

Invariants of Points Seen in Multiple Images [★]

Richard I. Hartley

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

Abstract

This paper investigates projective invariants of geometric configurations in 3 dimensional projective space \mathcal{P}^3 , and most particularly the computation of invariants from two or more independent images. A basic tool in this investigation is the essential matrix defined by Longuet-Higgins ([10]), for this matrix 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 \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 essential matrix is a fundamental tool in the computation of these invariants, new methods of computing the essential matrix from 7 point correspondences in two images, 6 point correspondences in 3 images or 13 line correspondences in three images are described.

1 Introduction

Projective invariants of geometrical configurations in space have recently received much attention because of their application to vision problems ([12]). 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 ([5]) which can be computed from a single view. Unfortunately, it has been shown in [4] 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

[★]The research described in this paper has been supported by DARPA Contract #xxx

of points, or else to allow two independent views of the object. An example of the first approach is contained in [15] 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. A previous paper of Barrett et al. ([3]) contains a beginning to the investigation of this subject. One of the results of that paper is presented rather more simply in this paper (Theorem below). Another paper is in preparation ([16]) considering geometrical structures satisfying various constraints. The present paper considers invariants of unstructured lines and points in \mathcal{P}^3 , and shows that under certain circumstances, invariants may be computed.

It has been shown by Longuet-Higgins ([10]) 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 4.10 of this paper shows, however, that the point locations may be computed 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 views².

For sets of lines, the situation is not quite so favourable. It may be seen that virtually no information can be got from two views of a set of lines in space. This is because given two images of a line and two arbitrary cameras, there is always a line in space which corresponds to the two images. In other words, two images of an unknown line do not in any way constrain the cameras. This point is discussed in [14]. On the other hand, if sufficiently many point matches are known as well, then it is possible to determine the locations of the lines, once more up to a collineation of \mathcal{P}^3 . This paper discusses an invariant of four lines in space and how it may be computed. The invariants of four lines may be defined either in algebraic or geometric terms, and greater insight into the properties of the invariants is achieved by considering both styles of definition.

The invariant described in the previous paragraph can not be computed from two views given line matches alone. It is shown in [14] that three views of a set of thirteen lines are sufficient to determine the placement of calibrated cameras, given at least 13 lines. As with Longuet-Higgins results, this result may be extended also to the case of uncalibrated cameras as is shown in section 8 of this paper. The cameras and the corresponding point locations may be computed up to a projective transformation of \mathcal{P}^3 . This allows the computation of invariants of sets of 13 or more lines appearing in three or more views.

Finally, to obtain invariants of point sets in \mathcal{P}^3 , from two views, it is necessary to have at least eight matching points, so as to be able to compute the essential matrix defined by Longuet-Higgins ([10]). On the other hand, projective invariants for sets of 6 points in \mathcal{P}^3 may be defined – they just may not be computed from two views. It is shown, however that from three views of six points invariants may be defined.

One further topic discussed in this paper is the “transfer” problem (section 3). This problem was discussed in [3]. Given a set of eight points as seen in three

²This theorem was discovered at about the same time and independently by Faugeras ([6]) and by the present author ([9]). The two proofs were given within three weeks of each other at separate conferences.

images, and one further point seen in just two of the images, it is possible to compute its position in the third image. I give a somewhat simpler formulation of the method derived in [3] as well as showing how the construction may be generalized to the case where only seven (instead of eight) point matches are known.

1.1 Notation

Consider a set of points $\{\mathbf{x}_i\}$ as seen in two images. Normally, unprimed quantities will be used to denote data associated with the **first image**, whereas primed quantities will denote data associated with the **second image**.

The set of points $\{\mathbf{x}_i\}$ will be visible at image locations $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ in the 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.

Since all vectors are represented in homogeneous coordinates, their values may be multiplied by any arbitrary non-zero factor. The notation \approx is used to indicate equality of vectors or matrices up to multiplication by a scale factor.

Given a vector, $\mathbf{t} = (t_x, t_y, t_z)^\top$ it is convenient to introduce the skew-symmetric matrix

$$[\mathbf{t}]_\times = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} \quad (1)$$

This definition is motivated by the fact that for any vector \mathbf{v} we have $[\mathbf{t}]_\times \mathbf{v} = \mathbf{t} \times \mathbf{v}$ 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} .

The notation A^* represents the adjoint of a matrix A , that is, the matrix of cofactors. If A is an invertible matrix, then $A^* \approx (A^\top)^{-1}$.

1.2 Camera Models

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 $\mathbf{u}_i = P\mathbf{x}_i$.

For convenience it will be assumed throughout this paper that the camera placements are not at infinity, that is, that the projections are not parallel projections. In this case, a camera matrix may be written in the form

$$P = (M \mid -M\mathbf{t})$$

where M is a 3×3 non-singular matrix and \mathbf{t} is a column vector $\mathbf{t} = (t_x, t_y, t_z)^\top$ representing the location of the camera in object space.

2 The Essential Matrix

For sets of points viewed from two cameras, Longuet-Higgins [10] 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 camera parameters were known. It is not hard to show, as was done explicitly in [8] that most of the results also apply to uncalibrated cameras of the type considered in this paper.

2.1 Existence and Characterization

The following basic theorem is proven in [10].

Theorem (Longuet-Higgins) *Given a set of image correspondences $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$ there exists a 3×3 real matrix Q such that*

$$\mathbf{u}'_i{}^\top Q \mathbf{u}_i = 0$$

for all i .

Notice that each image correspondence gives rise to a linear equation in terms of the entries of the matrix Q . Suppose that 8 or more image correspondences are given, then they give rise to a system of 8 or more linear equations which may be expressed as a matrix equation

$$A\mathbf{q} = 0 \tag{2}$$

where

$$\mathbf{q} = (q_{11}, q_{12}, q_{13}, q_{21}, q_{22}, q_{23}, q_{31}, q_{32}, q_{33})^\top \tag{3}$$

and A is a matrix with 9 columns, and one row for each image correspondence. Specifically, writing $\mathbf{u} = (u_i, v_i, 1)^\top$ and $\mathbf{u}' = (u'_i, v'_i, 1)$, the i -th row of the matrix A is equal to the vector

$$(u_i u'_i, v_i u'_i, u'_i, u_i v'_i, v_i v'_i, v'_i, u_i, v_i, 1) . \tag{4}$$

The set of image correspondences will be called **non-degenerate** if the rank of the matrix A is at least 8 (that is, 8 or 9). Geometrical conditions for a set of image correspondences to be non-degenerate were discussed in [10]. The existence of a solution to (2) gives rise to the following observation.

Proposition (Barrett et al. [3]) *Let A be the matrix derived from a set of image correspondences $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$, the i -th row of A being given by (4). Then $\text{rank}(A) \leq 8$. In particular, if the number of correspondences equals 9, then $\det(A) = 0$.*

Proof. If $\text{rank}(A) > 8$ then there is no non-zero solution to the equation $A\mathbf{q} = 0$. \square

If $\text{rank}(A) = 8$ then there is a non-zero solution to the equation $A\mathbf{q} = 0$ unique up to an arbitrary scale factor. In particular, a non-degenerate set of 8 image correspondences determines a unique (up to scale) essential matrix Q . If a set of more than 8 correspondences is given, then in general, due to numerical error, it will not be the case that $\text{rank}(A) \leq 8$. In this case, the matrix Q should be determined by finding the least-squares solution to the equation $A\mathbf{q} = 0$. More specifically, the problem becomes : minimize $\|A\mathbf{q}\|$ subject to the condition $\|\mathbf{q}\| = 1$. The solution \mathbf{q} is the eigenvector corresponding to the least eigenvalue of $A^T A$.

3 Loci of Matched Points

If Q is known and \mathbf{u} a fixed point in the first image then the equation $\mathbf{u}'^T Q \mathbf{u} = 0$ may be viewed as specifying the set of possible points \mathbf{u}' in the second image that are possible matches for \mathbf{u} . This set of points is an epipolar line in the second image. In other words, \mathbf{u}' is on the epipolar line corresponding to \mathbf{u} if and only if $\mathbf{u}'^T Q \mathbf{u} = 0$. This leads to the following interpretation of the essential matrix.

Proposition 3.1. *If Q is an essential matrix corresponding to a pair of images and \mathbf{u} is a point in the first image, then $Q\mathbf{u}$ is the epipolar line in the second image corresponding to \mathbf{u} .*

Barrett et al. [3] applied Proposition to solve the “transfer problem”. In particular, suppose that three images are given, and the essential matrices for each of the image pairs are known. This will be the case if sufficiently many matched points in the three images are known. Suppose that the image of a further point is known in the first two images. Then it is possible to determine its image in the third image. The method of Barrett et al. , though expressed differently, reduces to the following construction. Let \mathbf{x} be a point in space and let \mathbf{u} and \mathbf{u}' be its location as seen in the first two images. Let Q_{13} be the essential matrices corresponding to the first and third images and Q_{23} the one corresponding to the second and third images. According to Proposition 3.1 $Q_{13}\mathbf{u}$ and $Q_{23}\mathbf{u}'$ are the epipolar lines in the third image corresponding to the points \mathbf{u} and \mathbf{u}' . The intersection of these two lines is the location of the point \mathbf{u}'' where \mathbf{x} is seen in the third image. Since line intersection is given by the cross product, we have

$$\mathbf{u}'' = Q_{13}\mathbf{u} \times Q_{23}\mathbf{u}' .$$

This construction depends on the fact that when eight points \mathbf{u}_i in one image are matched with their corresponding points \mathbf{u}'_i in the second image, then the locus of the point \mathbf{u}' matching a further point \mathbf{u} in the first image is a straight line given by Proposition 3.1. Our goal in the rest of this section is, given only seven matched points $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$ and a further point \mathbf{u} in the first image, to determine the locus of the possible locations of the matching point \mathbf{u}' in the second image.

In order to be able to do this, we need the following characterization of essential matrices.

Proposition 3.2. *A 3×3 real matrix Q is an essential matrix if and only if $\text{rank}(Q) = 2$.*

A proof of Proposition 3.2 is given in [7].

Now, suppose that 7 non-degenerate image point correspondences are given and \mathbf{u} is an eighth point in the first image. The corresponding point \mathbf{u}' is unknown. Letting $\mathbf{u}' = (r', s', 1)^\top$, writing down the equations $A\mathbf{q} = 0$ given by (3) and (4) and solving for $\mathbf{q} = (q_{11}, \dots, q_{33})$ results in an essential matrix Q with entries that are linear expressions in r' and s' . Now the condition $\det(Q) = 0$ derived from Proposition 3.2 gives rise to a cubic equation in r' and s' . This equation describes the locus in the second image of all points that may correspond to \mathbf{u} . The form of the cubic equation is somewhat special, however, as will now be shown.

Since multiplication of Q by a non-zero scale factor is insignificant, and $\det(Q) = 0$, an essential matrix Q has 7 degrees of freedom. Because of this, it is possible to determine Q from only 7 image correspondences. A method is given in [7], and will be briefly described here.

From 7 image correspondences, we obtain 7 linear equations in the entries of Q . Since the scale of Q is arbitrary, a further equation $q_{11} = 1$ is available. (The difficulty that q_{11} may equal zero is discussed in [7] and need not concern us here). From these eight equations in nine variables (the entries of Q) a solution may be found of the form

$$q_{ij} = a_{ij}\mu + b_{ij} \quad (5)$$

where μ is unknown and each a_{ij} and b_{ij} is known. Substituting into the equation $\det(Q) = 0$ gives rise to a cubic equation in the variable μ . This equation has three solutions, including complex solutions. Substituting the values of μ back into (5), three possible solutions for the essential matrix Q are found. There are two cases. Either there are three real solutions for Q , or there are one real and two conjugate complex solutions. Let the solutions be Q_0 , Q_1 and Q_2 .

Now, considering the eighth correspondence $\mathbf{u} \leftrightarrow \mathbf{u}'$, it follows that $\mathbf{u}'^\top Q_i \mathbf{u} = 0$ where Q_i is one of Q_0 , Q_1 and Q_2 . Multiplying these relationships together gives an equation

$$(\mathbf{u}'^\top Q_0 \mathbf{u})(\mathbf{u}'^\top Q_1 \mathbf{u})(\mathbf{u}'^\top Q_2 \mathbf{u}) = 0.$$

This is just the cubic equation described previously, namely the locus of the point \mathbf{u}' . As can be seen, the cubic equation factors into linear factors over the complex field. Either there are three real factors, or there are one real and two conjugate complex factors. In other words, the locus of \mathbf{u}' is either three real lines in the plane, or one real line and two complex lines.

Let us investigate complex lines. Writing as before $\mathbf{u}' = (r', s', 1)^\top$, consider a line $\alpha r' + \beta s' + 1 = 0$ where α and β are complex. It is easily seen that there exists either one or no real points $(r', s', 1)^\top$ satisfying this equation. That is, a complex line contains at most one real point. Now we sum up this discussion.

Theorem 3.3. *Let $\{\mathbf{u}_i\} \leftrightarrow \{u'\}$ be a set of 7 non-degenerate image correspondences and let \mathbf{u} be a further point. The locus of (real) points \mathbf{u}' in the second image corresponding to the point \mathbf{u} in the first image consists either of three straight lines, or of a single straight line and a single isolated point (counted*

twice). The single isolated point corresponds to a complex essential matrix and is not realizable.

4 Projective Invariants

For calibrated cameras, Longuet-Higgins showed that the external camera parameters and the point placements may be determined from the essential matrix. This is not true in the case of uncalibrated cameras. It will be shown, however, that the camera transformation matrices and the point placements may be determined up to a collineation of projective 3-space, \mathcal{P}^3 .

4.1 Realization of the Essential Matrix.

First, we consider the inverse question of determining the essential matrix given the two camera transformation matrices. The following result was proven in [8].

Proposition 4.4. *The essential matrix corresponding to a pair of camera matrices $P = (M \mid -M\mathbf{t})$ and $P' = (M' \mid -M'\mathbf{t}')$ is given by*

$$Q \approx M'^* M^\top [M(\mathbf{t}' - \mathbf{t})]_\times .$$

For a proof of Proposition 4.4 see [8].

Definition 4.5. A pair of camera transformations $P = (M \mid -M\mathbf{t})$ and $P' = (M' \mid -M'\mathbf{t}')$ are called a *realization* of the essential matrix Q if $Q \approx M'^* M^\top [M(\mathbf{t}' - \mathbf{t})]_\times$.

Our present goal is to characterize all possible realizations of a given essential matrix. As is indicated by Proposition 4.4, an essential matrix Q factors into a product $Q = RS$, where R is a non-singular matrix and S is skew-symmetric. The next lemma shows to what extent this factorization is unique.

Lemma 4.6. *Let the 3×3 matrix Q factor in two different ways as $Q \approx R_1 S_1 \approx R_2 S_2$ where each S_i is a non-zero skew-symmetric matrix and each R_i is non-singular. Then $S_2 \approx S_1$. Furthermore, if $S_i = [\mathbf{t}]_\times$ then $R_2 \approx R_1 + \mathbf{a}\mathbf{t}^\top$ for some vector \mathbf{a} .*

Proof. Since R_1 and R_2 are non-singular, it follows that $Q\mathbf{t} = 0$ if and only if $S_i\mathbf{t} = 0$. From this it follows that the null-spaces of the matrices S_1 and S_2 are equal, and so $S_1 \approx S_2$. For the second statement, assume that $Q = R_1[\mathbf{t}]_\times = R_2[\mathbf{t}]_\times$. Then, $(R_1 - R_2)[\mathbf{t}]_\times = 0$, and so $R_1 - R_2 = \mathbf{a}\mathbf{t}^\top$ as required. \square

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

Theorem 4.7. *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 essential matrix Q if and only if there exists a 4×4 non-singular matrix H such that $P_1 H \approx P_2$ and $P'_1 H \approx P'_2$.*

Proof. First we prove the **if** part of this theorem. To this purpose, let $\{\mathbf{x}_i\}$ be a set of at least 8 points in 3-dimensional space and let $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ be the corresponding image-space points as imaged by the two camera P_1 and P'_1 . By the definition of the essential matrix, Q satisfies the condition $\mathbf{u}'_i{}^\top Q \mathbf{u}_i = 0$ for all i . We may assume that the points $\{\mathbf{x}_i\}$ have been chosen in such a way that the matrix Q is uniquely defined up to scale by the above equation. The point configurations that defeat this definition of the essential matrix are discussed in [10]. Suppose now that there exists a 4×4 matrix H taking P_1 to P_2 and P'_1 to P'_2 in the sense specified by the hypotheses of the theorem. For each i let $\mathbf{x}_i^{(2)} = H^{-1}\mathbf{x}_i$. Then we see that

$$P_2\mathbf{x}_i^{(2)} = P_1HH^{-1}\mathbf{x}_i = P_1\mathbf{x}_i = \mathbf{u}_i$$

and

$$P'_2\mathbf{x}_i^{(2)} = P'_1HH^{-1}\mathbf{x}_i = P'_1\mathbf{x}_i = \mathbf{u}'_i .$$

In other words, the image points $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ are a matched point set with respect to the cameras P_2 and P'_2 . Thus the essential matrix for this pair of cameras is defined by the same relationship $\mathbf{u}'_i{}^\top Q \mathbf{u}_i = 0$ that defines the essential matrix of the pair P_1 and P'_1 . Consequently, the two camera pairs have the same essential matrix.

Now, we turn to the **only if** part of the theorem and assume that two pairs of cameras have the same essential matrix, Q . First, we consider the camera pair $\{(M_1 | -M_1\mathbf{t}_1), (M'_1 | -M'_1\mathbf{t}'_1)\}$. It is easily seen that the 4×4 matrix

$$\begin{pmatrix} M_1^{-1} & \mathbf{t}_1 \\ 0 & 1 \end{pmatrix}$$

transforms this pair to the camera pair

$$\{(I | 0), (M'_1M_1^{-1} | -M'_1(\mathbf{t}'_1 - \mathbf{t}_1))\}$$

where I and 0 are identity matrix and zero column vector respectively. Furthermore by the **if** part of this theorem (or as verified directly using Lemma 4.4), this new camera pair has the same essential matrix as the original.

Applying this transformation to each of the camera pairs

$$\{(M_1 | -M_1\mathbf{t}_1), (M'_1 | -M'_1\mathbf{t}'_1)\}$$

and

$$\{(M_2 | -M_2\mathbf{t}_2), (M'_2 | -M'_2\mathbf{t}'_2)\}$$

we see that there is a 4×4 matrix transforming one pair to the other if and only if there is such a matrix transforming

$$\{(I | 0), (M'_1M_1^{-1} | -M'_1(\mathbf{t}'_1 - \mathbf{t}_1))\}$$

to

$$\{(I | 0), (M'_2M_2^{-1} | -M'_2(\mathbf{t}'_2 - \mathbf{t}_2))\}$$

Thus, we are reduced to proving the theorem for the case where the first cameras, P_1 and P_2 of each pair are both equal to $(I | 0)$. Thus, let $\{(I | 0), (M_1 | -M_1\mathbf{t}_1)\}$ and $\{(I | 0), (M_2 | -M_2\mathbf{t}_2)\}$ be two pairs of cameras corresponding to the same essential matrix. According to Lemma 4.4, the Q -matrices corresponding to the two pairs are $M_1^*[\mathbf{t}_1]_\times$ and $M_2^*[\mathbf{t}_2]_\times$ respectively, and these must be

equal (up to scale). According to Lemma 4.6, $\mathbf{t}_1 \approx \mathbf{t}_2$ and $M_2^* \approx M_1^* + \mathbf{a}\mathbf{t}_1^\top$ for some vector \mathbf{a} . Taking the transpose of this last relation yields

$$M_2^{-1} \approx M_1^{-1} + \mathbf{t}_1\mathbf{a}^\top \quad (6)$$

At this point we need to interpolate a lemma.

Lemma 4.8. *For any column vector \mathbf{t} and row vector \mathbf{a}^\top , if $I + \mathbf{t}\mathbf{a}^\top$ is invertible then*

$$(I + \mathbf{t}\mathbf{a}^\top)^{-1} = I - k\mathbf{t}\mathbf{a}^\top$$

where $k = 1/(1 + \mathbf{a}^\top\mathbf{t})$.

Proof. The proof is done by simply multiplying out the two matrices and observing that the product is the identity. One might ask what happens if $\mathbf{a}^\top\mathbf{t} = -1$ in which case k is undefined. The answer is that in that case, $I + \mathbf{t}\mathbf{a}^\top$ is singular, contrary to hypothesis. Details are left to the reader. \square

Now we may continue with the proof of the theorem. Referring back to (6), it follows that

$$\begin{aligned} M_2 &\approx (M_1^{-1} + \mathbf{t}_1\mathbf{a}^\top)^{-1} \\ &\approx (M_1^{-1}(I + M_1\mathbf{t}_1\mathbf{a}^\top))^{-1} \\ &\approx (I - kM_1\mathbf{t}_1\mathbf{a}^\top)M_1 \\ &\approx M_1 - kM_1\mathbf{t}_1(\mathbf{a}^\top M_1) \end{aligned}$$

and

$$\begin{aligned} M_2\mathbf{t}_1 &\approx M_1\mathbf{t}_1 - kM_1\mathbf{t}_1(\mathbf{a}^\top M_1\mathbf{t}_1) \\ &\approx k'M_1\mathbf{t}_1 \approx M_1\mathbf{t}_1 \end{aligned} \quad (7)$$

where $k' = 1 - k\mathbf{a}^\top M_1\mathbf{t}_1$. Since $\mathbf{t}_2 \approx \mathbf{t}_1$ according to Lemma 4.6, $M_2\mathbf{t}_2 \approx M_1\mathbf{t}_1$. From these results, it follows that

$$(M_2 \mid -M_2\mathbf{t}_2) \approx (M_1 \mid -M_1\mathbf{t}_1) \begin{pmatrix} I & 0 \\ k\mathbf{a}^\top M_1 & k'' \end{pmatrix}$$

for some constant k'' .

This completes the proof of the theorem. \square

4.2 Choosing a Realization of Q .

Given a set of image correspondences $\mathbf{u}_i \leftrightarrow \mathbf{u}'_i$ defining an essential matrix Q , Theorem 4.7 shows that one cannot unambiguously determine the position of the cameras, or the corresponding object-space points from Q . Since Q contains all the information that is available from the point correspondences, it follows that the position of the cameras and the object points can be determined only up to a 3-dimensional projective transform as specified by the matrix H . In order to determine the positions of the object-space points $\{\mathbf{x}_i\}$ unambiguously, it is necessary for some ground-control points to be specified as discussed in [9].

In this paper, we will not be interested in absolute determination of the points $\{\mathbf{x}_i\}$.

Our strategy, therefore, is to select *any* pair of camera placements consistent with the essential matrix, Q .

Provided we can factor an essential matrix Q into a product $Q = RS$ as promised by Proposition 3.2, then we can find a realization of Q as follows:

Proposition 4.9. *If $Q = R[\mathbf{t}]_{\times}$ is a factorization of an essential matrix into a product of a non-singular matrix R and a skew-symmetric matrix $[\mathbf{t}]_{\times}$, then one realization of Q is given by the pair of camera matrices $P = (I \mid 0)$ and $P' = (R^* \mid -R^*\mathbf{t})$.*

It is in no way intended that this should represent the true placement of the cameras. Nevertheless, according to Theorem 4.7 it is related to the true camera placement by a 3-dimensional projective transformation.

Thus finding a realization of Q comes down to finding a factorization. To this purpose, suppose that the singular value decomposition ([2]) of Q is given by $Q = UDV^{\top}$, where D is the diagonal matrix $D = \text{diag}(r, s, 0)$. In a practical case, the smallest singular value of Q will not be exactly equal to 0 because of numerical inaccuracies. However, setting the smallest singular value to 0 gives the matrix closest to Q in Euclidean norm that has the required rank 2. The following factorization of Q may now be verified by inspection.

$$Q = RS ; \quad R = U\text{diag}(r, s, \gamma)EV^{\top} ; \quad S = VZV^{\top}$$

where

$$E = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} ; \quad Z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and γ is any non-zero number, but is best chosen to lie between r and s so that the condition number [2] of R is as good as possible.

4.3 Computation of 3-D Points.

The point in the object space that projects on to $\mathbf{u}_i = (u_i, v_i, 1)^{\top}$ and $\mathbf{u}'_i = (u'_i, v'_i, 1)^{\top}$ in the two images, under the transforms P and P' , can be computed by solving the equations

$$(w_i u_i, w_i v_i, w_i)^{\top} = P(x_i, y_i, z_i, 1)^{\top}$$

$$(w'_i u'_i, w'_i v'_i, w'_i)^{\top} = P'(x_i, y_i, z_i, 1)^{\top} .$$

The values of u_i, v_i, u'_i, v'_i, P and P' are known, whereas x_i, y_i, z_i, w_i and w'_i are unknown. Thus we have 6 equations in 5 unknowns and the vector $\mathbf{x}_i = (x_i, y_i, z_i)^{\top}$ that minimizes the error can be computed.

4.4 Definition of Invariants

Next, invariants will be considered that can be derived from a set of corresponding image points. Consider a set of image correspondences $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$

sufficient to allow the computation of an essential matrix Q . The Q matrix may be obtained from 7 or 8 point correspondences, or from various other configurations such as 6 points correspondences of which 4 points are in the plane ([16]) or from 13 line correspondences in 3 images (see section 8).

According to the Proposition 4.9, it is easy to find a realization of Q as a pair of camera matrices, P_0 and P'_0 . Once P_0 and P'_0 are given, it is an easy matter to compute the actual coordinates of object space points \mathbf{x}_i . The computed positions \mathbf{x}_i satisfy the equations $\mathbf{u}_i = P_0\mathbf{x}_i$ and $\mathbf{u}'_i = P'_0\mathbf{x}_i$ and they are uniquely determined by these conditions as long as \mathbf{u}_i , \mathbf{u}'_i and the two camera matrices P_0 and P'_0 are given. If a different pair of camera matrices are given, then the computed values of the points \mathbf{x}_i will change. Thus, let P_1 and P'_1 be a different realization of Q . According to Theorem 4.7 there exists a non-singular 4×4 matrix H such that $P_1 = P_0H$ and $P'_1 = P'_0H$. One now verifies that $P_1(H^{-1}\mathbf{x}_i) = P_0\mathbf{x}_i = \mathbf{u}_i$ and similarly $P'_1(H^{-1}\mathbf{x}_i) = \mathbf{u}'_i$. This means that $\{H^{-1}\mathbf{x}_i\}$ are the locations of the object space points corresponding to the new realization of Q by P_1 and P'_1 . This gives the following result.

Theorem 4.10. (Faugeras [6], Hartley et al. [9]) *Given a set of image correspondences $\{\mathbf{u}_i\} \leftrightarrow \{\mathbf{u}'_i\}$ sufficient to determine the essential matrix, the corresponding object space coordinates $\{\mathbf{x}_i\}$ may be computed up to a collineation of projective 3-space \mathcal{P}^3 .*

This theorem allows us to compute projective invariants associated with the image correspondences. The general method is as follows

1. Use the image correspondences to compute the essential matrix Q
2. Select some realization of Q by camera matrices P and P' . The realization given in Proposition 4.9 is a possible choice.
3. Compute the object space coordinates $\{\mathbf{x}_i\}$ corresponding to the given camera matrices using for instance the method of Section 4.3.
4. Compute a projective invariant of the points $\{\mathbf{x}_i\}$ in \mathcal{P}^3 .

5 Invariants of point sets in \mathcal{P}^3

In this section some of the projective invariants of point sets in \mathcal{P}^3 will be investigated. In particular, a projective invariant of a set of six points $\{\mathbf{x}_i\}$ in \mathcal{P}^3 will be described.

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 .³

This is the analogy one dimension higher of the well known invariant of 5 points in a plane. Given 5 points \mathbf{x}_i in \mathcal{P}^2 , an invariant may be defined by selecting one of the points \mathbf{x}_0 and joining it to each of the other points in the plane. The cross ratio of the set of four lines so formed is a projective invariant of the original five points.

There is another way in which invariants may be defined. Five points in general position in the plane may be used to define a unique conic. The conic may be parametrized by a parameter θ and this parametrization may be done in such a way that three of the points have fixed known parameter values, 0, 1 and ∞ . The parameters for the other two points may be denoted by α and β , and these two values are independent invariants of the set of five points.

An analogous method of describing the invariants of six points in \mathcal{P}^3 also holds. In particular, given 6 points in \mathcal{P}^3 in general position, there exists a unique twisted cubic c that passes through the six points ([13]), and c may be parametrized by a parameter θ in such a way that three of the points receive parameters 0, 1 and ∞ . The parameters of the other three points will then be α, β and γ , and these values are projective invariants of the set of six points.

6 Line Invariants

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 locations of the lines in \mathcal{P}^3 from their images in two views. In general, this is impossible as remarked in [14] unless other information is available. Therefore, it will be assumed here that the essential matrix Q corresponding to the two images is known. This may be derived from a sufficient number of point correspondences, or else from line correspondences, as shown in section 8. From the matrix Q , two camera transformations M and M' realizing Q can be computed as in section 4.2.

Lines in the image plane are represented as 3-vectors. For instance, a vector $\mathbf{l} = (l, m, n)^\top$ represents the line in the plane given by the equation $lu + mv + nw =$

³Both 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.

0. Similarly, planes in 3-dimensional space are represented in homogeneous coordinates as a 4-dimensional vector $\pi = (p, q, r, s)^\top$.

The relationship between lines in the image space and the corresponding plane in object space is given by the following lemma.

Lemma 6.11. *Let λ be a line in \mathcal{P}^3 and let the image of λ as taken by a camera with transformation matrix P be \mathbf{l} . The locus of points in \mathcal{P}^3 that are mapped onto the image line \mathbf{l} is a plane, π , passing through the camera centre and containing the line λ . It is given by the formula $\pi = P^\top \mathbf{l}$.*

Proof. A point \mathbf{x} lies on π if and only if it is mapped to a point on the line \mathbf{l} by the action of the transformation matrix. This means that $P\mathbf{x}$ lies on the line \mathbf{l} , and so

$$\mathbf{l}^\top P\mathbf{x} = 0 . \tag{8}$$

On the other hand, a point \mathbf{x} lies on the plane π if and only if $\pi^\top \mathbf{x} = 0$. Comparing this with (8) lead to the conclusion that $\pi^\top = \mathbf{l}^\top P$ or $\pi = P^\top \mathbf{l}$ as required. \square

Now, given two images \mathbf{l} and \mathbf{l}' of a line λ in space as taken by two cameras with camera matrices P and P' , the line λ is the intersection of the planes $P^\top \mathbf{l}$ and $P'^\top \mathbf{l}'$. This line was computed assuming a particular realization of the essential matrix Q by P and P' . As with points, the choice of a different realization of Q will correspond to a collineation of \mathcal{P}^3 . The positions of a set of lines seen in the two images will be determined by Q up to a collineation.

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 (10) and (11) 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

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

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

$$|\lambda_i \lambda_j| = \det \begin{pmatrix} a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ b_{i1} & b_{i2} & b_{i3} & b_{i4} \\ a_{j1} & a_{j2} & a_{j3} & a_{j4} \\ b_{j1} & b_{j2} & b_{j3} & b_{j4} \end{pmatrix}. \quad (9)$$

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

$$I_1(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \frac{|\lambda_1 \lambda_2| |\lambda_3 \lambda_4|}{|\lambda_1 \lambda_3| |\lambda_2 \lambda_4|} \quad (10)$$

and

$$I_2(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \frac{|\lambda_1 \lambda_2| |\lambda_3 \lambda_4|}{|\lambda_1 \lambda_4| |\lambda_2 \lambda_3|}. \quad (11)$$

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 (10) and (11), 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 (10) and (11) unchanged. This completes the proof that I_1 and I_2 defined by (10) and (11) are indeed projective invariants of the set of four lines.

An alternative invariant may be defined by

$$I_3(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \frac{|\lambda_1 \lambda_4| |\lambda_2 \lambda_3|}{|\lambda_1 \lambda_3| |\lambda_2 \lambda_4|}. \quad (12)$$

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

$$I(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (|\lambda_1\lambda_2| \quad |\lambda_3\lambda_4|, |\lambda_1\lambda_3| \quad |\lambda_2\lambda_4|, |\lambda_1\lambda_4| \quad |\lambda_2\lambda_3|) \quad . \quad (13)$$

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 (10) or (11).

The definitions of I_1 and I_2 are similar to the definition of the cross-ratio of points on a line. It is well known that for four points on a line, there is only one independent invariant. It may be asked whether I_1 may be obtained from I_2 by some simple arithmetic combination. This is not the case, as will become clearer when the connection of these algebraic invariants with geometric invariants is shown.

6.3 Degenerate Cases

The determinant $|\lambda_i\lambda_j|$ as given in (9) 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 (13) 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

Consider four lines λ_i in general position (which means that they are not coincident) in \mathcal{P}^3 . It will be shown that there exist exactly two further lines τ_1 and τ_2 , called *transversals*, which meet each of the four lines. Once this is established, it is easy to define invariants.

The points of intersection of each of the four lines λ_i with one of the transversals τ_j constitute a set of four points on a line in \mathcal{P}^3 . The cross ratio of these points is an invariant of the four lines λ_i . In this way, two invariants may be defined, one for each of the two transversals.

Invariants may be defined in a dual manner as follows. Given a transversal, τ_j , meeting each of the lines λ_i , there exists, for each λ_i a plane denoted $\langle \tau_j, \lambda_i \rangle$,

containing τ_j and λ_i . This gives rise to a set of four planes meeting in a common line τ_j . The cross-ratio of this set of planes is an invariant of the lines λ_i .

It is easy to see that this dual construction does not give rise to any new invariant. Specifically, consider the cross-ratio of the four planes meeting at τ_1 . The cross-ratio of four planes meeting along a line is equal to the cross-ratio of the points of intersection of the planes with any other non-coincident line in space. The line τ_2 is such a line. Hence, the cross ratio of the planes $\langle \tau_1, \lambda_i \rangle$ is equal to the cross-ratio of the points $\langle \tau_1, \lambda_i \rangle \cap \tau_2$, where the symbol \cap denotes the point of intersection. However, plane $\langle \tau_1, \lambda_i \rangle$ meets τ_2 in the point $\lambda_i \cap \tau_2$. In other words, the cross-ratio of the four planes meeting along τ_1 is equal to the cross-ratio of the four points along τ_2 , and vice-versa.

6.5 Existence of Transversals

To prove the existence of transversals, we start by considering three lines in space.

Lemma 6.12. *There exists a unique quadric surface containing three given lines λ_1, λ_2 and λ_3 in general position in \mathcal{P}^3 .*

Proof. For a reference to properties of quadric surfaces, the reader is referred to [13]. It is shown there that a quadric surface is a doubly ruled surface containing two families of lines A and B . Two lines from the same set A or B do not meet, whereas any two lines chosen one from each set will always meet. Assuming that the lines λ_i lie on a quadric surface, since they do not meet, they must all come from the same family, which we assume to be A . Now consider any point \mathbf{x} on the quadric surface. There is a unique line passing through \mathbf{x} and belonging to the class B . This line must meet each of the lines λ_i , which belong to class A .

We are led therefore to consider the locus of all points \mathbf{x} in \mathcal{P}^3 for which there exists a line passing through \mathbf{x} meeting all the lines λ_i . To this end, let $\mathbf{x} = (x, y, z, t)^\top$ be a point on this locus. For each of the lines λ_i we may define a plane π_i passing through \mathbf{x} and λ_i . The condition that there exists a line passing through \mathbf{x} meeting each λ_i means that the three planes π_i meet along that line.

Next, we formulate this last condition algebraically and give a method of computing the formula for the quadric surface. As before, letting $(a_{i1}, a_{i2}, a_{i3}, a_{i4})^\top$ and $(b_{i1}, b_{i2}, b_{i3}, b_{i4})^\top$ be two points on the line λ_i , the plane π_i passing through $\mathbf{x} = (x, y, z, t)^\top$ and the line λ_i may be computed as follows. Consider the matrix

$$\begin{pmatrix} a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ b_{i1} & b_{i2} & b_{i3} & b_{i4} \\ x & y & z & t \end{pmatrix} \quad (14)$$

The plane π_i is given by the homogeneous vector $(p_{i1}, p_{i2}, p_{i3}, p_{i4})^\top$ where $(-1)^j p_{ij}$ is the determinant of the 3×3 matrix obtained by deleting the j -th column of (14). Consequently, each p_{ij} is a homogeneous linear expression in x, y, z and t . Furthermore, since point $(x, y, z, t)^\top$ lies on this plane it follows

that

$$xp_{i1} + yp_{i2} + zp_{i3} + tp_{i4} = 0 . \quad (15)$$

Now the fact that the three planes π_j meet along a common line translates into the algebraic fact that the rank of the matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix}$$

is 2. This is equivalent to the condition

$$\det(P^{(j)}) = 0 \quad \text{for all } j , \quad (16)$$

where $P^{(j)}$ is the matrix obtained by removing the j -th column of P . Since each entry p_{ij} of P is a linear homogeneous expression in the variables x, y, z and t , the determinant $\det(P^{(j)})$ is a cubic homogeneous polynomial. A point on the required locus must satisfy the condition $\det(P^{(j)}) = 0$ for $j = 1, \dots, 4$. However, because of condition (15) these four equations are not independent. In particular, if \mathbf{p}_j represents the j -th column of P , then (15) implies a relation

$$x\mathbf{p}_1 + y\mathbf{p}_2 + z\mathbf{p}_3 + t\mathbf{p}_4 = 0$$

Then

$$\begin{aligned} x \det(P^{(4)}) &= x \det(\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3) \\ &= \det(x\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3) \\ &= \det(-y\mathbf{p}_2 - z\mathbf{p}_3 - t\mathbf{p}_4 \ \mathbf{p}_2 \ \mathbf{p}_3) \\ &= \det(-t\mathbf{p}_4 \ \mathbf{p}_2 \ \mathbf{p}_3) \\ &= -t \det(\mathbf{p}_2 \ \mathbf{p}_3 \ \mathbf{p}_4) \\ &= -t \det(P^{(1)}) . \end{aligned} \quad (17)$$

This equation implies that x divides $\det(P^{(1)})$ and t divides $\det(P^{(4)})$. Furthermore, applying the same argument to other coordinates gives rise to an equation

$$\det(P^{(1)})/x = -\det(P^{(2)})/y = \det(P^{(3)})/z = -\det(P^{(4)})/t = R(x, y, z, t)$$

where $R(x, y, z, t)$ is some homogeneous degree-2 polynomial. Then the defining equations (16) of the locus become

$$xR(x, y, z, t) = yR(x, y, z, t) = zR(x, y, z, t) = tR(x, y, z, t) = 0 . \quad (18)$$

This implies that either $R(x, y, z, t) = 0$ or $x = y = z = t = 0$. The latter condition can be discounted, since $(0, 0, 0, 0)$ is not a valid set of homogeneous coordinates. Consequently, the desired locus is described by the degree-2 polynomial equation $R(x, y, z, t) = 0$, and is therefore a quadric surface. Since it is easily verified that the four original lines λ_i lie on this surface, the proof of the lemma is complete. \square

It is now a simple matter to prove the existence of transversals.

Theorem 6.13. *There exist exactly two transversals to four lines in general position in \mathcal{P}^3 .*

Proof. We choose three of the lines λ_1 , λ_2 and λ_3 and construct the quadric surface S that they all line on. Let \mathbf{x}_1 and \mathbf{x}_2 be the two points of intersection of the fourth line λ_4 with the quadric surface. The construction of S in Lemma 6.12 shows that any transversal to lines λ_1 , λ_2 and λ_3 must lie on S . Further, the lines λ_1 , λ_2 and λ_3 all belong to one of the families, A , of ruled lines on the quadric surface, S . Let τ_1 and τ_2 be the lines in the other family B passing through \mathbf{x}_1 and \mathbf{x}_2 . Then τ_1 and τ_2 are the two transversals to all four lines. \square

Of course, it is possible that λ_4 does not meet the surface S in any real point, or is tangent to S . The statement of the theorem must be interpreted as allowing complex or double solutions. In the case of four real lines in space, there are either two real transversals or two conjugate complex transversals. In the case of complex transversals, there is no conceptual difficulty in defining the invariants as in the real case. The cross-ratio of points of intersections of the lines with the two conjugate transversals will result in two invariants which are complex conjugates of each other.

Various degenerate sets of lines also allow two transversals. For instance suppose that λ_1 and λ_2 are coincident, and so are λ_3 and λ_4 . One transversal to the four lines passes through the two points of intersection of the pairs of lines. The other transversal is the line of intersection of the two planes defined by λ_1 , λ_2 and by λ_3 , λ_4 . The cross-ratio invariant corresponding to the first transversal is zero, but the invariant corresponding to the second transversal is in general non-zero and is a useful invariant for this geometric configuration. This is similar to what happens for the algebraically defined invariants (see Section 6.2).

6.6 Independence and Completeness

I shall now show that the two geometrically defined invariants are independent and together completely characterize the set of four lines up to a collineation of \mathcal{P}^3 .

To show independence, we start by selecting τ_1 and τ_2 , two arbitrary non-intersecting lines in space to serve as transversals. Next, we mark off points \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{a}_3 and \mathbf{a}_4 along τ_1 in such a way that their cross ratio is equal to any arbitrarily chosen invariant value. Similarly, mark off along τ_2 points \mathbf{b}_1 , \mathbf{b}_2 , \mathbf{b}_3 and \mathbf{b}_4 having another arbitrarily chosen cross-ratio invariant value. Now, joining \mathbf{a}_i to \mathbf{b}_i for each i gives a set of four lines having the two arbitrarily chosen invariants.

Next, it will be shown that the two invariants completely characterize the set of four lines up to a collineation. Consequently, let four lines in space have two given cross-ratio invariant values with respect to transversals τ_1 and τ_2 respectively. Let the points of intersection of the four lines with τ_1 be \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{a}_3 and \mathbf{a}_4 and the intersection points with τ_2 be \mathbf{b}_1 , \mathbf{b}_2 , \mathbf{b}_3 and \mathbf{b}_4 . Let a second set of lines with the same invariants be given, with transversals τ'_j and intersection points \mathbf{a}'_i and \mathbf{b}'_i . Our goal is to demonstrate that there is a collineation taking τ_j to τ'_j for $j = 1, 2$, taking points \mathbf{a}_i to \mathbf{a}'_i and \mathbf{b}_i to \mathbf{b}'_i for $i = 1, \dots, 4$. It will follow that the collineation takes one set of lines λ_i onto the other set.

Choosing two points on each of τ_1 and τ_2 , four points in all, and two points on

each of τ'_1 and τ'_2 a further four points, there exists a collineation taking the first set of four points to the second set, and hence taking τ_1 to τ'_1 and τ_2 to τ'_2 . Suppose that this collineation takes \mathbf{a}_i to \mathbf{a}''_i and \mathbf{b}_i to \mathbf{b}''_i , it remains to be shown that there exists a collineation preserving τ'_1 and τ'_2 and taking \mathbf{a}''_i to \mathbf{a}'_i and \mathbf{b}''_i to \mathbf{b}'_i . Without loss of generality it may be assumed that τ'_1 is the line $x = y = 0$ and that τ'_2 is the line $z = t = 0$. With this choice, we see that a collineation of \mathcal{P}^3 represented by a matrix of the form $\begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix}$, where each H_j is a 2×2 block, maps each τ'_j to itself. Furthermore each H_j represents a homography of the line τ'_j . Since the points \mathbf{a}'_i and \mathbf{a}''_i on τ'_1 have the same cross-ratio, there is a homography of τ'_1 taking \mathbf{a}'_i to \mathbf{a}''_i for $i = 1, \dots, 4$, and the same can be said for the points \mathbf{b}'_i and \mathbf{b}''_i on τ_2 . Hence by independent choice of the two 2×2 matrices H_1 and H_2 , both mappings can be carried out simultaneously and the proof is complete.

6.7 Existence of an Isotropy

Four lines in \mathcal{P}^3 can be represented by a total of 16 independent parameters. On the other hand, there are 15 degrees of freedom for collineations of \mathcal{P}^3 . This suggests that there should be only one invariant for four lines in space, but we have seen that there are two. The discrepancy arises because of the existence of an isotropy ([11]). To understand this, we need to determine the subgroup of all collineations of \mathcal{P}^3 that fix four given lines. Any such collineation will also fix the two transversals as well as the four points of intersection of the lines with each transversal. Since four points on each transversal are fixed, every point on the transversal must be fixed. This shows that a collineation of \mathcal{P}^3 fixes four given lines if and only if it fixes the two transversals pointwise. Assuming as before that the two transversals are the lines $x = y = 0$ and $z = t = 0$, it is easily seen that a collineation fixes the transversals pointwise if and only if it is represented by a matrix of the form $\text{diag}(k_1, k_1, k_2, k_2)$ where k_1 and k_2 are two independent constants. Allowing for an arbitrary scale factor in the matrix, this implies that there is a one-parameter subgroup of collineations fixing the four lines. This reduces the number of degrees of freedom of the group action of collineations of \mathcal{P}^3 on sets of four lines in space to 14, and explains why there are two independent invariants.

6.8 Relationship of Geometric to Algebraic Invariants

The fact that for real lines the algebraic invariants defined in Section 6.2 must be real whereas the geometric invariants may be complex indicates that they are not the same. However, since the geometric invariants completely determine the four lines up to collineation, it must be possible to determine the algebraic invariants given the values of the geometric ones. Consider four lines with geometric invariants α and β . We desire to determine the values of the algebraic invariants given by (13). To this end, we may assume that the transversals are the lines $x = y = 0$ and $z = t = 0$ and that the points of intersections of the

four lines with the transversals have coordinates

$$\begin{aligned}\mathbf{a}_2 &= (0, 0, 0, 1)^\top \\ \mathbf{a}_1 &= (0, 0, \alpha, 1)^\top \\ \mathbf{a}_3 &= (0, 0, 1, 1)^\top \\ \mathbf{a}_4 &= (0, 0, 1, 0)^\top\end{aligned}$$

and

$$\begin{aligned}\mathbf{b}_2 &= (0, 1, 0, 0)^\top \\ \mathbf{b}_1 &= (\beta, 1, 0, 0)^\top \\ \mathbf{b}_3 &= (1, 1, 0, 0)^\top \\ \mathbf{b}_4 &= (1, 0, 0, 0)^\top .\end{aligned}$$

These points have cross-ratio invariants α and β on the transversal lines $x = y = 0$ and $z = t = 0$ respectively.

From this it is easy to compute the value of the invariant (13) to be

$$I = (\alpha\beta, 1, 1 + \alpha\beta - \alpha - \beta) . \quad (19)$$

Hence, it is easy to compute the algebraic invariants from the geometric ones. Similarly, given I , it is easy to solve (19) for α and β , which indicates that the algebraic invariant (13) is complete.

7 Other Configurations

Since projective invariants exist for six points in \mathcal{P}^3 , it would be convenient if such invariants could be computed from just two view of six points. The method described requires the computation of the essential matrix in order to compute invariants of point configurations. The computation of the essential matrix requires eight points, or at the very least seven points, with possible ambiguity as described in section 3. This does not mean that invariants can not be computed in other ways. This question will be investigated now.

We begin by considering six points viewed in a single image. Let the points in space be denoted $\mathbf{x}_1, \dots, \mathbf{x}_6$, and their coordinates in the image be $\mathbf{u}_1, \dots, \mathbf{u}_6$. 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

$$\begin{aligned}w_i u_i &= \mathbf{p}_1^\top \mathbf{x}_i \\ w_i v_i &= \mathbf{p}_2^\top \mathbf{x}_i \\ w_i &= \mathbf{p}_3^\top \mathbf{x}_i .\end{aligned}$$

Cancelling w_i in the usual way leads to two equations

$$\begin{aligned}u_i \mathbf{p}_3^\top \mathbf{x}_i &= \mathbf{p}_1^\top \mathbf{x}_i \\ v_i \mathbf{p}_3^\top \mathbf{x}_i &= \mathbf{p}_2^\top \mathbf{x}_i .\end{aligned} \quad (20)$$

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 7.14. *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 $\mathbf{x}_1, \dots, \mathbf{x}_5$ 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 .*

If the set of points are seen in two views, then in a canonic coordinate system, the sixth point must lie on the intersection of two quadric surfaces, which in general will be a fourth-degree curve. For three views, the sixth point must lie on the intersection of three quadric surfaces. In general three quadric surfaces will meet in 8 points, including complex points. The points can be computed by solving a set of three simultaneous second degree equations. This gives the following corollary to Proposition 7.14.

Proposition 7.15. *The spatial locations of almost all sets of six points in \mathcal{P}^3 , can be determined up to collineations of \mathcal{P}^3 and up to 8-fold ambiguity by their locations in three images.*

Once the points x_i are determined, equations (20) can be used to solve for the camera matrices, and then the essential matrices for each pair can be computed from Theorem 4.4.

7.1 Degrees of Freedom

The previous argument can be formulated in terms of degrees of freedom. Suppose that the images of n points are known in k views. As was shown above, this gives rise to $2nk$ equations. 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 to be determined, we need more equations than degrees of freedom. In summary :

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

To solve for the point locations,

$$2nk \geq 3n + 11k - 15 . \quad (21)$$

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. These results confirm the previous results of this paper.

For lines, the situation is not so favourable. Suppose that n lines are visible in k views. As with points, each line in each view gives rise to two equations. In particular, suppose λ is a line in \mathcal{P}^3 and \mathbf{l} is the image of that line as seen by a camera with camera matrix P . Let \mathbf{x} be a point on λ , then as shown in (8) $\mathbf{l}P\mathbf{x} = 0$. Since the line λ can be specified by two points, two independent equations arise.

On the other hand, each line in \mathcal{P}^3 has four degrees of freedom, so up to collineations, n lines have a total of $4n - 15$ degrees of freedom, as long as $n \geq 5$.⁴ In summary :

$$\begin{aligned} \# \text{ D.O.F} &= 4n - 15 + 11k , \\ \# \text{ equations} &= 2nk . \end{aligned}$$

To solve for the point locations,

$$2nk \geq 4n + 11k - 15 . \quad (22)$$

In particular for 6 lines at least 9 views are necessary. On the other hand, for just 3 views, at least 9 lines are necessary.

As with points, once the lines are known, the camera matrices may be computed using (8), and the essential matrices of each pair may be computed using Theorem 4.4.

I have shown that being able to compute locations of points and lines up to collineation of \mathcal{P}^3 is equivalent to being able to compute the essential matrix for each pair of views. Consequently the bounds given in (21) and (22) are minimum requirements for the computation of the essential matrices of all the views. The necessity for at least 9 lines in 3 views just demonstrated should be compared with section 8 in which a linear method is given for computing Q from 13 lines in 3 views.

8 Determination of the Essential Matrix from Line Correspondences

This section will investigate the computation of the essential matrix of an uncalibrated camera from a set of line correspondences in three views. As discussed in [14], no information whatever about camera placements may be derived from any number of line-to-line correspondences in two views. In [14] the motion and structure problem from line correspondences is considered. An assumption made in that paper is that the camera is calibrated so that a pixel in each image corresponds to a uniquely specified ray in space relative to the location and placement of the camera. It will be shown in this section that this assumption is not necessary and that in fact the same approach can be adapted to apply to the computation of the essential matrix for uncalibrated cameras.

⁴As shown in section 6.7 four lines have two degrees of freedom

It will be assumed that three different views are taken of a set of fixed lines in space. That is, it is assumed that the cameras are moving and the lines are fixed, which is opposite to the assumption made in [14]. It will not even be assumed that the images are taken with the same camera. Thus the three cameras are uncalibrated and possibly different. The notation used in this section will be similar to that used in [14]. Since we are now considering three cameras, the different cameras will be distinguished using subscripts rather than primes. Consequently, the three cameras will be represented by matrices

$$(M_0 | 0) , (M_1 | -M_1 \mathbf{t}_1) \quad \text{and} \quad (M_2 | -M_2 \mathbf{t}_2)$$

where \mathbf{t}_1 and \mathbf{t}_2 are the positions of the cameras with respect to the position of the zero-th camera, and M_i is a non-singular matrix for each i . For convenience, the coordinate system has been chosen so that the origin is at the position of the zero-th camera, and so $\mathbf{t}_0 = 0$.

Now, consider a line in space passing through a point \mathbf{x} and with direction given by a vector ℓ . Let N_i be the normal to the plane passing through the center of the i -th camera and the line. Then, N_i is given by the expression

$$N_i = (\mathbf{x} - \mathbf{t}_i) \times \ell = \mathbf{x} \times \ell - \mathbf{t}_i \times \ell .$$

Then for $i = 1, 2$,

$$\begin{aligned} N_0 \times N_i &= (\mathbf{x} \times \ell) \times (\mathbf{x} \times \ell - \mathbf{t}_i \times \ell) \\ &= -(\mathbf{x} \times \ell) \times (\mathbf{t}_i \times \ell) \\ &= -((\mathbf{x} \times \ell) \cdot \ell) \mathbf{t}_i - ((\mathbf{x} \times \ell) \cdot \mathbf{t}_i) \ell \\ &= (N_0 \cdot \mathbf{t}_i) \ell \end{aligned} \tag{23}$$

However, for $i = 1, 2$,

$$\begin{aligned} N_i \cdot \mathbf{t}_i &= ((\mathbf{x} - \mathbf{t}_i) \times \ell) \cdot \mathbf{t}_i \\ &= (\mathbf{x} \times \ell) \cdot \mathbf{t}_i - (\mathbf{t}_i \times \ell) \cdot \mathbf{t}_i \\ &= N_0 \cdot \mathbf{t}_i \end{aligned}$$

Combined with the result of (23) this yields the expression

$$N_0 \times N_i = (N_i \cdot \mathbf{t}_i) \ell \tag{24}$$

for $i = 1, 2$. From this it follows, as in [14] that

$$(N_2 \cdot \mathbf{t}_2) N_0 \times N_1 = (N_1 \cdot \mathbf{t}_1) N_0 \times N_2 \tag{25}$$

Now, let \mathbf{n}_i be the representation in homogeneous coordinates of the image of the line ℓ in the i -th view. According to Lemma 6.11, N_i is the normal to the plane $(M_i | -M_i \mathbf{t}_i)^\top \mathbf{n}_i$. Consequently,

$$N_i = M_i^\top \mathbf{n}_i .$$

Applying this to (25) lead to

$$(\mathbf{n}_2^\top M_2 \mathbf{t}_2)(M_0^\top \mathbf{n}_0 \times M_1^\top \mathbf{n}_1) = (\mathbf{n}_1^\top M_1 \mathbf{t}_1)(M_0^\top \mathbf{n}_0 \times M_2^\top \mathbf{n}_2) \tag{26}$$

We now state without proof a simple formula concerning cross products :

Lemma 8.16. *If M is any 3×3 matrix, and \mathbf{u} and \mathbf{v} are column vectors, then*

$$(M\mathbf{u}) \times (M\mathbf{v}) = M^*(\mathbf{u} \times \mathbf{v}) . \quad (27)$$

Applying (27) to each of the two cross products in (26) leads to

$$M_0^{-1}(\mathbf{n}_2^\top M_2 \mathbf{t}_2)(\mathbf{n}_0 \times M_0^* M_1^\top \mathbf{n}_1) = M_0^{-1}(\mathbf{n}_1^\top M_1 \mathbf{t}_1)(\mathbf{n}_0 \times M_0^* M_2^\top \mathbf{n}_2) . \quad (28)$$

Now, cancelling M_0^{-1} from each side and combining the two cross products into one gives

$$\mathbf{n}_0 \times ((\mathbf{n}_2^\top M_2 \mathbf{t}_2) M_0^* M_1^\top \mathbf{n}_1 - (\mathbf{n}_1^\top M_1 \mathbf{t}_1) M_0^* M_2^\top \mathbf{n}_2) = 0 . \quad (29)$$

As in [14], we write

$$B = (\mathbf{n}_2^\top M_2 \mathbf{t}_2) M_0^* M_1^\top \mathbf{n}_1 - (\mathbf{n}_1^\top M_1 \mathbf{t}_1) M_0^* M_2^\top \mathbf{n}_2 \quad (30)$$

then $\mathbf{n}_0 \times B = 0$. Now, writing

$$\begin{aligned} M_0^* M_1^\top &= \begin{pmatrix} \mathbf{r}_1^\top \\ \mathbf{r}_2^\top \\ \mathbf{r}_3^\top \end{pmatrix} \\ M_0^* M_2^\top &= \begin{pmatrix} \mathbf{s}_1^\top \\ \mathbf{s}_2^\top \\ \mathbf{s}_3^\top \end{pmatrix} \\ M_1 \mathbf{t}_1 &= \mathbf{t} \\ M_2 \mathbf{t}_2 &= \mathbf{u} \end{aligned} \quad (31)$$

vector B can be written in the form

$$B = \begin{pmatrix} \mathbf{n}_1^\top (\mathbf{r}_1 \mathbf{u}^\top - \mathbf{t} \mathbf{s}_1^\top) \mathbf{n}_2 \\ \mathbf{n}_1^\top (\mathbf{r}_2 \mathbf{u}^\top - \mathbf{t} \mathbf{s}_2^\top) \mathbf{n}_2 \\ \mathbf{n}_1^\top (\mathbf{r}_3 \mathbf{u}^\top - \mathbf{t} \mathbf{s}_3^\top) \mathbf{n}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{n}_1^\top E \mathbf{n}_2 \\ \mathbf{n}_1^\top F \mathbf{n}_2 \\ \mathbf{n}_1^\top G \mathbf{n}_2 \end{pmatrix} . \quad (32)$$

Where E , F and G are defined by this formula. Therefore, we have the basic equation

$$\mathbf{n}_0 \times \begin{pmatrix} \mathbf{n}_1^\top E \mathbf{n}_2 \\ \mathbf{n}_1^\top F \mathbf{n}_2 \\ \mathbf{n}_1^\top G \mathbf{n}_2 \end{pmatrix} = 0 . \quad (33)$$

This is essentially the same as equation (2.13) in [14], derived here, however, for the case of uncalibrated cameras. As remarked in [14], for each line ℓ , equation (33) gives rise to two linear equations in the entries of E , F and G . Given 13 lines it is possible to solve for E , F and G , up to a common scale factor.

We now define a matrix Q_{01} by

$$Q_{01} = (\mathbf{t} \times \mathbf{r}_1, \mathbf{t} \times \mathbf{r}_2, \mathbf{t} \times \mathbf{r}_3)$$

This may be written as $Q_{01} = [\mathbf{t}]_\times (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. Then, we see that

$$Q_{01}^\top = - \begin{pmatrix} \mathbf{r}_1^\top \\ \mathbf{r}_2^\top \\ \mathbf{r}_3^\top \end{pmatrix} [\mathbf{t}]_\times$$

and in view of the definitions of \mathbf{r}_i and \mathbf{t} given in (31), we have

$$Q_{01}^\top = M_0^* M_1^\top [M_1 \mathbf{t}_1]_\times$$

from which it follows, using Proposition 4.4 that Q_{01} is the essential matrix corresponding to the (ordered) pair of transformation matrices $(M_0 | 0)$ and $(M_1 | -M_1 \mathbf{t}_1)$.

From the definition of $E = \mathbf{r}_1 \mathbf{u}^\top - \mathbf{t} \mathbf{s}_1^\top$ it follows that $E^\top (\mathbf{t} \times \mathbf{r}_1) = 0$. If E has rank 2, then $(\mathbf{t} \times \mathbf{r}_1)$ can be determined up to an unknown scale factor. In the same way, if F and G have rank 2, then $(\mathbf{t} \times \mathbf{r}_i)$ can be similarly determined. Since these three vectors are the columns of the essential matrix Q_{01} , it means that Q_{01} can be determined up to individually scaling its columns. How to handle the case where E , F or G does not have rank 2 is discussed in [14].

Now, by interchanging the roles of the first and second cameras in this analysis, it is possible to determine the matrix Q_{10} up to individual scalings of its columns. However, since $Q_{01} = Q_{10}^\top$ the matrix Q_{01} can be determined up to scale.

9 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 Figures 1, 2, and 3. Corresponding edges and vertices were selected by hand from among those detected automatically. The edges and vertices shown in Fig. 4 were chosen. 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 essential matrices Q_{12} for the first and second images and Q_{23} for the second and third images were computed from the point matches.

9.1 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

$$d(\mathbf{v}_1, \mathbf{v}_2) = \left(1 - \left| \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\| \|\mathbf{v}_2\|} \right| \right)^{1/2}. \quad (34)$$

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

9.2 Invariants of 6 points

The invariants of six points $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_6\}$ were computed by finding a projective coordinate frame in which the points $\mathbf{x}_1, \dots, \mathbf{x}_5$ 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 of 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 (34). 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 lines chosen as in the following table which shows the indices of the lines as given in Fig. 4.

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

Table (35) 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

(35)

The (i, j) -th entry of the table shows the distance according to the metric (34) 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.

9.3 Invariants of 4 lines

The same experiment was carried out with six sets of four lines. First the essential matrices were computed using point matches and then the line invariant (13) was computed for each pair of line sets and compared using the metric (34).

The sets of lines chosen are given in the following table (refer to Fig. 4).

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

Table (36 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.

$$\begin{array}{cccccc}
 \mathbf{0.0128906} & 0.674135 & 0.302728 & 0.688589 & 0.642501 & 0.449448 \\
 0.646976 & \mathbf{0.0337898} & 0.741489 & 0.83827 & 0.706921 & 0.221636 \\
 0.0619738 & 0.691264 & \mathbf{0.229193} & 0.707536 & 0.708276 & 0.461339 \\
 0.286604 & 0.607681 & 0.182331 & \mathbf{0.890303} & 0.855833 & 0.383939 \\
 0.656635 & 0.72182 & 0.899625 & 0.718942 & \mathbf{0.00349575} & 0.694361 \\
 0.473184 & 0.239022 & 0.555218 & 0.947915 & 0.719282 & \mathbf{0.0332098}
 \end{array}
 \tag{36}$$

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

Acknowledgement I am indebted to Joe Mundy for introducing me to the subject of projective invariants, and for many enlightening conversations during the preparation of this paper.

Figure 1. First view of houses

Figure 2. Second view of houses

Figure 3. Third view of houses

Figure 4. Selected vertices and edges

References

- [1] S. S. Abhyankar, “*Invariant Theory and Enumerative Combinatorics of Young Tableaux*”, In [12], pp. 53–90, (1992).
- [2] K.E. Atkinson, “An Introduction to Numerical Analysis,” John Wiley and Sons, 2nd Edition, 1989.
- [3] E. B. Barrett, Michael H. Brill, Nils N. Haag and Paul M. Payton, “*Invariant Linear Methods in Photogrammetry and Model Matching*” In [12], pp. 319–336, (1992).
- [4] J. B. Burns, R. S. Weiss and E. M. Riseman, “*The Non-existence of General-case View-invariants*”, In [12], pp. 143–156, (1992).
- [5] C. Coelho, A. Heller, J. Mundy, D. Forsyth and A. Zisserman, “*An Experimental Evaluation of Projective Invariants*”, In [12], pp. 103–124, (1992).
- [6] Faugeras, O., “What can be seen in three dimensions with an uncalibrated stereo rig?”, Proc. of ECCV-92, G. Sandini Ed., LNCS-Series Vol. 588, Springer- Verlag, 1992, pp. 563 – 578.
- [7] R. Hartley, “An investigation of the essential matrix”, GE internal report, available upon request, preparing for publication.
- [8] R. Hartley, “Estimation of Relative Camera Positions for Uncalibrated Cameras,”, Proc. of ECCV-92, G. Sandini Ed., LNCS-Series Vol. 588, Springer- Verlag, 1992, pp. 579 – 587.
- [9] R. Hartley, R. Gupta and Tom Chang, “Stereo from Uncalibrated Cameras” Proceedings of CVPR92.
- [10] H.C. Longuet-Higgins, “*A computer algorithm for reconstructing a scene from two projections*,” Nature, Vol. 293, 10 Sept. 1981.
- [11] J. L. Mundy and A. Zisserman. “Introduction – towards a new framework for vision” In [12], pp. 1–49.
- [12] J. L. Mundy and A. Zisserman (editors), “Geometric Invariance in Computer Vision,”, MIT Press, Cambridge Ma, 1992.
- [13] J.G. Semple and G. T. Kneebone “*Algebraic Projective Geometry*” Oxford University Press, (1952), ISBN 0 19 8531729.
- [14] J. Weng, T.S. Huang, and N. Ahuja, “Motion and Structure from Line Correspondences: Closed-Form Solution, Uniqueness and Optimization”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 14, No. 3, March, 1992.
- [15] A. Zisserman, D. A. Forsyth, J. L. Mundy and C. A. Rothwell, “*Recognizing General Curved Objects Efficiently*”, In [12], pp. 265–290, (1992).
- [16] A. Zisserman, R. Hartley, J. Mundy, P. Beardsley, “Projective Structure From Multiple Views”, in preparation.