



A Target Recognition Algorithm of Multi-Source Remote Sensing Image Based on Visual Internet of Things

Xue-jun Sun¹ · Jerry Chun-Wei Lin²

Accepted: 13 December 2021 / Published online: 5 February 2022
© The Author(s) 2022

Abstract

Multi-source remote sensing images have the characteristics of large differences in texture and gray level. Mismatch and low recognition accuracy are easy to occur in the process of identifying targets. Thus, in this paper, the target recognition algorithm of multi-source remote sensing image based on IoT vision is investigated. The infrared sensor and SAR radars are set in the visual perception layer of the iVIOT. The visual perception layer transmits the collected remote sensing image information to the application layer through the wireless networks. The data processing module in the application layer uses the normalized central moment idea to extract the features of multi-source remote sensing image. Contourlet two-level decomposition is performed on the image after feature extraction to realize multi-scale and multi-directional feature fusion. A two-step method of primary fineness is used to match the fused features and the random sampling consensus algorithm is used to eliminate false matches for obtaining the correct match pairs. After the image feature matching is completed, the BVM target detection operator is used to complete the target recognition of multi-source remote sensing image. Experimental results show that the use of the IoT to visually recognizing the desired remote sensing image target has low communication overhead, and the recognition reaches 99% accuracy.

Keywords IoT vision · Multi-source remote sensing image · Target recognition algorithm · Infrared sensor · SAR radar · Image registration

1 Introduction

Remote sensing refers to comprehensive information science technology that realizes detection and analysis at a long distance. Through sensors, we do not directly touch the surface of an object. Remote sensing technology is an important data source for military information and intelligence acquisition [1, 2]. With the continuous enrichment of remote sensing image data resources [3], the efficiency of remote sensing image processing has become an important factor restricting the development of remote sensing technology. With the continuous improvement of aerospace technologies such as

data communication and sensors, the acquisition of data in the field of remote sensing has gradually developed towards multi-source [4]. Therefore, to comprehensively utilize the information of heterogeneous remote sensing images, image registration is the first condition [5].

The Internet of things is the Internet that connects "things", and its essence is mainly reflected in three aspects: Internet communication characteristics, that is, people or things are connected to the Internet to achieve interconnection [6]; identification characteristics, that is, it can automatically identify people or things connected to the network; Intelligent features, that is, the network system can self-feedback and intelligent control [7].

In recent years, with the continuous development of sensor technology, the resolution of remote sensing images has been continuously improved, and the application of multi-source remote sensing information for the target level has received more and more attention. Melonakos et al. [8] proposed an image segmentation technique based on directional information augmented conformal active contour frame. Wei et al. [9] proposed a ship detection method based on the

✉ Jerry Chun-Wei Lin
jerrylin@ieee.org
Xue-jun Sun
fffx@lyu.edu.cn

¹ Feixian Campus of Linyi University, Linyi, Shandong Province 273400, China

² Western Norway University of Applied Sciences, Bergen, Norway

local power spectrum of SAR images. The core idea of this method is to detect ships through the power spectrum distortion of the local area of the SAR image. Although the energy of ships is distributed on the SAR intensity image, their spectral energy was quite concentrated, which may cause the power spectrum of the local area of the SAR image to deviate from the sea background, and analyze the local power spectrum of the moving target on the SAR image. The method of obtaining the detection threshold through the probability density function of the power spectrum is presented. An et al. [10] used land masking strategy, appropriate sea clutter model and neural network as a recognition scheme to detect ships in SAR images, and used a fully convolutional network to separate the ocean from the land. By analyzing the distribution of sea clutter in the SAR image, based on the comprehensive consideration of sea clutter modeling accuracy and computational complexity, the probability distribution model of the constant false alarm rate detector was selected from the three aspects of K distribution, gamma distribution and Rayleigh distribution, and the neural network was used to re-check the result as the recognition result. Yue et al. [11] studied the detection of special targets in the visual image of low-pixel surveillance systems. This method was used to detect special targets contained in the low-pixel surveillance system; Fan [12] studied the multi-target recognition of fuzzy remote sensing images based on Euclidean feature matching. The algorithm applied the European feature matching method to the multi-target recognition of remote sensing images. It has high recognition accuracy, but it cannot adapt to the complex and changeable detection environment. Therefore, in this paper, the target recognition algorithm of multi-source remote sensing image based on iVIOT is researched. Infrared sensors and SAR radar sensors are applied to the iVIOT to realize the accurate target recognition of multi-source remote sensing image. The experimental verification of the research method has high remote sensing image target recognition effectiveness and high applicability, improves the recognition accuracy, and reduces the time consuming, which plays an important role in the development of multi-source remote sensing image target recognition.

2 Multi-source remote sensing image target recognition algorithm based on IoT vision

2.1 The logical design of the iVIOT

The intelligent vision IoT consists of four levels: visual perception layer, network layer and application layer, and storage layer. The structure model of the iVIOT is shown in Fig. 1.

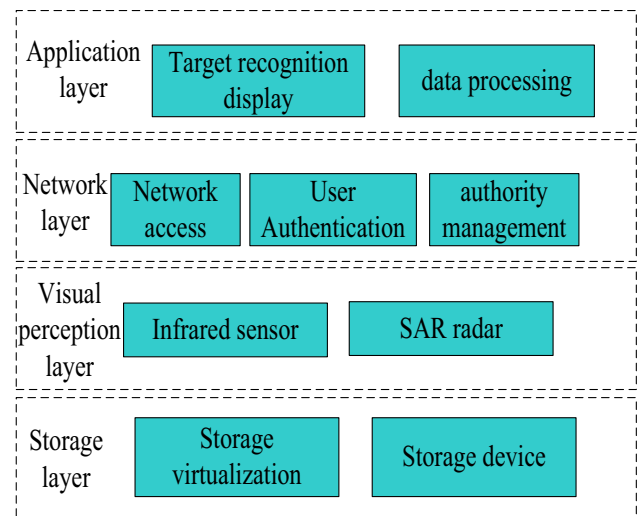


Fig. 1 The structure model of the iVIOT

The iVIOT uses wireless transmission and cloud storage technology to set infrared sensors and SAR radars in the visual perception layer of the iVIOT. The image information collected by the visual perception layer is transmitted to the application layer through the wireless network, and the application layer realizes accurate recognition of multi-source remote sensing image targets through three parts: feature extraction, image registration and target recognition of the data processing module [13]. During data processing, the data is synchronized to the network hard disk and the cloud storage of the iVIOT is realized. Users can access the stored data anytime and anywhere through various intelligent terminals to prevent the leakage and loss of target recognition data of remote sensing image, and improve the security and convenience of the iVIOT.

The visual perception layer mainly solves the data collection problem of the external physical world through various sensor devices. The perception layer of iVIOT acquires information captured by sensors through various images. Wi-Fi wireless network camera IPCAM is used as the information collection point, it is a device that transmits dynamic remote sensing images through the network [14], which can transmit local dynamic remote sensing images to the Internet via Wi-Fi, and which is convenient for users to view at any time.

IPCAM is a camera that collects and transmits dynamic video through a wireless network. It is designed based on user-friendly ideas. It is a new generation of video recording products that combine traditional cameras and network video technologies. It integrates video servers and cameras, wireless transmission and other technologies. It has built-in server and GUI, supports IE browsing mode, and can transmit video images based on TCP/IP protocol. Users can easily install at home, office, factory and any other places, access, configure, maintain and supervise through client video management software or log in to Web pages; where there is

network coverage in the world, users can view the monitored target environment at any time through the local area network or the Internet. Without being restricted by time and space, users can view the dynamic remote sensing images of the target environment captured by infrared sensors and SAR radar anytime and anywhere.

2.2 Feature extraction of multi-source remote sensing image

Whether it is an optical remote sensing image or a SAR remote sensing image, the image is often deformed due to external factors, which will cause certain external interference to the target. When constructing multi-source and multi-feature vectors for target detection, some features that are not affected by noise, light, shadow, deformation, etc. should be added [15]. By extracting seven Hu invariant moments and three affine invariant moments, a feature vector of moment invariants is constructed for target detection in remote sensing images. Discretization defines the digital image as follows:

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q I'(x, y) \quad (1)$$

$$\mu_{pq} = \sum_{x=1}^M \sum_{y=1}^N (x - \bar{x})^p (y - \bar{y})^q I'(x, y) \quad (2)$$

where, $p, q = 0, 1, 2, \dots$. The moment invariant feature m_{pq} will change with the change of the image, and the affine coefficient μ_{pq} will change with the rotation of the image.

In order to prevent the moment feature from changing with the image, the idea of normalized central moment is introduced.

$$y_{pq} = \frac{\mu_{pq}}{\mu_{00}^r} \quad (3)$$

where, $r = \frac{p+q+2}{2}$, $p + q = 2, 3, \dots$.

Seven Hu invariant moments are constructed by using the second-order normalized central moment and the third-order normalized central moment, which are invariant to translation, scaling and rotation. The specific definition is as follows:

$$I_1 = y_{20} + y_{02} \quad (4)$$

$$I_2 = (y_{20} + y_{02})^2 + 4y_{11}^2 \quad (5)$$

$$I_3 = (y_{30} + 3y_{12})^2 + (y_{03} + 3y_{21})^2 \quad (6)$$

$$I_4 = (y_{30} + y_{12})^2 + (y_{21} + y_{30})^2 \quad (7)$$

$$I_5 = (y_{30} - y_{12})(y_{30} + y_{12}) \left[(y_{30} + y_{12})^2 - 3(y_{21} + y_{03})^2 \right] + (3y_{31} - y_{03})(y_{31} + y_{03}) \left[3(y_{30} + y_{12})^2 - (y_{21} + y_{03})^2 \right] \quad (8)$$

$$I_6 = (y_{20} - y_{02}) \left[(y_{30} + y_{12})^2 - (y_{21} + y_{03})^2 \right] + 4y_{11}(y_{30} + y_{12})(y_{21} + y_{03}) \quad (9)$$

Fig. 2 Feature extraction part of the code

```
Physical training intensity test
#coding:utf-8
import sys,os
from PIL import Image,ImageDraw

def getPixel(image, x, y, G, N):
    L = image.getpixel((x, y))
    if L > G:
        L = True
    else:
        L = False

    nearDots = 0
    if L == (image.getpixel((x - 1, y - 1)) > G):
        nearDots += 1
    if L == (image.getpixel((x - 1, y)) > G):
        nearDots += 1
    if L == (image.getpixel((x - 1, y + 1)) > G):
        nearDots += 1
    if L == (image.getpixel((x, y - 1)) > G):
        nearDots += 1
    if L == (image.getpixel((x, y + 1)) > G):
        nearDots += 1
    if L == (image.getpixel((x + 1, y - 1)) > G):
        nearDots += 1
    if L == (image.getpixel((x + 1, y)) > G):
        nearDots += 1
    if L == (image.getpixel((x + 1, y + 1)) > G):
        nearDots += 1

    if nearDots < N:
        return image.getpixel((x, y-1))
    else:
        return None
```

$$I_7 = (3y_{21} + y_{03})(y_{03} + y_{12}) \left[(y_{30} + y_{12})^2 - 3(y_{21} + y_{03})^2 \right] + (y_{30} - 3y_{12})(y_{21} + y_{30}) \left[3(y_{30} + y_{12})^2 - (y_{21} + y_{03})^2 \right] \tag{10}$$

When the image is distorted due to different shooting angles, the translation, scale and rotation invariance of Hu moment invariants can not meet the actual requirements, and a moment invariance under the condition of affine transformation of the target is needed to deal with distortion and other deformations [16].

Three affine invariant moments are defined as follows:

$$A_1 = \frac{(\mu_{20}\mu_{02} - \mu_{11}^2)}{\mu_{00}^4} \tag{11}$$

$$A_2 = \frac{(\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{03}\mu_{21}\mu_{12} + 4\mu_{03}\mu_{12}^3)}{\mu_{00}^{10}} \tag{12}$$

$$A_3 = \frac{(\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{12}\mu_{30} - \mu_{21}^2))}{\mu_{00}^7} \tag{13}$$

Through the above process, the feature extraction of infrared remote sensing image and SAR remote sensing image is realized, the code is as follows (Fig. 2).

2.3 Design of automatic registration method for multi-source remote sensing image features

2.3.1 Multi-scale and multi-directional feature fusion

Contourlet two-level decomposition is performed on the image after feature extraction, to obtain multi-scale σ_1, σ_2 low-frequency sub-bands and d_1 - d_{12} -multi-directional high-frequency sub-bands. The Gaussian kernel function is added to the moment definition, and σ is the scale factor.

The discrete features of order $p + q$ of the image are defined as follows:

$$m_{pq} = \frac{4}{(k-1)^2} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \left(\frac{x}{\sigma}\right)^p \left(\frac{y}{\sigma}\right)^q \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) f(x, y) \tag{14}$$

The formula of discrete feature center distance is as follows:

$$\mu_{pq} = \frac{4}{(k-1)^2} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \left(\frac{x-\bar{x}}{\sigma}\right)^p \left(\frac{y-\bar{y}}{\sigma}\right)^q \exp\left(-\frac{(x-\bar{x})^2 + (y-\bar{y})^2}{2\sigma^2}\right) f(x, y) \tag{15}$$

where, (\bar{x}, \bar{y}) is the center coordinate of the $(n \times n)$ window.

the moment eigenvector of the secondary low-frequency subband is $f_L = [\zeta_1^1, \zeta_2^1, \zeta_3^1, \zeta_1^2, \zeta_2^2, \zeta_3^2]$, ζ_1, ζ_2 and ζ_3 are used to represent the Gaussian combined invariant moments. For multi-directional high-frequency sub-bands, the four parameters of energy f_{ene} , contrast f_{con} , correlation f_{cor} and entropy f_{ent} of the structural texture feature are extracted, and $T(i, j)$ is the gray-level co-occurrence matrix. The gray-level co-occurrence moment feature vector of the high-frequency sub-band is $f_H = [f_{ene}, f_{con}, f_{cor}, f_{ent}]$, the weighting coefficients of the four parameters of the high-frequency sub-band are calculated according to the contrast sensitivity function of the spatial activity degree [17], and the weighted high-frequency sub-band gray-level co-occurrence can be obtained. The moment feature vector is $f'_H = [f'_{ene}, f'_{con}, f'_{cor}, f'_{ent}]$.

2.3.2 Primary-fitness feature registration

A two-step method of primary-fitness is used to achieve feature matching. Firstly, the 6-dimensional moment feature vector $f_L = [\zeta_1^1, \zeta_2^1, \zeta_3^1, \zeta_1^2, \zeta_2^2, \zeta_3^2]$ of low-frequency subband is used for the initial matching of the similarity measure, and on this basis, the weighted high-frequency subband $f'_H = [f'_{ene}, f'_{con}, f'_{cor}, f'_{ent}]$ is used for the second fine matching. The matching formula is as follows:

$$S_{ij} = \exp(-|f_{L1}(i) - f_{L2}(j)|) \tag{16}$$

$$S_{ij} = \exp(-|f'_{H1}(i) - f'_{H2}(j)|) \tag{17}$$

2.3.3 Random sampling consensus algorithm

Random sampling consensus algorithm is used to eliminate wrong matches as the final correct matching pair. The random sampling consensus algorithm divides the feature points into correct matching point pairs and incorrect matching point pairs [18]. The random sampling consensus algorithm calculates the coordinate conversion relationship between the feature points of the reference image and the corresponding feature points of the image to be matched, that is, the transformation matrix H . Four pairs of matching points are randomly selected from the initial matching point pairs, and the transformation matrix H is calculated. The HX_i value of the reference image point $X_i(x, y)$ in the remaining matching pairs is calculated, and the d_i between this value and the matched point $X'_i(x', y')$ in the image to be matched is calculated. If d_i is less than the preset threshold T , the feature point is regarded as a correct match,

otherwise it is regarded as an error point. Four pairs of matching pairs are re-selected randomly, to repeat the above steps, and the corresponding calculated transformation matrix H of the most correct matching pairs is used as the final transformation matrix to eliminate the wrong matching point pairs.

2.4 Design of target recognition algorithm of multi-source remote sensing image

BVM target detection operator is used to identify multi-source remote sensing image targets. The result image obtained by the target detection operator should be able to distinguish the background and target information significantly, that is, the probability of the target or abnormal situation is large, so that the image information tends to be highly certain, and the target information is prominent and easy to distinguish [19]. According to Shannon's definition of information, information is a description of the uncertainty of the movement state or the way of existence of things, so the self-information of the image can reflect the uncertainty of the information, and the self-information I_i is small, indicating that the resulting image is less uncertain, that is, the detected target is prominent. The rest of the background is suppressed.

L operators are used to detect the target in n bands of hyper-spectral image, to obtain L target detection images. Assuming that the total variance is constant, the normalized variance coefficient of each image is calculated [20]. σ_i^2 is used to represent a monotonically increasing function, that is, the variance σ_i^2 is the minimum, ρ_i is the maximum, and I_i is the minimum. The target can be detected based on the smallest variance of the detection result image.

If $\{r_1, r_2, \dots, r_N\}$ is the pixel vector in the remote sensing image, N is the total number of pixels in the image, and each pixel $r_i = [r_{i1}, r_{i2}, \dots, r_{iL}]^T$ is the L -dimensional column vector, where L is the number of bands and $1 \leq i \leq N$.

Assuming that the priori information d is the spectrum signal of the target to be detected, and the vector w passed through the target detection operator corresponds to the output of the input pixel r_i as y_i , namely:

$$y_i = \sum_{l=1}^L w_l r_{ij} = w^T r_i = r_i^T w \quad (18)$$

The original image covariance matrix is expressed as Σ , and the resulting image variance is:

$$v(w^T \circ r_i) = w^T \Sigma w \quad (19)$$

The filter vector w of the BVM operator needs to meet the following conditions:

$$\begin{cases} \min(w^T \Sigma w) \\ d^T w = 1 \end{cases} \quad (20)$$

The auxiliary function $\varphi(w)$ is constructed as (λ is called the Lagrangian multiplier):

$$\varphi(w) = w^T \Sigma w + \lambda(d^T w - 1) \quad (21)$$

Lagrangian multiplier method is used to make the partial derivative of w as zero, and obtain the solution formula as follows:

$$w = -\frac{1}{2} \lambda \Sigma^{-1} d \quad (22)$$

At this time, the covariance matrix is a symmetric matrix, and the optimal solution is obtained as follows:

$$w^{BVM} = \frac{\Sigma^{-1} d}{d^T \Sigma^{-1} d} \quad (23)$$

The eigenvector product of BVM constraint operator and the known spectral is 1. The BVM operator is based on the smallest variance. After data processing is performed using the covariance matrix, the small target and the background are easier to separate, and accurate multi-source image target recognition results can be obtained.

3 Experimental results and analysis

In order to verify the effectiveness and practicability of the proposed method, a port is taken as the research object and experimental comparative analysis is carried out through various methods [21]. Firstly, a port target recognition based on Internet of things visual recognition remote sensing image infrared image and SAR image is studied, and infrared sensor and SAR iVIOT are selected to be used. Radar is used as the sensor of the visual perception layer to collect remote sensing image information, and the data processing module of the iVIOT is used to realize the processing of image target recognition. Figure 3 shows the results of infrared remote sensing images collected by the iVIOT.

The results of SAR remote sensing images collected by the iVIOT are shown in Fig. 4.

Using the method of this paper, the method of literature [11] and the method of literature [12], the feature points contained in the remote sensing images in Fig. 3 and Fig. 4 are extracted, and the comparative analysis results are shown in Fig. 5.

According to Fig. 5, compared with the literature method, the method in this paper can extract more feature points in a very short time. When iterating about 70 times, basically all the feature points are proposed, and the number of feature points extracted is far Much higher than the literature method.



Fig. 3 Infrared remote sensing image results

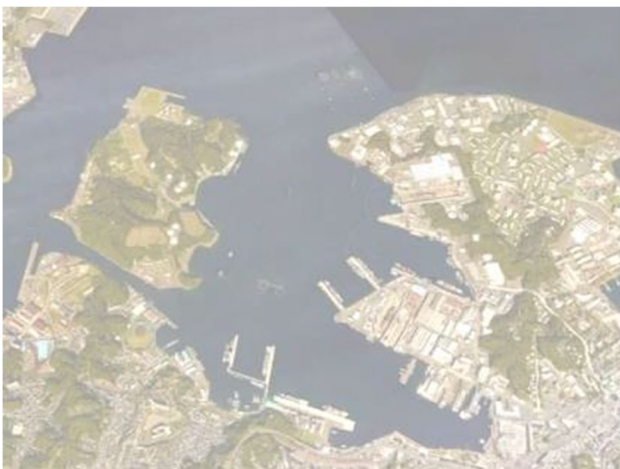


Fig. 4 SAR remote sensing image results

The collected infrared remote sensing images and SAR remote sensing images are used to implement image recognition. The image result after the recognition is shown in Fig. 6.

It can be seen from the experimental results in Fig. 6 that the method in this paper can be used to achieve high-efficiency recognition of infrared remote sensing images and SAR remote sensing images, and the use of recognition remote sensing images can improve the accuracy of target detection.

The results of using the method in this paper to identify ship targets in multi-source remote sensing images based on iVIOT are shown in Fig. 7.

The experimental target recognition results in Fig. 7 show that the method used in this paper can achieve effective recognition of targets in remote sensing images. The method

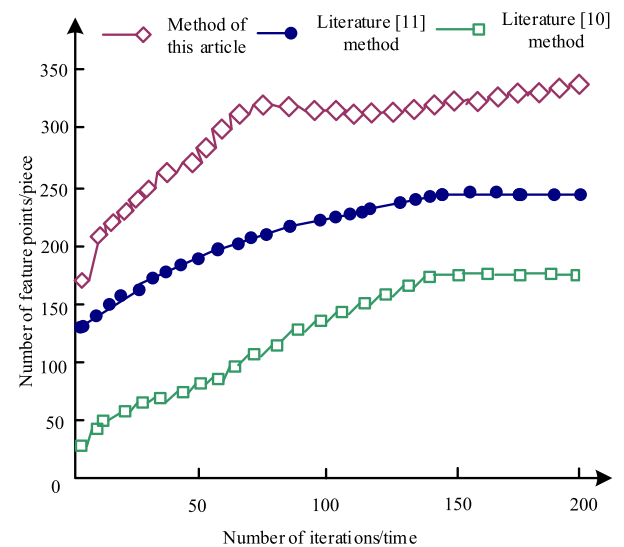


Fig. 5 Feature point extraction analysis comparison



Fig. 6 The recognition result of the method in this paper

in this paper uses iVIOT to effectively process the target recognition information of remote sensing image and efficiently obtain target recognition results of remote sensing image. The target in the remote sensing image is small, the method in this paper can accurately identify the small target in the remote sensing image, and verify the effectiveness of the target recognition of the research method.

The communication overhead of the method in this paper using the iVIOT to identify multi-source remote sensing image targets is made statistics. The iVIOT uses cloud computing technology to realize data transmission, processing and storage. In the different remote sensing image sizes, different environmental conditions, the comparison result of

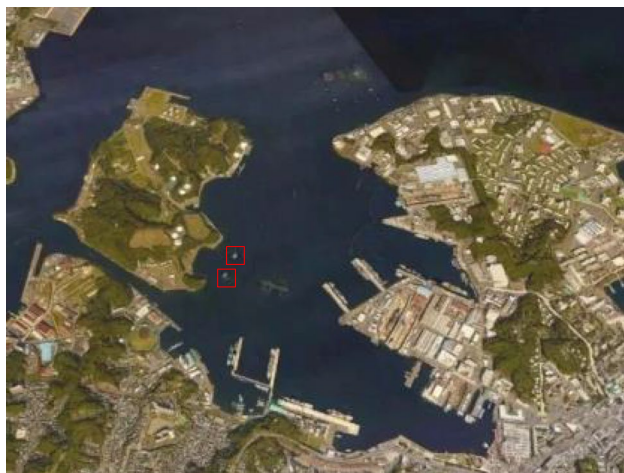


Fig. 7 Target recognition results

the communication overhead in target recognition of multi-source remote sensing image is shown in Fig. 8.

It can be seen from Fig. 8 that when the detection environment does not change, that is, when the target recognition environment is a normal environment, the iVIOT is used to realize the target recognition of remote sensing image, which has lower computational complexity and storage complexity. When the target recognition of multi-source remote sensing image is performed in a complex environment, because the need to verify and store data involved in operations such as updating data, modifying data, and deleting data, high computational complexity is required, and when the size of the remote sensing image is 128 MB, the communication overhead is small in different environments. When the size of the

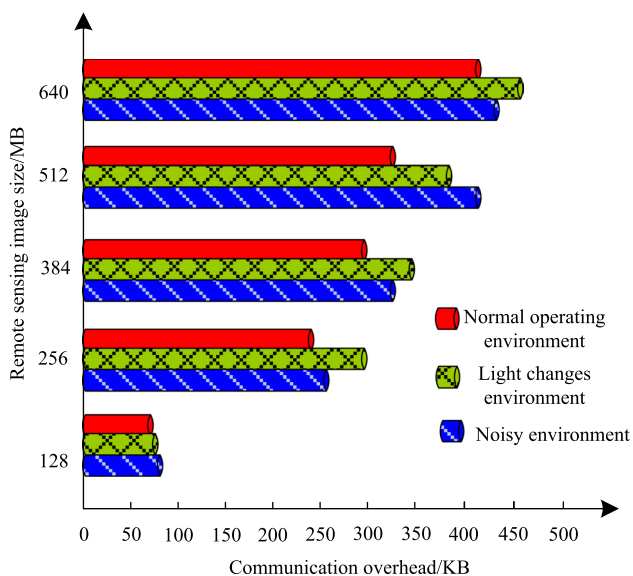


Fig. 8 Comparison of communication overhead

remote sensing image exceeds 128 MB, the communication overhead increases rapidly, showing a linear growth trend. The iVIOT adopts cloud computing technology to protect user privacy information, has good computing performance, and is of great significance for improving remote sensing image target recognition.

Suppose the imaging interval of multi-source remote sensing images is 0.32 h, and the results of ship motion state obtained by the proposed method at each imaging time are shown in Table 1.

From the experimental results in Table 1, it can be seen that the method in this paper can effectively obtain the motion state of the target in the remote sensing image. In order to verify the effectiveness of the method in identifying the remote sensing image target, the method in this paper used to track the position mean square and speed root mean square error of the target ship 1 and the target ship 2 is made statistics. The results are shown in Fig. 9.

The experimental results in Fig. 9 show that the method in this paper makes full use of the target features of remote sensing image recognition, adopts primary-fine feature registration and random sampling consensus algorithm to achieve the association of target data, has strong recognition stability, and can obtain better targets. The recognition result meets the requirements of moving target tracking of multi-source remote sensing image.

The special target detection method (Reference [11]) and the Euclidean feature matching method (Reference [12]) are selected as the comparison methods, and the correct recognition rate of the three methods used to register the remote sensing image under different target recognition

Table 1 Recognition results of ship motion state

Image number	filming time/s	Target coordinates/km		Target heading angle/rad	Driving speed/(km/h)
		x	y		
1	0	3.564	4.481	6.542	30.512
2	0.15	6.945	7.462	5.481	38.456
3	0.34	3.548	4.153	5.162	37.158
4	0.58	4.841	7.845	4.852	36.452
5	0.76	9.582	8.416	4.461	35.464
6	1.25	8.561	6.854	4.158	37.851
7	1.85	9.584	7.482	3.896	41.584
8	2.64	3.518	6.158	3.564	38.465
9	3.85	2.485	5.846	3.284	42.845
10	4.15	1.468	11.584	3.165	39.546
11	4.86	3.645	16.852	2.914	32.546
12	5.34	11.258	17.456	2.854	28.645
13	5.95	16.584	15.645	1.564	26.845
14	6.85	38.645	19.852	1.318	27.415
15	6.76	41.584	20.456	0.945	29.854

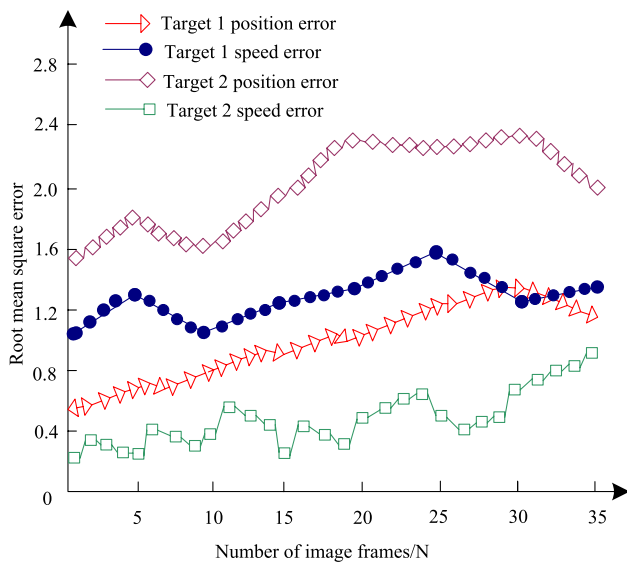


Fig. 9 Root mean square error results of target position and velocity

environments is carried out statistics. The statistical results are shown in Fig. 10.

Three methods are used for the statistics of the time required to register the remote sensing image under normal environment, lighting environment, noise environment, mobile environment and rotating environment. The statistical results are shown in Fig. 11.

Analyzing the experimental results of Fig. 8 and Fig. 11, we can see that using the method in this paper to identify remote sensing image targets can get a higher correct matching rate and higher matching accuracy.

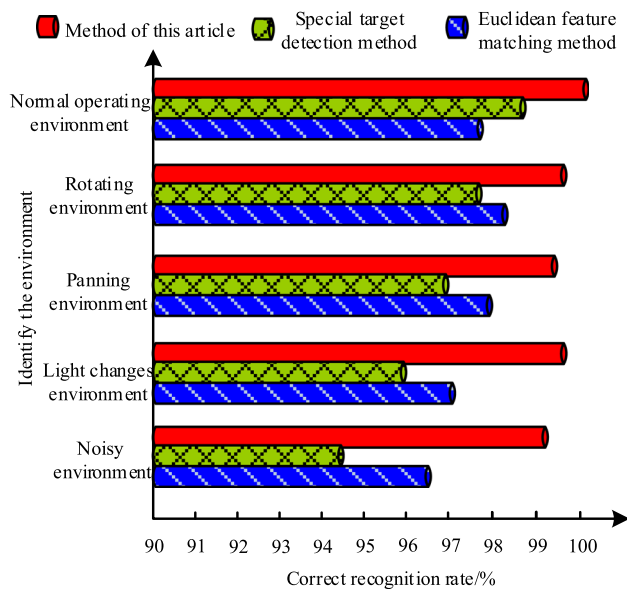


Fig. 10 Comparison results of correct recognition rate

The total recognition time of the method in this paper is significantly higher than that of the special target detection method and the Euclidean feature matching method. Using this method to identify multi-source remote sensing image targets, the correct matching rate in different environments is higher than 99%, the recognition time is less than 10 s, the correct matching rate and the recognition time are significantly better than the other two methods, verifying that the method in this paper has high recognition performance. It can provide a good foundation for target recognition of multi-source remote sensing image.

The method in this paper is used to identify remote sensing image targets, and the statistics are based on the target recognition accuracy and recognition recall rate using only infrared remote sensing images, SAR remote sensing images, and multi-source remote sensing images. The statistical results are shown in Fig. 12 and Fig. 13.

Analyzing the detection results of Figs. 12 and 13, it can be seen that compared with the pure use of optical remote sensing images and SAR radar remote sensing image extraction features for target detection, the target recognition method of multi-source remote sensing image proposed in this paper can greatly improve the accuracy of target detection and reduce the possibility of false targets being detected, on the basis of a certain improvement in the target detection recall rate. The method in this paper applies the iVIOT to target recognition of multi-source remote sensing image, uses efficient computing performance to achieve accurate target detection, and has high applicability.

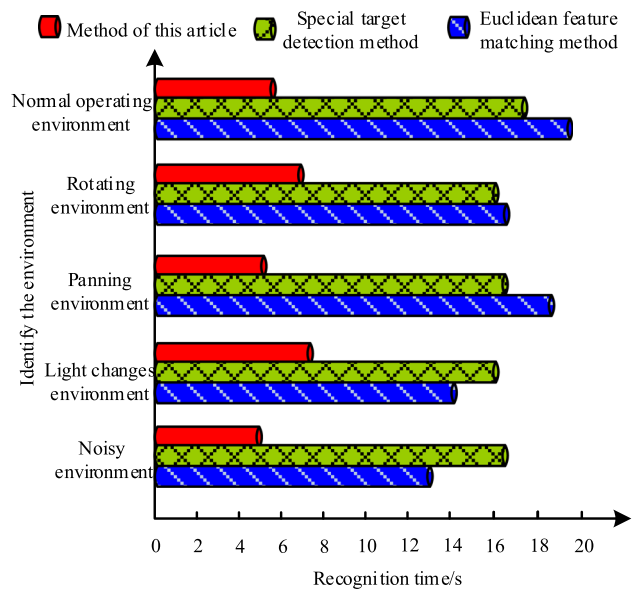


Fig. 11 Comparison of registration time

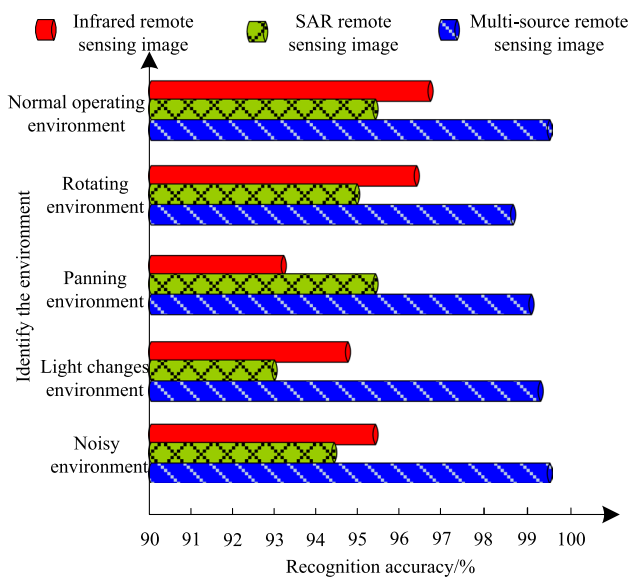


Fig. 12 Comparison of recognition accuracy

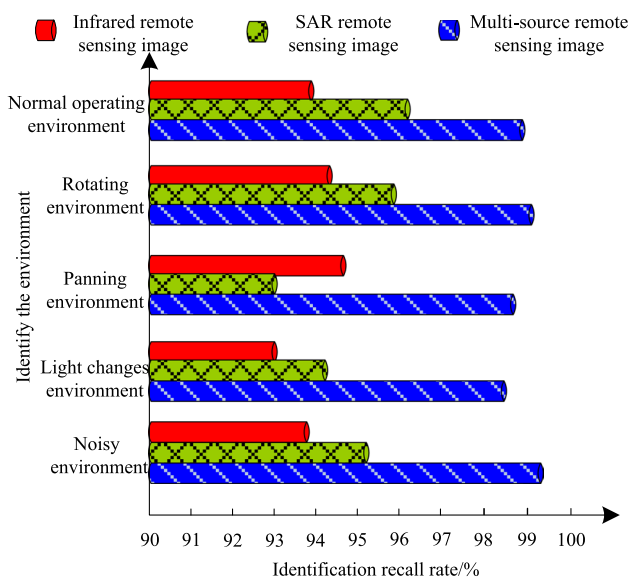


Fig. 13 Comparison of recognition recall rates

4 Discussion

The application range of remote sensing images is extremely wide, and the accurate targets detection of remote sensing image can enhance its application range. Research on target detection based on optical and SAR remote sensing images is carried out, and the high computing performance of the iVIOT is used to improve the accuracy of remote sensing image target detection. Because of their different imaging principles, the two have their own advantages in earth

observation. SAR sensors have all-weather and all-weather detection capabilities, can penetrate clouds, fog and are not affected by shadow occlusion and light time, but their texture and ground object radiation information are not enough, and it is difficult to interpret. Optical remote sensing images can intuitively reflect information on texture, color, and shape to users, but due to the limitations of light and weather, the ability to acquire data is limited. Optical remote sensing images can extract rich spectral information in radiation characteristics, which is more beneficial for classification and interpretation. Different types of remote sensing data such as optics and SAR are increasing at a rate of thousands of GB every day, which provides a rich source of data for multi-source processing of remote sensing images. How to achieve the interpretation of the specified target from the massive high-resolution remote sensing images and fully excavate the multi-source information will become a key link in the application of remote sensing information. Carrying out target interpretation of multi-source remote sensing image fusion is not only of great significance for the development of multi-source remote sensing image fusion and target interpretation processing theory, but also conducive to the full mining of massive remote sensing data and the realization of target-level multi-source information interpretation, to provide target information support in military fields such as military strikes and intelligence analysis, and civilian fields such as urban planning, aviation control, and traffic navigation.

5 Conclusion

Remote sensing images contain a large amount of target information. In order to accurately identify targets in remote sensing images, a target recognition algorithm of multi-source remote sensing image based on IoT vision is studied, and iVIOT structures are used to achieve accurate targets recognition of remote sensing image. After the feature extraction from the multi-source remote sensing image is completed, a robust and stable multi-source remote sensing image registration method is used to achieve accurate remote sensing image registration. The feature registration method combining elementary-precision is used to realize the registration of multi-source remote sensing images affected by noise and illumination changes. Experimental results show that this method is robust to noise and illumination changes, and has a good application prospect for tracking moving targets in remote sensing images. iVIOT is usually used in occasions with high real-time requirements. The target recognition of remote sensing image usually includes large-scale matrix decomposition, repeated convolution, solving large-scale equations, and many non-linear optimization problems, which are computationally intensive

and time-consuming. IVIOT can effectively solve the above problems, improve the computing performance of remote sensing image target detection, and has extremely high applicability. Due to the limitation of time, the research on image blur in this paper is not in-depth enough. In the future, further research will be made on image blur processing, so as to improve target recognition in multi-source remote sensing images.

Author contributions Xuejun Sun coded the algorithm and wrote the paper. Jerry Chun-wei Lin provide the idea, discussed the result, revised the paper.

Funding Open access funding provided by Western Norway University Of Applied Sciences.

Data Availability Not applicable.

Declarations

Ethics Approval Not applicable.

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Cexus JC, Toumi A, Riahi M (2020) Target recognition from ISAR image using polar mapping and shape matrix// 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)(4):250–259.
2. Liu S, Liu D, Srivastava G et al (2021) Overview and methods of correlation filter algorithms in object tracking. *Complex Intell Syst* 7:1895–1917
3. Tu Y, Lin Y, Wang J et al (2018) Semi-Supervised Learning with Generative Adversarial Networks on Digital Signal Modulation Classification[J]. *CMC-Comput Mater Continua* 55(2):243–254
4. Roh KA, Jung JY, Song SC (2019) Target Recognition Algorithm Based on a Scanned Image on a Millimeter-Wave(Ka-Band) Multi-Mode Seeker. *J Korea Inst Electromagn Eng Sci* 30(2):177–180
5. Baby RM (2018) New Automatic Target Recognition Approach based on Hough Transform and Mutual Information. *Int J Image Graph Signal Process* 10(3):18–24
6. Shuai W, Xinyu L, Shuai L et al (2021) Human Short-Long Term Cognitive Memory Mechanism for Visual Monitoring in IoT-Assisted Smart Cities. *IEEE Internet Things J*. online first. <https://doi.org/10.1109/JIOT.2021.3077600>
7. Hsia CH, Yen SC, Jang JH (2019) An intelligent iot-based vision system for nighttime vehicle detection and energy saving. *Sens Mater* 31(6):1803–1814
8. Melonakos J, Pichon E, Angenent S (2008) Finsler Active Contours. *IEEE Trans Pattern Anal Mach Intell* 30(3):412–423
9. Wei X, Wang X, Chong J (2018) Local region power spectrum-based unfocused ship detection method in synthetic aperture radar images. *J Appl Remote Sens* 12(1):1–5
10. An Q, Pan Z, You H (2018) Ship Detection in Gaofen-3 SAR Images Based on Sea Clutter Distribution Analysis and Deep Convolutional Neural Network. *Sensors* 18(2):334–337
11. Yue P, Zhao LP, Zhang W (2018) Visual Image Special Target Detection Simulation of Low Pixel Monitoring System. *Comput Simul* 035(007):452–455
12. Fan H (2018) Multi-object recognition algorithm based on euclidean feature match in fuzzy remote sensing images. *IOP Conf Ser Mater Sci Eng* 466(1):012100–012108
13. Li X, Monga V, Mahalanobis A (2020) Multiview automatic target recognition for infrared imagery using collaborative sparse priors. *IEEE Trans Geosci Remote Sens* 58(10):6776–6790
14. Bolourchi P, Moradi M, Demirel H, Uysal S (2020) Improved SAR target recognition by selecting moment methods based on Fisher score. *SIViP* 14(1):39–47
15. Zhou B, Duan X, Ye D, Wei W, Damaševičius R (2019) Multi-level features extraction for discontinuous target tracking in remote sensing image monitoring. *Sensors* 19(22):4855–4861
16. Kumudham R, Rajendran V (2018) Classification performance assessment in side scan sonar image while underwater target object recognition using random forest classifier and support vector machine. *Int J Eng Technol* 7(2):21–30
17. Nasrabadi NM (2019) DeepTarget: An Automatic Target Recognition using Deep Convolutional Neural Networks. *IEEE Trans Aerosp Electron Syst* 55(6):2687–2697
18. Liu S, Wang S, Liu X et al (2021) Fuzzy Detection aided Real-time and Robust Visual Tracking under Complex Environments. *IEEE Trans Fuzzy Syst* 29(1):90–102
19. Bina K, Lina J, Tonga X, Zhang X, Luoa S (2021) Moving Target Recognition With Seismic Sensing: A Review. *Measurement* 181(7802):109584–109584
20. Gokaraju JSAV, Song WK, Ka MH, Kaitwanidvilai S (2021) Human and bird detection and classification based on doppler radar spectrograms and vision images using convolutional neural networks. *Int J Adv Rob Syst* 18(3):172988142110105
21. Shuai L, Chunli G, Fadi A, et al (2020) Reliability of Response Region: A Novel Mechanism in Visual Tracking by Edge Computing for IIoT Environments. *Mech Syst Signal Process* 138:106537

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.