Research on Image Detection and Matching Based on SIFT Features

Feng Guo, Jie Yang*, Yilei Chen, Bao Yao
Key Laboratory of Fiber Optic Sensing Technology and Information Processing, Ministry of Education,
Wuhan University of Technology, Wuhan, China
e-mail: * 466802324@qq.com

***Abstract*—Scale-invariant feature transform (SIFT) is a kind of computer vision algorithm used to detect and describe Local characteristics in images. It finds extreme points in scale-space and gets its coordinate, scale, orientation, which in final come into being a descriptor. This paper studied the theory of SIFT matching, use Euclid distance as similarity measurement of key points and use RANSAC to eliminate mismatches. The result shows that SIFT algorithm is invariant on rotations, translations and scaling and SIFT features have strong matching robustness for radiation transformation, perspective changes, illumination changes and noises. This paper also compare different results obtained by different ratio threshold and finally set 0.6 as the best value considering the balance number of matched points and matching accuracy. It is important to image recognition application.**

Keywords—SIFT；Features points；Match；RANSAC

I. INTRODUCTION

David Lowe proposed Scale Invariant Feature Transform algorithm (SIFT) for extracting local features. Mikolajczyk and Schmid proved that the SIFT operator is suitable for different situations such as illumination changes, resolution, geometric deformation, rotation and so on[1]. It has the best performance and robust feature matching results and it is widely used for the feature extraction of target recognition, panoramic stitching, robot navigation mapping, 3-d scene modeling and other fields[2].

There are few SIFT algorithms that can be applied to cases with high real-time requirements like image registration. Image registration has a wide range of applications in the fields of target detection, model reconstruction, motion estimation, feature matching[3], tumor detection, lesion localization, angiography, geological exploration and aerial reconnaissance[4]. It is generally divided into two categories: one is based on the frequency domain such as Fourier transform and wavelet transform; the other is based on the spatial domain which mainly includes region-based registration algorithm and feature-based registration algorithm[5]. However, the existing feature-based image registration methods have a common problem that the feature points[6] do not usually have the invariance of affine or perspective projection transformations. The SIFT feature points are not only invariant to image scaling, translation and rotation, but also partly invariant to illumination changes, affine and projection transformations[7], which is very suitable for the registration of remote sensing images.

In this paper, the Euclidean distance between SIFT feature points is used to measure the similarity, and the RANSAC algorithm is used to eliminate the mismatch[8]. Through the analysis and research of SIFT algorithm, the SIFT feature points of the two images are obtained, and the Euclidean distance between the two feature points is measured to realize image registration which is applied to ship image recognition[9].

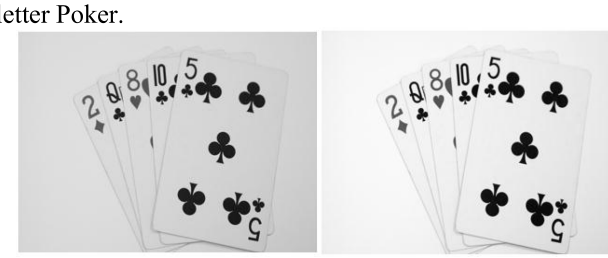
II. IMAGE BASED SIFT FEATURE EXTRACTION

Scale-invariant feature transform (SIFT) is a computer vision algorithm to detect and describe the local features in the image. It looks for the extreme point in the spatial scale and extracts its position, scale and rotation invariants for feature descriptors. The SIFT feature extraction process includes four steps of extreme value detection in scale space, key point location, direction assignment and eigenvector generation[10-11]. An example extract the image SIFT features to visualize the specific process with the parameter definition: inter-group spacing equals 2, low contrast threshold equals 0.04/3, and edge response point threshold equals 10. The following steps are running under matlabR2014b environment.

1. Initialize the image Poker. As shown in Fig.1, Convert the color image to a grayscale image, turn the image size to 240 rows and 320 columns, and then normalize pixel value.

2. Establish Gaussian scale space with a total of 4 groups as 6 layers each. As shown in Fig.2, the size of each layer within the same group becomes more and more blurred as the scale increases; Among the four groups, the size of the next group is 1/4 of the group.

3. Establish Gaussian differential pyramid, a total of 4 groups as 6 layers each. As shown in Fig.3, we can see that the image size gradually decreases as the number of groups increases, and the left part is at the edge of the image like the letter Poker.



(a) (b)

Figure 1. Initialization of original image: (a)Grayscale image; (b)Normalized image.

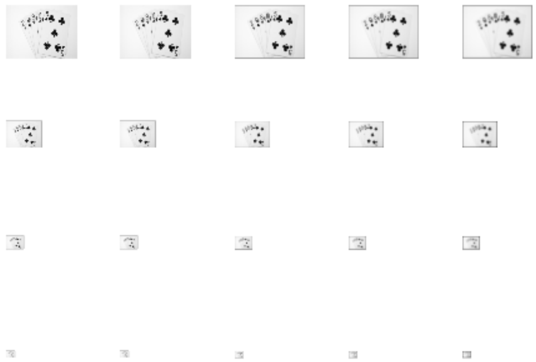


Figure 2. Gaussian scale space diagram.

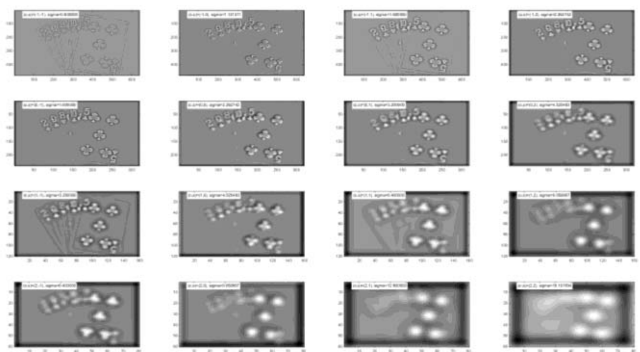


Figure 3. Gaussian differential pyramid.

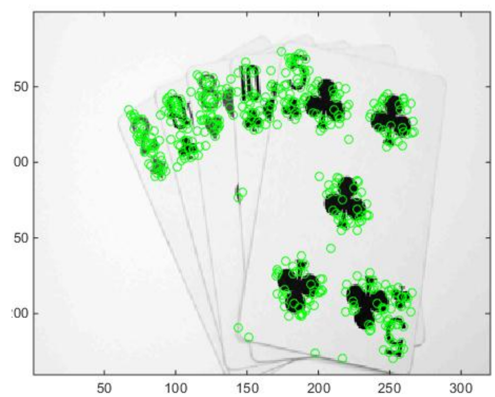


Figure 4. Local extreme points diagram.

4. Take the local extreme points. The specific distribution of extreme points are shown in Fig.4. It can be found that the local extreme points are mainly distributed in the contours.

5. Position the key points. The specific distribution of the key points are shown in Fig.5. Compare Fig.4 with Fig.5, some unstable points and low contrast points are filtered out, and the position of the remaining key points closer to the outline and slightly deviated from the extreme points.

6. Assign the direction of key points. The direction of each key point distribution is shown in Fig.6. Center of the circle is the key point, and the radius direction is the direction as the length is the gradient modulus.

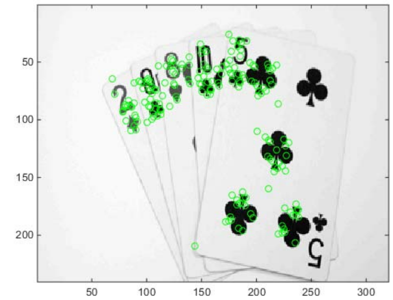


Figure 5. Key points distribution diagram.

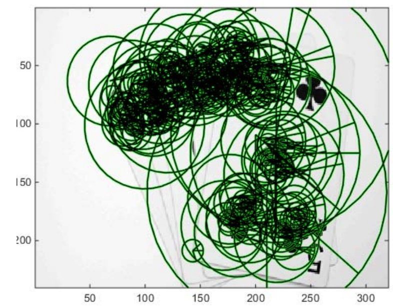
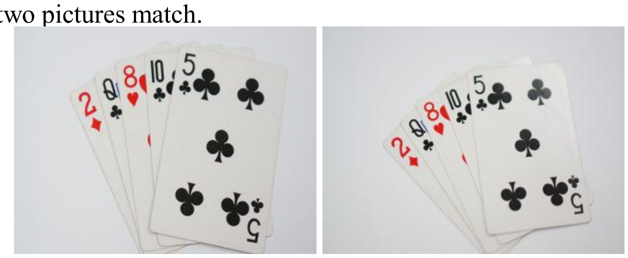


Figure 6. Direction distribution diagram of key points.

III. Image Recognition Based on SIFT Feature Registration

A. SIFT Feature Point Matching

After generating the SIFT feature vectors of the two images, the Euclidean distance can be used as the similarity measurement of the key points in the two images. Take one of the key points in one image and find the shortest two European distance of key points in the other image. If the result that shortest distance divides by the second shortest distance is less than some proportional threshold, accept the pair of matching points. Lowering this threshold can reduce SIFT matching points and stabilize the matching. In order to eliminate the key points caused by image occlusion and background clutter, Lowe proposed a method of comparing the nearest-neighbor distance to the second-nearest neighbor, and the distance ratio which is less than a certain threshold is considered as the correct match. For the mismatch, due to the high-dimensional feature space, similar distances may have a large number of other mismatches, resulting in a higher ratio value. There are a large number of studies on arbitrary existence of the scale, rotation and brightness changes in the two pictures match.



(a) (b)

Figure 7. Original images: (a) Image Poker1; (b).Image Poker2.

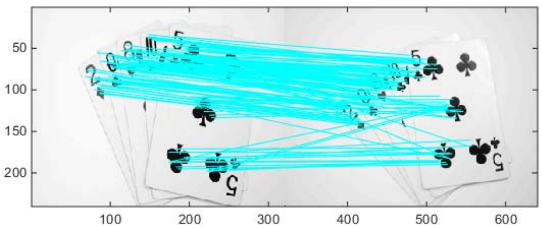


Figure 8. Result of matching diagram.

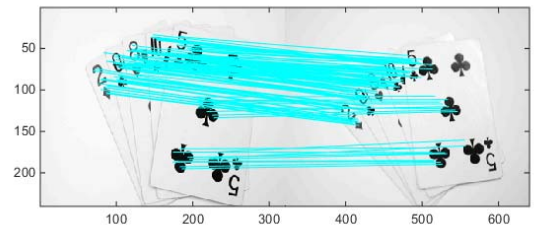


Figure 9. Result after mismatch elimination diagram.

Use vector angles to approximate the Euclidean distance: firstly normalize the 128-dimensional SIFT feature vector to a unit vector (each number divided by the square root of the sum of squares), then multiply it by the vector angle cosine value, and finally use the arccosine to determine the angle between the vectors. The original images named Poker1 and Poker2 are shown in Fig.7.

The result of the matching is shown in Fig.8. It can be seen that SIFT eigenvectors are invariant to image translation, scaling and rotation, but there are a few mismatches.

B. RANSAC Algorithm to Eliminate Mismatches

Simple registration can produce some mismatches, use random sample consensus (RANSAC) algorithm to solve this problem. Use iterative methods to randomly extract several structural models from a set of data and then determine whether other data are suitable for the model. If the data volume is larger than the initial one, continue using these data as the basis for the construction of a new model, and then repeat the above steps until the proportion of data reaches the default standard.

Data have valid data and invalid data. Valid data refers to less offset data, and invalid data refers to more offset data. If the proportion of valid data is larger than the proportion of invalid data, the parameters and errors of the model can be finally determined by least-squares method; Otherwise, a new algorithm is needed to replace the failed least-squares method.

RANSAC has three basic assumptions: The ingroup data can be modelling, the data that satisfy the model can also be ingroup data, while the outlier data is not suitable for modeling; Data may be corrupted by various outliers; For a set of given ingroup data, there is a function that can calculate and estimate to get the best interpretation or the most suitable parameters for this data model.

Based on above assumptions, RANSAC algorithm process is roughly as the seven steps:

(1) Select a few matching points randomly and set them as ingroup, we select 4 points in this paper;

(2) Estimate the model that fits the ingroup;

(3) Put all the unselected points into the model established before to test whether it can be classified to the ingroup;

(4) Note the amount of data of the ingroup;

(5) Repeat above steps several times;

(6) Compare the calculation and the built model with the largest number of ingroup data is the solution;

(7) Keep the matching point pairs in the solution group, and delete the other points.

The result after RANSAC mismatch elimination is shown in Fig.9. The mismatches is basically eliminated.

C. RANSAC Algorithm Applied to SIFT Feature Matches

Enter two color images with size 320x240, the original image shown in Fig.7. Compare two images, you can see the size changes, perspective changes, rotation changes. SIFT feature points distribution of original images has shown in Fig.6. The center of circle is the feature point, the radius is the gradient value, and the direction of the radius represents the gradient direction. As a key point retain more than 80% directions of main peak value, there are two gradient directions of some of key points in the graph. Since there are many key points, information like direction and the modulo values can be seen by using above expressions. In order to describe the position of the key points clearly, the key point diagram including the coordinate information is shown in Fig.10. Table 1 shows the feature SIFT related data.

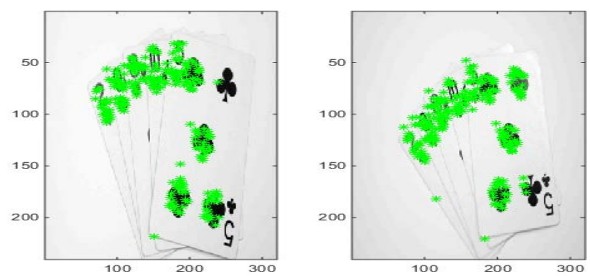
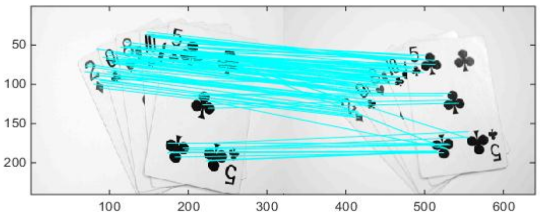


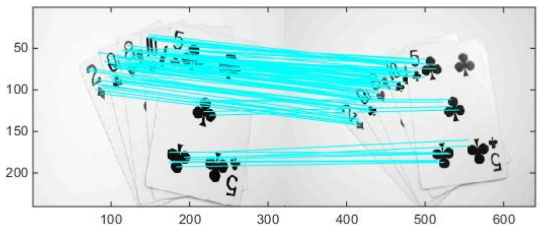
Figure 10. Key points coordinate diagram.

TABLE I. SIFT FEATURE POINTS DATA

Image	**Octave**	**Extreme point**	**Remove low contrast and edge response**	**Direction**	**Total**
Poker1	-1	462	352	123	488
	0	132	82	83	676
	1	44	43	34	792
	2	3	3	3	807
Poker2	-1	423	405	123	509
	0	112	105	73	727
	1	37	35	31	810
	2	5	5	4	821

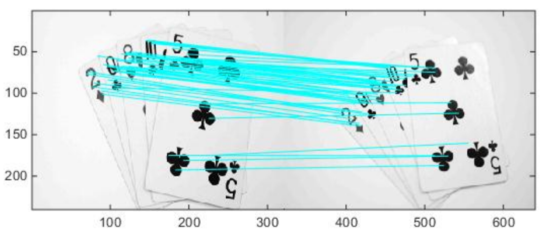


(a)

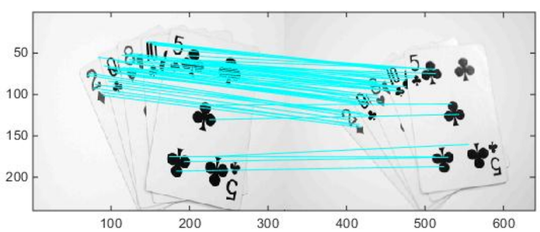


(b)

Figure 11. SIFT matching diagram: (a) Before RANSAC; (b) After RANSAC (ratio threshold is 0.7).

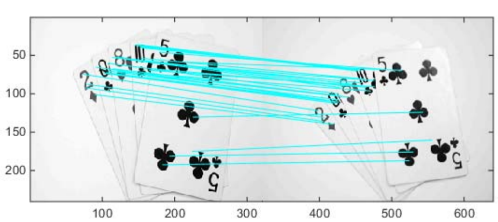


(a)

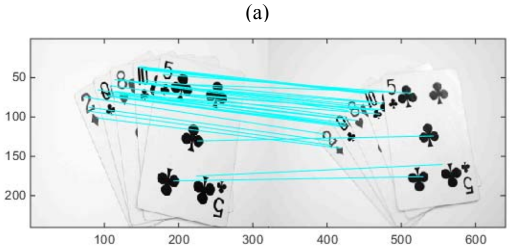


(b)

Figure 12. SIFT matching diagram: (a) Before RANSAC; (b) After RANSAC (ratio threshold is 0.6).

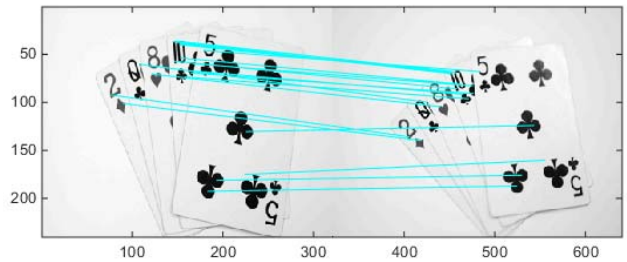


(a)

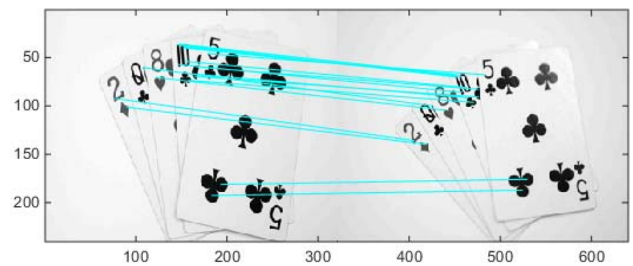


(b)

Figure 13. SIFT matching diagram: (a) Before RANSAC; (b) After RANSAC (ratio threshold is 0.55).



(a)



(b)

Figure 14. SIFT matching diagram: (a) Before RANSAC; (b) After RANSAC (ratio threshold is 0.5).

TABLE II. RELATED MATCHING DATA

Item	**Matching Data**				
Matched points	0.5	0.55	0.6	0.7	0.8
Before ransac	25	44	76	149	237
After ransac	23	43	75	142	211

In order to study the setting of the proportional threshold for matching, make research on the matching results of two images with different thresholds, it is shown in Fig.11-14 and Fig.8-9. Image matching data is shown in Tab.2. Compare Fig.13 (a) and (b), a small amount of mismatches exists before the mismatch elimination, and the wrong match points are deleted after RANSAC. Compare Fig.13-15, the matching point pairs decrease as the proportion threshold decreases, meanwhile the probability of mismatching decreases. Therefore, it is moderate to set the proportion threshold 0.7. Running mismatch results can have various number of matching pairs as RANSAC makes random sampling and generates templates, according to which to get the best template. It can also be seen from the matching results that the SIFT feature invariants are invariant to the size, translation, light, rotation, and so on.

To verify the application of SIFT in image recognition, choose two ship images that is shown in Fig.15.

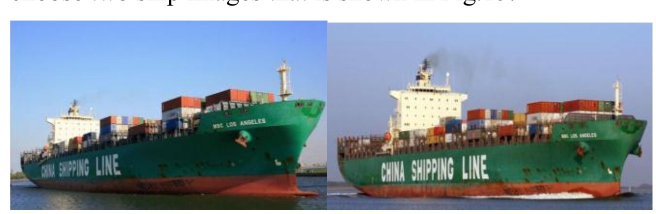


Figure 15. Original images: (a) Image Ship1; (b).Image Ship2.

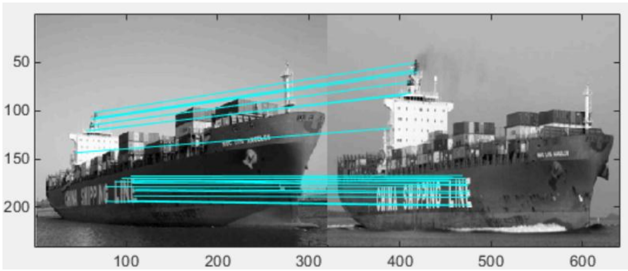


Figure 16. Ship image recognition results diagram.

The matching result of two images is shown Fig.16. SIFT features are invariant to the viewing angle and rotation. Through the matching of feature points, recognition can be well achieved and the method can be well used in ship inspection.

IV. CONCLUSION

Image matching is widely used in target recognition, intelligent identification and other fields and has always played a very important role. In this paper, the SIFT algorithm is analyzed in great detail, and the SIFT features are used to match the two images. The image matching based on the SIFT feature and the mismatch elimination of the image is studied. Then the influence of the ratio of the nearest distance and the second nearest distance for matching is discussed, and the effect is better when the threshold ratio is 0.6. Meantime, the two ships image is used to verify the application of image matching in ship detection.

Although SIFT has a very good invariance on rotation, scale scaling, brightness and a certain degree of stability on changes in perspective, affine transformation, noise, it is local feature points and weak performance for objects in complex background. SIFT needs to be improved at this point. The image similarity measure in this paper only uses one of the most commonly used Euclidean distances. Many measurements of similarity for SIFT features will be studied in the next step.

ACKNOWLEDGMENT

This work was supported by National Science Foundation of China (NSFC) under Grand 51479159 and The Excellent Dissertation Cultivation Funds of Wuhan University of Technology under Grand 2016-YS-063.

REFERENCES

[1] CHEN Yong, SHANG Lei. Improved SIFT image registration algorithm on characteristic statistical distributions and consistency constraint[J]. Optik-International Journal for Light and Electron Optics,2016,127(2).

[2] XI Chao-jian, GUO San-xue. Image Target Identification of UAV Based on SIFT[J]. Procedia Engineering,2011,15:3205-3209.

[3] SUN Tongfeng, DING Shifei, REN Zihui Ren. Novel Image Recognition Based on Subspace and SIFT[J]. Journal of Software,2013,8(5):1109-1116.

[4] DONG Yanli, CUI Yan. An approach to the partial shoeprints image matching based on SIFT and RANSAC[J]. Journal of Henan University of Engineering(Natural Science Edition),2017,29(1):71-75.

[5] ZENG Luan, WANG Yuan-qin, TAN Jiu-bin. Improved algorithm for SIFT feature extraction and matching[J]. Optics and Precision Engineering,2011,(6):1391-1397.

[6] CHEN Shu; WANG Lei. An improved SIFT feature matching based on RANSAC algorithm[J]. Information Technology,2016,40(12):39-43.

[7] ZHAO Ye; JIANG Jianguo; HONG Richang. An Optimized SIFT Matching Based on RANSAC[J]. Opto-Electronic Engineering,2014,(8):58-65.

[8] MA Li-li; CAO Chun-mei; CHEN Jin-guang; WANG Xing-hui. Feature point matching algorithm based on RANSAC[J]. Computer Engineering and Design,2016,(7):1794-1797,1856.

[9] YANG Yu-wei1 , ZHANG Ya-ping. An improved SIFT image detection and feature matching algorithm[J]. Journal of Yunnan University（Natural Sciences）,2017,39(3):376-384.

[10] Xiangqian Wu,Qiushi Zhao,Wei Bu.A SIFT-based contactless palmprint verification approach using iterative RANSAC and local palmprint descriptors[J]. Pattern Recognition,2014.

[11] J. Novatnack,K. Nishino.Scale-dependent 3D geometric features[C]. Proceedings of the 11th IEEE International Conference on Computer Vision,2007.