An Adaptive Harris Corner Detection Algorithm for Image Mosaic

Haixia Pan, Yanxiang Zhang, Chunlong Li, and Huafeng Wang

Colleage of Software, Beihang University, Beijing, China
{haixiapan,wanghuafeng}@buaa.edu.cn,
stdcoutzyx@163.com, lcl426@gmail.com

Abstract. Image Stitching refers to the technology fusing more than one images with overlapping part into a large field of view image. Image mosaic consists of image preprocessing, image registration and image fusion. To solve problems of serious clustering phenomenon and fewer corner points in the texture region caused by traditional Harris Corner detection algorithm, this paper proposes an improving adaptive threshold setting algorithm by calculating the second-order value of the corner response function, avoiding effects of the selection of scale factor k and threshold T on corner detection. To overcome the weakness of obvious traces in the jointing places caused by traditional weighted average method for image fusion, this paper enhances the weighted average method with trigonometric functions. Experimental results show our proposed algorithms can effectively eliminate the gap generated by image mosaic, with a better speed and precision.

Keywords: Image Stitching, Image Registration, Corner Detection, Image Fusion.

1 Introduction

Image Stitching [1] refers to the technology fusing more than one images with overlapping part into a large field of view image. Nowadays, with the popularization of the intelligent equipment, the high definition and wide-angle image is becoming increasingly urgent. Image Stitching technology, known as one of the newest achievements in the image processing, has gained lots of attentions from researchers and is developing at a high speed recent years. Image Stitching technology has played an important role in aeronautics, astronautics, geological exploration, video session, medicine and military, and also stands as a hotspot in computer visual analogue, computer effects and augmented reality research [2,3,4,5,6].

Image mosaic consists of image preprocessing, image registration and image fusion [7]. As the base of image mosaic is the preprocessing, cores are the registration and fusion, deciding the precision, speed and visual quality. The image mosaic is divided into registration based on feature and registration based on gray level, depending on the image information used in registration. The registration based on feature [8] shows stronger adaptability in gray level transformation, deformation and

S. Li et al. (Eds.): CCPR 2014, Part II, CCIS 484, pp. 53–62, 2014.

exposure discrepancy, which additionally can locate the matching positions easily and accurately. Although with highly accuracy, the one on gray level, which is also named as correlation matching algorithm, is hard to meet the demand of instantaneity because of its large computation and complexity [9]. The normal features of the registration are points, lines, close-contoured and other advanced ones such as Gaussian Sphere [10]. More attentions are attracted by registration based on feature points on account of its characteristics as being easily for fetching and less sensitive from image deformation. In the field of registration on feature points, the Harris Corner Detection algorithm [11] is famous for its high stability and robustness, but it has failed to consider the impact of the selection for the scale factor k and the threshold T of the algorithm. This paper proposes an improved Harris algorithm for this. While the traditional weighted average method is intuitionistic, fast, and less sensitive from discrepancy of exposure, shutter phenomenon [12] appears if there exists objects in the overlap region, we enhances weighted average method with trigonometric functions for image fusion, effectively reducing the obvious traces in the image splicing place.

2 Pipeline for Image Mosaic

Image Mosaic consists of image preprocessing, image registration and image fusion. For the possibility that there will be some effects from image rotation, image scaling and disparity of exposure, images are needed to be preprocessed before registration, whose keys are speed and precision. Finally images are stitched into one, in which no obvious traces are allowed for a good result. Figure 1 shows how the process flows.

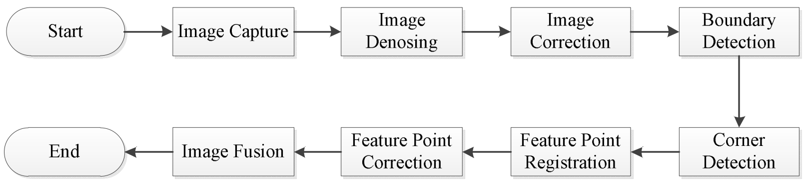

Fig. 1. Flow chart for image mosaic

2.1 Image Preprocessing

Because of the flaw on image obtaining equipment and noise from outer condition, images are interfered during the digitization and transmission, resulting in the noisy output. Median filter is often used to denoise the image. As a nonlinear smoothing technique, it can help maintain the boundary information effectively. The principle of median filter is to replace the value of specified point with the mid-value of its neighborhoods, making the surrounding pixels close to the real value, thus eliminating the isolated noise points. Median filter is calculated as formula 1.

$$g(x,y) = med\{f(x-k, y-l)\}, (k, l \in W) \tag{1}$$

In formula 1, $f(x,y)$, $g(x,y)$ stand for the original image and the processed respectively. W is the two-dimension template, usually regions of $2 \times 2, 3 \times 3$ are used, any other different shapes like threadiness, roundness, cross, annulus can also serve the same purpose.

During the generating process, because of the nonlinearity of imaging system, or change of the shooting visual angle, the images will have discrepancy with each other, mainly on the gray level and geometrical deformation. Normalization [9] is able to make the amendment. After that most discrepancy of the gray level between images can be eliminated, thus decreasing the deviation in image registration.

2.2 Feature Point Extraction

As one of the most important parts in image mosaic, the precision of image registration affects the image fusion directly. Although the one based on gray level is more intuitionistic and easy to implement, it focuses on the specified template of image for stitching, so the result will be interfered by the smooth cover caused by the similar measurement. Additionally, the registration is also affected by the scaling transformation, rotation, overlapping, especially illumination. Registration deviation emerges since the nonlinear inhomogeneous illumination. Therefore, the registration based on gray level isn't widely used in image mosaic. However, the image registration based on feature points can overcome the shortcomings of the one on gray level, making it widely used. Extraction of image feature greatly speeds up with reducing the amount of calculation and can maintain the image during the image displacement, scaling, rotation, and other transformations. Properties of unique and easily distinguish of the image features make it easy to detect the positions' changes, thus highly increasing the precision of registration. This paper mainly focuses on the image registration based on feature points. In this section traditional Harris Corner Detection Algorithm is discussed, and then comes the algorithm mentioned in this paper. Experiments and analysis of those results are given out in the next section.

Traditional Harris Corner Detection Algorithm

Harris Corner Detection [10] algorithm was proposed in 1988 by C. Harris and M. J. Stephens. The brightness variation $E(\mu, \nu)$ is the local autocorrelation function of a pixel $\mathrm{I}(\mathrm{x}, \mathrm{y})$'s local offset (μ, ν), as showed in formula 2.

$$\mathrm{E}(\mu, \nu) = \sum_{x,y} w(x, y)[I(x + \mu, y + \nu) - I(x, y)]^2 \tag{2}$$

In formula 2, the $\mathrm{I}(\mathrm{x}, \mathrm{y})$ stands for the gray level function. $[\mathrm{I}(\mathrm{x} + \mu, \mathrm{y} + \nu) - \mathrm{I}(\mathrm{x}, \mathrm{y})]$ refers to the gradient value of gray level. $\mathrm{w}(\mathrm{x}, \mathrm{y})$ is the window function, which represents the weight of each pixel. There are two different window functions, namely two-valued and Gaussian. The influence of the noise from rectangular window can be reduced by Gaussian function, whose two-dimension version is given.

$$w(x, y) = \exp\left[-(x^2 + y^2)/2\sigma^2\right] \tag{3}$$

When the partial excursion(μ, ν) is small, Second-order Taylor series expansion of $I(x + \mu, y + \nu)$ on (x, y) can be made in formula 4 after taking out the leading term.

$$\mathrm{E}(\mu, \nu) \cong [\mu\ \nu]\mathrm{M}\begin{bmatrix}\mu \\ \nu\end{bmatrix} \tag{4}$$

M is a 2×2 symmetric matrix, also called autocorrelation matrix. It can be represented by formula 5, in which $\otimes$ refers to the convolution.

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = e^{-\frac{(x^2+y^2)}{2\sigma^2}} \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \tag{5}$$

$$X = I \otimes [-1\ 0-1] = \frac{\partial I}{\partial x} = I_x;\quad Y = I \otimes [-1\ 0-1]^T = \frac{\partial I}{\partial y} = I_y \tag{6}$$

In formula 6, I_x and I_y refer to the gradient on horizon and verticality. The extremum curvature of the gray level autocorrelation function for particular pixel can approximation be represented by the eigenvalue of matrix M. Supposing two eigenvalues are λ_1 and λ_2. If both of them are large, namely both the extremum curvature of the horizontal and vertical autocorrelation function for the pixel are large, the pixel can be treated as a corner. On the other hand, if both are small, the pixel is in a planar region. If one is large while the other is small, the region is boundary. In practice, in order to avoid solving the eigenvalue directly to increase efficiency, the response function of corner can be defined as follows.

$$R = det\ M - k(traceM)^2 \tag{7}$$

In formula 7, $\det M = \lambda_1\lambda_2 = AB - C^2$, $traceM = \lambda_1 + \lambda_2 = A + B$. The scale factor k depends on experience. In most condition, it is selected from 0.04 to 0.06. If R is larger than the default threshold T, the point can be judged a corner point.

Nobel has improved the function to avoid the deviation causing by the scale factor k [13]. He defined the function R in formula 8.

$$R = \frac{det\ M}{(traceM)^2} = \frac{AB - C^2}{A+B} \tag{8}$$

Improved Harris Corner Detection Algorithm

Although Harris operator is classic, there are still some deficiencies. For example, though the scale factor k is defined in a particular range, the result of feature detection is poor when large differences exist between images. After getting the partial extremum, comparison with the threshold T and corner function R need to be made to determine the corner points, during which it's hard to select the threshold T. Too small of T leads to too many feature points and cluster phenomenon while too large of T corresponds to too few feature points, which decreases the accuracy of the results. This paper puts forward an adaptive Harris Corner Detection algorithm for that.

Because of the difficulty for the selection of threshold T, an improvement based on the Noble algorithm is made in formula 9, in which M is the autocorrelation matrix.

$$\mathrm{R}(\mu, \nu, \sigma_I, \sigma_D) = \frac{\det(M(\mu,\nu,\sigma_I,\sigma_D))}{trace(M(\mu,\nu,\sigma_I,\sigma_D))} \tag{9}$$

$$M(\mu, \nu, \sigma_I, \sigma_D) = \sigma_D^2 G(\sigma_I) \otimes \begin{bmatrix} L_\mu^2(\mu,\nu,\sigma_D) & L_\mu L_\upsilon(\mu,\nu,\sigma_D) \\ L_\mu L_\upsilon(\mu,\nu,\sigma_D) & L_\upsilon^2(\mu,\nu,\sigma_D) \end{bmatrix} \tag{10}$$

$$L(\mu, \nu, \sigma_D) = G(\sigma_I) \otimes I \tag{11}$$

In the formula above, σ_I is the integral scale factor, σ_D is the differential scale factor, $G(\sigma)$ represents the Gaussian function with variance of σ and mean value of 0, L_μ and L_υ refer to the partial derivative of direction U and direction V, det and trace are the abbreviation for the determinant and trace of matrix.

Template is calculated by mean circulation filter with the corner function R, integral scale factor σ_I and the differential scale factor σ_D, which can be used for second-order statistic to get the max value Max and the great value tmpMax. The pixel should be the maximum coordinate I_{max} when the corner function R is equal to the max value Max and Max isn't equal to the tmpMax. The threshold T can be calculated as formula 12, in which R is from formula 9.

$$T = R \cdot \max(I_{max}) \tag{12}$$

The steps for the algorithm are as follows.

1. Calculate the integral scale σ_I and the differential scale σ_D;
2. Figure out the partial derivative for each pixel in x axis and y axis;
3. Calculate the corner response function R based on formula 9;
4. Calculate partial maximum pixel I_{max} ;
5. Calculate the adaptive threshold T from formula 12.

When I_{max} is larger than the threshold T, the pixel can be treated as a corner.

2.3 Image Registration

After extracting feature points with the improved Harris Corner Detection algorithm, relationships between the points can be figured out by the registration algorithm. In this chapter, image registration based on corners points is in usage. The corner registration aims to find out the corresponding corner point pairs with one unique point in I_1 and the other unique point in I_2. In this paper, image registration for two adjacent images is accomplished by corner registration algorithm based on the SVD(singular value decomposition) in three steps [14].

First, the cross-correlation function of regions in the two images is computed. Supposing $I_1$and I_2 are the two adjacent images, after processing of corner detection algorithm, the two images will have M and N feature points respectively. For each feature point $m(u_1, v_1)$ in image I_1 , a $(2p+1)\times(2q+1)$ rectangle region A is selected as the window function in which the feature point is the center. Rectangle region B is selected the same way for each feature point $n(u_2, v_2)$ in image I_2. The cross-correlation function for A and B from the $(2p+1)\times(2q+1)$ neighborhood constituted by the corner m in image I_1 and the corner n in image I_2 is defined in formula 13.

$$c(m,n) = \sum_{i=-p}^{p}\sum_{j=-q}^{q}\frac{(I_1(u_1+i,v_1+j)-\bar{A})\times(I_2(u_2+i,v_2+j)-\bar{B})}{(2p+1)\times(2q+1)\sigma(A)\sigma(B)} \tag{13}$$

The $\overline{A}$ means the average value of region A centering m in image I_1, and the value can be calculated in formula 14. $\sigma(A)$ is the standard deviation of the $(2p+1) \times (2q+1)$ neighborhood for the region A, with the value gained by formula 15.

$$\overline{A} = \overline{I_1(u_1, v_1)} = \sum_{i=-p}^{p} \sum_{j=-q}^{q} \frac{I_1(u_1+i, v_1+j)}{[(2P+1)\times(2q+1)]} \tag{14}$$

$$\sigma(A) = \sqrt{\frac{\sum_{i=-p}^{p} \sum_{j=-q}^{q} I_1(u_1, v_1)}{(2p+1)\times(2q+1)} - \overline{I_1(u_1, v_1)}} \tag{15}$$

And the $\overline{B}$ means the average value of region B centering n in I_2, $\sigma(B)$ is the standard deviation of the $(2p+1) \times (2q+1)$ neighborhood for the region B. $\overline{B}$ and $\sigma(B)$ can be calculated in similar ways of $\overline{A}$ and $\sigma(A)$.

Formula 13 illustrates the two regions' relevancy will change from -1(means that the two regions are different) to 1(means that the two regions are the same). Then the similar matrix can be calculated:

$$G(m, n) = \frac{C(m,n)+1}{2} e^{-r(m,n)^2 / 2\sigma^2} \tag{16}$$

In which G(m, n) is the Gaussian weighted distance range from 0 to 1. And r(m, n) is the Euclidean distance from m to n. Finally, the SVD is applied to G(m, n).

$$G(m, n) = TDU^T \tag{17}$$

The T is the orthogonal matrix of M rows. The U is the orthogonal matrix of N columns. The D means the diagonal matrix of M rows and N columns.

The elements on the diagonal line of diagonal matrix D should be in descending order. The identity matrix E is constructed by setting elements which are not equal to zero to one on the diagonal line of D, so that it comes to the matrix P in formula 18.

$$P(m, n) = TEU^T \tag{18}$$

The same form of the matrix P and G makes the suited feature points stand out easily. If P(m, n) is the maximum in row as well as the maximum in column, then we consider the feature points m and n as a pair of matching points.

2.4 Image Fusion

As one of the core techniques in image mosaic, image fusion aims to eliminate the traces in image splicing place. In this part the weighted mean method will be used to image fusion [15]. Firstly weighted mean calculation based on gray level is applied to the feature points, and then final pixel value is determined by superimposing the pixels. Supposing the images to be mosaicked are f_1 and f_2, f is the result image, then weighted mean method can be achieved as formula 19.

$$f(x, y) = \begin{cases} f_1(x, y) & (x, y) \in f_1 \\ \omega_1 f_1(x, y) + \omega_2 f_2(x, y) & (x, y) \in (f_1 \cap f_2) \\ f_2(x, y) & (x, y) \in f_2 \end{cases} \tag{19}$$

The ω_1 and ω_2 are the weights of corresponding pixels in the overlay region of the two images. And they are conditioned by $0 < \omega_1, \omega_2 < 1$ as well as $\omega_1 + \omega_2 = 1$. Though this method is simple and intuitive, fast for image fusion, can easily deal with the discrepancy of exposure, shutter appears when there exist objects in the overlay region. According to the Weber's Law, responses of the HVS (Human Visual System) for stimulus signals are based on the luminance comparison between signal and background (the average of signal) rather than on the absolute luminance. The sensitivity of HVS for gray level is in direct proportion to the logarithm of practical luminance. Usually the HVS is more sensitive to the medium gray level, and the sensitivity is more stable in this region, while nonlinear declines in two directions of high and low gray level region. So the sensitivity of human eyes is not linear as the traditional weight method considers. When the discrepancy of luminance in two images is large, it is not ideal to use the factor d in formula 20 for image fusion.

$$d = (x_2 - x)/(x_2 - x_1) \tag{20}$$

In order to solve the problem mentioned above, weight factor k is displaced with trigonometric function and the segmented weighted function.

$$d = \begin{cases} 1 - 0.5\sin\left(\frac{x-x_1}{x_2-x_1}\right)\pi & x_1 \le x \le x_1 + \frac{x_2-x_1}{4} \\ \frac{(\sqrt{2}-2)(x-x_1)}{x_2-x_1} + \frac{3-\sqrt{2}}{2} & x_1 + \frac{x_2-x_1}{4} \le x \le x_1 + \frac{3(x_2-x_1)}{4} \\ 0.5\sin\left(\frac{x-x_1}{x_2-x_1}\right)\pi & x_1 + \frac{3(x_2-x_1)}{4} \le x \le x_2 \end{cases} \tag{21}$$

3 Experimental Results and Discussion

3.1 Experimental Results

In this chapter the traditional method and the improved method are compared on multiple sets of images. The comparisons focus on many aspects, including effect of image mosaic, time of algorithms and the number of suited points in image registration.

Firstly the source and result images of image mosaic are showed in figure 2. For convenience of the observation, the corresponding feature points in two images are connected by a line.

In the first set of images, the (a1, b1) is the original image. Image (c1) is the result using traditional method of image registration while image (d1) is the one using the improved method. In the second set of images, the (a2, b2) is the original image. Image (c2) is the result of traditional method and (d2) uses the improved method. (e1) and (e2) are final results of image mosaic.

Secondly, five different groups of images are tested to obtain time of algorithms, including traditional image registration method and the improved one. Time of extraction and registration is measured separately. The result is shown in table 1.

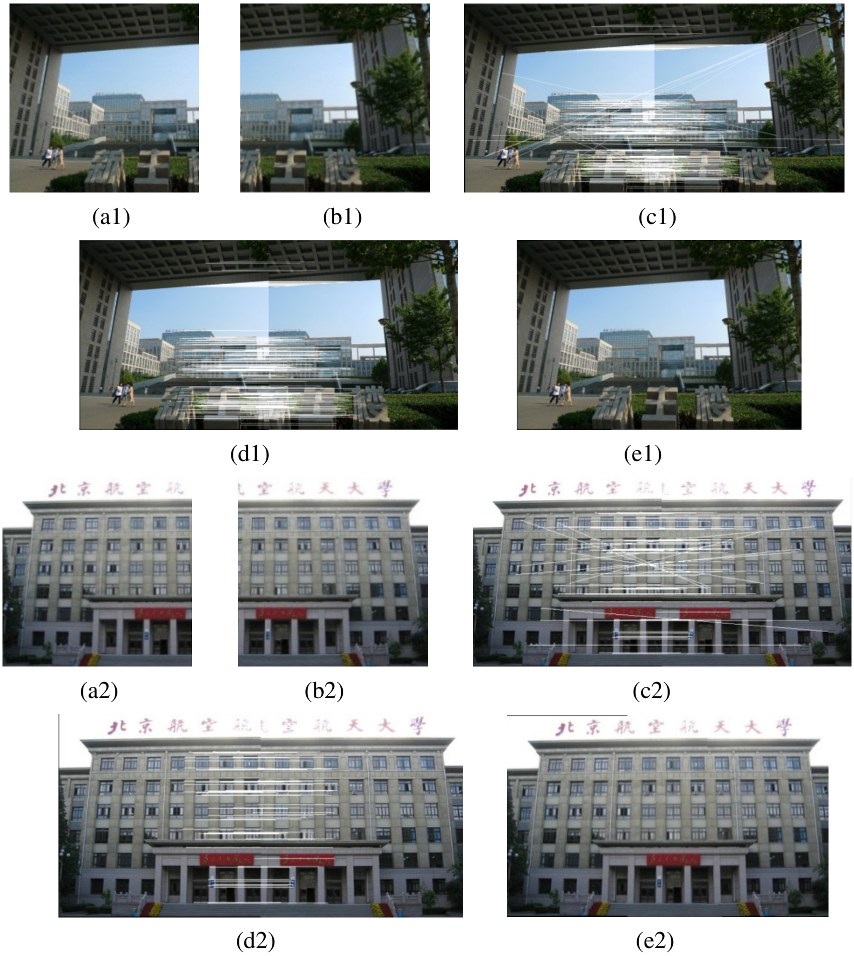

Fig. 2. Two sets of image mosaic experiment result.

Because of the different algorithms, the number of suited point pairs and missing match point pairs are different, thus raising discrepancy on the accuracy of registration. So lastly comparisons of the number of suited points and erroneous suited points and the accuracy of registration between the two algorithms are conducted. Results are shown in table 2 and table 3.

3.2 Discussion

Compare to the different results, we can find out that the improved algorithm in this paper can detect the feature points more accurately and obtaining the suited features more stably. The incorrect registration in our algorithm is less than the traditional one.

Generally, the comparisons between the traditional Harris algorithm and the improved one based on the extraction time of feature points, time for registration, number of erroneous suited points and accuracy can give us a clear look on the discrepancy of the two algorithms. And we can conclude that the improved Harris corner detection algorithm can get better result and improve the accuracy. The algorithm mentioned in our paper is effective and utility.

Table 1. Comparison of time for Extraction and registration

No.	Traditional algorithm			Improved algorithm		
	Extraction (Image1)/s	Extraction (Image2)/s	Registration /s	Extraction (Image1)/s	Extraction (Image2)/s	Registration /s
1	1.95	1.84	1.99	1.21	1.07	1.27
2	1.42	1.65	1.16	1.18	1.29	0.89
3	1.03	1.07	1.07	0.56	0.64	0.78
4	3.36	3.90	4.44	2.84	3.15	3.92
5	1.88	1.76	1.95	1.10	0.96	1.23

Table 2. Comparison on number of suited points

No.	Traditional algorithm			Improved algorithm		
	Points (Image1)	Points (Image2)	suited points	points (Image1)	points (Image2)	suited points
1	971	959	323	950	948	305
2	330	275	125	301	246	105
3	141	233	91	125	211	65
4	3015	4680	2262	2709	4299	2236
5	709	830	201	654	689	112

Table 3. Comparison on the erroneous suited points and accuracy

	Traditional algorithm		Improved algorithm	
No.	erroneous suited points	Accuracy	erroneous suited points	Accuracy
1	62	81.0%	46	85%
2	21	83.2%	14	86.7%
3	14	85.1%	6	91.1%
4	189	91.6%	156	93.0%
5	18	91.0%	6	94.4%

4 Conclusion

For serious clustering phenomenon and fewer corner points in the texture region of traditional Harris corner detection algorithm, this paper proposed an improved adaptive threshold setting algorithm by calculating the corner response function of second-order, avoiding the impact of the scale factor k and the threshold T on corner detection. For obvious traces of weighted average method in the jointing places, this paper proposed a new weighted average method based on trigonometric functions for image fusion. Experimental results show that the modified algorithm can effectively eliminate the gap generated by image mosaic, and improve stitching accuracy and speed.

References

1. Peleg, S., Herman, J.: Panoramic mosaics by manifold projection. In: Proceedings of 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 338–343. IEEE (June 1997)
2. Liu, J., Li, L.: A Brief Introduction of Reconstruction Technology of Distant Medical Consultation for Large Image(远程医疗会诊中拼接大型医学图像技术简介). China Contemporary Medicine 6(10), 62－63 (2000)
3. Nie, S.D., Si, J.Y.: Methodological Study of Automatically Mosaicing for Medical Microscopic Images(医学显微图像自动拼接的方法研究). Chinese Journal of Biomedical Engineering 24(2), 173－178 (2005)
4. Cai, Y., Hu, X.: Short wave infrared imaging technology and its defense application (短波红外成像技术及其军事应用). Infrared and Laser Engineering 35(6), 643－647 (2006)
5. Wen, H. Y.: Creating image-based VR using a self-calibration fisheye lens (遥感图像拼接算法研究). Doctoral Dissertation, 华中科技大学 (2009)
6. Xiong, Y., Turkowski, K.: Creating image-based VR using a self-calibrating fisheye lens. In: Proceedings of 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 237–243. IEEE (June 1997)
7. Wang, J., Shi, J., Wu, X.X.: Survey of image mosaics techniques. Application Research of Computers 25(7), 1940–1943 (2008)
8. Szeliski, R.: Image alignment and stitching: A tutorial. Foundations and Trends® in Computer Graphics and Vision 2(1), 1–104 (2006)
9. Li, Q., Zhang, B.: A fast matching algorithm based on image gray value. Journal of Software 17(2), 216–222 (2006)
10. Harris, C., Stephens, M.: A combined corner and edge detector. In: Alvey Vision Conference, vol. 15, p. 50 (1988)
11. Gumustekin, S., Hall, R.W.: Mosaic image generation on a flattened Gaussian sphere. In: Proceedings 3rd IEEE Workshop on Applications of Computer Vision, WACV 1996, pp. 50–55. IEEE (December 1996)
12. Burt, P.J., Adelson, E.H.: The Laplacian pyramid as a compact image code. IEEE Transactions on Communications 31(4), 532–540 (1983)
13. Noble, J.A.: Finding corners. Image and Vision Computing 6(2), 121–128 (1988)
14. Feng, Y.P., Dai, M.: An Image Mosaic Algorithm Based on Corner Features (一种基于角点特征的图像拼接融合算法). Microelectronics & Computer (7), 21－23 (2009)
15. Szeliski, R.: Video mosaics for virtual environments. IEEE Computer Graphics and Applications 16(2), 22–30 (1996)