Estimating the Fundamental Matrix by Transforming Image Points in Projective Space[1]

Zhengyou Zhang and Charles Loop
Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA
E-mail: {zhang,cloop}@microsoft.com

This paper proposes a novel technique for estimating the fundamental matrix by transforming the image points in projective space. We therefore only need to perform nonlinear optimization with one parameterization of the fundamental matrix, rather than considering 36 distinct parameterizations as in previous work. We also show how to preserve the characteristics of the data noise model from the original image space.

Key Words: fundamental matrix, epipolar geometry, projective transformation

1. INTRODUCTION

Two perspective images of a single rigid object/scene are related by the so-called epipolar geometry, which can be described by a 3×3 singular matrix known as fundamental matrix. Robust and accurate estimation of the fundamental matrix is very important in many applications of 3D computer vision, and has been the focus of many researchers [4, 5, 6, 7]. Previous work on nonlinear estimation of the fundamental matrix requires the consideration of 36 distinct parameterizations to account for the fact that an epipole may be at infinity and an element of the epipolar transformation may be equal to 0. This leads to a cumbersome implementation of the optimization procedure.

This work proposes a novel technique for estimating the fundamental matrix. Instead of using 36 maps to parameterize the fundamental matrix, we transform the image points in projective space, and in turn we only need to perform nonlinear optimization with one parameterization of the fundamental matrix. We also show how to preserve the characteristics of the data noise model from the original image space.

2. PARAMETERIZATION OF THE FUNDAMENTAL MATRIX

This section reviews the parameterization of the fundamental matrix.

[1] Published in **Computer Vision and Image Understanding**, Vol.82, No.2, pages 174–180, May 2001

2.1. Epipolar geometry

The epipolar geometry exists between any two camera systems. For a point $\mathbf{m}_i$ in the first image, its correspondence in the second image, $\mathbf{m}'_i$, must lie on the epipolar line in the second image. This is known as the *epipolar constraint*. Algebraically, in order for $\mathbf{m}_i$ and $\mathbf{m}'_i$ to be matched, the following equation must be satisfied:

$$\widetilde{\mathbf{m}}_i'^T \mathbf{F} \widetilde{\mathbf{m}}_i = 0 \; , \tag{1}$$

where $\mathbf{F}$, known as the *fundamental matrix*, is a 3×3 matrix of rank 2 (i.e., $\det(\mathbf{F}) = 0$), defined up to a scale factor; $\widetilde{\mathbf{m}}_i$ and $\widetilde{\mathbf{m}}'_i$ are homogeneous coordinates of points $\mathbf{m}_i$ and $\mathbf{m}'_i$, i.e., $\widetilde{\mathbf{m}}_i = [\mathbf{m}_i^T, 1]^T$ and $\widetilde{\mathbf{m}}'_i = [\mathbf{m}_i'^T, 1]^T$.

2.2. Parameterization based on the epipolar transformation

There are several possible parameterizations for the fundamental matrix [4], e.g., we can express one row (or column) of the fundamental matrix as the linear combination of the other two rows (or columns). One possibility is the following:

$$\mathbf{F} = \begin{bmatrix} a & b & -ax - by \\ c & d & -cx - dy \\ -ax' - cy' & -bx' - dy' & F_{33} \end{bmatrix} \tag{2}$$

$$\text{with} \quad F_{33} = (ax + by)x' + (cx + dy)y' \; .$$

Here, x and y are the coordinates of the first epipole, x' and y' are the coordinates of the second epipole, and a, b, c and d, defined up to a scale factor, parameterize the epipolar transformation mapping an epipolar line in the first image to its corresponding epipolar line in the second image.

However, this parameterization does not work if an epipole is at infinity. This is because in that case at least one of x, y, x' and y' has a value of infinity. In order to overcome this problem, INRIA group [1, 7] has proposed to use in total 36 maps to parameterize the fundamental matrix, as summarized below.

2.3. 36 maps to parameterize the fundamental matrix

Let us denote the columns of $\mathbf{F}$ by the vectors $\mathbf{c}_1$, $\mathbf{c}_2$ and $\mathbf{c}_3$. The rank-2 constraint on $\mathbf{F}$ is equivalent to the following two conditions:

$$\exists \lambda_1, \lambda_2 \quad \text{such that} \quad \mathbf{c}_{j_0} + \lambda_1 \mathbf{c}_{j_1} + \lambda_2 \mathbf{c}_{j_2} = 0 \tag{3}$$

$$\nexists \lambda \quad \text{such that} \quad \mathbf{c}_{j_1} + \lambda \mathbf{c}_{j_2} = 0 \tag{4}$$

for $j_0, j_1, j_2 \in \{1, 2, 3\}$, where λ_1, λ_2 and λ are scalars. Condition (4), as a non-existence condition, cannot be expressed by a parameterization: we shall only keep condition (3) and so extend the parameterized set to all the 3×3-matrices of rank strictly less than 3. Indeed, the rank-2 matrices of, for example, the following forms:

$$[\mathbf{c}_1 \; \mathbf{c}_2 \; \lambda\mathbf{c}_2] \quad \text{and} \quad [\mathbf{c}_1 \; \mathbf{0}_3 \; \mathbf{c}_3] \quad \text{and} \quad [\mathbf{c}_1 \; \mathbf{c}_2 \; \mathbf{0}_3]$$

do not have any parameterization if we take $j_0 = 1$. A parameterization of $\mathbf{F}$ is then given by $(\mathbf{c}_{j_1}, \mathbf{c}_{j_2}, \lambda_1, \lambda_2)$. This parameterization implies to divide the parameterized set among three maps, corresponding to $j_0 = 1$, $j_0 = 2$ and $j_0 = 3$.

If we construct a 3-vector such that λ_1 and λ_2 are the ${j_1}^{\text{th}}$ and ${j_2}^{\text{th}}$ coordinates and 1 is the ${j_0}^{\text{th}}$ coordinate, then it is obvious that this vector is the eigenvector of $\mathbf{F}$, and is thus the epipole in the case of the fundamental matrix. Using such a parameterization implies to compute directly the epipole which is often a useful quantity, instead of the matrix itself.

To make the problem symmetrical and since the epipole in the other image is also worth being computed, the same decomposition as for the columns is used for the rows, which now divides the parameterized set into 9 maps, corresponding to the choice of a column and a row as linear combinations of the two columns and two rows left. A parameterization of the matrix is then formed by the two coordinates x and y of the first epipole, the two coordinates x' and y' of the second epipole and the four elements a, b, c and d left by $\mathbf{c}_{i_1}$, $\mathbf{c}_{i_2}$, $\mathbf{l}_{j_1}$ and $\mathbf{l}_{j_2}$, which in turn parameterize the epipolar transformation mapping an epipolar line of the first image to its corresponding epipolar line in the second image. The parameterization shown in (2) corresponds to the case $i_0 = 3$ and $j_0 = 3$.

At last, to take into account the fact that the fundamental matrix is defined only up to a scale factor, the matrix is normalized by dividing the four elements (a, b, c, d) by the largest in absolute value. We have thus in total 36 maps to parameterize the fundamental matrix.

3. A NOVEL TECHNIQUE FOR ESTIMATING THE FUNDAMENTAL MATRIX

As said in the last section, in order to deal with all possibilities, we have to consider 36 distinct parameterizations, and this is a burden in implementing the optimization procedure. In this section, we propose to transform the image points in a projective space and work only with one parameterization such as (2) while ensuring the value of a is largest so we can set $a = 1$.

The idea is to find a projective transformation in each image, denoted by $\mathbf{P}$ and $\mathbf{P}'$, such that *in the transformed image space the first element of the fundamental matrix has the largest value and the epipoles are not at infinity*. Let the image points in the transformed space be

$$\widetilde{\widehat{\mathbf{m}}}_i = \mathbf{P}\tilde{\mathbf{m}}_i \quad \text{and} \quad \widetilde{\widehat{\mathbf{m}}}_i' = \mathbf{P}'\tilde{\mathbf{m}}_i' \ , \tag{5}$$

then the fundamental matrix in the transformed space is given by

$$\widehat{\mathbf{F}} = \mathbf{P}'^{-T}\mathbf{F}\mathbf{P}^{-1} \ . \tag{6}$$

Given an initial estimate of matrix $\mathbf{F}_0$ obtained for example with Hartley's normalized 8-point algorithm [3], we compute the epipoles $\widetilde{\mathbf{e}}_0$ and $\widetilde{\mathbf{e}}_0'$. The matrices $\mathbf{P}$ and $\mathbf{P}'$ are 3×3 permutation matrices determined as follows:

1. Initialize $\mathbf{P}$ and $\mathbf{P}'$ to be identity matrices.

2. Find the position of the largest element of $\mathbf{F}_0$, denoted by (i_0, j_0). (Index of a vector or a matrix starts with 0 as in C++)

3. If $j_0 \neq 0$, permute rows 0 and j_0 of matrix $\mathbf{P}$ and permute elements 0 and j_0 of epipole $\widetilde{\mathbf{e}}_0$.

4. If $i_0 \neq 0$, permute rows 0 and i_0 of matrix $\mathbf{P}'$ and permute elements 0 and i_0 of epipole $\widetilde{\mathbf{e}}_0'$.

5. If $|\widetilde{\mathbf{e}}_0[1]| > |\widetilde{\mathbf{e}}_0[2]|$, permute elements 1 and 2 of epipole $\widetilde{\mathbf{e}}_0$ and permute rows 1 and 2 of matrix $\mathbf{P}$.

6. If $|\widetilde{\mathbf{e}}'_0[1]| > |\widetilde{\mathbf{e}}'_0[2]|$, permute elements 1 and 2 of epipole $\widetilde{\mathbf{e}}'_0$ and permute rows 1 and 2 of matrix $\mathbf{P}'$.

Steps 3 and 4 ensure that the first element of the fundamental matrix in the transformed space has the largest value, while steps 5 and 6 ensure that the epipoles are not at infinity. Note that $\mathbf{P}^{-1} = \mathbf{P}^T$ and $\mathbf{P}'^{-1} = \mathbf{P}'^T$ because they are permutation matrices. The reader is referred to [2, page 109] for efficiently representing a permutation matrix with an integer vector.

We can now use the parameterization (2) to estimate $\widehat{\mathbf{F}}$ from the transformed image points $\widetilde{\widetilde{\mathbf{m}}}_i$ and $\widetilde{\widetilde{\mathbf{m}}}'_i$. The fundamental matrix in the original image space is given by $\mathbf{F} = \mathbf{P}'^T\widehat{\mathbf{F}}\mathbf{P}$.

4. PRESERVING INFORMATION OF THE ORIGINAL NOISE DISTRIBUTION

Points are detected in the original image space. All nonlinear optimization criteria (see [7]) are derived from the noise distribution of these points. One reasonable assumption is that the image points are corrupted by independent and identically distributed Gaussian noise with mean zero and covariance matrices given by

$$\mathbf{\Lambda}_{\mathbf{m}_i} = \mathbf{\Lambda}_{\mathbf{m}'_i} = \sigma^2 \operatorname{diag}(1,1) \ , \tag{7}$$

where σ is the noise level, which is usually unknown. This assumption is clearly no longer reasonable in the transformed image space. In the following, we show how to estimate the transformed fundamental matrix while preserving information of the original noise distribution, and we will illustrate this by using the gradient-weighted criterion.

From (1), (5) and (6), we see that the transformed fundamental matrix and the original image points are related by

$$f_i \equiv \widetilde{\mathbf{m}}_i'^T \mathbf{P}'^T \widehat{\mathbf{F}} \mathbf{P} \widetilde{\mathbf{m}}_i = 0 \ . \tag{8}$$

The least-squares technique produces an optimal solution if each term has the same variance. Therefore, we can estimate the transformed fundamental matrix by minimizing the following weighted sum of squares (the chi-square χ^2):

$$\min_{\widehat{\mathbf{F}}} \sum_i f_i^2/\sigma_{f_i}^2 \ , \tag{9}$$

where σ_{f_i} is the variance of f_i, and its computation is given below.

Let us introduce the notation

$$\mathbf{Z} = \operatorname{diag}(1,1,0) \ ,$$

then from (7), the covariances of the *homogeneous coordinates* of the image points are given by

$$\mathbf{\Lambda}_{\widetilde{\mathbf{m}}_i} = \mathbf{\Lambda}_{\widetilde{\mathbf{m}}'_i} = \sigma^2 \mathbf{Z} \ . \tag{10}$$

Under first order approximation, the variance of f_i is then given by

$$\begin{aligned} \sigma_{f_i}^2 &= \left(\frac{\partial f_i}{\partial \widetilde{\mathbf{m}}_i}\right)^T \mathbf{\Lambda}_{\widetilde{\mathbf{m}}_i} \frac{\partial f_i}{\partial \widetilde{\mathbf{m}}_i} + \left(\frac{\partial f_i}{\partial \widetilde{\mathbf{m}}'_i}\right)^T \mathbf{\Lambda}_{\widetilde{\mathbf{m}}'_i} \frac{\partial f_i}{\partial \widetilde{\mathbf{m}}'_i} \\ &= \sigma^2 (\hat{\mathbf{l}}_i^T \mathbf{P} \mathbf{Z} \mathbf{P}^T \hat{\mathbf{l}}_i + \hat{\mathbf{l}}_i'^T \mathbf{P}' \mathbf{Z} \mathbf{P}'^T \hat{\mathbf{l}}'_i) \end{aligned}$$

where $\hat{\mathbf{l}}'_i = \widehat{\mathbf{F}}\widetilde{\widetilde{\mathbf{m}}}_i$ and $\hat{\mathbf{l}}_i = \widehat{\mathbf{F}}^T\widetilde{\widetilde{\mathbf{m}}}_i'$ are epipolar lines in the transformed image space. Since multiplying each term by a constant does not affect the minimization, the problem (9) becomes

$$\min_{\widehat{\mathbf{F}}} \sum_i \frac{(\widetilde{\widetilde{\mathbf{m}}}_i'^T \widehat{\mathbf{F}} \widetilde{\widetilde{\mathbf{m}}}_i)^2}{\hat{\mathbf{l}}_i^T \mathbf{P}\mathbf{Z}\mathbf{P}^T \hat{\mathbf{l}}_i + \hat{\mathbf{l}}_i'^T \mathbf{P}'\mathbf{Z}\mathbf{P}'^T \hat{\mathbf{l}}_i'} . \tag{11}$$

The denominator is simply the gradient of f_i. The minimization can be conducted by means of, for example, the Levenberg-Marquardt algorithm.

5. EXPERIMENTAL RESULT

In this section, we show an example with real data as displayed in Fig. 1. The automatic robust image-matching algorithm described in [8] was used to find point matches across the two views. 217 matches have been found, which are shown in Fig. 1. The matching algorithm also provides an estimate of the fundamental matrix based on the least-median-squares (LMedS) technique. This estimate is used as the initial guess for the experiment described in this section.

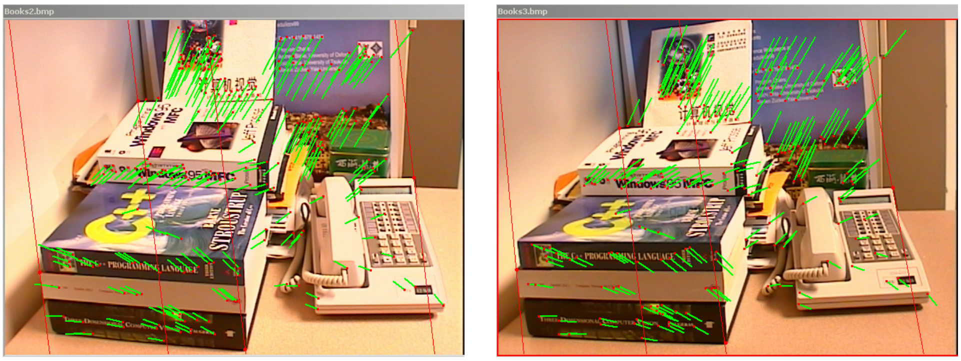

FIG. 1. Automatically matched points and four pairs of epipolar lines. A point match is indicated by a green line segment, with one end, indicated by a red dot, being the point in the current image, and the other end being the matched point in the other image.

TABLE 1

Comparison of different methods, in pixels

Method	epipolar distances image 1	image 2	epipole 1 (u,	v)	epipole 2 (u,	v)
Initial	0.587	0.582	−918	−11272	−169	−3319
Method 1	0.471	0.467	−2797	−26210	−315	−4104
Method 2	0.504	0.499	−682	−8430	−140	−2915
Method 3	0.469	0.465	22973	189570	−497	−5308

We have compared four methods, all based on the gradient-weighted criterion.

- Method 1: only one parameterization of the fundamental matrix, as shown in (2), is used.
- Method 2: the novel technique described in Sect. 3 is used; however, the original data noise information is not preserved.

- Method 3: the novel technique together with the preservation of the original noise information as described in Sect. 4 is used.
- Method 4: 36 distinct parameterizations, as described in [7], are used.

Not surprisingly, Method 4 produced exactly the same result as Method 3, and the results are therefore not shown. The comparison of the other three methods is shown in Table 1, where we show the average distances from a point to its corresponding epipolar line in each image, and the position of both epipoles. The first row displays the results provided by the LMedS technique during image matching. We see that Method 1 gives quite a good result, but it is also clear that the parameterization used has difficulty to represent epipoles located far away from the image. Method 2 does not produce very good result because it does not take into account appropriately the original noise information. Method 3 gives the best result.

6. CONCLUSION

In this paper, we have proposed a novel technique for estimating the fundamental matrix. We transform the image points in the projective space, and therefore only need to perform nonlinear optimization with one parameterization of the fundamental matrix, rather than considering 36 maps of parameterization as in previous work. In addition, we have presented a method for preserving the characteristics of the data noise model from the original image space. The implementation of the proposed technique is much easier than with the previous approach, and the same quality has been obtained.

REFERENCES

1. G. Csurka, C. Zeller, Z. Zhang, and O. Faugeras. Characterizing the uncertainty of the fundamental matrix. *Computer Vision and Image Understanding*, 68(1):18–36, October 1997.
2. G.H. Golub and C.F. van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, Maryland, 3 edition, 1996.
3. R.I. Hartley. In defence of the 8-point algorithm. In *Proceedings of the 5th International Conference on Computer Vision*, pages 1064–1070, Boston, MA, June 1995. IEEE Computer Society Press.
4. Quang-Tuan Luong. *Matrice Fondamentale et Calibration Visuelle sur l'Environnement-Vers une plus grande autonomie des systèmes robotiques*. PhD thesis, Université de Paris-Sud, Centre d'Orsay, December 1992.
5. Quang-Tuan Luong and Olivier Faugeras. The fundamental matrix: theory, algorithms, and stability analysis. *The International Journal of Computer Vision*, 17(1):43–76, January 1995.
6. Phil Torr, Andrew Zisserman, and Steven Maybank. Robust detection of degenerate configurations for the fundamental matrix. In *Proceedings of the 5th International Conference on Computer Vision*, pages 1037–1042, Boston, MA, June 1995. IEEE Computer Society Press.
7. Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *The International Journal of Computer Vision*, 27(2):161–195, 1998.
8. 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*, 78:87–119, October 1995.