∎ 336

Image Mosaic Method Based on Gaussian Second-order Difference Feature Operator

Yong Chen*, Yu-bin Hao, Di Zhan
Key Laboratory of Industrial Internet of Things& Network Control, MOE,
Chongqing University of Posts and Telecommunications, Chongqing, China, 400065
*Corresponding author, e-mail: chenyong@cqupt.edu.cn

Abstract

To compose the wide visual angle and high resolution image from the sequence of images which have overlapping region in the same scene quickly and correctly, an improved SIFT algorithm which is based on D2oG interest point detector was proposed. It extracted the image feature points and generated corresponding feature descriptors by improved SIFT algorithm. Then, using the random consistency (RANSAC) algorithm purified feature point matching pairs and calculating the transformation matrix H. Last, complete the seamless mosaic of images by using the image fusion algorithm of slipping into and out. It respectively process the images which had the four typical transformations with the traditional SIFT and the proposed method. The result indicated that the number of feature pairs is fewer than SIFT algorithm and the mosaic time is shorter, and then the matching efficiency is higher than the later. This proposed method reduces the complexity of operation and improves real-time of image mosaic simultaneously.

Keywords: sequence image, image matching, SIFT algorithm, first-order difference pyramid, second-order difference pyramid

1. Introduction

The perspective of a single image obtained by image acquisition device is limited. It is needed to compose the wide visual angle and high resolution image from the sequence of images which have overlapping region in some applications [1]. With the development of computer and image processing technology, image stitching technology is widely used in space exploration, remote sensing image processing, medical image analysis and video retrieval, etc. Image registration is the key technology of image mosaic, the method main includes area-based approach and feature-based approach. In the area-based approach, there is a classic algorithm which is called registration algorithm based on a template. The method of based on gray correlation is proposed by literature [2], however, due to the limitations of time-consuming affected, and the algorithm's practical is constrained, this method's errors is large. Feature-based approach is matched by the characteristics which derived from the pixel values. Such as SUSAN corner detection algorithm which doesn't need gradient operator, so the efficiency of algorithm is raised. And this algorithm has integral characteristics which make it have been greatly improved in terms of computing speed and noise immunity. However, an excessive number of corners of this algorithm is not conducive to corner matching. The method of Moravec corner detection is proposed by literature [3]. This method's calculation is small but sensitive to the influence of the noise. When Harris operator is used in filtering and the first order difference of gray scale, the corner features extracted by it are uniform and reasonable. Although, the algorithm can detect corners in a single scale, the positioning accuracy of detection is poor [4]. According to different scenarios, Mikolajczyk etc take a test for a variety of the most representative descriptors [5]. The SIFT descriptor performance is good, but the complexity of the algorithm is high, and the calculation of the algorithm is large.

To compose the wide visual angle and high resolution image from the sequence of images which have overlapping region in the same scene quickly and correctly, an improved SIFT algorithm which is based on D2oG interest point detector was proposed. It extracted the image feature points and generated corresponding feature descriptors by improved SIFT algorithm [6]. Then, using the random consistency (RANSAC) algorithm purified feature point

matching pairs and calculating the transformation matrix **H**. The result indicated that the number of feature pairs is fewer than SIFT algorithm and the mosaic time is shorter, and then the matching efficiency is higher than the later.

And the rest part of this is organized as follows, part 2 introduces the principle of Multi-scale feature extraction, and in section 3 we give a very clear elaboration of the algorithm which we proposed. In part 4, we evaluated the effect of different algorithms and come to a conclusion.

2. Multi-scale Feature Extraction

The feature point is an area which is different in color and gray level with the surrounding. SIFT feature point is to find the extreme point in the spatial scale, and extracts its position, scale and rotation invariant [7].

SIFT feature extraction is mainly divided into three steps:
(1) Build scale-invariant space;
(2) Feature point positioning;
(3) Generating feature point descriptor [8].

SIFT feature point location is shown in Figure 1. In DoG space.Then take a further description of the feature points and form a 128-dimensional feature vector. Finally, take subsequent matching to the feature descriptor and get the features corresponding matching point pair of the image [9-10].

In three steps of SIFT feature extraction, there are more layers of LoG and DoG pyramids need to be constructed. And there is a need to use the local information of the three layers image in DoG Pyramid to locate the feature points. Therefore, feature detection occupies about 80% of the time of SIFT algorithm [11]. And splicing takes too much time, it is difficult to realize real-time performance. This paper reduced the computational complexity by simplifying the pyramid structure and changed the method of feature point positioning. Thus proposed SIFT feature extraction algorithm is based on D2oG.

3. The Principle of SIFT Feature Extraction with D2oG
3.1. D2oG Feature Detection Operator

In this paper, to get D2oG function by taking a differential operation to DoG functions is shown in Equation (1). And replace extreme point differential of DoG function with zero-crossing detection of D2oG function. Then determine the characteristic scale and detect feature points.

$$D^2(x,y,\sigma) = D(x,y,k\sigma) - D(x,y,\sigma)$$
(1)

Where $D^2(x,y,\sigma)$ is D2oG function, $D(x,y,\sigma)$ is DoG function, (x,y) is spatial coordinates, σ is scale coordinates, k is scale factor as Equation (2) shown.

$$\frac{\partial D}{\partial \sigma} \approx \frac{D(x,y,k\sigma) - D(x,y,\sigma)}{k\sigma - \sigma} = \frac{D^2(x,y,\sigma)}{k\sigma - \sigma}$$
(2)

$$D^2(x,y,\sigma) = D(x,y,k\sigma) - D(x,y,\sigma) \approx \frac{\partial D}{\partial \sigma}(k\sigma - \sigma)$$
(3)

According to Equation (3), where:

$$k\sigma - \sigma \neq 0 \text{ ,}$$
$$D^2(x,y,\sigma) = 0 \Leftrightarrow \frac{\partial D}{\partial \sigma} = 0$$
(4)

By shows of the derivation of Equation (2), (3) and (4) that the zero crossing of D2oG function is the zero-point of the first derivative of DoG function, which is the local extreme points

of DoG function. So the zero detection of D2oG pyramid can be used to replace local extreme point detection of DoG pyramid to achieve feature point extraction.

3.2. D2oG Feature Detection Steps

(1) Establish D2oG pyramid

First, get the Gauss image at different scales by Gauss kernel function for image convolution.

Then, get the DoG pyramid by doing subtraction between adjacent two layers within the same order of Gaussian pyramid. Finally, get the first layer of D2og pyramid by doing subtraction between adjacent two layers within the same order of Dog pyramid. In the establishment of the first and second order Gaussian second-order are difference pyramids for such as shown in Figure 2. There are 5 layers of Gaussian pyramid which have been established. Get Gaussian difference pyramid by doing subtraction between adjacent two layers of Gaussian pyramid. And finally, get 3 layers Gaussian second-order difference pyramids by doing subtraction between adjacent two layers of 4 layers Gaussian difference pyramid.

(2) Detection zero-point of each layer of D2oG pyramid.

By setting the zero point detection threshold T, detecting pixels which Gaussian second-order difference absolute value close to zero in each layer of D2oG pyramid. Pixels is considered as feature points when their Gaussian second-order difference absolute values are less than or equal to T. And then record the location and scale of the point (x_i, y_i, σ_i).Through experimental statistics, comprehensive consideration of both accuracy and speed of the algorithm, this paper set the threshold T =0.000099.

(3) Precise positioning feature points.

Mapping the feature points of D2oG space back to DoG space, that is: the feature point (x_m, y_m, σ_m) locate in order i layer j of D2oG pyramid is corresponding to the pixel (x_m, y_m, σ_m) locate in order 1 layer 2 of D2oG pyramid which has the same parameter with the feature point. So the precise positioning of feature points in DoG space can be transformed to the precise positioning of the in DoG pyramid.

(4) Extract edge feature points. the feature points in the edge are determined by the characteristics of the the Hessian matrix's Eigen values ,which are proportional to the main curvature values of the Gaussian difference function D. The Hessian matrix is as shown in Equation (5).

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{5}$$

Function on the x direction. Take derivative on the x and y direction and take second derivative on where D_{xx}, D_{xy}, D_{yy} respectively the second derivative of Gaussian difference y direction.

$$Tr(H) = D_{xx} + D_{yy} = \frac{\partial^2 D}{\partial x^2} + \frac{\partial^2 D}{\partial y^2} = \alpha + \beta \tag{6}$$

And,

$$Det(H) = D_{xx}D_{yy} - D_{xy}^2 = \frac{\partial^2 D}{\partial x^2} \times \frac{\partial^2 D}{\partial y^2} - \left(\frac{\partial^2 D}{\partial x \partial y}\right)^2 = \alpha\beta \tag{7}$$

Then,

$$\frac{Tr^2(H)}{Det(H)} = \frac{(\alpha + \beta)^2}{(\alpha\beta)} = \frac{(k+1)^2}{k} \tag{8}$$

Where $Tr(H)$ is the trace of the H matrix, $Det(H)$ is the determinant of H matrix, α is the larger eigen value, β is the smaller eigen value.

$$\alpha = k\beta$$

To determine whether the point is the edge point by setting the size of *k*. If Equation (8) is bigger than this threshold, then this point is extracted as the edge point. In this paper, $k = 10$.
Matching accuracy will drop because the dimension of feature descriptor is too low. And the time-consuming of matching calculation will increase because the dimension is too high. The original 128-dimensional descriptor is a compromise value. In the case of less accuracy and less time-consuming, the dimension of descriptor can be reduced to increase the of the speed of stitching. After extracting the feature points by the proposed method, 128-dimensional feature descriptors are still used for its expression and subsequent registration operation.

4. Feature Points Matching and Fusion
4.1. Determining of the Transformation Matrix H

$$\begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} h_0 & h_3 & h_6 \\ h_1 & h_4 & h_7 \\ h_2 & h_5 & h_8 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$
(9)

Where x_i', y_i' are the coordinates of the reference image, x_i, y_i are the coordinates of the image which is to be spliced. $h_0 \cdots h_8$ are parameters of perspective matrix.

There are eight independent linear equations can be obtained by four pairs of matching points. And the perspective transformation matrix from Equation(9) can be determined by $h_0 \cdots h_8$ which are obtained by equation. In order to avoid the equation has no solution because of the four pairs of randomly selected feature points which constitute the equation located in the same plane, it is necessary to take the method of solving initial value of model transformation parameters by selecting four pairs of feature points randomly. After the initial model parameters are calculated, it should be used to check other matching point. And then all the matching pairs which meet the modeling a certain tolerance can be obtained. Finally, to obtain parameters of the image transformation model by using these matched pair.

4.2. The Extraction of Feature Matching Points
Using the 128-dimensional feature vector of feature descriptor in the image. Matching the SIFT detector by the Euclidean distance similarity judgment method as shown in Equation (10) [12-13]. And finding out the minimum distance $D_{\min}$ and the second smallest distance D_{scn} by Equation (10). Then calculating the ratio $R = D_{\min}/D_{scn}$. The corresponding feature points will be the matching point when R is less than a threshold value. The threshold is set to 0.75 by comparison. Find the nearest neighbor and next nearest neighbor by the Best-Bin-First (BBF) algorithm.

$$D = \sum_{j=1}^{128} (X_{2i}[j] - X_{1i}[j])^2$$
(10)

Where $X_{2i}[j]$ is the eigenvectors of feature points in the reference image. $X_{1i}[j]$ is the i-th feature point vectors in image which is to be registered. D is the Euclidean distance. Excluding false matching points by the RANSAC algorithm [14] and calculating the transformation matrix H. Steps are as follows
To calculate matrix H by selecting four groups of matching points randomly. And then to calculate the distance d_i of all the remaining matching points. The matching points are interior point of H if the value of d_i is in the scope of the error threshold.

(1) To recalculate H by the set of points which include the largest number of interior point. To minimize errors by using the least squares method. Then to calculate the average error of a new set of points after removing a few "outside points".

(2) Repeat the above two steps until achieve satisfied effect. To calculate the final homo graph H with the point set which has the smallest average error. The transform relation of Image I_1 and I_2 can be expressed as:

$$I_2 = H \times I_1 \tag{11}$$

4.3. Image Fusion

To realize the seamless Mosaic with Gradual fade out fusion method [15] after registration. The pixel values of the non-overlapping area keep R, G, B value of their pixels remains the same in the two images. And the pixel values of the overlapping area to obtain R, G, B value of the new pixels values with the weight value. As shown in Equation (12).

$$I(x,y) = d_1 I_1(x,y) + d_2 I_2(x,y) \tag{12}$$

Where $I_1(x,y)$ $I_2(x,y)$ are the fused image, the reference image and the image stitching. d_1 d_2 are weight values. (x,y) is the pixel coordinate.

Assuming that the abscissa of current pixel is, x_i and the abscissa of left and right boundary of the overlap region are x_l, x_r .So d_1, d_2 are calculated as followed:

$$d_1 = \frac{x_r - x_i}{x_r - x_l}, d_2 = 1 - d_1 = \frac{x_i - x_l}{x_r - x_l} \tag{13}$$

5. Results and Discussion

The lat form for experiments are VS2010 and Open CV. The size of image is 340 280.In this paper, we choose the number of image matching points, splicing time, matching accuracy and matching efficiency as the evaluation. And the matching efficiency [16-17] is shown in Equation (14).

$$\eta(Matching\ efficiency) = \frac{\gamma(Matching\ accuracy)}{t(Computation\ time)} \tag{14}$$

Where,

$$\gamma = \frac{Consistent\ focus\ on\ match\ point\ pairs}{All\ of\ the\ matching\ point\ pairs}$$

Our experiment chooses four kinds of typical test chart for testing. Figure 3(a) is the image of the vertical translation. Figure 3(b) is an image with a substantial change in the brightness and contrast. Figure 3(c) is the image which camera viewing angle changes 40 degree. Figure 3(d) is the image which is rotated by 45 degree and reduced by half. Especially, the images on the left of the 4 groups are the reference images and on the right of the 4 groups are the stitching images. Choosing the module-based image mosaic and image mosaic based on SIFT feature points compare with the image mosaic method based on zero detection we proposed.
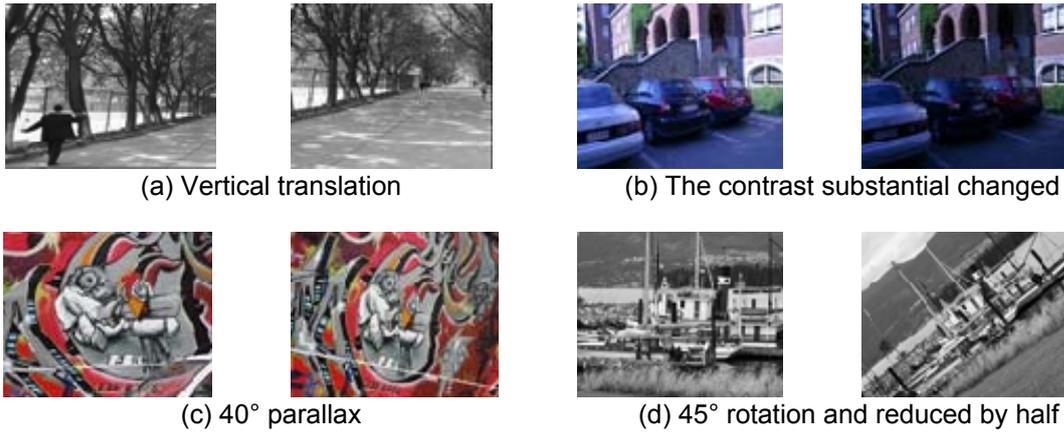
(a) Vertical translation

(b) The contrast substantial changed

(c) 40° parallax

(d) 45° rotation and reduced by half

Figure 3. Test chart

5.1. The Experimental Results
(1) Module-based image mosaic

The rendering after stitching of 3(a) which is processed by module-based image mosaic [18-20] as shown in Figure 4.

Figure 4. The rendering of module-based

The phenomenon of dislocation of template-based registration method can be seen form Figure 4. It means that the image can't find the best match position well when it is in rotation.

(2) The SIFT algorithm

(a) Vertical translation

(b) The contrast substantial changed

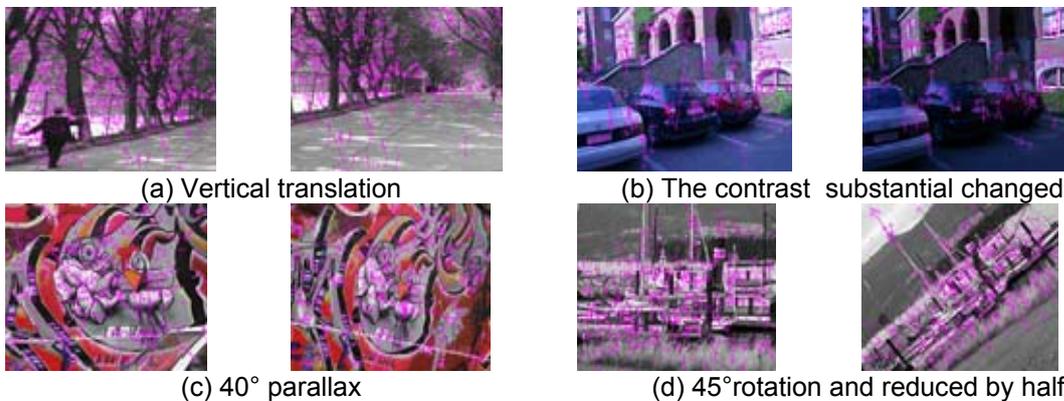(c) 40° parallax

(d) 45°rotation and reduced by half

Figure 5. SIFT feature descriptor

We use the SIFT algorithm to do the feature extraction of the four photos in the Figure 3. And to determine the number of SIFT feature points in Figure 5(a) are 1376 and 1071 respectively. The number of SIFT feature points in Figure 5(b) are 350 and 246. The number of SIFT feature points in Figure 5(c) are 698 and 839. The number of SIFT feature points in Figure 5(d) are 773 and 607.

Then we calculate Figure 5 feature points by kd-tree algorithm to obtain crude matching point pairs: 566 pairs (Figure 6(a)), 121 pairs (Figure 6(b)), 101 pairs (Figure 6(c)), 230pairs (Figure 6(d)).

| (a) Vertical translation | (b) Changes of brightness and contrast | (c) 40° parallax | (d) 45° rotation and reduced by half |

Figure 6. Feature points pairs

To eliminate false matching points by RANSAC algorithm and calculate the perspective transformation matrix H. To transform the splicing image into the coordinate system of the reference image. And then get the mosaic image by image fusion as shown in Figure 7.

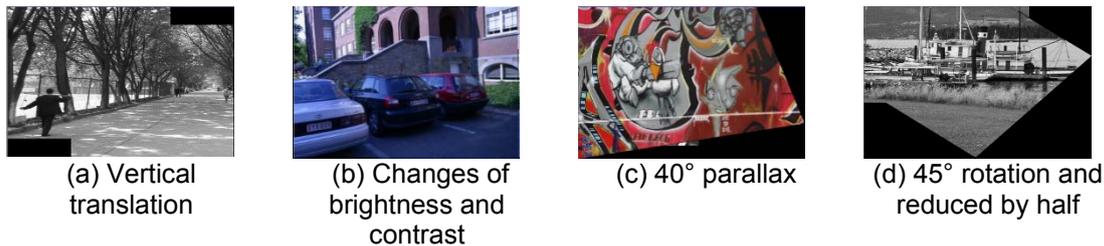| (a) Vertical translation | (b) Changes of brightness and contrast | (c) 40° parallax | (d) 45° rotation and reduced by half |

Figure 7. Image of SIFT fusion method

The above calculation shows that the number of feature points in Figure 5, Figure 6 detection are large. However, we need only 4 pairs of feature points pair to calculate the transmission transformation matrix. It will lead to long computation time and memory consumption.

(3) SIFT algorithm of D2oG zero detection

We use the SIFT algorithm of D2oG zero detection to extract image feature points, image registration and image fusion. Then obtain the crude match points after coarse match: 199 pairs (Figure 8(a)), 15 pairs (Figure 8(b)), 18 pairs (Figure 8(c)), 18pairs (Figure 8(d))

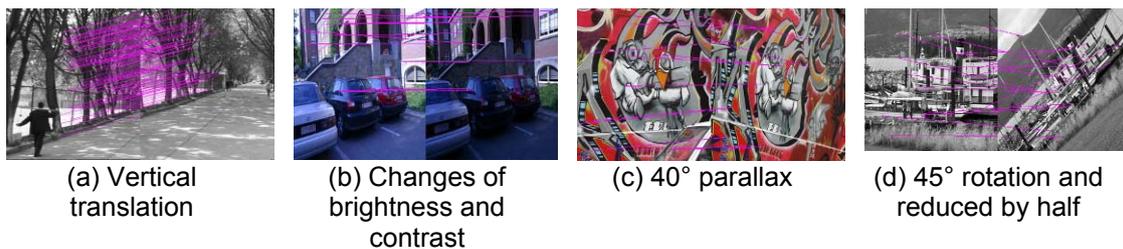| (a) Vertical translation | (b) Changes of brightness and contrast | (c) 40° parallax | (d) 45° rotation and reduced by half |

Figure 8. Feature points pairs

To eliminate false matching points by RANSAC algorithm and calculate the parameter of coordinate transformation matrix H of stitching image in Figure 3 by Equation (9).

$$H = \begin{bmatrix} -0.99907 & 0.000040 & 132.001421 \\ 0.000040 & 132.001421 & -0.000017 \\ 132.001421 & -0.000017 & -0.999958 \end{bmatrix}$$

$$H = \begin{bmatrix} 0.996046 & -0.017415 & -1.302306 \\ -0.017415 & -1.302306 & 0.000472 \\ -1.302306 & 0.000472 & 0.991628 \end{bmatrix}$$

$$H = \begin{bmatrix} 1.051323 & -0.858201 & 108.183310 \\ -0.858201 & 108.183310 & 0.871204 \\ 108.183310 & 0.871204 & 1.086682 \end{bmatrix}$$

After the parameter of transformation matrix H, r is determined, we can get the transmission diagram of Figure 3.

(a) Vertical translation   (b) Changes of brightness and contrast   (c) 40° parallax   (d) 45° rotation and reduced by half
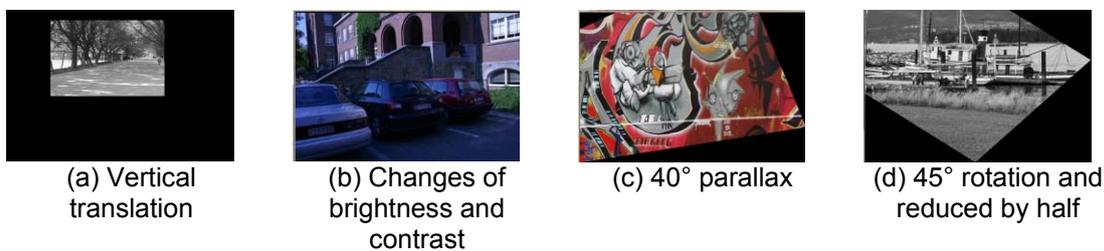
Figure 9. Transmission diagram

To fuse the transmission diagram shown in Figure 9 and the reference image shown in Figure 3 by the gradual fade out method, we can get the fused image as shown in Figure 10.

(a) Vertical translation   (b) Changes of brightness and contrast   (c) 40° parallax   (d) 45°rotation and reduced by half
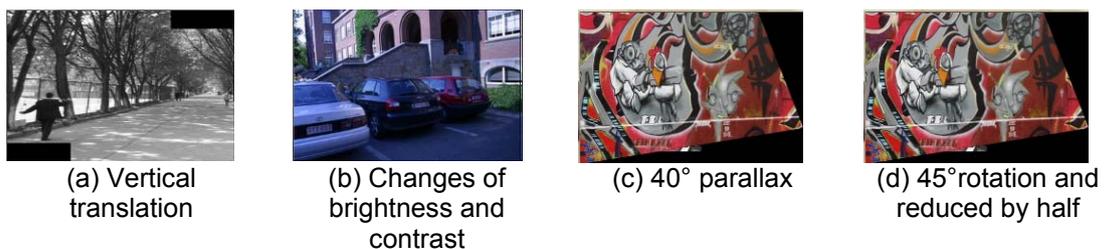
Figure 10. The D2oG method image

To verify the effectiveness of the proposed method, this article will analyse and discussion both subjective and objective evaluation.

5.2. Analysis and Discussion
It can be seen from Figure 10 that the mosaic effect of the method we proposed is a superior module-based registration method used in Figure 4. It better solves the problem of image dislocation. The module-based mosaic method has great limitations when deal with the rotation scaled image. So the method based on feature points is more practical. By observing Figure 7 and 10, we can find that the difference of splicing effect between the method which is based on the SIFT algorithm and the proposed method is not obvious. To verify the effect of the proposed method, we compare the two methods by the following objective evaluation. Feature

points extraction, matching accuracy and efficiency of image stitching is the key to real-time of splicing. We use the number of all the matching points pairs (ie coarse matching points), the number of consistent focus matching points, stitching time, matching accuracy and matching efficiency as the evaluation to compare the two methods.

Table 1 shows the experimental calculation data of splicing. It can be seen from Table 1 that the number of match points extracted by the proposed method, stitching time is significantly lower than the data obtained by SIFT algorithm while matching accuracy, efficiency is higher than the latter. It proves the feasibility and effectiveness of the proposed method is better than SIFT algorithm.

Table 1

Stitching Images Evaluation Index	Fig3(a) SIFT	Our Method	Fig3(b) SIFT	Our Method	Fig3(c) SIFT	Our Method	Fig3(d) SIFT	OurMethod
Rough Matching	566	199	121	15	101	18	230	18
Fine Matching	536	196	96	12	65	15	199	14
Matching Accuracy（γ/%）	93.9%	99.0%	79.4%	93.3%	77.5%	65.4%	85.8%	77.8%
Splicing Time(t/s)	7.003	3.422	2.057	1.384	4.572	2.431	4.298	2.554

5. Conclusion

Based on study of the SIFT feature matching algorithm, the improved SIFT features stitching algorithm based on D2oG feature detection operator is proposed in this paper. The improved algorithm improve the speed by simplifying the structure of pyramid which means reduce the number of layers of Gaussian pyramid and replacing the extremum of pixels of the three-dimensional plane in DoG space with the zero pixels in a single layer of a two-dimensional plane of D2oG space. And the D2oG pyramid still retains the DoG pyramid effective layer during construction which is to ensure the effectiveness of image stitching. Experimental shows that the proposed method has a better effect of stitching to color, grayscale image. It improves the effectiveness of stitching. And it has a better effect of stitching to image with translation, small angle differences and small deformation. So it has a certain reference value.

References

[1] Lin Jinmei, Zhou Fugen, Jin Ting. Stabilization algorithm based on SIFT feature for dynamic airborne video. *Infrared and Laster Engineering.* 2011; 40(12): 2552-2557.
[2] Yang Zhanlong. Research on image registration and mosaic based on feature point. Shanxi: Graduate XIDIAN.
[3] Zhang Yujie. Study on the Detecting and Matching Technique of Local Invariant Feature and Its Applications. Jiangsu: Graduate Nanjing University of Science and Technology. 2010.
[4] Wang Wei, Tang Yiping, Reng Juali. A improved algorithm for Harris corner detection. *Optics and Precision Engineering.* 2008; 16(10): 1995-2001.
[5] JIANG DG, YI JK. *Comparison and study of classic feature point detection algorithm.* Computer Science & Service System. Nanjing, P.R. China. 2012: 2307-2309.
[6] JB Ryu, HH Park. Log-log scaled Harris corner detector. *Electronics Letters.* 2010; 46(24): 1602-1604.
[7] LOWE DG. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision.* 2004; 60(20): 91-110.
[8] Wang Mei, Tu Daxiong, Zhou Xuchao. Moving object detection by combing SIFT and differential multiplication. *Optics and Precision Engineering.* 2011; 19(4): 1391-1397.
[9] Zeng Luan, Wang Yuanqin, TAN Jiubin. Improved algorithm for SIFT feature extraction and matching. *Optics and Precision Engineering.* 2011; 19(6): 1391-1397.
[10] Liu Xiangzeng, Tian Zheng, Shi Zhenguang, CHEN ZHSH. SAR image multi-scale registration based on KFICA-SIFT features. *Optics and Precision Engineering.* 2011; 19(9): 2186-2195.

[11] ZHU ZHW, SHEN ZHF, LUO JCH. Parallel remote sensing image registration based on improved SIFT point feature. *Journal of Remote Sensing*. 2011; 15(5): 1024-1031.

[12] Zhang H, Ling J, Lian Y. *Fast mosaic detection for real-time video based on template matching strategy*. Image Analysis and Signal Processing, International Conference on IEEE. 2012: 1-4.

[13] WAN LZH, CHONG-WAH NGO. Flip-Invariant SIFT for copy and object detection. *IEEE Transaction on Image Processing*. 2013; 22(3): 980-991.

[14] ZHU XF, MA CW, LIU B. Target classification using SIFT sequence scale invariants. *Systems Engineering and Electronics*. 2012; 23(5): 633-639.

[15] YANG BLST. Multifocus image fusion and restration with sparse representation. *IEEE TRANSACTIONSON INSTRUMENTATION AND MEASUREMENT.* 2010; 59(4): 884-892.

[16] Liu Li, Peng Fuyuan, Zhao Kun. Simplified SIFT algorithm for fast image matching. *Infrared and Laser Engineering.* 2008; 37(1): 181-184.

[17] Zhang Weizhong, Yang Houjun, Zhang Liyan. A subpixel matching algorithm for curve correspondences from images. *Journal of Beijing University of Posts and Telecommunications*. 2008; 31(4): 66-69.

[18] Ding Xiaoli. Study on image stitching technology. Nanjing: Graduate Southeast University. 2009.

[19] Yong Chen, Jie Xiong, Huan-lin Liu, Qiang Fan. Combine Target Extraction and Enhancement Methods to Fuse Infrared and LLL Images. *TELKOMNIKA Indonesian Journal of Telecommunication, Computing, Electronics and Control.* 2014; 12(3): 605-612.

[20] Jianhua L, Jianguo Y. Multifocus image fusion by SML in the shearlet subbands. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2014; 12(1): 618-626.