Robust Fundamental Matrix Determination without Correspondences

Stefan Lehmann,[†] I. Vaughan L. Clarkson,[*†] Andrew P. Bradley,[‡] John Williams,[†] and Peter J. Kootsookos[ℭ]
[‡]Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP)
[†]School of Information Technology and Electrical Engineering
The University of Queensland, 4072 AUSTRALIA
[ℭ] UTRC, MS129-15, 411 Silver Lane, CT 06108, USA
(lehmann,v.clarkson,bradley, jwilliams,kootsoop)@itee.uq.edu.au

Abstract

Estimation of the fundamental matrix is key to many problems in computer vision as it allows recovery of the epipolar geometry between camera images of the same scene. The estimation from feature correspondences has been widely addressed in the literature, particularly in the presence of outliers. In this paper, we propose a new robust method to estimate the fundamental matrix from two sets of features without any correspondence information. The method operates in the frequency domain and the underlying estimation process considers all features simultaneously, thus yielding a high robustness with respect to noise and outliers. In addition, we show that the method is well-suited to widely separate viewpoints.

1. Introduction

One of the main objectives of computer vision is the recovery of structure and motion information from a sequence of camera images. The determination of the fundamental matrix plays a key role in this context since it allows the computation of the underlying epipolar geometry. A variety of methods have been proposed to compute the fundamental matrix from point correspondences in stereo images. A comprehensive overview is given in [4]. However, the identification of these correspondences remains a fundamental problem. The sensitivity to noise and outliers of classical approaches to the estimation of the fundamental matrix is well-known [12].

The estimation of the fundamental matrix without correspondences remains largely unaddressed in the literature [3]. Some methods deal with the case of correct but incomplete correspondence information by extending a minimum set of features into a complete set covering all reconstructible features [9]. Alternatively, occluded features are artificially generated by projecting computed 3D feature coordinates onto computed camera positions [10]. However, both of these methods rely on the prior knowledge of a correct set of initial correspondences. Other approaches tackle the correspondence problem by using geometrical constraints, such as in [5], where geometric rank constraints are used to facilitate the optical flow computation over closely-spaced views. In [3], a method is proposed that relies on the Expectation-Maximization (EM) algorithm to iteratively estimate structure and motion without correspondences. At each iteration, a new structure from motion problem is solved for virtual measurements that are derived from a probability distribution. This probability distribution is iteratively refined over the set of correspondences. It is acknowledged that results for occlusions or spurious features have not been demonstrated and that the EM algorithm can converge to a local minimum.

In this paper, we propose a method to estimate the fundamental matrix from two sets of features without the need for correspondences. The two sets of features are the 2D orthographic projections of a set of 3D object features from different viewpoints. Our method deduces motion parameters without correspondences by evaluating the frequency spectra of the 2D feature spaces. The approach is based on an integral projection model and has previously been applied to estimating 3D rigid body transformations based on raw images [7]. Here, we extend this work to feature correspondences. The estimation process considers all features simultaneously, making the method robust with respect to noise and outliers.

2. Integral Projection of Sparse Features

2.1. Concept and relationship to parallel projection

We will illustrate the integral projection scheme based on a set of 2D features that we project into 1D. The integral projection model determines the 1D projection values by integrating the 2D feature scene along lines that run parallel to the view axis. Due to the duality between *Structure from Motion* and *Motion from Structure*, recording static scenes with multiple cameras from different viewpoints is equivalent to recording dynamic scenes with one static camera. Suppose we have a 2D object that is represented by a number of 2D feature points in both the original and the transformed position. Integrating along lines that are parallel to the view axis results in the 1D feature projections. Figures 1(a) and 1(b) depict this situation for five features. In these figures, the view axis is the z-axis of the scene coordinate system. The integral projections are denoted by

[*] Vaughan Clarkson is currently on study leave at the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

$I(y)$ and $I'(y)$ respectively. Let us assume that all 2D features are located on the object surface and are in the field of view of an orthographic camera. Under these assumptions, integral projection is identical to parallel projection.

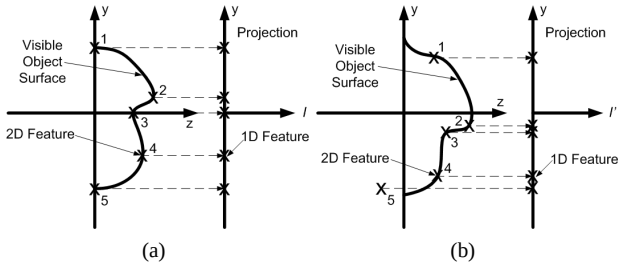


Figure 1. Integral projection of feature set (a) before and (b) after transformation

For general object geometries, object faces including their corresponding features might become obscured or revealed after the scene transformation. Therefore, the visible object surfaces before and after the transformation are generally not only transformed versions but differently shaped. There will consequently be deviations between applying an integral projection or a parallel projection to the transformed scene as Feature 5 in Figure 1(b) illustrates.

2.2. Mathematical Model

In our model, we describe each 3D feature by a Dirac function at the appropriate feature location. Thus, assuming N features, our 3D feature space is represented by:

$$f_3(x,y,z) = \sum_{k=1}^{N} \delta(x-x_k, y-y_k, z-z_k) \qquad (1)$$

where (x_k, y_k, z_k) are the individual feature locations. Integral projection determines the 2D feature projections by integrating $f_3(x,y,z)$ along lines that are running parallel to the view axis of the camera. Using our integral projection approach and assuming that the z-axis of our scene coordinate system is aligned with the view axis of the camera, the corresponding 2D projection data will be

$$f_2(x,y) = \int_{\mathbb{R}} f_3(x,y,z)\, dz = \sum_{k=1}^{N} \delta(x-x_k, y-y_k) \quad (2)$$

The Fourier spectra of $f_2(x,y)$ and $f_3(x,y,z)$ can be denoted as

$$\begin{aligned} F_2(\xi,\eta) &= \int_{\mathbb{R}^2} f_2(x,y) e^{-j(\xi x+\eta y)}\, dx\, dy \\ &= \sum_{k=1}^{N} e^{-j(\xi x_k+\eta y_k)} \qquad (3) \\ F_3(\xi,\eta,\zeta) &= \int_{\mathbb{R}^3} f_3(x,y,z) e^{-j(\xi x+\eta y+\zeta z)}\, dx\, dy\, dz \\ &= \sum_{k=1}^{N} e^{-j(\xi x_k+\eta y_k+\zeta z_k)} \qquad (4) \end{aligned}$$

Thus, the relationship between the Fourier spectra $F_2(\xi,\eta)$ and $F_3(\xi,\eta,\zeta)$ is:

$$\begin{aligned} F_2(\xi,\eta) &= \int_{\mathbb{R}^3} f_3(x,y,z) e^{-j(\xi x+\eta y)}\, dx\, dy\, dz \\ &= F_3(\xi,\eta,0). \qquad (5) \end{aligned}$$

Equation 5 is known as the projection-slice theorem and is commonly used in X-ray tomographic reconstruction [1]. As (5) shows, the spectrum of the projected feature data is the $\xi,\eta-$plane (where $\zeta = 0$) of the corresponding spectrum of the 3D feature data. Next, we will discuss what this relationship means for stereo projections of a 3D feature set.

2.3. Effect on stereo projections

Let us now consider that we have two 2D integral projections of our 3D feature data from two cameras at different viewpoints. However, to simplify the derivations, we consider the equivalent case of projecting the original and a transformed set of 3D features onto one camera projection plane instead by applying the rigid-body transformation T_{scene}. We decompose T_{scene} into a rotation matrix R and a translation vector Λ with

$$\Lambda = (x_0, y_0, z_0)^T, \qquad (6)$$

where x_0, y_0 and z_0 are the translational components of the scene transformation with respect to the x,y,z coordinate axes of the scene coordinate system. Therefore, T_{scene} transforms each 3D feature point $P = (x,y,z)^T$ into $P' = (x',y',z')^T$ according to the following equation:

$$P' = RP + \Lambda. \qquad (7)$$

We now introduce the vector

$$\Delta = (\xi,\eta,\zeta)^T \qquad (8)$$

where ξ,η and ζ represent the 3D frequency components of $F_3(\xi,\eta,\zeta)$ in (4). Using the vectors from (6) and (8), the 3D spectrum that corresponds to the transformed scene is

$$F_3'(\Delta) = e^{-j(\Lambda^T\Delta)} F_3(R^T\Delta). \qquad (9)$$

According to the projection-slice theorem (5), the 2D spectrum $F_2(\xi,\eta)$ is the $\zeta = 0$ plane of the 3D spectrum of the object. Therefore, the two spectra of the projections before and after the scene transformation show matching lines that run through the origin of the coordinate systems of the spectra. The magnitudes of the two spectra along these lines will be identical, while the phases will show an offset which depends upon the translational component of the transformation. Here we propose a method for detecting matching lines in the 2D Fourier spectra as this give us valuable information on the 3D scene transformation T_{scene} that has taken place.

2.4. Analysis of the transformation parameters

Let us assume that (ξ, η) and (ξ', η') are the corresponding frequency locations along the matching lines of the spectra $F_2(\xi, \eta)$ and $F_2'(\xi', \eta')$ respectively. Equation 5 yields the following relationship

$$F_2'(\xi', \eta') = F_3'(\xi', \eta', 0) \tag{10}$$

and finally with (9)

$$F_2'(\xi', \eta') = e^{j(\xi' x_0 + \eta' y_0)} F_2(\xi, \eta). \tag{11}$$

Let us introduce the matching line angle pair $(\alpha,\ \alpha')$ with respect to the ξ- and ξ'-axes of the frequency spectra $F_2(\xi, \eta)$ and $F_2'(\xi', \eta')$ respectively. The values of the 2D spectra $F_2(\xi, \eta)$ and $F_2'(\xi', \eta')$ along the matching lines can now be transformed into one-dimensional representations $F_1(\rho)$ and $F_1'(\rho)$ where ρ denotes a 1D frequency index. Thus, (11) can be transformed into:

$$F_1'(\rho) = e^{j\rho\,\sigma}\, F_1(\rho) \tag{12}$$

where we define the displacement, σ, as

$$\sigma = x_0\, \cos\alpha' + y_0\, \sin\alpha'. \tag{13}$$

Detecting the matching lines in the two 2D spectra therefore yields two types of information. Firstly, by recovering the displacement σ we gain information about the translational components x_0 and y_0. Even though x_0 and y_0 can not be isolated from σ, we can reveal information about their relationship. Secondly, additional information about the scene transformation is contained in the angle pair $(\alpha,\ \alpha')$ itself.

To discuss this in more detail, we will examine how (α, α') depend on the rotation of the 3D feature scene. Let us assume that the scene has been rotated by the angles θ, ϕ and ρ around the x-, y- and z axis respectively according to the following rotation matrix:

$$\begin{aligned} R &= R_z^{(\rho)}\, R_y^{(\phi)}\, R_x^{(\theta)} \\ &= \begin{pmatrix} \cos\phi\cos\rho & -\cos\theta\sin\rho + \sin\theta\sin\phi\,\cos\rho & \sin\theta\sin\rho + \cos\theta\sin\phi\cos\rho \\ \cos\phi\sin\rho & \cos\theta\cos\rho + \sin\theta\,\sin\phi\sin\rho & -\sin\theta\cos\rho + \cos\theta\sin\phi\sin\rho \\ -\sin\phi & \sin\theta\cos\phi & \cos\theta\cos\phi \end{pmatrix} \end{aligned} \tag{14}$$

The rotation matrix R transforms each 3D feature $(x, y, z)^T$ into a corresponding feature $(x', y', z')^T$ according to (7). Equation 9 shows that R also establishes the transformation between corresponding frequency indices in the 3D Fourier spaces of the original and the transformed scene. According to (5) and (10), the transformation of the frequency pair ξ and η into the corresponding matched frequencies ξ', η' is described by

$$(\xi', \eta', 0)^T = R \cdot (\xi, \eta, 0)^T \tag{15}$$

This yields a relationship between ξ and η along the matching line dependent upon the angles θ, ϕ and ρ. Thus, the angle α of the matching line can be found. Similarly, the angle α' can be determined from

$$(\xi, \eta, 0)^T = R^T \cdot (\xi', \eta', 0)^T. \tag{16}$$

Since the two equations for (α, α') depend on three rotation angles θ, ϕ, ρ, this problem is not invertible in the general case. In other words, various sets of 3D rotation angles yield the same matching line angles (α, α'). However, the rotation matrix R can be determined from the matching line angle pair (α, α') up to an unknown rotation parameter τ as follows:

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} = R_z^{\alpha'}\, R_x^{\tau}\, R_z^{-\alpha} \tag{17}$$

where $R_z^{\alpha'}, R_z^{-\alpha}$ are rotation matrices that rotate around the z-axis at angles α' and $(-\alpha)$ respectively and R_x^{τ} rotates around the x-axis about an unknown angle τ. Therefore, assuring that the orientations of the rotations are consistent with the ones in (14), the parameters of R specified in (17) are:

$$\begin{aligned} r_{11} &= \cos\alpha'\cos\alpha + \sin\alpha'\cos\tau\sin\alpha \\ r_{12} &= \cos\alpha'\sin\alpha - \sin\alpha'\cos\tau\cos\alpha \\ r_{13} &= \sin\alpha'\sin\tau \\ r_{21} &= \sin\alpha'\cos\alpha - \cos\alpha'\cos\tau\sin\alpha \\ r_{22} &= \sin\alpha'\sin\alpha + \cos\alpha'\cos\tau\cos\alpha \\ r_{23} &= -\cos\alpha'\sin\tau \\ r_{31} &= \sin\tau\sin\alpha \\ r_{32} &= \sin\tau\cos\alpha \\ r_{33} &= \cos\tau \end{aligned} \tag{18}$$

3. Estimation of the Fundamental Matrix

3.1. Determination of the Epipolar Lines

In order to derive equations for the epipolar lines, we need to examine the relationship between two orthographic projections of a 3D point under variation of the unknown depth value of this point. This variation of the feature depth corresponds to a back-projection of a 2D feature point into the 3D space and results in a line which naturally includes the original 3D feature. The projection of this line onto the second projection plane is defined as the epipolar line. For orthographic projections, all epipolar lines are parallel. The epipolar geometry for 2D parallel projection stereos has been studied in [2].

Under pure rotations of the feature scene, we find the 2D projection $(x_r', y_r')^T$ of a 3D feature $(x, y, z)^T$ from (7) with Λ being the null vector and R being given by (17) and (18). Varying the feature depth z yields the epipolar line equation for pure rotations:

$$\cos(\alpha')\, x_r' + \sin(\alpha')\, y_r' = \cos(\alpha)\, x + \sin(\alpha)\, y \tag{19}$$

For general scene transformations, the displacement σ from (13) has to be added to account for translations. This results in the equation

$$\cos(\alpha')\,x' + \sin(\alpha')\,y' = \cos(\alpha)\,x + \sin(\alpha)\,y + \sigma \quad (20)$$

for the epipolar lines of a feature (x, y).

3.2. Derivation of the Fundamental Matrix

The fundamental matrix F is defined by the equation

$$\mathbf{p'}^T F\,\mathbf{p} = 0 \quad (21)$$

for any pair of matching points

$$\mathbf{p} = (u, v, w)^T,\ \mathbf{p'} = (u', v', w')^T \quad (22)$$

in two images [4], where the 2D points are denoted in homogenous coordinates as

$$x = \frac{u}{w}, \quad y = \frac{v}{w}, \quad x' = \frac{u'}{w'}, \quad y' = \frac{v'}{w'}. \quad (23)$$

Geometrically, F represents a mapping between a point and its epipolar line. Thus, the fundamental matrix can be regarded as the algebraic representation of the epipolar geometry. In Structure and Motion from stereo views, classical methods such as the 8-point algorithm [4, 8] use the following procedure to determine F: First, feature points are identified in the stereo images. Then, point matches are established conventionally based on proximity and similarity of their intensity neighbourhood. Finally, the unknown matrix F is computed from (21).

Having identified feature points, our approach pursues a different strategy to find F. We first established epipolar geometry constraints based on the proposed integral projection scheme. We then use the resulting epipolar line equation to construct the fundamental matrix. For this, we write F as

$$F = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \quad (24)$$

Using this notation, the epipolar line equation can be directly derived from (21) and (22):

$$(F_{11}\,u + F_{12}\,v + F_{13}\,w)u' + (F_{21}\,u + F_{22}\,v + F_{23}\,w)v' + (F_{31}\,u + F_{32}\,v + F_{33}\,w)w' = 0 \quad (25)$$

Without loss of generality, we let $w = 1$ and $w' = 1$ in which case with (23), (25) becomes

$$(F_{11}\,x + F_{12}\,y + F_{13})x' + (F_{21}\,x + F_{22}\,y + F_{23})y' = -F_{31}\,x - F_{32}\,y - F_{33} \quad (26)$$

A comparison of (26) and (20) yields

$$F_{11}\,x + F_{12}\,y + F_{13} = \cos(\alpha') \quad (27)$$
$$F_{21}\,x + F_{22}\,y + F_{23} = \sin(\alpha') \quad (28)$$
$$F_{31} = -\cos\alpha \quad (29)$$
$$F_{32} = -\sin\alpha \quad (30)$$
$$F_{33} = -\sigma \quad (31)$$

Equations 27 and 28 must be fulfilled for all possible values of x and y which implies

$$F_{11} = F_{12} = F_{21} = F_{22} = 0 \quad (32)$$

Thus, (27) and (28) become

$$F_{13} = \cos(\alpha') \quad (33)$$
$$F_{23} = \sin(\alpha'). \quad (34)$$

It should be noted that there is a remaining a degree of freedom in the construction of the fundamental matrix. That is, since the epipolar line equation in (20) can be multiplied with an arbitrary scalar on both sides, multiplying the elements of F in (29) to (34) with an arbitrary scalar would still yield a valid fundamental matrix.

4. An Algorithm for Estimating the Fundamental Matrix

4.1. Estimation of the parameters

The accurate estimation of $(\alpha, \alpha', \sigma)$ is crucial for the accurate determination of the epipolar geometry and the fundamental matrix. We have designed an algorithm that relies on a maximum likelihood model to robustly extract the matching line angles. Letting two vectors b and c denote the sampled frequency data along the matching lines in the first and second spectrum respectively, the maximum likelihood model leads to the following objective function:

$$d = \frac{\max\left|\Re(\mathcal{F}^{-1}\{b \cdot c^*\})\right|}{|b||c|} \quad (35)$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transformation. An iterative Levenberg-Marquardt search is then used to find the maximum of the resulting objective function.

The estimation of the displacement σ is based on a Lank-Reed-Pollon frequency estimator [6]. We derive a vector r from the complex frequency vectors b and c such that

$$r_k = \frac{b_k \cdot c_k^*}{|b_k||c_k|}. \quad (36)$$

This leads to the following estimate $\hat{\sigma}$ for the displacement

$$\hat{\sigma} = \frac{\arg\left(\sum_k r_{k+1}\,r_k^*\right)}{2\pi\Delta_f}, \quad (37)$$

where Δ_f denotes the frequency resolution along the matching lines. Potential overruns of the 2π range in the phase of the sum in the numerator of (37) can cause ambiguities. However, this can be avoided by choosing a sufficiently small frequency resolution, *i.e.*

$$\Delta_f < \frac{1}{2(|x_{\max}| + |y_{\max}|)} \quad (38)$$

where $x_{\max}$ and $y_{\max}$ are the maximum allowable scene translations in x- and y- direction respectively, parameters which are assumed to be known a priori.

4.2. The overall algorithm

The proposed algorithm involves four basic steps:

1. Given two sets of features that represent the 2D orthographic projections of a set of 3D features from different viewpoints, find starting values for the matching line angles $(\hat{\alpha}_s, \hat{\alpha}'_s)$. These starting values can be found by first extracting the discrete frequency vectors b_k and c_k along the matching lines of the spectra $F_2(\xi, \eta)$ and $F_2(\xi', \eta')$ respectively, evaluating the objective function given in (35) for all vector combinations and choosing the angle pair that corresponds to the vector pair that maximizes the objective function.

2. Using $(\hat{\alpha}_s, \hat{\alpha}'_s)$ as initial values, perform a Levenberg-Marquardt search to iteratively approximate the matching line angle pair $(\alpha_{\max}, \alpha'_{\max})$ that maximizes the objective function in (35). Each iteration yields new estimates $(\hat{\alpha}, \hat{\alpha}')$. Exit when the search algorithm converges to a solution or a maximum number of iterations has been reached.

3. Select the frequency resolution according to the constraint in (38). Then, extract the discrete frequency vectors b_k and c_k along the lines with the angles $(\hat{\alpha}, \hat{\alpha}')$ in the spectra $F_2(\xi, \eta)$ and $F_2(\xi', \eta')$. Compute r_k using (36) and finally the displacement estimate $\hat{\sigma}$ using (37).

4. Using the final estimates $(\hat{\alpha}, \hat{\alpha}', \hat{\sigma})$, either retrieve the epipolar line that corresponds to a feature location (x, y) from (20) or compute the elements of the fundamental matrix F using (29) to (34).

5. Experimental Results

As test data, we generated a random set of N 3D features and projected these features onto a 2D plane using orthographic projection. We synthetically generated a variable percentage of random mismatches during this process.

In our first test, we compared two epipolar lines in the second projection plane that correspond to an arbitrary feature (x, y) in the first projection plane. The first epipolar line was generated with the proposed algorithm. The second epipolar line was generated from a conventional, correspondence-based linear approach [11]. We provided the conventional method with the correspondence information which is hard to obtain in practice [3], therefore putting the conventional method at an advantage. Figure 2 shows the 2D features that result from the projections of thirty 3D features that were randomly generated in a $2000 \times 2000 \times 2000$ pixel sized cube. The 2D features that correspond to the 3D features before and after the scene transformation are depicted by crosses and circles respectively. The solid arrows represent the displacement vectors of the 2D feature correspondences. We used a relatively large scene transformation of $(\phi = 20°, \theta = 45°)$ in azimuth and elevation and translations of $(x_0 = 15,\ y_0 = 10)$. In addition, we incorporated integer rounding of the 2D features to model the effect that, in practice, features are not known with perfect accuracy. We synthetically generated a single mismatch by shifting one of the thirty features in the first 3D feature set into a random position within the 3D feature cube. This step was performed after the 3D scene transformation and before the feature projection. The dashed arrow in Figure 2 shows this mismatch.

The epipolar lines were determined for two arbitrarily selected features that are marked by squares in Figure 2. It can be clearly seen from the solid epipolar lines that results from the proposed integral projection approach are far more precise than the dashed epipolar lines generated with the conventional method. The fact that we provided the conventional method with the correspondence information that was correct apart from minor rounding noise and one single mismatch even further corroborates the advantage of the proposed approach.

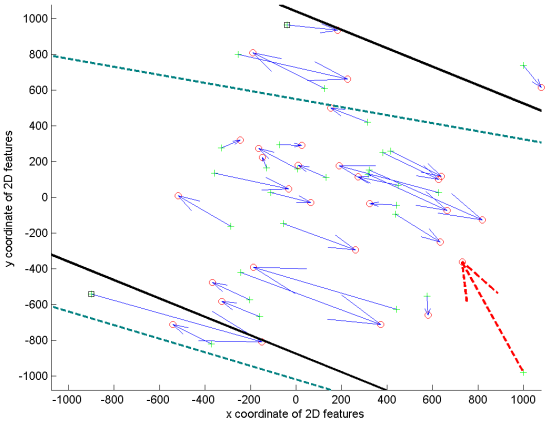
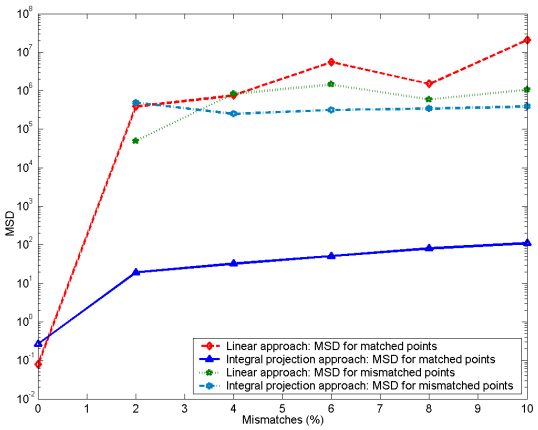


Figure 2. Comparison of epipolar lines

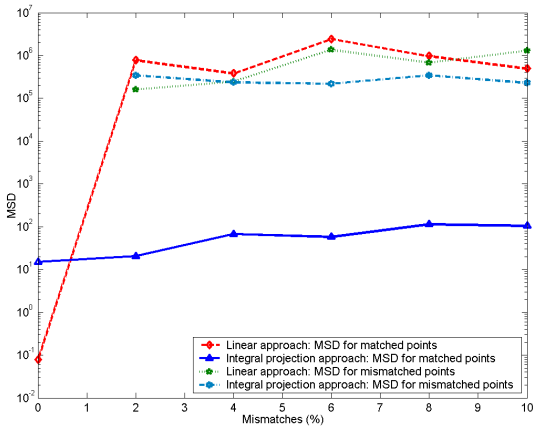
In a second test, we quantitatively examined the performance of our approach with respect to the conventional method under various percentages of mismatches. Specifically, we generated $N = 100$ random features identically to the first test. Both the original large and a small scene transformation $(\phi = 5°, \theta = 2°, x_0 = 7,\ y_0 = 8)$ were applied. In both cases, we synthetically generated mismatches as a percentage of the total number of features $([0, 2, 4, 6, 8, 10]\%)$.

As a quantitative performance measure, two pairs of mean square distances (*MSDs*) were computed. The first *MSD* pair was derived from the distances of those correspondences that were only subject to rounding noise, *i.e.* correctly corresponding feature points (matched points), to the estimated epipolar lines. The computation of the second *MSD* pair was based on the distances of the synthetically generated outliers (mismatched points) to their respective epipolar lines. For each percentage mismatch, we performed ten independent tests of all *MSDs*, each time using a different random set of 3D points. We then averaged these *MSDs* to obtain statistically reliable results. This data is

shown in Figures 3(a) and 3(b) for both the large and the small scene transformations.



(a)



(b)

Figure 3. Mean squared distances for various percentages of mismatches using (a) large transformation parameters ($\phi = 45°, \theta = 20°, x_0 = 15, y_0 = 10$) and (b) small transformation parameters ($\phi = 5°, \theta = 2°, x_0 = 7, y_0 = 8$)

We can make the following observations for both the large and the small scene transformations: If there are no outliers in the correspondence data, the linear method shows slightly smaller matched *MSDs* than the integral projection approach. However, in the presence of mismatches, which is typically the case in practice, the *MSDs* of the matched points are smaller by several orders of magnitude for the integral projection than for the linear method. For the linear approach, the *MSDs* of the mismatched points are smaller than the *MSDs* of the matched points. This highlights the sensitivity of the linear method to outliers and the robustness of the proposed method. The *MSDs* of the mismatched points reach minima at 2% mismatches for the linear approach. This shows the undesirable effect that if only few outliers are present, the linear method estimates the epipolar lines for these mismatched points relatively well. In contrast, the *MSDs* of the mismatches for the integral projection approach are significantly larger than the *MSDs* of the matches. This shows the robustness of the integral projection approach towards outliers in the correspondence data for both small and large scene transformations.

6. Conclusions

In this paper, we have proposed an approach to determine the fundamental matrix from feature points, without any correspondences, that is robust to mismatched points. This can be seen as a major advantage over classical correspondence-based approaches, since establishing correspondences is a problematic task and mismatches have a significant impact on accuracy. Results have been presented to show that the proposed method is robust in the presence of outliers in the feature data. In particular, the *MSDs* of the integral projection approach for matches are smaller by several orders of magnitude than the corresponding *MSDs* of the linear method when outliers are present.

References

[1] R. N. Bracewell. *Two-Dimensional Imaging*. Prentice-Hall, 1995.

[2] J.-X. Chai and H.-Y. Shum. Parallel projections for stereo reconstruction. In *IEEE Conf. Computer Vision Pattern Recognition*, volume 2, pages 493–500, June 2000.

[3] F. Dellaert, S. M. Seitz, C. E. Thorpe, and S. Thrun. Structure from motion without correspondences. *IEEE Computer Vision and Pattern Recognition Proceedings*, 2:557–564, June 2000.

[4] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[5] M. Irani. Multi-frame optical flow estimation using subspace constraints. In *IEEE International Conference on Computer Vision (ICCV)*, pages 626–633, Sept. 1999.

[6] G. W. Lank, I. S. Reed, and G. E. Pollon. A semicoherent detection statistic and doppler estimation statistic. *IEEE Transactions on Aerospace and Electronic Systems*, AES-9(2):151–165, 1973.

[7] S. Lehmann, I. V. L. Clarkson, and P. Kootsookos. An integral projection approach to 3d rigid body transformations. In *Proceedings of the 38th Asilomar Conference on Signals, Systems, and Computers (to appear)*.

[8] Y. Ma, S. Soatto, J. Košecká, and S. S. Sastry. *An Invitation to 3-D Vision*. Springer, 2004.

[9] S. Seitz and C. Dyer. Complete structure from four point correspondences. In *Proc. Fifth Int. Conf. on Computer Vision*.

[10] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization method. *International Journal of Computer Vision*, 9(2):137–154, Nov. 1992.

[11] P. H. S. Torr. A structure and motion toolkit in matlab. Technical Report MS-TR-2002-56, Microsoft Research, June 2002.

[12] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *International Journal of Computer Vision*, 27(2).