PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Estimating tensors for matching over multiple views

The Royal Society

Phil. Trans. R. Soc. Lond. A 1998 **356**, 1267-1282
doi: 10.1098/rsta.1998.0221

The Royal Society

Estimating tensors for matching over multiple views

By Joan Lasenby[1] and Anthony N. Lasenby[2]

[1] *Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK*
[2] *Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK*

In this paper we give purely geometric derivations of the constraints between point and line correspondences in multiple views of a static field. The analysis is carried out using *geometric algebra*, a system which provides a useful tool in many computer vision applications. It is shown that with a straightforward geometric interpretation it is simple to derive the degrees of freedom of such tensors and to understand their structure. Given such information, minimal parametrizations of the tensors are possible. Such parametrizations may be useful for estimation of the tensors and subsequent matching of points.

Keywords: geometry; tensors; geometric algebra; trifocal tensor; projective geometry; minimization

1. Introduction

Estimating the fundamental matrix, F, has become an important part of many feature matching schemes. For example, in RANSAC schemes, Fs are estimated for a number of randomly selected groups of seven potential point matches. Each F is then applied to all points apart from those from which it was formed. The F which shows the best behaviour (i.e. gives a small value of $F_{ij}x_ix'_j$ for the largest number of potential matches) is taken as the estimated F and outliers are rejected on the basis of this F. Thus it is clear that we must have a means of estimating F in the presence of noisy data which is accurate and robust. Such estimation techniques have been the subject of much recent research (Luong & Faugeras 1995; Hartley 1995).

More recently, the *trifocal tensor*, T, which relates points and lines in three views, has been used for matching in a similar way as described above for F (Beardsley *et al*. 1996). T may have many advantages over F; it is claimed that matching over three views is more robust and being able to use points and lines simultaneously provides additional flexibility. It may also be the case that T has no critical surfaces (Sashua & Maybank 1996), or that the critical surfaces prevent less of a problem than they do for F. This issue of the critical surfaces of T is a current research area. Once again, for good matching, we need a means of estimating T which is accurate and robust in the presence of noise.

In the following sections we will derive the bilinear and trilinear constraints relating two and three views in purely geometric terms and from this redefine the tensors arising from the constraints as linear functions. Using such a framework we will analyse the characteristics of the linear functions and develop minimal parametrizations. The use of such minimal parametrizations in minimization routines will be discussed.

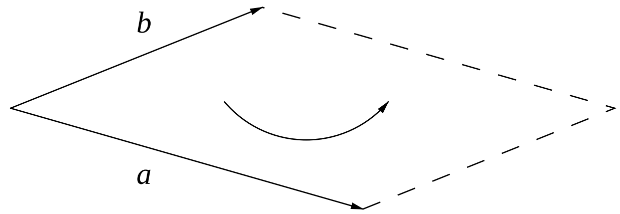

Figure 1. The directed area, or bivector, $\boldsymbol{a} \wedge \boldsymbol{b}$.

2. Geometric algebra: a brief outline

The algebras of Clifford (1878) and Grassmann (1877), are well known to pure mathematicians, but were abandoned by physicists in favour of the vector algebra of Gibbs, which is that used today in most areas of physics. The approach to Clifford algebra we adopt here was pioneered in the 1960s by Hestenes (1966), who has worked on developing his version of Clifford algebra, which will be referred to as *geometric algebra*, into a unifying language for mathematics and physics.

Let $\mathcal{G}_n$ denote the geometric algebra of n-dimensions; this is a vector space in which the elements have a property called grade (see later). As well as vector addition and scalar multiplication we have a non-commutative product which is associative and distributive over addition; this is the *geometric* or *Clifford* product. A further distinguishing feature of the algebra is that any vector squares to give a scalar. The geometric product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is written $\boldsymbol{ab}$ and can be expressed as a sum of its symmetric and antisymmetric parts which we write as

$$\boldsymbol{ab} = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \wedge \boldsymbol{b}. \tag{2.1}$$

We are therefore able to define the inner product $\boldsymbol{a} \cdot \boldsymbol{b}$ and the outer product $\boldsymbol{a} \wedge \boldsymbol{b}$ in terms of the more fundamental geometric product as follows:

$$\boldsymbol{a} \cdot \boldsymbol{b} = \tfrac{1}{2}(\boldsymbol{ab} + \boldsymbol{ba}), \qquad \boldsymbol{a} \wedge \boldsymbol{b} = \tfrac{1}{2}(\boldsymbol{ab} - \boldsymbol{ba}). \tag{2.2}$$

The inner product of two vectors is the standard *scalar* or *dot* product and produces a scalar. The outer or wedge product of two vectors is a new quantity we call a *bivector*. We think of a bivector as a directed area in the plane containing $\boldsymbol{a}$ and $\boldsymbol{b}$, formed by sweeping $\boldsymbol{a}$ along $\boldsymbol{b}$; see figure 1.

Thus $\boldsymbol{b} \wedge \boldsymbol{a}$ will have the opposite orientation making the wedge product anticommutative. The outer product is immediately generalizable to higher dimensions; for example, $(\boldsymbol{a} \wedge \boldsymbol{b}) \wedge \boldsymbol{c}$, a *trivector*, is interpreted as the oriented volume formed by sweeping the area $\boldsymbol{a} \wedge \boldsymbol{b}$ along vector $\boldsymbol{c}$. The outer product of k vectors is a k-vector, and has *grade* k. A general element of the geometric algebra of n-dimensions is a *multivector*, which is a linear combination of objects of any grade. If a multivector possesses only terms of a single grade it is termed *homogeneous*. The geometric algebra provides a means of manipulating multivectors which allows us to keep track of different grade objects simultaneously. We are already familiar with such a process in dealing with complex numbers: there one has two different types of object (real and imaginary), but the algebra is such that we can manipulate the complex number in a way which gives us the correct behaviour in the real and imaginary domains. In a space of three dimensions we can construct a trivector $\boldsymbol{a} \wedge \boldsymbol{b} \wedge \boldsymbol{c}$, but no 4-vectors exist since there is no possibility of sweeping the volume element $\boldsymbol{a} \wedge \boldsymbol{b} \wedge \boldsymbol{c}$ over a fourth dimension. The highest grade element in a space is called the *pseudoscalar*. The unit pseudoscalar is denoted by I or i in two and three dimensions.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

We can generalize the definitions of inner and outer products given in equation (2.2). For two homogeneous multivectors A_r and B_s (i.e. multivectors of grades r and s respectively), we define the inner and outer products as

$$A_r \cdot B_s = \langle A_r B_s \rangle_{|r-s|}, \tag{2.3}$$

$$A_r \wedge B_s = \langle A_r B_s \rangle_{r+s}, \tag{2.4}$$

where $\langle M \rangle_t$ denotes the t-grade part of the multivector M. Thus the inner product produces an $|r-s|$-vector, which means it effectively reduces the grade of B_s by r; and the outer product gives an $(r+s)$-vector, therefore increasing the grade of B_s by r. This is an extension of the general principle that dotting with a vector lowers the grade of a multivector by 1 and wedging with a vector raises the grade of a multivector by 1. To avoid the inclusion of too many brackets, the convention we shall use here is that the inner and outer products take precedence over the geometric product in any expression.

Another concept which will be used elsewhere in this paper is that of the *reciprocal frame*. Given a set of linearly independent vectors $\{\boldsymbol{e}_j\}$ (where no assumption of orthonormality is made), we can form a *reciprocal frame*, $\{\boldsymbol{e}^j\}$, which is such that

$$\boldsymbol{e}^j \cdot \boldsymbol{e}_k = \delta_{jk}. \tag{2.5}$$

For details of the explicit construction of such a reciprocal frame in n-dimensions see Hestenes & Sobczyk (1984). In three dimensions this is a very simple operation and the reciprocal frame vectors for a linearly independent set of vectors $\{\boldsymbol{e}_j\}$, $j = 1, \ldots, 3$, are as follows

$$\boldsymbol{e}^1 = \frac{1}{\alpha} i \boldsymbol{e}_2 \wedge \boldsymbol{e}_3, \quad \boldsymbol{e}^2 = \frac{1}{\alpha} i \boldsymbol{e}_3 \wedge \boldsymbol{e}_1, \quad \boldsymbol{e}^3 = \frac{1}{\alpha} i \boldsymbol{e}_1 \wedge \boldsymbol{e}_2, \tag{2.6}$$

where $i\alpha = \boldsymbol{e}_3 \wedge \boldsymbol{e}_2 \wedge \boldsymbol{e}_1$. Note that, according to our defined order of precedence for operators, the above expressions mean $(1/\alpha)i(\boldsymbol{e}_2 \wedge \boldsymbol{e}_3)$, etc. We can express any vector $\boldsymbol{a}$ in terms of the frame or the reciprocal frame:

$$\boldsymbol{a} = a^j \boldsymbol{e}_j = a_k \boldsymbol{e}^k. \tag{2.7}$$

Here and in the following sections the summation convention will be used unless otherwise stated, i.e. repeated indices are summed over.

3. Projective space and the projective split

Points in real three-dimensional (3D) space will be represented by vectors in $\mathcal{E}^3$, a 3D space with a Euclidean metric. We take as our model a pinhole camera so that any point in space is projected down onto an image plane which is at a distance f (the focal length) from the centre of projection O. Since any point on a line through O will be mapped to a single point in the image plane, we can see why the approach of associating a point in $\mathcal{E}^3$ with a line in a 4D space, R^4, might be a reasonable thing to do. Suppose we define basis vectors: $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ in R^4 and $(\sigma_1, \sigma_2, \sigma_3)$ in $\mathcal{E}^3$ and use the geometric algebras of four and three dimensions on these spaces. We require that vectors, bivectors and trivectors in R^4 will represent points, lines and planes in $\mathcal{E}^3$. Choosing γ_4 as a selected direction in R^4, we can then define a mapping which associates the bivectors $\gamma_i \gamma_4$, $i = 1, 2, 3$, in R^4 with the vectors σ_i, $i = 1, 2, 3$, in $\mathcal{E}^3$. This process of association is called the *projective split*. To ensure $\sigma_i^2 = +1$ we are forced to assume a non-Euclidean metric for the basis vectors in R^4. We choose to use $\gamma_4^2 = +1$, $\gamma_i^2 = -1$, $i = 1, 2, 3$.

For a vector $\boldsymbol{X} = X_1\gamma_1 + X_2\gamma_2 + X_3\gamma_3 + X_4\gamma_4$ in R^4 the projective split is obtained by taking the geometric product of $\boldsymbol{X}$ and γ_4. This leads to the association of the vector $\boldsymbol{x}$ in $\mathcal{E}^3$ with the bivector $\boldsymbol{X} \wedge \gamma_4 / X_4$ in R^4 so that

$$\boldsymbol{x} = \frac{X_1}{X_4}\gamma_1\gamma_4 + \frac{X_2}{X_4}\gamma_2\gamma_4 + \frac{X_3}{X_4}\gamma_3\gamma_4 = \frac{X_1}{X_4}\sigma_1 + \frac{X_2}{X_4}\sigma_2 + \frac{X_3}{X_4}\sigma_3, \tag{3.1}$$

which implies $x_i = X_i / X_4$, for $i = 1, 2, 3$. The process of representing $\boldsymbol{x}$ in a higher dimensional space can therefore be seen to be equivalent to using a vector of *homogeneous coordinates*, $\boldsymbol{X}$, for $\boldsymbol{x}$.

(*a*) *Projective geometry and algebra in projective space*

We now look at the basic projective geometry operations of *meet* and *join*, and briefly discuss the algebra of incidence in projective space. For more detail the reader is referred to Hestenes & Ziegler (1991) and Bayro *et al*. (1996).

Any pseudoscalar P can be written as $P = \alpha I$ where α is a scalar and we introduce the notation

$$PI^{-1} = \alpha I I^{-1} = \alpha \equiv [P]. \tag{3.2}$$

This *bracket* is precisely the bracket of the Grassmann–Cayley algebra. We then define the dual, A^*, of an r-vector A as

$$A^* = AI^{-1}. \tag{3.3}$$

The *join* $J = A \bigwedge B$ of an r-vector A and an s-vector B is defined by

$$J = A \wedge B \quad \text{if } A \text{ and } B \text{ are linearly independent}, \tag{3.4}$$

and it can be shown that the *meet* of A and B can be written as

$$A \vee B = (A^* \wedge B^*)I = (A^* \wedge B^*)(I^{-1}I)I = (A^* \cdot B). \tag{3.5}$$

The join and meet can be used to describe lines and planes and to intersect these quantities. Consider three non-collinear points, P_1, P_2, P_3, represented by vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ in $\mathcal{E}^3$ and by vectors $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$ in R^4. The line L_{12} joining points P_1 and P_2, and the plane Φ_{123} passing through points P_1, P_2, P_3, can be expressed in R^4 by the following bivector and trivector respectively:

$$L_{12} = \boldsymbol{X}_1 \wedge \boldsymbol{X}_2, \qquad \Phi_{123} = \boldsymbol{X}_1 \wedge \boldsymbol{X}_2 \wedge \boldsymbol{X}_3. \tag{3.6}$$

In $\mathcal{E}^3$, the intersection of lines and planes, of planes and planes, and of lines and lines, can be dealt with entirely using the meet operation. Details and derivations are given in Lasenby & Bayro (1998).

Any point P, represented in R^4 by $\boldsymbol{X}$, on the line through P_1 and P_2, will satisfy

$$\boldsymbol{X} \wedge L_{12} = \boldsymbol{X} \wedge \boldsymbol{X}_1 \wedge \boldsymbol{X}_2 = 0. \tag{3.7}$$

This is therefore the equation of the line in R^4. In general such an equation is telling us that $\boldsymbol{X}$ belongs to the subspace spanned by $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. Similarly, $\boldsymbol{X}$ lies in the plane Φ_{123} if

$$\boldsymbol{X} \wedge \Phi_{123} = \boldsymbol{X} \wedge \boldsymbol{X}_1 \wedge \boldsymbol{X}_2 \wedge \boldsymbol{X}_3 = 0. \tag{3.8}$$

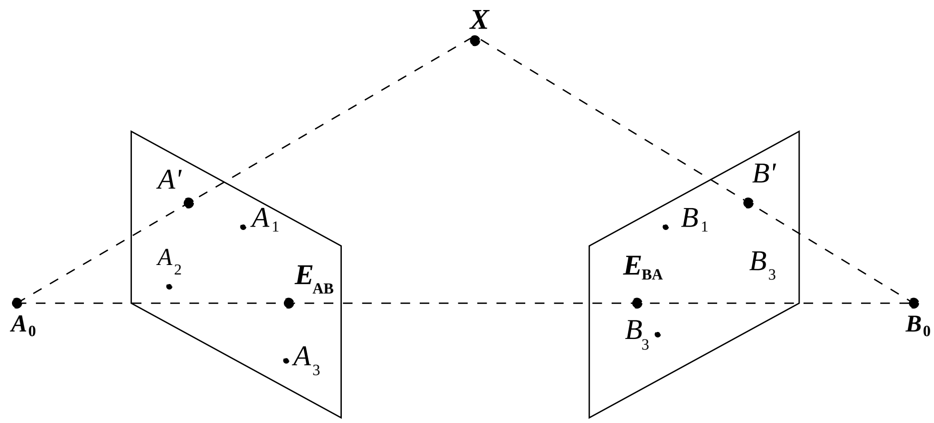

Figure 2. The projections of a world point $\boldsymbol{X}$ onto two image planes A and B are shown together with the epipoles, $\boldsymbol{E}_{AB}$ and $\boldsymbol{E}_{BA}$, and three points which define each plane. All vectors are in R^4.

4. The fundamental matrix

In this section we will look at how the fundamental matrix, F, can be regarded as a linear function taking two vectors to a scalar and how this enables us to extract the structure of F. In the following section we will then see how to apply exactly the same techniques to the trifocal tensor relating three views. As before, we will use lower case letters to denote vectors in $\mathcal{E}^3$ and upper case letters to denote vectors in R^4. Let us first define two image planes by specifying three points in each plane: $(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$ and $(\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3)$, with the R^4 representations of these points being $\boldsymbol{A}_i, \boldsymbol{B}_i$, $i = 1, 2, 3$. We will work mainly in projective space, therefore dealing with the R^4 vectors. As shown in figure 2, a world point $\boldsymbol{X}$ projects onto points $\boldsymbol{A}'$ and $\boldsymbol{B}'$ in the two image planes. In figure 2 the epipoles (the intersections of the line joining the optical centres with the image planes) in the A and B planes are denoted by $\boldsymbol{E}_{AB}$ and $\boldsymbol{E}_{BA}$. It is clear that the points $\boldsymbol{A}_0, \boldsymbol{B}_0, \boldsymbol{A}', \boldsymbol{B}'$ are coplanar. The wedge of these four vectors must therefore vanish:

$$\boldsymbol{A}_0 \wedge \boldsymbol{B}_0 \wedge \boldsymbol{A}' \wedge \boldsymbol{B}' = 0. \tag{4.1}$$

Now, if we let $\boldsymbol{A}' = \alpha_i \boldsymbol{A}_i$ and $\boldsymbol{B}' = \beta_j \boldsymbol{B}_j$, then equation (4.1) can be written as

$$\alpha_i \beta_j \{\boldsymbol{A}_0 \wedge \boldsymbol{B}_0 \wedge \boldsymbol{A}_i \wedge \boldsymbol{B}_j\} = 0. \tag{4.2}$$

Defining $F_{ij} = \{\boldsymbol{A}_0 \wedge \boldsymbol{B}_0 \wedge \boldsymbol{A}_i \wedge \boldsymbol{B}_j\} I^{-1}$ gives us

$$F_{ij} \alpha_i \beta_j = 0, \tag{4.3}$$

which is the well-known relationship between the components of the fundamental matrix, F, and the image coordinates. This therefore suggests that we might cast F as a linear function mapping two vectors onto a scalar:

$$F(\boldsymbol{A}, \boldsymbol{B}) = \{\boldsymbol{A}_0 \wedge \boldsymbol{B}_0 \wedge \boldsymbol{A} \wedge \boldsymbol{B}\} I^{-1}, \tag{4.4}$$

so that $F_{ij} = F(\boldsymbol{A}_i, \boldsymbol{B}_j)$. Now, in order to investigate the number of degrees of freedom of F we can consider what F looks like under a change of basis. Let us use a new basis in each image plane; instead of $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$, we have $\boldsymbol{E}\boldsymbol{A}_i$ and $\boldsymbol{E}\boldsymbol{B}_i$, where

$$\boldsymbol{E}\boldsymbol{A}_1 = \boldsymbol{E}_{AB}, \qquad \boldsymbol{E}\boldsymbol{A}_2 = \boldsymbol{A}_2, \qquad \boldsymbol{E}\boldsymbol{A}_3 = \boldsymbol{A}_3, \tag{4.5}$$

$$\boldsymbol{E}\boldsymbol{B}_1 = \boldsymbol{E}_{BA}, \qquad \boldsymbol{E}\boldsymbol{B}_2 = \boldsymbol{B}_2, \qquad \boldsymbol{E}\boldsymbol{B}_3 = \boldsymbol{B}_3, \tag{4.6}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

provided $\boldsymbol{E}_{AB}, \boldsymbol{A}_2, \boldsymbol{A}_3$ and $\boldsymbol{E}_{BA}, \boldsymbol{B}_2, \boldsymbol{B}_3$ are not collinear. Now, because $\boldsymbol{E}\boldsymbol{A}_1 \wedge \boldsymbol{A}_0 \wedge \boldsymbol{B}_0 = 0$ since the points are collinear, and similarly $\boldsymbol{E}\boldsymbol{B}_1 \wedge \boldsymbol{A}_0 \wedge \boldsymbol{B}_0 = 0$, we see that

$$F(\boldsymbol{E}\boldsymbol{A}_1, \boldsymbol{B}) = 0 \quad \forall \boldsymbol{B}, \tag{4.7}$$

$$F(\boldsymbol{A}, \boldsymbol{E}\boldsymbol{B}_1) = 0 \quad \forall \boldsymbol{A}, \tag{4.8}$$

which imply that $F_{11} = F_{12} = F_{13} = 0$ and $F_{11} = F_{21} = F_{31} = 0$. In this *epipole basis*, F therefore has $(9-5) = 4$ non-zero entries. Thus the total number of degrees of freedom is $4 + (2 \times 2) - 1 = 7$, where the 2×2 comes from the unknown epipole coordinates (up to a scale factor) and the -1 is because F can only be determined up to an overall scale. Thus a possible minimal parametrization of F would be the vector $\boldsymbol{X}$:

$$\boldsymbol{X} = [F_{22}, F_{23}, F_{32}, \boldsymbol{E}_{ab}, \boldsymbol{E}_{ba}], \tag{4.9}$$

or

$$\boldsymbol{X} = [F_{22}, F_{23}, F_{32}, \boldsymbol{e}_{ab}, \boldsymbol{e}_{ba}], \tag{4.10}$$

where we have assumed that $F_{33} = 1$. Now, in order to use this minimal parametrization in any useful way, we will need to express the function as acting on 3D vectors. It can be shown that

$$F(\boldsymbol{X}_1, \boldsymbol{X}_2) = \{\boldsymbol{A}_0 \wedge \boldsymbol{B}_0 \wedge \boldsymbol{X}_1 \wedge \boldsymbol{X}_2\} I_4^{-1} = (\boldsymbol{a}_0 - \boldsymbol{b}_0) \wedge (\boldsymbol{x}_1 - \boldsymbol{a}_0) \wedge (\boldsymbol{x}_2 - \boldsymbol{b}_0) I_3^{-1}, \tag{4.11}$$

where I_3 and I_4 are respectively the pseudoscalars in the 3D and 4D spaces. We see from this that we can write $F(\boldsymbol{X}_1, \boldsymbol{X}_2)$ as $F(\boldsymbol{x}_1, \boldsymbol{x}_2)$ which satisfies

$$F(\boldsymbol{e}\boldsymbol{a}_1, \boldsymbol{b}) = 0 \quad \forall \boldsymbol{b}, \tag{4.12}$$

$$F(\boldsymbol{a}, \boldsymbol{e}\boldsymbol{b}_1) = 0 \quad \forall \boldsymbol{a}. \tag{4.13}$$

We must therefore also have $F(\boldsymbol{x}, \boldsymbol{x}') = 0$ if $\boldsymbol{x}$ and $\boldsymbol{x}'$ are projections of the same world point. Given N matching point sets we might try to minimize the following quantity:

$$S = \sum_{p=1}^{N} \frac{[F(\boldsymbol{x}_p, \boldsymbol{x}'_p)]^2}{{\sigma_p}^2}, \tag{4.14}$$

where the ${\sigma_p}^2$ are weighting factors to be determined through statistical considerations. We can use the reciprocal frame vectors (equation (2.7)) to write the quantities we are summing in terms of our observations and our minimal parameter set:

$$\begin{aligned} F(\boldsymbol{x}, \boldsymbol{x}') &= F(x^j \boldsymbol{e}_j, x'^k \boldsymbol{e}_k) \\ &= (\boldsymbol{x} \cdot \boldsymbol{e}^j)(\boldsymbol{x}' \cdot \boldsymbol{e}^k) F_{jk}. \end{aligned} \tag{4.15}$$

Here we have used the fact that $\boldsymbol{e}^i \cdot \boldsymbol{e}_j = \delta^i_j$ to express the coefficients of the points in the $\{\boldsymbol{e}\}$ basis in terms of the vectors and the reciprocal frame vectors.

It may be shown that this procedure is effectively equivalent to some methods for determining F which have appeared in the literature in recent years (Luong & Faugeras 1995). But, we can now apply almost exactly the same techniques to determining T. This is discussed in the following sections.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

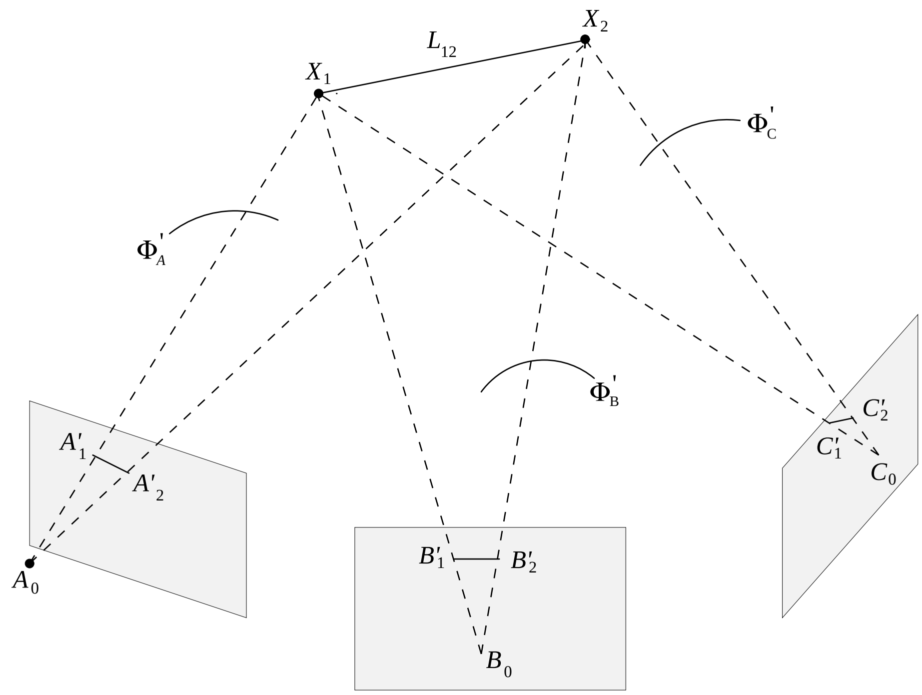

Figure 3. The projections of two world points, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, in each of three image planes and the lines joining these projected points are shown.

5. The trifocal/trilinear tensor

We can derive the constraints between points and lines in three views via the same geometric approach that was used for two views. Figure 3 shows three image planes A, B and C, with optical centres $\boldsymbol{A}_0, \boldsymbol{B}_0, \boldsymbol{C}_0$, and the intersections of two world points, $\boldsymbol{X}_1, \boldsymbol{X}_2$, with these planes. $\boldsymbol{X}_i$ intersects the planes at points $\boldsymbol{A}'_i, \boldsymbol{B}'_i, \boldsymbol{C}'_i$, $i = 1, 2$, and we denote the lines in the image planes joining the two intersection points by $L_A = \boldsymbol{A}'_1 \wedge \boldsymbol{A}'_2$, $L_B = \boldsymbol{B}'_1 \wedge \boldsymbol{B}'_2$ and $L_C = \boldsymbol{C}'_1 \wedge \boldsymbol{C}'_2$. The line joining the world points is $L_{12} = \boldsymbol{X}_1 \wedge \boldsymbol{X}_2$. Let us firstly define three planes:

$$\Phi'_A = \boldsymbol{A}_0 \wedge \boldsymbol{A}'_1 \wedge \boldsymbol{A}'_2, \quad \Phi'_B = \boldsymbol{B}_0 \wedge \boldsymbol{B}'_1 \wedge \boldsymbol{B}'_2, \quad \Phi'_C = \boldsymbol{C}_0 \wedge \boldsymbol{C}'_1 \wedge \boldsymbol{C}'_2. \tag{5.1}$$

It is clear that L_{12} can be formed by intersecting Φ'_B and Φ'_C,

$$L_{12} = \Phi'_B \vee \Phi'_C = (\boldsymbol{B}_0 \wedge L_B) \vee (\boldsymbol{C}_0 \wedge L_C). \tag{5.2}$$

Letting $L_1 = \boldsymbol{A}_0 \wedge \boldsymbol{A}'_1$ and $L_2 = \boldsymbol{A}_0 \wedge \boldsymbol{A}'_2$, we can easily see that L_1 and L_2 intersect with L_{12} and $\boldsymbol{X}_1$ at $\boldsymbol{X}_2$ respectively. We therefore have

$$L_1 \wedge L_{12} = 0 \quad \text{and} \quad L_2 \wedge L_{12} = 0. \tag{5.3}$$

$L_i \wedge L_{12} = 0$ can then be written as

$$(\boldsymbol{A}_0 \wedge \boldsymbol{A}'_i) \wedge \{(\boldsymbol{B}_0 \wedge L_B) \vee (\boldsymbol{C}_0 \wedge L_C)\} = 0. \tag{5.4}$$

This therefore suggests that we define a linear function T which maps a point and two lines onto a scalar:

$$T(\boldsymbol{A}, L_B, L_C) = (\boldsymbol{A}_0 \wedge \boldsymbol{A}) \wedge \{(\boldsymbol{B}_0 \wedge L_B) \vee (\boldsymbol{C}_0 \wedge L_C)\}. \tag{5.5}$$

Now, if we define bases $\{\boldsymbol{A}_i\}$, $\{\boldsymbol{B}_i\}$ and $\{\boldsymbol{C}_i\}$ for each plane, then it is possible to define a line basis for the B and C planes as follows:

$$LB_1 = \boldsymbol{B}_2 \wedge \boldsymbol{B}_3, \quad LB_2 = \boldsymbol{B}_3 \wedge \boldsymbol{B}_1, \quad LB_3 = \boldsymbol{B}_1 \wedge \boldsymbol{B}_2, \tag{5.6}$$

and similarly for LC_i and LA_i. Thus we can write

$$\boldsymbol{A} = \alpha_i \boldsymbol{A}_i, \quad L_B = l_j^B LB_j, \quad L_C = l_k^C LC_k. \tag{5.7}$$

Defining the components of a tensor as $T_{ijk} = T(\boldsymbol{A}_i, LB_j, LC_k)$, then if $\boldsymbol{A}, L_B, L_C$ are all derived from projections of the same two world points, equation (5.4) tells us that we can write

$$T_{ijk}\alpha_i l_j^B l_k^C = 0. \tag{5.8}$$

This is the constraint arrived at in Hartley (1994), where it was produced via a consideration of camera matrices. Equation (5.8) is simply saying that two planes intersect in a line which intersects with another line.

It is easy to see how the relationship between the line coordinates in the three image planes comes about using this framework. We can express the line in image plane A joining $\boldsymbol{A}_1'$ and $\boldsymbol{A}_2'$ as the intersection of the plane joining As optical centre to the world line L_{12} with the image plane $\Phi_A = \boldsymbol{A}_1 \wedge \boldsymbol{A}_2 \wedge \boldsymbol{A}_3$, e.g.

$$L_A = \boldsymbol{A}_1' \wedge \boldsymbol{A}_2' = (\boldsymbol{A}_0 \wedge L_{12}) \vee \Phi_A. \tag{5.9}$$

Expressing L_{12} as the meet of two planes $\Phi_B' \vee \Phi_C'$ and using the expansions of L_A, L_B, L_C given in equation (5.7), we can rewrite this equation as

$$l_i^A LA_i = l_j^B l_k^C \{[\boldsymbol{A}_0 \wedge \{(\boldsymbol{B}_0 \wedge LB_j) \vee (\boldsymbol{C}_0 \wedge LC_k)\}] \vee \Phi_A\}. \tag{5.10}$$

We can then expand parts of this equation using the projective geometry relations to give

$$l_i^A LA_i = l_j^B l_k^C \{[(\boldsymbol{A}_0 \wedge \boldsymbol{A}_n) \wedge \{(\boldsymbol{B}_0 \wedge LB_j) \vee (\boldsymbol{C}_0 \wedge LC_k)\}] LA_n, \tag{5.11}$$

which, when we equate coefficients, gives

$$l_i^A = T_{ijk} l_j^B l_k^C \tag{5.12}$$

for lines in the three images produced by the same world line.

Now, since we know the geometry of the linear function, T, we can begin to investigate its structure as we did with F. We firstly look to characterize T in terms of the epipoles in the problem. It is clear that in the three camera case, each image plane has two epipoles, in plane A we will call the epipoles $\boldsymbol{E}_{AB}$ and $\boldsymbol{E}_{AC}$, and similarly for the B and C planes as illustrated in figure 4. In each plane we can now define an '*epipole basis*', $\{\boldsymbol{EA}_i\}$, $\{\boldsymbol{EB}_i\}$ and $\{\boldsymbol{EC}_i\}$ for $i = 1, 2, 3$, where

$$\left.\begin{aligned} \boldsymbol{EA}_1 &= \boldsymbol{E}_{AB}, & \boldsymbol{EA}_2 &= \boldsymbol{E}_{AC}, \\ \boldsymbol{EB}_1 &= \boldsymbol{E}_{BA}, & \boldsymbol{EB}_2 &= \boldsymbol{E}_{BC}, \\ \boldsymbol{EC}_1 &= \boldsymbol{E}_{CA}, & \boldsymbol{EC}_2 &= \boldsymbol{E}_{CB}, \end{aligned}\right\} \tag{5.13}$$

and $\boldsymbol{EA}_3$ is any other point in the A plane which does not lie on the line joining the epipoles, and similarly for the other planes. From these bases it is then possible to form a set of *basis bivectors* in each plane, LA_i, LB_i, LC_i, given by

$$LA_1 = \boldsymbol{EA}_2 \wedge \boldsymbol{EA}_3, \quad LA_2 = \boldsymbol{EA}_3 \wedge \boldsymbol{EA}_1, \quad LA_3 = \boldsymbol{EA}_1 \wedge \boldsymbol{EA}_2, \tag{5.14}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

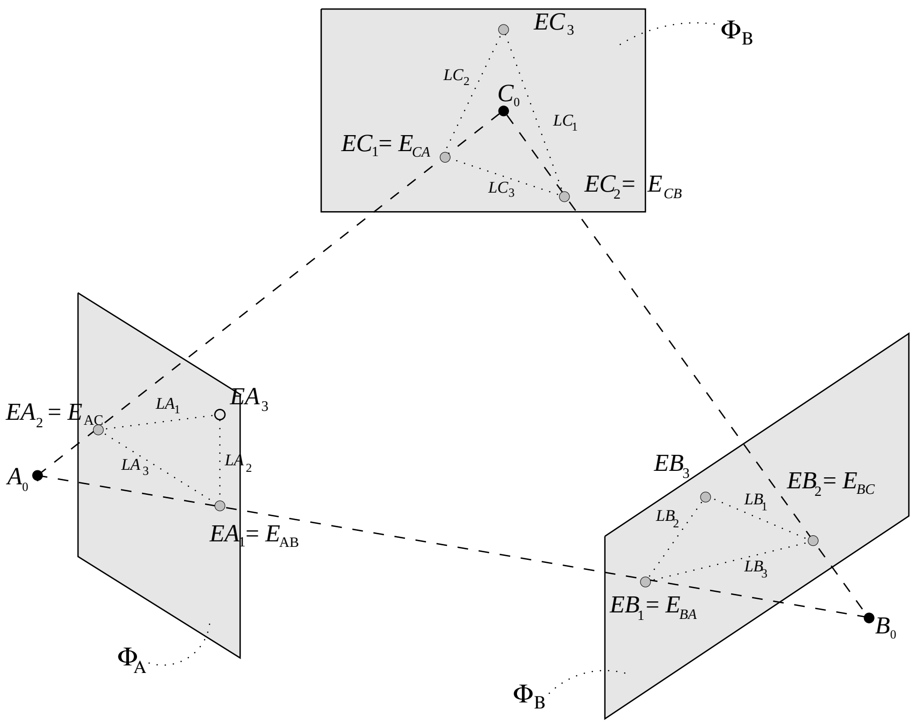

Figure 4. The plane containing the epipoles of three image planes is shown. In each of the image planes a basis is indicated (which includes the epipoles) and from this basis bivectors are constructed.

and similarly for the B and C-plane basis bivectors. If we now define the elements of a third rank tensor T_{ijk} by

$$T_{ijk} = (\boldsymbol{A}_0 \wedge \boldsymbol{EA}_i) \wedge \{(\boldsymbol{B}_0 \wedge LB_j) \vee (\boldsymbol{C}_0 \wedge LC_k)\}, \tag{5.15}$$

it will be straightforward to count the number of degrees of freedom of T. This is best done in several stages:

1. We note that if $\boldsymbol{A} = \boldsymbol{EA}_1$, or $\boldsymbol{EA}_2$ and $L_B = LB_1$, or LB_3 and $L_C = LC_1$, or LC_3 in $T(\boldsymbol{A}, L_B, L_C)$, then we obtain $T_{ijk} = 0$. We can see this from figure 4 since in this case the quantity $\{(\boldsymbol{B}_0 \wedge LB_j) \vee (\boldsymbol{C}_0 \wedge LC_k)\}$ will always be the line $\boldsymbol{B}_0 \wedge \boldsymbol{C}_0$ and so the resultant expression will then contain three vectors that are collinear and will therefore be zero, i.e. $\boldsymbol{A}_0 \wedge \boldsymbol{EA}_{(1 \text{ or } 2)} \wedge \boldsymbol{B}_0 \wedge \boldsymbol{C}_0$. This implies

$$T_{111} = T_{113} = T_{131} = T_{133} = T_{211} = T_{213} = T_{231} = T_{233} = 0. \tag{5.16}$$

   Thus we know that at least eight components are zero in this basis.

2. Now consider putting $\boldsymbol{A} = \boldsymbol{EA}_1$ and $L_B = LB_2$ or LB_3 with any LC_k. Since it is possible to rearrange the right-hand side of equation (5.15) as

$$T_{ijk} = (\boldsymbol{C}_0 \wedge LC_k) \wedge \{(\boldsymbol{A}_0 \wedge \boldsymbol{EA}_i) \vee (\boldsymbol{B}_0 \wedge LB_j)\}, \tag{5.17}$$

   we see that the substitutions cause the meet in the above equation to be zero (the line and the plane meet in a line rather than a point). This therefore gives

$$T_{121} = T_{122} = T_{123} = T_{131} = T_{132} = T_{133} = 0 \tag{5.18}$$

giving another four zero entries.

3. We can also rearrange equation (5.15) as

$$T_{ijk} = (\boldsymbol{B}_0 \wedge LB_j) \wedge \{(\boldsymbol{A}_0 \wedge \boldsymbol{E}\boldsymbol{A}_i) \vee (\boldsymbol{C}_0 \wedge LC_k)\}, \tag{5.19}$$

which tells us in the same way that if we put $\boldsymbol{A} = \boldsymbol{E}\boldsymbol{A}_2$ and $L_C = LC_2$ or LC_3 with any LB_k, then the T_{ijk} so formed is zero. We can then deduce that

$$T_{212} = T_{222} = T_{232} = T_{213} = T_{223} = T_{233} = 0. \tag{5.20}$$

A further four entries are therefore zero.

4. Putting $\boldsymbol{A} = \boldsymbol{E}\boldsymbol{A}_3$ and $L_B = LB_2$ or LB_3 **and** $L_C = LC_2$ or LC_3 in $T(\boldsymbol{A}, L_B, L_C)$, will also give zero. This can be seen from equations (5.17),(5.19) in which the meet will then be the point $\boldsymbol{A}_0$, thus causing the whole expression to contain three collinear vectors which makes it zero. Thus

$$T_{322} = T_{323} = T_{332} = T_{333} = 0, \tag{5.21}$$

providing a further four zero entries.

We therefore have $(8+4+4+4) = 20$ zero entries, giving seven non-zero entries of T in this epipole basis. The total number of degrees of freedom is therefore given by

$$7 + (6 \times 2) - 1 = 18. \tag{5.22}$$

The (6×2) comes from the coordinates (up to scale) of the six unknown epipoles and the (-1) is because T is only defined up to some overall scale factor.

Thus we have an explicit minimal parametrization of T given by the vector

$$\boldsymbol{X_T} = [T_{112}, T_{221}, T_{311}, T_{312}, T_{313}, T_{321}, \boldsymbol{E}_{ab}, \boldsymbol{E}_{ac}, \boldsymbol{E}_{ba}, \boldsymbol{E}_{bc}, \boldsymbol{E}_{ca}, \boldsymbol{E}_{cb}], \tag{5.23}$$

or

$$\boldsymbol{X_T} = [T_{112}, T_{221}, T_{311}, T_{312}, T_{313}, T_{321}, \boldsymbol{e}_{ab}, \boldsymbol{e}_{ac}, \boldsymbol{e}_{ba}, \boldsymbol{e}_{bc}, \boldsymbol{e}_{ca}, \boldsymbol{e}_{cb}], \tag{5.24}$$

where we have taken $T_{331} = 1$. The transition from four to three dimensions is brought about in the same way as we did for F. Thus, one possible means of estimating T given a number of point/line matches would be to minimize the following expression

$$SS = \sum_{p=1}^{N} \frac{[T(\boldsymbol{x}_p, \boldsymbol{l}_p^b, \boldsymbol{l}_p^c)]^2}{{\sigma_p}^2}, \tag{5.25}$$

where, again, the σ_ps are weighting factors to be determined by statistical considerations. We can proceed in precisely the same way as was outlined for F, we can write $T(\ldots)$ in terms of the parameters in $\boldsymbol{X_T}$ via use of the reciprocal frames in each image plane.

In order to minimize the number of iterations that are necessary it would be useful to have reasonable estimates of the epipoles; the next section will be concerned with extracting such estimates.

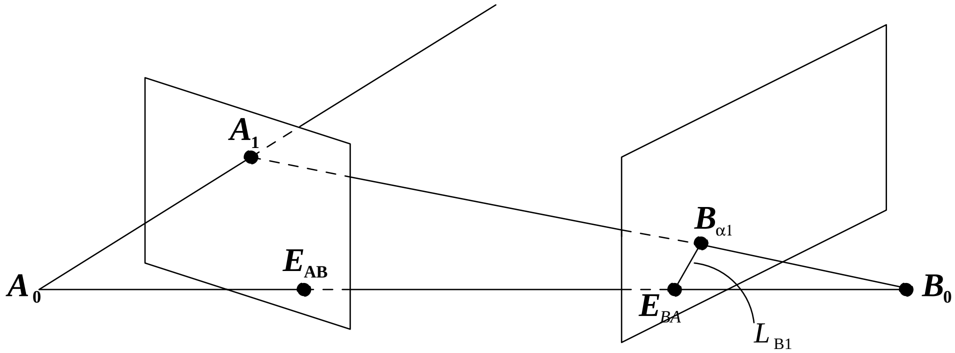

Figure 5. Sketch showing the epioles $\boldsymbol{E}_{AB}$ and $\boldsymbol{E}_{BA}$ and the intersection of the line joining $\boldsymbol{B}_0 \wedge \boldsymbol{A}_1$ with the B-plane.

6. Extracting the epipoles from T

In this section we will look first at the problem of extracting the epipoles from a given trifocal tensor. For some epipoles there are recent indications in the literature as to how to extract them (Hartley 1997; Shashua 1997). Here we will show how to obtain all six epipoles. In the case of a noiseless T the extraction method is, of course, exact, while for noisy Ts the extractions should be done in the most robust way possible.

(a) *Epipoles* $\boldsymbol{E}_{BA}$ *and* $\boldsymbol{E}_{CA}$

The first stage in finding epipoles $\boldsymbol{E}_{BA}$ and $\boldsymbol{E}_{CA}$ is to take the third rank tensor T and form from it three 3×3 matrices, $T1, T2, T3$ defined by

$$T1_{jk} = T_{1jk}, \quad T2_{jk} = T_{2jk}, \quad T3_{jk} = T_{3jk}. \tag{6.1}$$

The corresponding linear functions are given by $T1 = T(\boldsymbol{A}_1, L_B, L_C)$, $T2 = T(\boldsymbol{A}_2, L_B, L_C)$ and $T3 = T(\boldsymbol{A}_3, L_B, L_C)$. Since we know that

$$T(\boldsymbol{A}_1, L_B, L_C) = (\boldsymbol{C}_0 \wedge L_C) \wedge \{(\boldsymbol{A}_0 \wedge \boldsymbol{A}_1) \vee (\boldsymbol{B}_0 \wedge L_B)\}, \tag{6.2}$$

we can see that if $L_B \equiv L_{B1} = \boldsymbol{E}_{BA} \wedge \boldsymbol{B}_{\alpha_1}$, where $\boldsymbol{B}_{\alpha_1}$ is the point at which the line joining $\boldsymbol{B}_0$ and $\boldsymbol{A}_1$ intersects the B plane (see figure 5), then $T(\boldsymbol{A}_1, L_{B1}, L_C) = 0$ since $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ and $\boldsymbol{B}_0 \wedge L_{B1}$ meet in a line rather than a point. L_{B1} is therefore a left null eigenvector of $T1$. Similarly, $L_{B2} = \boldsymbol{E}_{BA} \wedge \boldsymbol{B}_{\alpha_2}$ and $L_{B3} = \boldsymbol{E}_{BA} \wedge \boldsymbol{B}_{\alpha_3}$ (where $\boldsymbol{B}_{\alpha_2}$ and $\boldsymbol{B}_{\alpha_3}$ are the points at which the lines joining $\boldsymbol{B}_0$ to $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$ respectively intersect the B plane) are left null eigenvectors of $T2$ and $T3$. It is clear that we have

$$L_{Bi} \vee L_{Bj} = \boldsymbol{E}_{BA} \tag{6.3}$$

for any $i \neq j$. Thus, a way of determining the epipole $\boldsymbol{E}_{BA}$ using all three null vectors is to form a matrix M whose rows are the three null eigenvectors. This matrix will be singular and its right null eigenvector will be the epipole $\boldsymbol{E}_{BA}$, or more correctly, its $\mathcal{E}^3$ equivalent, $\boldsymbol{e}_{ba}$. We see that this is the case since if $M\boldsymbol{x} = 0$, then $[\boldsymbol{u}_1 \cdot \boldsymbol{x}, \boldsymbol{u}_2 \cdot \boldsymbol{x}, \boldsymbol{u}_3 \cdot \boldsymbol{x}]^{\mathrm{T}} = 0$, where $\boldsymbol{u}_i$ is the left null eigenvector of Ti, and that this is satisfied if $\boldsymbol{x} = \boldsymbol{e}_{ba}$.

Similarly, the epipole $\boldsymbol{E}_{CA}$ or $\boldsymbol{e}_{ca}$ is found by considering the right null eigenvectors of $T1, T2, T3$ and proceeding in the same way as for $\boldsymbol{E}_{BA}$.

(b) Epipoles $\boldsymbol{E}_{AB}$ and $\boldsymbol{E}_{AC}$

Given the epipoles $\boldsymbol{E}_{BA}$ and $\boldsymbol{E}_{CA}$ we can now find the next two epipoles. We first form the following six matrices from T:

$$\left.\begin{aligned} U1_{ik} &= T_{i1k}, & V1_{ij} &= T_{ij1},\\ U2_{ik} &= T_{i2k}, & V2_{ij} &= T_{ij2},\\ U3_{ik} &= T_{i3k}, & V3_{ij} &= T_{ij3}. \end{aligned}\right\} \tag{6.4}$$

Consider $V1$; the equivalent linear function is $V1(\boldsymbol{A}, L_B)$. We can write this as

$$V1(\boldsymbol{A}, L_B) = T(\boldsymbol{A}, L_B, LC_1) = (\boldsymbol{C}_0 \wedge LC_1) \wedge \{(\boldsymbol{A}_0 \wedge \boldsymbol{A}) \vee (\boldsymbol{B}_0 \wedge L_B)\}, \tag{6.5}$$

Letting $\boldsymbol{A} = \boldsymbol{E}_{AB}$ then gives $T(\boldsymbol{E}_{AB}, L_B, LC_1) = 0$ if $L_B = \boldsymbol{E}_{BA} \wedge \boldsymbol{B}_j$, for any $j = 1, 2, 3$; see the figures. Thus, if we form a vector $\boldsymbol{m}_j$ by contracting $V1$ with $\boldsymbol{E}_{BA} \wedge B_j$ (and we can do this because we have already found $\boldsymbol{E}_{BA}$), then we know that $\boldsymbol{E}_{AB}$ is orthogonal to this vector. Thus, one way of extracting $\boldsymbol{E}_{AB}$ is to form the matrix M whose three rows are the vectors $\boldsymbol{m}_j$, $j = 1, 2, 3$; then, as above, $\boldsymbol{E}_{AB}$ will be the right null eigenvector of M.

$\boldsymbol{E}_{AC}$ is found in the same way as above but using $U1$ contracted with $L_C = \boldsymbol{E}_{CA} \wedge \boldsymbol{C}_j$. Note that in these derivations we could equally well have used $V2$ or $V3$ and $U2$ or $U3$; this suggests that there will be a better way of extracting these epipoles using all the available information.

(c) Epipoles $\boldsymbol{E}_{BC}$ and $\boldsymbol{E}_{CB}$

Now that we know four of the six epipoles, the final two are relatively easy to extract. Suppose we contract T with $\boldsymbol{E}_{AB}$, to form a matrix $S1$:

$$S1(L_B, L_C) = T(\boldsymbol{E}_{AB}, L_B, L_C) = (\boldsymbol{B}_0 \wedge L_B) \wedge \{(\boldsymbol{A}_0 \wedge \boldsymbol{E}_{AB}) \vee (\boldsymbol{C}_0 \wedge L_C)\}. \tag{6.6}$$

If we let $L_C = \boldsymbol{E}_{CB} \wedge \boldsymbol{C}$ for any $\boldsymbol{C}$, then $(\boldsymbol{A}_0 \wedge \boldsymbol{E}_{AB}) \vee (\boldsymbol{C}_0 \wedge L_C) = \boldsymbol{B}_0$, which makes $T(\boldsymbol{E}_{AB}, L_B, L_C) = 0$. This tells us that we must be able to write the matrix $S1$ in the form $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$, so that it is rank 1; therefore, the epipole $\boldsymbol{E}_{CB}$ will correspond to the vector $\boldsymbol{x}$.

Similarly, the epipole $\boldsymbol{E}_{BC}$ is found by writing the matrix $W1$, which is formed by contracting T with $\boldsymbol{E}_{AC}$, in the form $\boldsymbol{x}'^{\mathrm{T}}\boldsymbol{y}'$ and equating $\boldsymbol{E}_{BC}$ with the vector $\boldsymbol{x}'$.

In any real situation we will, of course, have noise and will therefore not be able to find exact null eigenvectors in the first stage of the epipole extraction. The procedure would then be to project $T1, T2, T3$ down onto the closest determinant zero matrices and work with these to find the first two epipoles.

(d) Estimation of T

Given a sufficient number of point and/or line matches, we can estimate T linearly via the tensor equations. Although straightforward, this will often be very inaccurate in the presence of significant noise since there is nothing in this estimation process which preserves the structure of T. Now that we have a means of extracting the epipoles from T we can make an initial guess at the epipoles from the linear guess for T and then estimate T in the epipole basis as described previously. This will have the advantage of always producing a T with the correct structure. In such a minimization we will be optimizing over 18 parameters; the alternative is to minimize

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

a cost function with 27 degrees of freedom and impose the constraints via Lagrange multipliers.

7. Conclusions

This paper shows how the constraints relating point and line correspondences in multiple views emerge in a geometrically intuitive manner within the geometric algebra framework. For the case of three views, the trilinear constraints are derived using only the intersections of planes and lines; there is no introduction of matrices, Plücker coordinates, etc. When viewed as a linear function the structure of the trifocal tensor in terms of the six epipoles in the construction becomes transparent.

References

Bayro-Corrochano, E., Lasenby, J. & Sommer, G. 1996 Geometric algebra: a framework for computing point and line correspondences and projective structure using n-uncalibrated cameras. *Proc. Int. Conf. on Pattern Recognition (ICPR'96), Vienna, August 1996.*

Beardsley, P., Torr, P. & Zisserman, A. 1996 3D model acquisition from extended image sequences. In *Proc. European Conf. on Computer Vision, 1996*, pp. 683–695.

Clifford, W. K. 1878 Applications of Grassmann's extensive algebra. *Am. J. Math.* **1**, 350–358.

Luong, Q.-T. & Faugeras, O. 1995 The fundamental matrix: theory, algorithms, and stability analysis. *Int. J. Computer Vision* **17**(1), 43–76.

Grassmann, H. 1877 Der Ort der Hamilton'schen Quaternionen in der Ausdehnungslehre. *Math. Ann.* **12** 375.

Hartley, R. 1994 Lines and points in three views – a unified approach. In *ARPA Image Understanding Workshop*. Monterey, California.

Hartley, R. 1995 In defence of the 8-point algorithm. In *Proc. Int. Conf. on Computer Vision*, pp. 882–887.

Hartley, R. 1997 Lines and points in three views and the trifocal tensor. *Int. J. Computer Vision* **22**, 125–140.

Hestenes, D 1966 *Space-time algebra*. Gordon and Breach.

Hestenes, D. & Sobczyk, G. 1984 *Clifford algebra to geometric calculus: a unified language for mathematics and physics*. Dordrecht: D. Reidel.

Hestenes, D. & Ziegler, R. 1991 Projective geometry with Clifford algebra. *Acta Applicandae Mathematicae* **23**, 25–63.

Lasenby, J. & Bayro-Corrohano, E. 1998 Computing invariants in computer vision using geometric algebra. Cambridge University Engineering Department Technical Report, CUED/F-INENG/TR.244

Shashua, A. 1997 Trilinear tensor: the fundamental construct of multiple-view geometry and its applications. In *Algebraic frames for the perception-action cycle* (ed. G. Sommer & J. J. Koenderink). Lecture Notes in Computer Science 1315, pp. 190–206.

Shashua, A. & Maybank, S. J. 1996 Degenerate n-point configurations of three views: do critical surfaces exist? Technical Report TR 96-19. Hebrew University of Jerusalem.

Discussion

M. Sabin (*Numerical Geometry, Cambridge, UK*). Why not cheat? There must be a lot of information coming from approximate camera calibration, approximate camera positions, etc., which should give something good for the epipoles.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

J. Lasenby. Yes, in practical cases one might certainly do this. However, what I wanted to do here was to establish an optimal estimation process to compare with more computationally feasible methods.

A. Fitzgibbon (*Department of Engineering, University of Oxford, UK*). It is conceivable that the problem with the minimization is not one of having the starting point close to the solution, but because the minimal parametrization is too tangled up, too tight, leading to a highly non-smooth error surface. Why not loosen this parametrization (see the Discussion following Hartley, this volume)? This would still guarantee the generation of valid trifocal tensors by, for example, parametrizing the epipoles by three parameters rather than two. And the minimization surface might then smooth out.

J. Lasenby. Yes, I am willing to believe that this may be true, but would that not give it too many degrees of freedom?

A. Fitzgibbon. It doesn't. The numerical minimizer sort of notices that the curvature is zero in that direction and doesn't bother heading off there. What is important is that although there are more than the minimal number of degrees of freedom, the parametrization never produces invalid trifocal tensors.

W. Triggs (*INRIA, Grenoble, France*). I've experimented with an algorithm which simply represents the trifocal tensor as 27 components, and does constrained optimization using all the constraints on it. It is neither slow nor unstable if you implement it reasonably. Each iteration is about twice as fast as the linear method I use, with convergence in 3–10 iterations, even from an arbitrary initialization.

J. Lasenby. This is obviously making the same point as the previous question. Provided the correct constraints are being used, then the numerical issues of minimization may well mean that more degrees of freedom may improve the search surface. How are the constraints imposed? Is it via Lagrange multipliers?

W. Triggs. It is done by sequential quadratic programming, Lagrange multiplier-based constrained optimization. Practically, I think it is often best to take the simplest possible representations and throw something numerical at them.

J. Lasenby. Given there are now 27 parameters, how quick is the minimization?

W. Triggs. If you do a reasonable job with the numerical side, it seems fast enough. Approaches like Richard Hartley's are also very good. But you have to be careful if you want good numerical stability. As soon as you start to do algebraic elimination, such as using the epipoles to eliminate tensor coordinates, the effective degree rises very rapidly and numerical stability is lost.

P. H. S. Torr (*Department of Engineering Science, University of Oxford, UK*). I agree with Bill Triggs and Andrew Fitzgibbon that the constraints can often be over-parametrized if a good minimization algorithm is used, e.g. the algorithm of Powell, that throws away redundant directions as the minimization progresses. Here one can actually put the burden of the work within the numerical algorithm; and if you get a good algorithm, this will in fact converge very quickly.

J. Lasenby. Yes, this could tie in with what I have been finding. Perhaps an over-parametrized surface is much flatter making a search more stable.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

W. Triggs. Moving on to another topic, I'm afraid I don't like the use of the Clifford algebra here. Although the scalar product made sense in the original physical applications of the Clifford algebra, it doesn't in the projective geometry case. If you have calibrated cameras you may have a scalar product. In the projective case the scalar product does not mean anything.

J. Lasenby. In fact it does mean something. The scalar product enters in two well-defined ways; in the projective split and in the computation of the meet. In most systems which try to avoid the use of a scalar product (e.g. differential forms), disguised forms of the scalar product often have to be invented, e.g. the Hodge $*$ operator.

W. Triggs. You can *fake* a scalar product: you can say I've taken a basis; I sum the squares of the coordinates; that's my scalar product.

J. Lasenby. You may wish to do this but I have no need to fake a scalar product. It is an intrinsic part of the framework and one which is very important in many aspects of the analysis, as mentioned previously.

W. Triggs. The scalar product is not projectively invariant. The result is that whenever you do a projective calculation using Clifford algebra, you have to carefully get rid of all those extra scalar product terms you shouldn't have put in in the first place, to get back to the projective, Grassmann–Cayley domain.

J. Lasenby. I do not follow what Mr Triggs means here. I never have to get rid of any unwanted parts. Is he sure he has actually implemented this correctly?

W. Triggs. I think everything you have presented here is essentially Grassmann–Cayley algebra. I do not see any application of the Clifford part of it.

J. Lasenby. Mr Triggs has obviously misunderstood my message here. I would have hoped to have shown that the Grassmann–Cayley algebra is simply contained in the geometric algebra, and that the geometric algebra is a very powerful system which can be used for much more than just projective geometry. In addition, the Clifford approach gives concrete computational tools which are very useful in an actual application.

W. Triggs. I come to a more general criticism of the Clifford approach. Because you have this mixed scalar-plus-vector product you always end up with quantities which do not have a simple transformation law under changes of coordinates. They are formal sums of scalars, vectors, matrices, etc., each with different transformation laws.

J. Lasenby. No, this last assertion is simply wrong. The fundamental product in this system is the geometric product and most elements of the algebra share the same well-defined transformation properties. I get the feeling that Mr Triggs's criticisms arise from feelings he has about the system from a slight knowledge of it, rather than from any serious attempts to use it.

W. Triggs. Whenever you do some sort of calculation, the physically meaningful objects are always invariants or covariants. In other words *tensors* with well-defined transformation laws under changes of image of 3D coordinates.

J. Lasenby. It is completely wrong to think that all meaningful quantities obey the tensor-transformation laws. Spinors happen to be extremely important entities and do not obey the tensor transformation law, but do obey other transformation laws. Any of these quantities can be dealt with in the geometric algebra.

W. Triggs. Can Dr Lasenby give me an example of any quantity used in vision which has this mixed character? That is, a quantity which has two essential parts with different transformation laws?

J. Lasenby. A rotation is an example of something with scalar and bivector parts. Here, the two parts combine to give a single object with the transformation properties of a spinor.

W. Triggs. A rotation is a matrix which transforms homogeneously in a particular way.

J. Lasenby. I'm not sure that this is a useful description. People who do any kind of computer aided design use quaternions to rotate. Hamilton spent many years trying to work out the generalization of the complex numbers which could rotate objects in 3D. He came up with quaternions, and quaternions are nothing more than our 3D rotors with a scalar and bivector part. The advantage of the geometric algebra is that it smoothly extends all this to 4D and projective space.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES