo types is realized by weight detection, and its main process is shown in the following figure (Fig. 1, Fig. 2, Fig. 3, and Fig. 4).

Fig. 1. System flow chart

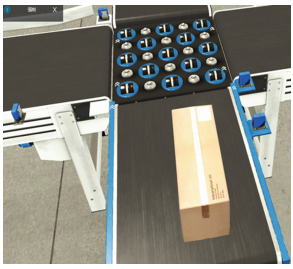


Fig. 2. A goods

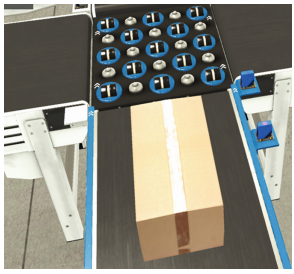


Fig. 3. B goods

2.2 Scheme Design

2.2.1 Software Platform

In this design, factory IO is selected as the virtual simulation platform. This software is a 3D simulation software, which supports PLC drivers of Siemens, Schneider, abb

and other brands [7]. Factory IO will have 21 basic scenes, which can be selected for simulation according to their own needs, or built in the scene by themselves.

Factory IO provides a variety of connection options, which can connect s7–1200/s7–1500 PLCs through plcsim advanced connection or IP addressing. At present, plcsim advanced only supports s7–1500 PLC for programming control. If you need to realize control through virtual PLC, you need to establish s7–1500 virtual PLC. Using this simulation software, you only need to make simple network card selection and IP settings to realize communication connection. In this design, s7-1500plc is selected as the carrier, TIA portal V16 software is used for programming, and factory IO software is combined to simulate and test the compiled program.

2.2.2 Sort

The sorting operation will be realized by measuring the weight. In the simulation software, there will be two different kinds of goods, namely, goods a and goods B, which will be transported to the designated path during the sorting operation. The sorting of goods is based on weight detection, that is, a weight detection device is installed on the main conveyor belt. When the materials are transported to this position, the materials are weighed, and the materials are sorted through the steering device according to the set weight value range.

2.2.3 Transshipment

In the aspect of industrial production and transportation, it is necessary to place the target at the designated position through the handling robot to carry out the task of the next link. There are mainly two types of handling robots: right angle type and joint type. The right angle type handling robots can be divided into gantry type, cantilever type, side wall type, swing arm type and other different forms. This design adopts the gantry type handling robot to grab the workpiece to the pallet, and carry the goods on the belt conveyor belt to the pallet of the roller conveyor belt, and then transport them to the three-dimensional shelf.

2.2.4 Warehousing

The warehousing operation adopts the combination of three-dimensional shelves and stackers to achieve the utilization of space. Stacker is the main access equipment in the automatic storage system. Its function is to automatically access goods on fixed shelves, which is also known as roadway stacker. The stacker can only move along the x-axis, Y-axis and z-axis on the fixed guide rail of the roadway, and all its motion tracks are based on a three-dimensional coordinate system, which is easy to stably place the workpiece at the designated position of the shelf (Fig. 5 and Fig. 6).

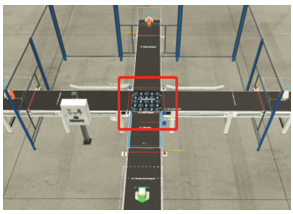


Fig. 4. Weight detection

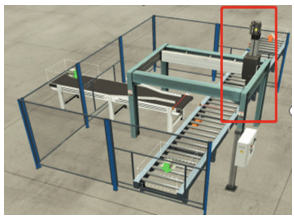


Fig. 5. Gantry handling robot

Fig. 6. Shelves and Stackers

3 System Design and Implementation

Based on the previous scheme design, a two-way path sorting and warehousing system scenario is built in factory IO (as shown in Fig. 7), and all links of the built automatic warehousing system are designed. First, the sorting system adopts weight sorting to complete the two-way path transportation according to the weight of goods; The second is the handling unit, which adopts the gantry type handling robot to carry the goods from the transportation belt to the pallet to realize the next level of transportation; Finally, the warehousing link, using the stacker to complete the warehousing of goods, and place the items in the designated position of the three-dimensional shelf. After the construction of the working scene is completed, the storage system design is realized by using TIA Portal software and selecting Siemens 1500plc for programming.

Fig. 7. Storage system simulation model

Sorting workflow: when starting the program, power on initialization will be carried out. When the corresponding sensor detects materials, the issuing motor and the weighing forward motor will start. When the materials reach the middle position of the weighing machine, they will stay and weigh to get an accurate value. The obtained value will make the steering device operate. The materials are transported according to the specified route. When passing the sensors on the sub line, the steering device will reset for the next operation.

Transfer workflow: the writing method is completed by moving instructions. When starting the PLC, the program will be powered on for initialization. When the chassis is detected at the delivery end of the chassis, the conveyor belt will run, and the chassis will be transported to the designated place for loading. When the belt conveyor end detects goods, the handling robot will run to the designated position, pick up the goods, move above the chassis, release the goods, and complete the handling (as shown in Fig. 8) (Fig. 9 and Fig. 10).

Warehousing process: there are 54 storage locations on the three-dimensional shelf. The objects are stored in automatic order from storage location 1 to storage location 54 (as shown in Fig. 11).

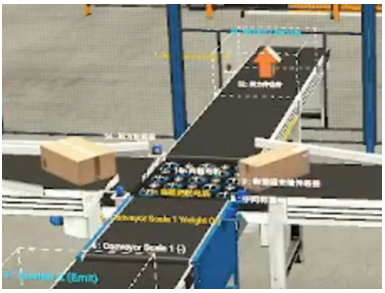


Fig. 8. Detect and Sort

Fig. 9. Transport

Fig. 10. Palletizer Handling

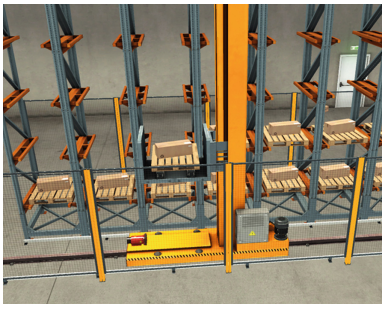


Fig. 11. Warehousing

4 Conclusion

Combined with virtual simulation software and Siemens PLC programming software, this paper designs and develops an intelligent sorting and storage system controlled by PLC. The system can realize the sorting and warehousing tasks of different types of workpieces, and has practical significance for the design and construction of actual warehousing. In addition, combined with the application of virtual simulation technology in teaching, it can be developed into a virtual simulation experiment teaching project, improve the quality of PLC practice teaching, and improve students' professional ability.

Acknowledgements. This work was financially supported by the project of natural science research in colleges and universities of Anhui Provincial (No. KJ2019A1153), program for outstanding young talents in colleges and universities in Anhui Province (No: gxyq2020233), academic support project for top talent in disciplines in colleges and universities in Anhui Province (No: gxbjZD2020108).

References

1. Shao, N., Jiang, G., Yang, H.: Design of automatic storage system based on PLC. Manufact. Automat. **35**(03), 122–124+128 (2013)
2. Ye, J., Xu, X,, Hao, X., Zhang, J.: Research on PLC three-dimensional storage simulation system based on unity 3D. J. Jinggangshan Univ. (Nat. Sci. Ed.), **43**(03), 54–60 (2022)
3. Zhang, Z., Liang, L.L., Wang, L., Lei, Z., Yin, C.: Virtual simulation experiment design of material sorting and assembly based on factory IO. Experim. Sci. Technol. **19**(03), 122–126 (2021)
4. Zhu, H.: Research on PLC Virtual simulation control system based on factory IO. Inf. Comput. (Theoret. Ed.), **32**(14), 224–226 (2020)
5. LIu, J., Zhao, D.-j, Yang, X.: components and operation mechanism of warehouse adaptive automatic execution system. China Circulat. Econ. **33**(2), 3–10 (2019)
6. Xie, Y., Liu, Z.-h, Wu, H.-c: preliminary study on the development model of the transformation from Chinese manufacturing to chinese intellectual manufacturing. Ordnance Automat. **36**(2), 1–7 (2017)
7. Liu, Z.: research on engineering practice based on virtual simulation factory (Factory IO) [J] Tianjin Sci. Technol. **49**(04), 43–47 (2022)

Design and Research of a Tower Type Intelligent Storage Cabinet

Yechao Shen(✉) and Yong Ge

Anhui Technical College of Mechanical and Electrical Engineering, Wuhu, Anhui 241002, China
shenyechao2006@126.com

Abstract. In order to study the intelligent auxiliary storage management system with rapid material positioning, automatic recording of access and multi-mode collaborative control. Taking PLC 1212 (DC/DC/DC) V4.2 as the control core, with the voice control module, WINCC touch screen system and remote cooperation terminal, through the linkage operation of servo motor, servo driver and electromagnet system, voice and voice control, on-site The all-round, three-dimensional intelligent operation of touch screen and remote monitoring is suitable for the operation needs of all age groups in different occasions. The simultaneous introduction of facial recognition, identity login verification and other systems enhances the reliability of the intelligent storage system. The unique polygonal pyramid configuration of the tower-type intelligent storage, with the cooperation of the positioning angle rotation mechanism, solves the problem of inconvenient access to materials inside the traditional large storage cabinet, and automatically generates a storage material ledger stored in the cloud, which greatly facilitates the material account. Inquiry and storage location locking have certain practical value.

Keywords: programmable controller · voice control · touch screen design · human-computer interaction · remote collaboration

1 Introduction

With the continuous improvement of productivity and living standards, the tools, work materials and daily necessities used by people have become more abundant. Taking tmall mall as an example, of the 30 common commodities, at least 5387 products have been sold [1]. On the one hand, the great variety of items has facilitated our life, and the resulting storage and management problems have become increasingly prominent. People tend to forget the specific placement of items and tools, and can't get them quickly and accurately when they are in urgent need of use, which is time-consuming and labor-consuming. For example, in the emergency repair scenario, different types of repair kits are equipped due to different levels of faults [2], and the rapid access of tools puts forward higher requirements for repair personnel. An optimized storage management device can reduce the trouble of searching for us and facilitate our production and life. At the same time, the safety protection performance of the existing storage cabinets is insufficient, and the storage layers of different storage boxes need to be locked with

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 338–348, 2023.
https://doi.org/10.1007/978-981-99-2653-4_42

independent traditional keys. Under the condition of insufficient security level, a large number of storage cabinet doors or drawer keys increase the storage cost of users. In recent years, experts and scholars at home and abroad have tried to explore and study the improvement of storage methods in different ways. Literature [3–5] has studied the characteristic needs and auxiliary improvement methods of children, the elderly, the disabled and other different groups in the use of storage cabinets through different ways such as functional module adaptation design, spatial regional preference division, auxiliary power drive, etc. literature [6–9] has used risk map, genetic algorithm The introduction of semi-structured bar code and UML unified modeling language tool have studied the risk management elements and the ways to improve the storage efficiency in the process of goods storage management. Literature [10–12] has studied the main effects of human factors such as cognitive ability, psychological characteristics and sports habits in the process of goods sorting and storage. However, at present, the design and research of multifunctional and intelligent storage cabinets suitable for the majority of the general population is still relatively rare.

Based on the shortcomings of the current technical level, change the traditional structure of the single-sided door opening of the ordinary storage cabinet, design the polygonal pyramid rotating tower body, and support the configuration of multiple storage drawers in the upper and lower directions, take pc/plc controller, servo driver, servo motor as the main control core, and use manual control buttons, touch screen, voice control to realize the multi port execution and release of goods storage and access, and cooperate with the intelligent control system, It not only enhances the realization of cloud monitoring and electronic security, but also makes it possible for users to remotely authorize access to specific items, enriches the application occasions of lockers, and provides a new method for classified storage of different items, rapid access of user authorization, and real-time recording of storage locations.

2 Overall Structure and Control Principle of Tower Type Intelligent Storage Cabinet

The purpose of tower shaped intelligent storage cabinet is to facilitate the intelligent storage and access of materials, and its shape and structure also need to adapt to the functional characteristics of intelligent access of materials. Traditional storage cabinets are mostly set in a single direction. The biggest disadvantage of this structure is that once the storage objects are large and there are many materials, the storage cabinets are bound to be designed and made larger, wider and deeper. When the depth of the storage cabinet exceeds the single arm extension of ordinary users, the access of materials placed on the inner side often needs to remove the obstruction of materials on the outer side first, and even users need to probe and bend down to penetrate part of the body into the storage cabinet, so as to increase the reach range of materials taken and placed by arms. This kind of access mode often brings inconvenience to people, and when the storage cabinet is placed at the corner of the wall of a home or factory, its access inconvenience will be greatly enhanced. In order to solve the problem of inconvenient access to materials on the inside, the shape of the storage cabinet with traditional cube structure is changed, and a structure of polygonal pyramid rotating tower body is designed. When the accessed

materials are placed on the back of the storage cabinet, the control cabinet is driven to rotate at an appropriate angle through the control system, so that the current access area is located on the side facing the user. The touch screen and control buttons can be set in the appropriate position according to the user's needs. Its shape structure is shown in Fig. 1 below.

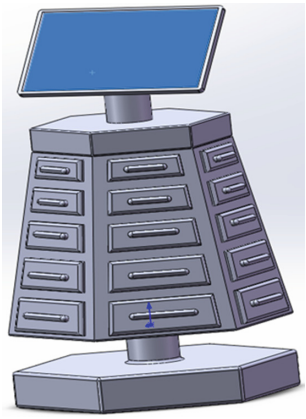


Fig. 1. Outline drawing of tower type intelligent storage cabinet

The control system of the tower type intelligent storage cabinet is composed of PC /plc controller, servo motor, servo driver and other components, which cooperate with the mechanical inching button, the touch screen interface made by WinCC, the app input and output terminal of the network module linkage control and the voice module of the field command control. It can realize the multi position integrated control linkage mode of manual control, voice control, touch screen control and remote control. In order to avoid the failure of the storage cabinet in the power-off state, the UPS power supply with uninterrupted power supply is introduced to cooperate with the low-power operation state in the power-off mode to ensure the emergency access of materials in the power-off state. The control logic relationship is shown in Fig. 2 below.

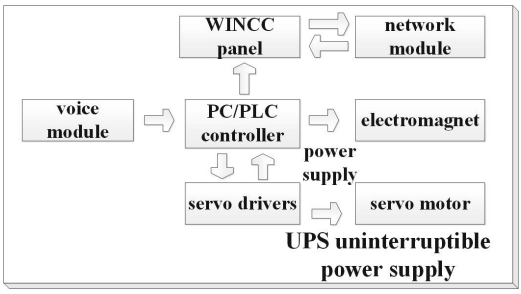
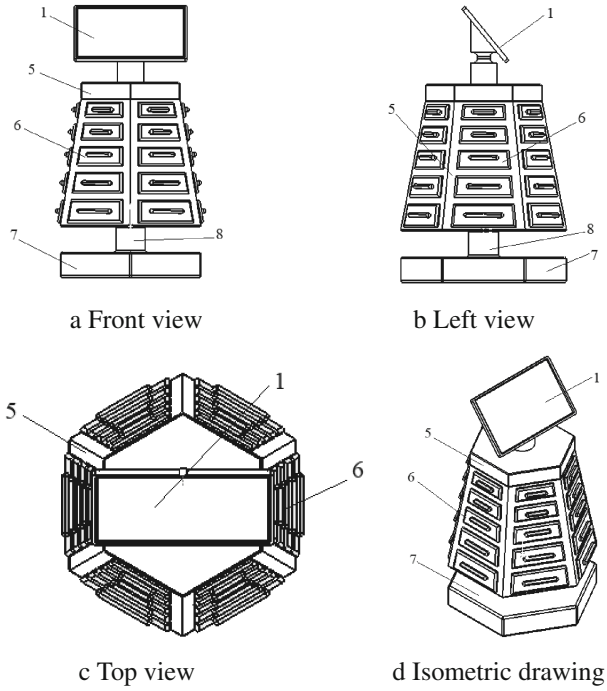


Fig. 2. Schematic diagram of control logic relationship

3 Unit Selection and Structure Analysis of Tower Type Intelligent Storage Cabinet

Combined with the design drawing of the tower type intelligent storage cabinet, the structural function analysis is carried out, and its shape and structure design are shown in Fig. 3 below.

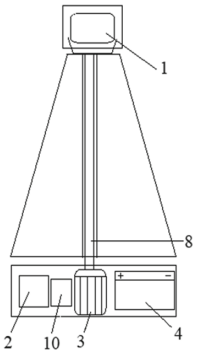


1-display 5-rotating tower 6-storage drawer 7-base 8-column

Fig. 3. Outline structure design of tower type intelligent storage cabinet

The tower body of the tower type intelligent storage cabinet adopts the multilateral pyramid design. In order to give consideration to the relative fixity of the tower body in the static state and the rotatable displacement when taking materials, the rotating tower body is fitted on the vertical column by sliding on the same axis. In order to maintain the relative stability of the column, the column is fixedly connected with the base. On each side of the polygonal pyramid, a plurality of storage drawers are arranged in the up and down directions to realize the partition storage of different materials. In Fig. 3, the polygonal pyramid adopts hexagonal pyramid to form six sides. In the actual use process, other polygonal forms can be used to replace the polygonal pyramid according to the needs of users. The number of layers and numbers of drawers can also be increased or decreased according to the actual needs.

In order to ensure that the main structure of the tower type intelligent storage cabinet has excellent impact strength, electrical performance, wear resistance and chemical resistance, the rotating tower body is made of ABS engineering plastic (engineering plastic alloy). The base of the storage cabinet is a hollow structure, and the extended connecting rod of the rotating tower passes through the column into the hollow base, which is connected with the pre customized servo motor in the base. Under the action of the servo driver, the servo motor drives the whole tower body to rotate and position along the central axis of rotation. The PC or PLC controller is connected with the rotation sensor that detects the rotation angle to verify whether the precise control of the rotation angle meets the established requirements. The internal structure of the base and the connection mode of components are shown in Fig. 4 below.



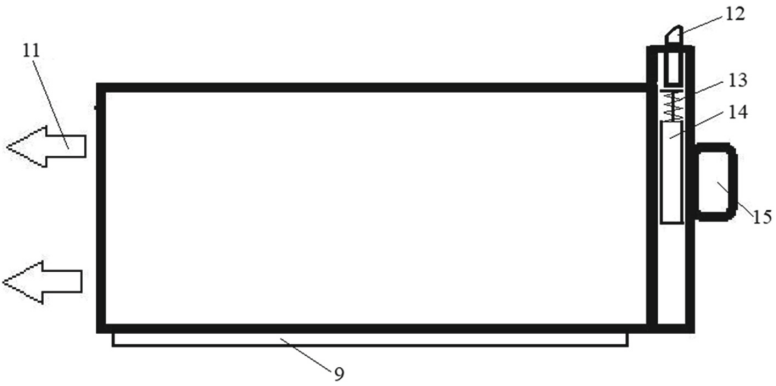
1-display 2-plc controller 3-servo motor 4-ups power supply 8-column 10-servo driver

Fig. 4. Schematic diagram of the internal structure of the base of the tower type intelligent storage cabinet

In order to overcome the limitation of the tower type intelligent storage cabinet relying on electric energy and avoid the emergency taking and placing of materials in the storage cabinet in the state of power failure. Prepare the UPS uninterruptible power supply system, which is set in the hollow base of the storage cabinet. Under normal conditions, the storage cabinet uses the conventional power supply system to work. In case of an emergency, the tower type intelligent storage cabinet temporarily turns the unnecessary functions into the sleep state (the user can temporarily restore the full function state as needed), and only retains the function of finding and taking and placing materials. In order to obtain sufficient tolerance redundancy, when the UPS system works independently, it should at least ensure that the material access action of the storage cabinet works independently for more than 50 times. When the conventional power supply system is powered on, the ups battery will be charged automatically.

The storage drawers are also made of ABS engineering plastics, and the drawers are contracted inside the rotating tower in daily state. The storage drawer panel is provided with a latch which is vertically upward, with the upper end facing inward and inclined downward. A latch spring is arranged below the latch, and an electromagnet is arranged below the spring. The control circuit of the electromagnet is connected with a PC or

PLC controller. The inner end spring of the drawer is set on the outer surface of the inner end plate of the storage drawer, and the drawer slide rail is set on the bottom surface of the drawer bottom plate and forms a sliding fit with the guide rail built in the rotating tower. Photoelectric sensors are preset on each drawer to detect the storage status of items in the drawer. The data in the photoelectric sensor is also transmitted to the PC or PLC controller through the communication line. The drawer structure of the tower type intelligent storage cabinet is shown in Fig. 5 below.



9 drawer slide rail 11 drawer inner end spring 12 latch 13 latch spring 14 electromagnet 15 drawer handle

Fig. 5. Drawer structure diagram of tower type intelligent storage cabinet

The introduction of the storage drawer adopts the principle that the electromagnetic iron is energized to generate magnetic attraction unlocking. When the storage cabinet is positioned to the items required by the user, the servo motor controlled by PLC will automatically turn the found storage drawer to the opposite side of the user, and then the unlocking electromagnetic iron on the storage drawer is energized. Under the condition that the magnetic traction is lost during storage and power on, the storage drawer is ejected by the built-in spring, so as to realize the accurate positioning and automatic access of materials.

4 Operation Control Programming of 3 Tower Type Intelligent Storage Cabinet

When the hexagonal pyramid is used, the servo motor needs to have six working positions, so the program needs to have six absolute positioning commands, and the reference commands are shown in Fig. 6 below.

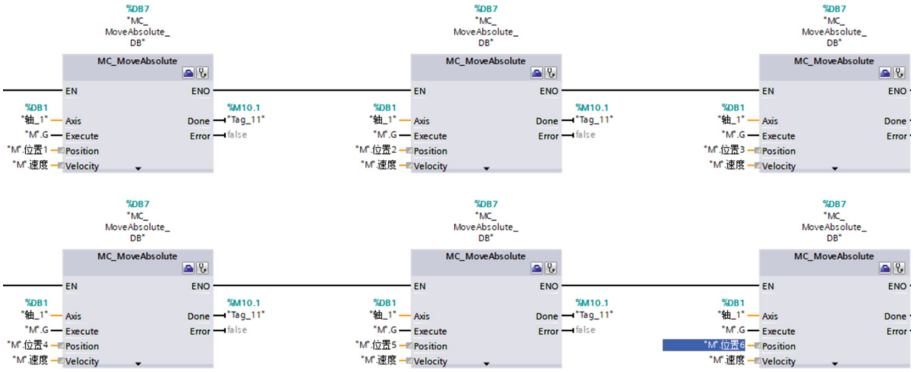


Fig. 6. Absolute positioning command

The signal of the rotary position sensor is sent to the controller or computer to achieve accurate positioning. Here, PLC (or PC) can be used for programming control. Taking plc1212 (dc/dc/dc) v4.2 controller as an example, botu is used for shaft configuration debugging. As shown in Fig. 7 below.

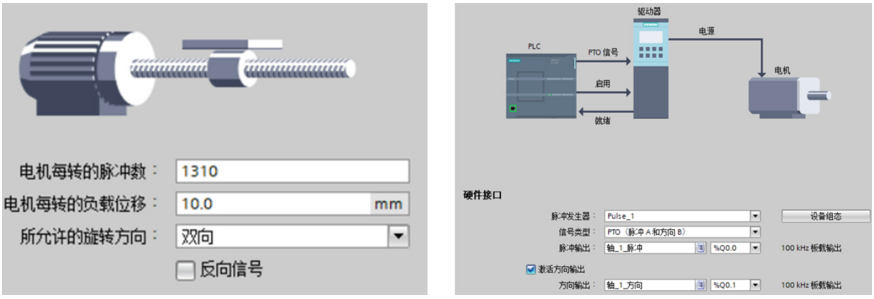


Fig. 7. Configuration of axis

After the configuration debugging is completed, the basic programming is carried out. The start of the servo motor requires an MC power command as the enable start, allowing the start of the shaft, and then the absolute positioning shaft is used, as shown in Fig. 8 below.

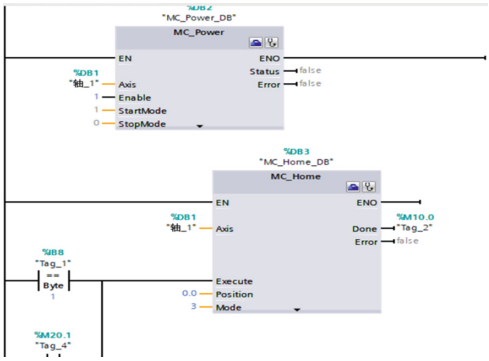


Fig. 8. Start of shaft

5 Realization of Human-Computer Interaction of 4-Tower Intelligent Storage Cabinet

After the motion control configuration is completed, it is necessary to realize the accurate call of motion control commands through the human-computer interaction system, cooperate with the accurate security system, and ensure the safety and reliability of material storage. The commonly used human-computer interaction can be realized through the multi-path of the storage cabinet airborne touch screen, remote control system, face-to-face speech recognition and so on. The airborne touch screen can be made with corresponding supporting tools according to the selection of different touch screens. Take the touch screen tool WinCC of botu software as an example, and make the touch screen interface as shown in Fig. 9 below.

Fig. 9. Human computer interaction interface

According to the needs of the security level of the storage, security systems with different warning standards can be configured as appropriate, such as face recognition, abnormal alarm, login password setting, etc. After entering the system, you can query the current storage configuration information and storage unit occupancy through the system. The corresponding information has been generated when the materials are stored, and can be uploaded to the cloud through the network system. Users can query the inventory of materials in real time. Common consumables can also set low inventory warning

values to help users make replenishment decisions. The material storage information of the storage cabinet is shown in Fig. 10 below.

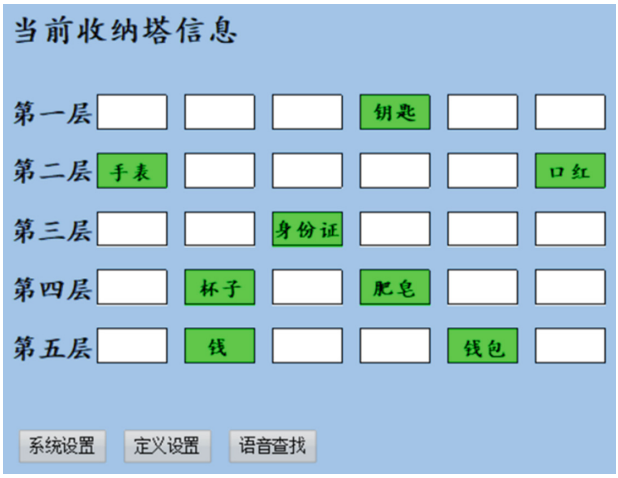


Fig. 10. Storage information of tower storage cabinet

The storage system can also integrate voice recognition buttons. In this functional state, the intelligent voice control system and sensor module customized according to user needs should be used, and the control purpose of intelligent voice automation can also be achieved. Users can tell the storage cabinet their current material access and placement requirements. After the information is confirmed, it can quickly rotate and position, and push out the corresponding storage drawer in place. Different forms of control instructions can call the same motion control program.

Using the remote control system and cloud server, the current state of the storage cabinet can be monitored in real time through the smart phone, as shown in Fig. 11 below. When the storage cabinet is damaged artificially or accessed with abnormal violence, it can push the warning information to the user terminal. On some specific occasions, when the administrator is not on site and it is necessary for others to take the materials on behalf of others, the receiver sends an application on the airborne control terminal, and the designated storage drawer pops up after the administrator's remote terminal confirmation. The video monitoring can be turned on during the agent picking process to record whether the picking status of the picking operator matches the application authorization. This mode is especially suitable for off-site tool requisition and finished product and parts management.

Fig. 11. Mobile phone remote control interface

6 Conclusion

The design of the tower type intelligent storage cabinet realizes the rapid picking up and memory storage of materials. Through the multi prism tower body design, under the action of the control unit, with the positioning and rotation of the servo driver and servo motor, the problem of inconvenient access to materials inside the larger traditional storage cabinet is solved, which has a certain practical value. Its main features are as follows:

(1) The use of intelligent voice, touch screen, remote monitoring and other automatic control methods greatly facilitates people's daily life, and is suitable for the elderly, young people, children and other people of different ages.
(2) The increased security alarm system makes the storage of goods more safe and effective. When encountering illegal intrusion from the outside world, it can give an active or passive alarm to remind users of the occurrence of theft and help users reduce the loss of theft as much as possible.
(3) Compact structure, multi-function integration, modular function design is suitable for a variety of occasions. As an auxiliary and extension of the new home and intelligent office, the tower type intelligent storage cabinet has good adaptability.

Fund Project. Anhui university scientific research project - research and application of life cycle sustainable remanufacturing technology of bridge housing products based on strength redundancy (KJ2020A1106), Anhui provincial quality engineering project of colleges and Universities - strengthening the documentary evidence integration and Qualification Framework - pilot research and practice of "1+x" certificate system of industrial robot technology (2020jyxm0306). Anhui Province University excellent young talents support program project(gxyq2022275).

References

1. Yang, X., Dang, W., Wu, J.: Analysis on the relationship between the number of product reviews and the number of purchases based on symbolic regression. J. Syst. Eng. **35**(03), 289–300 (2020)
2. Zu, G., Pei, Y., Hou, P.: Review of aircraft battle damage assessment and repair estimation and design technology. Acta Aeronautica et Astronautica Sinica **41**(06), 87–108 (2020)
3. Ye, F., Song, J., Liang, J.: Platform design method based on storage products for kids. Packaging Eng. **41**(22), 121–129 (2020)
4. Li, X., Zhang, Y., Wang, Y.: Storage strategy of elderly cabinet items based on individual preference. Packaging Eng. **42**(16), 174–181+187 (2021)
5. Wang, Y., Gao, L., Ye, Z., et al.: An intelligent locker convenient for the disabled. Hubei province: CN209983780U,2020–01–24
6. Kulińska, E., Giera, J.: Identification and analysis of risk factors in the process of receiving goods into the warehouse. Foundat. Manag. **11**(1), 103–118 (2019)
7. Kordos, M., Boryczko, J., Blachnik, M., et al.: Optimization of warehouse operations with genetic algorithms. Appl. Sci. **10**(14), 1–28 (2020)
8. Istiqomah, N.A., Nadya, A.I., Putri, F.S., et al.: the implementation of barcode on warehouse management system for warehouse efficiency. J. Phys. Conf. Ser. **1573**(1), 1–6 (2020)
9. Fauzan, R., Raddlya, M.F., Raddlya, N.R.: The designing of warehouse management information system. In: IOP Conference Series: Materials Science and Engineering, vol. 879(1), pp. 1–7 (2020)
10. Shan, X., Gao, W., Liu, J., et al.: Research on the design of intelligent toys to cultivate the finishing habits of preschool children. Home Drama **31**, 88–89 (2019)
11. Lyu, Y., Zeng, W., Yan, J.: Rom "Thing" to "Heart": Psychological effect of dormitory arrangement. J. Ningbo Inst. Educ. **19**(06), 55–58 (2017)
12. Wu, M.: Labor is glorious, starting from learning to tidy up. Guangdong Educ. (Compreh. Ed) **06**, 66–67 (2021)

Research on Double Closed Loop Control Method of Single-Phase Inverter

Hao Yang(✉)

Department of Electronic Engineering, Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241002, Anhui, China
etbeat@163.com

Abstract. This paper presents a double-closed-loop PWM design and control method for single-phase inverter current inner loop and voltage outer loop. By establishing the mathematical model of the single-phase inverter, the current inner loop control can obtain rapid dynamic performance, and the voltage outer loop control can improve the steady-state performance of the system. Secondly, using the pole configuration method, the parameters of the double closed-loop PI can be obtained. Finally, the model is built by SIMULINK. The simulation results verify that the dual-loop control can improve and improve the steady-state performance and dynamic performance of single-phase inverter power supply.

Keyword: PWM inverter · Pole allocation method · filter · MATLAB simulation

1 Introduction

With the rapid development of technology, the fields of inverter applications are becoming more and more extensive, and the demand for inverter and UPS with inverters as the core of many fields is increasing. A indicator of the device. The traditional UPS control method is based on voltage single -closed control, and the dynamic response effect is not good. Therefore, this article uses a dual -closed control method to control the single -phase voltage PWM inverter. The rapid control of the output can improve the dynamic and stable performance of the system.

2 Single-Phase Full-Bridge Voltage Inverter

In single-phase inverter circuit, the most widely used or single-phase full-bridge voltage inverter circuit, IGBT as a switching device of the single-phase voltage inverter circuit as shown in Fig. 1, and LC filter and then connected to the load.

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 349–357, 2023.
https://doi.org/10.1007/978-981-99-2653-4_43

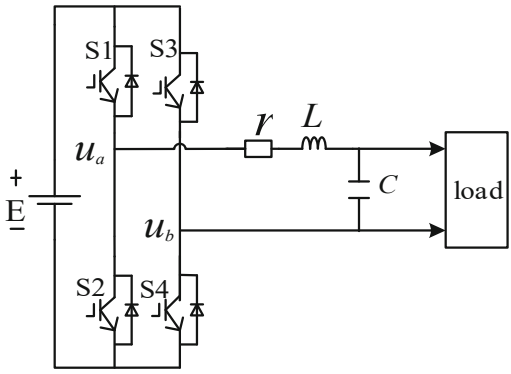


Fig. 1. Topology of single-phase voltage-type inverter

3 Modeling of Inverter

In order to study the model of inverter, single-phase half-bridge inverter is taken as the model, as shown in Fig. 2. Udc is dc bus voltage, S1 and S2 are IGBT switching devices, L and C are filter inductance and filter capacitance of filter respectively. R is the equivalent resistance of the line and inductance, r is the load, and $E = 2U_{dc}$ is assumed.

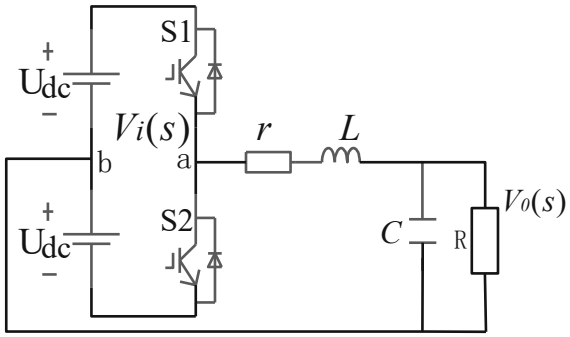


Fig. 2. Single-phase half-bridge inverter circuit

In SPWM control, sinusoidal wave Ur = Vrsin (ω T) is used as the modulation wave (Ur) and triangular wave as the carrier wave (Vc). The pulse obtained by Vr and Uc are compared to control the switching device. Since the switching state is discontinuous, the state space average method is adopted for analysis, that is, when the output frequency is far less than the switching frequency, the average value of variables is used to replace the instantaneous value within a switching cycle, that is, the average model under continuous state is obtained.

According to Fig. 2, assume that the voltage at two points A and B is Vi(s) and the output voltage is V0(s), then the transfer function G(s) from Vi(s) to V0(s) is:

$$\mathrm{G(s)} = \frac{V0(s)}{Vi(s)} = \frac{1}{LC\mathrm{s}^2 + (\frac{L}{R} + rC)s + 1 + \frac{r}{R}} \tag{2-1}$$

When r is ignored, (2–1) is simplified as:

$$\mathrm{G(s)} = \frac{1}{LC\mathrm{s}^2 + \frac{L}{R}s + 1} \tag{2-2}$$

When using bipolar SPWM modulation,

$$V_i(\mathrm{S}) = \mathrm{U_{dc}}(2\mathrm{S} - 1) \tag{2-3}$$

where S is the switching function.

That is, when S1 = conduction, s = 1; When S2 turns on, s is equal to 0.

Due to the existence of the switching function S, Vi of Eq. (2–3) is discontinuous.

Therefore, the average switching period of (2–3) can be obtained as follows:

$$< \mathrm{V_i} >_{\mathrm{ts}} = \mathrm{U_{dc}}(2 < \mathrm{s} >_{\mathrm{ts}} - 1) \tag{2-4}$$

<Vi> Ts indicates the switching period of Vi. And the average switching period of S is:

$$< \mathrm{S} >_{\mathrm{ts}} = \mathrm{D(t)} \tag{2-5}$$

where, D(t) is duty cycle. Obtained from literature [2]

$$\mathrm{D} = 0.5(1 + \mathrm{u_r}/\mathrm{V_c}) \tag{2-6}$$

Substitute (2–6) into (2–4) to get:

$$\frac{\langle \mathrm{V}_i \rangle_{Ts}}{Vr} = \frac{U\mathrm{dc}}{Vc} \tag{2-7}$$

Therefore, the transfer function from the modulation input to the inverter bridge output is:

$$\mathrm{K_{PWM}} = \frac{U_{\mathrm{dc}}}{V_c} \tag{2-8}$$

As can be seen from (2–8), in SPWM, When the output frequency is far less than the switching frequency, the inverter bridge can be regarded as a proportional link, and the proportional coefficient is KPWM.

Then, the transfer function from modulator Vr to inverter output Vo is

$$Go = \frac{V_{\mathrm{o}}}{V_r} = frac V_{\mathrm{o}} V_i \frac{V_i}{V_r} = \frac{1}{LCs^2 + (\frac{L}{R} + rC)\mathrm{s} + 1 + \frac{r}{R}} \frac{U_{dc}}{V_c} \tag{2-9}$$

4 Filter Parameter Settings

In this paper, second-order lc filter is adopted to design, and the equivalent resistance r of filter inductor is ignored. Then, the transfer function of filter output can be obtained as shown in Fig. 2–1:

$$G(s) = \frac{V_O}{V_i} = \frac{\frac{1}{LC}}{s^2 + \frac{1}{RC}s + \frac{1}{LC}} = \frac{\omega n^2}{s^2 + 2\zeta\omega n + \omega n^2} \tag{3-1}$$

where, $\omega_n = \frac{1}{\sqrt{LC}}$ is the undamped natural oscillation angular frequency and $\zeta = \frac{1}{2R}\sqrt{\frac{L}{C}}$ is the damping ratio.

Formula (3–1) is a typical second-order oscillation system, and the frequency characteristic is

$$\begin{aligned} G(j\omega) &= \frac{\omega n^2}{(j\omega)^2 + 2\xi\omega\omega n j + \omega_n^2} \\ &= \frac{\omega_n^2}{\omega n^2 - \omega^2 + j2\xi\omega\omega n} \\ &= A(\omega)e^{j\varphi(\omega)} \end{aligned} \tag{3-2}$$

Among them:

$$\varphi(\omega) = -\arctan\left(\frac{2\zeta\omega\omega n}{\omega n^2 - \omega^2}\right) \tag{3-3}$$

$$A(\omega) = \frac{1}{\sqrt{(1 - \frac{\omega^2}{\omega n^2})^2 + (\frac{2\zeta\omega}{\omega n})^2}} \tag{3-4}$$

Figure 3 is the second-order LC low-pass filter drawn by Matlab.

Can be seen in Fig. 3, when ω << ωn is a low-frequency band, A (ω) ≈0; when ω >>ωn is a high-frequency band, A (ω) ≈−40 lg (ω/ωn), which indicates that the high frequency gradient of L (ω) is a straight line with a slope of −40 db/dec. ωn is the turning frequency of the oscillating link.When ξ <0.707, the peak value appears on several frequency curves. From Fig. 3–1 and the above analysis values, the factors that affect the filtering effect are the turning frequency ωn and the damping ratio, so the turning frequency of the LC filter of the SPWM inverter fn = ωn/2π is far lower than the switch frequency FS, and the frequency FS, which is far lower than the switch frequency FS.Then the filter has inhibitory effects on the switching frequency and the frequency of nearby, so the turning frequency of the LC filter should be taken as (3–5).

$$10f_0 \le \frac{1}{2\pi\sqrt{LC}} \le \frac{f_s}{10} \tag{3-5}$$

f_0 is the fundamental frequency.

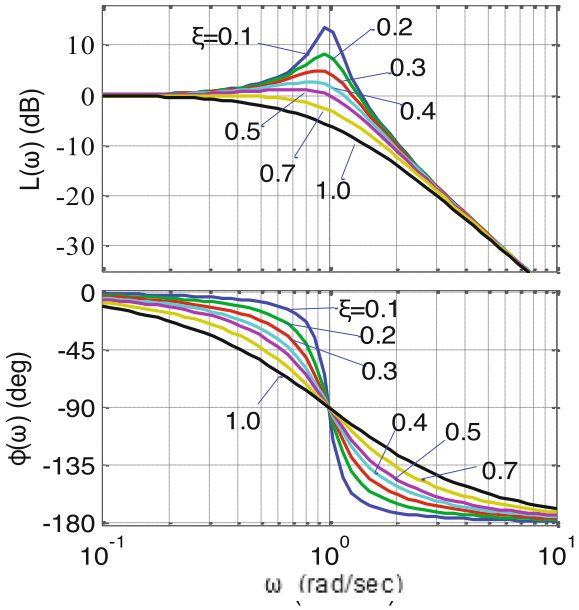


Fig. 3. The second-order LC low-pass filter of Berdu

Literature [3] gives the formula for the value of L, namely, Eq. (3–6).

$$\Delta i_{\mathrm{LMAX}} = \frac{E}{8Lf_s} \tag{3-6}$$

Among them, Δi_{LMAX} is the maximum pulse current in L, and its value is 10%to 20%of the output rated current.

5 PWM Dual-Closed Control Strategy

5.1 Dual Closed-Loop Control

Drawing 4 is a single-phase inverter dual-closed control box diagram. The current inner ring uses capacitor current IC because IC advance capacitor voltage UC can correct UC in advance, so the capacity of the loading capacity canpowerful. In the figure, the GV is the voltage regulator, the GI is the current regulator, and the KPWM is the equivalent gain of the inverter bridge.

The output voltage u_0 is compared with a feedback signal and a given voltage signal u_{ref} to form a instant error regulating signal u_{e}, and then the current as a current given value i_{ref} after passing the PI regulator. It is different from the current feedback value i_{C}, (Fig. 4)

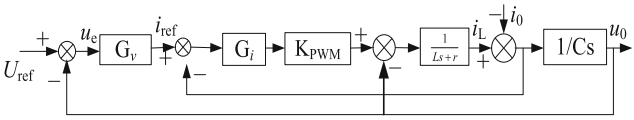


Fig. 4. Dual closed-loop control box diagram

The instantaneous error signal is formed, and the current error control signal is generated by the current PI regulator. The signal is compared with the triangular carrier signal to generate SPWM signal to control the switching device of the inverter (S1~S4). Finally, the output sinusoidal voltage $u0$ is filtered by LC filter.

5.2 The Polar Configuration and Performance Analysis of the Control System

For the design method of controlling parameters, you can be designed according to the principle of automatic control, but you need to consider the interaction and cooperation of the response speed and frequency bandwidth between the two regulators. However, this method is complicated.EssenceAt the same time, it can also be designed according to the theory of modern control, and the polar point configuration method [4] is used to design the parameters of PID. This method is to calculate the PID controller parameter calculation with the closed -loop control system performance indicators, and the parameter calculation is more accurate.When the capacitor current inner ring is adjusted by proportion, the voltage outer ring uses a proportional integral controller:

$$\begin{cases} G_v{=}k_{vp} + \dfrac{k_{vi}}{s} \\ G_i{=}k_{ip} \end{cases} \tag{4-1}$$

The characteristic equation of the inverter closed -loop system is:

$$D(s) = LCs^3 + (rC + k_{ip})Cs^2 + (k_{vp} + k_{ip})s + k_{vi}k_{ip} \tag{4-2}$$

The method of designing the polar configuration design controller:

From the formula (4–2), we can see that there are 3 roots of the closed-loop feature equation, that is, there are 3 poles of the closed-loop transmission function. According to the principle of automatic control, the distribution of the polar point of the closed -loop system in the S plane determines the stability of the system. The dynamic characteristics of the system determine the system's closed -loop pole and closed -loop zero point.However, the dynamic characteristics of the system are basically determined by the dominant pole point. The leading poles are defined as a closed -loop pole that leads the leading role throughout the time, that is, the closed -loop pole that is close to the virtual axis and is not very close to the zero zero point of the closed loop.

Assuming that the expectations of the closed -loop system at this time are: $s_{1,2} = -\xi\omega_n \pm j\omega_n\sqrt{1-\xi^2}$, non -dominated poles are $s_3 = n\xi\omega_n$, and generally n take 5–10. At this time, the expected characteristic equation of the dual -closed control system is:

$$D_r(s) = (s^2 + 2\xi\omega_n s\omega_n^2) \cdot (s + n\xi\omega_n) \tag{4-3}$$

By comparing Eqs. (4–2) and (4–3) and making corresponding coefficients equal, we can obtain:

$$\begin{cases} k_{vp} = \dfrac{(1+2n\xi 2)\omega_n^2 LC - 1}{k_{ip}} \\ k_{ip} = (2+n)\xi\omega_n L - r \\ k_{vi} = \dfrac{n\xi\omega_n^3 LC}{k_{ip}} \end{cases} \tag{4-4}$$

6 Simulation Experiment Analysis

The simulation model parameters are mainly as follows: DC bus voltage E = 380V, rated output voltage UO = 220V, output rated frequency f = 50 Hz, rated output power p = 10 kW, filtering inductance L = 660 μH, filter capacitor C = 200 μF, equivalent damping resistance R = 0.1Ω, switch frequency fs = 10 kHz, expects the natural oscillation frequency ωn = 3200 rad/s, n = 10, from the formula (4–4):

$$k_{vp} = 0.56,\ k_{vi} = 1441,\ k_{ip} = 0.01237$$

single phase voltage all-phase voltage all-phase voltage all-phase voltage allThe simulation model of the bridge inverter is shown in Fig. 5.

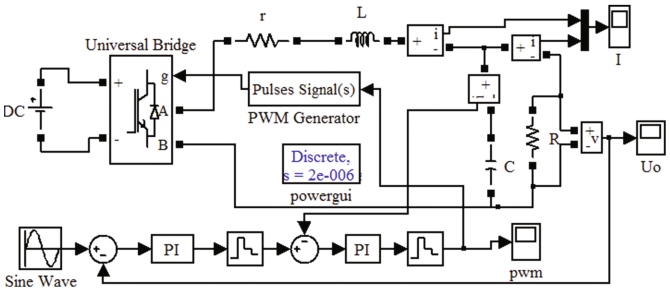


Fig. 5. The simulation model of a single-phase voltage inverter

First simulates the voltage and current output waveform at full load, the voltage and current output waveforms in Fig. 6.

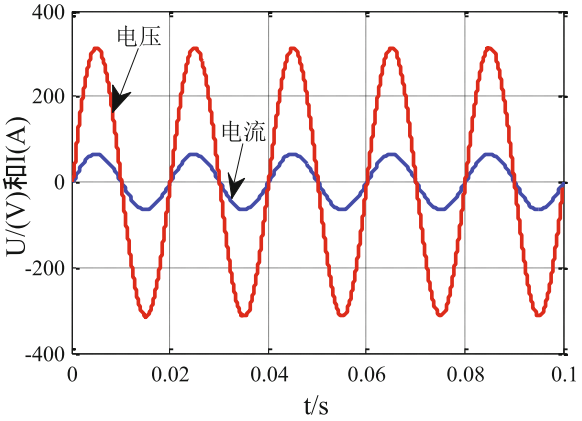


Fig. 6. The voltage and current waveforms at full load

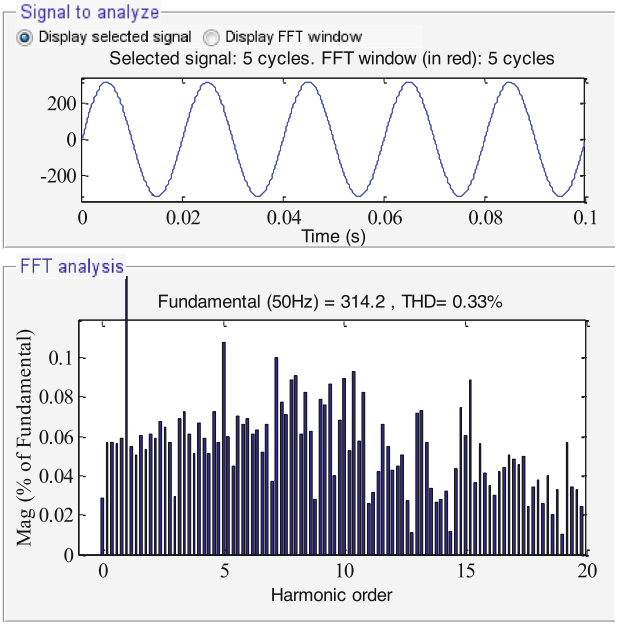


Fig. 7. The voltage THD analysis at full loadIt

Can be seen from Fig. 7 that the output voltage and current waveforms are sine waves, and THD is less than 5%than the standards prescribed by the state.

7 Conclusion

This article first establishes a model of a single -phase inverter, a dual -closed -loop control PI inverter control system is designed, and parameters of the PI design are used

to design PI.Secondly, the simulation waveform of the single -phase inverter was used to use Matlab. It can be obtained through the simulation waveform. Dual -closed controls make the inverter's output voltage steady stability accuracy, good dynamic performance, and small deformation.

Acknowledgments. This work is supported by the 2022 Excellent top-notch Talents Training program of Anhui Universities--Academic program for top-notch talents in disciplines (majors) in Colleges and Universities.

References

1. Wang, Z.: Power electronics technology. China Machine Press two thousand and nine, Beijing
2. Xudehong Modeling and control of power electronic system. China Machine Press two thousand and six, Beijing
3. Wang, B.: Double loop control design of single-phase SPWM inverter based on LC filter. Electrical Eng. (04), 21–25+34 (2017)
4. Deng, J., Cong, S.: Application and performance comparison of several self-tuning controls in control engineering. In: Proceedings of the 2006 Academic Exchange on System Simulation Technology and Its Applications, pp. 759–762 (2006)
5. Hu, S.: Automatic control principle 6th Edition. Science Press two thousand and eighteen, Beijing

Remote Sensing Image Mosaic Technology Based on Improved SIFT Algorithm

Meng Wang(✉) and Feng-kun Wang

Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241002, Anhui, China
ahjdwm@ahcme.edu.cn

Abstract. To solve the problem of a large amount of computation and low efficiency of the traditional remote sensing image mosaic algorithm, an Improved SIFT feature matching image mosaic algorithm is proposed. Aiming at the disadvantage of too many dimensions of feature point descriptors in SIFT algorithm, a dimension reduction scheme is proposed. Centered on the feature point, concentric circles are constructed with the radius of 2 pixels, 4 pixels and 6 pixels respectively. The region is divided according to the distance from the feature point. If the distance is close, the weight is high, and if the distance is far, the weight is low. The region is divided into 4, 8, and 16 partitions, so as to generate a 56-dimensional descriptor. Furthermore, the feature points in the image matching are divided into two data sets, namely, maximum and minimum. The best matching point is selected by calculating the distance between the maximum and minimum, which reduces the amount of calculation of image matching and saves the algorithm time. Finally, the weighted smoothing method is used for image mosaic. Experimental tests were carried out on the three situations of remote sensing image scaling, rotation, and gray change. The results show that the proposed algorithm can effectively reduce the calculation amount and accelerate the stitching speed, and the stitching effect is ideal.

Keywords: SIFT algorithm · Feature point descriptor · Feature matching · Image mosaic · Remote sensing image

1 Introduction

Satellite remote sensing technology is widely used in military reconnaissance, surface analysis, resource survey, and other fields. In the process of satellite remote sensing image processing, due to the limited perspective of a single image, it is often necessary to splice two or more images to obtain a wider field of vision. Remote sensing image mosaic refers to the merging of two or more images with overlapping areas by overlapping the same parts while keeping the other parts unchanged, to form a whole image. The key to fusing the same part into a mosaic of satellite remote sensing images is how to make the computer recognize the same part in different remote sensing images. At the same time, in the shooting process, because the camera motion is not completely stable, it is necessary to solve the problems of translation, scaling, and distortion of the same area

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 358–365, 2023.
https://doi.org/10.1007/978-981-99-2653-4_44

of the two images. Therefore, it is necessary to make corresponding adjustments to the image, to achieve better image stitching effect [1].

At present, the most commonly used method for image feature matching in image mosaic is SIFT(scale-invariant feature transform) algorithm. SIFT algorithm can adapt to different scales, illumination, and rotation angles, and has affine invariance and high matching accuracy. However, SIFT algorithm also has obvious shortcomings: the feature point description sub-dimension is too large, increasing the amount of data, time-consuming, and seriously affecting the efficiency of the algorithm.

2 Related Works

The literature proposed the PCA-SIFT algorithm for principal component analysis (PCA) of SIFT algorithm. Compared with SIFT, the performance of this algorithm in running time and affine change is improved, but it is more sensitive to scale transformation and smooth blur, and the amount of calculation is larger [2]. Aiming at the shortcomings of SIFT algorithm and binary descriptor, a binary local feature descriptor based on SIFT is proposed. The Hamming distance is used to replace the Euclidean distance to measure the similarity between the quantized binary descriptors to improve the matching efficiency. In the matching stage, the descriptors are divided into two parts to match them respectively. In this way, some invalid matching feature points can be eliminated through the initial matching, and the matching time is further shortened [3]. Euclidean distance is used to adjust the ratio threshold of the nearest neighbor (NN) to the next nearest neighbor (SCN), to reduce the mismatch, but it is also easy to lose some correct matching points, which cannot improve the matching rate [4]. A median filtering algorithm based on feature points is proposed, but the algorithm cannot eliminate the error of matching feature points and is time-consuming [5]. The initial iterative feature point pair of RANSAC is detected by median filtering, but the mismatched pair is not eliminated, and the efficiency of the algorithm is not significantly improved [6].

To reduce the complexity of SIFT algorithm in image mosaic, improve the correct matching rate and reduce the algorithm time, this paper proposed to improve the SIFT algorithm. Firstly, a new feature descriptor is designed to reduce the dimension of the high-dimensional descriptor in the original algorithm; Then, the feature points are divided into maximum value and minimum value. Calculate the distance between the feature points of the same font between two images to reduce the amount of image matching calculation. The weighted average fusion method is used for image fusion to improve the effect of image mosaic.

3 Methodology

Image mosaic based on SIFT algorithm includes the following steps. Firstly, SIFT algorithm is used for image feature extraction and feature matching. Secondly, the RANSAC algorithm is used to eliminate outlier matching pairs. Finally, an image mosaic is performed [7].

3.1 SIFT Algorithm

SIFT algorithm in image matching, need to find two images with characteristic critical point, that is, the key point. The purpose of doing this is to accurately locate the position of key points, which are guaranteed to remain unchanged in the process of image scaling, rotation and gray level change. Further processing and screening of key points will serve as the basis for the next processing. The SIFT algorithm generates descriptors for each key point in the primary and secondary directions. Furthermore, the similarity is judged according to the Euclidean distance between descriptors, so as to match the feature points.

3.2 RANSAC Algorithm

RANSAC algorithm steps are as follows:

(1) according to confidence probability P and data error rate ω, and the minimum amount of data m required to calculate the model parameters, and calculate the sampling quantity m to be selected from Eq. (2):

$$1 - (1 - (1 - \varepsilon^m))^M = P \tag{1.1}$$

(2) randomly select the original data samples, estimate the sample number of each sample as the minimum data required by the model parameters, and calculate the corresponding model parameters of the samples.

(3) use all the original data to test the parameters of the model and get the correct number of parameters of each model; Repeat steps (2) and (3) until the processing of group M sampling is completed.

(4) Find all the interior points corresponding to the optimal model parameters, and use these data to calculate the final parameters of the model.

3.3 Algorithm Improvement

3.3.1 Dimensionality Reduction of Feature Descriptor

The original sift algorithm uses a total of 256 pixels in the 16 * 16 rectangular area around the feature points to determine the descriptor of the feature points. Finally, each feature point generates a 4 * 4 * 8 = 128-dimensional descriptor, which requires a large amount of computation and takes a long time to implement the algorithm. The feature descriptor in SIFT algorithm is improved.

(1) The feature points are divided into 3 circular regions with a radius of 2 pixels, 4 pixels and 6 pixels. The smaller the radius, the higher the weight of the feature points in this area. Based on this, three rings are divided into regions of different sizes. The ring with a radius of 2 pixels is divided into 16 regions, 4 pixels into 8 regions, and 6 pixels into 4 regions.

(2) In order to reduce the influence of illumination, the gray level in the region is normalized and the average gray level is obtained.

$$\bar{g}_i = \frac{g_i}{\sqrt{\sum_{j=1}^{8} g_i}} \quad i = 1, 2, \ldots, 8 \tag{1.2}$$

(3) Then the gray difference value in each region is calculated.

$$d\,\bar{g}_i = \begin{cases} \left|\bar{g}_i - \bar{g}_{i+1}\right| & i = 1 \\ |2g_i - g_{i-1} - g_{i+1}| & 1 < i < 8 \\ \left|\bar{g}_i - \bar{g}_{i-1}\right| & i = 8 \end{cases} \tag{1.3}$$

(4) Finally, the feature point descriptor is composed of the gray mean value and gray difference value of each region.

$$g = (g_1 + g_2 + \ldots + g_{28} + \bar{g}_1 + \bar{g}_2 + \ldots + \bar{g}_{28}) \tag{1.4}$$

Each feature descriptor has: $(4 + 8 + 16) * 2 = 56$ dimensions. Compared with SIFT algorithm, each feature descriptor reduces data by 56%.

3.3.2 Feature Point Matching and Selection

The distance calculation of feature points in SIFT algorithm is improved. The feature points are the corresponding extreme points in DoG. Therefore, the maximum and minimum values of the extreme points are put into two data sets respectively. In the process of calculating the distance between the feature points of two images, we only need to calculate the distance between the maximum data set and the minimum data set, so that we can only calculate 1/2 times.

3.4 Image Fusion

After image registration, if it is only a simple superposition, the image will be blurred and have obvious stitches, resulting in poor splicing. This paper uses the weighted smoothing method.

The specific idea is: suppose that $I_1\,(x, y)$ and $I_2\,(x, y)$ is the image sequence to be spliced, $I\,(x, y)$ is the fused image, and $d_1\,(x, y)$ and $d_2\,(x, y)$ are the weights of the image to be spliced, so the formula of the algorithm is expressed as follows:

$$I(x, y) = \begin{cases} I_1(x, y), (x, y) \in I_1 \\ d_1(x, y) * I_1(x, y) + d_2(x, y) * I_2(x, y) \\ (x, y) \in (I_1 \cap I_2) \\ I_2(x, y), (x, y) \in I_2 \end{cases} \tag{1.5}$$

where d_1 and d_2 represent the weight value, which is different from the width of the overlapping area, and $d_1 + d_2 = 1$, $0 < d_1, d_2 < 1$. In the image overlapping area, the

weight d_1 *(x, y)* slowly changes from 1 to 0, and d_2 *(x, y)* slowly changes from 0 to 1, thus realizing the smooth transition between the two images to be spliced I_1 and I_2.

d_1 and d_2 are calculated as follows: let the abscissa of the current pixel be x_i, and the abscissa of the left and right boundaries of the overlapping area be x_1 and x_r respectively.

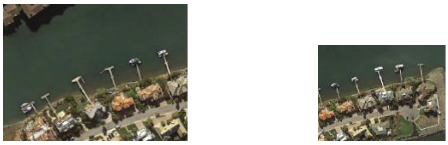
$$d_1 = \frac{x_r - x_i}{x_r - x_1}, d_2 = 1 - d_1 = \frac{x_r - x_1}{x_r - x_1} \tag{1.6}$$

4 Results and Discussion

This algorithm is based on Intel corei7–4510, 2G 4-core CPU and 8GB memory windows10 operating system, OpenCV 4.5 platform, and python 3.06 language development and implementation. The data of remote sensing images are selected from the DOTA-v1.5 dataset.

Image mosaicking experiments are carried out for rotation, scaling, and gray change. Figure 1 shows three groups of test images, including left and right test images of rotation, zoom, and gray transformation (Table 1).

(a) Rotation (left-right comparison)

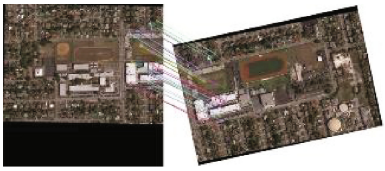


(b)Zoom (left-right comparison)

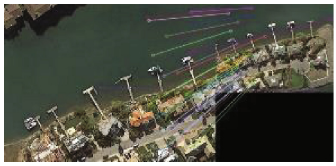
(c) Gray transformation (left-right comparison)

Fig. 1. Three groups of test images

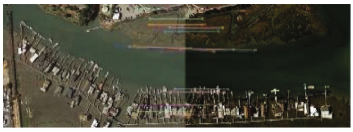
The Improved SIFT algorithm is used for feature matching, and the results are shown in Fig. 2.



(a) Rotation



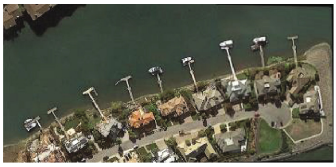
(b)Zoom



(c) Gray transformation

Fig. 2. Feature matching image

(a) Rotation



(b)Zoom

(c) Gray transformation

Fig. 3. Results of image mosaic

The image mosaic results are shown in Fig. 3.

Table 1. Experimental results of the original algorithm and the improved algorithm

Test images	Number of feature points		Correct matching rate		Time-consuming (S)	
	Original algorithm	Improved algorithm	Original algorithm	Improved algorithm	Original algorithm	Improved algorithm
Rotation	864	631	65.25%	74.31%	20.15	17.62
Zoom	316	186	76.35%	81.25%	9.75	7.14
Gray transformation	1101	842	68.73%	84.56%	24.31	19.50

The matching accuracy of the improved algorithm is about 7% higher than that of the traditional algorithm. In the aspect of matching speed, the improved algorithm saves 19% more time than the traditional algorithm.

5 Conclusions

The image mosaic algorithm based on SIFT feature matching makes up for the defects of the traditional image mosaic algorithm in the changes of scale and illumination. However, the matching feature points extracted by this algorithm contain a large number of mismatched feature points, so the matching workload is large and time-consuming. Therefore, two improvements are proposed. One is to propose a scheme to reduce the sub-dimension of feature points. The other is to reduce the amount of distance calculation by dividing the feature points into two groups: the maximum value and the minimum value. Finally, the weighted smoothing method is used for image mosaic.

The three remote sensing image mosaic scenarios of rotation, scaling and gray change were tested. The results show that the proposed algorithm can effectively reduce the calculation amount and accelerate the stitching speed, and the stitching effect is ideal.

Acknowledgements. Key project of natural science research in colleges and universities of Anhui Province (KJ2020A1102).

Anhui teachers' teaching innovation team (2021jxtd065).

References

1. Tong, L., Hu, L., Wang, Y., CHu, J.: Satellite remote sensing image mosaic based on convolutional neural network. J. Jilin Univ. (Sci. Ed.) **60**(1), 99–108 (2022)
2. Ke, Y., Sukthankar, R.: PCA-SIFT: a more distinctive representation for local image descriptors. In: IEEE Computer Society Conference on Computer Vision & Pattern Recognition (2004)
3. Qian, L., Jiang, Z.: Binary quantized SIFT feature descriptors for optimized image stitching. J. Image Graph. **21**(12), 1593–1601 (2016)

4. Kasar, T., Ramakrishnan, A.: Block-based feature detection and matching for mosaicing of camera-captured document images. In: TENCON 2007 - 2007 IEEE Region 10 Conference (2007)
5. Beiji, Z., Peng, R., Yao, X., Jiabei, H.: An automatic panoramic images mosaic algorithm with precise matching, Comput. Eng. Sci. **32**(8), 60–63 (2010)
6. Nicolas, H.: New methods for dynamic mosaicking. IEEE Trans. Image Process. **10**(8), 1239–1251 (2001)
7. Fei, L., Weiqiong, Y., Pengcheng, W., Jinjin, W., Keke, X.: Image mosaic algorithm based on SIFT feature matching. J. Hunan Univ. Technol. **28**(1), 71–75 (2014)
8. Sunan, D., Qiuju, Z.: Image registration method based on improved SIFT algorithm. Transducer Microsyst. Technol. **39**(10), 45–50 (2020)

Open-Set Human Micro-doppler Action Recognition by Extreme Value Theory

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

University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave., Chengdu 611731, China
sixu0210@gmail.com

Abstract. In this paper, an open-set radar micro-Doppler human action recognition method based on extreme value theory is proposed, which aims to solve the task of rejecting unknown actions in real-world application. Our method uses the idea of extreme value theory from statistics to achieve the recognition of unknown classes, and the loss function of metric learning is used to further improve the recognition ability of open-set tasks. Experimental results on public datasets show that our method can effectively solve the micro-Doppler open-set recognition problem compared with the traditional methods.

Keywords: open-set · radar micro-Doppler · action recognition · metric learning · extreme value theory

1 Introduction

Human activity recognition (HAR) has been used more and more frequently in the fields of recreation, security and medicine in the past few decades. Radar micro-Doppler spectrum can represent the characteristics of object movement to a certain extent and due to the advantages of radar's robustness to the environment, penetrability and visual privacy protection, radar has been increasingly used for HAR and replacing the traditional optical image recognition in some extreme environments [1–3].

Deep learning has been widely used in radar HAR tasks due to its powerful ability to extract features from complex data. Recently most methods based on deep learning can perform well in close sets, that is, the training label is consistent with the test label. However, the closed set assumption barely holds in real-world applications. In the real world, the types of activities are nearly infinite, so the number of unknown actions received in the practical application may be far greater than the known actions used in the training process. Therefore, it is important to develop an effective open-set radar human activity recognition method, which can not only correctly classify the known actions, but also recognition the unknown actions. Figure 1. shows a simple comparison between close-set and open-set micro-Doppler HAR tasks. The open-set recognition task

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 366–373, 2023.
https://doi.org/10.1007/978-981-99-2653-4_45

has been studied to some extent in the field of computer vision [4–6], but it has not been applied more in the field of human micro-Doppler action recognition. Yang [7] introduced the open-set recognition problem into radar HAR task for the first time based on Generative Adversarial Network.

In this paper, we apply a deep learning network based on CNN to extract the features of radar human action and apply the idea of extreme value theory to the model classification task, so as to realize the open-set radar HAR function.

Our main contributions are as follows:

1. We construct an open-set recognition model base on extreme value theory, which realize the capability of radar HAR in an open environment.
2. The loss function based on metric learning is applied in our model to further improve the accuracy of action classification and the ability to distinguish unknown classes in open-set recognition.

The rest of this paper is arranged as follows. Section 2 describes an open-set recognition algorithm we used in our paper and how to further improve the accuracy of the algorithms by using metric learning. Section 3 discusses the accuracy of the proposed algorithm on public micro-Doppler human action datasets and compares our method with some other HAR methods. Section 4 is the conclusion.

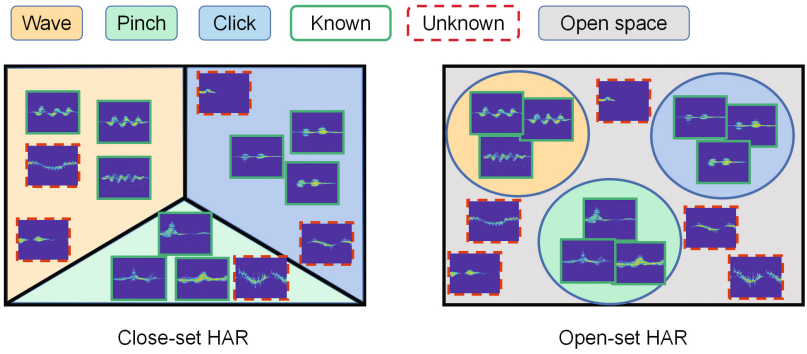


Fig. 1. Close-set and Open-set HAR

2 Open-Set HAR Algorithm

In this section, we propose an open-set micro-Doppler HAR framework. This framework can be divided into two stages: First, a CNN-based deep learning network is introduced to extract the feature from the time-frequency images of radar echo signals. Then, a metric learning loss function is used to make the distribution of extracted features more suitable for open-set scenes. Last, our model will use the extracted features to fit an EVT model and by this model, the framework can refuse the unknown sample and properly identify the known sample.

2.1 Metric Representation Learning

For open-set recognition tasks, having a powerful representation learning model to extract sample features can improve the ability of the network to distinguish between known and unknown samples. In our method, we introduced Large Margin Cosine Loss (LMCL) [8] from face recognition field to help us build this model.

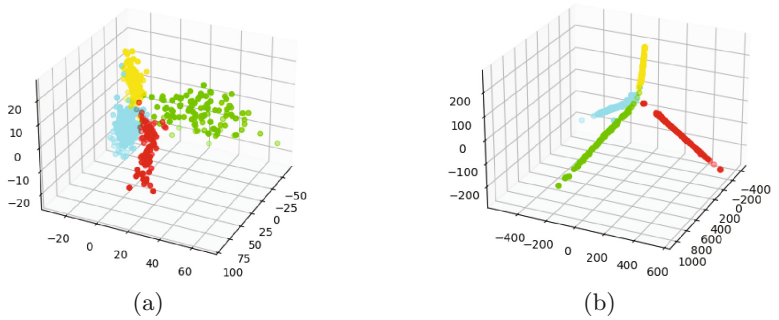
LMCL is improved on the original Softmax loss to make the feature of samples more separate and discriminative in the embedding space. First, normalize the weight of the original Softmax loss with L2 normalization, setting the bias to 0 and fixing the input to S. In this case, the normalized Softmax loss function is as follows:

$$L_{ns} = \frac{1}{N}\sum_{i} -log\frac{e^{scos(\theta_{y_i},i)}}{\sum_j e^{scos(\theta_j,i)}} \tag{1}$$

As we can see from the above formula, the loss function converts the Softmax from optimizing inner product to optimizing cosine angle. At the same time, in order to make the embedded features more discriminative, LMCL introduced cosine margin to make the sample of the same category still belong to this category after subtracted a margin, and the final LMCL is:

$$L_{lmc} = \frac{1}{N}\sum_{i} -log\frac{e^{s(cos(\theta_{y_i},i)-m)}}{e^{s(cos(\theta_{y_i},i)-m)} + \sum_{j\neq y_i} e^{scos(\theta_j,i)}} \tag{2}$$

In Fig. 2, PCA dimension reduction is used to show the comparison of feature distribution between Softmax loss and LMCL in the embedding space. The visualization result shows LMCL can extract more separate and discriminative features.



(a) (b)

Fig. 2. Visualization result of the embedded feature: (a) Softmax loss (b) LMCL

2.2 Open-Set Action Recognition Based on Extreme Value Theory

EVT theory is applied to the analysis of the abnormal distribute data in statistics. In the work of Bendale [5], the author replaces Softmax layer with Openamx layer using the concept of EVT, and in Openmax layer the known sample can be fitted into a Weibull distribution, and the fitted Weibull model can determine the probability of failure to classify in close-set classification and use this probability to adjust known sample score and calculate unknown sample score. Also, EVT model can effectively detect out-of-distribution samples when it is applied to the feature representations learned by deep learning models [9]. In this paper, we propose a deep learning framework to extract the feature of radar HAR time-frequency image, and fitted into a Weibull distribution to achieve open-set recognition.

Algorithm 1. Open-set recognition based on EVT

STEP 1:
INPUT: feature of training sample X_i^s, tail-size μ
for c=1...N **do**
Compute class center by mean of features in one class, $m_c = mean(X_{i,c}^s)$
Compute distance between each sample and class center, $d_{i,c}^s = 1 - cos(m_c, x_{i,c}^s)$
Fit Weibull model using $d_{i,c}^s$ and μ, get Weibull distribution parameters $(\varsigma_c, \kappa_c, \lambda_c)$
end for
STEP 2:
INPUT: feature of test sample Y_j^t, threshold θ
for c=1...N **do**
Compute distance between test sample and class center $d_{j,c}^t = 1 - cos(m_c, Y_{j,c}^t)$
Weibull CDF: $\omega_c(d_{j,c}^t) = 1 - exp(-\frac{||d_{j,c}^t - \varsigma_c||}{\lambda_c})^{\kappa_c}$
end for
Reject input if $\omega_c(d_{j,c}^t) > \theta$ for any class c.

The algorithm to fitting the model and determine the unknown class we used in our paper is shown in Algorithm 1. In step one, for each class, we use the distance between known samples and class center to fit Weibull distribution, the class center is the mean of all the samples in the class and we use cosine distance to replace the Euclidean distance in the original paper [9] because we use LMCL training the embedding network. In step two, we calculate the cosine distance between test samples and every class centers, using this distance we can calculate the CDF of Weibull distribution and get the probability of test samples belong to this class, then decide whether this sample belong to unknown class according to the threshold manually set. Tail size and threshold need to be set before building the model.

3 Experiment and Analysis on Open-Set HAR

3.1 Data Preparation and Evaluation Metrics

In this paper, we use FMCW human activity dataset from University of Glasgow [10] and FMCW human gesture dataset from University College London [11]. UoG dataset contains five types of human motion echo signals including walking, sitting down, standing up, picking up an object and drinking water, UCL dataset contains swipe, click, pinch and wave. Short-time Fourier transform (STFT) can extract the frequency characteristic and get the micro-Doppler diagram. Four micro-Doppler spectrums from two dataset has shown in Fig. 3.

The spectrums from two datasets have similar characteristic to some extent. Therefore, we use the combination of two datasets as the experiment data. We choose UCL's four actions as the known class and UoG's five actions as the unknown class.

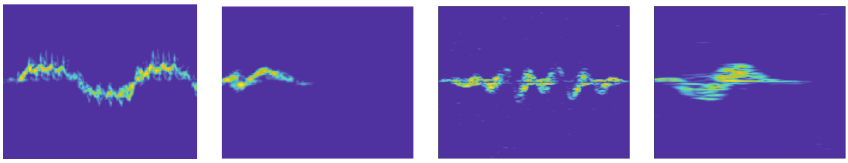


Fig. 3. Human micro-Doppler spectrum: (a)walking (b)setting down (c)wave (d) click

In the open-set recognition problem, [4] proposes "Openness" to measure the grade of an open-set problem which can be calculated using the following formula:

$$O = 1 - \sqrt{\frac{2 * |C_{TR}|}{|C_{TR}| + |C_{TE}|}} \tag{3}$$

where C_{TR} is the number of known class in training stage and C_{TE} is the number of all the class in testing stage. The increasing number of unknown class in test set will cause the Openness increasing. And in our experiment, the Openness will increase from 0 to 0.22.

3.2 Implementation

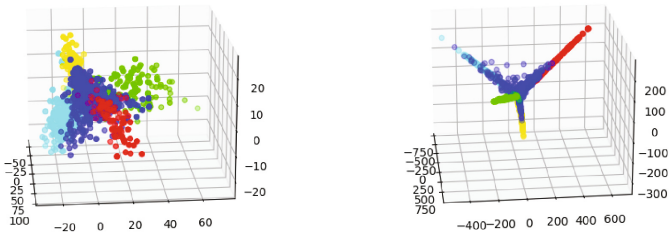
In this method, we use Vgg16 [12] as the backbone to extract sample features and it will output a feature with 128 dimensions. In the design of our EVT model, we set tail size as 20 and the threshold is set to 0.9.

Figure 4 shows the distribution of micro-Doppler spectrum features in the embedding space using Softmax and LMCL in open-set scene. It can be seen that LMCL can obtain more separate and discriminative feature distribution, so

it can provide more space for unknown class in embedding space and improves the ability of the network to distinguish known class from unknown class.

In our experiment, we compare our method with close-set Softmax and LMCL, Softmax with EVT and the original OpenMax [5]. To make a fair comparison, we test all these methods on the same backbone network and training parameters. The test result is shown in Fig. 5.

From our test result, we can see that methods without open-set modeling have a linear decreasing accuracy with Openness, illustrate the traditional HAR methods can not reject any unknown sample in open-set scene. And our method achieves 87% accuracy at 0.22 Openness which is 7% better than Softmax with EVT modeling and 14% better than Openmax. Also, our method has a smaller accuracy decrease with increasing Openness which shows the robustness of our model.



(a) (b)

Fig. 4. Visualization result of the embedded feature in open-set (blue is the unknown samples): (a) Softmax loss (b) LMCL

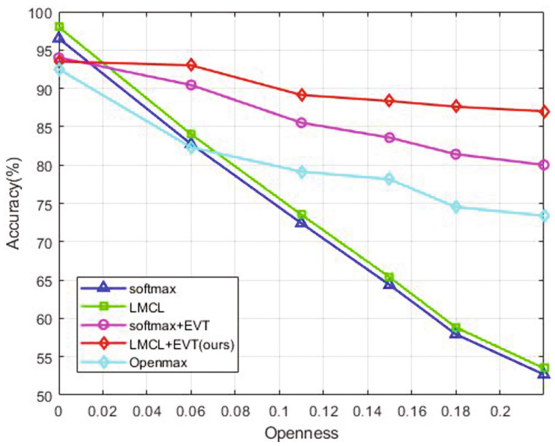


Fig. 5. Comparison results of different models

4 Conclusion

In this paper, we analyze radar micro-Doppler HAR task in open world scene and propose a open-set micro-Doppler recognition methods. Our method use LMCL loss function from metric learning to extract more separate and discriminative feature in embedding space. Then we use EVT method to fit these features into a Weibull distribution and build a model to achieve the ability of rejecting unknown samples. The experiment result shows our method achieve more than 30% higher accuracy compare with traditional close-set micro-Doppler HAR methods and 7%–14% higher than other open-set recognition methods. In the future, we will focus on how to make our network dynamically study from the environment and recognition unknown actions from open world.

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

References

1. Kim, Y., Moon, T.: Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks. IEEE Geosci. Remote Sens. Lett. **13**(1), 8–12 (2016). https://doi.org/10.1109/LGRS.2015.2491329
2. Javier, R.J., Kim, Y.: Application of linear predictive coding for human activity classification based on micro-doppler signatures. IEEE Geosci. Remote Sens. Lett. **11**(10), 1831–1834 (2014). https://doi.org/10.1109/LGRS.2014.2311819
3. Bai, X., Hui, Y., Wang, L., Zhou, F.: Radar-based human gait recognition using dual-channel deep convolutional neural network. IEEE Trans. Geosci. Remote Sens. **57**(12), 9767–9778 (2019)
4. Geng, C., Huang, S.-J., Chen, S.: Recent advances in open set recognition: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **43**(10), 3614–3631 (2021). https://doi.org/10.1109/TPAMI.2020.2981604
5. Bendale, A., Boult, T.E.: Towards open set deep networks. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1563–1572 (2016). https://doi.org/10.1109/CVPR.2016.173
6. Oza, P., Patel, V.M.: C2AE: class conditioned auto-encoder for open-set recognition. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2302–2311 (2019). https://doi.org/10.1109/CVPR.2019.00241
7. Yang, Y., Hou, C., Lang, Y., Guan, D., Huang, D., Xu, J.: Open-set human activity recognition based on micro-Doppler signatures. Pattern Recogn. **85**, 60–69 (2019). ISSN 0031-3203. https://doi.org/10.1016/j.patcog.2018.07.030
8. Wang, H., et al.: CosFace: large margin cosine loss for deep face recognition. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition, vol. 2018, pp. 5265–5274 (2018). https://doi.org/10.1109/CVPR.2018.00552
9. Mundt, M., Pliushch, I., Majumder, S., Ramesh, V.: Open set recognition through deep neural network uncertainty: does out-of-distribution detection require generative classifiers?. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), pp. 753–757 (2019). https://doi.org/10.1109/ICCVW.2019.00098

10. Fioranelli, D.F., Shah, D.S.A., Li, H., Shrestha, A., Yang, D.S., Kernec, D.J.L.: Radar sensing for healthcare. Electron. Lett. **55**(19), 1022–1024 (2019)
11. Ritchie, M., Capraru, R., Fioranelli, F.: Dop-NET: a micro-Doppler radar data challenge. Electron. Lett. **56**, 568–570. https://doi.org/10.1049/el.2019.4153
12. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 1409.1556

Author Index

Q. Liang et al. (Eds.): CSPS 2022, LNEE 872, pp. 375–376, 2023.
https://doi.org/10.1007/978-981-99-2653-4