International Journal of Computer Vision 65(3), 131–145, 2005

Motion Recovery by Integrating over the Joint Image Manifold

LIRAN GOSHEN AND ILAN SHIMSHONI
Faculty of Industrial Engineering, Technion, Haifa, Israel 32000
lirang@techunix.technion.ac.il
ilans@ie.technion.ac.il

P. ANANDAN
Microsoft Research India, Bangalore, India
anandan@microsoft.com

DANIEL KEREN
Department of Computer Science, Haifa University, Haifa, Israel 31905
dkeren@cs.haifa.ac.il

Received April 12, 2004; Revised September 24, 2004; Accepted September 24, 2004

Abstract. Recovery of epipolar geometry is a fundamental problem in computer vision. The introduction of the "joint image manifold" (JIM) allows to treat the recovery of camera motion and epipolar geometry as the problem of fitting a manifold to the data measured in a stereo pair. The manifold has a singularity and boundary, therefore special care must be taken when fitting it.

Four fitting methods are discussed—direct, algebraic, geometric, and the integrated maximum likelihood (IML) based method. The first three methods are the exact analogues of three common methods for recovering epipolar geometry. The more recently introduced IML method seeks the manifold which has the highest "support," in the sense that the largest measure of its points are close to the data. While computationally more intensive than the other methods, its results are better in some scenarios. Both simulations and experiments suggest that the advantages of IML manifold fitting carry over to the task of recovering epipolar geometry, especially when the extent of the data and/or the motion are small.

Keywords: fundamental matrix estimation, epipolar geometry estimation, motion recovery, integrated maximum likelihood

1. Introduction and Previous Work

Given a stereo pair with point correspondences, one seeks to recover the epipolar geometry, which is dependent on the camera motion and internal calibration. This is a fundamental problem in computer vision, and there exists a huge body of research tackling it; see Hartley and Zisserman (2000) for a thorough treatment.

In the pioneering work Longuet-Higgins (1981), a simple algebraic relation between the corresponding points and the epipolar geometry was derived, which allows to recover the essential matrix given eight matching points in a stereo pair. We refer to this as the *direct solution*. In Weng et al. (1989), it was assumed that more matching pairs are given, and that there are errors in the coordinates. In this scenario, the problem cannot be solved exactly as in Longuet-Higgins

(1981), therefore one seeks an approximate solution by minimizing the sum of squares of the aforementioned algebraic relation. We will henceforth refer to this method by the commonly used name *algebraic method*.

More recent work has roughly followed two other directions:

- The *geometric method*. The idea here is to find a "legal" geometric configuration (i.e., one satisfying the epipolar geometry constraints), such that the sum of squared distances of the matching pairs from it is minimal. This problem is numerically more challenging, but it yields better results (Hartley and Zisserman (2000).
- The *IML (Integrated Maximum Likelihood) method*, sometimes referred to as the *Bayesian approach*. Here, the idea is to recover the epipolar geometry G which maximizes the probability $Pr(G \mid D)$, where D is the measured data (in this case the matching pairs). Some work in this direction is presented in Torr (2002) Kanatani (1993) Kanatani (1994) Torr and Zisserman (1998) Forsyth et al. (1999) Werman and Keren (2001) Keren et al. (2003) Goshen et al. (2003) Ohta (2003) Nestares et al. (2000) Okatani and Deguchi (2002). The paper falls in this category, but is different in the method used to compute the probability (see next section).

1.1. The Joint Image Manifold

The *joint image space* (JIS) Anandan and Avidan (2000) Triggs (1995) is the cartesian product of point pairs in two images. The *joint image manifold* (JIM) for a given epipolar configuration consists of the set of matching pairs which adhere to the epipolar geometry. The notion of the JIM allows to interpret the epipolar geometry problem as the problem of fitting an algebraic manifold. One may work in projective or Euclidean space; we will use the latter, in which the JIM is a three dimensional manifold in $\mathcal{R}_4$ which happens to be an algebraic manifold of degree two Anandan and Avidan (2000).

The key observation in this paper is that since the JIM is an algebraic manifold, the JIM (and epipolar geometry) recovery problem can be represented as a problem of fitting an algebraic manifold, i.e., an implicit polynomial, to the data. While this idea is not new Anandan and Avidan (2000) Triggs (1995), this work suggests to use a fitting method which obtains the IML estimate by integrating out the entire space of nuisance parameters.

1.2. Fitting Algebraic Manifolds

Given a set of points p_i, $1 \leq i \leq n$ in Euclidean space, one may seek a polynomial P such that its zero set (i.e., the points in which it obtains a value of zero), approximates p_i, $1 \leq i \leq n$ (Keren et al. 1994; Taubin et al., 1994). The zero set is commonly called an *algebraic manifold*. Obviously, this is useful when one seeks a polynomial relation which has to be satisfied by some measured data—but this is exactly the situation we face when trying to recover the epipolar geometry. An explanation follows, as well as an interpretation of the four aforementioned methods as fitting techniques.

In Faugeras (1992) Faugeras (1992) Hartley (1992) Hartley et al. (1992), the following equation was derived: $(x_1, y_1, 1)F(x_2, y_2, 1)^T = 0$, where F is the fundamental matrix and $\{(x_1, y_1), (x_2, y_2)\}$ a pair of matching points. This is a linear constraint on F's elements, and if we look at the JIM in $\mathcal{R}^4$ space, which is defined by $(x_1, y_1, 1)F(x_2, y_2, 1)^T = 0$, the problem reduces to fitting such a manifold (defined by F) to the data. How should this be done? Let us proceed to review some methods for recovering epipolar geometry, and compare them to work done in the realm of manifold fitting:

- ***Direct solution.*** If it is assumed that no error is present in the data, it is possible to recover F by directly solving the equations $(x_1, y_1, 1)F(x_2, y_2, 1)^T = 0$. Clearly, if eight pairs are available, there results a system of eight linear equations in eight variables (Longuet-Higgins, 1981). Alas, usually the data is susceptible to measurement errors.
- ***Algebraic method.*** This method minimizes the sum of squares of the residuals (Weng et al., 1989; Bruckstein et al., 2001). It is a common method for treating noisy data and the case in which there are more degrees of freedom in the data than in the model. In the context of algebraic manifold fitting, this is equivalent to finding a polynomial such that the values of the polynomial at the data points should be closest to zero. The method finds the polynomial P which minimizes $\sum_{i=1}^{n} P^2(p_i)$; but this is a notoriously weak method for fitting manifolds (Keren et al., 1994; Taubin et al., 1994; Ahn et al., 2001).

- ***Geometric method.*** This is equivalent to fitting a manifold by minimizing the sum of squared distances of each data point from the point on the manifold which is closest to it. Statistically this method recovers the joint maximum likelihood of both the manifold and the noiseless sources of the measured data points. While computationally more challenging, it yields better results (Taubin et al., 1994; Ahn et al., 2001; Hartley and Zisserman, 2000). This method is also known as the *profile maximum likelihood* method Berger et al. (1999) Severini (2001), but we will stick with the terminology common in the computer vision community. This method requires non-linear optimization with its known problems. To overcome these problems and to correct the statistical bias associated with various approximations, several iterative methods have been proposed (Sampson, 1982; Kanatani, 1996; Leedan and Meer, 2000; Sullivan et al., 1994; Nestares and Fleet, 2003).
- ***Integrated Maximum Likelihood (IML) method*** [Newsam and Redding (1997) Berger et al. (1999) Werman and Keren (2001) Severini (2001)]. The idea is to recover the manifold V, given the data $D = p_1, \ldots, p_n$, by maximizing the probability $Pr(V \mid D)$. Here, p_i will be a point in $\mathcal{R}^4$ obtained by concatenating two matching points in a stereo pair, and V a manifold defining the epipolar geometry, as will be explained shortly. Contrary to the geometric approach, the method suggested here allows each manifold point (and not only the closest one) to be the source of each data point. Bayes' formula states that

$$Pr(V \mid D) = \frac{Pr(D \mid V)Pr(V)}{Pr(D)}.$$

We can ignore $Pr(D)$ since it is fixed. Next we assume a uniform prior on the space of all manifolds V. The justification for using such is prior is twofold: firstly, we do not have any prior knowledge on V (and if such knowledge exists we can incorporate it into our framework). Secondly, if there are many data points, then the effect of such priors is minor. Under these assumptions the V obtained is the maximum likelihood estimate of the data given the manifold. Assuming that the data points are independent $Pr(D \mid V) = \prod_{i=1}^{n} Pr(p_i \mid V)$. But, as opposed to the geometric method—which assumes that $Pr(p_i \mid V)$ is proportional to $\exp(-\frac{d^2(p_i,V)}{2\sigma^2})$, where σ^2 is the noise variance and $d(p_i, V)$ the distance from p_i to V—the IML method seeks an estimate which uses the full probability distribution over V, which under the Gaussian noise model and up to a normalizing factor equals

$$Pr(p_i \mid V) = \int_V \exp\left(-\frac{d^2(p_i, v)}{2\sigma^2}\right)\mu(dv) \quad (1)$$

where the integration is with respect to the usual Lebesgue measure $\mu(dv)$, which assigns identical measures to regions with identical area. For brevity's sake hereafter we will simply write dv. Here we further assume that all points on the manifold have equal probability and therefore their prior probability can be dropped. We make this assumption because a-priori we have no information on the probability that a pair of points in the two images are in correspondence. This choice of prior makes the estimator shift invariant. Even though V is infinite, the integral is finite because the integrand decreases exponentially when moving away from p_i. This is also true for the rest of the integrals appearing in this paper. In Werman and Keren (2001) it was empirically demonstrated that while IML fitting is time consuming, it yields good results especially when

- The manifold is small with respect to the noise.
- The manifold is strongly curved.
- The manifold has a boundary.
- The manifold has a singular point.

In all these cases, $\exp(-\frac{d^2(p_i,V)}{2\sigma^2})$ is a poor approximation to Eq. (1), especially if there's data close to the singularity or the boundary. This possible pitfall was noted in Torr (2002), however there it was assumed that the JIM is "locally linear," and it was proved (as in Werman and Keren (2001)) that in this case $Pr(p_i \mid V) \approx \exp(-\frac{d^2(p_i,V)}{2\sigma^2})$. However, as we shall demonstrate, the JIM is not locally linear, and therefore the IML method is expected to perform better, especially in scenarios in which there is data close to the singularity or the boundary of the JIM.

It should be noted that the geometric method can also be viewed as a maximum likelihood estimate—not of the manifold alone, but of the manifold and the "true" (denoised) sources of the measurement points simultaneously. The IML method used here integrates out the "true" points, yielding the probability of the manifold only.

The paper is laid out as follows: in Section 2 we demonstrate the importance of the IML method for a simple example. In Section 3 we describe the Focus of Expansion (FOE) estimation problem and its solution, and Section 4 deals with the fundamental matrix estimation problem and compares the IML method with the geometric method. Section 5 shortly studies the non-linearity of the JIM. Section 6 presents simulations and experimental results, and Section 7 concludes the paper.

2. IML Estimation—Simple Example

In order to demonstrate the performance of the suggested IML method in the presence of measurement data near a singularity, we study a very simple case — a cone in $\mathcal{R}^2$, which consists of two straight lines: $y = ax, y = -ax$ for some a. This example was chosen since the JIM is also a cone (albeit higher-dimensional—see Section 4).

Denote the lines $y = ax$ and $y = -ax$ by L_1 and L_2 respectively. Then clearly the cone is the union of L_1 and L_2. Its implicit equation is $(y - ax)(y + ax) = 0$. Following the previous study of the three methods, and noting that Werman and Keren (2001)

$$\int_{L_i} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^2 \exp\left(-\frac{d^2(p,l)}{2\sigma^2}\right) dl = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{d^2(p,L_i)}{2\sigma^2}\right) \qquad (2)$$

it follows that the below cost functions have to be optimized for the various methods, in order to recover the optimal slope a when given measured data $p_i = (x_i, y_i)$, $1 \le i \le n$ (assuming that the noise variance satisfies $2\sigma^2 = 1$):

- Algebraic method: minimize

$$\sum_{i=1}^{n} ((y_i - ax_i)(y_i + ax_i))^2$$

- Geometric method: minimize

$$\sum_{i=1}^{n} \min\{d^2(p_i L_1), d^2(p_i, L_2)\}$$

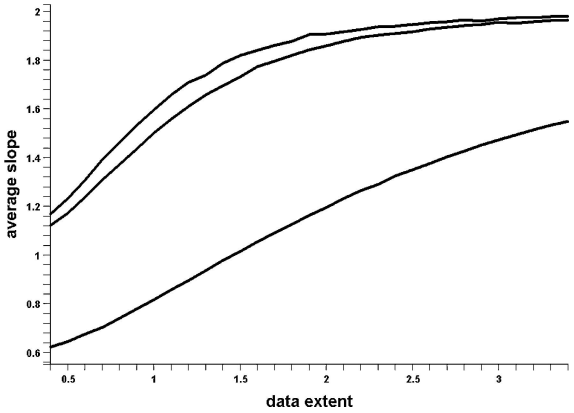


Figure 1. Simulation results of fitting a 1D cone. The horizontal axis represents the data's extent and the vertical axis the average estimate of the slope—upper graph shows the IML results, middle graph the geometric method results, and lower graph the algebraic method results. The correct slope is 2.

- IML method: maximize

$$\sum_{i=1}^{n} \log(\exp(-d^2(p_i, L_1)) + \exp(-d^2(p_i, L_2)))$$

(note that the integration is carried over the entire manifold – i.e., the cone's two branches – hence the exponentials are added).

For points far away from the origin (which is the cone's apex) the geometric method and IML criteria are nearly equivalent, since such points, will be much closer to L_1 than to L_2 or vice-versa (unless the slope is very large); in that case, one of the expressions $\exp(-d^2(p_i, L_1))$, $\exp(-d^2(p_i, L_2))$ is much smaller than the other, hence the cost function will be well represented by $\min\{d^2(p_i, L_1), d^2(p_i, L_2)\}$. However, this is not the case for points near the apex. The strength of the IML method is that it does not force us to decide from what branch of the cone—L_1 or L_2—the point came from; both options are considered. Even in this simple case, the IML method yields better results than the geometric method, and both are far superior to the algebraic method. In Fig. 1 simulation results of fitting a cone to 200 data points generated by adding Gaussian noise of unit variance to a cone of slope 2 are presented. The horizontal axis represents the data's extent (meaning that it ranged uniformly between $-x$ and x), and the vertical axis is the average estimate of the slope—upper graph shows the IML results, middle graph the geometric method results, and lower graph the algebraic method results. When the extent of the data reaches

3.4 (which means that there are more data points away from the apex), both the IML and geometric methods converge to the correct slope, but for smaller extents of the data IML consistently performs better than the geometric method. The algebraic method performs very poorly. In all experiments, the same number of data points were used and the distribution was the same for the $y = 2x$ and $y = -2x$ branches. Note that there is no analog for the manifold's boundary in this simple case.

3. FOE Estimation

Consider next a relatively simple problem—FOE estimation. Given are two images I_1 and I_2 whose centers of projection are O and O'. In this case the camera undergoes pure translation, $\vec{T} = O' - O$, and every pair of corresponding points is collinear with an epipole point, v. So, estimation of the epipolar geometry reduces to estimation of the epipole point. We will now describe the different solutions that can be given to this problem. In the sequel we will represent points in the image using homogeneous coordinates (i.e., the third coordinate is always 1). In this formulation the computation of distance between points as in Eq. (3) is correct as the 1's cancel out.

3.1. The Algebraic Method

The algebraic approach to determine the epipole has a geometric interpretation which is to find the point closest to all lines passing through the pairs of corresponding points. Let l_i be the normalized line through two corresponding points p_i and p'_i such that $|l_i \cdot v|$ is the distance from the line to v. We wish to find the point which satisfies

$$\hat{v} = \arg\min_{v} \sum_{i=1}^{n} (l_i \cdot v)^2,$$

where $\hat{v}$ is an estimate of v·$\hat{v}$ can be easily computed using linear least squares. The problem with this method is that instead of assuming that there are measurement errors in the corresponding points, it is assumed that the estimated epipole is inaccurate and we are trying to minimize this inaccuracy. This problem is rectified by the geometric method.

3.2. The Geometric Method

Given measured corresponding pairs $\{p_i \leftrightarrow p'_i\}$, and assuming that image measurement errors occur in both images, one asks how to "correct" the measurements in each of the two images in order to obtain a perfectly matched set of image points. Formally, we seek an epipole, v, and a set of correspondences $\{\hat{p}_i \leftrightarrow \hat{p}'_i\}$, which satisfy the epipolar geometry constraint and minimize the total error function

$$\hat{v}, \{\hat{p}_i \leftrightarrow \hat{p}'_i\} = \arg\min_{v, \{\hat{p}_i \leftrightarrow \hat{p}'_i\}} \sum_i (\|p_i - \hat{p}_i\|^2 + \|p'_i - \hat{p}'_i\|^2) \quad (3)$$

$$\text{subject to} \quad \hat{p}'_i \cdot (v \times \hat{p}_i) = 0 \quad \forall_i$$

Minimizing this cost function involves determining both v and a set of subsidiary correspondences $\{\hat{p}_i \leftrightarrow \hat{p}'_i\}$. In general, the minimization of this cost function involves non-linear optimization.

3.3. Interpretation of the Geometric Method as Manifold Fitting

The estimation of the epipole point can be thought of as fitting a manifold to points in $\mathcal{R}^4$. Each correspondence of image points $p_i \leftrightarrow p'_i$ defines a single point in the JIS, formed by concatenating the coordinates of p_i and p'_i. For every candidate epipole point v, the image correspondences $p_i \leftrightarrow p'_i$ that satisfy $p_i \cdot (v \times p'_i) = 0$ define a quadratic manifold in $\mathcal{R}^4$, which consists of all the points satisfying

$$p'_{xi} p_{yi} - p'_{yi} p_{xi} - p'_{xi} v_y = p'_{yi} v_x + p_{xi} v_y - p_{yi} v_x = 0.$$

Given measured point matches $\{P_i\} = \{(p_{xi}, p_{yi}, p'_{xi}, p'_{yi})\}$ the task of estimating an epipole point becomes the task of finding a 3D manifold V (defined by v) that approximates the points $\{P_i\}$. In general, of course, it will not be possible to find a manifold which precisely goes through the points; so, the geometric method proceeds as follows: let V be the manifold corresponding to a candidate epipole v, and for each point P_i, let $\hat{P}_i$ be the closest point to P_i lying on the manifold V. One immediately sees that $|P_i - \hat{P}_i\|^2 = \|p_i - \hat{p}_i\|^2 + \|p'_i - \hat{p}'_i\|^2$.

So, the geometric distance in $\mathcal{R}^4$ is equivalent to the geometric distance in the images; and since the IML

method improves over the geometric method for the problem of manifold fitting, it is reasonable to assume that it will also improve over the geometric method for the problem of motion recovery. We now explore this venue.

3.4. The IML Method

The IML method proceeds as follows. Let

$$\hat{v} = \arg\max_{v} f(p_i \leftrightarrow p_i' \mid v)$$

where f is the probability density function (pdf). It remains now to calculate f. First, we take into consideration the "ordering constraint", i.e., either the points in the first image are closer to the epipole than the points of the second image or vice-versa. The direction of the motion T, with a forward or a backward component T_z, determines this ordering. Since we have no prior information about the motion of the camera, we have to compute

$$f(\{p_i \leftrightarrow p_i'\} \mid v) = f(\{p_i \leftrightarrow p_i'\} \mid v, T_z > 0)$$
$$Pr(T_z > 0) + f(\{p_i \leftrightarrow p_i'\} \mid v, T_z < 0) Pr T_z < 0)$$

Because we have no prior information, we assume that $Pr(T_z > 0) = Pr(T_z < 0) = \frac{1}{2}$. Assuming independency over the measurement points we get

$$f(\{p_i \leftrightarrow p_i'\} \mid v, T_z > 0) = \prod_{i=1}^{n} f(p_i \leftrightarrow p_i' \mid v, T_z > 0)$$

and the opposite translation term can be computed similarly.

In order to write down the pdf for a candidate v, integrate out the "nuisance" parameters (Severini, 2001; Berger et al., 1999) to obtain

$$f(p_i \leftrightarrow p_i' \mid v, T_z > 0) = \int \int_v \int f(\{p_i \leftrightarrow p_i'\} \mid \bar{p}_i \leftrightarrow \bar{p}_i', v, T_z > 0) f(\bar{p}_i \leftrightarrow \bar{p}_i' \mid v, T_z > 0) dv$$

where $\{\bar{p}_i \leftrightarrow \bar{p}_i'\}$ are the "real" (i.e., "nuisance") points which have been corrupted by noise, yielding $\{p_i \leftrightarrow p_i'\}$. The integration is over the manifold of "legal" pairs $(\bar{p}_{xi}, \bar{p}_{yi}, \bar{p}_{xi}', \bar{p}_{yi}')$, which satisfy $\bar{p}_i' \cdot (v \times \bar{p}_i) = 0$ and which also correspond to $T_z > 0$. The geometric interpretation of the condition that $T_z > 0$ is that the point $\bar{p}_i$ is constrained to lie on the line segment between v and $\bar{p}_i'$. It will be shown in the sequel that V is a three-dimensional manifold in $\mathcal{R}^4$ with a boundary and a singular point. The integration uses the following parametrization of V

$$\bar{p}_i = \lambda \bar{q}_i' + (1 - \lambda) v, \quad \bar{p}_i' = \bar{q}_i'$$

where $\lambda \in [0, 1]$ parameterizes $\bar{p}_i$ to lie on the segment $[v, \bar{p}_i']$. In this transformation we replace the variables: $\bar{p}_{ix}, \bar{p}_{iy}, \bar{p}_{ix}'$ and $\bar{p}_{iy}'$ by the variables $\lambda, \bar{q}_{ix}'$ and $\bar{q}_{iy}'$. Let $\mathbf{M}$ be the matrix of partial derivatives:

$$\mathbf{M} = \begin{pmatrix} \frac{\partial \bar{p}_{ix}}{\partial \bar{q}_{ix}'} & \frac{\partial \bar{p}_{ix}}{\partial \bar{q}_{iy}'} & \frac{\partial \bar{p}_{ix}}{\partial \lambda} \\ \frac{\partial \bar{p}_{iy}}{\partial \bar{q}_{ix}'} & \frac{\partial \bar{p}_{iy}}{\partial \bar{q}_{iy}'} & \frac{\partial \bar{p}_{iy}}{\partial \lambda} \\ \frac{\partial \bar{p}_{ix}'}{\partial \bar{q}_{ix}'} & \frac{\partial \bar{p}_{ix}'}{\partial \bar{q}_{iy}'} & \frac{\partial \bar{p}_{ix}'}{\partial \lambda} \\ \frac{\partial \bar{p}_{iy}'}{\partial \bar{q}_{ix}'} & \frac{\partial \bar{p}_{iy}'}{\partial \bar{q}_{iy}'} & \frac{\partial \bar{p}_{iy}'}{\partial \lambda} \end{pmatrix} = \begin{pmatrix} \lambda & 0 & \bar{q}_{ix}' - v_x \\ 0 & \lambda & \bar{q}_{iy}' - v_y \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

The Jacobian for a non-square matrix is $J = \sqrt{\det(M^T M)} = \sqrt{1 + \lambda^2} \| v - \bar{q}_i' \|$, and

$$f(p_i \leftrightarrow p_i' \mid v, T_z > 0) = \int_0^1 \left[\iint f(p_i \leftrightarrow p_i' \mid \bar{p}_i = \lambda \bar{q}_i' + (1 - \lambda) v, \bar{p}_i' = \bar{q}_i', v, T_z > 0) J \, d\bar{q}_i' \right] d\lambda \qquad (4)$$

assuming Gaussian noise

$$f(p_i \leftrightarrow p_i' \mid \bar{p}_i \leftrightarrow \bar{p}_i', v, T_z > 0) = \left(\frac{1}{2\pi\sigma^2} \right)^2 \exp\left(- \frac{\| p_i - \bar{p}_i \|^2 + \| p_i' - \bar{p}_i' \|^2}{2\sigma^2} \right).$$

The integral in Eq. (4) was computed numerically. The two infinite integrals over $\bar{q}'$ can be evaluated efficiently by the Gauss-Hermite integration method [Press et al. (1986), Chap. 4.5]. We are then left with an integral over λ whose integrand is effectively non-zero only for a small interval of λ values (an example of such a function is shown in Fig. 2). For functions of this type, some numerical integration procedures do not work well. We therefore find the interval on which the function does not vanish and apply to it the Gauss-Legendre integration method. This interval is located

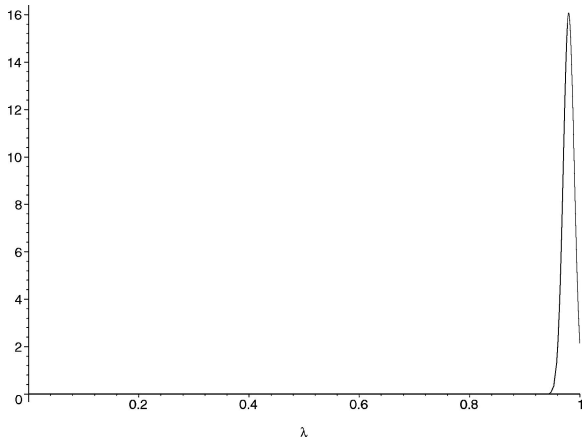


Figure 2. An example of a typical function which is a function of λ for values of $p = (149.5, 150.5)$, $p' = (147.5, 148.5)$, $\sigma = 1$, $v = (50, 50)$.

as follows: first, find an initial value for λ, for which the function does not vanish. This is done by first using the geometric method, which yields an estimate of $\hat{p}_i$ and $\hat{p}'_i$. The corresponding value of λ, i.e., $\lambda = \frac{\|\hat{p}_i - v\|}{\|\hat{p}'_i - v\|}$, is then taken to be the center of the λ interval over which integration is carried out; the interval's limits are estimated by two binary searches in the segments $[0, \lambda]$ and $[\lambda, 1]$.

The integral measures the pdf of a particular v. The optimal v was recovered by applying the Nelder-Mead optimization method (Press et al., 1986).

3.5. *Geometric Interpretation of the IML Method*

The IML method proposed here seeks to find a manifold which has the largest "support", in a sense that there is a large measure of corresponding points on the manifold – i.e., $\bar{p}_i \leftrightarrow \bar{p}'_i$ which satisfy $\bar{p}_i \cdot (v \times \bar{p}'_i)$ – that are close to the measured corresponding pairs $p_i \leftrightarrow p'_i$. The IML method takes into consideration a large volume of information; it considers the entire manifold and its exact structure instead of only the point on the manifold closest to the measured point pair. Therefore it is expected to provide higher accuracy, and it does so, as will be demonstrated in Section 6.

Figure 3 demonstrates the difference between the geometric and IML methods, in the simplest scenario possible—two matching point pairs, $p_1 \leftrightarrow p'_1$ and $p_2 \leftrightarrow p'_2$. The geometric method will prefer v_1 over v_2 as the FOE, since the distance of the manifold corresponding to v_1 from the measured data is zero. However, if the measurement noise is not too small

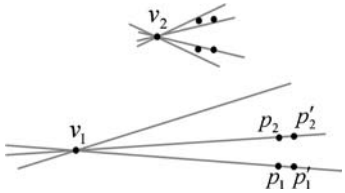


Figure 3. A simple scenario demonstrating the difference between the geometric and IML methods. The geometric method prefers v_1, the IML prefers v_2 which has a larger support (see text preceding figure).

with respect to the size of the point configuration, the IML method will prefer v_2, since the manifold defined by it carries a larger *support* for the data. This may appear odd – after all, v_1 seems like a perfect epipole. However, the v_1 manifold carries very little support for the data, as only a very narrow range of line angles through v_1 are close to the data, as opposed to v_2. v_1 is therefore an unstable choice and it overfits the measurement noise.

4. The Fundamental Matrix Case

In this section the general case—the fundamental matrix estimation problem—is addressed. The fundamental matrix F depends on the internal calibration matrix K, the rotation matrix between two views R, and the translation vector $\mathbf{e}$ (the epipole). When optimizing over all possible fundamental matrices, we have to optimize over these components. As before, assume that a set of correspondences $\{p_i \leftrightarrow p'_i\}$ is given.

Next, we integrate over the nuisance parameters and then maximize over the manifold's parameters. The manifold $\hat{F}$ is sought (where f is the probability density function):

$$\hat{F} = \arg\max_F f(\{p_i \leftrightarrow p'_i\} \mid F).$$

In order to calculate f, it is necessary to integrate over the manifold that corresponds to a candidate F. We start with some simple observations about the manifold's shape. As noted in Anandan and Avidan (2000), this manifold is a *cone*.

4.1. *The Cone*

We now take a closer look at the cone which constitutes the JIM. Using the well-known notion of the fundamental matrix F, the epipolar constraint can be

written as $(x_1, y_1, 1)F(x_2, y_2, 1)^T = 0$ for matching points $(x_1, y_1), (x_2, y_2)$. It is also well-known that F is of rank 2 (see Hartley and Zisserman (2000) for discussion and references). Now follow a few lemmas:

Lemma 1. *Under the transformation*

$$(x_1, y_1, x_2, y_2) \to (x_1 - a_1, y_1 - b_1, x_2 - a_2, y_2 - b_2)$$

where $(a_1, b_1, 1)F = 0, F(a_2, b_2, 1)^T = 0^T$, *the fundamental matrix assumes the form*

$$F = \begin{pmatrix} F_{11} & F_{12} & 0 \\ F_{21} & F_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Note that this transformation is achieved simply by moving the origin of the left and right images to the epipole points. The proof is immediate.

Lemma 2. *In the notation of Lemma* 1, *the constraint* $(x_1, y_1, 1)F(x_2, y_2, 1)^T$ *can be expressed as*

$$(x_1, y_1, x_2, y_2)F_4(x_1, y_1, x_2, y_2)^T = 0$$

where (Anandan and Avidan, 2000)

$$F_4 = \begin{pmatrix} 0 & 0 & F_{11} & F_{12} \\ 0 & 0 & F_{21} & F_{22} \\ F_{11} & F_{21} & 0 & 0 \\ F_{12} & F_{22} & 0 & 0 \end{pmatrix}$$

The proof is immediate.

Lemma 3. *There is a rotation of coordinates such that if* F_4 *is in the form of Lemma* 2, $(x_1, y_1, x_2, y_2)F_4(x_1, y_1, x_2, y_2)^T = 0$ *is equal to up to a scale factor*

$$\left(\mathbf{x}_1^2 - \mathbf{y}_1^2\right) + \gamma^2\left(\mathbf{x}_2^2 - \mathbf{y}_2^2\right) = 0$$

$\mathbf{x}_i, \mathbf{y}_i, i = 1 \ldots 2$ *are the new coordinates.*

The proof follows simply by diagonalizing the 4×4 matrix of Lemma 2. It turns out that it has two pairs of eigenvalues with opposite signs, $\pm\gamma_1, \pm\gamma_2$ given by the following expressions:

$$e_1 = 2F_{11}^2 + 2F_{12}^2 + 2F_{21}^2 + 2F_{22}^2$$

$$\begin{aligned} e_2 = {} & F_{11}^4 + F_{12}^4 + F_{21}^4 + F_{22}^4 + 2F_{22}^2F_{21}^2 + 2F_{22}^2F_{12}^2 \\ & -2F_{22}^2F_{11}^2 - 2F_{12}^2F_{21}^2 + 2F_{11}^2F_{21}^2 \\ & +2F_{11}^2F_{12}^2 + 8F_{11}F_{12}F_{21}F_{22} \end{aligned}$$

$$\gamma_1 = \frac{1}{2}\sqrt{e_1 + 2\sqrt{e_2}} \quad \gamma_2 = \frac{1}{2}\sqrt{e_1 - 2\sqrt{e_2}} \quad \gamma = \frac{\gamma_1}{\gamma_2}$$

We note, however, that the rotation required for diagonalizing F is not separable in the images—i.e., it cannot be represented as a combination of separate rotations in (x_1, y_1) and (x_2, y_2), but it "mixes" all the four coordinates $\{x_1, y_1, x_2, y_2\}$. However, as far as the fitting is concerned, this makes no difference, as long as we apply the same transformations to the cone and to the data points. After the transformations of Lemmas 1 and 3, Eq. (5), evidently describes a cone in $\mathcal{R}^4$ whose apex is at the origin.

4.2. The Cone Boundaries and Singularity

As will be described shortly, the IML method is more computationally expensive than other methods, due to the numerical integration. When is it important to suffer this overhead? As noted in Werman and Keren (2001) Torr (2002), if the data is near a locally linear region of the cone, not much is gained byintegrating. However, in the following two cases, the local linearity assumption is strongly violated:

- The (transformed) data points are close to the cone's apex (0, 0, 0, 0). Clearly, local linearity is violated. This can happen, for example, when the object is small, and the camera is moving towards it (as in tracking).
- For the sake of simplicity, assume for now that there is only camera translation present, and that it is forward or backward relative to the center of the scene, which we assume to be at the origin of the coordinate system. It is clear that if the matching point pairs are denoted $(p_1^{(i)}, p_2^{(i)})$, then either all the "true" $p_1^{(i)}$'s are closer to the origin than all the corresponding $p_2^{(i)}$'s, or vice-versa—the "order constraint."

What does this mean, in terms of the manifold? If we disregard the order constraint, then the only restriction on the matching pairs (in the very simple scenario described above), is that each $p_2^{(i)}$ is the product of $p_1^{(i)}$ by a certain scalar. So, the corresponding cone is equal to

$$C = \{(x_1, y_1, \delta x_1, \delta y_1) \mid x_1, y_1 \in \mathcal{R}, \delta \in \mathcal{R}^+\}$$

However, the order constraint implies that the legal configurations of the "true" points (that is, the denoised measurement points) are in the union of the "half-cones" C_1, C_2, where

$$C_1 = \{(x_1, y_1, \delta x_1, \delta y_1) \mid x_1, y_1 \in \mathcal{R}, 0 \leq \delta \leq 1\}$$

and

$$C_2 = \{(x_1, y_1, \delta x_1, \delta y_1) \mid x_1, y_1 \in \mathcal{R}, 1 \leq \delta \leq \infty\}$$

note that C_1, C_2 are *manifolds with boundary*; the boundary of both is $\{(x_1, y_1, x_1, y_1) \mid x_1, y_1 \in \mathcal{R}\}$. When can there be data points close to the boundary? If the disparity between the matching points is large relative to the noise, then the noised "true points" will be close to each other (and hence to the boundary) only with low probability. However, if the motion is small (as can be the case in a video sequence), data will lie by the boundary.

What does this mean, intuitively? Suppose that the camera motion is forward, hence $p_2^{(i)}$ is farther from the origin than $p_1^{(i)}$. If we simply integrate over the entire cone, we are allowing *illegal* configurations in which $p_1^{(i)}$ is farther from the origin than $p_2^{(i)}$. If the disparity is small, these illegal configurations are assigned relatively high probabilities, as even a small noise can switch the order of the corresponding points.

In light of this, we have to integrate over C_1 and C_2, and multiply the resulting probabilities. It should be clear that the problem of violating the order constraint for small motions occurs in all scenarios, not only the simple one discussed here.

We next address the problem of fitting such a cone in the IML approach based on integrating over it.

4.3. Integration Over the Cone

In order to deal with the boundary we used the following parametrization of the fundamental matrix: $F = [e]_\times H_\infty$, where $H_\infty = KRK^{-1}$ is the infinite homography (Hartley and Zisserman, 2000), the epipole is $\mathbf{e} = (e_1, e_2, e_3)^T$ and

$$[\mathbf{e}]_\times = \begin{pmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{pmatrix}.$$

When we use this parametrization we can apply the ordering constraint. The correspondences satisfy the relation $p_i'^T F p_i = 0$. Note that the transformed points $\bar{q}_i = H_\infty^{-1} p_i$ satisfy $p_i'^T [e]_\times \bar{q}_i = 0$ which is similar to the FOE case and we can enforce the constraint in a similar manner.

Let

$$\hat{F} = \arg\max_{F=[e]_\times H_\infty} f(\{p_i \leftrightarrow p_i'\} \mid F = [e]_\times H_\infty)$$

If there is no prior information about the motion of the camera, the following should be computed:

$$f(\{p_i \leftrightarrow p_i'\} \mid F) = f(\{p_i \leftrightarrow p_i'\} \mid F, T_z > 0) Pr(T_z > 0) + f(\{p_i \leftrightarrow p_i'\} \mid F, T_z < 0) Pr(T_z < 0)$$

Assuming $Pr(T_z > 0) = Pr(T_z < 0) = \frac{1}{2}$ and independency over the measurement points

$$f(\{p_i \leftrightarrow p_i'\} \mid F, T_z > 0) = \prod_{i=1}^{n} f(p_i \leftrightarrow p_i' \mid F, T_z > 0)$$

The opposite translation term can be computed in a similar manner. We again integrate over the "nuisance" parameters, yielding

$$f(p_i \leftrightarrow p_i' \mid v, T_z > 0) = \int\int_V \int f(\{p_i \leftrightarrow p_i'\} \mid \bar{p}_i \leftrightarrow \bar{p}_i', F, T_z > 0) f(\bar{p}_i \leftrightarrow \bar{p}_i' \mid F, T_z > 0) dv$$

where $\bar{p}_i$ and $\bar{p}_i'$ are the 'real' points. We use the following change of variables:

$$\frac{H_\infty \bar{p}_i}{H_{\infty 3} \bar{p}_i} = \lambda \bar{q}_i' + (1 - \lambda) e \quad \bar{p}_i' = \bar{q}_i'$$

where $H_{\infty 3}$ is the last row of H_∞ and $\lambda \in [0, 1]$ parameterizes $\frac{H_\infty \bar{p}_i}{H_{\infty 3} \bar{p}_i}$ to lie on the segment $[e, \bar{p}_i']$ which constrains the proper order of the correspondence. The rest of the derivation is very similar to the FOE case—insert the proper Jacobian and then deal with the resulting triple integral.

Note that the above derivation is general. In the case of FOE estimation or calibrated translation and rotation estimation, all we need is to change the H_∞ transformation to the identity transformation or the rotation matrix R respectively.

5. Non-Linearity of the Manifold

In Kanatani (1994) Kanatani (1996), the problems which arise when fitting a curved manifold are

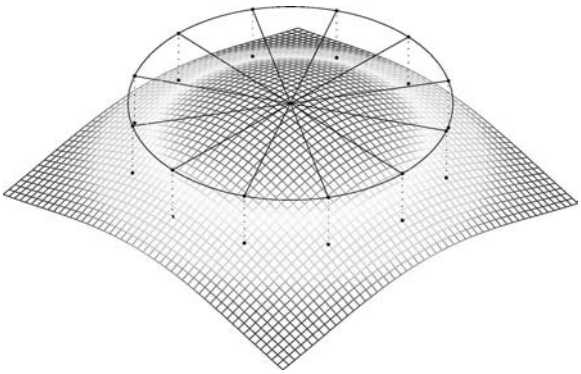


Figure 4. An illustration of the non-linearity measure for a surface in 3D. The average distance from points on a circle on the tangent plane to the surface is computed. In 4D the circle is replaced by a sphere.

addressed. In Torr (2002) it is demonstrated that the geometric method can be viewed as an approximation to the IML method when the manifold portions close to the data are linear. The IML method is not restricted to locally linear manifolds. We now take a closer look at the relation between the non-linearity of the manifold and the computer vision problem at hand.

Due to the fact that there is no single measure for curvature of surfaces of dimension more than one we use the following method to estimate the non-linearity of the surface at a point. Let V be a three dimensional manifold in $\mathcal{R}^4$ and let P be a point on V. We estimate the non-linearity of the manifold using the following method illustrated in Fig. 4 for a surface in 3D. Compute the normal $\vec{N}$ at P. $\vec{N}$ is perpendicular to the tangent hyper-plane of V at P, which we denote $\Pi(V, P)$. Then, construct an orthogonal basis of the $4D$ space in which one of the vectors is $\vec{N}$, and the other three span $\Pi(V, P)$. In $\Pi(V, P)$ construct a unit dodecahedron over the three basis vectors. From each point of the dodecahedron project a line l to V, where l is perpendicular to $\Pi(V, P)$. The non-linearity of V at P is defined as the mean length of these l's. In the pure translation case, the parameter which carries the largest influence on the non-linearity is the distance of the correspondences from the epipole. In Fig. 5 the non-linearity is depicted as a function of that distance. The non-linearity is much larger when the correspondences are close to the epipole. This indicates that as suggested before, the strength of the IML method is more noticeable in this case.

In the general case we could not find any simple relation between the parameters of the fundamental matrix and the non-linearity of the manifold; the relation seems to be quite complex.

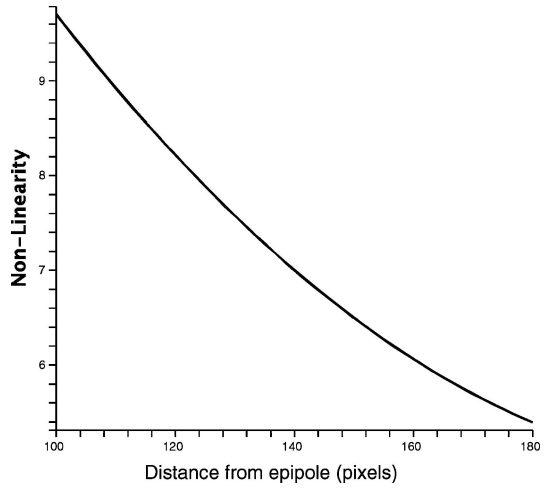


Figure 5. Manifold's non-linearity in the pure translation case. The epipole is at the origin, p varies, and p' is in a distance of 10 pixels from p'. The point P at which the non-linearity was calculated is the concatenation of p, p'.

6. Experimental Results

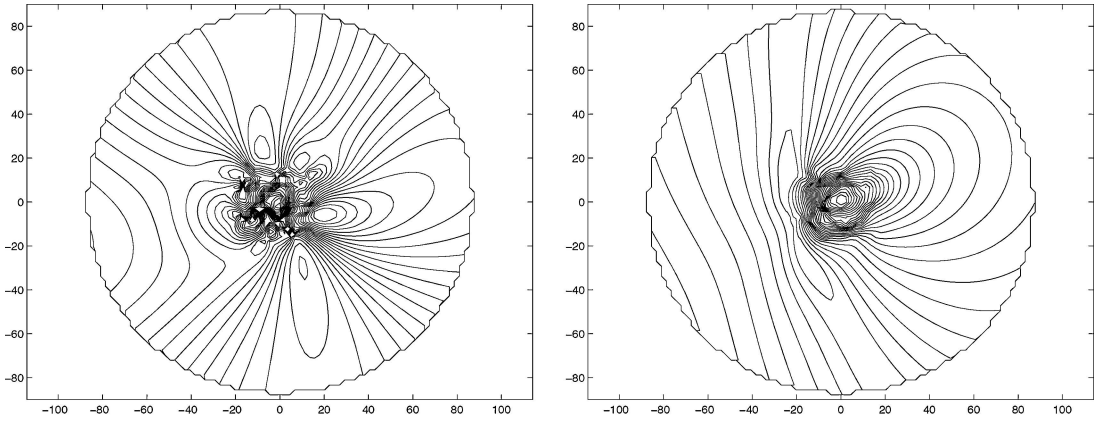
In this section we present results obtained on simulated and real images.

6.1. Simulations

Several simulations were performed to compare the IML method to the geometric method. Pure translation as well as translation and rotation (with known camera calibration) were studied.

In each simulation, $3D$ points were chosen at random in the common field of view. We have added Gaussian noise with zero mean and variance σ^2 to every image coordinate. In each set of experiments, the epipolar geometry (FOE or essential matrix) was estimated 100 times with different instances of points and noise.

For each simulation the estimation errors of the epipole and the rotation angle were calculated. The estimation error of the epipole was calculated as follows: Let v be the "real" epipole point in the simulation and let $\hat{v}$ be the estimated epipole. Let $\vec{v} = (v_x, v_y, f_o)$ and $\vec{\hat{v}} = (\hat{v}_x, \hat{v}_y, f_o)$, where f_o is the focal length. The estimation error is the angle v_{err} between these two vectors. i.e., $v_{\text{err}} = \arccos\left(\frac{\vec{v}\cdot\vec{\hat{v}}}{|\vec{v}||\vec{\hat{v}}|}\right)$. The estimation error of the rotation matrix is the rotation angle about



(a) (b)

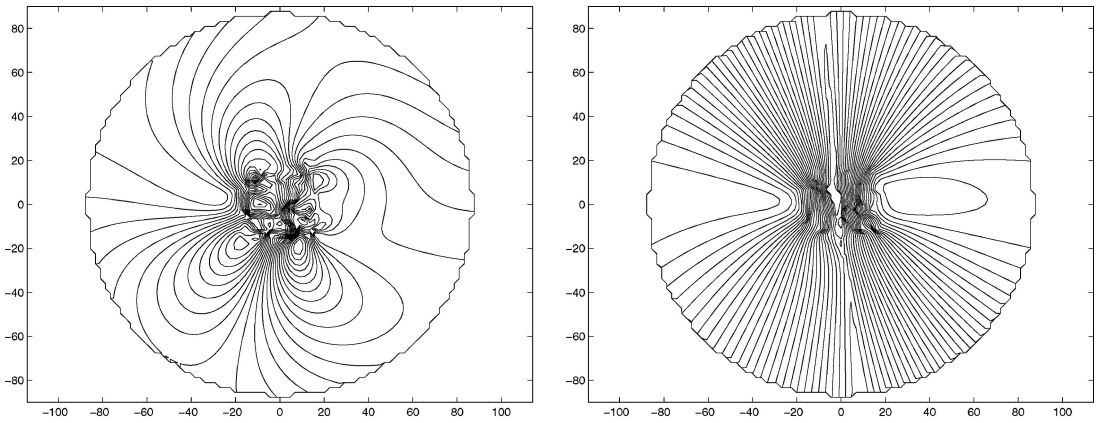
Figure 6. (a) Likelihood function for the geometric method, 100 point correspondence. The correct FOE is at (5, 5). (b) IML likelihood function for the same scenario as (a). The function is much more stable and obtains its global minimum close to the correct location.

the axis of rotation corresponding to the rotation matrix $R_{\text{err}} = R^{-1}\hat{R}$, where R is the ground truth rotation matrix in the simulation and $\hat{\text{R}}$ the estimated rotation matrix.

The numerical parameters used in the simulations were: image size 600 × 800 pixels and the internal calibration matrix

$$K = \begin{pmatrix} 1000 & 0 & 0 \\ 0 & 1000 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

6.1.1. Pure Translation. In the following experiment we compared the likelihood functions of the IML method Eq. (4) to the geometric method Eq. (3). In this simulation there were 100 point correspondences and a very small translation, resulting in a disparity of 1.8 pixels on the average. The simulated motion was nearly forward, with the (normalized) motion vector (0.086, 0.086, 0.992). The noise was Gaussian with standard deviation 1. Typical examples of the values of the likelihood functions are shown in Fig. 6 as contour maps. These examples demonstrate that the IML likelihood function has less local minima than the geometric



(a) (b)

Figure 7. (a) Likelihood function for the geometric method, with parameters similar to those in Fig. 6, except that the motion has a stronger sideways component and the correct FOE is at (60, 5). (b) IML likelihood function for the same scenario as (a). The function is much more stable and obtains its global minimum close to the correct location.

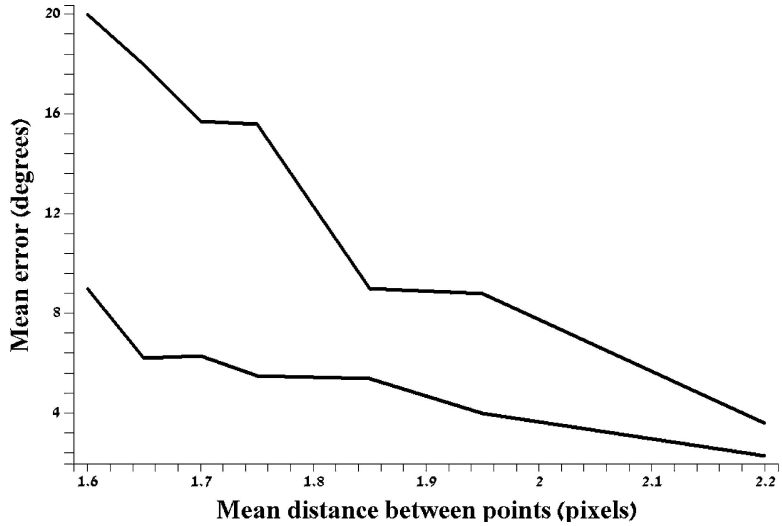
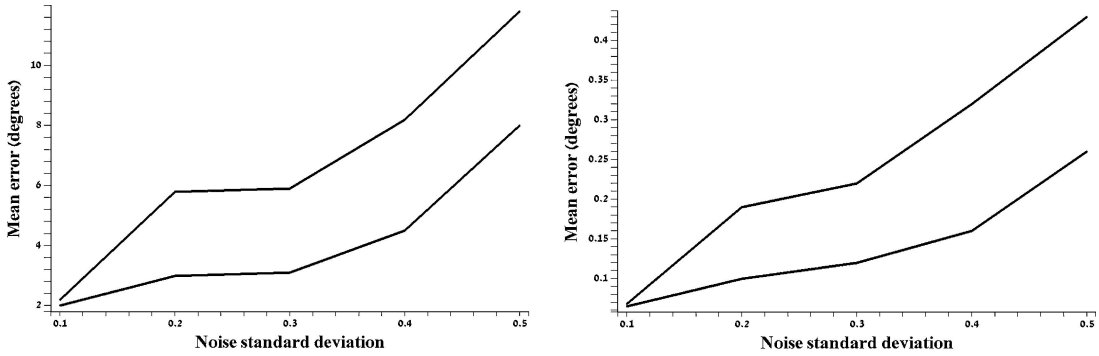


Figure 8. Performance of the IML (Lower graph) and geomatric (upper graph) methods for pure translation.



(a) (b)

Figure 9. (a) Results for FOE in the calibrated rotation and translation case. (b) Results for the rotation angle in the calibrated rotation and translation case. In both figures, IML error is lower graph, geometric method error upper graph.

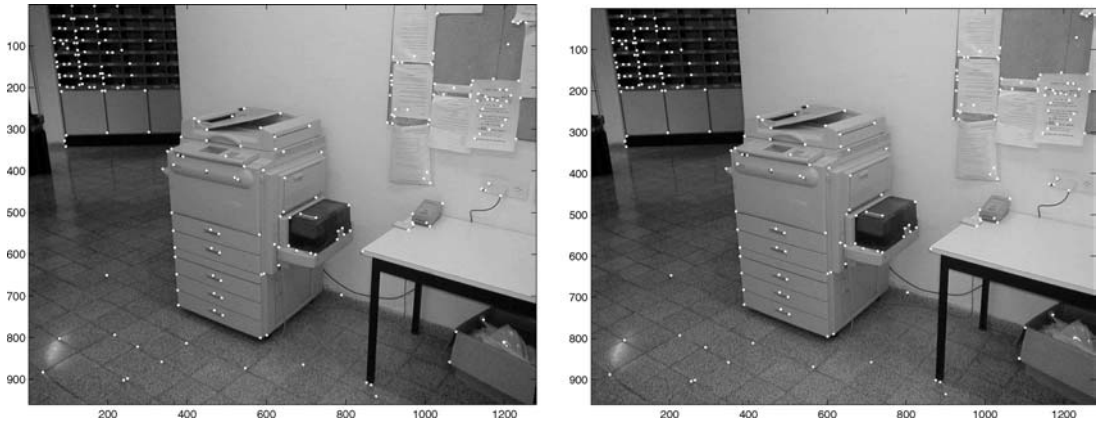
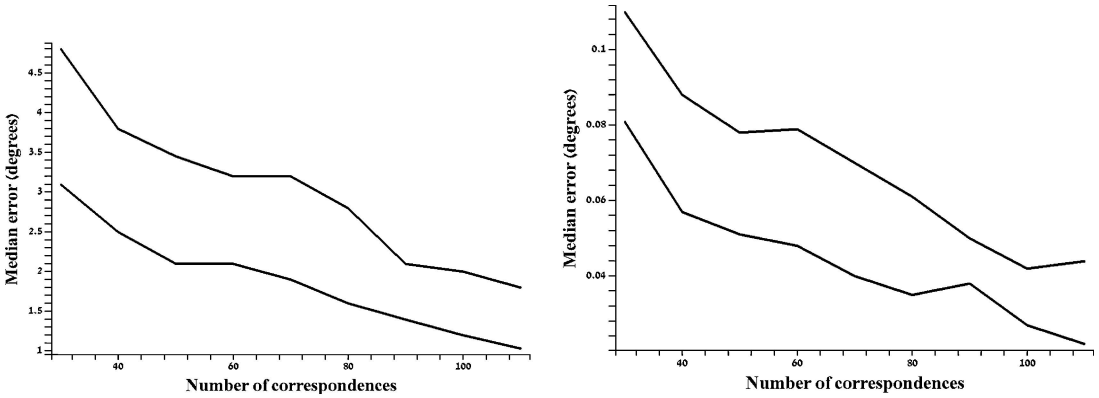


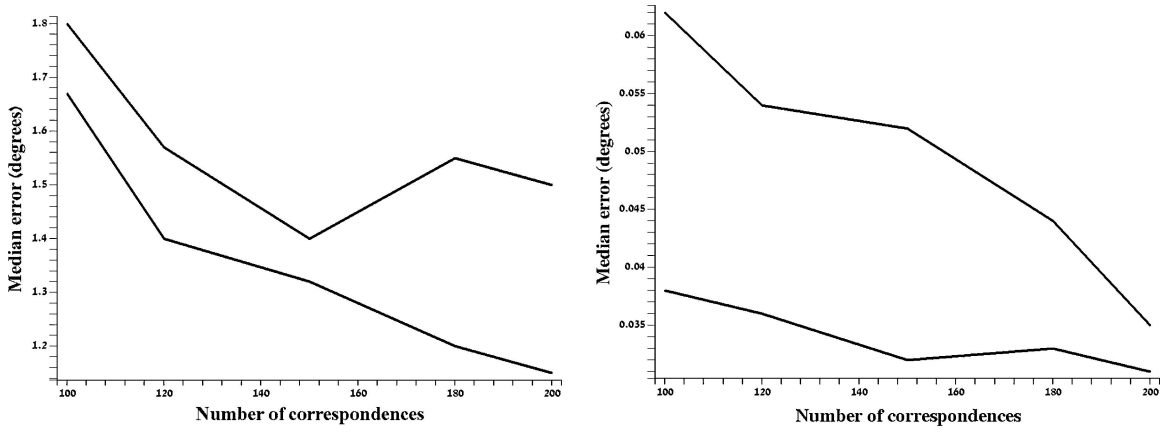
Figure 10. First Image pair, with matching points marked.



(a) (b)

Figure 11. (a) Error in FOE for IML (lower graph) and geometric (upper graph) methods, vs. the number of point correspondences. (b) Error in rotation for IML (lower graph) and geometric (upper graph) methods, vs. the number of point correspondences.

Figure 12. Second image pair, with matching points marked.



(a) (b)

Figure 13. (a) Error in FOE for IML (lower graph) and geometric (upper graph) methods, vs. the number of point correspondences. (b) Error in rotation for IML (lower graph) and geometric (upper graph) methods, vs. the number of point correspondences.

likelihood function. Thus the optimization procedure will usually not get trapped in local minima.

In Fig. 7 the parameters are similar to those in Fig. 6, but the motion has a stronger sideways component, and the correct FOE is at (60°, 5°). Again, the IML likelihood function is more stable, less local minima exist and the location of its global minimum is closer to the correct location. Next, the estimates of the geometric and IML methods for the parameters corresponding to Fig. 6 are compared; the results are shown in Fig. 8, with the average error plotted as a function of the disparity between corresponding points. The IML method performs better, especially for small translations.

6.1.2. Calibrated Translation and Rotation. In this simulation the IML and geometric method were compared for translation and rotation (essential matrix recovery). There were 10 point correspondences. The translation was such that the mean disparity between the corresponding points, due to the translation alone, was 8 pixels. The rotation angle was 3°. The results are presented in Fig. 9. As for the translation only case, the IML method outperforms the geometric method, especially when the noise increases.

The IML method is more significant for the problem of essential matrix recovery. It gives superior results also in configurations in which the noise is relatively small and translation relatively large. In such configurations for the pure translation case, the geometric and IML methods yield very similar results.

6.2. Real Images

Next we show results obtained for two real image pairs. The image pairs consist of two images of office scenes shown in Figs. 10 and 12. The camera motion was very small (a few centimeters). The corners were recovered using the Harris corner detector (Harris, and Stephens, 1988).

In this case the internal calibration matrix K was known, and the goal was to recover the essential matrix E by optimizing over all values of the rotation matrix R and the epipole e.

When taking the entire field of view, and using all the point correspondences, an accurate estimate of the essential matrix was found using the geometric method, with the Nelder-Mead optimization method (Press et al., 1986); this was regarded as the ground truth. Then, the performance of the IML and geometric methods were tested on small random subsets of the matched pairs.

In the first scene (Fig. 10) the total number of correspondences was 240 and the rotation angle was 1.1°. Figure 11 compares the results of the geometric and IML method for 30 to 110 corresponding pairs. In the second scene (Fig. 12) the total number of correspondences was 330 and the rotation angle was 2.6°. Figure 13 compares the results of the geometric and IML method for 100 to 250 corresponding pairs. In both cases the IML method outperformed the geometric method. The running times of the algorithm for the two examples is 26–55 min and 96–120 min for 30 and 100 pairs respectively. The algorithm's running time is linear in the number of point pairs given.

7. Summary and Conclusions

We have described an IML estimation method for the recovery of epipolar geometry. The introduction of the joint image space manifold allows to treat the problem of recovering camera motion and epipolar geometry as the problem of fitting a manifold to the data measured in a stereo pair.

If the camera motion is small, and/or the objects are small relative to their distance from the camera, the IML method has the potential to significantly improve on the geometric method. This is because the manifold which represents the epipolar geometry has a singularity and boundary; hence the local linearity assumption, under which the geometric method is a reasonable approximation, may well be violated—since the points may in these cases be close to the singularity and to the manifold's boundary. The IML method can handle these situations better than the geometric method.

Planned future work includes further developing the numerical integration and optimization techniques, as well as extending the ideas presented here to more than two images.

Acknowledgments

Part of this research took place while D. Keren was a visitor at the Vision Technology group at Microsoft Research, Redmond. The generous support of the Israel Science Foundation grant no. 591-00/10.5 is gratefully acknowledged. We are very grateful to the reviewers for their many helpful comments and corrections.

References

Ahn, S.J., Rauh, W., and Warnecke, H.J. 2001. Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola, and parabola. *Pattern Recognition*, 34(12):2283–2303.

Anandan, P. and Avidan, S. 2000. Integrating local affine into global projective images in the joint image space. In *European Conference on Computer Vision*, pp. 907–921.

Berger, J., Liseo, B., and Wolpert, R. 1999. Integrated likelihood methods for eliminating nuisance parameters. *Statistical Science*, 14(1):1–28.

Bruckstein, A.M., Holt, R.J., Jean, Y.D., and Netravali, A.N. 2001. On the use of shadows in pose recovery. *International Journal of Imaging Systems and Technology*, 11:315–330.

Faugeras, O.D. 1992. What can be seen in three dimensions with an uncalibrated stereo rig? In *European Conference on Computer Vision*, pp. 563–578.

Faugeras, O.D., Luong, Q.T., and Maybank, S.J. 1992. Camera self-calibration: Theory and experiments. In *European Conference on Computer Vision*, pp. 321–334.

Forsyth, D.A., loffe, S., and Haddon, J. 1999. Bayesian structure from motion. In *International Conference on Computer Vision*, pp. 660–665.

Goshen, L., Shimshoni, I., Anandan, P., and Keren, D. 2003. Recovery of epipolar geometry as a manifold fitting problem. In *International Conference on Computer Vision*, pp. 1321–1328.

Harris, C. and Stephens, M.J. 1988. A combined corner and edge detector. In *Proc. of Alvey Vis. Conf.*, pp. 147–152.

Hartley, R.I. 1992. Estimation of relative camera positions for uncalibrated cameras. In *European Conference on Computer Vision*, pp. 579–587.

Hartley, R.I., Gupta, R., and Chang, T. 1992. Stereo from uncalibrated cameras. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, pp. 761–764.

Hartley, R.I. and Zisserman, A. 2000. *Multiple View Geometry in Computer Vision*, Cambridge.

Kanatani, K. 1993. Geometric computation for machine vision. In Oxford University Press.

Kanatani, K. 1994. Statistical-analysis of geometric computation. *Comp. Vis. Graph. Im. Proc.*, 59(3):286–306.

Kanatani, K. 1994. Statistical bias of conic fitting and renormalization. *IEEE Trans. Patt. Anal. Mach. Intell.* 16(3):320–326.

Kanatani, K. 1996. *Statistical Optimization for Geometric Computation: Theory and Practice*. North-Holland.

Keren, D., Cooper, D., and Subrahmonia, J. 1994. Describing complicated objects by implicit polynomials. *IEEE Trans. Patt. Anal. Mach. Intell.* 16:38–53.

Keren, D., Shimshoni, I., Goshen, L., and Werman, M. 2003. All points considered: A maximum likelihood method for motionrecovery. In *Theoretical Foundations of Computer Vision*, pp. 72–85.

Leedan, Y. and Meer, P. 2000. Heteroscedastic regression in computer vision: Problems with bilinear constraint. *International Journal of Computer Vision*, 37(2):127–150.

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

Nestares, O., Fleet, D.J., and Heeger, D.J. 2000. Likelihood functions and confidence bounds for total-least-squares problems. In *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*, pp. I:523–530.

Nestares, O. and Fleet, D.J. 2003. Error-in-variables likelihood functions for motion estimation. In *International Conference on Image Processing*, vol. 2, pp. III-77–80.

Newsam, G. and Redding, N. 1997. Fitting the most probable curve to noisy observations. In *International Conference on Image Processing*, pp. II:752–755.

Ohta, N. 2003. Motion parameter estimation from optical flow without nuisance parameters. In *Third IEEE Workshop on Statistical and Computational Theories of Vision*.

Okatani, T. and Deguchi, K. 2002. Is there room for improving estimation accuracy of the structure and motion problem? In *Proc. Statistical Methods in Video Processing Workshop*, pp. 25–30.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. 1986. *Numerical Recipes*. Cambridge University Press.

Sampson, P.D. 1982. Fitting conic sections to 'very scattered' data: An iterative refinement of the bookstein algorithm. *Computer Vision, Graphics, and Image Processing*, 18:97–108.

Severini, T.A. 2001. *Likelihood Methods in Statistics*. Oxford University Press.

Sullivan, S., Sandford, L., and Ponce, J. 1994. Using geometric distance fits for 3D object modeling and recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(12):1183–1196.

Taubin, G., Cukierman, F., Sullivan, S., Ponce, J., and Kriegman, D.J. 1994. Parameterized families of polynomials for bounded algebraic curve and surface fitting. *IEEE Trans. Patt. Anal. Mach. Intell.* 16(3):287–303.

Torr, P.H.S. 2002. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. *International Journal of Computer Vision*, 50(1):35–61.

Torr, P.H.S. and Zisserman, A. 1998. Concerning bayesian motion segmentation, model averaging, matching and the trifocal tensor. In *European Conference on Computer Vision*, pp. 511–527.

Triggs, B. 1995. Matching constraints and the joint image. In *International Conference on Computer Vision*, pp. 338–343.

Weng, J., Huang, T.S., and Ahuja, N. 1989. Motion and structure from two perspective views: Algorithms, error analysis, and error estimation. *IEEE Trans. Patt. Anal. Mach. Intell.* 11(5):451–476.

Werman, M. and Keren, D. 2001. A Bayesian method for fitting parametric and nonparametric models to noisy data. *IEEE Trans. Patt. Anal. Mach. Intell.* 23(5):528–534.