ISSN: 2664-9640

Infrared and optical image registration algorithm based on stable edge characteristics

Hai Xia*, Shiwei Chen, Wenxin Xia

Rocket Force University of Engineering, Xi' an, China

*Corresponding author: xiahai12121 @163.com

Abstract

According to the matching difficulties of multi-sensor remote sensing images, a new infrared and optical image registration approach based on style transfer invariable features is proposed. First, the style transfer network of infrared image to optical image is trained. Then, the artificial optical image and the difference image are generated based on the image style transfer network, and the edge regions of the artificial optical image and the difference image are enhanced by using the multi-scale characteristics of wavelet. After binary segmentation, the stable edge features of the artificial optical image. And the edge features of the optical reference image are also extracted. Finally, the edge features are matched by the cross-correlation criterion to achieve the accurate registration between the original infrared image and the optical reference image. The experiments demonstrate that the artificial optical image can be accurately registered with the optical reference image even under the condition of insufficient training samples. This enhances the adaptability of the algorithm.

Keywords

Multi-sensor images matching; Image style transfer; Difference image; Stable edge features.

1. Introduction

Satellite remote sensing is an effective means for human to observe the earth's environment. With the development of satellite remote sensing technology, heterogeneous image registration based on multi-sensor imaging have attracted extensive interest [1]. In practical engineering applications, due to the specific imaging principle of infrared image, It is not affected by light and has strong contrast. On the contrary, because the optical sensor measures the solar radiation reflected by objects on the ground, the interpretation of the image is easier, but it is greatly affected by light, cloud, season, shadow and other conditions [2]. Through the above analysis and comparison, it can be found that infrared image and optical image can complement each other in some aspects. Therefore, effective registration between two kinds of images is of great significance for multi-information fusion. The common methods of infrared and optical image matching can be divided into intensity based and invariant feature based methods. Strength based methods usually use similarity measures, such as normalized cross-correlation (NCC) [3], mutual information or cross cumulative residual entropy [4]. On the other hand, features such as points, lines, contours or regions are widely used in matching methods based on invariant features [5] [6]. Common invariant features include scale invariant feature transform (SIFT) [7], maximum stable extreme regions (MSER) [8], etc. However, different imaging mechanisms make great style differences between infrared and optical images, which makes it difficult to extract invariant features. If styles between different images can be transferred successfully, it will transform the different image matching into the same image matching, which will greatly reduce the difficulty of matching. With the rise of deep

ISSN: 2664-9640

learning, this method has attracted more and more attention. N Merkle et al. [9] trained an image style migration network to transfer the image style of optical satellite to generate high fidelity artificial infrared image, and then used common matching methods (NCC, SIFT, etc.) to register the artificial infrared image and target infrared image, and achieved high accuracy and precision. However, this method needs a large number of training samples to ensure the fidelity of the generated artificial infrared image. If the training sample size is small, the fidelity of the generated artificial infrared image will decline. It is difficult to ensure the accuracy of registration by using common matching methods.

On the basis of reference [9], this paper focuses on the problem of infrared and optical image registration under the condition of insufficient training samples. Although there are great differences between infrared and optical images in vision, there are still some stable edge features. If these stable edge features can be extracted, accurate registration of infrared and optical images can be achieved. Therefore, a infrared and optical image registration algorithm based on image style shift invariant edge features is proposed in the paper. Considering that the inherent speckle noise of infrared image will affect the edge features, so the infrared image style is transferred to optical image mode to reduce the noise interference. Algorithm principle: firstly, the original infrared image is transformed by image style migration network; The image is converted into an artificial optical image; then, the difference image is generated from the artificial optical image and the original infrared image, and the stable edge features are extracted based on wavelet multi-scale edge enhancement and Canny operator; finally, the accurate matching between the artificial optical image and the optical reference image is achieved through the conventional matching algorithm, and the algorithm flow chart is shown in Figure 1. The experimental results show that the proposed method can achieve accurate matching between infrared and optical images based on image style transfer technology and edge invariant features with few training samples.

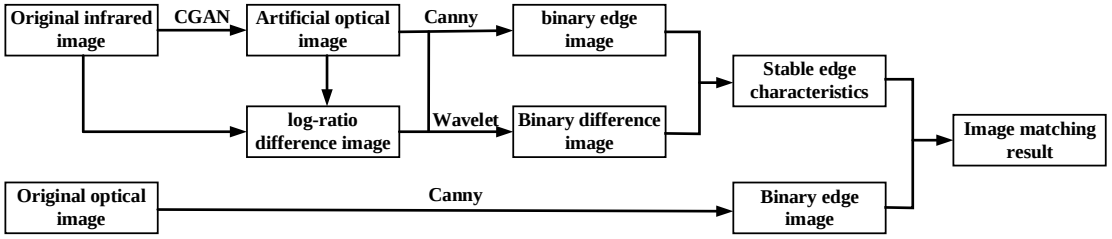


Fig. 1 algorithm flow chart

2. Image style migration network

With the rise of deep learning, Gatys[10] and others creatively proposed an image style transfer method based on convolutional neural network. Compared with the traditional nonparametric image style transfer method, only the low-level features of the image can be extracted, and the image content features and style features can be separated and extracted, and these high-level abstract features can be processed independently[11]

In 2014, Ian J. Goodflow et al [12] proposed the Generative Adversarial Networks (GAN). Since then, based on Gan, researchers have proposed some multiple image style migration networks such as Pix2Pix,CycleGan, StarGan et al. [13] ~ [15]. The images generated can achieve the effect of confusing the real with the fake in face replacement, image restoration, image transformation and other application scenarios.

2.1. Generative Adversarial Networks

Generative confrontation network is a deep learning model, which is one of the most promising unsupervised learning methods on complex distribution in recent years. From the perspective of image generation, Gan is a generation model, whose goal is to train a generator G(Generation

ISSN: 2664-9640

Network) to map random noise z to output image G(z). The training is realized through a confrontational process, and a discriminator D (Discrimination Network) is trained at the same time. The task of D is to distinguish the real image y and the image G(z) generated by G as much as possible, however, G tries to "cheat" D by generating more real G(z) as much as possible, so as to achieve a balance after repeated games. The loss function can be expressed by formula (1).

$$\min_G \max_D \Gamma_{GAN}(G,D) = E_{y\sim p_{real}(y)}[\log D(y)] + E_{y\sim p_{real}(y),z\sim p_z(z)}[\log(1-D(G(z)))] \tag{1}$$

Among this formula, $\Gamma_{GAN}(G,D)$ represents the difference degree between the generated sample and the real sample, and can use the cross entropy loss of binary classification. E represents the expected value. $p_{real}(y)$ represents the real data distribution of the output image. $p_z(z)$ represents noise distribution. $\max_D \Gamma(G,D)$ means that the parameters of discriminator D are updated by maximizing the cross entropy loss $\Gamma_{GAN}(G,D)$ when the generator is fixed. $\min_G \max_D \Gamma(G,D)$ means that the generator should minimize the cross entropy loss while the discriminator maximizes the cross entropy loss $\Gamma_{GAN}(G,D)$.

Conditional Generative Adversarial Nets(CGAN) controls the categories of Gan generated data by adding restrictions on the basis of Gan,. Its principle is: during training, the classification label X which controls the generated category is sent to the input of the generator together with the noise, so that in prediction, the generator will also generate pictures of the specified category according to the input label. Discriminator processing is the same, just in the input with a category label. The loss function of conditional generation antagonism network can be expressed by formula (2).

$$\min_G \max_D \Gamma_{CGAN}(G,D) = E_{y\sim p_{real}(y)}[\log D(y/x)] + E_{y\sim p_{real}(y),z\sim p_z(z)}[\log(1-D(G(z/x)))] \tag{2}$$

2.2. Artificial image generation based on image style migration network

Zhu et al. [13] proposed a classical model, called Pix2Pix model, which applies CGAN to supervised image style transfer. Referring to Pix2Pix model, the control condition of input generator is changed from "classification label" to infrared image. The purpose is to generate artificial optical image with infrared image geometric characteristics and optical image radiation characteristics. In the same way, the control condition of the input discriminator should be changed from "classification label" to infrared image, and as a "condition", it should be spliced with the real optical image or the generated artificial optical image and sent to the discriminator. Therefore, the essence of the heterogeneous image style transfer algorithm in this paper is: the infrared image is input into the training model as a "constraint condition", the model fits the Pixel probability distribution of the optical image in the training sample, and after the model is trained, the infrared image is input and the artificial optical image is output. The loss function of heterogeneous image style migration network is shown in formula (3).

$$G^* = \min_G \max_D \Gamma_{CGAN}(G,D) + \lambda\Gamma_{L1}G \tag{3}$$

The loss function G* is composed of two parts: the first part is CGAN loss, which is the same as ordinary CGAN loss function, see formula (2); the second part is L1 loss, which is calculated by averaging the absolute value of the pixel by pixel difference between the real optical image y and the artificial optical image G(x) generated by the generator, as shown in formula (4).

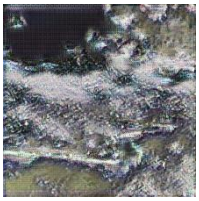
$$\Gamma_{L1}G = E_{x,y}[\|y-G(x)\|_2] \tag{4}$$

Among them, CGAN loss mainly represents the image content features, L1 loss mainly represents the image style features. Combining the two kinds of loss can make the artificial

ISSN: 2664-9640

image have higher fidelity. In addition, when using pix2pix model for image style conversion, a large number of training samples are needed to train the model. If the training samples are insufficient, the fidelity of the generated artificial image will decline.


(a) Original infrared image


(b) A artificial optical image


(c) B artificial optical image


(d) Optical target image


(e) C artificial optical image


(f) C optical target image

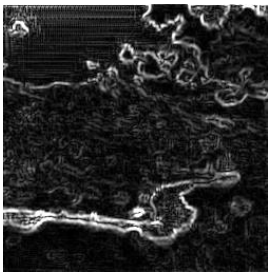
Fig. 2 The influence of training set size on the production of human image

Two trained image style migration network models are used： The generated artificial optical image is shown in Fig. 2(a) and (d). It can be seen that with the increase of the number of training samples, the artificial image is closer to the original image (target image). However, the maximum training set is 300, even if all the training is involved, the artificial image still has a certain visual gap with the target image. Figure 2(e) and (g) shows the image conversion results provided by reference [9] (assuming its training model is c). 69900 pairs of infrared and optical images are used in the training set. It can be seen that the generated artificial optical image is almost the same as the target image in vision. It also fully proves that the size of training set has an important influence on the fidelity of artificial image. In reality, it is difficult to have a large enough training sample set for many matching application scenarios, which restricts the promotion and application of heterogeneous image matching method based on style transfer.

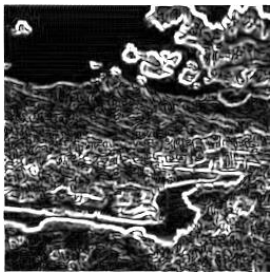
3. Heterogeneous image matching based on edge invariant features

Edge features have better adaptability in remote sensing image matching because of its rich image information. However, for infrared and optical image, the edge features will change greatly, making it difficult to match. From the visual point of view, we can find that there are still some edge invariant features between infrared and optical images. If these edge features can be extracted, it is very useful for heterogenous matching.

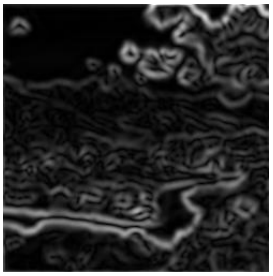

(a) LR difference image


(b) Edge enhanced image

ISSN: 2664-9640



(c) Wavelet small scale edge enhancement for artificial optical image



(d) Wavelet large scale edge enhancement for artificial optical image

Fig. 3 Wavelet multiscale image edge enhancement results

3.1. Image edge enhancement based on multi scale wavelet

The artificial optical image generated by image style transfer network can be regarded as an intermediate state between the original infrared image and the target optical image. The more fully the model training, the closer the target optical image is. However, even if the original infrared image is completely converted into an artificial optical image, some invariant features, especially the edge invariant features, are retained, as shown in Fig. 2. If the two images before and after conversion are compared to generate difference pictures by pixel, the stable edge features in the unchanged region can be highlighted. To suppress noise, the Log-Ratio (LR) operator is used to obtain the difference graph [18], as shown in Fig. 3(a).

The contrast of LR difference image is weak and the vision is dark, so it is very difficult to segment the LR difference image directly. There are many image enhancement algorithms. Considering that LR difference image is still affected by the residual noise of the original infrared image, wavelet multi-scale image edge enhancement algorithm is adopted here [19]. The principle is as follows: firstly, B-spline wavelet is used for multi-scale decomposition of LR difference image; secondly, modulus maximum edge detection method is used to suppress noise in large scale, identify edge, and accurately locate edge in small scale; finally, edge information in different scales is integrated to obtain edge enhancement image of difference image, as shown in Figure 3 (b). The principle of wavelet modulus maximum multi-scale edge detection is as follows：

Let the two-dimensional smoothing function be $g(x,y)$. For smoothing function, we can find partial derivative of X and Y directions respectively, and as basic wavelet, we have：

$$\begin{cases} w_x(x,y) = \dfrac{\partial g(x,y)}{\partial x} \\ w_y(x,y) = \dfrac{\partial g(x,y)}{\partial y} \end{cases} \tag{5}$$

The two functions are horizontal wavelet function and vertical wavelet function at X and Y respectively. Then the scale functions of image $g(x,y)$ in the two directions of wavelet transform with scale time are recorded as follows:

$$\begin{cases} w_s^x(x,y) = \dfrac{1}{s^2} w^x(\dfrac{x}{s},\dfrac{y}{s}) \\ w_s^y(x,y) = \dfrac{1}{s^2} w^y(\dfrac{x}{s},\dfrac{y}{s}) \end{cases} \tag{6}$$

From the above derivation, we can get the expression of wavelet transform of two-dimensional image $g(x,y)$ as follows:

ISSN: 2664-9640

$$\begin{bmatrix} w_s^x f(x,y) \\ w_s^y f(x,y) \end{bmatrix} = s \begin{bmatrix} \frac{\partial}{\partial x}(f \cdot g_s)(x,y) \\ \frac{\partial}{\partial y}(f \cdot g_s)(x,y) \end{bmatrix} = s\nabla(f \cdot g_s)(x,y) \tag{7}$$

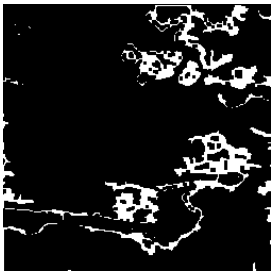
$(f \cdot g_s)(x,y)$ represents the image of $f(x,y)$ smoothed by $g_s(x,y)$. It can be seen from the above formula that the gradient of $(f \cdot g_s)(x,y)$ is proportional to the two components of wavelet transform. Therefore, when the scale is s, the mode and phase angle of the gradient are as follows:

$$\begin{cases} M = \sqrt{\left|w_x f(x,y)\right|^2 + \left|w_y f(x,y)\right|^2} \\ A = \arctan\left[\frac{w_y f(x,y)}{w_x f(x,y)}\right] \end{cases} \tag{8}$$

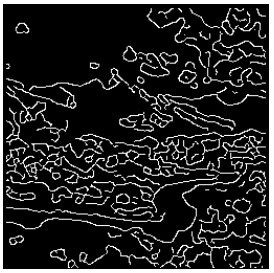
The size of the modulus value reflects the gray change degree of the image on the pixel. The point where the modulus value takes the local maximum along the gradient direction corresponds to the mutation point of the gray level of the image, that is, the edge point of the image. However, the artificial optical image contains the noise left by the original infrared image, and some of the noise also has wavelet modulus maxima, so the detected gray mutation points are not necessarily edge feature points. The edge detected by a single scale contains a lot of noise, and wavelet transform has multi-scale. The modulus maximum of noise signal generally decreases with the increase of scale because of its randomness, while the modulus maximum of edge signal generally increases with the increase of scale. Therefore, in the small scale, it can accurately locate and identify the edge position, as shown in Fig. 3 (c); in the large scale, it can well suppress the noise and identify the edge, but it will lose the detail information, as shown in Fig. 3 (d).

3.2. Edge invariant feature matching

LR difference image is enhanced by wavelet multiscale edge, and the binary difference edge image can be obtained by using the maximum inter class difference method. However, the resolution of binary difference edge image is low, and the distinction between the changed and unchangeable regions is not obvious. Therefore, it is necessary to synthesize binary difference edge image and artificial optical edge image for re segmentation, and then the binary image of unchanged region can be obtained, wherein white area represents the unchanged region, as shown in Fig. 4 (a).



(a) Binary image of unchanged region



(b) Artificial optical image edge

ISSN: 2664-9640


(c) Artificial optical image stable edge (d) Optical reference image edge

Fig. 4 Stable edge feature extraction results

The binary image of unchanged region is the result of edge enhancement and binary region segmentation of artificial optical image, which will lose a lot of detail information and cannot directly extract stable edge features [20]. In this paper, Canny operator is used to extract the edge features of the artificial optical image, and then the stable edge features are separated from the binary image of the unchanged region, as shown in Fig. 4 (b) and (c). It can also be seen from Fig. 4 (b) ~ (d) that most of the stable edge features of the artificial optical image overlap with the edge features of the optical reference image, which lays a good foundation for feature matching.

Because of the strong correlation between the stable edge feature and the edge feature of optical reference image, a conventional normalized correlation index NCC is used as the matching measure, that is, NCC algorithm. The correlation index between the two feature regions can be obtained by formula (9).

$$Ncc = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N} f(x,y)\cdot g(x,y)}{(\sum_{x=1}^{M}\sum_{y=1}^{N} f^2(x,y))^{1/2}\cdot(\sum_{x=1}^{M}\sum_{y=1}^{N} g^2(x,y))^{1/2}} \tag{9}$$

Where $f(x,y)$ represents the pixel value of any point in the stable edge feature image of size $M \times N$, and $g(x,y)$ represents the pixel value of any point in an edge feature area of size $M \times N$ in the optical reference image.

The registration algorithm of infrared and optical image based on style transfer invariant features is as follows:

(1) Training the model of image style transfer network based on training sample set;

(2) Inputting the original infrared image into the style transfer network model to generate the artificial optical image;

(3) Generating LR difference graph is based on original infrared image and artificial optical image;

(4) Enhancing the LR difference image and artificial optical image by wavelet multi-scale;

(5) Using the specific threshold segmentation to obtaining the binary image of the unchanged region;

(6) Using Canny operator to extract the edge features of the artificial optical image, and extracting the edge invariant feature from the results of (5);

(7) Using Canny operator to extract the edge features of optical reference image;

(8) sliding the edge invariant feature template in the edge feature image of optical reference, and calculating the correlation index. The maximum region center of NCC is the location of matching center of heterogenous image.

ISSN: 2664-9640

4. Results

4.1. Experimental preparation

In order to comprehensively measure the performance of the algorithm, three kinds of infrared and optical remote sensing image data sets are selected in the experiment: the first data set is city scene with 300 pairs of training images; the second data set is harbor scene with 100 pairs of training images; the third data set is mountain scene with 50 pairs of training images. Before the matching experiment, the image style transfer network training is carried out, and the experimental environment is as follows: the hardware platform GPU is NVIDIA Tesla P40 24GB, and the memory is 128GB; the specific convolutional neural network training is realized by using pytorch deep learning framework. For the same type of data set, different sizes of data samples are used for training, and different image style migration network models are obtained. The training results are shown in Table 1.

Tab.1 Training results of image style transfer net

No.	Type of sence	Dataset size	Training time（h）	Model name
1	Remote sensing image of city	30	6	City1
2		100	20	City2
3		300	65	City3
4	Remote sensing image of harbor	100	18	Harbor
5	Remote sensing image of mountainous area	50	8	Mountain

4.2. Matching experiment results and analysis

The experimental environment is: Intel Core 2.4G processor, 8G memory, windows 10 operating system, Matlab R2014a computing platform.

Experimental data: three sets of infrared and optical image pairs with different scene types are selected. The first set is urban scene with 50 pairs of images, the second set is harbor scene with 30 pairs of images, and the third set is mountain scene with 20 pairs of images. The size of optical reference image is 800 × 800, the size of infrared image is 512 × 512, and the image format is TIF. Firstly, the corresponding training model is used to generate the artificial optical image for three groups of images in different scenes; then, the conventional matching algorithms (NCC, SIFT) are used to carry out the matching experiment according to the method in reference [9]; finally, the algorithm in this paper is used to carry out the matching experiment, and the experimental results are shown in Table 2. Matching success rate (MSR), matching precision (MP) and matching average precision (MAP) were used to evaluate the matching effect. MP is the distance (in pixels) between the matching center position and the center position marked in advance. If MP is less than 5 pixels, the matching is considered successful; MSR is number of matching successful image pairs divide total number of matching image pairs; MAP is sum of matching successful image pairs and MP divide total number of matching successful image pairs.

Tab.2 Experimental results of image matching

No.	Method (algorithm + model)	Type of sence	MSR	MAP
1	NCC+City1	City	12%	4.68
2	NCC+City2		23.2%	3.24
3	NCC+City3		38%	2.95

ISSN: 2664-9640

4	SIFT+City1		20%	3.23
5	SIFT+City2		42%	2.47
6	SIFT+City3		52%	1.18
7	The algorithm in this paper+City1		30%	2.78
8	The algorithm in this paper+City2		58%	2.34
9	The algorithm in this paper+City3		66%	1.92
10	NCC+Harbor		26.6%	4.53
11	SIFT+Harbor	Harbor	43.3%	3.76
12	The algorithm in this paper+Harbor		78%	2.68
13	NCC+Mountain		5%	4.12
14	SIFT+Mountain	Mountain Area	15%	3.09
15	The algorithm in this paper+Mountain		35%	2.83

The analysis of the experimental results: (1) from the overall matching results, the algorithm in this paper has a higher matching success rate than other algorithms for three groups of scenes; (2) from the previous nine groups of data, we can see that the more training data, the more sufficient the model training, the higher the matching success rate; (3) from the matching results of three types of scenes, we can know that the algorithm in this paper is more suitable for harbor and city scenes, Because these two kinds of scenes will have a lot of stable edge features; (4) from the average matching accuracy results, we can get a conclusion that SIFT algorithm has the highest accuracy, the accuracy of NCC algorithm is the lowest, this algorithm in this paper is in the middle, this is because SIFT algorithm is based on a single pixel for matching, while this algorithm and NCC are based on the statistical information of pixels on the line and region for matching.

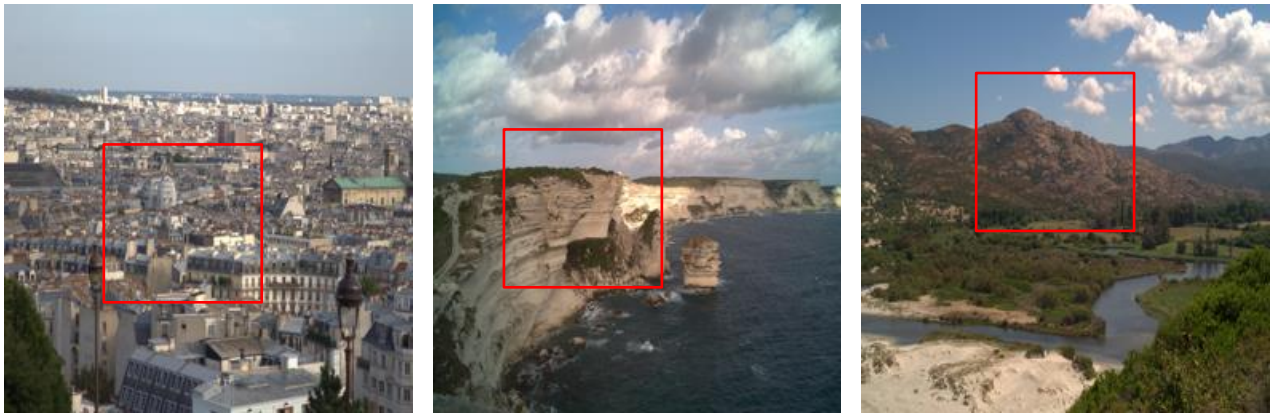
Figure 5 shows a set of experimental results of the proposed algorithm for three scenes, in which the upper part of each column is infrared image, the lower part is optical reference image, and the red rectangle represents the matching position. the MP of urban scenes in Figure 5 (a1) and figure 5 (a2) is 2.3, the MP of harbor scenes in Figure 5 (b1) and figure 5 (b2) is 1.2, and the MP of mountain scenes in Figure 5 (c1) and figure 5 (c2) is 4.5. It is obvious that the two scenes match each other.

(a1) Urban infrared remote image

(b1) Harbor infrared remote image

(c1) Mountain infrared image

ISSN: 2664-9640



(a2) Urban optical reference image (b2) Harbor optical reference image (c2) Mountain optical reference image

Fig. 5 SAR and optical image matching results

5. Conclusion

In the case of insufficient training samples, the limitation of heterologous image matching based on image style transfer is verified; the method of generating difference map based on image style transfer before and after image is proposed; the method of extracting unchanged area of artificial optical image based on difference map of style transfer is proposed; the method of accurately extracting infrared and optical image based on wavelet multi-scale edge enhancement and Canny operator is summarized In style transfer, the method of edge invariant feature is used to realize the accurate registration of infrared and optical images under the condition of insufficient training samples. The experimental results show that the proposed algorithm can achieve the same matching accuracy of infrared and optical images by using conventional matching algorithm when the training samples are far lower than those in reference [9]. The application scope of the heterogeneous image matching algorithm based on image style transfer is expanded. However, this algorithm does not consider the distortion between the images to be matched, which will be the focus of the next step.

Acknowledgements

The authors gratefully acknowledge the financial support from National Natural Science Foundation of China.

References

[1] Hao Shuai, Wu Yingqi, Ma Xu, He Tian, Wen Hu, Wang Feng. Visible and infrared image matching based on cyclegan-sift [J]. Optical precision engineering, 2022,30 (05): 602-614.

[2] TAO Z, JIAN J P, PEI Y Z, et al. Mapping Winter Wheat with Multi-temporal SAR and Optical Images in an Urban Agricultural Region[J]. Sensors, 2017, 17(6):1210-1215.

[3] Brown L G. A Survey of Image Registration Techniques[J]. ACM Computing Surveys, 1999,24(4): 325-376.

[4] SURI S, REINARTZ P. Mutual-Information-Based Registration of TerraSAR[J]. Geoscience & Remote Sensing IEEE Transactions on, 2010, 48(2):939 - 949.

[5] HONG T D, SCHOWENGERDT R A. A Robust Technique for Precise Registration of Radar and Optical Satellite Images[J]. Photogrammetric Engineering & Remote Sensing, 2005,71(5):585-593.

[6] AN Y, SUN W. A Contour-Based Approach to Multisensor Image Registration Using Narrow Angle Features[C]// Proceedings of the 2012 Second International Conference on Electric Information and Control Engineering - Volume 01. IEEE Computer Society, 2012:325-332.

ISSN: 2664-9640

[7] LUO J, OUBONG G. A Comparison of SIFT, PCA-SIFT and SURF[J]. International Journal of Image Processing, 2009:131-183.

[8] MATAS J, CHUM O, URBAN M, et al. Robust wide-baseline stereo from maximally stable external regions. Image Vision Computing, 2004, 22(10): 761-769.

[9] MERKLE N, AUER S, MULLER R , et al. Exploring the Potential of Conditional Adversarial Networks for Optical and SAR Image Matching[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2018:1-10.

[10] GATYS L A , ECKER A S. Image Style Transfer Using Convolutional Neural Networks[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016:309-312.

[11] CHEN S H, WEI Y K, XU L, et al. Survey of image style transfer based on deep learning[J]. Application Research of Computers, 2019, 36(8):2250-2255.

[12] GOODFELLOW I J, POUGHT A J, MIRZA M, et al. Generative Adversarial Networks[J]. Advances in Neural Information Processing Systems, 2014, 3:2672-2680.

[13] ISOLA P, ZHU J Y, ZHOU T, et al. Image to Image Translation with Conditional Adversarial Networks [C]// 2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017:167-173.

[14] ZHU J Y, PARK T, ISOLA P, et al. Unpaired Image to Image Translation using Cycle-Consistent Adversarial Networks[C]// 2017 IEEE International Conference on Computer Vision.2017:33-43.

[15] ChOI Y , ChOI M , KIM M , et al. StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-Image Translation[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018:122-130.

[16] MIRZA M, OSINDERO S . Conditional Generative Adversarial Nets[J]. Computer Science,2014:2672-2680.

[17] DU Z X, YIN J Y, YANG J. Remote Sensing Aircraft Image Detection Based on Semi-Supervised Learning[J]. Laser & Optoelectronics Progress, 2020, 57(6):123-131.

[18] LI D R. Change Detection from Remote Sensing images[J]. Geomatics and Information Science of Wu Han University, 2003,28(s1):7-12.

[19] MAO T Q, LIU W, HANG J, et al. Change Detection of SAR Images Using Dyadic Wavelet Enhancement and Edge Local Information FCM[J]. Journal of Signal Processing, 2018, 34(1):54-61.

[20] WANG Z S, YANG F B, JI L E, et al. Optical and SAR image registration based on cluster segmentation and mathematical morphplogy[J]. Acta Optica Sinica, 2014, 34(2):176-182.