ACCV2002: The 5th Asian Conference on Computer Vision, 23--25 January 2002, Melbourne, Australia.

On the Automatic Estimation of Fundamental Matrix

C. Lei[1], H. T. Tsui[2], Z. Y. Hu[1]

1. National Laboratory of Pattern Recognition, CAS, Beijing, 100080, P.R. China
2. Department of Electronic Engineering, Chinese University of Hong Kong

{leic, huzy}@nlpr.ia.ac.cn; httsui@ee.cuhk.edu.hk

Abstract

Accurate and robust estimation of the Fundamental matrix is very important for many computer vision applications, such as 3D reconstruction from 2D images. In this paper, a complete analysis and comparison of three important algorithms for estimating the Fundamental matrix when three views are given. The six-point algorithm originally used for three-view geometry is implemented for the automatic estimation of the Fundamental Matrix — the two-view geometry tensor, assuming that more than two views are available. An extensive comparison to other two popular methods: the seven-point and eight-point algorithms are made through carefully designed experiments with synthetic and real images, assuming three views are available. The comparisons show that the new-implemented six-point algorithm, which makes use of the three-view constraint, is more accurate and has higher ability to handle the outliers.

1. Introduction

The Fundamental Matrix encapsulates the whole epipolar geometry between two views of a static scene and surely plays a "fundamental" role in various applications of Computer Vision. Due to its importance, much effort has been exerting on implementing more robust and accurate methods for the Fundamental Matrix estimation. According to Zhang's review [4], the existing methods for the Fundamental Matrix estimation can be mainly classified into linear ones and non-linear ones. They are all based on the basic constraint from the epipolar geometry. The differences lie in their particular strategies to cope with the noise and/or outliers, the specific definitions of the adopted minimization criteria in the optional non-linear optimization procedure, the specific parameterizations or measures to ensure the rank-2-ness of the estimated Fundamental Matrix and so on.

To estimate the Fundamental Matrix automatically, some robust-statistics-based techniques must be used to reject the outliers (mismatches) from automatic image matching. In the review of Torr [5] on the development history and performance comparison of the robust estimation methods, RANSAC (Random sample consensus) paradigm [8] was recommended, which has been successfully used for many other parameter estimation problems.

The objective of this paper is to determine the most accurate and robust method for estimating the Fundamental Matrix automatically when three views are available. Specifically, the eight-point algorithm (normalized) [7, 9] is a linear method that works with two images. And the seven-point algorithm [13] is a non-linear method that also works with two images. With three views given, the Fundamental Matrices between three image pairs can be estimated pair-wisely by implementing any one of these two algorithms as the corresponding RANSAC search engine. The estimation accuracy may be improved by an additional non-linear optimization taking into account the constraints derived from the two-view geometry. Alternatively, a six-point algorithm to estimate directly the geometry of 3 views proposed in [1] can also be used. In fact, an equivalent six-point minimal subset based RANSAC method has been successfully implemented by Torr to estimate the tri-focal tensor [6]. Other successful applications of the six-point algorithm have been reported in many papers recently, such as [2, 3], with its robustness proven and applicability recommended. In this paper, we similarly implement the six-point algorithm as the search engine of the RANSAC for the automatic estimation of Fundamental Matrix. An extensive comparison of the above three algorithms is performed through experiments using synthetic and real images.

2. A short review of the methods for estimating Fundamental Matrix

2.1 Epipolar geometry

The Epipolar geometry arises from any bi-ocular system. As illustrated in Fig.1, two views $\boldsymbol{I}$ and $\boldsymbol{I}'$ of the same scene are captured by two/single camera(s) located at two optical centers $\boldsymbol{C}$ and $\boldsymbol{C}'$ simultaneously or respectively. Assume that two imaged points of a same 3D space point $\boldsymbol{M}$ in view $\boldsymbol{I}$ and $\boldsymbol{I}'$ are $\boldsymbol{m}$ and $\boldsymbol{m}'$ (3-vector of the homogenous coordinates) respectively, then the epipolar geometry relation:

$$\boldsymbol{m}'^T \boldsymbol{F} \boldsymbol{m} = 0 \tag{1}$$

is satisfied, where the 3×3 singular matrix $\boldsymbol{F}$ is the so-called Fundamental Matrix. Geometrically, it means that the corresponding point $\boldsymbol{m}'$ ($\boldsymbol{m}$) of point $\boldsymbol{m}$ ($\boldsymbol{m}'$) is lying on the corresponding epipolar line $\boldsymbol{Fm}$ ($\boldsymbol{F}^T\boldsymbol{m}'$). All the epipolar lines in view $\boldsymbol{I}'$ ($\boldsymbol{I}$) pass a same point, the so-called epipole $\boldsymbol{e}'$ ($\boldsymbol{e}$), which is the projection of the optical center $\boldsymbol{C}$ ($\boldsymbol{C}'$) in view $\boldsymbol{I}'$ ($\boldsymbol{I}$) respectively. This property reflects in the imaged two views in that all the epipolar lines in one view $\boldsymbol{I}$ ($\boldsymbol{I}'$) will interest at the epipole $\boldsymbol{e}$ ($\boldsymbol{e}'$), while in the algebraic description is the rank-2-ness of $\boldsymbol{F}$.

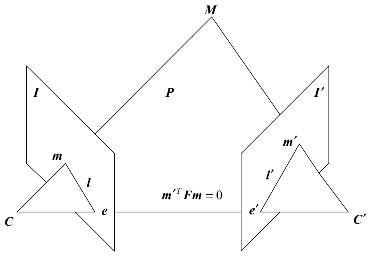


Fig. 1 The epipolar geometry

2.2 Methods for estimating Fundamental Matrix

The basic constraint for estimating the Fundamental Matrix $\boldsymbol{F}$ is from equation (1). Specifically, give n matches between two images, by rewriting the to-be-estimated $\boldsymbol{F}$ as a 9-vector $\boldsymbol{f}$, we have

$$\boldsymbol{A}\,\boldsymbol{f} = 0 \tag{2}$$

with $\boldsymbol{A} = \begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix}$,

$\boldsymbol{f} = [\boldsymbol{F}_{11}, \boldsymbol{F}_{12}, \boldsymbol{F}_{13}, \boldsymbol{F}_{21}, \boldsymbol{F}_{22}, \boldsymbol{F}_{23}, \boldsymbol{F}_{31}, \boldsymbol{F}_{32}, \boldsymbol{F}_{33}]^T$,

where $\boldsymbol{m}_i = [x_i, y_i, 1]^T$, $\boldsymbol{m}'_i = [x'_i, y'_i, 1]^T$ $(i = 1, \cdots n)$ are the n matches between two images.

Since the Fundamental Matrix $\boldsymbol{F}$ is singular and only defined up to an arbitrary scale, it has only 7 degrees of freedom. Therefore, seven matches are enough to determine $\boldsymbol{F}$, that is the so-called seven-point algorithm. With more matches available, the normalized eight-point algorithm [9] can be used. These two methods are both sensitive to the noise since the minimized residual of equation (2) has no physical meaning. Accordingly, many non-linear methods capable of minimizing some criteria having the physical meaning are proposed. Such criterions include, the symmetric distance of the image points to their corresponding epipolar lines:

$$\sum_{i=1}^{n} (d(\boldsymbol{m}_i, F^T \boldsymbol{m}'_i)^2 + d(\boldsymbol{m}'_i, F\boldsymbol{m}_i)^2) = \sum_{i=1}^{n} \frac{(\boldsymbol{m}'^T_i \boldsymbol{F}\boldsymbol{m}_i)^2}{(\boldsymbol{F}\boldsymbol{m}_i)_1^2 + (\boldsymbol{F}\boldsymbol{m}_i)_2^2} + \frac{(\boldsymbol{m}'^T_i \boldsymbol{F}\boldsymbol{m}_i)^2}{(\boldsymbol{F}^T\boldsymbol{m}'_i)_1^2 + (\boldsymbol{F}^T\boldsymbol{m}'_i)_1^2} \tag{3}$$

the Sampson distance:

$$\sum_{i=1}^{n} \frac{(\boldsymbol{m}'^T_i \boldsymbol{F}\boldsymbol{m}_i)^2}{(\boldsymbol{F}\boldsymbol{m}_i)_1^2 + (\boldsymbol{F}\boldsymbol{m}_i)_2^2 + (\boldsymbol{F}^T\boldsymbol{m}'_i)_1^2 + (\boldsymbol{F}^T\boldsymbol{m}'_i)_1^2} \tag{4}$$

or directly the re-projection error distance:

$$\sum_{i=1}^{n} (d(\boldsymbol{m}_i, \widehat{\boldsymbol{m}}_i)^2 + d(\boldsymbol{m}'_i, \widehat{\boldsymbol{m}}'_i)^2) \tag{5}$$

where $\widehat{\boldsymbol{m}}_i$ and $\widehat{\boldsymbol{m}}'_i$ are the estimated re-projection points determined by a appropriate triangulation method.

However, all the methods mentioned so far can't cope with the case where the inputted data are contaminated with outliers. Corresponding to this limitation, the robust estimation techniques such as M-Estimator, Case Deletion Diagnostics, Least Median of Squares (LMedS), and RANSAC, are introduced.(The readers are referred to the paper [5] for more details.) In this paper, we mainly focus on the RANSAC-based methods. According to its specific application in the Fundamental Matrix automatic estimation, the best minimal subset maximizing the number of the inliers in the inputted data, which are identified by thresholding a specific physical distance with respect to the Fundamental Matrix derived from the iteratively sampled minimal subset, is first traced out by repetitious subset samplings for a specifiable number of iterations. Then all the identified inliers, together with the obtained best Fundamental Matrix from the RANSAC optimization, are fed into a non-linear optimization procedure as initial values to reach the finally refined Fundamental Matrix. The necessary number of the minimal subset samplings is related to the size of the minimal subset and the percentage of the outliers by

$$\gamma = 1 - (1 - (1 - \varepsilon)^p)^m \tag{6}$$

where γ is the probability to find a good subset, ε is the percentage of the outliers, p is the size of the subset sampled in each iteration, and m is the needed sampling times.

From equation (6), we can see that, the smaller the size of the subset sampled, the fewer the sampling times are required for a specified level of confidence, usually 95%. Given two images, to estimate the Fundamental Matrix, the size of the minimal subset is at least seven. In this paper, two popular implementations in which the seven-point algorithm and the eight-point algorithm (normalized) are adopted as the search engine of RANSAC respectively are involved. For simplicity, we name these two RANSAC based Fundamental Matrix estimation methods as RANSAC-F-7 and RANSAC-F-8 respectively.

While in many applications needing to estimate the Fundamental Matrix, the number of the available images is often more than two. Furthermore, in some cases, the number of the "well-configured" matches among the available images is also possibly less than seven. If so, the popular estimation methods may fail, while yet much other useful information cannot be utilized. In the following section, we will show that, by resorting to the six-point algorithm that is originated from the three-view geometry, the above limitation can be counteracted to some extent.

3. Six-point algorithm and its application in fundamental matrix automatic estimation

In this section, we start our discussion of the so-called six-point algorithm from Quan's well-known work on the projective invariant between six space points and their image points [1]. According to [1], given three projective views, only six points are enough to obtain almost all the projective geometry information. Therefore, returning to our specific task, with more than two images available, it's natural for us to consider taking use of the six-point algorithm to estimate the Fundamental Matrix. In this paper, we try to implement the six-point algorithm as the search engine of a RANSAC based Fundamental Matrix estimation technique, similarly named as RANSAC-F-6 for simplicity.

3.1 Fundamentals of the six-point algorithm

According to [1], for a single image, there exists a simple invariant relationship between the invariant of the set of 6 points in space and the invariant of their projected image points. With three images given, computing the invariant in space is possible, which further enables us to compute invariant for projective reconstruction, epipolar geometry estimation and self-calibration.

Specifically, given any six points $\boldsymbol{X}^{(i)}(i=1,2,\cdots 6)$ in the 3D projective space $\boldsymbol{P}^3$, any five of them (no 3 of them collinear and no 4 of them coplanar) can be transformed into the canonical projective basis: { $(1,0,0,0)^T$, $(0,1,0,0)^T$, $(0,0,1,0)^T$, $(0,0,0,1)^T$ and $(1,1,1,1)^T$ }by a space collineation $\boldsymbol{H}_{4\times 4}$ (ranking 4). We assume the sixth point is correspondingly transformed into a new projective coordinate (X, Y, Z, T) by this collineation $\boldsymbol{H}_{4\times 4}$. Similarly, considering the six projected image points $\boldsymbol{x}^{(i)}(i=1,2,\cdots 6)$ in one view in 2D projective space $\boldsymbol{P}^2$, any 4 of them (no 3 of them collinear) can also be transformed into the canonical projective basis of $\boldsymbol{P}^2$ as follows: $(1,0,0)^T$, $(0,1,0)^T$, $(0,0,1)^T$ and $(1,1,1)^T$ by a plane collineation $\boldsymbol{H}_{3\times 3}$, with the left two image points transformed to their corresponding new coordinates $(u_5,v_5,w_5)^T$ and $(u_6,v_6,w_6)^T$ respectively.

With the camera imaging relationship regarded as a perspective projection from 3D projective space to the 2D projective space, we can get the following homogenous equation between (X, Y, Z, T) and $(u_i,v_i,w_i)^T\ (i=5,6)$:

$$\lambda_1 XY+\lambda_i XZ+\lambda_3 XT+\lambda_4 YZ+\lambda_5 YT -(\lambda_1+\lambda_2+\lambda_3+\lambda_4+\lambda_5)ZT=0 \quad (7)$$

where: $\lambda_1 = w_6(u_5-v_5)$, $\lambda_2 = v_6(w_5-u_5)$, $\lambda_3 = u_5(v_6-w_6)$, $\lambda_4 = u_6(v_5-w_5)$, $\lambda_5 = v_5(w_6-u_6)$ by eliminating the entries of the projective matrix in the $\boldsymbol{P}^3 \to \boldsymbol{P}^2$ imaging relation between the above transformed six points.

Given three images, we can obtain three such similar equations from which a cubic equation relating X and T can be obtained. Then we can get at most three solutions for X:T, and further corresponding solutions of Y:T and Z:T for each solution of X:T. Then for each group of solutions (X:T, Y:T, Z:T), since the projective homogenous coordinates of all the six space and image points are all determined, the projective matrices of the three images can be determined (with corresponding inverse transformations from the canonical coordinates to the original ones applied.) With the projective matrices $\boldsymbol{P}_i(i=1,2,3)$ for the three views determined, many kinds of useful information can be retrieved. As for our specific case, we can retrieve the Fundamental Matrices $\boldsymbol{F}_{ij}$ between the i-th and j-th views $(i,j=1,2,3\ and\ i\neq j)$ as:

$$\boldsymbol{F}_{ij} = [\boldsymbol{P}_j\boldsymbol{C}_i]_\times \boldsymbol{P}_j\boldsymbol{P}_i^+ \quad (8)$$

where $\boldsymbol{P}_i^+$ is the pseudo-inverse of $\boldsymbol{P}_i$ and $\boldsymbol{C}_i$ is the optical center of the i-th view. Furthermore, $\boldsymbol{C}_i$ is computed by $\boldsymbol{C}_i = (\boldsymbol{I}-\boldsymbol{P}_i^+\boldsymbol{P}_i)\boldsymbol{\omega}$, where $\boldsymbol{I}$ is the 4×4 identity matrix and $\boldsymbol{\omega}$ is an arbitrary 4-vector.

Therefore, the above six-point algorithm provides an alternative approach to estimating Fundamental Matrix, whose specific implementation will be detailed as follow.

3.2 Implementation of RANSAC-F-6

According to last sub-section, it's easy for us implement a RANSAC-F-7 (or RANSAC-8)-like algorithm using the six-point algorithm given more than two views.

The specific implementation is summarized as follows:

(1) Perform initial matching between each pair of the three images using correlation-based matching method [10].

(2) Extract the matches of all the three images by finding the common ones in the two-view matches obtained in Step (1).

(3) Invoke the RANSAC optimization, in which a specifiable $\boldsymbol{N}$ number of minimal subset samplings are repeated. While in each sampling,

(a) Select a random sample of 6 matches between three images and determine the projective matrices of the three images by using the six-point algorithm as described in sub-section 3.2.

(b) Identify the inliers in the putative matches by thresholding the re-projection error of the triangulated 3D projective space points in each image with respect to the determined projective matrices.

Remark: Since it's possible to obtain three real solutions from the cubic constraint over X and T, then at most three sets of projective matrices for the three images will be tested to identify the inliers. The one corresponding to the maximum number of the inliers is retained as the correct solution of the projective matrices corresponding to that sampling.

After N samplings, the set of projective matrices with the largest number of inliers is retrieved. In the case of ties, the one having lowest standard deviation of the re-projection errors is selected.

(4) Taking the obtained projective matrices and inliers as initial values, perform the bundle-adjustment to minimize the sum of the re-projection errors.

(5) Then from the optimized projective matrices, the fundamental matrices between each image pair are retrieved using equation (8).

4. Experiments and comparisons

To assess and compare the performance of the above three Fundamental Matrix automatic estimation methods, extensive experiments are performed.

4.1 Simulation

To conduct simulations, a synthetic image simulator is constructed by appropriately setting the positions and orientations of three cameras and the camera intrinsic parameters. Three synthetic images are generated by the projection the randomly distributed space points into the image planes of specified size. To account for the erroneous effect in practice, such as feature extraction, the coordinates of the image points are corrupted by the Zero-mean Gaussain noise. Furthermore, to investigate the ability of the tested methods to cope with outliers, a specified percentage of the originally correct matches between the three images are artificially changed into the outliers by disturbing the original matching order. It should be noted that, if the epipolar geometry constraint is used to guide the practical matching process, such as in Zhang's matching routine [10], some of the putative matches are outliers although they maybe still satisfy the epipolar geometry between the matched two images. As shown later, with respect to this kind of outliers, RANSAC-F-7 and RANSAC-F-8 will fail to identify them, while RANSAC-F-6 can.

With the synthetic images are generated, the three investigated RANSAC based Fundamental Matrix estimation methods are invoked in turn to estimate three Fundamental matrices between each image pair simultaneously or pair-wisely. According to the simulation results, we compared their respective performance of the investigated three methods with respect to the image noise level, the ratio of the outliers and other factors. One criterion adopted for evaluating the accuracy of the estimated fundamental matrix is the average symmetric epipolar line distance (hereafter is simply named as residual):

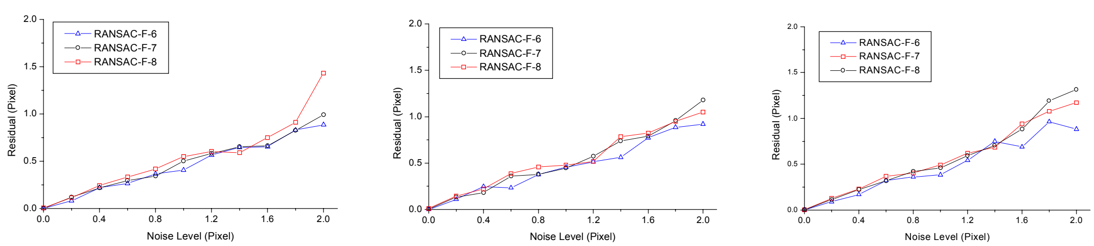
$$\frac{1}{2N}\sum_{i=1}^{N}\left(d(\boldsymbol{m}_i, \boldsymbol{F}^T\boldsymbol{m}_i')^2 + d(\boldsymbol{m}_i', \boldsymbol{F}\boldsymbol{m}_i)^2\right) \tag{9}$$

with respect to N "real image points" as suggested in [5]. Here by "real image points" we mean the original image points that are not contaminated by noise and outliers.

First we conduct the simulations for the cases of fewer image points and outliers. The simulation conditions are:

Synthetic image size: 640×480
Image points number: 100
Outlier percentage: 30%
Gassian noise level: 0.0 ~ 2.0 pixels
Experiments times at each noise level: 100

From the simulation results shown in Fig.2, we can see that, with image noise level increasing, the residuals of the three methods are all increasing linearly. RANSAC-F-6 is the best and RANSAC-F-8 algorithm is the worst. Even with the noise up to 2.0 pixels, the residual of RANSAC-F-6 can still be limited in 1.0 pixel in most cases.



$\boldsymbol{F}_{12}$ between Image 1 and 2 $\boldsymbol{F}_{23}$ between Image 2 and 3 $\boldsymbol{F}_{31}$ between Image 3 and 1

Fig.2 Curves of residual w.r.t the noise level

Just from the above comparison, the differences among such three methods are not so significant. Therefore, we take use of another maybe more appropriate criterion — the difference between the estimated Fundamental Matrix and its ground truth. A measure originally proposed by Stephane Laveau and recommended by Zhang in [4] is adopted. For details, please refer to [4]. Simply to say, this measure defines the difference between two Fundamental Matrices in terms of physical image distances. Smaller the "distance" between two Fundamental Matrices, more similar they are. For this, we directly take use of the software of "FDiff" from INRIA with the trail number set as 50000 to calculated the distances between the estimated Fundamental Matrices and their ground truths with respect to the three investigated methods respectively. The statistical results are illustrated in Fig. 3.

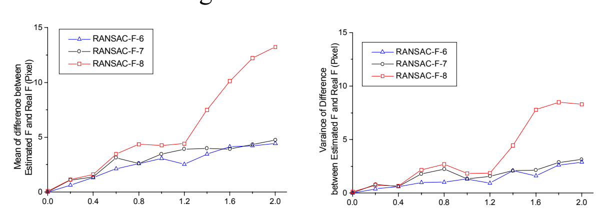


Fig.3 Distance between the estimated and the real F

From Fig.3, it is apparent that RANSAC-F-6 is the best, while RANAC-F-8 is the worst. For RANSAC-F-6, even at the noise level of 2.0 pixels, the mean distance can still limited in 3.5 pixels, which indicates that the estimation results are very closer to the ground truths. Additional extensive simulations with respect to many other different configurations are also performed, with the similar results obtained.

4.2 Real image experiments

After simulations, we conducted extensive experiments with real images. The matches between each image pairs are extracted by using Zhang's correlation based robust matching algorithm [10] first. Then common matches between all the three images are retrieved as the initial putative matches. After that, the three investigated methods are invoked in turn to estimate the Fundamental Matrices between image pairs (1-2, 2-3, 3-1) and identify the inliers (outliers).

First, we tested these three methods with the Valbonne Church image sequence as illustrated in Fig.4. In Fig.4, the first row is for illustrating the putative matches between the tested three frames. In each of them, the matches indicated by the colorful hollow squares are indexed. In the left-most one, the putative matches are additionally indicated through the lines linking the corresponding points. Then the lower two rows are for illustrating the estimation results of RANSAC-F-6 and RANSAC-F-7 respectively, in which the yellow lines are the epipolar lines, the solid red squares are the sampled minimal point set corresponding to the maximum inliers number, and the yellow solid rhombuses indicate the identified outliers. For the space reasons, we don't show the images corresponding to the estimation results of RANSAC-F-8, only with some measurements reflecting its accuracy listed in Table 1.

Another issue should be noticed is the way to overlap the epipolar lines. That is, for two images i and j, the epipolar lines corresponding to the Fundamental Matrix $\boldsymbol{F}_{ij}$ is overlapped on image j. In addition, since the three Fundamental Matrices between image pair 1-2, 2-3 and 3-1 are estimated pair-wisely when using RANSAC-F-7, different from RANSAC-F-6, which estimates the Fundamental Matrices for three images simultaneously, the identified outliers by RANSAC-F-7 in each image might be inconsistent.

By investigating the images in the first row of Fig.4, 7 outliers (indexed as 32, 50, 51, 83, 88, 96, 98) out of 98 putative matches are identified by naked eyes carefully. Then as shown in the second row, these outliers are all identified correctly by RANSAC-F-6. But it also should note that three intuitively correct matches (indexed as 62, 80, 87) are also identified as outliers, maybe caused by the maybe a little stringent threshold. Then for RANSAC-F-7, since the estimations are performed pair-wisely, the identified outliers are not consistent for each image. It's apparent that some obvious outliers are missed to identify, for example, the 32-th image point in the first image and the 98-th one in the second image.

In addition, from some experiment results collected in Table1, from which we can also find that RANSAC-F-8 is worse than both of RANSAC-F-6 and RANSAC-F-7 in ability to minimize the residual and to identify the outliers.

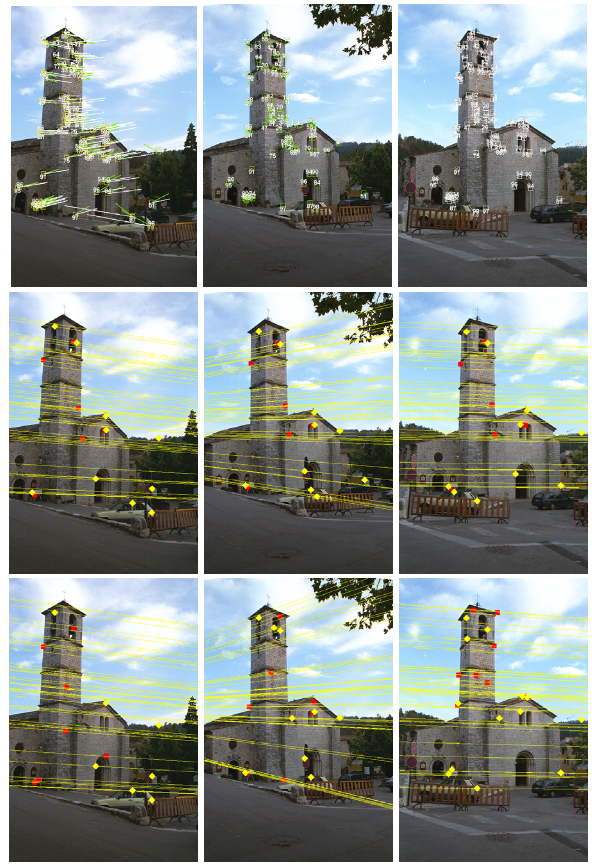


Fig.4 Results of the Valbonne Church image sequence

After the above experiment, we also conducted similar experiments with many other image sequences, one of which is shown in Fig.5, with same illustration theme used. By investigating the putative matches shown on the first row of Fig.5, the putative matches indexed as 9, 31, 34, 37 40, 42 48, 56 are found being matched wrongly. For RANSAC-F-6, these outliers are all identified correctly. While on the other hand, for RANSAC-F-7, it still can't identify all the outliers correctly, even with a more stringent inliers identification threshold. For example, for the image pair 1-2, there is even no a single outlier is identified. So it can be seen again that RANSAC-F-6 has a better ability to identify the outliers than RANSAC-F-7, especially when the images are captured with considerable camera rotation motions. However, for RANSAC-F-7, we can see that the estimated epipolar geometry is not so different from the one estimated by RANSAC-F-6. This is because that the putative matches are obtained with epipolar geometry considered for guidance in the matching process. To some extent, the outliers are fairly well consistent to the real epipolar geometry.

		Residual (pixel)			Number of Identified Inliers		
		$\boldsymbol{F}_{12}$	$\boldsymbol{F}_{23}$	$\boldsymbol{F}_{31}$	Image1	Image2	Image3
RANSAC-F-6	Sequence1	0.19	0.21	0.30	88	88	88
	Sequence2	0.27	0.26	0.30	60	60	60
RANSAC-F-7	Sequence1	0.24	0.26	0.29	89	85	91
	Sequence2	0.39	0.27	0.28	70	64	64
RANSAC-F-8	Sequence1	0.28	0.25	0.32	94	92	89
	Sequence2	0.32	0.31	0.34	69	62	63

Table. 1 Results of the experiments of the two image sequence

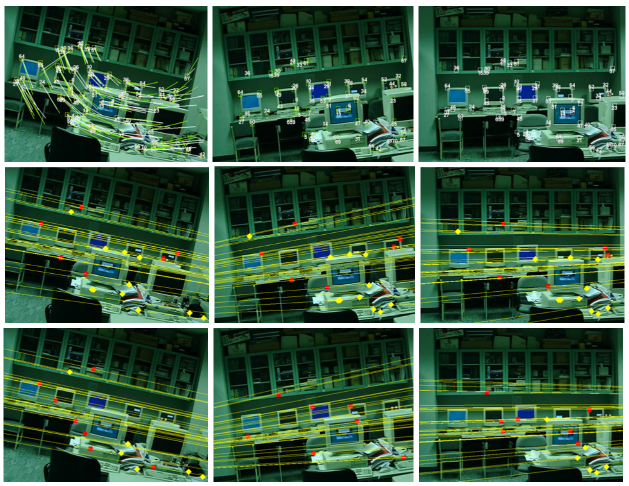


Fig.8 Results of a indoor scene image sequence

5. Conclusions and future works

The main contribution of this paper is to determine a most accurate and robust method for the automatic estimation of the Fundamental Matrix when three views are available, through extensive analysis and comparison of three popular methods by carefully designed experiments. These methods include the seven-point method, the eight-point method and the six-point method. The six-point algorithm, which originally arises from three-view geometry, is implemented to estimate the Fundamental Matrix in this paper. RANSAC is used in all these methods to reject outliers. The six-point algorithm is found to be the best. This conclusive and convincing result is a valuable piece of information for researchers in computer vision. Further more, it should also be noted that the above comparison is not an entirely fair comparison as only two views are used in the seven-point algorithm and the eight-point algorithm while three views are used in the six-point algorithm. This could be rectified in our future research. As we know, in addition to the direct three-view based approach using the six-point algorithm, the three-view geometry can also be determined from the three pair-wisely estimated Fundamental Matrices in a two-step style, with the seven-point algorithm and eight-point algorithm used in the first step. In fact, it is obvious that this future work can also be formulated as a comparison of the different approaches for robust estimation of a three-view geometry.

6. Acknowledgements

This research is supported by Hong Kong RGC research grand CRC 4/98, Chinese “973” Program (G1998030502) and the National Science Foundation of China (60033010, 69975021).

7. References

[1] L. Quan, *Invariants of six points and projective reconstruction from three uncalibrated images*, IEEE Transactions on Pattern Analysis and Machine Intelligence", Vol. 17, pp. 34-46, 1995.

[2] F. Schaffalitzky, A. Zisserman, R. I. Hartley, P. H. S. Torr, *A Six Point Solution for Structure and Motion*, ECCV (1) 2000: 632-648

[3] R. I. Hartley, N. Y. Dano, *Reconstruction from six-point sequences*, CVPR(2)2000:480-486.

[4] Z. Zhang, *Determining the epipolar geometry and its uncertainty: A review*, Tech. Rep. Technical Report 2927, INRIA, Sophia-Antipolis, France, July 1996.

[5] P. H. S. Torr and D. W. Murray, *The development and comparison of robust methods for estimating the fundamental matrix*, International Journal of Computer Vision, 24(3): 271--300, 1997.

[6] P. Torr and A. Zisserman, *Robust parameterization and computation of the trifocal tensor*, Image and Vision Computing, pages 591--605, 1997.

[7] H.C. Longuet-Higgins, *A computer algorithm for reconstructing a scene from two projections*, Nature, 293:133--135, 1981.

[8] Fischler, M. A. and Bolles, R. C, *Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography*, Communications of the ACM, 24:381—395, 1981.

[9] R. I. Hartley, *In Defense of the 8-point Algorithm*, IEEE-T PAMI 19:6, pp.580-593, 1997.

[10] 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, 75(1-2):87--120, 1995.

[11] Stefan Carlsson and Daphna Weinshall, *Dual computation of projective shape and camera positions from multiple images*, International Journal of Computer Vision, V27 (3), 1998.

[12] R. Hartley and P. Sturm, *Triangulation*, Computer Vision and Image Understanding, Vol. 68, No. 2, 1997.

[13] Press: R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, 2000.