Projective Reconstruction and Invariants from Multiple Images *

Richard I. Hartley

G.E. CRD, Schenectady, NY, 12301.

Abstract

This paper investigates projective reconstruction of geometric configurations seen in two or more perspective views, and the computation of projective invariants of these configurations from their images. A basic tool in this investigation is the fundamental matrix which describes the epipolar correspondence between image pairs. It is proven that once the epipolar geometry is known, the configurations of many geometric structures (for instance sets of points or lines) are determined up to a collineation of projective 3-space $\mathcal{P}^3$ by their projection in two independent images. This theorem is the key to a method for the computation of invariants of the geometry. Invariants of 6 points in $\mathcal{P}^3$ and of four lines in $\mathcal{P}^3$ are defined and discussed in detail. An example with real images shows that they are effective in distinguishing different geometrical configurations.

Since the fundamental matrix is a basic tool in the computation of these invariants, new methods of computing the fundamental matrix from 7 point correspondences in two images or 6 point correspondences in 3 images are given.

1   Introduction

Projective invariants of geometrical configurations in space have recently received much attention because of their application to vision problems ([11]). Although invariants of a wide range of objects in the 3-dimensional projective space $\mathcal{P}^3$ do exist ([1]), one is restricted in the field of vision to considering those that may be computed from two-dimensional projections (images). For point sets and more structured geometrical objects lying in planes in $\mathcal{P}^3$, many invariants exist ([3]) which can be computed from a single view. Unfortunately, it has been shown in [8] that no invariants of arbitrary point sets in 3-dimensions may be computed from a single image. One is led either to consider constrained sets of points, or else to allow two independent views of the object. An example of

the first approach is contained in [17] which considers solids of revolution. This paper takes the second course and considers invariants that can be derived from two views of an object. Very little previous work has been done in this area.

It has been shown by Longuet-Higgins ([9]) that for calibrated cameras, the relative locations of a set of points in $\mathcal{P}^3$ may be computed from two views using a non-iterative algorithm. This is not quite true of uncalibrated cameras. Theorem 3.5 of this paper shows, however, that the point locations may be computed uniquely up to collineation of $\mathcal{P}^3$, as long as sufficiently many points (at least 8) are given. This is one of the basic results of this paper, since it allows us to compute invariants of point sets in $\mathcal{P}^3$ from two views [2].

As examples of invariants that may be computed from multiple views, invariants of six points and of four lines in $\mathcal{P}^3$ are defined. Experiments are carried out with real images to determine their effectiveness at distinguishing projectively different configurations.

Notation : Vectors are usually represented in homogeneous coordinates, and may be considered as elements of projective 2-space, $\mathcal{P}^2$ or projective 3-space, $\mathcal{P}^3$. Projective transformations are represented by matrices. Two homogeneous vectors, or transformation matrices that differ by a non-zero constant factor are considered to be the same.

Consider a set of points $\{\mathbf{x}_i\}$ in space, visible at image locations $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ in two images. In normal circumstances, the correspondence $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$ will be known, but the location of the original points $\{\mathbf{x}_i\}$ will be unknown.

Nothing will be assumed about the calibration of the two cameras that create the two images. The camera model will be expressed in terms of a general projective transformation from three-dimensional real projective space, $\mathcal{P}^3$, known as object space, to the two-dimensional real projective space $\mathcal{P}^2$ known as image space. The transformation may be expressed in homogeneous coordinates by a 3×4 matrix P known as a camera matrix and the correspondence between points in object space and image space is given by $\mathbf{u}_i = P\mathbf{x}_i$. It is often convenient to write a camera matrix in the form $P = (M \mid \mathbf{t})$ where M is a 3×3 and $\mathbf{t}$ is a column vector.

2 The Fundamental Matrix

For sets of points viewed from two cameras, Longuet-Higgins [9] introduced a matrix that has subsequently become known as the essential matrix. In Longuet-Higgins's treatment, the two cameras were assumed to be calibrated, meaning that the internal cameras parameters were known. It is not hard to show (for instance, see [6]) that this matrix may also be defined for uncalibrated cameras, in which context it is usually referred to as the fundamental matrix. The following basic theorem is proven in [9].

---

[2]This theorem was discovered at about the same time and independently by Faugeras ([4]) and by the present author ([5]).

Theorem (Longuet-Higgins) *Given a set of image correspondences $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}_i'\}$ there exists a 3×3 real matrix F of rank 2 such that $\mathbf{u}_i'^\top F \mathbf{u}_i = 0$ for all i.*

Notice that each image correspondence gives rise to a linear equation in terms of the entries of the matrix F. If 8 or more image correspondences are known, then they lead to a system of 8 or more homogeneous linear equations in the entries of F, from which F may be computed up to an insignificant scale factor.

Factorization of the fundamental matrix : Suppose that the singular value decomposition ([2]) of F is given by $F = UDV^\top$, where D is the diagonal matrix $D = \mathrm{diag}(r, s, 0)$. The following factorization of F may now be verified by inspection.

$$F = SM \; ; \;\; S = UZU^\top \; ; \;\; M = UE\mathrm{diag}(r, s, \alpha)V^\top$$

where

$$E = \left(\begin{array}{ccc} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right) \; ; \;\; Z = \left(\begin{array}{ccc} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

and α is an arbitrary number. The matrix S is skew-symmetric.

Any arbitrary 3×3 skew-symmetric matrix is of the form

$$\left(\begin{array}{ccc} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{array} \right) \tag{1}$$

and it will be convenient to denote this matrix as $[\mathbf{t}]_\times$, where $\mathbf{t}$ is the vector $\mathbf{t} = (t_1, t_2, t_3)^\top$. This notation is motivated by the fact that for any vector $\mathbf{v}$ we have $[\mathbf{t}]_\times \mathbf{v} = \mathbf{t} \times \mathbf{v}$ (the cross product) and $\mathbf{v}[\mathbf{t}]_\times = \mathbf{v} \times \mathbf{t}$. For any non-zero vector $\mathbf{t}$, matrix $[\mathbf{t}]_\times$ has rank 2. Furthermore, the null-space of $[\mathbf{t}]_\times$ is generated by the vector $\mathbf{t}$. This means that $\mathbf{t}^\top [\mathbf{t}]_\times = [\mathbf{t}]_\times \mathbf{t} = 0$ and that any other vector annihilated by $[\mathbf{t}]_\times$ is a scalar multiple of $\mathbf{t}$.

It is evident since α may be arbitrarily chosen that the factorization of F as a product $[\mathbf{t}]_\times M$ is not unique. The following lemma makes this fact clearer.

Lemma 2.1. *Let the fundamental matrix F factor in two different ways as $F = [\mathbf{t}]_\times M = [\mathbf{t}']_\times M'$. Then $\mathbf{t} = \mathbf{t}'$ and $M' = M + \mathbf{t}\mathbf{a}^\top$ for some vector $\mathbf{a}$.*

Proof. First, note that $\mathbf{t}F = \mathbf{t}[\mathbf{t}]_\times M = 0$, and similarly, $\mathbf{t}'F = 0$. Since F has rank 2, it follows that $\mathbf{t} = \mathbf{t}'$ as required. Next, from $[\mathbf{t}]_\times M = [\mathbf{t}]_\times M' = F$ it follows that $[\mathbf{t}]_\times (M' - M) = 0$, and so $M' - M = \mathbf{t}\mathbf{a}^\top$ for some $\mathbf{a}$. Hence, $M' = M + \mathbf{t}\mathbf{a}^\top$ as required. □

3 Projective Reconstruction

For calibrated cameras, Longuet-Higgins showed that the external camera parameters and the point placements may be determined from the fundamental matrix. This is not true in the case of uncalibrated cameras. A basic result of

this paper, however, is that the camera transformation matrices and the point placements may be determined up to a collineation of projective 3-space, $\mathcal{P}^3$.

A matrix F is said to be the fundamental matrix corresponding to a pair of camera matrices P and P' if for every point $\mathbf{x}$ in space with corresponding image points $\mathbf{u} = P\mathbf{x}$ and $\mathbf{u}' = P'\mathbf{x}$ the equation $\mathbf{u}'^\top F\mathbf{u} = 0$ is satisfied. Written in terms of the original point $\mathbf{x}$, this gives $\mathbf{x}^\top P'^\top FP\mathbf{x} = 0$. Since this relationship must hold for all $\mathbf{x}$ we have the following criterion.

Lemma 3.2. *The matrix F is the fundamental matrix corresponding to a pair of camera matrices P and P' if and only if $P'^\top FP$ is skew-symmetric.*

Under this condition we say that the pair (P, P') is a *realisation* of the fundamental matrix F.

Two camera matrices P and P' with different centres of projection, uniquely determine the fundamental matrix F, since one can choose sufficiently many points $\mathbf{x}_i$ not in a critical configuration ([10]). The matrix F is then determined by the equation $\mathbf{u}_i'^\top F\mathbf{u}_i = 0$ where $\mathbf{u}_i = P\mathbf{x}_i$ and $\mathbf{u}_i' = P'\mathbf{x}_i$. This may also be deduced explicitly from the condition in Lemma 3.2.

On the other hand, given matrix F, the matrices P and P' are not uniquely determined. Specifically, if H is a non-singular 4×4 matrix, and $P'^\top FP$ is skew-symmetric, then so is $H^\top P'^\top FPH$. This shows that (P, P') and $(P'H, PH)$ are both realizations of the matrix F. It will be shown that this is the only ambiguity in the realization of a fundamental matrix. Before showing this we need to find explicitly at least one realization of a fundamental matrix F. This is given by the following result

Proposition 3.3. *The fundamental matrix corresponding to a pair of camera matrices $P = (I \mid 0)$ and $P' = (M \mid \mathbf{t})$ is $F = [\mathbf{t}]_\times M$.*

Proof. One simply verifies that

$$P'^\top FP = \begin{pmatrix} M^\top \\ \mathbf{t}^\top \end{pmatrix} [\mathbf{t}]_\times M(I \mid 0) = \begin{pmatrix} M^\top [\mathbf{t}]_\times M & 0 \\ 0 & 0 \end{pmatrix}$$

is skew symmetric. $\qquad\qquad\square$

We now prove our main theorem which indicates when two pairs of camera matrices correspond to the same fundamental matrix.

Theorem 3.4. *Let (P_1, P_1') and (P_2, P_2') be two pairs of camera transforms. Then (P_1, P_1') and (P_2, P_2') correspond to the same fundamental matrix F if and only if there exists a 4×4 non-singular matrix H such that $P_1 H = P_2$ and $P_1' H = P_2'$.*

Proof. The *if* part of this theorem has already been proven, so we turn to the *only if* part. Since each of the matrices P_1 and P_2 has rank 3, we can multiply them (on the right) by suitable matrices H_1 and H_2 to transform them each to the matrix $(I \mid 0)$. If the matrices P_1' and P_2' are also multiplied by H_1 and H_2

respectively, then the fundamental matrix corresponding to the camera matrix pairs are unchanged, as seen previously. Thus, we have reduced the problem to the case where $P_1 = P_2 = (I \mid 0)$.

Suppose therefore, that $P_1 = P_2 = (I \mid 0)$ and that $P_1' = (M_1' \mid \mathbf{t}_1')$ and $P_2' = (M_2' \mid \mathbf{t}_2')$. By proposition 3.3 we have $F = [\mathbf{t}_1']_\times M_1' = [\mathbf{t}_2']_\times M_2'$. According to 2.1 this implies that $\mathbf{t}_1' = \mathbf{t}_2' = \mathbf{t}$ and that $M_2' = M_1' + \mathbf{t}\mathbf{a}^\top$ for some vector $\mathbf{a}$. Let H be the matrix $\begin{pmatrix} I & 0 \\ \mathbf{a}^\top & 1 \end{pmatrix}$. Then one verifies that $(I \mid 0) = (I \mid 0)H$, so $P_2 = P_1 H$. Furthermore, $P_1'H = (M_1' \mid \mathbf{t})H = (M_1' + \mathbf{t}\mathbf{a}^\top \mid \mathbf{t}) = (M_2' \mid \mathbf{t}) = P_2'$. Thus H is the matrix required for the conclusion of theorem 3.4. $\qquad\square$

3.1 Point set reconstruction

Given a pair of camera matrices P and P' and a pair of matched points $\mathbf{u} \leftrightarrow \mathbf{u}'$ it is evident that the space point $\mathbf{x}$ that gives rise to the two matching image points is uniquely defined, and may be obtained by intersecting two rays from the camera centres. Here is a simple way of computing the point $\mathbf{x}$.

Suppose that the fundamental matrix factors as as $F = [\mathbf{t}']_\times M'$, and let $P = (I \mid 0)$ and $P' = (M' \mid \mathbf{t}')$ be a realization of the matrix F. Let $\mathbf{u} \leftrightarrow \mathbf{u}'$ be a pair of matched points in the two images. We wish to find a point $\mathbf{x}$ in space such that $\mathbf{u} = P\mathbf{x}$ and $\mathbf{u}' = P'\mathbf{x}$. From the relation $\mathbf{u}'^\top F\mathbf{u} = \mathbf{u}'^\top[\mathbf{t}']_\times M'\mathbf{u} = \mathbf{u}'^\top(\mathbf{t}' \times M'\mathbf{u}) = 0$, it follows that $\mathbf{u}'$, $M'\mathbf{u}$ and $\mathbf{t}'$ are linearly dependent. If in particular $M'\mathbf{u} = \beta\mathbf{u}' - \alpha\mathbf{t}'$ then we define the corresponding object space point $\mathbf{x}$ to be the point $\begin{pmatrix} \mathbf{u} \\ \alpha \end{pmatrix}$. It is now easily verified that $P\mathbf{x} = (I \mid 0)\mathbf{x} = \mathbf{u}$ and $P'\mathbf{x} = (M' \mid \mathbf{t}')\mathbf{x} = M'\mathbf{u} + \alpha\mathbf{t}' = \mathbf{u}'$ This verifies that the given values of P, P' and $\mathbf{x}_i$ constitute a projective reconstruction of the data.

As shown, $\mathbf{x}$ is determined by the two camera matrices P and P' and the matched points $\mathbf{u} \leftrightarrow \mathbf{u}'$. If we choose a different pair of camera matrices PH and $P'H$ realizing the same fundamental matrix F, then in order to preserve the same pair of matched image points, the point $\mathbf{x}$ must be replaced by $H^{-1}\mathbf{x}$. Thus, changing to a different realization of F results in a projective transformation (namely H^{-1}) of the scene. This proves the following theorem

Theorem 3.5. (Faugeras [4], Hartley et al. [5]) *Given a set of image correspondences $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}_i'\}$ sufficient to determine the fundamental matrix, the corresponding object space coordinates $\{\mathbf{x}_i\}$ may be computed up to a collineation of projective 3-space $\mathcal{P}^3$.*

3.2 Refinement of the Reconstruction

The linear techniques described here for projective reconstruction do not give optimal results, and are sensitive to noise. If better accuracy is required, then the linear solution may be refined using iterative techniques to find an optimal reconstruction. The method used is to use Levenberg-Marquardt iteration ([13]) to minimize the sum-of-square difference between the measured image coordinates $\mathbf{u}_i$ and $\mathbf{u}_i'$, and the predicted values, $\hat{\mathbf{u}}_i = P\mathbf{x}_i$ and $\hat{\mathbf{u}}_i' = P\mathbf{x}_i$. The entries

of the matrix P' as well as the image coordinates $\mathbf{x}_i$ are treated as variable parameters in seeking this best fit, but P is held equal to $(I \mid 0)$. Under an assumption of zero-mean Gaussian noise in measurement of image coordinates, this gives the maximum likelihood reconstruction.

For the most efficient implementation, it is important to take advantage of the block structure of the normal equations to separate the estimation of incremental changes to point coordinates from the estimation of the changes to the camera matrices. This technique is described in [15], and gives an enormous speed increase, particularly when there are many matched points. Convergence of this algorithm is rapid and reliable given an initial guess derived from the linear solution.

4   Reconstruction from Other Configurations

Suppose that the images of n points are known in k views. Since each image point provides two constraints on the location of the object points, this gives a total of $2nk$ constraints. On the other hand, up to collineations of $\mathcal{P}^3$, n points in space have a total of $3n - 15$ degrees of freedom. In addition, the k views have $11k$ degrees of freedom. In order for the positions of the points and the cameras to be determined, we need more equations than degrees of freedom. In summary :

$$\begin{aligned} \# \text{ D.O.F} &= 3n - 15 + 11k \ , \\ \# \text{ constraints} &= 2nk \ . \end{aligned}$$

To solve for the point locations,

$$2nk \geq 3n + 11k - 15 \ . \tag{2}$$

Particular cases show that with $n = 7$ points, $k = 2$ views will suffice, for $n = 6$ points, $k = 3$ views are sufficient. and for $n = 5$ no solution is possible however many views are given. We will now consider the former two cases, and give specific algorithms.

4.1   Seven Points in Two Views

Since multiplication of F by a non-zero scale factor is insignificant, and $\det(F) = 0$, a fundamental matrix F has 7 degrees of freedom. Because of this, it is possible to determine F from only 7 image correspondences.

From 7 image correspondences, we obtain a set of 7 homogeneous linear equations in the entries of F. The solution of these equations has the form

$$q_{ij} = a_{ij}\mu + b_{ij}\nu \tag{3}$$

where μ and ν are unknown and each a_{ij} and b_{ij} is known. Substituting into the equation $\det(F) = 0$ gives rise to a homogeneous cubic equation in the variables μ and ν. Up to scale, this equation has three solution, including complex solutions. Substituting the values of μ and ν back into (3), three possible solutions for the fundamental matrix F are found. Either there are

three real solutions for F, or there are one real and two conjugate complex solutions.

An alternative solution to the problem of 7 points in two images was given in [16], but this solution seems simpler. An implementation and results of reconstruction from seven points has been reported recently in [12].

4.2 Six points in three views.

We begin by considering six points viewed in a single image. Let the points in space be denoted $\mathbf{x}_1, \ldots, \mathbf{x}_6$, and their coordinates in the image be $\mathbf{u}_1, \ldots, \mathbf{u}_6$. If the camera matrix is given by P, then the basic relationship is $\mathbf{u}_i = P\mathbf{x}_i$. We assume that $\mathbf{u}_i = (w_i u_i, w_i v_i, w_i)^\top$ where each u_i and v_i is known, but w_i is not. Further, let the rows of P be vectors $\mathbf{p}_1^\top$, $\mathbf{p}_2^\top$ and $\mathbf{p}_3^\top$. Each point gives rise to three equations

$$w_i u_i = \mathbf{p}_1^\top \mathbf{x}_i \; ; \quad w_i v_i = \mathbf{p}_2^\top \mathbf{x}_i \; ; \quad w_i = \mathbf{p}_3^\top \mathbf{x}_i \; .$$

Cancelling w_i in the usual way leads to two equations

$$u_i \mathbf{p}_3^\top \mathbf{x}_i = \mathbf{p}_1^\top \mathbf{x}_i \; ; \quad v_i \mathbf{p}_3^\top \mathbf{x}_i = \mathbf{p}_2^\top \mathbf{x}_i \; . \tag{4}$$

These equations are linear in the entries of P, and so six points generate a set of 12 equations in 12 unknowns which may be written in the form $A\mathbf{p} = 0$. The vector $\mathbf{p}$ is made up of the entries of the matrix P, and the coefficient matrix A has entries which are linear expressions in the coordinates $(x_i, y_i, z_i, t_i)^\top)$ of the various points $\mathbf{x}_i$. Since this system of equations must have a non-zero solution for $\mathbf{p}$, it follows that $\det(A) = 0$. This gives rise to a polynomial of degree 12 in the x_i, y_i, z_i and t_i. Any set of points which are mapped onto the observed image points by an unknown camera must satisfy this polynomial equation.

Now, by an appropriate choice of projective coordinates, it may be assumed that the first five points $\mathbf{x}_i$ have coordinates $(1,0,0,0)^\top$, $(0,1,0,0)^\top$, $(0,0,1,0)^\top$, $(0,0,0,1)^\top$ and $(1,1,1,1)^\top$. The position of the sixth point $\mathbf{x}_6 = (x, y, z, t)^\top$ is not determined. The equation $\det(A) = 0$ now reduces to a second degree polynomial. This proves the following result.

Proposition 4.6. *Suppose a set of six points $\mathbf{x}_i$ are mapped to points $\mathbf{u}_i$ in an image. If projective coordinates are chosen for $\mathcal{P}^3$ such that points $\mathbf{x}_1, \ldots, \mathbf{x}_5$ have given canonic coordinates, then the sixth point must lie on a quadric surface, determined only by the coordinates of the image points $\mathbf{u}_i$.*

Given three views, the location of the point $\mathbf{x}_6$ is determined as the intersection of three quadric surfaces. There may be as many as 8 solutions. It has been remarked recently by Quan [14] and by Ponce [12] that five of these solutions are the points $\mathbf{x}_1, \ldots, \mathbf{x}_5$, so that there are in fact only three possible solutions for $\mathbf{x}_6$. Results of reconstruction from 6 points in three views are reported in [12, 14].

Once the points $\mathbf{x}_i$ are determined, equations (4) can be used to solve for the camera matrices, and then the fundamental matrices for each pair can be computed from Theorem 3.3.

5 Projective Invariants of Point Sets

As has been shown, although it is impossible to determine the exact geometry of a scene from multiple views, it is in general possible to reconstruct the scene up to an unknown projective transformation of space. Suppose that some mathematical quantity, defined as a function of the scene geometry, may be computed, and that this quantity is unchanged under projective transformations in $\mathcal{P}^3$. Such a quantity is called a projective invariant of the scene geometry. If this quantity is computed from a projective reconstruction of the scene, as discussed in this paper, then it will have the same value as if it were constructed from the actual scene. Such projective invariants do not include such scene properties as angles and length ratios, which are not invariant under projective transformations of the scene. There are, however several projective invariants of certain 3D geometrical configurations, as will be seen shortly. Such invariants may be computed from two or more views of the scene and used for object recognition or for distinguishing between different objects. The general strategy of computing these invariants is as follows.

1. Use image correspondences to compute the fundamental matrix F. Then find a factorization $F = [\mathbf{t}]_\times M$, and hence two camera matrices $P = (I \mid 0)$ and $P' = (M \mid \mathbf{t})$.

2. Compute the projective reconstruction of the scene, using for instance the method of Section 3.1.

3. Compute a projective invariant of the reconstructed scene in $\mathcal{P}^3$.

5.1 An invariant of 6 points

Given a set of six points $\{\mathbf{x}_i\}$ in $\mathcal{P}^3$, a coordinate system may be selected in which the first five points have coordinates $(1, 0, 0, 0)^\top$, $(0, 1, 0, 0)^\top$, $(0, 0, 1, 0)^\top$, $(0, 0, 0, 1)^\top$ and $(1, 1, 1, 1)^\top$. The coordinates of the sixth point give rise to three independent projective invariants of the six points.

Another formulation of these invariants is given by selecting $\mathbf{x}_0$ and $\mathbf{x}_1$ as base points. Given any other point in $\mathcal{P}^3$, not collinear with $\mathbf{x}_0$ and $\mathbf{x}_1$, there exists a unique plane passing through that point and the two base points $\mathbf{x}_0$ and $\mathbf{x}_1$. In this way, the four points $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ and $\mathbf{x}_5$ give rise to four planes all containing the line joining $\mathbf{x}_0$ to $\mathbf{x}_1$. From the four planes it is possible to define a cross ratio. In particular, if λ is any line in space, skew to the line passing through $\mathbf{x}_0$ and $\mathbf{x}_1$, then λ intersects the four planes at points $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$ and $\mathbf{p}_5$. The cross ratio of these four points on the line λ is a projective invariant of the six original points in $\mathcal{P}^3$. Different invariants result from different choices of $\mathbf{x}_0$ and $\mathbf{x}_1$.

Both these definitions of invariants fail if three of the points happen to be collinear, however, this case will be ignored for the sake of simplicity.

6 Projective Invariants of Lines

In this section, invariants of lines in space will be described. It will be shown that four lines in the 3-dimensional projective plane, $\mathcal{P}^3$ give rise to two independent invariants under collineations of $\mathcal{P}^3$. Two different ways of defining invariants will be described, one algebraic and one geometric.

6.1 Computing Lines in Space

To be able to compute invariants of lines in space, it is necessary to be able to compute the location of a line in $\mathcal{P}^3$ from its images in two or more views; Lines in the image plane are represented as 3-vectors. For instance, a vector $\boldsymbol{\lambda} = (\lambda, \mu, \nu)^\top$ represents the line in the plane given by the equation $\lambda u + \mu v + \nu w = 0$. Similarly, a plane in 3-dimensional space is represented in homogeneous coordinates as a 4-dimensional vector. The relationship between lines in the image space and corresponding planes in object space is given by the following lemma.

Lemma 6.7. *The set of all points in $\mathcal{P}^3$ that are mapped by a camera with matrix P onto a line $\boldsymbol{\lambda}$ in the image is the plane π with coordinates $P^\top \boldsymbol{\lambda}$.*

Proof. A point $\mathbf{x}$ lies on π if and only if $P\mathbf{x}$ lies on the line $\boldsymbol{\lambda}$, and so $\boldsymbol{\lambda}^\top P \mathbf{x} = 0$. On the other hand, a point $\mathbf{x}$ lies on the plane π if and only if $\pi^\top \mathbf{x} = 0$. Comparing these two conditions leads to the conclusion that $\pi^\top = \boldsymbol{\lambda}^\top P$ or $\pi = P^\top \boldsymbol{\lambda}$ as required. $\square$

If a line in space is seen in two or more views, then it may be found by computing the intersection of the corresponding planes in space.

6.2 Algebraic Invariant Formulation

Consider four lines λ_i in space. A line may be given by specifying either two points on the line or dually, two planes that meet in the line. It does not matter in which way the lines are described. For instance, in the formulae (6) and (7) below certain invariants of lines are defined in terms of pairs of points on each line. The same formulae could be used to define invariants in which lines are represented by specifying a pair of planes that meet along the line. Since the method of determining lines in space from two view given in section 6.1 gives a representation of the line as an intersection of two planes, the latter interpretation of the formulae is most useful.

Nevertheless, in the following description, of algebraic and geometric invariants of lines, lines will be represented by specifying two points, since this method seems to allow easier intuitive understanding. It should be borne in mind, however, that the dual approach could be taken with no change whatever to the algebra, or geometry.

In specifying lines, each of two points on the line will be given as a 4-tuple of homogeneous coordinates, and so each line λ_i is specified as a pair of 4-tuples

$$\lambda_i = \big((a_{i1}, a_{i2}, a_{i3}, a_{i4})(b_{i1}, b_{i2}, b_{i3}, b_{i4})\big)$$

Now, given two lines λ_i and λ_j, one can form a 4×4 determinant, denoted by

$$|\lambda_i \lambda_j| = \det \begin{pmatrix} a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ b_{i1} & b_{i2} & b_{i3} & b_{i4} \\ a_{j1} & a_{j2} & a_{j3} & a_{j4} \\ b_{j1} & b_{j2} & b_{j3} & b_{j4} \end{pmatrix} . \tag{5}$$

Finally, it is possible to define two independent invariants of the four lines by

$$I_1(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \frac{|\lambda_1 \lambda_2| \ |\lambda_3 \lambda_4|}{|\lambda_1 \lambda_3| \ |\lambda_2 \lambda_4|} \tag{6}$$

and

$$I_2(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \frac{|\lambda_1 \lambda_2| \ |\lambda_3 \lambda_4|}{|\lambda_1 \lambda_4| \ |\lambda_2 \lambda_3|} . \tag{7}$$

It is necessary to prove that the two quantities so defined are indeed invariants under collineations of $\mathcal{P}^3$. First, it must be demonstrated that the expressions do not depend on the specific formulation of the lines. That is, there are an infinite number of ways in which a line may be specified by designating two points lying on it, and it is necessary to demonstrate that choosing a different pair of points to specify a line does not change the value of the invariants. To this end, suppose that $(a_{i1}, a_{i2}, a_{i3}, a_{i4})^\top$ and $(b_{i1}, b_{i2}, b_{i3}, b_{i4})^\top$ are two distinct points lying on a line λ_i, and that $(a'_{i1}, a'_{i2}, a'_{i3}, a'_{i4})^\top$ and $(b'_{i1}, b'_{i2}, b'_{i3}, b'_{i4})^\top$ are another pair of points lying on the same line. Then, there exists a 2×2 matrix D_i such that

$$\begin{pmatrix} a'_{i1} & a'_{i2} & a'_{i3} & a'_{i4} \\ b'_{i1} & b'_{i2} & b'_{i3} & b'_{i4} \end{pmatrix} = D_i \begin{pmatrix} a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ b_{i1} & b_{i2} & b_{i3} & b_{i4} \end{pmatrix} .$$

Consequently,

$$\begin{pmatrix} a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ b_{i1} & b_{i2} & b_{i3} & b_{i4} \\ a_{j1} & a_{j2} & a_{j3} & a_{j4} \\ b_{j1} & b_{j2} & b_{j3} & b_{j4} \end{pmatrix} = \begin{pmatrix} D_i & 0 \\ 0 & D_j \end{pmatrix} \begin{pmatrix} a'_{i1} & a'_{i2} & a'_{i3} & a'_{i4} \\ b'_{i1} & b'_{i2} & b'_{i3} & b'_{i4} \\ a'_{j1} & a'_{j2} & a'_{j3} & a'_{j4} \\ b'_{j1} & b'_{j2} & b'_{j3} & b'_{j4} \end{pmatrix} .$$

Taking determinants, it is seen that the net result of choosing a different representation of the lines λ_i and λ_j is to multiply the value of $|\lambda_i \lambda_j|$ by a factor $\det(D_i) \det(D_j)$. Since each of the lines λ_i appears in both the numerator and denominator of the expressions (6) and (7), the factors will cancel and the values of the invariants will be unchanged.

Next, it is necessary to consider the effect of a change of projective coordinates. If H is a 4×4 invertible matrix representing a coordinate transformation of $\mathcal{P}^3$, then it may be applied to each of the points used to designate the four lines. The result of applying this transformation is to multiply the determinant $|\lambda_i \lambda_j|$ by a factor $\det(H)$. The factors on the top and bottom cancel, leaving the values of the invariants (6) and (7) unchanged. This completes the proof

that I_1 and I_2 defined by (6) and (7) are indeed projective invariants of the set of four lines.

An alternative invariant may be defined by

$$I_3(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \frac{|\lambda_1 \lambda_4| \ |\lambda_2 \lambda_3|}{|\lambda_1 \lambda_3| \ |\lambda_2 \lambda_4|} \ . \tag{8}$$

It is easily seen, that $I_3 = I_1/I_2$. However, if $|\lambda_1 \lambda_2|$ vanishes, then both I_1 and I_2 are zero, but I_3 is in general non-zero. This means that I_3 can not always be deduced from I_1 and I_2. A preferable way of defining the invariants of four lines is as a homogeneous vector

$$I(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (|\lambda_1 \lambda_2| \ |\lambda_3 \lambda_4|, |\lambda_1 \lambda_3| \ |\lambda_2 \lambda_4|, |\lambda_1 \lambda_4| \ |\lambda_2 \lambda_3|) \ . \tag{9}$$

Two such computed invariant values are deemed equal if they differ by a scalar factor. Note that this definition of the invariant avoids problems associated with vanishing or near-vanishing of the denominator in (6) or (7).

6.3   Degenerate Cases

The determinant $|\lambda_i \lambda_j|$ as given in (5) will vanish if and only if the four points involved are coplanar, that is, exactly when the two lines are coincident (meet in space). If all three components of the vector $I(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ given by (9) vanish, then the invariant is undefined. Enumeration of cases indicates that there are two essentially different configurations of lines in which this occurs.

1. Three of the lines lie in a plane.

2. One of the lines meets all the other three.

The configuration where one line meets two of the other lines is not degenerate, but does not lead to very much useful information, since two of the components of the vector vanish. Up to scale, the last component may be assumed to equal 1, which means that two such configurations can not be distinguished. In fact any two such configurations are equivalent under collineation.

6.4   Geometric Invariants of Lines

It is also possible to define projective invariants of sets of four lines geometrically. In particular, given four lines in space in general position, there will exist exactly two *transverse* lines that meet all four of these lines. The cross ratio of the points of intersection of lines with each of the transverse lines give two independent projective invariants of the set of four lines. These invariants may take real or complex values. The relationship of these invariants to the algebraic invariants is clarified in [7]. In particular, it is shown that there are just two independent projective invariants of four lines in space.

7 Experimental Results

Three images of a pair of wooden blocks representing houses were acquired and vertices and edges were extracted. The images are shown in Fig 1. Corresponding edges and vertices were selected by hand from among those detected automatically. There were 13 edges and 15 lines extracted from each of the images. The dotted edges were not visible in all images and were not chosen. Vertices are represented by numbers and edges by letters in the figure. Because of the way edges and vertices were found by the segmentation algorithm, the edges do not always pass precisely through the indicated vertices, but sometimes through a closely neighboring vertex. On other occasions, the full edge was not detected as a single, but was broken into several pieces. This is usual with most edge detection algorithms, and is a source of error in the computation of invariants.

The fundamental matrices F_{12} for the first and second images and F_{23} for the second and third images were computed from the point matches.

7.1 Comparison of Invariant Values

The invariants described in this paper are represented as homogeneous vectors. Two such vectors are considered equivalent if they differ by a non-zero scale factor. Because of arithmetic error and image noise, two computed invariant values will rarely be exactly proportional. In order to compare two such computed invariant values (perhaps when attempting to match an object with a reference object), each homogeneous vector is multiplied by a scale factor chosen to normalize its length to 1. This normalization determines the vector up to a multiplication by a factor ± 1. Two such normalized homogeneous vector invariants $\mathbf{v}_1$ and $\mathbf{v}_2$ are deemed *close* if $\mathbf{v}_1$ is close to $\mathbf{v}_2$ or to $-\mathbf{v}_2$ using a Euclidean norm. Correspondingly, a metric may be defined by

$$d(\mathbf{v}_1, \mathbf{v}_2) = \left(1 - \left|\frac{\mathbf{v}_1 . \mathbf{v}_2}{||\mathbf{v}1|| \, ||\mathbf{v}_2||}\right|\right)^{1/2}. \tag{10}$$

For any $\mathbf{v}_1$ and $\mathbf{v}_2$, distance $d(\mathbf{v}_1, \mathbf{v}_2)$ lies between 0 and 1.

7.2 Invariants of 6 points

The invariants of six points $\{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_6\}$ were computed by finding a projective coordinate frame in which the points $\mathbf{x}_1, \ldots, \mathbf{x}_5$ have coordinates $(1,0,0,0)^\top$, $(0,1,0,0)^\top$, $(0,0,1,0)^\top$, $(0,0,0,1)^\top$ and $(1,1,1,1)^\top$ respectively. The homogeneous coordinates or the sixth point, $\mathbf{x}_6$ in that frame are the desired invariants of the original set of points. Two points are compared using the metric (10). Six sets of six points were chosen for computation of invariants. The sets of points were chosen arbitrarily by hand. The six sets of six points chosen as in the following table which shows the indices of the lines as given in

Fig. 1.

$$S_1 = \{1, \quad 2, \quad 3, \quad 6, \quad 9, \quad 10\} \quad , \quad S_2 = \{2, \quad 4, \quad 6, \quad 8, \quad 10, \quad 12\} \quad ,$$
$$S_3 = \{1, \quad 3, \quad 5, \quad 7, \quad 9, \quad 11\} \quad , \quad S_4 = \{1, \quad 2, \quad 3, \quad 6, \quad 7, \quad 8\} \quad ,$$
$$S_5 = \{1, \quad 4, \quad 7, \quad 10, \quad 13, \quad 12\} \quad , \quad S_6 = \{2, \quad 5, \quad 8, \quad 11, \quad 12, \quad 13\}$$

Table (11) shows the invariant of the sets of six points as computed from the first and second and from the second and third images.

0.0266367	0.970462	0.975994	0.619897	0.847914	0.823575
0.995416	**0.0155304**	0.0648768	0.841029	0.252926	0.548214
0.967114	0.066834	**0.0136234**	0.863063	0.276384	0.516868
0.617346	0.830651	0.873538	**0.0166752**	0.704992	0.752215
0.861618	0.238502	0.289846	0.708237	**0.00561718**	0.590905
0.828638	0.54423	0.519272	0.719518	0.574651	**0.0263892**

$$(11)$$

The (i, j)-th entry of the table shows the distance according to the metric (10) between the invariant of set S_i as computed from images 1 and 2 with that of set S_j as computed from images 2 and 3. The diagonal entries of the matrix (in bold) should be close to 0.0, which indicates a match. The matrix should be approximately symmetric, which is in fact the case.

The off-diagonal entries are not close to zero, except for the $(2, 3)$ entry – but even that entry is greater than the diagonal entries. This indicates that the six-point invariant is very good at discriminating between sets of points with different geometrical structure. Evidently, sets of points S_2 and S_3 are quite similar in arrangement, at least up to collineation.

7.3 Invariants of 4 lines

The same experiment was carried out with six sets of four lines. First the fundamental matrices were computed using point matches and then the line invariant (9) was computed for each pair of line sets and compared using the metric (10).

The sets of lines chosen are given in the following table (refer to Fig. 1).

$$S_1 = \{B, C, J, K\} \quad , \quad S_2 = \{B, G, J, N\}$$
$$S_3 = \{A, B, H, I\} \quad , \quad S_4 = \{B, D, E, G\}$$
$$S_5 = \{A, C, O, J\} \quad , \quad S_6 = \{B, I, L, N\}$$

Table (12 shows the results. The only bad entry in this matrix is in the position $(4, 4)$. This is because of the fact that the four lines chosen contained three coplanar lines (lines B, D and E). This causes the values of the invariant to be indeterminate (that is $(0, 0, 0)$), and shows that such instances must be detected and avoided.

0.0128906	0.674135	0.302728	0.688589	0.642501	0.449448
0.646976	**0.0337898**	0.741489	0.83827	0.706921	0.221636
0.0619738	0.691264	**0.229193**	0.707536	0.708276	0.461339
0.286604	0.607681	0.182331	**0.890303**	0.855833	0.383939
0.656635	0.72182	0.899625	0.718942	**0.00349575**	0.694361
0.473184	0.239022	0.555218	0.947915	0.719282	**0.0332098**

$$(12)$$

figure=imm1.ps,width=2.5in   figure=imm2.ps,width=2.5in
figure=imm3.ps,width=2.5in   figure=houses.ps,width=2.5in

Figure 1: Three views of houses, and numbered selected vertices and lines

Once again, the four-line invariant is shown to be a powerful discriminator between sets of four lines.

References

[1] S. S. Abhyankar. Invariant theory and enumerative combinatorics of young tableaux. In J.L. Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, pages 45 – 76. MIT Press, Boston, MA, 1992.

[2] K.E. Atkinson. *An Introduction to Numerical Analysis, 2nd Edition.* John Wiley and Sons, New York, 1989.

[3] C. Coelho, A. Heller, J. Mundy, D. Forsyth, and A. Zisserman. An experimental evaluation of projective invariants. In J.L. Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, pages 87 – 104. MIT Press, Boston, MA, 1992.

[4] O. D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Computer Vision - ECCV '92, LNCS-Series Vol. 588, Springer-Verlag*, pages 563 – 578, 1992.

[5] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 761–764, 1992.

[6] R. I. Hartley. Estimation of relative camera positions for uncalibrated cameras. In *Computer Vision - ECCV '92, LNCS-Series Vol. 588, Springer-Verlag*, pages 579 – 587, 1992.

[7] R. I. Hartley. Camera calibration using line correspondences. In *Proc. DARPA Image Understanding Workshop*, pages 361–366, 1993.

[8] R. S. Weiss J. B. Burns and E. M. Riseman. The non-existence of general-case view-invariants. In J.L. Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, pages 120 – 134. MIT Press, Boston, MA, 1992.

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

[10] S. J. Maybank. The projective geometry of ambiguous surfaces. *Phil. Trans. R. Soc. Lond.*, A 332:1 – 47, 1990.

[11] J.L. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision.* MIT Press, Boston, MA, 1992.

[12] Jean Ponce, Todd Cass, and David Marimont. Relative stereo and motion reconstruction. Report UIUC-BI-AI-RCV-93-07, Beckman Institute, University of Illinois, 1993.

[13] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, 1988.

[14] Long Quan. Invariants of 6 points from 3 uncalibrated images. In *Computer Vision - ECCV '94, Volume II, LNCS-Series Vol. 801, Springer-Verlag*, pages 459–470, 1994.

[15] C. C. Slama, editor. *Manual of Photogrammetry.* American Society of Photogrammetry, Falls Church, VA, fourth edition, 1980.

[16] Rudolf Sturm. Das problem der projectivität and seine anwendung auf die flachen zweiten grades. *Math. Ann.*, 1:533–574, 1869.

[17] A. Zisserman, D. A. Forsyth, J. L. Mundy, and C. A. Rothwell. Recognizing general curved objects efficiently. In J.L. Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, pages 228 – 251. MIT Press, Boston, MA, 1992.