IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 2, 2019, accepted November 11, 2019, date of publication November 14, 2019,
date of current version November 25, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2953539

Method for Fused Phase and PCA Direction Based on a SIFT Framework for Multi-Modal Image Matching

HAIQIAO LIU[1], SHIBIN LUO[2], JIAZHEN LU[2], AND JING DONG[2]
[1]School of Automation, Central South University, Changsha 410083, China
[2]School of Aeronautics and Astronautics, Central South University, Changsha 410083, China

Corresponding author: Jiazhen Lu (ljzbuaa@163.com)

This work was supported in part by the National Natural Science Foundations of China under Grant 61802423 and in part by the Fundamental Research Funds for the Central Universities of Central South University under Grant 2019zzts871.

ABSTRACT The high noise and local deformation of multi-modal images reduce the accuracy of scale invariant feature transform (SIFT) image matching. To solve this problem, a new method based on the SIFT framework, which fuses the phase consistency optimization strategy and the gradient direction of principal component analysis (PCA) with the 8 direction of latitude reduction, is proposed in this paper. This method fuses the histogram of the orientated phase congruency (HOPC) method to extract the direction of the image, and adopts PCA to extract the main direction, which effectively solves the problem that the matching accuracy decreases due to the inversion of the direction of the image. Using the image phase instead of gradient intensity, the difficult problem of direction extraction is effectively solved when image edge characteristics are not obvious. Finally, the random sample consistency (RANSAC) algorithm is used to eliminate false match points. Simulation and experiments show that compared with the SIFT algorithm and PCA-SIFT algorithm, the proposed method improved the number of match points and matching accuracy, significantly reduced the mismatching rate. The statistical results show that the number of match points raised in this paper increases by 20.1% and 200% respectively compared with the former two algorithms.

INDEX TERMS Phase consistency, SIFT, HOPC, image matching.

I. INTRODUCTION

Multi-modal image matching has a wide range of applications in the fields of medicine, remote sensing, and navigation [1]–[3]. Due to the significant difference in image contrast, image brightness and image texture of multi-modal images [4]. Multi-modal image matching has always been a major problem and has not been completely solved [5], [6].

Multi-modal image matching can be classified into two types according to methods: template matching and feature matching [7], [8]. Based on template matching, gray or edge information of the entire template area is matched, which mainly includes gray similarity matching, gradient similarity matching and the mutual information correlation method [9], [10]. Due to the different generation mechanisms of different image matching, methods based on template matching have high computational complexity, low robustness and unstable performance [11]–[13].

Multi-modal image matching based on feature matching uses feature similarity of images to match images, including points, lines and edges [8], [14], [15]. The literature [16] used shape context descriptions of feature points to match infrared and visible images, which improves the matching stability. The literature [17] proposed a matching algorithm based on linear features and virtual intersection points to match airport images. The literature [18] used edge features and improved Hausdorff distance to achieve infrared and visible image matching. However, due to the distortion of different images, it is difficult to extract the above features to correspond with each other [19]. Higher requirements are proposed for the adaptability and stability of the features when they are selected [20].

Point features are widely used for image matching. According to different methods of selecting points, there is an approach based on machine learning. Some recent works

The associate editor coordinating the review of this manuscript and approving it for publication was Xiao-Yu Zhang.

try to solve the image matching problems with machine learning methods [21]–[25]. However, all the methods are developed for homologous images, none of them can handle the complicate nonlinear gray difference between multi-modal images. Besides, there is another approach based on traditional method. David Lowe proposed a SIFT method to conduct image matching [26]. The SIFT feature has invariants of scale, rotation, brightness, and affine, and is widely used in multi-modal image matching. In 2006, Bay et al. improved the site selection process on the basis of SIFT and proposed the speeded up robust features (SURF) algorithm, which improved the computational speed of the algorithm [27]. The literature [28] achieved the matching of infrared and visible images by Canny and SURF algorithms. The following year, the literature [29] matched the SURF algorithm with the improved thrice-b spline edge detection algorithm. Although their method had some effect, it did not fully consider the similarities and differences in imaging properties of different images.

In addition, although the SURF algorithm has advantages over the SIFT algorithm in speed, the SIFT algorithm is not as good as the SIFT algorithm in image processing when light and angle of view change. Proposed by Yan the paper uses PCA to SIFT feature descriptor from 128 dimensions to 20 dimensions, which optimized the descriptor memory and simultaneously improved the precision of matching [30]. The algorithm greatly improved the performance, but the problem of direction reversal and noise-induced image direction extraction errors still was not considered [31], [32].

In this paper, the SIFT algorithm fusing phase consistency and PCA direction is proposed to solve the problem of the traditional SIFT algorithm not being accurate in extracting gradient direction due to high noise and local deformation, resulting in low matching accuracy. In this paper, the algorithm is applied to multi-modal image matching using the image phase instead of the image gradient intensity. The HOPC [33] method is used to extract the image direction, and the dimension is reduced by PCA. Finally, the RANSAC algorithm [34] is used to eliminate false match points.

Under the current conditions, there is no universal method to solve the Multi-modal image matching problem. For the problem of noisy multi-modal images and the reversed image direction, the gradient intensity and gradient direction in the traditional SIFT algorithm are replaced with the gradient direction extracted by the phase and HOPC method, resulting in higher accuracy and faster speed than the traditional SIFT algorithm and PCA-SIFT image matching results. The main innovation points of this paper include the following points:

1) Image angle is used instead of image gradient, which effectively solves the problem of image direction extraction under noise.

2) The HOPC method is adopted to extract the image direction, and PCA is combined to replace the gradient direction in the traditional SIFT algorithm to describe the image main direction, effectively solving the image direction reversal problem. This method is hereinafter referred to as HP-SIFT.

The chapter structure of this paper is arranged as follows. Chapter I studies the SIFT algorithm of the fusion phase and the PCA direction step by step. In chapter II, the SIFT algorithm with phase fusion and PCA direction is tested and discussed. Finally, chapter III is the conclusion.

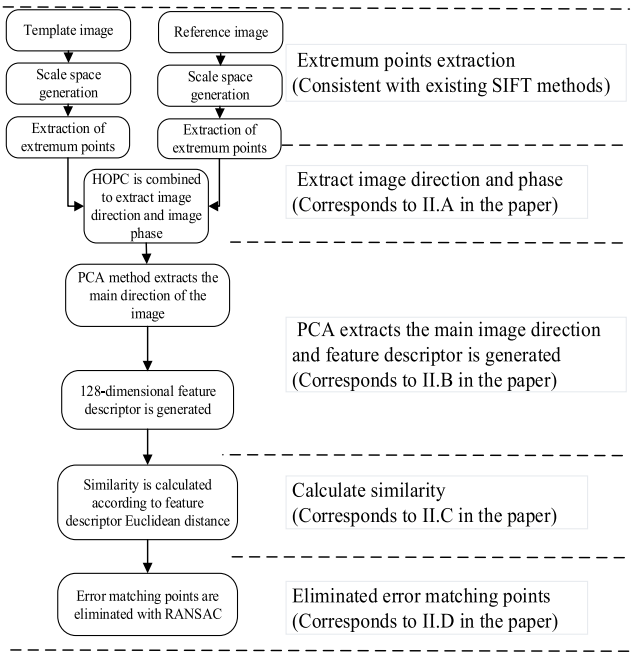


FIGURE 1. Flow chart of the SIFT algorithm combining phase and PCA direction.

II. SIFT ALGORITHM COMBINING PHASE AND PCA DIRECTION

The overall flow chart of the HP-SIFT algorithm is shown in Fig 1 below. The method mainly includes the following steps:

1) Image extremum point extraction is performed according to the method of SIFT extreme point extraction;

2) According to the Gaussian image of the SIFT algorithm and the HOPC method, the phase and direction of the different images are extracted;

3) The main direction of the new constructed feature direction is extracted with PCA, and the gradient direction is projected to 0-180 degrees;

4) The feature descriptor of the image is generated;

5) The similarity of the feature description is calculated, and the matching pairs of a reference image and a real-time image is obtained;

6) The matching pairs are combined with RANSAC, the false matching pairs are eliminated, and the results are compared to obtain the experimental results of the algorithm.

The first step of the algorithm is to extract the extreme points according to the extreme point extraction method of the SIFT algorithm. The second step of the algorithm corresponds to II.A in the paper, the third and fourth step of the algorithm corresponds to II.B in the paper, the fifth step of

the algorithm corresponds to II.C in the paper, and the sixth step of the algorithm corresponds to II.D in the paper.

A. PHASE AND GRADIENT DIRECTION CALCULATION WITH THE HP-SIFT ALGORITHM

The HP-SIFT algorithm first needs to calculate the phase and direction of the template image and the real-time image. The phase calculation method is as shown in Equation 1:

$$PC(x,y) = \frac{\sum_0 \sum_n W_o(x,y) \lfloor A_{no}(x,y)\Delta\Phi_{no}(x,y) - T \rfloor}{\sum_o \sum_n A_{no}(x,y) + \varepsilon} \tag{1}$$

$PC(x, y)$ is the size of phase congruence, (x, y) represents the coordinates of the point in the image, $W_o(x, y)$ is the weighting factor of the frequency propagation, $A_{no}(x, y)$ is the amplitude at the wavelet scale n and direction o at (x, y), $\Delta\Phi_{no}(x, y)$ is a sensitive phase deviation, and T is a noise threshold and a small constant to avoid being divided by zero. $\lfloor \rfloor$ indicates that when the value is positive, the amount contained is equal to itself, otherwise, it is zero.

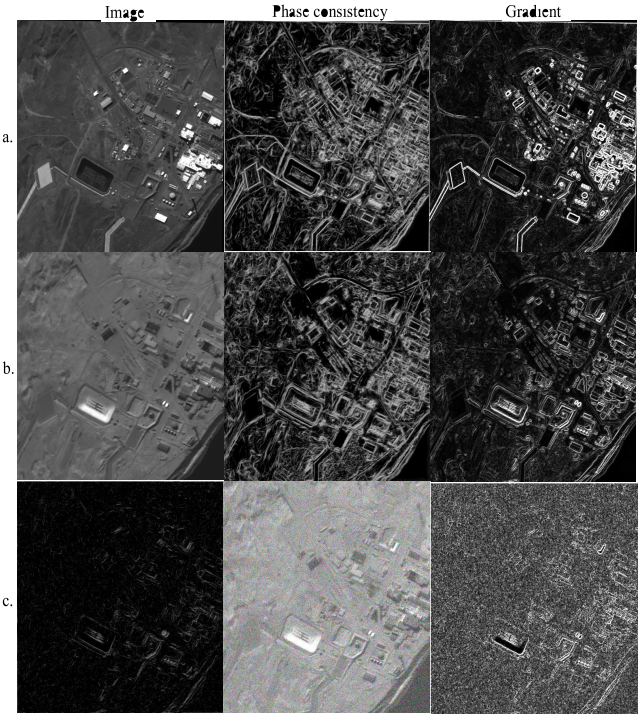


FIGURE 2. (a) SAR image phase and gradient (b) Infrared image phase and gradient (c) Infrared image with noise phase and gradient.

A weighted average of noise based on image phase consistency can effectively adapt to noise and brightness. Compared with the gradient intensity, as shown in Fig 2, the left side of Fig 2 shows the original image, which is generally a template image and a real-time image. The intermediate image is a phase-consistent extraction result, and the right image is a gradient image of the original image. The results show that the image phase can effectively adapt to illumination changes. According to the contour information of the extracted image, it can be clearly seen that the phase can better solve the gray distortion and geometric distortion.

The HP-SIFT algorithm requires image gradient direction calculation. The direction calculation formula is shown in the following 2-4, where a and b represent the gradient directions in the x and y directions, respectively:

$$a = \sum_{\theta} (O_{no}(\theta)cos(\theta)) \tag{2}$$

$$b = \sum_{\theta} (O_{no}(\theta)sin(\theta)) \tag{3}$$

$$\Phi = arctan(b, a) \tag{4}$$

where Φ is the direction of phase consistency and $o_{no}(\theta)$ is the result of wavelet transform convolution calculation.

B. HP-SIFT ALGORITHM GRADIENT DIRECTION PCA MAIN DIRECTION EXTRACTION ALGORITHM AND FEATURE DESCRIPTOR GENERATION

The PCA algorithm is the most commonly used main direction extraction method for reducing data complexity and identifying the most important features. It maps high-dimensional data to low-dimensional space, which uses fewer data dimensions and has the largest variance in smaller dimensions. This is the least loss method of raw data information.

The PCA-SIFT compresses the SIFT descriptor data. First, we collect all the characteristics of the data, transform the data, observe the important components of the data and classify them to reduce the data. To express an object, the object has many attributes, we can observe the importance of each attribute by transforming the data, and then select several important attributes to describe the object, which plays a role in compressing the data. Specific steps are as follows:

(1) The construction description subregion is selected as a 41*41 rectangle centered on the feature point (which has been aligned with the main direction of the feature point).

(2) Since the outermost pixel does not calculate the partial derivative, the partial derivative of the horizontal and vertical directions of each pixel of the 39*39 matrix is calculated, and a vector of 39*39*2 = 3042 dimensions is obtained and normalized.

(3) Assuming that there are N feature points, vectors that describe all feature points form a matrix, and then calculate the covariance matrix of N vectors.

(4) Calculate the eigenvector corresponding to the first m maximum eigenvalue of the N*N covariance matrix, and m vectors constitute a 3024*m projection matrix.

(5) Multiply the N*3024 descriptor submatrix with the 3024*m projection matrix to obtain an N*m matrix composed of descriptor vectors. At this time, the descriptor vectors of the N feature points are all m dimensions, and the data are converted into the new space constructed by the above m feature vectors.

The main direction of phase is extracted by the PCA method. The extraction result is shown in Fig 3 below.

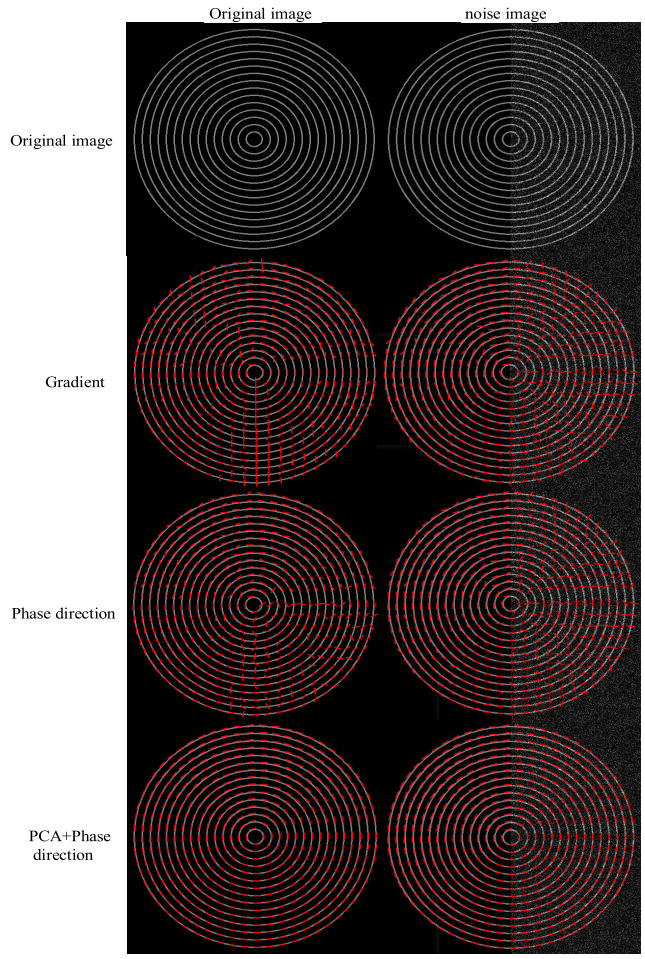


FIGURE 3. Image phase and image gradient extraction image direction comparison chart.

The left side shows the original picture, the middle is the phase direction extraction result, and the right side is the extraction result of the phase direction combined with the PCA image. It can be seen from the figure that there is no large difference between the three methods in the absence of noise. When there is noise in the image, the method proposed in this passage is more accurate.

The calculated phase and phase directions are combined with the SIFT algorithm step to generate the feature descriptor, which includes the following.

Description of key points. Considering that each key point has three pieces of information: position, scale and direction, then a descriptor is created for each key point, and a set of vectors is used to describe the key points so that they do not change with the environment, such as lighting changes and viewing angle changes. The descriptor includes not only the key points, but also the pixels that contribute to the key points. The description is unique to improving the probability of correct matching of the feature vectors.

The pixels in the neighborhood are divided into 16×16 subdomains and further divided into 4×4 blocks (each block is composed of 4×4 small fields). For details, see 8 directions in each block. Gradient direction histogram. Finally, a descriptor vector of $4 \times 4 \times 8 = 128$ is obtained. The process is shown in Fig 4 below.

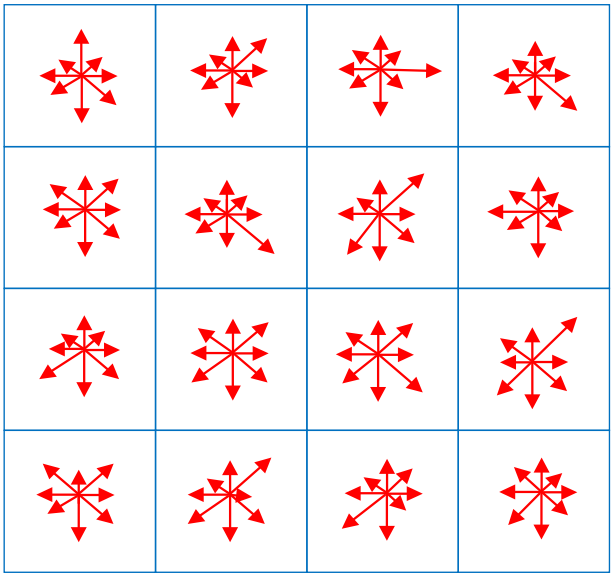


FIGURE 4. 4∗4 Direction map.

According to the above diagram, the direction of each extreme point is calculated to obtain the direction of 8 directions. Forms a 128-dimensional feature descriptor. Then, histogram statistics are performed on the 128-dimensional feature vector.

Finally, the feature descriptors are normalized according to the following formula.

$$W = (w_1, w_2, \ldots, w_{128}) \tag{5}$$

$$L = (l_1, l_2, \ldots l_{128}) \tag{6}$$

$$l_j = w_j / \sqrt{\sum_{i=1}^{128} w_i} \quad j = 1, 2, \ldots 128 \tag{7}$$

where W represents a 128-dimensional feature descriptor and L represents a normalized feature descriptor.

C. HP-SIFT ALGORITHM SIMILARITY MEASUREMENT METHOD

In this paper, the similarity calculation method of the L2 norm is adopted. The specific calculation method is as shown in formula 8:

$$\text{HP-SIFT} = \frac{\sum_{k=1}^{n} (V_A(k) - \bar{V}_A)(V_B(k) - \bar{V}_B)}{\sqrt{\sum_{k=1}^{n} (V_A(k) - \bar{V}_A)^2 \sum_{k=1}^{n} (V_B(k) - \bar{V}_B)^2}} \tag{8}$$

The similarity calculation of HP-SIFT, V_A and V_B represent the feature descriptors of the two images HP-SIFT, $\bar{V}_A$ and $\bar{V}_B$ represent the means of V_A and V_B.

D. RANSAC-BASED HP-SIFT MISMATCH ELIMINATING

The parameter estimation of the RANSAC algorithm model is completed by iterative calculation and repeated testing through the inner and outer points. The initial model parameters are calculated by randomly extracting the sample data, so there is great uncertainty. The quality of initial parameters directly determines the number of iterations and the cost of calculation. The distance ratio threshold is a bond that establishes a matching relationship between feature points. Lowe's experiment gives the relationship between the distance ratio and matching rate. That is, the smaller the threshold, the larger the correct probability of matching points, but the fewer the number of matching points; the larger the threshold, the more matching points, but the smaller the matching accuracy.

In this paper, the method of simplifying RANSAC is to select a few optimal matching points as representative of the observation data set, so that the range of observation data is greatly reduced and the overall quality is improved. The initial model parameters obtained by randomly extracting from the initial sample data are closer to the true value, and the most realistic homography matrix parameters can be obtained by minimizing the number of iterations.

The entire RANSAC geometry verification process is to fit the image transformation matrix through the sample dataset. The relationship between the original image and the image to be matched is as follows:

$$\begin{bmatrix} \omega x_2 \\ \omega y_2 \\ \omega \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \tag{9}$$

This is a projection transformation matrix of 8 parameters. At least four matching pairs are needed to obtain the matrix parameters. The least squares method is used to solve these 8 parameters, so that:

$$\mathrm{H} = [\mathrm{h}_1 \quad \mathrm{h}_2 \quad \mathrm{h}_3 \quad \mathrm{h}_4 \quad \mathrm{h}_5 \quad \mathrm{h}_6 \quad \mathrm{h}_7 \quad \mathrm{h}_8] \tag{10}$$

$$\mathrm{L} = -\frac{1}{\omega}[\mathrm{x}_1 \mathrm{y}_2]^{\mathrm{T}} \tag{11}$$

$$\mathrm{C} = \frac{1}{\omega}\begin{bmatrix} -\mathrm{x}_1 & \mathrm{y}_1 & 1 & 0 & 0 & 0 & -\mathrm{x}_2\mathrm{x}_1 & -\mathrm{x}_2\mathrm{y}_1 \\ 0 & 0 & 0 & \mathrm{x}_1 & \mathrm{y}_1 & 1 & -\mathrm{y}_2\mathrm{x}_1 & -\mathrm{y}_2\mathrm{y}_1 \end{bmatrix} \tag{12}$$

Thus, the formula changes as follows:

$$\mathrm{H} = -[\mathrm{C}^{\mathrm{T}}\mathrm{C}]^{-1}\mathrm{C}^{\mathrm{T}}\mathrm{L} \tag{13}$$

Let a have an initial value of 1 first, and a set of H values is obtained. Then, the ω value is obtained from this value. After many iterations, a stable H is obtained. The initial sample data number N can be determined as follows:

$$N = min\{N_0, max\{N_s, N_s log_2 \mu N_0\}\} \tag{14}$$

Here N_0 is the number of matching points determined according to the nearest neighbor ratio method, and $\mathrm{N}_0 \geq 4$. N_s is the number of sample steps. μ is a scaling factor.

III. COMPUTATIONAL COMPLEXITY ANALYSIS

The SIFT algorithm and the improved algorithm based on SIFT, the complexity of the algorithm is time-consuming: extreme point extraction, feature descriptor and similarity calculation.

The extraction of the extreme points of the algorithm is consistent with the SIFT and PCA-SIFT methods, and it takes the same time. The establishment of feature descriptions is different. The specific analysis is as follows: N extreme points are extracted for each image; similar to the traditional SIFT algorithm, the HP-SIFT algorithm simply replaces the gradient strength calculation with the phase calculation, and the gradient direction calculation is equivalent. It only increases the time consumption of the PCA calculation; the implementation of the PCA algorithm does not have a large number of calculated sliding window operations, and the complexity of the algorithm is limited. The complexity of the HP-SIFT algorithm is equivalent to the SIFT algorithm.

IV. EXPERIMENT RESULTS AND ANALYSIS

To verify the performance of the algorithm, this simulation experiment was designed. The image data used in the experiment were based on network collection and partial self-photographing. The experiment verifies the algorithm from two types of data:

1) Heterogenous image matching without significant noise;

2) Multi-modal image matching with artificially added non-mean, Gaussian noise with a variance of 0.1.

In the comparison process, the gradient directions of the three algorithms are compared. The correct matching pairs of the three algorithms are compared and the correct matching pairs are compared. The proportion of the increase is calculated according to the following formula:

$$\eta = \frac{N_1 - N_2}{N_2} \tag{15}$$

where N_1 and N_2 represent the correct matching pair of the comparison algorithm and the correct matching pair of the improved algorithm, respectively, and η indicates the improvement ratio of the improved algorithm.

A. EXPERIMENTAL PREPARATION

1) INTRODUCTION OF EXPERIMENTAL DATA

Part of the multi-modal image used in the experiment came from the experimental map of the reference and partly from the image downloaded from a Google map. Four pairs of representative images were selected, as shown in Fig 5 below:

2) INTRODUCTION OF THE COMPARISON METHOD

The method of this paper was compared with the SIFT and PCA-SIFT algorithms. The running code of the SIFT algorithm is provided in [26]. The PCA-SIFT method in [30] is used to perform experiments for comparison.

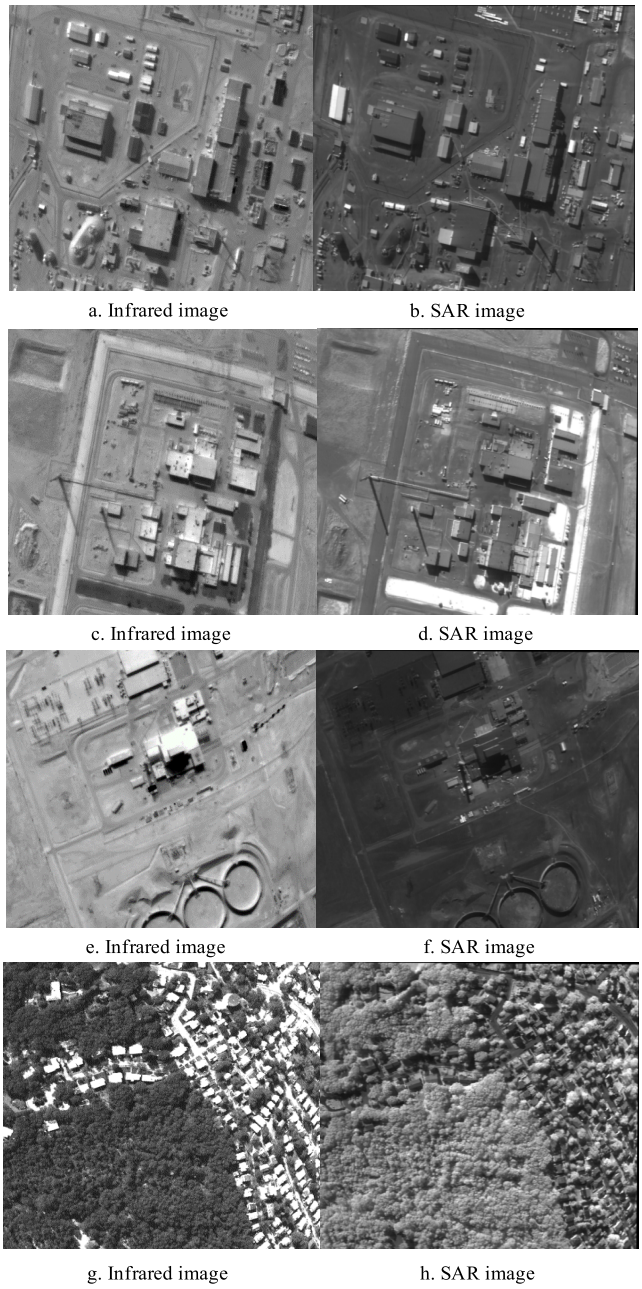


FIGURE 5. Experimental data.

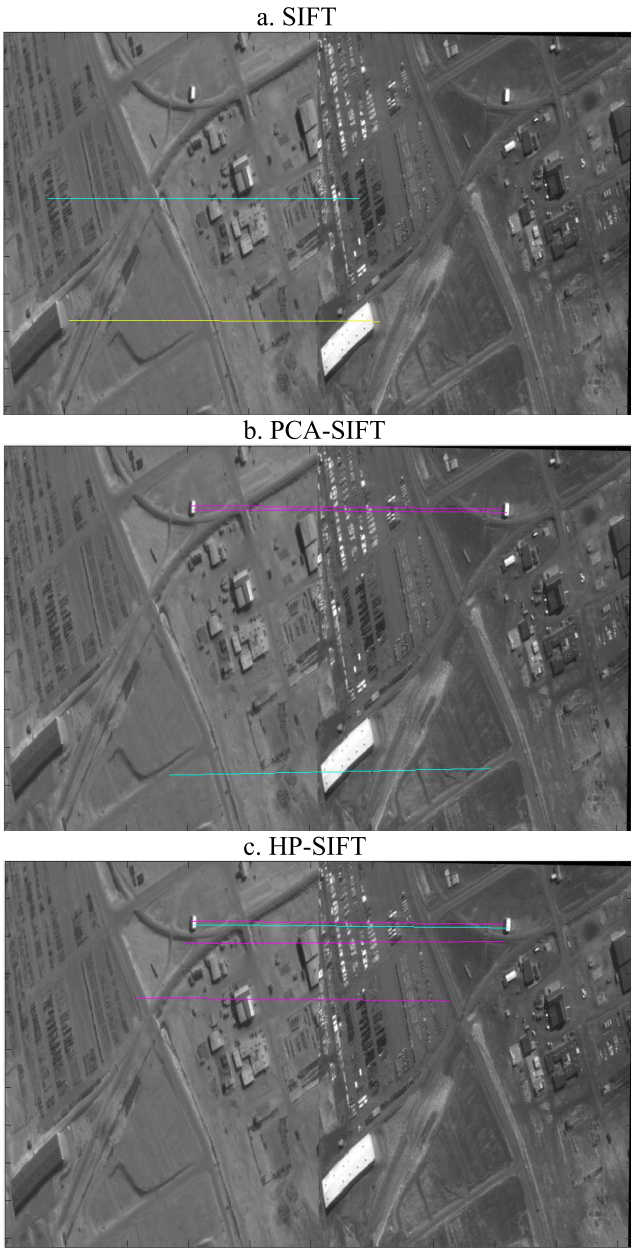


FIGURE 6. Comparison of SIFT(a), PCA-SIFT(b) and HP-SIFT(c) effects under normal conditions.

B. EXPERIMENT 1 - EXPERIMENTAL VERIFICATION UNDER NORMAL CONDITIONS

The 20 pairs of image libraries prepared for the experiment were plotted to make the image an ideal matching pair without affine transformation. Experiments were performed using SIFT, PCA-SIFT, and HP-SIFT. The matching results of the SIFT, PCA-SIFT and HP-SIFT algorithms are shown in Fig 6 below. A representative set of image matching pairs was selected for display, as shown in Fig 6 below:

It can be seen from Fig 6 that using the SIFT algorithm to match in the infrared and SAR images only obtained two correct pairs, the same image using the PCA-SIFT algorithm to match had three correct results, and the correct rate was increased by 50% compared with the SIFT algorithm. The HP-SIFT algorithm had six correct matching pairs, which was 200% better than the traditional SIFT algorithm and 100% higher than the HP-SIFT algorithm.

In Fig 7, the blue line indicates the matching results of 20 sets of data using the SIFT algorithm; the orange line indicates the matching results of 20 sets of data using the PCA-SIFT algorithm; the gray line indicates the matching results of 20 sets of data using the HP-SIFT algorithm; it can be seen from the figure that the results of the HP-SIFT algorithm were always higher than those of the SIFT and PCA-SIFT algorithms, and the improved matching accuracy was basically maintained above 60%.

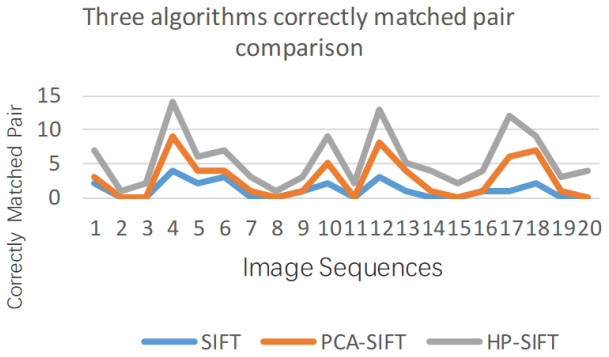


FIGURE 7. Comparison of statistical results between SIFT, PCA-SIFT and HP-SIFT in normal conditions.

C. EXPERIMENT 2 - EXPERIMENTAL VERIFICATION UNDER NOISE CONDITIONS

The 20 pairs of image libraries prepared for the experiment were plotted, adding Gaussian noise with a variance of 0.1. Experiments were performed using SIFT, PCA-SIFT, and HP-SIFT. The matching results of the SIFT, PCA-SIFT and HP-SIFT algorithms are shown in Fig 9 below. A representative set of image matching pairs was selected for display, as shown in Fig 8 below:

It can be seen from Fig 8 that using the SIFT algorithm to match the infrared and SAR images obtained no correct results. The same image using the PCA-SIFT algorithm had two correct matching pairs; the HP-SIFT algorithm had five correct matching pairs. Compared with the traditional SIFT algorithm, it had a significant improvement, which was 150% higher than that of the HP-SIFT algorithm.

In Fig 9, the blue line indicates the results of 20 sets of data using the SIFT algorithm in the case of noise; the orange line indicates the results of 20 sets of data using the PCA-SIFT algorithm in the case of noise; the gray line indicates the results of 20 sets of data using the HP-SIFT algorithm in the case of noise. The results of the HP-SIFT algorithm were always higher than those of the SIFT and PCA-SIFT algorithms, and the improved matching accuracy was maintained above 80%, indicating that the HP-SIFT algorithm has better antinoise characteristics.

D. EXPERIMENT 2 - EXPERIMENTAL VERIFICATION UNDER NOISE CONDITIONS

Three algorithms use the same device for calculation, experimental image size is 640*512. This paper compares the time consumption of different algorithms. The results are shown in the following table:

TABLE 1. Time consumption comparison of the three algorithms.

Algorithm	SIFT	PCA-SIFT	HP-SIFT
Time consuming / s	1.3	1.15	1.56

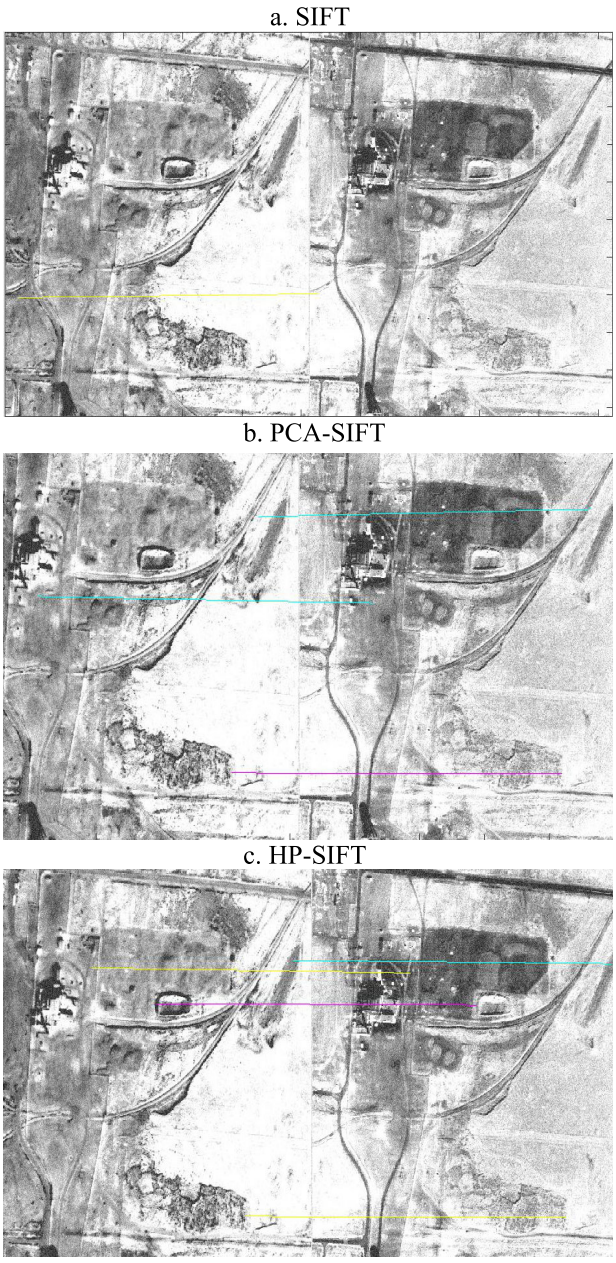


FIGURE 8. Comparison of SIFT(a), PCA-SIFT(b) and HP-SIFT(c) effects in the case of noise.

As seen in the Table, to improve the accuracy of the algorithm, the method adds some time consumption to the algorithm.

E. DISCUSSION

This paper is aimed at the problem of low precision and poor robustness in heterogeneous image matching in the traditional SIFT algorithm. The phase-replacement gradient strength was used to calculate the gradient direction based on the phase intensity, and the HP-SIFT was obtained by combining the SIFT framework. Compared with the traditional SIFT and PCA-SIFT algorithms, the results showed that the HP-SIFT algorithm extracts 10% more points than the SIFT algorithm, 5% more points than the PCA-SIFT algorithm, and the correct

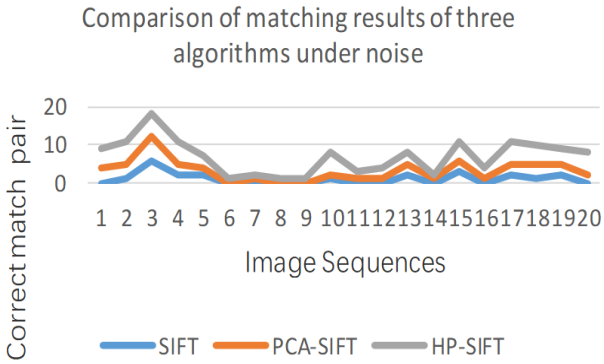


FIGURE 9. Comparison of SIFT, PCA-SIFT and HP-SIFT correct matching statistics in the case of noise.

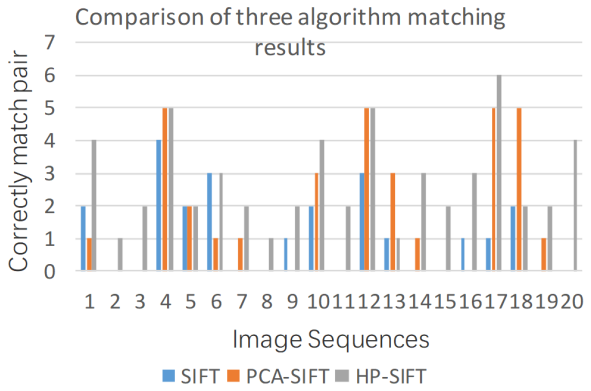


FIGURE 10. Comparison of statistical results between SIFT, PCA-SIFT and HP-SIFT in normal conditions.

matching point pairs increased by 12% and 11% respectively. The computational complexity increased, but there was no increase in magnitude.

In the absence of affine transformation and noise, the SIFT algorithm had certain effects in infrared and visible image matching due to its invariance to scale and rotation, but the performance was poor in SAR images. The PCA-SIFT algorithm was obtained from the SIFT algorithm, but the principal components of the feature vector were extracted, which had a certain improvement compared with the SIFT algorithm. The HP-SIFT algorithm better solved the illumination variation and geometry distortion and other issues by replacing the gradient features with phase consistency. The matching results of the 20 sets of data are shown in Fig 10 below:

In Fig 10, the blue histogram shows the matching results of the SIFT algorithm of 20 sets of data; the orange histogram shows the matching results of the PCA-SIFT algorithm of 20 sets of data; the gray histogram shows the matching results of the HP-SIFT algorithm of 20 sets of data; it can be seen from the figure that the results of the HP-SIFT algorithm are always higher than that of the SIFT algorithm, and the improved matching accuracy is basically maintained above 30%, indicating that the HP-SIFT algorithm has better antinoise characteristics.

In the noisy case, the SIFT algorithm is inaccurate due to the inaccurate extraction gradient direction; the PCA-SIFT

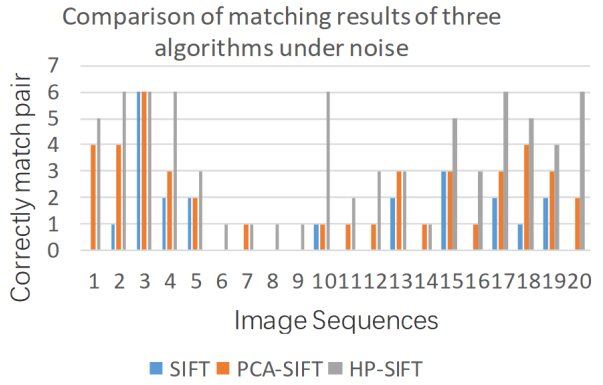


FIGURE 11. Comparison of SIFT, PCA-SIFT and HP-SIFT correct matching statistics in the case of noise.

algorithm improved on the basis of the SIFT algorithm, and the principal component of the gradient direction was extracted, which had a certain improvement compared with the SIFT algorithm; the HP-SIFT algorithm had a better effect by replacing the gradient features with phase consistency and performing PCA principal component extraction. The matching result statistics of the 20 sets of data are shown in Fig 11 below:

In Fig 11, the blue histogram shows the matching results of the SIFT algorithm of 20 sets of data in the case of noise; the orange histogram shows the matching results of the PCA-SIFT algorithm of 20 sets of data in the case of noise; the line indicates the matching results of the SIFT algorithm of 20 sets of data in the case of noise; as seen from the figure, the results of the HP-SIFT algorithm are basically higher than those of the SIFT and PCA-SIFT algorithms. The improved matching accuracy can be increased by 30% at the lowest and 500% at the highest, indicating that the HP-SIFT algorithm has better adaptability to noise.

V. CONCLUSION

In this paper, we propose an HP-SIFT heterogeneous image matching method based on SIFT. First, the phase and gradient direction of the image are extracted from the heterogeneous image to replace the gradient intensity and gradient direction in the traditional SIFT algorithm. Then, combined with the PCA to extract the gradient direction main direction, the accuracy of the gradient direction improves, and the feature descriptor is constructed. Finally, the RANSAC algorithm is used to eliminate the mismatch. The experimental results show that the matching performance of this method proposed in this paper is better than the traditional SIFT and PCA-SIFT algorithms.

REFERENCES

[1] W. Ouyang, F. Tombari, S. Mattoccia, L. Di Stefano, and W.-K. Cham, "Performance evaluation of full search equivalent pattern matching algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 127–143, Jan. 2012.

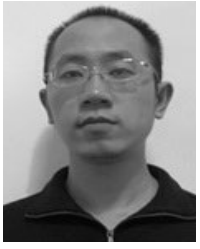
[2] W. Ouyang, T. Zhao, W.-K. Cham, and L. Wei, "Fast full-search-equivalent pattern matching using asymmetric Haar wavelet packets," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 4, pp. 819–833, Apr. 2018.

[3] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A SIFT-like algorithm for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 453–466, Jan. 2015.
[4] J. Fan, Y. Wu, F. Wang, Q. Zhang, G. Liao, and M. Li, "SAR image registration using phase congruency and nonlinear diffusion-based SIFT," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 562–566, Mar. 2015.
[5] I. Papila, S. Kent, and M. Kartal, "An automated SAR image registration approach using hidden Markov scale invariant feature transform algorithm," in *Proc. 10th Eur. Conf. Synth. Aperture Radar*, Jun. 2014, pp. 1–4.
[6] F. Wang, H. You, and X. Fu, "Adapted anisotropic Gaussian SIFT matching strategy for SAR registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 160–164, Jan. 2015.
[7] Q. Zhang, X. Shen, L. Xu, and J. Jia, "Rolling guidance filter," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 815–830.
[8] Z. Yang, Y. Yang, K. Yang, and Z.-Q. Wei, "Non-rigid image registration with dynamic Gaussian component density and space curvature preservation," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2584–2598, May 2019.
[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst (NIPS)*, 2012, pp. 1097–1105.
[10] G. Lv, "ℓm_p: A novel similarity measure for matching local image descriptors," *IEEE Access*, vol. 6, pp. 55315–55325, 2018.
[11] G. Lv, "A novel correspondence selection technique for affine rigid image registration," *IEEE Access*, vol. 6, pp. 32023–32034, 2018.
[12] X. Fengguang and H. Xie, "A 3D surface matching method using keypoint—Based covariance matrix descriptors," *IEEE Access*, vol. 5, pp. 14204–14220, 2017.
[13] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust registration of multimodal remote sensing images based on structural similarity," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2941–2958, Mar. 2017.
[14] C. Leng, H. Zhang, B. Li, G. Cai, Z. Pei, and L. He, "Local feature descriptor for image matching: A Survey," *IEEE Access*, vol. 7, pp. 6424–6434, 2019.
[15] G. Wang, Q. Zhou, and Y. Chen, "Robust non-rigid point set registration using spatially constrained Gaussian fields," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1759–1769, Apr. 2017.
[16] C. Guo, M. Rana, M. Cisse, and L. van der Maaten, "Countering adversarial images using input transformations," Oct. 2017, *arXiv:1711.00117*. [Online]. Available: https://arxiv.org/abs/1711.00117
[17] F. Tramèr, A. Kurakin, N. Papernot, I. Goodfellow, D. Boneh, and P. McDaniel, "Ensemble adversarial training: Attacks and defenses," May 2017, *arXiv:1705.07204*. [Online]. Available: https://arxiv.org/abs/1705.07204
[18] F. Tramèr, N. Papernot, I. Goodfellow, D. Boneh, and P. McDaniel, "The space of transferable adversarial examples," Apr. 2017, *arXiv:1704.03453*. [Online]. Available: https://arxiv.org/abs/1704.03453
[19] H. Wang, F. Zhang, X. Xie, and M. Guo, "DKN: Deep knowledge-aware network for news recommendation," in *Proc. World Wide Web Conf.*, Apr. 2018, pp. 1835–1844.
[20] D. Jakubovitz and R. Giryes, "Improving DNN robustness to adversarial attacks using jacobian regularization," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, 2018, pp. 525–541.
[21] X.-Y. Zhang, H. Shi, X. Zhu, and P. Li, "Active semi-supervised learning based on self-expressive correlation with generative adversarial networks," *Neurocomputing*, vol. 345, pp. 103–113, Jun. 2019.
[22] X. Zhang, S. Wang, X. Zhu, X. Yun, G. Wu, and Y. Wang, "Update vs. upgrade: Modeling with indeterminate multi-class active learning," *Neurocomputing*, vol. 162, pp. 163–170, Aug. 2015.
[23] X.-Y. Zhang, S. Wang, and X. Yun, "Bidirectional active learning: A two-way exploration into unlabeled and labeled data set," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 12, pp. 3034–3044, Dec. 2015.
[24] X. Y. Zhang, H. C. Shi, and C. S. Li, "Learning transferable self-attentive representations for action recognition in untrimmed videos with weak supervision," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, Feb. 2019, pp. 9227–9234.
[25] Z. Yang, T. Dan, and Y. Yang, "Multi-temporal remote sensing image registration using deep convolutional features," *IEEE Access*, vol. 6, pp. 38544–38555, 2018.
[26] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[27] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, 2008.
[28] Y. Ege, S. Nazlibilek, A. Kakilli, H. Citak, O. Kalender, D. Karacor, K. L. Erturk, and G. Sengul, "A study on the performance of magnetic material identification system by SIFT–BRISK and neural network methods," *IEEE Trans. Magn.*, vol. 51, no. 8, Aug. 2015, Art. no. 6201316.
[29] S. Wang, H. You, and K. Fu, "Bfsift: A novel method to find feature matches for SAR image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 649–653, Jul. 2012.
[30] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proc. CVPR*, vol. 2, Jul. 2004, pp. 506–513.
[31] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823.
[32] A. M. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 427–436.
[33] Y. Ye and L. Shen, "Hopc: A novel similarity metric based on geometric structural properties for multi-modal remote sensing image matching," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 3, pp. 9–16, Jul. 2016.
[34] T. Sattler, B. Leibe, and L. Kobbelt, "SCRAMSAC: Improving RANSAC's efficiency with a spatial consistency filter," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 2090–2097.

HAIQIAO LIU was born in 1989. He received the master's degree from the School of Electronic and Information Engineering, Guizhou University, Guiyang, China, in 2014. He is currently pursuing the Ph.D. degree with Central South University, Changsha, China. His current research interests are visual navigation and integrated navigation.

SHIBIN LUO was born in 1976. He received the Ph.D. degree from the School of Aeronautics and Astronautics, National University of Defense Technology, Changsha, China, in 2004. From 2004 to 2015, he was a Teacher with the National University of Defense Technology. He is currently a Professor with the School of Aeronautics and Astronautics, Central South University, Changsha. His current research interest is advance craft design.

JIAZHEN LU was born in 1982. He received the Ph.D. degree from the School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing, China, in 2009. From 2009 to 2019, he was a Lecturer with Beihang University. He is currently a Professor with the School of Aeronautics and Astronautics, Central South University, Changsha, China. His current research interests are integrated navigation and control engineering.



JING DONG was born in 1984. He received the M.S. and Ph.D. degrees from the National University of Defense Technology, Changsha, China, in 2015. He is currently a Supervisor with Central South University and a Lecturer with the College of Aerospace Science and Engineering, National University of Defense Technology. His research interests include image registration, moving object detection, and video stabilization.