Image Matching Based on Singular Value Decomposition

Feng Zhao[1,2]

[1] Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China
[2] Graduate School of the Chinese Academy of Sciences, Beijing 100039, China
fzhao@jdl.ac.cn

Abstract. This paper presents a simple and effective method for matching two uncalibrated images. First, corner points are extracted in both images separately. Then the initial set of point matches is obtained by singular value decomposition of a well-designed correspondence strength matrix. A new expression of this matrix is introduced to get more reliable initial matches. Last, the epipolar geometry constraint is imposed to reject the false matches. Experimental results on real images show this method to be effective for general image matching.

1 Introduction

Matching two uncalibrated images is a classic problem in computer vision. The topic has been researched for several decades. One effective strategy is using interest point matching technology. The interest point matching technology first extracts points of interest such as corners in the two images and then uses robust matching technology to establish point correspondences between the two images [1,2,3,4].

Corners are considered as good candidates for interest points in many computer vision applications such as motion correspondence, object tracking and stereo matching, etc. Using corner points as interest points has been proved to be effective in image matching [1,2]. Among the most popular corner detectors, Harris corner detector [5] is known to be robust against rotation and illumination changes. And its results have high repeatability under different imaging conditions, which is of great advantage to image matching task [6].

In the recent literature some effective strategies for interest point correspondence between image pairs have emerged. A robust technique for matching two uncalibrated images has been proposed by Zhang et al. [1], which finds initial matches using correlation and relaxation methods followed by the LMedS technique to discard false matches. Pilu [2] used a direct method based on singular value decomposition for feature correspondence, which was originally proposed by Scott and Longuet-Higgins [7]. The method first sets up a correlation-weighted proximity matrix and then performs singular value decomposition calculation on

K. Aizawa, Y. Nakamura, and S. Satoh (Eds.): PCM 2004, LNCS 3333, pp. 119–126, 2004.

the matrix to get initial matches. We improve Pilu's method by giving a new expression of the matrix. Experiments on real images show the improved method can handle more complicated cases and obtain more reliable initial matches.

The initial set of interest point matches usually contains some false matches due to the bad locations of corners and the improper matches in the establishment of correspondence. Constraints like epipolar geometry or disparity gradient limit can be used to reject the false matches [1,8,9].

This paper presents a simple and effective method to solve the problem of image matching. The method is based on singular value decomposition. The initial set of point matches is directly obtained by singular value decomposition of a well-designed correspondence strength matrix. A new expression of the correspondence strength matrix is introduced. The comparison of matching results on real images demonstrates the new matrix outperforms that used in previous works. Experimental results on real images show the method is effective for matching image pairs under different imaging conditions.

The remainder of the paper is organized as follows. Section 2 describes extracting corners as interest points with slightly modified Harris corner detector. Section 3 introduces the matching approach based on singular value decomposition and presents the new expression of the correspondence strength matrix. Section 4 describes rejecting the false matches by imposing epipolar geometry constraint. Section 5 presents some experimental results on real images and Section 6 concludes the paper.

2 Extracting Corners as Interest Points

Corners are highly informative image locations. They are very useful features for many computer vision applications. Corners can be automatically detected without any prior knowledge and they are stable features for image matching task. Many algorithms for detecting corners have been reported up to now. Among the most popular corner detectors, Harris corner detector [5] is known to be robust against rotation and illumination changes. The results of Harris corner detector have high repeatability under different imaging conditions, which is very important for image matching.

Harris corner detector is based on the auto-correlation matrix, which is built as follows:

$$M = \exp - \frac{x^2 + y^2}{2\sigma^2} \otimes \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} , \qquad (1)$$

where $\otimes$ is the convolution operation. I_x and I_y indicate the x and y directional derivative respectively. The auto-correlation matrix performs a smoothing operation on the products of the first derivatives by convolving with a Gaussian window. Two sufficient large eigenvalues of M indicate the presence of a corner. To avoid the explicit eigenvalues decomposition of M, a corner response function is defined by the following expression:

$$C_H = \det(M) - k * \mathrm{trace}(M)^2 , \qquad (2)$$

where k is usually set to 0.04. Then a threshold t_h can be used to select corner points. If $C_H > t_h$ then the point is identified as a corner.

In the original Harris detector, I_x and I_y are computed by convolution with the mask [-1 0 1] and the corner response function requires setting a parameter k. In our implementation, the mask [-2 -1 0 1 2] is used to compute the first derivatives. And a corner measure function without additional parameter [10] is adopted. The measure function is defined as follows:

$$C_F = \frac{\det(M)}{\mathrm{trace}(M)} \ . \tag{3}$$

A threshold t_f is given and a point is considered as a corner if $C_F > t_f$. In our implementation, t_f is set to 1% of the maximum observed C_F.

3 Matching Based on Singular Value Decomposition

The problem of matching interest points in two images is fundamental in computer vision. A direct method for establishing feature correspondences between two images based on singular value decomposition (SVD) has been proposed in Scott and Longuet-Higgins [7] and further developed in Pilu [2]. The basic idea of this method is obtaining feature correspondences with singular value decomposition of a correspondence strength matrix. In the following, the algorithm will be briefly described.

Let I_1 and I_2 be two images. I_1 contains m features $I_{1i}(i = 1 \dots m)$ and I_2 contains n features $I_{2j}(j = 1 \dots n)$. To get one-to-one feature correspondences between the two images, a proximity matrix $G \in M_{m,n}$ is set up first:

$$G_{ij} = e^{-r_{ij}^2/2\sigma^2} \ , \tag{4}$$

where $r_{ij} = \| I_{1i} - I_{2j} \|$ is the Euclidean distance between the two features if they are regarded as being on the same plane. G is a positive definite matrix and the element of G decreases monotonically from 1 to 0 with the increase of the distance between the two features. The degree of interaction between the two sets of features is controlled by the parameter σ.

The next step of the algorithm is to perform the singular value decomposition of G:

$$G = UDV^T \ , \tag{5}$$

where $U \in M_{m,m}$ and $V \in M_{n,n}$ are orthogonal matrices and the diagonal matrix $D \in M_{m,n}$ contains the singular values along its diagonal elements D_{ii} in descending order.

Then a new matrix E is converted from the diagonal matrix D by replacing every diagonal element of D with 1. Computing the following product will get the new matrix H:

$$H = UEV^T \ . \tag{6}$$

The matrix $H \in M_{m,n}$ has the same shape as the proximity matrix G. Two features I_{1i} and I_{2j} are pairing up if H_{ij} is both the greatest element in its row and the greatest element in its column. With selecting all such elements in H, the feature correspondences between the two images can be established.

Scott and Longuet-Higgins algorithm only takes spatial location into account for establishing feature correspondences. To include similarity information in feature matching with above algorithm, Pilu [2] uses normalized cross-correlation score as local measurement to quantify feature similarity. Adding similarity constraint can eliminate rogue features, which shouldn't be similar to anything. The elements of G can then be transformed as follows:

$$G_{ij} = \frac{(C_{ij}+1)}{2} \, e^{-r_{ij}^2/2\sigma^2} , \tag{7}$$

where C_{ij} is the normalized cross-correlation score between the two features.

In our method, a new expression of the correspondence strength matrix G is introduced in order to get more reliable initial matches under different imaging conditions. The new definition of the correspondence strength matrix G is given as follows:

$$G_{ij} = (C_{ij}+1)^3 \, e^{-r_{ij}/2\sigma^2} , \tag{8}$$

where σ is set to 50 in our system. C_{ij} can be calculated as follows:

Let $m_1 = I_1(x_i, y_i)$ be the i-th interest point in the first image I_1 and $m_2 = I_2(x_j, y_j)$ be the j-th interest point in the second image I_2. W_1 and W_2 are two windows of size $(2w+1) \times (2w+1)$ centered on each point. The normalized cross-correlation score C_{ij} is defined as:

$$C_{ij} = \frac{\sum_{u=-w}^{w} \sum_{v=-w}^{w} [I_1(x_i+u, y_i+v) - \overline{I_1(x_i, y_i)}\,][I_2(x_j+u, y_j+v) - \overline{I_2(x_j, y_j)}\,]}{(2w+1)(2w+1)\sigma(m_1)\sigma(m_2)} , \tag{9}$$

where $\overline{I_1(x_i, y_i)}$ ($\overline{I_2(x_j, y_j)}$) is the average and $\sigma(m_1)$ ($\sigma(m_2)$) is the standard deviation of all the pixels in the correlation window. The size of the correlation window is 11 in our system.

The new definition of G gives more weight to the similarity measurement (i.e. normalized cross-correlation score) and weakens the spatial location measurement. Experiments on complicated real images show the new expression of G yields better results. The initial set of matches contains less false matches than the former methods and the number of good matches increases simultaneously.

4 Rejecting False Matches

The initial set of interest point matches usually contains some false matches. Constraints like epipolar geometry can be used to reject the false matches. The epipolar geometry constraint can be described as follows.

Suppose F is the fundamental matrix. Point $p\,(x, y)$ in the image can be represented as:

$$\widetilde{p} = [x, y, 1]^T . \tag{10}$$

For a point match (p, q), the epipolar line of point p in the first image is defined as:

$$l_p = F\widetilde{p} \,. \tag{11}$$

If the match is perfect, then point q in the second image should lie on the epipolar line l_p exactly. The distance d_q of point q to the epipolar line l_p is calculated by

$$d_q = \frac{|\, \widetilde{q}^T F \widetilde{p} \,|}{\sqrt{(F\widetilde{p})_1^2 + (F\widetilde{p})_2^2}} \,, \tag{12}$$

where $(F\widetilde{p})_i$ is the i-th component of vector $F\widetilde{p}$. The distance d_p of point p to the epipolar line l_q is calculated similarly. Then a threshold t_e can be used to find the good matches. A point match is identified as a good match if $\max(d_p\,, d_q) \leq t_e$.

In our implementation, the epipolar geometry constraint is imposed based on RANSAC [11]. The method is described as follows. First, eight matches are randomly chosen from the initial set of point matches and the fundamental matrix F is calculated from them. Then we can find all good matches that are consistent with the epipolar geometry constraint according to this fundamental matrix. The threshold t_e is 1.5 in our system. Repeat the above steps to get the largest set of good matches.

5 Experimental Results

Experiments on real images of various content have been performed. Some of the results are reported here. These images are under different imaging conditions such as rotation, scale changes, illumination changes and the combination of them.

All the results here are obtained with the same parameter setting as mentioned in the forgoing sections. These results are the final results after applying RANSAC on the initial matches. Image pairs **Boat**, **Residence**, **Car** are from

Fig. 1. Matching result for image pair **Boat** with rotation and scale changes

Fig. 2. Matching result for image pair **Residence** with rotation and scale changes

Fig. 3. Matching result for image pair **Car** with illumination changes

Fig. 4. Matching result for image pair **Gate** with translation and rotation

Table 1. The columns of "Final Matches" contain the number of good matches after rejecting the false matches in the initial matches with epipolar geometry constraints. In addition, the algorithm using original G fails in matching the image pair **Gate**, therefore the corresponding number of initial matches and final matches are small

	Original G		**Improved G**	
	Initial Matches	**Final Matches**	**Initial Matches**	**Final Matches**
Boat	88	65	109	91
Residence	89	51	133	102
Car	85	55	100	72
Gate	45	10	133	96

INRIA[1], and image pair **Gate** is from our lab. More reliable matching results have been obtained by using the new expression of the correspondence strength matrix G . Table 1 gives the comparison of matching results between original definition of G (Equation (7)) and the improved G (Equation (8)).

6 Conclusions

We have presented a simple and effective method for matching two uncalibrated images. The method exploits singular value decomposition, which is one of the stablest numerical matrix operations. The kernel of this method is establishing interest point correspondences by singular value decomposition of a well-designed correspondence strength matrix. We introduce a new expression of this matrix to get more reliable initial matches. The comparison of matching results on real images demonstrates the new matrix outperforms that used in previous works. One characteristic of this algorithm is that it does not require a correlation threshold for a candidate point match to be accepted. This contributes to the simplicity of the method. Experimental results on real images show the method is effective for matching image pairs under different imaging conditions.

Acknowledgements. This work is supported by National Hi-Tech Development Programs of China under grant No. 2003AA142140.

References

1. Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry. *Artificial Intelligence Journal*, Vol. 78, pp. 87–119, 1995
2. M. Pilu. A direct method for stereo correspondence based on singular value decomposition. *IEEE CVPR*, pp. 261–266, 1997

[1] http://lear.inrialpes.fr/people/Mikolajczyk/Database/index.html

3. A. Baumberg. Reliable feature matching across widely separated views. *IEEE CVPR*, pp. 774–781, 2000
4. D. Tell and S. Carlsson. Combining Appearance and Topology for Wide Baseline Matching. *7th European Conference on Computer Vision (ECCV)*, pp. 68–81, 2002
5. C. Harris and M. Stephens. A combined corner and edge detector. *Fourth Alvey Vision Conference*, pp. 148–151, 1988
6. C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of Interest Point Detectors. *International Journal of Computer Vision*, Vol. 37, pp. 151–172, 2000
7. G. Scott and H. Longuet-Higgins. An algorithm for associating the features of two patterns. *Proceedings of the Royal Statistical Society of London*, Vol. B244, pp. 21–26, 1991
8. G. Xu, E. Nishimura, and S. Tsuji. Image correspondence and segmentation by epipolar lines: Theory, algorithm and applications. Technical report, Dept. of Systems Engineering, Osaka University, Japan, 1993
9. S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. PMF: A stereo correspondence algorithm using a disparity gradient limit. *Perception*, Vol. 14, pp. 449–470, 1985
10. W. Faustner. A feature based correspondence algorithm for image matching. *International Archives of Photogrammetry and Remote Sensing*, Vol. 26, pp. 150–166, 1986
11. M. A. Fischler and R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, Vol. 24, pp. 381–395, 1981