Camera Calibration Using Line Correspondences ⋆

Richard I. Hartley

G.E. CRD, Schenectady, NY, 12301.
Ph: (518)-387-7333
Fax : (518)-387-6845
Email : hartley@crd.ge.com

Abstract

In this paper, a method of determining the essential matrix for uncalibrated cameras is given, based on line matches in three images. The three cameras may have different unknown calibrations, and the essential matrices corresponding to each of the three pairs of cameras may be determined. Determination of the essential matrix for uncalibrated cameras is important, forming the basis for many algorithms such as computation of invariants image rectification, camera calibration and scene reconstruction.

In the case where all the three cameras are assumed to have the same unknown calibration, the method of Faugeras and Maybank ([3, 4]) may be used to calibrate the camera. The scene may then be reconstructed exactly (up to a scaled Euclidean transformation). This extends previous results of Weng, Huang and Ahuja ([12]) who gave a method for scene reconstruction from 13 line correspondences using a calibrated camera. The present paper shows that the the camera may be calibrated at the same time that the scene geometry is determined.

⋆The research described in this paper has been supported by DARPA Contract #MDA972-91-C-0053

1 Introduction

A traditional approach to analysis of perspective images in the field of Computer Vision has been to attempt to measure and model the camera that took the image. A large body of literature has grown up seeking to calibrate the camera and determine its parameters as a preliminary step to image understanding. The papers [1] and [14] represent two of the latest approaches to camera calibration. A recent view ([2]) is that camera calibration is not desirable or necessary in many image understanding situations. Many authors have been led to consider uncalibrated cameras. The study of projective invariants ([11]) is an example of a growing field based on the philosophy of avoiding camera calibration. In fact, study of uncalibrated cameras is intimately linked with the study of projective invariants, for a result of Faugeras ([2]), Hartley et. al. ([6]) and Mohr shows that under most conditions a scene can be determined up to a projective transform of projective 3-space $\mathcal{P}^3$by a pair of images taken by uncalibrated cameras.

Central to the study of pairs of images is the essential matrix, introduced by Longuet-Higgins ([10]) for calibrated cameras, but easily extended to uncalibrated cameras. The essential matrix encodes the epipolar correspondences between two images. It has been shown to be a key tool in scene reconstruction from two uncalibrated views ([2, 6]) as well as for the computation of invariants ([7]). The task of image rectification, which seeks to line up epipolar lines in a pair of images as a preliminary step to finding image correspondences, can be accomplished using the uncalibrated essential matrix ([9]) where previous methods have relied on camera modelling. It is particularly interesting that the essential matrix may be used for the calibration of a camera, and consequent scene reconstruction, given three or more views ([3, 4]). This result provides a strong argument for not assuming camera calibration *a priori*, and underlines the central rôle or the essential matrix.

A recent paper Weng, Huang and Ahuja ([12]) gave an algorithm for reconstructing a scene from a set of at least 13 line correspondences in three images. They assumed a calibrated camera in their algorithm. It is the purpose of the present paper to extend their result to uncalibrated cameras, showing that the essential matrices can be computed from three uncalibrated views of a set of lines. It is not assumed that the three cameras all have the same calibration. In fact, the essential matrices corresponding to each of the three image pairs may be computed. In the case where all three cameras are the same, however, the result of Faugeras and Maybank ([3, 4]) may be applied to obtain the complete calibration of the three cameras and reconstruct the scene up to a scaled Euclidean transformation. Thus, it is shown that the assumption of calibrated cameras is unnecessary in [12], for the cameras may be calibrated at the same time that the scene is reconstructed.

One unfortunate aspect of the algorithm [12] is that 13 line correspondences in three images are necessary, compared with eight point correspondences (and with some effort only six, [8]). Although nothing can be done with two views or fewer (see [12]), a counting argument shows that as few as nine lines in three views may be sufficient, although it is extremely unlikely that a linear or closed form algorithm can be found in this case. It is shown in section 4 of this paper that if four of the lines are known to lie in a plane, then a linear solution exists with only nine lines.

2 Preliminaries

2.1 Notation

Consider a set of points $\{\mathbf{x}_i\}$ as seen in two images. 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. 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*.

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

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

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 $\mathbf{v}[\mathbf{t}]_\times = \mathbf{v} \times \mathbf{t}$.

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

2.2 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 $\mathbf{u}_i \approx P\mathbf{x}_i$.

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.

2.3 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 cameras parameters were known. It is not hard to show (for instance see[5]) that most of the results also apply to uncalibrated cameras of the type considered in this paper.

The following basic theorem is proven in [10].

Theorem 2.1. (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*

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

for all i*.*

The matrix Q is called the essential matrix. Next, we consider the question of determining the essential matrix given the two camera transformation matrices. The following result was proven in [5].

Proposition 2.2. *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 2.2 see [5].

2.4 Computing Lines in Space

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 $lu + mv + nw = 0$. 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 2.3. *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

$$\mathbf{l}^{\top} P \mathbf{x} = 0 \ . \qquad (2)$$

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 (2) lead to the conclusion that $\pi^{\top} = \mathbf{l}^{\top} P$ or $\pi = P^{\top}\mathbf{l}$ as required. □

2.5 Degrees of Freedom

In this section, we compute how many views of a set of lines are necessary to determine the positions of the lines in space. Suppose that n unknown lines are visible in k views with unknown camera matrices. Suppose that the images of the lines in each of the k views are known. 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 (2) $\mathbf{l}^\top P\mathbf{x} = 0$. Since the line λ can be specified by two points, two independent equations arise. The total number of equations is therefore equal to $2nk$.

On the other hand, each line in $\mathcal{P}^3$has four degrees of freedom, so up to projectivity, n lines have a total of $4n - 15$ degrees of freedom, as long as $n \geq 5$.[2] Furthermore, each camera matrix has 11 degrees of freedom. In summary :

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

To solve for the line locations,

$$2nk \geq 4n + 11k - 15 \;. \tag{3}$$

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.

Once the lines are known, the camera matrices may be computed using (2), and the essential matrices of each pair may be computed using Theorem 2.2.

The bounds given by (3) 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 3 in which a linear method is given for computing Q from 13 lines in 3 views. Also, compare with section 4 in which a linear method is given for computing Q under the assumption that four of the lines are coplanar.

3 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 [12], no information whatever about camera placements may be derived from any number of line-to-line correspondences in two views. In [12] 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.

[2]As shown in [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 [12]. 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 [12]. 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

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

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 $i = 1, 2$,

$$\begin{aligned} N_0 \times N_i &= (\mathbf{x} \times \ell) \times (\mathbf{x} \times \ell - \mathbf{t}_i \times \ell) \\ &= -(\mathbf{x} \times \ell) \times (\mathbf{t}_i \times \ell) \\ &= -((\mathbf{x} \times \ell) \,.\, \ell)\mathbf{t}_i - ((\mathbf{x} \times \ell) \,.\mathbf{t}_i)\, \ell) \\ &= (N_0.\mathbf{t}_i)\, \ell \end{aligned} \tag{4}$$

However, for $i = 1, 2$,

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

Combined with the result of (4) this yields the expression

$$N_0 \times N_i = (N_i.\mathbf{t}_i)\, \ell \tag{5}$$

for $i = 1, 2$. From this it follows, as in [12] that

$$(N_2.\mathbf{t}_2) N_0 \times N_1 = (N_1.\mathbf{t}_1) N_0 \times N_2 \tag{6}$$

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 2.3, N_i is the normal to the plane $(M_i \mid -M_i\mathbf{t}_i)^\top \mathbf{n}_i$. Consequently,

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

Applying this to (6) lead to

$$({\mathbf{n}_2}^\top M_2\mathbf{t}_2)({M_0}^\top \mathbf{n}_0 \times {M_1}^\top \mathbf{n}_1) = ({\mathbf{n}_1}^\top M_1\mathbf{t}_1)({M_0}^\top \mathbf{n}_0 \times {M_2}^\top \mathbf{n}_2) \tag{7}$$

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

Lemma 3.4. *If M is any 3×3 matrix, and $\mathbf{u}$ and $\mathbf{v}$ are column vectors, then*

$$(M\mathbf{u}) \times (M\mathbf{v}) = M^*(\mathbf{u} \times \mathbf{v}) \ . \tag{8}$$

Applying (8) to each of the two cross products in (7) leads to

$$M_0^{-1}(\mathbf{n}_2{}^\top M_2 \mathbf{t}_2)(\mathbf{n}_0 \times M_0^* M_1{}^\top \mathbf{n}_1) = M_0^{-1}(\mathbf{n}_1{}^\top M_1 \mathbf{t}_1)(\mathbf{n}_0 \times M_0^* M_2{}^\top \mathbf{n}_2) \ . \tag{9}$$

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

$$\mathbf{n}_0 \times \big((\mathbf{n}_2{}^\top M_2 \mathbf{t}_2) M_0^* M_1{}^\top \mathbf{n}_1 - (\mathbf{n}_1{}^\top M_1 \mathbf{t}_1) M_0^* M_2{}^\top \mathbf{n}_2\big) = 0 \ . \tag{10}$$

As in [12], we write

$$B = (\mathbf{n}_2{}^\top M_2 \mathbf{t}_2) M_0^* M_1{}^\top \mathbf{n}_1 - (\mathbf{n}_1{}^\top M_1 \mathbf{t}_1) M_0^* M_2{}^\top \mathbf{n}_2 \tag{11}$$

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

$$\begin{aligned} M_0^* M_1{}^\top &= \begin{pmatrix} \mathbf{r}_1{}^\top \\ \mathbf{r}_2{}^\top \\ \mathbf{r}_3{}^\top \end{pmatrix} \\ M_0^* M_2{}^\top &= \begin{pmatrix} \mathbf{s}_1{}^\top \\ \mathbf{s}_2{}^\top \\ \mathbf{s}_3{}^\top \end{pmatrix} \\ M_1 \mathbf{t}_1 &= \mathbf{t} \\ M_2 \mathbf{t}_2 &= \mathbf{u} \end{aligned} \tag{12}$$

vector B can be written in the form

$$B = \begin{pmatrix} \mathbf{n}_1{}^\top(\mathbf{r}_1 \mathbf{u}^\top - \mathbf{t}\mathbf{s}_1{}^\top)\mathbf{n}_2 \\ \mathbf{n}_1{}^\top(\mathbf{r}_2 \mathbf{u}^\top - \mathbf{t}\mathbf{s}_2{}^\top)\mathbf{n}_2 \\ \mathbf{n}_1{}^\top(\mathbf{r}_3 \mathbf{u}^\top - \mathbf{t}\mathbf{s}_3{}^\top)\mathbf{n}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{n}_1{}^\top E \mathbf{n}_2 \\ \mathbf{n}_1{}^\top F \mathbf{n}_2 \\ \mathbf{n}_1{}^\top G \mathbf{n}_2 \end{pmatrix} \ . \tag{13}$$

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

$$\mathbf{n}_0 \times \begin{pmatrix} \mathbf{n}_1{}^\top E \mathbf{n}_2 \\ \mathbf{n}_1{}^\top F \mathbf{n}_2 \\ \mathbf{n}_1{}^\top G \mathbf{n}_2 \end{pmatrix} = 0 \ . \tag{14}$$

This is essentially the same as equation (2.13) in [12], derived here, however, for the case of uncalibrated cameras. As remarked in [12], for each line ℓ, equation (14) 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 (12), we have

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

from which it follows, using Proposition 2.2 that Q_{01} is the essential matrix corresponding to the (ordered) pair of transformation matrices $(M_0 \mid 0)$ and $(M_1 \mid -M_1\mathbf{t}_1)$.

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. If 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 [12].

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.

4 Computation from 9 lines

If the lines are known to satisfy certain geometric constraints, then it is possible to compute the essential matrix using fewer lines in three views. The general idea is that if the projective geometry of some plane in the image can be fixed, then the determination of the epipolar geometry is simplified. This observation was applied to the determination of Q from point correspondences in [13]. Instead of considering the configuration of 9 lines of which four are coplanar, we consider four points in a plane and five lines not in the plane. From four lines in a plane it is easy to identify four points as the intersections of pairs of lines. Thus, let $\mathbf{x}_1, \ldots \mathbf{x}_4$ be four points lying in a plane π in $\mathcal{P}^3$. Let the images of these points as seen in three images be $\mathbf{u}_i$, $\mathbf{u}'_i$ and $\mathbf{u}''_i$. We suppose for convenience that the images have been subjected to appropriate projective transforms so that $\mathbf{u}_i = \mathbf{u}'_i = \mathbf{u}''_i$ for all i. Then, a necessary and sufficient condition for any further point $\mathbf{x}$ to lie in the plane π is that $\mathbf{x}$ projects to the same point in all three images.

This observation may be viewed in a different way. We may assume that the image planes of the three images are all identical with the plane π itself, since by an appropriate choice of projective coordinates in each of the image planes, it may be ensured that the projective mapping from plane π to each of the image planes is the identity coordinate map. The projective mapping associated with each camera maps a point $\mathbf{x}$ in space to the image point $\mathbf{u}$ in which the line through $\mathbf{x}$ and the camera centre pierces the image plane. Coordinates for $\mathcal{P}^3$may be chosen so that the plane π is the plane at infinity and the first camera is placed at the point $(0,0,0,1)^\top$. Let the other two cameras be placed at the points $\begin{pmatrix} \mathbf{a} \\ 1 \end{pmatrix}$ and $\begin{pmatrix} \mathbf{b} \\ 1 \end{pmatrix}$. The three camera transformation matrices are then $P = (I \mid 0)$, $P' = (I \mid -\mathbf{a})$ and $P'' = (I \mid -\mathbf{b})$. If we can compute the vectors $\mathbf{a}$ and $\mathbf{b}$, then the essential matrices can be computed using Theorem 2.2.

Now consider a line λ in $\mathcal{P}^3$which does not lie in the image plane. Let the projections of λ with respect to the three cameras be $\boldsymbol{\ell}$, $\boldsymbol{\ell}'$ and $\boldsymbol{\ell}''$. Since λ does not lie in the image plane, its three images will be distinct lines. However, lines $\boldsymbol{\ell}$, $\boldsymbol{\ell}'$ and $\boldsymbol{\ell}''$ must all meet at a common point, namely the point at which λ

meets the image plane.

Given $\boldsymbol{\ell}$, $\boldsymbol{\ell}'$ and $\boldsymbol{\ell}''$ the line λ may be retrieved as the intersection of the three planes defined by each line and its corresponding camera centre. Each such plane may be computed explicitly. In particular from (2) the three planes are equal to $P^\top \boldsymbol{\ell} = \begin{pmatrix} \boldsymbol{\ell} \\ 0 \end{pmatrix}$, $P'^\top \boldsymbol{\ell}' = \begin{pmatrix} \boldsymbol{\ell}' \\ \mathbf{a}^\top \boldsymbol{\ell}' \end{pmatrix}$, and $P''^\top \boldsymbol{\ell}'' = \begin{pmatrix} \boldsymbol{\ell}' \\ \mathbf{b}^\top \boldsymbol{\ell}'' \end{pmatrix}$. The fact that these three planes meet in a common line implies that the 4×3 matrix

$$A = \begin{pmatrix} \boldsymbol{\ell} & \boldsymbol{\ell}' & \boldsymbol{\ell}'' \\ 0 & \boldsymbol{\ell}'^\top \mathbf{a} & \boldsymbol{\ell}''^\top \mathbf{b} \end{pmatrix} \ .$$

has rank 2. Hence, there must be a linear dependency between the columns of A.

As remarked above, the lines $\boldsymbol{\ell}$, $\boldsymbol{\ell}'$ and $\boldsymbol{\ell}''$ are coincident, so there is a relationship $\alpha\boldsymbol{\ell} + \beta\boldsymbol{\ell}' + \gamma\boldsymbol{\ell}'' = 0$. This gives a linear dependency between the first three rows of A. Since $\boldsymbol{\ell}$, $\boldsymbol{\ell}'$ and $\boldsymbol{\ell}''$ are known, the weights α, β and γ may be computed explicitly. Since A has rank 2, this dependency must also apply to the last row as well which means that

$$\beta\boldsymbol{\ell}'^\top \mathbf{a} + \gamma\boldsymbol{\ell}''^\top \mathbf{b} = 0 \ .$$

This is a single linear equation in the coordinates of the two vectors $\mathbf{a}$ and $\mathbf{b}$. Given five such equations, arising from five lines not lying in the plane π, it is possible to solve for $\mathbf{a}$ and $\mathbf{b}$ up to an unknown (but insignificant) scale factor.

Summary of the algorithm The algorithm for determining the essential matrices from four coplanar points and five lines in three images is as follows. We start with coordinates $\mathbf{u}_i$, $\mathbf{u}'_i$ and $\mathbf{u}''_i$, the images of the points in the three images and also $\boldsymbol{\ell}$, $\boldsymbol{\ell}'$ and $\boldsymbol{\ell}''$, the images of the lines. The steps of the algorithm are as follows.

1. Determine two-dimensional projective transformations represented by non-singular 3×3 matrices K' and K'' such that for each $i = 1, \ldots 4$ we have $\mathbf{u}_i = K'\mathbf{u}'_i = K''\mathbf{u}''_i$.
2. Replace each line $\boldsymbol{\ell}'_i$ by the transformed line $K'^*\boldsymbol{\ell}'_i$, and each $\boldsymbol{\ell}''_i$ by $K''^*\boldsymbol{\ell}''_i$.
3. For each $i = 1, \ldots, 5$ find coefficients α_i, β_i and γ_i such that $\alpha_i\boldsymbol{\ell}_i + \beta_i\boldsymbol{\ell}'_i + \gamma_i\boldsymbol{\ell}''_i = 0$.
4. Solve the set of five linear equations $\beta_i\boldsymbol{\ell}_i'^\top \mathbf{a} + \gamma_i\boldsymbol{\ell}_i''^\top \mathbf{b} = 0$ to find the vectors $\mathbf{a}$ and $\mathbf{b}$, up to an indeterminate scale.
5. The three essential matrices are $K'^\top[\mathbf{a}]_\times$, $K''^\top[\mathbf{b}]_\times$ and $K''^\top[\mathbf{b} - \mathbf{a}]_\times K'$.

The above discussion was concerned with the case in which the plane π was defined by four points. Any other planar object which uniquely defines a projective basis for the plane may be used just as well, for example four coplanar lines (as already noted). This shows that four coplanar lines plus five lines not in the plane are sufficient (in 3 views) to determine the essential matrices.

5 Conclusion

The two algorithms given above can be used to determine the essential matrices for the purposes of invariant computation, scene reconstruction, image rectification or some other purpose.

Most interesting would be the case in which the three cameras are assumed to be the same. Then the cameras can be calibrated and the entire scene reconstructed up to scaled Euclidean transform from line correspondences in three views. In order to implement this method, an efficient implementation of the calibration algorithm of Faugeras and Maybank ([3, 4]) would be required. At the present time, no such implementation is available, so the calibration method described in this paper also remains unimplemented. This paper, therefore represents a contribution to the theory of calibration and scene reconstruction. It seems likely, however, that an efficient implementation of the algorithms of this paper and [3, 4] will become available in the future.

References

[1] H. A. Beyer, *"Accurate Calibration of CCD Cameras,"* Proceedings of IEEE Computer Vision and Pattern Recognition Conference, CVPR-92, pp. 96–101, 1992.

[2] 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.

[3] O. D. Faugeras, Q.-T Luong and S. J. Maybank, *"Camera Self-Calibration: Theory and Experiments,"* Proc. of ECCV-92, G. Sandini Ed., LNCS-Series Vol. 588, Springer- Verlag, 1992, pp. 321 – 334.

[4] O. D. Faugeras and S. J. Maybank, *"A theory of Self-Calibration of a Moving Camera"* International Journal of Computer Vision, 8:2, pp. 123-151, (1992).

[5] 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.

[6] R. Hartley, R. Gupta and Tom Chang, *"Stereo from Uncalibrated Cameras"* Proceedings of CVPR92.

[7] R. Hartley, *"Invariants of Lines in Space,"* submitted to this conference.

[8] R. Hartley, *"Invariants of Points Seen in Multiple Images,"*, preprint submitted for publication.

[9] R. Hartley, *"Computing Matched-epipolar Projections"* preprint submitted to this conference.

[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 (editors), "Geometric Invariance in Computer Vision,", MIT Press, Cambridge Ma, 1992.

[12] 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.

[13] A. Zisserman, R. Hartley, J. Mundy, P. Beardsley, "Projective Structure From Multiple Views", in preparation.

[14] P. Beardsley, D. Murray and A. Zisserman, *"Camera Calibration Using Multiple Images"*, Proc. of ECCV-92, G. Sandini Ed., LNCS-Series Vol. 588, Springer- Verlag, 1992, pp. 312 – 320.