

High-Definition Video Image Mosaics in Real-Time Based on Adaptive H-SIFT

Xiaowei Han^{1, a}, Zhiyong Wang^{2, b} and Wenjun Yang^{3, c}

¹professor, Shenyang University, Shenyang, China

²postgraduate, Shenyang University, Shenyang, China

³Postgraduate, Xi'an Jiaotong University, Xi'an, China

^ahxw69@163.com, ^bwzy.ok@qq.com, ^cywj17593@stu.xjtu.edu.cn

Keywords: video mosaic; image mosaic; real-time mosaic; adaptive; SIFT algorithm

Abstract. High definition video image mosaic technology has been widely used in various fields, with high accuracy and good real-time performance. Based on the research of video image mosaic algorithm which based on SIFT, a new H-SIFT image mosaic algorithm is proposed, which is suitable for real-time mosaic of high definition video images. On the basis of the SIFT high precision registration algorithm, The algorithm can automatically recognize the image size and process it, and the stitching speed is improved while the quality of high-definition video images is guaranteed. Experimental tests and verify the effectiveness of the H-SIFT algorithm.

Introduction

With the development of digital image processing technology, video mosaic technology has been paid more and more attention. At present, there are two main algorithms for image mosaic, SIFT and SURF. SIFT has the best effect in scale and rotation transform, SURF has the best effect in brightness changes, in the aspects of fuzzy is better than SIFT, but in the rotation and scale changes a lot worse than SIFT[1-2]. The UAV aerial image sequence needs to guarantee the rotation without deformation, so SIFT image stitching algorithm is widely used [3]. However, the stitching speed of SIFT algorithm is about 3 times slower than that of SURF algorithm[4]. An algorithm with high stitching speed, high accuracy and fast stitching speed is needed.

In conclusion, a H-SIFT image stitching algorithm for HD video scene real-time image splicing requirements is presented in this paper, the algorithm based on SIFT algorithm, with high accuracy registration, to optimize the extraction of feature points, the matching speed of video image is greatly improved, suitable for high-definition video image processing.

H-SIFT Real Time Stitching Algorithm

H-SIFT Feature Extraction

Sift operator was first proposed by David.G.Lowe in 1999. It was mainly used for object recognition at that time[5]. In 2004, David.G.Lowe made a comprehensive summary of the operator, formally proposed a SIFT operator based on scale space of the image, zoom, rotation and affine transformation invariance to keep the local image features are described, namely scale invariant feature transform [6]. Improve and optimize the Sift algorithm, so that it can satisfy the stitching of ten million pixels or more images, which is called H-SIFT algorithm.

H-SIFT algorithm feature extraction steps are as follows:

1)Generation of scale spaces

Difference of Gaussian, detection extreme point in scale space.

2)Detection scale space extreme point

Gauss convolution kernel is the unique transformation kernel for scale transformation. The scale space of an image is $I(x, y)$ defined as $L(x, y, \sigma)$, which is generated by convolution of different scales of Gauss's function $G(x, y, \sigma)$ and the original image $I(x, y)$.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y). \quad (1)$$

$$\text{among: } G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \quad (2)$$

① in the above formula, σ calls scale factor, and the smaller it means, the less smooth the image is. Large scale corresponding to the overview of the image, small scale corresponding to the details of the image.

② In order to detect extreme points efficiently in scale space, David G. Lowe introduces the extreme value detection in scale space created by DoG (Difference of Gaussian) operators. The DoG operator is defined by subtracting two Gaussian images from adjacent scales in a scale space.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma). \quad (3)$$

The K in the formula is a scaling factor.

The scale space created by the DoG operator is shown in figure 1:

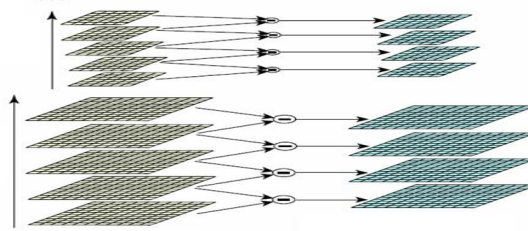


Fig. 1 Gaussian Pyramid and DoG Pyramid

3) Precise location of extreme points

The scale and position of the extreme points can be accurately located by fitting the three dimensional quadratic function to remove the low contrast and unstable edge response points. The marginal response point is removed by (4-6) formula.

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (4)$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (5)$$

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \quad (6)$$

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (7)$$

In the formula, H is the Hessian matrix, and R is the ratio of the larger and smaller values in the eigenvalues of the Hessian matrix.

4) Allocation direction of key point

In order to make the SIFT feature points have rotation invariance, the direction parameters of each key point are assigned by using the distribution characteristics of the neighborhood gradient pixels of the key points, and the formulas of the gradient and the direction of the points are as follows.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (8)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (9)$$

Where L is the criterion for each key point. In practice, sampled in neighborhood windows centered on feature points, and the gradient direction of neighborhood pixels is calculated with

histograms of gradient directions, so as to determine the dominant gradient direction of the extremum points.

5) Generating feature point descriptors

rotate axis to the main direction of key points, to ensure the rotation invariance, then with the key point as the center, take 16 pieces of uniform distribution, draw the gradient direction in 8 directions each piece on the tired value, forming a seed point information vector for each seed point has 8 directions, one feature with 16 seed points to describe, from the 128 dimension vector to describe.

In the first step of feature extraction of SIFT algorithm, the scale space is obtained by convolution of image and Gauss function. The higher the image resolution, the more pixels, the greater the amount of convolution calculation, the longer the time consuming. Because the SIFT algorithm has a high accuracy, does not need to match the feature points too can be successful for image registration, high resolution and high pixel image, extract the feature points matching is a large amount of waste of resources, waste of time. Therefore, the improved H-SIFT algorithm for feature extraction algorithm are optimized and improved, control the number of feature points extracted by the size of the image resolution, control operation times and shorten the matching time, and which has high registration rate.

H-SIFT feature point matching

The use of stitching algorithm based on feature points, this kind of method is to determine the registration parameters according to the geometric relations between the important characteristics of the image, so this kind of method we first extract features, such as edges, corners and curvature. Then, the corresponding relation between feature points is established, and the registration parameters are obtained.

The utility model has the advantages that the image feature points is much less than the pixel points of the image, thus greatly reduce the computational complexity of the matching process; feature point matching measure is sensitive, can greatly improve the accuracy of matching; feature extraction process can reduce the influence of noise, the gray change, image deformation and illumination inequality has good adaptability, so it is widely used in the field of image registration.

X and Y represent n-dimensional vectors, respectively, and represent their components, then Euclidean distances between vectors X and Y are shown in formula (10).

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \dots + (x_n - y_n)^2} \quad (10)$$

Take the SIFT feature point descriptor of Known image and Matched image, use the SIFT feature point descriptor in Matched image and each SIFT feature point descriptor in Known image to calculate Euclidean distance respectively, and find the coordinates of two characteristic points with the smallest distance. For each matching feature points in the image to find the nearest distance and close in the known image, for a feature point if its nearest distance and time distance ratio is less than a threshold ratio, respectively, to keep the feature points and feature point coordinates with its nearest distance in the corresponding image, which is a pair of matching points.

Experimental Result

To verify the performance of the algorithm, Microsoft Visual C++ 2013 is used as an experimental platform to select multiple sets of high-definition video images for image stitching verification.

The specific stitching steps are as follows:

1) the feature points of the image are extracted by the optimized H-SIFT algorithm, and the number of feature points is determined by the resolution of the image.

2) feature matching of the extracted feature points, and then estimate the geometric deformation parameters of the two images as well as the overlap region between the images.

3) through the RANSAC algorithm to remove mismatches, and then two images fusion, and image smoothing.

4) after two images are fused, edges are cut and output images are pieced together.

Figure 3 is a typical aerial image mosaic result of an unmanned aerial vehicle, and the results are shown in table 1.

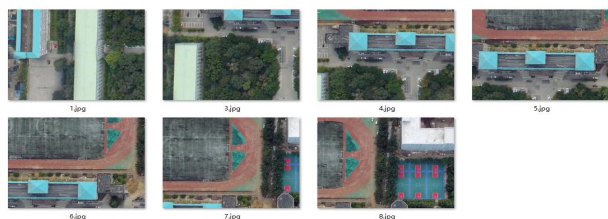


Fig. 2 Mosaic HD video image



Fig. 3 Splicing result

Table 1 experimental results

Total time (s)	SIFT algorithm	H-SIFT algorithm
800*600Resolution	1.73	1.72
1920*1080Resolution	7.14	3.86
4000*3000Resolution	34.83	6.55

Experimental environment:

the experimental results are Debug version statistics.

CPU:intel Xeon E5-2680 *2

GPU:NVIDIA GeForce GTX 1080

CUDA acceleration is not enabled.

Empirical Conclusion

Aiming at the problems of high complexity, large computation and long computing time of SIFT based image mosaic algorithm, this paper proposes a H-SIFT algorithm, which is mainly used in the stitching process of high definition video images. In the process of image preprocessing, the algorithm can adjust the image size intelligently. In the process of feature extraction, the algorithm can control the number of feature points and the number of matched feature points through the size of image resolution. In order to guarantee the SIFT algorithm with higher registration ratio, improve the speed of image registration. The experimental results show that the H-SIFT algorithm can shorten the time of image mosaic, achieve the real-time stitching of high definition video images, and achieve good image mosaic effect.

Acknowledgements

This work was financially supported by Shenyang Program for Tackling Key Problems in Science and Technology (F15-126-9-00).

References

- [1] Lin, Yuan-Sheng,Xu, Gang,Jiang, Ming,Zhou, Peng,Jiang, Juan-Juan. A Novel Image Matching Method Based on SIFT[J]. Sensors & Transducers,2014,171(5).
- [2] Zahra Hossein-Nejad,Mehdi Nasri. An adaptive image registration method based on SIFT features and RANSAC transform[J]. Computers and Electrical Engineering,2016.
- [3] WENYU CHEN,WENZHI XIE,RU ZENG. Method of item recognition based on SIFT and SURF †[J]. Mathematical Structures in Computer Science,2014,24(5).
- [4] Ali Nouman,Bajwa Khalid Bashir,Sablatnig Robert,Chatzichristofis Savvas A,Iqbal Zeshan,Rashid Muhammad,Habib Hafiz Adnan. A Novel Image Retrieval Based on Visual Words Integration of SIFT and SURF.[J]. PloS one,2016,11(6).
- [5] Sen Wang,Yin Hui Zhang,Zhong Hai Shi,Zi Fen He. Method of Footprint Image Stitching Based on Multiscale SIFT Feature Matching[J]. Applied Mechanics and Materials,2013,2308(303).
- [6] Zhen Wu,Zhe Xu,Rui Nian Zhang,Shao Mei Li. SIFT Feature Extraction Algorithm for Image in DCT Domain[J]. Applied Mechanics and Materials,2013,2560(347).