A METHOD OF SIFT FEATURE POINTS MATCHING FOR IMAGE MOSAIC

JIE ZHAO[1], HUI-JUAN ZHOU[1], GUO-ZUN MEN[2]

[1]College of Electronic and Information Engineering, Hebei University, Baoding 071002, China
[2]College of Economics, Hebei University, Baoding 071002, China
E-MAIL: zhaojie@mail.hbu.edu.cn, keaiduo612@163.com, gzmen@mail.hbu.edu.cn

Abstract:

This paper presents a approach of SIFT feature points matching for image mosaic. This method combines improved K-means clustering and simulated annealing algorithm to match SIFT feature points. Firstly, high robust points are extracted by SIFT algorithm; Secondly, cluster with the initial centers obtained by density function , and then optimize the results of clustering which are used as initial results of simulated annealing algorithm by perturbation; Thirdly, match feature points according to Nearest Neighbor algorithm; Finally, calculate the homography and realize image mosaic. This method does not need to traverse all feature points and avoid trapping in a local extremum. Experimental results prove that the method is only relative to geometric position of feature points, and is robust on scale invariant, arbitrary rotation and scaling.

Keywords:

K-means; SIFT; Simulated annealing; Homography; Image mosaic

1. Introduction

Image mosaic is an important branch of computer vision. By combining data from multiple images, it forms a large composite image or mosaic to enlarge our vision. Image mosaic technology is widely used in many fields including space exploration, undersea exploration, medicine, meteorology, geological surveying, military, video compression and transmission, digital files preservation, 3D reconstruction, security of evidence and so forth.

Generally, image mosaic includes four steps: image acquisition, image registration, homography calculation and image mosaic. Image registration is the most difficult and most critical procedure which not only extracts feature points from images, but also searches correct matching points. There are two main methods of image registration. One is based on image gray correlation[1,2,3], the other is based on image feature[4,5,6]. The former is a classical method of image registration, which selects a window in one image as templet, and searches the most similar matching window in the other image according to gray correlation. This method needs masses of calculations and can not deal with the occlusion problem. The latter compresses image information and reduces calculations by extracting dominate image feature, and solves the occlusion problem simultaneously by using the structure information of image feature. So the image registration method based on image feature becomes a focus recently. Generally speaking, people use Random Sample Consensus to match images, nevertheless this method needs to traverse all of feature points, which makes a large number of calculations. Somebody has proposed optimized methods based on neural network[7] and integer programming[8]. But these methods often get into a local extremum in searching space of multi-peak distribution because they execute single point searching in solution space.

In this paper, a SIFT feature point matching method based on improved K-means clustering[9,10] and simulated annealing[11] algorithm is proposed. Firstly, extract feature points from a serial of overlapped images by Low's improved SIFT(Scalable Invariant Feature Transform) algorithm[12]. Secondly, select high density points from the overlap as the initial seeds, compute the Euclidean distances between seeds and other points, classify the points into the nearest clustering, get a global extremum by simulated annealing, then use the Nearest Neighbor algorithm[13,14] to match the clusterings. Finally, estimate the homography between images by RANSAC, then use the fade-in and fade-out method to register images. This method combines improved K-means and simulated annealing algorithm so that not to traverse all of feature points, and avoids trapping in a local extremum. Experimental results prove that this method can also work well on the image sequence with large change of scale, parallax and illumination.

2. Feature extracted

There are some noises in visual angle and scale of the obtained image sequence because of the movement of

978-1-4244-3703-0/09/$25.00 ©2009 IEEE

camera. In order to avoid the impact of noises, improved SIFT algorithm by Low in 2004 is introduced to extract the feature points. Following are the major stages of computation used to generate the set of image features:

Ⅰ. Scale-space extrema detection: The first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation.

Ⅱ. Keypoint localization: At each candidate location, a detailed model is fit to determine location and scale. Keypoints are selected based on measures of their stability.

Ⅲ. Orientation assignment: One or more orientations are assigned to each keypoint location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.

Ⅳ. Keypoint descriptor: The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

3. Feature matching

This paper combines improved K-means clustering algorithm with simulated annealing algorithm to match feature points.

The K-means is a typical clustering algorithm based on distance. It adopts distance as the evaluative index of the similarity, deeming that closer the distance between two objects is, more similar they are. However, randomly selected initial center points of the traditional K-means algorithm would lead to a random result. To avoid this, density function is used to get the initial clustering centers.

Simulated annealing algorithm was proposed by Metropolis et al[11]. It is based on the similarity between annealing process of Solid State Physics and combinatorial optimization problem. Simulated annealing algorithm has the advantages as follows: get global optimization solution with a greater probability; manipulate various variables of optimum design including discrete, continuous and mixed various; has no need to any auxiliary information and so on.

The algorithm of the feature matching is as follows:

Ⅰ. Assume the sample points input are N, the number of the clusterings are K. From (1) we can calculate each sample points' density which is defined as the number of sample points in the hypersphere with a center X and radius σ.

$$M(X)=|\{W \in Z \mid d(X,W) \le \sigma\}| \quad (1)$$

Z——Sample aggregate

d—— Euclidean distance

σ——Radius

Ⅱ. Compute the Euclidean distances between seeds and other points by equation (2), classify the points into the nearest clustering.

$$D(X_{i,j}, K_{m,n}) = \sqrt{(X_{i,1}-K_{m,1})^2 + (X_{i,2}-K_{m,2})^2} \quad (2)$$

Where $X_{i,1}$, $X_{i,2}$, $X_{m,1}$, $X_{m,2}$ is abscissa and ordinate of sample point X_i ($0 \le i \le N-K$) and center point K_m ($0 \le m \le K$), respectively.

Ⅲ. Recount the value of clustering centers and the distances between other points and the centers. If cluster centers or data aggregate samples change, go back to stepⅢ. Otherwise stop the cluster.

Ⅳ. Take the results of cluster as the initial solution of simulated annealing algorithm, and compute the objective function as follow:

$$S_{um} = \sum_{\substack{0 \le i \le N-K \\ 0 \le j \le K}} D(X_{i,j}, K_{m,n}) \quad (3)$$

Ⅴ. The initial temperature is T_0, $T(t)= T_0A^t$, annealing at A rate.

Ⅵ.Randomly change the current category of one or several clusters at a certain temperature T, and then calculate the new objective function S'_{um}. If S'_{um} is the optimization function, save the cluster results. Otherwise judge function (4).

$$\Delta S = S'_{um} - S_{um} = \begin{cases} \ge 0, (P = e^{-\Delta S/KT}) \\ < 0, (S_{um} = S'_{um}) \end{cases} \quad (4)$$

Ⅶ.Do iterative step Ⅵ. If the end condition is satisfied, the current solution is the optimization. Export it and stop algorithm. Otherwise, reduce temperature and go on iterative.

The clustering result can be got from the algorithm stated above. Then the clustering matching can be performed with Nearest Neighbor algorithm: suppose a threshold value, below which the pair of clustering is accepted. It is clear that a low threshold means a high accuracy.

3.1. 2D Motion models

Parametric models are as the assumptions on the relation between two images.

Table 1 shows the matrix , D.O.F. , preserves and icon corresponding with each transformation including translation, Euclidean, similarity, affine and projective.

Table 1. Related information of 2D motion models

Name	Matrix	#D.O.F.	Preserves	Icon
translation	$[I \mid t]_{2\times3}$	2	orientation	
Euclidean	$[R \mid t]_{2\times3}$	3	lengths	
similarity	$[sR \mid t]_{2\times3}$	4	angles	
affine	$[A]_{2\times3}$	6	parallelism	
projective	$[\tilde{H}]_{3\times3}$	8	straight lines	

3.2. Homography calculation

A serial images captured by rotating the camera around the optical center can be associated with each other through a homography plane, and register by projective deformation interpolation. The procedures are as follow:

Ⅰ.Randomly select an image as reference image from the images sequence.

Ⅱ. Choose one in the rest images, use RNASAC[15,16] to compute the homography between the chosen image and the reference image.

$$\begin{bmatrix} x \\ y \\ s \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} (x) \\ (y) \\ (1) \end{bmatrix} = H \begin{bmatrix} (x) \\ (y) \\ (1) \end{bmatrix} \tag{5}$$

A. To each random minimum subset, calculate the corresponding H and the median of all clusters' the sum error.

$$E_{med} = \underset{i=1,2,\dots,n}{med} \left[d^2\left(x_{\xi1}, Hx_{\xi2}\right) + d^2\left(x_{\xi2}, H^{-1}x_{\xi1}\right) \right] \tag{6}$$

$x_{\xi1}$, $x_{\xi2}$ ——Corresponding cluster center

d——Euclidean distance

B. Repeat A, record H with minimum E_{med}.

C. Calculate the variance of robust estimate $\sigma^{\wedge}$ to reduce the influence of the noise.

$$\sigma^{\wedge} = 1.4826\left(1 + \frac{5}{n-4}\right)\sqrt{E_{med}} \tag{7}$$

Weight W_i :

$$W_i = \begin{cases} 1, r_t^2 \le (25\sigma^{\wedge})^2 \\ 0 \end{cases} \tag{8}$$

$$r_t = d^2(x_{\xi2}, Hx_{\xi1}) + d^2(x_{\xi1}, H^{-1}x_{\xi2})$$

D. Calculate new value H^* with known W_i by weighting least square:

$$H^* = \arg\min \sum_i w_i r_i^2 \tag{9}$$

Ⅲ.Make projective deformation interpolation with the homography H^* , and enlarge the reference image with the interpolation image.

Ⅳ. Repeat Ⅱ, Ⅲ on other images.

4. Image stitching

If we stitch the images directly, ghosting will appear due to the luminance difference. It's the problem that needs to be solved in image mosaic.

We employ the fade-in and fade-out method stitching images gaplessly and smoothly. The value of the pixel in the overlap is the weighted average value of the corresponding pixel in the two images. Assume pixel3 value of point A is pixel1, the pixel value of the corresponding point B in the neighbor image is pixel2. Thus after the image mosaic, the pixel value of the corresponding point is:

pixel3=kpixel+(1-k)pixel2, where k is weight factor, ranges from (0,1).

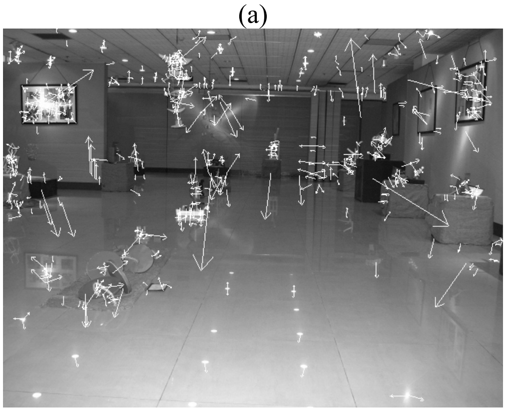
5. Experimental results

In experiment,we use a hand-hold camera Canon A480 capturing an image sequence of exhibition and campus scenery. The image resolution is 640x480. To avoid deflexion and pitching when the camera-handed wheel and to ensure that the pictures of the level of the center line with the focus of the camera center at the same horizontal surface, we support the camera on a tripod. After capturing an image, we rotate the camera at an angle to ensure that the images have at least twenty percent overlapping. There are two exhibition sample images shown in Figure 1.

Figure 1. Exhibition sample image

The results of extracting SIFT feature points are shown in Figure 2.

(a)



(b)

Figure 2. Results of extracting SIFT feature points

Firstly, transform the image into the gray-scale image and extract the feature points, shown in figure 2. There are 716 feature points extracted (a) and 680 feature points extracted (b). Secondly, use K-means algorithm and simulated annealing algorithm to match the feature points in overlapped area. It needs optimization on the clustering number in this process for more or less clustering would influence the homography between images. In this work, we found that the best number of clustering is 40.

We use this method and traditional traverse-point method to extract and match the two pictures shown in figure 1, respectively. The result shows the method in this paper is faster. The time of matching is 8.18 seconds and 10.89 seconds, respectively.

After getting the matching relationship between the neighbor images, the RANSAN algorithm is used to calculate the homography. Finally, perform seamless mosaic with fade-in and fade-out method. The result is shown in figure 3.

Figure 3. Results of image stitching

Figure 4. Results of 5 pictures image stitching

The picture shown in figure 4 is stitched with five overlapped pictures.

It can be seen from the experimental result that this algorithm does well on the cluster of the feature points got from SIFT, and makes wide vision, high resolution image after mosaic.

6. Conclusions

In this paper, a method of SIFT feature points matching for image mosaic is proposed. Firstly, use SIFT algorithm to extracted feature points from gray image. Then cluster features round corresponding cluster center by K-means and simulated annealing algorithm, and use Nearest Neighbor algorithm to match the clusterings. Through these procedures, we get correct matching points. Finally, use RANSAC algorithm to estimate the homography between images and stitch images by fade-in and fade-out method.

From the results of figure 3 and figure 4 we can see that this method does not need to traverse all feature points and avoid trapping in a local extremum. Experimental results prove that the method is only relative to geometric position of feature points, and is robust on scale invariant, arbitrary rotation and scaling.

Acknowledgements

This work is supported by the research and development projects of Science and technology of Hebei Province [Project No.06242188D-2], Natural Science Foundation of Hebei Province [Project No.F2007000221] and Main Item of Natural Science Research Plans of The Education Department of Hebei Province [Project No.ZH2006012].

References

[1] THORNTON J, and SAVV IDOS M. "A Bayesian approach to deformed pattern matching of iris image[J]". IEEE Pattern Analysis Machine Intelligence, Vol 29,No. 4, pp. 596-606, 2007.

[2] Fonseca L M G, and Manjunath B S. "Registration Techniques for Multisensor Remotely Sensed Imagery [J]". Photogrammetric Engineering and Remote Sensing, pp. 1049-1056, 1996.

[3] Althof R J , and Dobbins J T. "A Rapid and Automatic Image Registration Algorithm with Subpixel Accuracy [J]". IEEE Transactions on Medical Imaging, pp. 308-316, 1997.

[4] JA IN A K, and CHEN Yi. "Pores and ridges: high-resolution finger print match using level3 features[J]". IEEE Pattern Analysis Machine Intelligence, Vol 29, No. 1, pp. 15-27, 2007.

[5] Guest E, and Berry E. "Robust Point Correspondence Applied to Two-and Three-dimensional Image Registration[J]". IEEE Transaction on Pattern Analysis and Machine Intelligence, pp. 165-179, 2001.

[6] Yang Z, and Cohen F S. "Image Registration and Object Recognition Using Affine Invariants and Convex Hulls [J]". IEEE Transactions on Image Processing, pp. 934-946, 1999.

[7] Nasrabadi, and Choo C Y. "Hopfield network for stereo vision correspondence". IEEE trans NN, 1992.

[8] Vinod VV, Ghose S. "Point Matching Using Asymmetric Neural Networks". Pattern Recognition, pp. 1207, 1993.

[9] Kiri Wagstaff, and Claire Cardie. "Constrained K-means Clustering with Background Knowledge". Proceedings of the Eighteenth International Conference on Machine Learning, pp. 577-584, 2001.

[10] Yiu-Ming and Cheung. "K*-Means: A new generalized k-means clustering algorithm". Pattern Recognition Letters, pp. 2883-2893, April 2003.

[11] Kirkpatrick, and Gelatt, "Optimization by Simulated Annealing". Science, pp. 671-680, May. 1983.

[12] Lowe D G. "Distinctive image features from scale-invariant keypoints"，International Journal of Computer Vision, Jan. 2004.

[13] Arya, S., and Mount. "Approximate nearest neighbor queries in fixed dimensions". In Fourth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA’93), pp. 271-280.1993.

[14] Arya, S., and Mount. "An optimal algorithm for approximate nearest neighbor searching". Journal of the ACM, pp. 891-923. 1998.

[15] R.Hartley, and A.Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, ISBN:0521540518, second edition, 2004.

[16] Fishier M A , and Boles R C. "Random Sample Concensus :A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography[J]".Comm Assoc Comp Mach, Vol 24, No. 6, pp. 381-395, 1981.