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

# Grassman–Cayley algebra for modelling systems of cameras and the algebraic equations of the manifold of trifocal tensors

The Royal Society

*Phil. Trans. R. Soc. Lond. A* 1998 **356**, 1123-1152
doi: 10.1098/rsta.1998.0214

# Grassmann–Cayley algebra for modelling systems of cameras and the algebraic equations of the manifold of trifocal tensors

By Olivier Faugeras and Théodore Papadopoulo

*INRIA, 2004 Route des Lucioles, B.P. 93, 06902 Sophia-Antipolis Cedex, France*

We show how to use the Grassmann–Cayley algebra to model systems of one, two and three cameras. We start with a brief introduction of the Grassmann–Cayley or double algebra and proceed to demonstrate its use for modelling systems of cameras. In the case of three cameras, we give a new interpretation of the trifocal tensors and study in detail some of the constraints that they satisfy. In particular we prove that simple subsets of those constraints characterize the trifocal tensors, in other words, we give the algebraic equations of the manifold of trifocal tensors.

**Keywords: multiple-view geometry; trifocal tensor; Grassmann–Cayley algebra; fundamental matrix; double algebra**

## 1. Introduction

This article deals with the problem of representing the geometry of several (up to three) pinhole cameras. The idea that we put forward is that this can be done elegantly and conveniently using the formalism of the Grassmann–Cayley algebra. This formalism has already been presented to the computer vision community in several publications (e.g. Carlsson 1994; Faugeras & Mourrain 1995*a*), but no effort has yet been made to systematically explore its use for representing the geometry of systems of cameras.

The thread that is followed here is to study the relations between the three-dimensional (3D) world and its images obtained from one, two or three cameras as well as, when possible, the relations between those images, with the idea of having an algebraic formalism that allows us to compute and estimate things while keeping the geometric intuition which, we think, is important. The Grassmann–Cayley, or double algebra, with its two operators 'join' and 'meet' that correspond to the geometric operations of summing and intersecting vector spaces or projective spaces, was precisely invented to fill this need.

After a very brief introduction to the double algebra (more detailed contemporary discussions can be found, for example, in Doubilet *et al*. (1974) and Barnabei *et al*. (1985)), we apply the algebraic–geometric tools to the description of one pinhole camera in order to introduce such notions as the optical centre, the projection planes and the projection rays which appear later. This introduction is particularly dense and only meant to make the paper more or less self-contained.

We then move on to the case of two cameras and give a simple account of the fundamental matrix (Longuet-Higgins 1981; Faugeras 1992; Carlsson 1994; Luong & Faugeras 1995) which sheds some new light on its structure.

*Phil. Trans. R. Soc. Lond.* A (1998) **356**, 1123–1152

*Printed in Great Britain*

TEX Paper

The next case we study is the case of three cameras. We present a new way of deriving the trifocal tensors which appear in several places in the literature. It has been shown originally by Shashua (1994) that the coordinates of three corresponding points in three views satisfy a set of algebraic relations of degree 3 called the trilinear relations. It was later pointed out by Hartley (1994) that those trilinear relations were in fact arising from a tensor that governed the correspondences of lines between three views and which he called the trifocal tensor. Hartley also correctly pointed out that this tensor had been used, if not formally identified, by researchers working on the problem of the estimation of motion and structure from line correspondences (Spetsakis & Aloimonos 1990*b*). Given three views, there exist three such tensors and we introduce them through the double algebra.

Each tensor seems to depend upon 26 parameters (27 up to scale); these 26 parameters are not independent since the number of degrees of freedom of three views has been shown to be equal to 18 in the projective framework (33 parameters for the 3 perspective projection matrices minus 15 for an unknown projective transformation) (Luong & Viéville 1994). Therefore the trifocal tensor can depend upon at most 18 independent parameters, and its 27 components must satisfy several algebraic constraints, some of which have been elucidated (Shashua & Werman 1995; Avidan & Shashua 1996). We have given a slightly more complete account of those constraints in Faugeras & Papadopoulo (1998), used them to parametrize the tensors minimally (i.e. with 18 parameters) and to design an algorithm for their estimation given line correspondences. In this paper we explore those constraints in great detail and prove that two particular simple subsets are sufficient for a tensor to arise from three cameras (theorems 5.14 and 5.15).

We denote vectors and matrices with bold letters, e.g. $\boldsymbol{x}$ and $\boldsymbol{\mathcal{P}}$. The determinant of a square matrix $\boldsymbol{A}$ is noted $\det(\boldsymbol{A})$. When the matrix is defined by a set of vectors, e.g. $\boldsymbol{A} = [\boldsymbol{a}_1\, \boldsymbol{a}_2\, \boldsymbol{a}_3]$ we use $|\boldsymbol{a}_1\, \boldsymbol{a}_2\, \boldsymbol{a}_3|$. The canonical basis of $\mathbb{R}^3$ is noted $\boldsymbol{e}_i$, where $i = 1, 2, 3$. When dealing with projective spaces, such as $\mathbb{P}^2$ and $\mathbb{P}^3$, we occasionally make the distinction between a projective point, e.g. $x$ and one of its coordinate vectors, $\boldsymbol{x}$. The dual of $\mathbb{P}^n$, the set of projective points is the set of projective lines ($n = 2$) or the set of projective planes ($n = 3$); it is denoted by $\mathbb{P}^{*n}$. Let $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$ be four vectors of $\mathbb{R}^3$. In § 5 *e* we will use Cramer's relation (see Faugeras & Mourrain 1995*b*):

$$|\boldsymbol{bcd}|\boldsymbol{a} - |\boldsymbol{acd}|\boldsymbol{b} + |\boldsymbol{abd}|\boldsymbol{c} - |\boldsymbol{abc}|\boldsymbol{d} = \boldsymbol{0}.$$

## 2. Grassmann–Cayley algebra

Let $E$ be a vector space of dimension 4 on the field $\mathbb{R}$. The corresponding three-dimensional projective space is noted $\mathbb{P}^3$. We consider ordered sets of $k$, $k \leqslant 4$ vectors of $E$. Such ordered sets are called $k$-sequences. We first define an equivalence relation over the set of $k$-sequences as follows. Given two $k$-sequences $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$ and $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_k$, we say that they are equivalent when, for every choice of vectors $\boldsymbol{x}_{k+1}, \cdots, \boldsymbol{x}_n$, we have

$$|\boldsymbol{a}_1 \cdots \boldsymbol{a}_k \boldsymbol{x}_{k+1} \cdots \boldsymbol{x}_4| = |\boldsymbol{b}_1 \cdots \boldsymbol{b}_k \boldsymbol{x}_{k+1} \cdots \boldsymbol{x}_4|. \tag{2.1}$$

That this defines an equivalence relation is immediate. An equivalence class under this relation is called an *extensor* of step $k$ and is written as

$$\boldsymbol{a}_1 \triangledown \boldsymbol{a}_2 \triangledown \cdots \triangledown \boldsymbol{a}_k. \tag{2.2}$$

The product operator $\triangledown$ is called the *join* for reasons related to its geometric interpretation. Let us denote by $G_k(E)$, $1 \leqslant k \leqslant 4$ the vector set generated by all extensors of step $k$, i.e. by all linear combinations of terms like (2.2). It is clear from the definition that $G_1(E) = E$. To be complete one defines $G_0(E)$ to be equal to the field $\mathbb{R}$. The dimension of $G_k(E)$ is

$$\binom{4}{k}.$$

The join operator corresponds to the union of projective subspaces of $P(E)$. The exterior algebra is the direct sum of the vector spaces $G_k$, $k = 0, \ldots, 4$ with the join operator. For example, a point of $\mathbb{P}^3$ is represented by a vector of $E$, its coordinate vector, or equivalently by a point of $G_1(E)$. The join $\boldsymbol{M}_1 \triangledown \boldsymbol{M}_2$ of two distinct points $M_1$ and $M_2$ is the line $(M_1, M_2)$. Similarly, the join $\boldsymbol{M}_1 \triangledown \boldsymbol{M}_2 \triangledown \boldsymbol{M}_3$ of three distinct points $M_1$, $M_2$, $M_3$ is the plane $(M_1, M_2, M_3)$. It is an extensor of step 3. The set of extensors of step 3 represents the sets of planes of $\mathbb{P}^3$.

Let us study in more detail the case of the lines of $\mathbb{P}^3$. Lines are extensors of step 2 and are represented by six-dimensional vectors of $G_2(E)$ with coordinates $(L_{ij}, 1 \leqslant i < j \leqslant 4)$ which satisfy the well-known Plücker relation:

$$L_{12}L_{34} - L_{13}L_{24} + L_{14}L_{23} = 0. \tag{2.3}$$

This equation allows us to define an inner product between two elements $\boldsymbol{L}$ and $\boldsymbol{L}'$ of $G_2(\mathbb{R}^4)$:

$$[\boldsymbol{L} \mid \boldsymbol{L}'] = L_{12}L'_{34} + L'_{12}L_{34} - L_{13}L'_{24} - L'_{13}L_{24} + L_{14}L'_{23} + L'_{14}L_{23}. \tag{2.4}$$

We will use this inner product when we describe the imaging of 3D lines by a camera in § 4. Not all elements of $G_2(E)$ are extensors of step 2 and it is known that:

**Proposition 2.1.** *An element $\boldsymbol{L}$ of $G_2(E)$ represents a line if and only if $[\boldsymbol{L} \mid \boldsymbol{L}]$ is equal to* 0.

To continue our program to define algebraic operations which can be interpreted as geometric operations on the projective subspaces of $P(E)$, we define a second operator, called the meet, and noted $\vartriangle$, on the exterior algebra $G(E)$. This operator corresponds to the geometric operation of intersection of projective subspaces. If $\boldsymbol{A}$ is an extensor of step $k$ and $\boldsymbol{B}$ is an extensor of step $h$, $k+h \geqslant 4$, the meet $\boldsymbol{A} \vartriangle \boldsymbol{B}$ of $\boldsymbol{A}$ and $\boldsymbol{B}$ is an extensor of step $k+h-4$. For example, if $\boldsymbol{\Pi}_1$ and $\boldsymbol{\Pi}_2$ are non-proportional extensors of step 3, i.e. representing two planes, their meet $\boldsymbol{\Pi}_1 \vartriangle \boldsymbol{\Pi}_2$ is an extensor of step 2, representing the line of intersection of the two planes. Similarly, if $\boldsymbol{\Pi}$ is an extensor of step 3 representing a plane and $\boldsymbol{L}$ an extensor of step 2 representing a line, the meet $\boldsymbol{\Pi} \vartriangle \boldsymbol{L}$ is either 0 if $\boldsymbol{L}$ is contained in $\boldsymbol{\Pi}$ or an extensor of step 1 representing the point of intersection of $L$ and $\Pi$. Finally, if $\boldsymbol{\Pi}$ is an extensor of step 3, a plane, and $\boldsymbol{M}$ an extensor of step 1, a point, the meet $\boldsymbol{\Pi} \vartriangle \boldsymbol{M}$ is an extensor of step 0, a real number, which turns out to be equal to $\boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{M}$, the scalar product of the usual vector representation of the plane $\Pi$ with a coordinate vector of the point $M$ which we note $\langle \boldsymbol{\Pi}, \boldsymbol{M} \rangle$. The connection between a plane as a vector in $G_3(E)$ and the usual vector representation $\boldsymbol{\Pi}$ is through the Hodge operator and can be found, for example, in Barnabei *et al.* (1985).

We also define a special element of $G_4(E)$, called the integral. Let $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_4\}$ be a basis of $E$ such that $|\boldsymbol{a}_1 \cdots \boldsymbol{a}_4| = 1$ (the $4 \times 4$ determinant), a unimodal basis. The extensor $\boldsymbol{I} = \boldsymbol{a}_1 \triangledown \cdots \triangledown \boldsymbol{a}_4$ is called the integral. In § 4 we will need the following property of the integral:





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

**Proposition 2.2.** *Let $A$ and $B$ be two extensors such that* $\mathrm{step}(A)+\mathrm{step}(B)=4$. *Then*

$$\boldsymbol{A} \triangledown \boldsymbol{B} = (\boldsymbol{A} \vartriangle \boldsymbol{B}) \triangledown \boldsymbol{I} = |\boldsymbol{AB}|\boldsymbol{I}.$$

The inner product (2.4) has an interesting interpretation in terms of the join $\boldsymbol{L} \triangledown \boldsymbol{L}'$, an extensor of step 4:

**Proposition 2.3.** *We have the following relation:*

$$\boldsymbol{L} \triangledown \boldsymbol{L}' = [\boldsymbol{L} \mid \boldsymbol{L}']\boldsymbol{e}_1 \triangledown \boldsymbol{e}_2 \triangledown \boldsymbol{e}_3 \triangledown \boldsymbol{e}_4, \tag{2.5}$$

*where $\boldsymbol{e}_1, \dots, \boldsymbol{e}_4$ is the canonical basis of $\mathbb{R}^4$.*

In § 3, we will use the following results on lines:

**Proposition 2.4.** *Let $L$ and $L'$ be two lines. If the two lines are represented as the joins of two points $A$ and $B$ and $A'$ and $B'$, respectively, then*

$$[\boldsymbol{L} \mid \boldsymbol{L}'] = |\boldsymbol{ABA'B'}|.$$

*If the two lines are represented as the meets of two planes $P$ and $Q$ and $P'$ and $Q'$, then*

$$[\boldsymbol{L} \mid \boldsymbol{L}'] = |\boldsymbol{PQP'Q'}|. \tag{2.6}$$

*If one line is represented as the meet of two planes $P$ and $Q$ and the other as the join of two points $A'$ and $B'$, then*

$$[\boldsymbol{L} \mid \boldsymbol{L}'] = \langle \boldsymbol{P}|\boldsymbol{A}'\rangle\langle \boldsymbol{Q}|\boldsymbol{B}'\rangle - \langle \boldsymbol{Q}|\boldsymbol{A}'\rangle\langle \boldsymbol{P}|\boldsymbol{B}'\rangle. \tag{2.7}$$

We will also use in § 4 the following result:

**Proposition 2.5.** *Let $L$ and $L'$ be two lines. The inner product $[\boldsymbol{L} \mid \boldsymbol{L}']$ is equal to 0 if and only if the two lines are coplanar.*

## 3. Geometry of one view

We consider that a camera can be modelled accurately as a pinhole and performs a perspective projection. If we consider two arbitrary systems of projective coordinates, for the image and the object space, the relationship between 2D pixels and 3D points can be represented as a linear projective operation which maps points of $\mathbb{P}^3$ to points of $\mathbb{P}^2$. This operation can be described by a $3 \times 4$ matrix $\boldsymbol{\mathcal{P}}$, called the *perspective projection* matrix of the camera:

$$\boldsymbol{m} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \simeq \boldsymbol{\mathcal{P}} \begin{bmatrix} \mathcal{X} \\ Y \\ Z \\ T \end{bmatrix} = \boldsymbol{\mathcal{P}}\boldsymbol{M}. \tag{3.1}$$

This matrix is of rank 3. Its nullspace is therefore of dimension 1, corresponding to a unique point of $\mathbb{P}^3$, the optical centre $C$ of the camera. We give a geometric interpretation of the rows of the projection matrix. We use the notation,

$$\boldsymbol{\mathcal{P}}^{\mathrm{T}} = [\boldsymbol{\varGamma}^{\mathrm{T}}\ \boldsymbol{\varLambda}^{\mathrm{T}}\ \boldsymbol{\varTheta}^{\mathrm{T}}], \tag{3.2}$$





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

where $\boldsymbol{\Gamma}$, $\boldsymbol{\Lambda}$, and $\boldsymbol{\Theta}$ are the row vectors of $\boldsymbol{\mathcal{P}}$. Each of these vectors represents a plane in 3D. These three planes are called the *projection planes* of the camera. The projection equation (3.1) can be rewritten as

$$x : y : z = \langle \boldsymbol{\Gamma}, \boldsymbol{M} \rangle : \langle \boldsymbol{\Lambda}, \boldsymbol{M} \rangle : \langle \boldsymbol{\Theta}, \boldsymbol{M} \rangle,$$

where, for example, $\langle \boldsymbol{\Gamma}, \boldsymbol{M} \rangle$ is the dot product of the plane represented by $\boldsymbol{\Gamma}$ with the point represented by $\boldsymbol{M}$. This relation is equivalent to the three scalar equations, of which two are independent:

$$\left.\begin{aligned} x\langle \boldsymbol{\Lambda}, \boldsymbol{M} \rangle - y\langle \boldsymbol{\Gamma}, \boldsymbol{M} \rangle = 0, \\ y\langle \boldsymbol{\Theta}, \boldsymbol{M} \rangle - z\langle \boldsymbol{\Lambda}, \boldsymbol{M} \rangle = 0, \\ z\langle \boldsymbol{\Lambda}, \boldsymbol{M} \rangle - x\langle \boldsymbol{\Theta}, \boldsymbol{M} \rangle = 0. \end{aligned}\right\} \tag{3.3}$$

The planes of equation $\langle \boldsymbol{\Gamma}, \boldsymbol{M} \rangle = 0$, $\langle \boldsymbol{\Lambda}, \boldsymbol{M} \rangle = 0$ and $\langle \boldsymbol{\Theta}, \boldsymbol{M} \rangle = 0$ are mapped to the image lines of equations $x = 0$, $y = 0$, and $z = 0$, respectively. We have the proposition:

**Proposition 3.1.** *The three projection planes of a perspective camera intersect the retinal plane along the three lines going through the first three points of the standard projective basis.*

The optical centre is the unique point $C$ which satisfies $\boldsymbol{\mathcal{P}}\boldsymbol{C} = \boldsymbol{0}$. Therefore this point is the intersection of the three planes represented by $\boldsymbol{\Gamma}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\Theta}$. In the Grassmann–Cayley formalism, it is represented by the meet of those three planes $\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}$. This is illustrated in figure 1. Because of the definition of the meet operator, the projective coordinates of $C$ are the four $3 \times 3$ minors of matrix $\boldsymbol{\mathcal{P}}$:

**Proposition 3.2.** *The optical centre $C$ of the camera is the meet $\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}$ of the three projection planes.*

The three projection planes intersect along the three lines $\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}$, $\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}$ and $\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma}$ called the *projection rays*. These three lines meet at the optical centre $C$ and intersect the retinal plane at the first three points $e_1$, $e_2$ and $e_3$ of the standard projective basis. Given a pixel $m$, its optical ray $(C, m)$ can be expressed very simply as a linear combination of the three projection rays:

**Proposition 3.3.** *The optical ray $L_m$ of the pixel $m$ of projective coordinates $(x, y, z)$ is given by*

$$\boldsymbol{L}_m = x\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} + y\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} + z\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}. \tag{3.4}$$

*Proof*. Let consider the plane $x\boldsymbol{\Lambda} - y\boldsymbol{\Gamma}$. This plane contains the optical centre $\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}$ since both $\boldsymbol{\Lambda}$ and $\boldsymbol{\Theta}$ do. Moreover, it also contains the point $M$. To see this, let us take the dot product:

$$\langle x\boldsymbol{\Lambda} - y\boldsymbol{\Gamma}, \boldsymbol{M} \rangle = x\langle \boldsymbol{\Lambda}, \boldsymbol{M} \rangle - y\langle \boldsymbol{\Gamma}, \boldsymbol{M} \rangle,$$

but since $x : y = \langle \boldsymbol{\Gamma}, \boldsymbol{M} \rangle : \langle \boldsymbol{\Lambda}, \boldsymbol{M} \rangle$, this expression is equal to 0. Therefore, the plane $x\boldsymbol{\Lambda}–y\boldsymbol{\Gamma}$ contains the optical ray $(C, m)$. Similarly, the planes $z\boldsymbol{\Gamma}–x\boldsymbol{\Theta}$ and $y\boldsymbol{\Theta}–z\boldsymbol{\Lambda}$ also contain the optical ray $(C, m)$, which can therefore be found as the intersection of any of these two planes. Taking for instance the first two planes that we considered, we obtain

$$(x\boldsymbol{\Lambda} - y\boldsymbol{\Gamma}) \vartriangle (z\boldsymbol{\Gamma} - x\boldsymbol{\Theta}) = -x(x\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} + y\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} + z\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}).$$





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

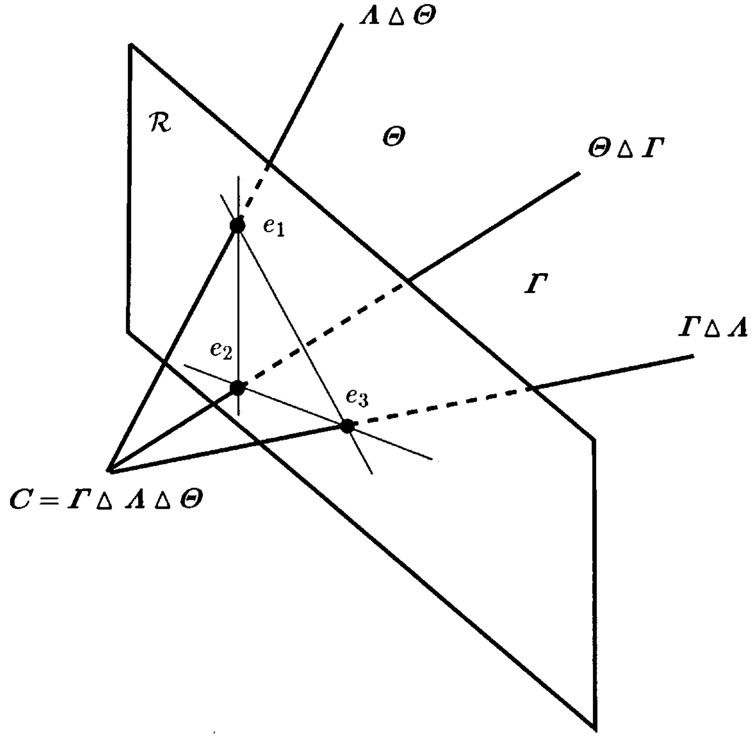


Figure 1. Geometrical interpretation of the three rows of the projection matrix as planes. The three projection planes $\boldsymbol{\Gamma}$, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Theta}$ are projected into the axes of the retinal coordinate system. The three projection rays intersect the retinal plane at the first three points of the retinal projective basis. The three projection planes meet at the optical centre.

The scale factor $x$ is not significant, and if it is zero, another choice of two planes can be made for the calculation. We conclude that the optical ray $L_m = (C, m)$ is represented by the line $x\boldsymbol{\Lambda} \triangle \boldsymbol{\Theta} + y\boldsymbol{\Theta} \triangle \boldsymbol{\Gamma} + z\boldsymbol{\Gamma} \triangle \boldsymbol{\Lambda}$ (see proposition 3.6) for another interesting interpretation of this formula). ■

We also have an interesting interpretation of matrix $\boldsymbol{\mathcal{P}}^{\mathrm{T}}$ which we give in the following proposition:

**Proposition 3.4.** *The transpose $\boldsymbol{\mathcal{P}}^{\mathrm{T}}$ of the perspective projection matrix defines a mapping from the set of lines of the retinal plane to the set of planes going through the optical centre. This mapping associates to the line $l$, represented by the vector $\boldsymbol{l} = [x, y, z]^{\mathrm{T}}$, the plane $\boldsymbol{\mathcal{P}}^{\mathrm{T}}\boldsymbol{l} = x\boldsymbol{\Gamma} + y\boldsymbol{\Lambda} + z\boldsymbol{\Theta}$.*

*Proof*. The fact that $\boldsymbol{\mathcal{P}}^{\mathrm{T}}$ maps planar lines to planes is a consequence of duality. The plane $x\boldsymbol{\Gamma} + y\boldsymbol{\Lambda} + z\boldsymbol{\Theta}$ contains the optical centre since it is contained by each projection plane. ■

Having discussed the imaging of points, let us tackle the imaging of lines, which plays a central role in subsequent parts of this paper. Given two 3D points $M_1$ and $M_2$, the line $\boldsymbol{L} \equiv \boldsymbol{M}_1 \triangledown \boldsymbol{M}_2$ is an element of $G_3(\mathbb{R}^4)$ represented by its Plücker coordinates. The image $l$ of that line through a camera defined by the perspective projection matrix $\boldsymbol{\mathcal{P}}$ is represented by the $3 \times 1$ vector:

$$\begin{aligned} \boldsymbol{l} &= \boldsymbol{\mathcal{P}}\boldsymbol{M}_1 \times \boldsymbol{\mathcal{P}}\boldsymbol{M}_2 \\ &= [\langle \boldsymbol{\Lambda}, \boldsymbol{M}_1\rangle\langle \boldsymbol{\Theta}, \boldsymbol{M}_2\rangle - \langle \boldsymbol{\Theta}, \boldsymbol{M}_2\rangle\langle \boldsymbol{\Lambda}, \boldsymbol{M}_1\rangle, \\ &\qquad \langle \boldsymbol{\Theta}, \boldsymbol{M}_1\rangle\langle \boldsymbol{\Gamma}, \boldsymbol{M}_2\rangle - \langle \boldsymbol{\Gamma}, \boldsymbol{M}_2\rangle\langle \boldsymbol{\Theta}, \boldsymbol{M}_1\rangle, \langle \boldsymbol{\Gamma}, \boldsymbol{M}_1\rangle\langle \boldsymbol{\Lambda}, \boldsymbol{M}_2\rangle - \langle \boldsymbol{\Lambda}, \boldsymbol{M}_2\rangle\langle \boldsymbol{\Gamma}, \boldsymbol{M}_1\rangle]^{\mathrm{T}}. \end{aligned}$$

Equation (2.7) of proposition 2.4 allows us to recognize the inner products of the projection rays of the camera with the line $L$:

$$\boldsymbol{l} \simeq [[\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} \mid \boldsymbol{L}], [\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} \mid \boldsymbol{L}], [\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} | \boldsymbol{L}]]^{\mathrm{T}}. \tag{3.5}$$

We can rewrite this in matrix form,

$$\boldsymbol{l} \simeq \tilde{\boldsymbol{\mathcal{P}}}\boldsymbol{L}, \tag{3.6}$$

where $\tilde{\boldsymbol{\mathcal{P}}}$ is the following $3 \times 6$ matrix:

$$\begin{bmatrix} \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} \\ \boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} \\ \boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} \end{bmatrix}.$$

The matrix $\tilde{\boldsymbol{\mathcal{P}}}$ plays for 3D lines the same role that the matrix $\boldsymbol{\mathcal{P}}$ plays for 3D points. Equation (3.6) is thus equivalent to

$$l_1 : l_2 : l_3 = [\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} \mid \boldsymbol{L}] : [\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} \mid \boldsymbol{L}] : [\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} | \boldsymbol{L}].$$

We have the following proposition:

**Proposition 3.5.** *The pinhole camera also defines a mapping from the set of lines of $\mathbb{P}^3$ to the set of lines of $\mathbb{P}^2$. This mapping is an application from the projective space $P(G_2(\mathbb{R}^4))$ (the set of 3D lines) to the projective space $P(G_2(\mathbb{R}^3))$ (the set of 2D lines). It is represented by a $6 \times 4$ matrix, noted $\tilde{\boldsymbol{\mathcal{P}}}$ whose row vectors are the Plücker coordinates of the projection rays of the camera:*

$$\tilde{\boldsymbol{\mathcal{P}}} = \begin{bmatrix} \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} \\ \boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} \\ \boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} \end{bmatrix}. \tag{3.7}$$

*The image $l$ of a 3D line $L$ is given by*

$$l_1 : l_2 : l_3 = [\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} | \boldsymbol{L}] : [\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} | \boldsymbol{L}] : [\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} | \boldsymbol{L}].$$

*The nullspace of this mapping contains the set of lines going through the optical centre of the camera.*

*Proof*. We have already proved the first part. Regarding the nullspace, if $L$ is a 3D line such that $\tilde{\boldsymbol{\mathcal{P}}}\boldsymbol{L} = \boldsymbol{0}$, then $L$ intersect all three projection rays of the camera and hence goes through the optical centre. ■

The dual interpretation is also of interest:

**Proposition 3.6.** *The $3 \times 6$ matrix $\tilde{\boldsymbol{\mathcal{P}}}^{\mathrm{T}}$ represents a mapping from $\mathbb{P}^2$ to the set of 3D lines, subset of $P(G_2(\mathbb{R}^4))$, which associates to each pixel $m$ its optical ray $L_m$.*





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

*Proof*. Since $\tilde{\boldsymbol{\mathcal{P}}}$ represents a linear mapping from $G_2(\mathbb{R}^4)$ to $G_2(\mathbb{R}^3)$, $\tilde{\boldsymbol{\mathcal{P}}}^{\mathrm{T}}$ represents a linear mapping from the dual $G_2(\mathbb{R}^3)^*$ of $G_2(\mathbb{R}^3)$ which we can identify to $G_1(\mathbb{R}^3)$, to the dual $G_2(\mathbb{R}^4)^*$ which we can identify to $G_2(\mathbb{R}^4)$ (see, for example, Barnabei *et al*. (1985) for the definition of the Hodge operator and duality). Hence it corresponds to a morphism from $\mathbb{P}^2$ to $P(G_2(\mathbb{R}^4))$. If the pixel $m$ has projective coordinates $x$, $y$ and $z$, we have:

$$\tilde{\boldsymbol{\mathcal{P}}}^{\mathrm{T}}\boldsymbol{m} \simeq x\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} + y\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} + z\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}$$

and we recognize the right-hand side to be a representation of the optical ray $L_m$. ■

### (a) Affine digression

In the affine framework we can give an interesting interpretation of the third projection plane of the perspective projection matrix:

**Proposition 3.7.** *The third projection plane* $\boldsymbol{\Theta}$ *of the perspective projection matrix* $\boldsymbol{\mathcal{P}}$ *is the focal plane of the camera.*

*Proof*. The points of the plane of equation $\langle \boldsymbol{\Theta}, \tilde{\boldsymbol{M}} \rangle = 0$ are mapped to the points in the retinal plane such that $z = 0$. This is the equation of the line at infinity in the retinal plane. The plane represented by $\boldsymbol{\Theta}$ is therefore the set of points in 3D space which do not project at finite distance in the retinal plane. These points form the *focal plane*, which is the plane containing the optical centre, and parallel to the retinal plane. ■

When the focal plane is the plane at infinity, i.e. $\boldsymbol{\Theta} \simeq \boldsymbol{e}_4$, the camera is called an affine camera and performs a parallel projection. Note that this class of cameras is important in applications, including the orthographic, weak perspective, and scaled orthographic projections.

## 4. Geometry of two views

In the case of two cameras, it is well-known that the geometry of correspondences between the two views can be described compactly by the fundamental matrix, noted $\boldsymbol{F}_{12}$, which associates to each pixel $m_1$ of the first view its epipolar line, noted $l_{m_1}$ in the second image:

$$\boldsymbol{l}_{m_1} \simeq \boldsymbol{F}_{12}\boldsymbol{m}_1.$$

Similarly, $\boldsymbol{F}_{21} = \boldsymbol{F}_{12}^{\mathrm{T}}$ associates to a pixel $m_2$ of the second view its epipolar line $l_{m_2}$ in the first view.

The matrix $\boldsymbol{F}_{12}$ (resp. $\boldsymbol{F}_{21}$) is of rank 2, the point in its null-space is the epipole $e_{1,2}$ (resp. the epipole $e_{2,1}$):

$$\boldsymbol{F}_{12}\boldsymbol{e}_{1,2} = \boldsymbol{F}_{21}\boldsymbol{e}_{2,1} = \boldsymbol{0}.$$

There is a very simple and natural way of deriving the fundamental matrix in the Grassmann–Cayley formalism. We use the simple idea that two pixels $m$ and $m'$ are in correspondence if and only if their optical rays $(C, m)$ and $(C', m')$ intersect. We then write down this condition using proposition 2.5 and obtain the fundamental matrix using the properties of the double algebra.

We will denote the rows of $\boldsymbol{\mathcal{P}}$ by $\boldsymbol{\Gamma}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\Theta}$, and the rows of $\boldsymbol{\mathcal{P}}'$ by $\boldsymbol{\Gamma}'$, $\boldsymbol{\Lambda}'$, $\boldsymbol{\Theta}'$. We have the following proposition:

**Proposition 4.1.** *The expression of the fundamental matrix $\boldsymbol{F}$ as a function of the row vectors of the matrices $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{P}}'$ is*

$$\boldsymbol{F} = \begin{bmatrix} |\boldsymbol{\Lambda\Theta\Lambda'\Theta'}| & |\boldsymbol{\Theta\Gamma\Lambda'\Theta'}| & |\boldsymbol{\Gamma\Lambda\Lambda'\Theta'}| \\ |\boldsymbol{\Lambda\Theta\Theta'\Gamma'}| & |\boldsymbol{\Theta\Gamma\Theta'\Gamma'}| & |\boldsymbol{\Gamma\Lambda\Theta'\Gamma'}| \\ |\boldsymbol{\Lambda\Theta\Gamma'\Lambda'}| & |\boldsymbol{\Theta\Gamma\Gamma'\Lambda'}| & |\boldsymbol{\Gamma\Lambda\Gamma'\Lambda'}| \end{bmatrix}. \tag{4.1}$$

*Proof*. Let $m$ and $m'$ be two pixels. They are in correspondence if and only if their optical rays $(C, m) = L_m$ and $(C', m') = L'_{m'}$ intersect. According to proposition 2.5, this is equivalent to the fact that the inner product $[\boldsymbol{L}_m \mid \boldsymbol{L}'_{m'}]$ of the two optical rays is equal to 0. Let us translate this algebraically. Let $(x, y, z)$ (resp. $(x', y', z')$) be the coordinates of $m$ (resp. $m'$). Using proposition 3.6, we write

$$\boldsymbol{L}_m \simeq \tilde{\boldsymbol{\mathcal{P}}}^{\mathrm{T}}\boldsymbol{m} = x\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} + y\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} + z\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}$$

and

$$\boldsymbol{L}'_{m'} \simeq \tilde{\boldsymbol{\mathcal{P}}}'^{\mathrm{T}}\boldsymbol{m}' = x'\boldsymbol{\Lambda}' \vartriangle \boldsymbol{\Theta}' + y'\boldsymbol{\Theta}' \vartriangle \boldsymbol{\Gamma}' + z'\boldsymbol{\Gamma}' \vartriangle \boldsymbol{\Lambda}'.$$

We now want to compute $[\boldsymbol{L}_m \mid \boldsymbol{L}'_{m'}]$. In order to do this, we use proposition 2.3 and compute $\boldsymbol{L}_m \triangledown \boldsymbol{L}'_{m'}$:

$$\boldsymbol{L}_m \triangledown \boldsymbol{L}'_{m'} \simeq (x\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} + y\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} + z\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}) \triangledown (x'\boldsymbol{\Lambda}' \vartriangle \boldsymbol{\Theta}' + y'\boldsymbol{\Theta}' \vartriangle \boldsymbol{\Gamma}' + z'\boldsymbol{\Gamma}' \vartriangle \boldsymbol{\Lambda}').$$

Using the linearity of the join operator, we obtain an expression which is bilinear in the coordinates of $m$ and $m'$ and contains terms such as

$$(\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}) \triangledown (\boldsymbol{\Lambda}' \vartriangle \boldsymbol{\Theta}').$$

Since $\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}$ and $\boldsymbol{\Lambda}' \vartriangle \boldsymbol{\Theta}'$ are extensors of step 2, we can apply proposition 2.2 and write

$$(\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}) \triangledown (\boldsymbol{\Lambda}' \vartriangle \boldsymbol{\Theta}') = |\boldsymbol{\Lambda\Theta\Lambda'\Theta'}|\boldsymbol{I},$$

where $\boldsymbol{I}$ is the integral defined in § 2. We have similar expressions for all terms in $\boldsymbol{L}_m \triangledown \boldsymbol{L}'_{m'}$. We thus obtain

$$\boldsymbol{L}_m \triangledown \boldsymbol{L}'_{m'} = (\boldsymbol{m}'^{\mathrm{T}}\boldsymbol{F}\boldsymbol{m})\boldsymbol{I},$$

where the $3\times 3$ matrix $\boldsymbol{F}$ is defined by equation (4.1). Since $\boldsymbol{L}_m \triangledown \boldsymbol{L}'_{m'} = [\boldsymbol{L}_m \mid \boldsymbol{L}'_{m'}]\boldsymbol{I}$, the conclusion follows. ∎

Let us determine the epipoles in this formalism. We have the following simple proposition:

**Proposition 4.2.** *The expression of the epipoles $e$ and $e'$ as a function of the row vectors of the matrices $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{P}}'$ is*

$$\boldsymbol{e} \simeq \begin{bmatrix} |\boldsymbol{\Gamma\Gamma'\Lambda'\Theta'}| \\ |\boldsymbol{\Lambda\Gamma'\Lambda'\Theta'}| \\ |\boldsymbol{\Theta\Gamma'\Lambda'\Theta'}| \end{bmatrix}, \qquad \boldsymbol{e}' \simeq \begin{bmatrix} |\boldsymbol{\Gamma'\Gamma\Lambda\Theta}| \\ |\boldsymbol{\Lambda'\Gamma\Lambda\Theta}| \\ |\boldsymbol{\Theta'\Gamma\Lambda\Theta}| \end{bmatrix}.$$

*Proof*. We have seen previously that $e$ (resp. $e'$) is the image of $C'$ (resp. $C$) by the first (resp. the second) camera. According to proposition 3.2, these optical centres are represented by the vectors of $G_1(\mathbb{R}^4)$ $\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}$ and $\boldsymbol{\Gamma}' \vartriangle \boldsymbol{\Lambda}' \vartriangle \boldsymbol{\Theta}'$. Therefore we have, for example, that the first coordinate of $e$ is

$$\langle \boldsymbol{\Gamma}, \boldsymbol{\Gamma}' \vartriangle \boldsymbol{\Lambda}' \vartriangle \boldsymbol{\Theta}' \rangle,$$

which is equal to $(\boldsymbol{\Gamma}' \vartriangle \boldsymbol{\Lambda}' \vartriangle \boldsymbol{\Theta}') \vartriangle \boldsymbol{\Gamma} = -|\boldsymbol{\Gamma\Gamma'\Lambda'\Theta'}|$. ∎

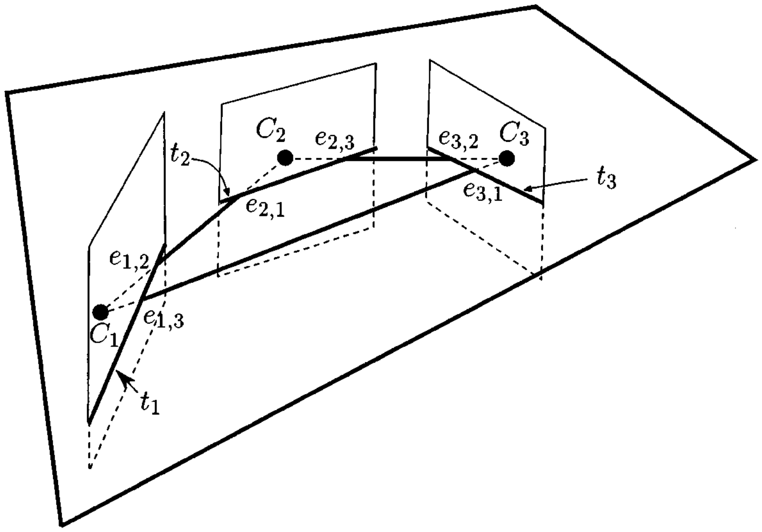


Figure 2. The trifocal geometry.

### (*a*) *Another affine digression*

Let us now assume that the two cameras are affine, i.e. $\boldsymbol{\Theta} \simeq \boldsymbol{\Theta}' \simeq \boldsymbol{e}_4$ (in fact it is sufficient that $\boldsymbol{\Theta} \simeq \boldsymbol{\Theta}'$). Because of the standard properties of determinants, it is clear from equation (4.1) that the fundamental matrix takes a special form:

**Proposition 4.3.** *The fundamental matrix of two affine cameras has the form:*

$$\boldsymbol{F} = \begin{bmatrix} 0 & 0 & |\boldsymbol{\Gamma\Lambda\Lambda'\Theta'}| \\ 0 & 0 & |\boldsymbol{\Gamma\Lambda\Theta'\Gamma'}| \\ |\boldsymbol{\Lambda\Theta\Gamma'\Lambda'}| & |\boldsymbol{\Theta\Gamma\Gamma'\Lambda'}| & |\boldsymbol{\Gamma\Lambda\Gamma'\Lambda'}| \end{bmatrix}. \tag{4.2}$$

## 5. Geometry of three views

### (*a*) *Trifocal geometry from binocular geometry*

When we add one more view, the geometry becomes more intricate; see figure 2. Note that we assume that the three optical centres $C_1$, $C_2$, $C_3$ are different, and call this condition the *general viewpoint assumption*. When they are not aligned they define a plane, called the *trifocal plane*, which intersects the three image planes along the *trifocal lines* $t_1$, $t_2$, $t_3$ which contain the epipoles $e_{i,j}$, $i \neq j$, $i = 1, \dots, 3$, $j = 1, \dots, 3$. The three fundamental matrices $\boldsymbol{F}_{12}$, $\boldsymbol{F}_{23}$ and $\boldsymbol{F}_{31}$ are not independent since they must satisfy the three constraints:

$$\boldsymbol{e}_{2,3}^{\mathrm{T}} \boldsymbol{F}_{12} \boldsymbol{e}_{1,3} = \boldsymbol{e}_{3,1}^{\mathrm{T}} \boldsymbol{F}_{23} \boldsymbol{e}_{2,1} = \boldsymbol{e}_{1,2}^{\mathrm{T}} \boldsymbol{F}_{31} \boldsymbol{e}_{3,2} = 0, \tag{5.1}$$

which arise naturally from the trifocal plane: for example, the epipolar line in view 2 of the epipole $e_{1,3}$ is represented by $\boldsymbol{F}_{12}\boldsymbol{e}_{1,3}$ and is the image in view 2 of the optical ray $(C_1, e_{1,3})$ which is identical to the line $(C_1, C_3)$. This image is the trifocal line $t_2$ which goes through $e_{2,3}$ (see figure 2), hence the first equation in (5.1).

This has an important impact on the way we have to estimate the fundamental matrices when three views are available: very efficient and robust algorithms are





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

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



now available to estimate the fundamental matrix between two views from point correspondences (Zhang *et al*. 1995; Torr & Zissermann 1997; Hartley 1995). The constraints (5.1) mean that these algorithms cannot be used blindly to estimate the three fundamental matrices independently because the resulting matrices will not satisfy the constraints causing errors in further processes such as prediction.

Indeed, one of the important uses of the fundamental matrices in trifocal geometry is the fact that they in general allow to *predict* from two correspondences, say $(m_1, m_2)$ where the point $m_3$ should be in the third image: it is simply at the intersection of the two epipolar lines represented by $\boldsymbol{F}_{13}\boldsymbol{m}_1$ and $\boldsymbol{F}_{23}\boldsymbol{m}_2$, when this intersection is well defined.

It is not well defined in two cases:

1. In the general case where the three optical centres are not aligned, when the 3D points lie in the trifocal plane (the plane defined by the three optical centres), the prediction with the fundamental matrices fails because, in the previous example both epipolar lines are equal to the trifocal line $t_3$.

2. In the special case where the three optical centres are aligned, the prediction with the fundamental matrices fails always since, for example, $\boldsymbol{F}_{13}\boldsymbol{m}_1 \simeq \boldsymbol{F}_{23}\boldsymbol{m}_2$, for all corresponding pixels $m_1$ and $m_2$ in views 1 and 2, i.e. such that $\boldsymbol{m}_2^{\mathrm{T}}\boldsymbol{F}_{12}\boldsymbol{m}_1 = 0$.

For those two reasons, as well as for the estimation problem mentioned previously, it is interesting to characterize the geometry of three views by another entity, the *trifocal tensor*.

The trifocal tensor is really meant for describing line correspondences and, as such, has been well-known under disguise in the part of the computer vision community dealing with the problem of structure from motion (Spetsakis & Aloimonos 1990*a*, *b*; Weng *et al*. 1992) before it was formally identified by Hartley (1994) and Shashua (1995).

### (*b*) *The trifocal tensors*

Let us consider three views, with projection matrices $\boldsymbol{\mathcal{P}}_n$, $n = 1, 2, 3$, a 3D line $L$ with images $l_n$. Given two images $l_j$ and $l_k$ of $L$, $L$ can be defined as the intersection (the meet) of the two planes $\boldsymbol{\mathcal{P}}_j^{\mathrm{T}}\boldsymbol{l}_j$ and $\boldsymbol{\mathcal{P}}_k^{\mathrm{T}}\boldsymbol{l}_k$:

$$\boldsymbol{L} \simeq \boldsymbol{\mathcal{P}}_j^{\mathrm{T}}\boldsymbol{l}_j \mathbin{\triangle} \boldsymbol{\mathcal{P}}_k^{\mathrm{T}}\boldsymbol{l}_k.$$

The vector $\boldsymbol{L}$ is the $6 \times 1$ vector of the Plücker coordinates of the line $L$.

Let us write the right-hand side of this equation explicitly in terms of the row vectors of the matrices $\boldsymbol{\mathcal{P}}_j$ and $\boldsymbol{\mathcal{P}}_k$ and the coordinates of $\boldsymbol{l}_j$ and $\boldsymbol{l}_k$:

$$\boldsymbol{L} \simeq (l_j^1\boldsymbol{\Gamma}_j + l_j^2\boldsymbol{\Lambda}_j + l_j^3\boldsymbol{\Theta}_j) \mathbin{\triangle} (l_k^1\boldsymbol{\Gamma}_k + l_k^2\boldsymbol{\Lambda}_k + l_k^3\boldsymbol{\Theta}_k).$$

By expanding the meet operator in the previous equation, it can be rewritten in the following less compact form, with the advantage of making the dependency on the projection planes of the matrices $\boldsymbol{\mathcal{P}}_j$ and $\boldsymbol{\mathcal{P}}_k$ explicit:

$$\boldsymbol{L} \simeq \boldsymbol{l}_j^{\mathrm{T}} \begin{bmatrix} \boldsymbol{\Gamma}_j \mathbin{\triangle} \boldsymbol{\Gamma}_k & \boldsymbol{\Gamma}_j \mathbin{\triangle} \boldsymbol{\Lambda}_k & \boldsymbol{\Gamma}_j \mathbin{\triangle} \boldsymbol{\Theta}_k \\ \boldsymbol{\Lambda}_j \mathbin{\triangle} \boldsymbol{\Gamma}_k & \boldsymbol{\Lambda}_j \mathbin{\triangle} \boldsymbol{\Lambda}_k & \boldsymbol{\Lambda}_j \mathbin{\triangle} \boldsymbol{\Theta}_k \\ \boldsymbol{\Theta}_j \mathbin{\triangle} \boldsymbol{\Lambda}_k & \boldsymbol{\Theta}_j \mathbin{\triangle} \boldsymbol{\Lambda}_k & \boldsymbol{\Theta}_j \mathbin{\triangle} \boldsymbol{\Theta}_k \end{bmatrix} \boldsymbol{l}_k. \tag{5.2}$$

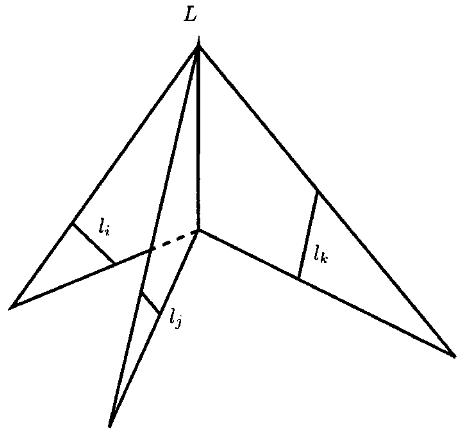


Figure 3. The line $l_i$ is the image by camera $i$ of the 3D line $L$ intersection of the planes defined by the optical centres of the cameras $j$ and $k$ and the lines $l_j$ and $l_k$, respectively.

This equation should be interpreted as giving the Plücker coordinates of $L$ as a linear combination of the lines defined by the meets of the projection planes of the perspective matrices $\boldsymbol{\mathcal{P}}_j$ and $\boldsymbol{\mathcal{P}}_k$, the coefficients being the products of the projective coordinates of the lines $l_j$ and $l_k$.

The image $l_i$ of $L$ is therefore obtained by applying the matrix $\tilde{\boldsymbol{\mathcal{P}}}_i$ (defined in § 3) to the Plücker coordinates of $L$, hence the equation:

$$\boldsymbol{l}_i \simeq \tilde{\boldsymbol{\mathcal{P}}}_i(\boldsymbol{\mathcal{P}}_j^{\mathrm{T}}\boldsymbol{l}_j \vartriangle \boldsymbol{\mathcal{P}}_k^{\mathrm{T}}\boldsymbol{l}_k), \tag{5.3}$$

which is valid for $i \neq j \neq k$. Note that if we exchange view $j$ and view $k$, we just change the sign of $\boldsymbol{l}_i$ and therefore we do not change $l_i$. A geometric interpretation of this is shown in figure 3. For convenience, we rewrite equation (5.3) in a more compact form:

$$\boldsymbol{l}_i \simeq \boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k). \tag{5.4}$$

This expression can be also put in a slightly less compact form with the advantage of making the dependency on the projection planes of the matrices $\boldsymbol{\mathcal{P}}_n$, $n = 1, 2, 3$ explicit:

$$\boldsymbol{l}_i \simeq \left[\boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{G}_i^1\boldsymbol{l}_k \quad \boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{G}_i^2\boldsymbol{l}_k \quad \boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{G}_i^3\boldsymbol{l}_k\right]^{\mathrm{T}}. \tag{5.5}$$

This is, in the projective framework, the exact analogue of the equation used in the work of Spetsakis & Aloimonos (1990*b*) to study the structure from motion problem from line correspondences.

The three $3 \times 3$ matrices $\boldsymbol{G}_i^n$, $n = 1, 2, 3$ are obtained from equations (5.2) and (5.3):

$$\boldsymbol{G}_i^1 = \begin{bmatrix} |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_j\boldsymbol{\Lambda}_k| & |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_j\boldsymbol{\Theta}_k| \\ |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Lambda}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Lambda}_j\boldsymbol{\Lambda}_k| & |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Lambda}_j\boldsymbol{\Theta}_k| \\ |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Theta}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Theta}_j\boldsymbol{\Lambda}_k| & |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Theta}_j\boldsymbol{\Theta}_k| \end{bmatrix}. \tag{5.6}$$

Note that equation (5.3) allows us to predict the coordinates of a line $l_i$ in image $i$, given two images $l_j$ and $l_k$ of an unknown 3D line in images $j$ and $k$, except in two cases where $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k) = \boldsymbol{0}$:

Figure 4

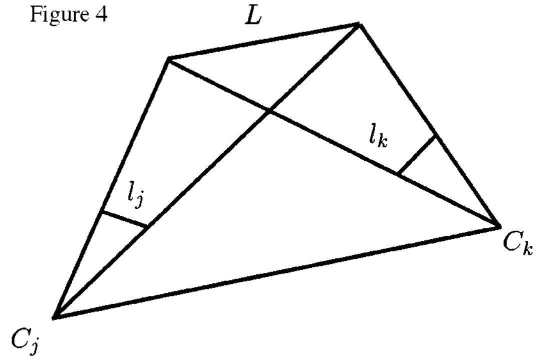


Figure 5

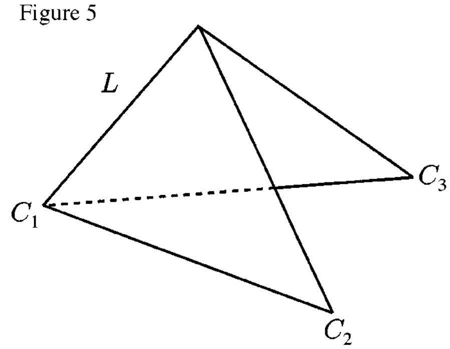


Figure 4. When $l_j$ and $l_k$ are corresponding epipolar lines, the two planes $\boldsymbol{\mathcal{P}}_j^{\mathrm{T}}\boldsymbol{l}_j$ and $\boldsymbol{\mathcal{P}}_k^{\mathrm{T}}\boldsymbol{l}_k$ are identical and therefore $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k) = \mathbf{0}$.

Figure 5. When $l_j$ and $l_k$ are epipolar lines with respect to view $i$, the line $l_i$ is reduced to a point, hence $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k) = \mathbf{0}$.

1. When the two planes determined by $l_j$ and $l_k$ are identical, i.e. when $l_j$ and $l_k$ are corresponding epipolar lines between views $j$ and $k$. This is equivalent to saying that the 3D line $L$ is in an epipolar plane of the camera pair $(j, k)$. The meet that appears in equation (5.3) is then 0 and the line $l_i$ is undefined; see figure 4. If $L$ is not in an epipolar plane of the camera pair $(i, j)$ then we can use the equation:
$$\boldsymbol{l}_k \simeq \tilde{\boldsymbol{\mathcal{P}}}_k(\boldsymbol{\mathcal{P}}_i^{\mathrm{T}}\boldsymbol{l}_i \vartriangle \boldsymbol{\mathcal{P}}_j^{\mathrm{T}}\boldsymbol{l}_j)$$
to predict $l_k$ from the images $l_i$ and $l_j$ of $L$. If $L$ is also in an epipolar plane of the camera pair $(i, j)$ it is in the trifocal plane of the three cameras and prediction is not possible by any of the formulas, such as (5.3).

2. When $l_j$ and $l_k$ are epipolar lines between views $i$ and $j$ and $i$ and $k$, respectively. This is equivalent to saying that they are the images of the same optical ray in view $i$ and that $l_i$ is reduced to a point (see figure 5).

Except in those two cases, we have defined an application $\mathcal{T}_i$ from $\mathbb{P}^{*2} \times \mathbb{P}^{*2}$, the Cartesian product of two duals of the projective plane, into $\mathbb{P}^{*2}$. This application is represented by an application $\boldsymbol{\mathcal{T}}_i$ from $\mathbb{R}^3 \times \mathbb{R}^3$ into $\mathbb{R}^3$. This application is bilinear and antisymmetric and is represented by the three matrices $\boldsymbol{G}_i^n$, $n = 1, 2, 3$. It is called the *trifocal tensor* for view $i$. The properties of this application can be summarized in the following theorem:

**Theorem 5.1.** *The application* $\mathcal{T}_i : \mathbb{P}^{*2} \times \mathbb{P}^{*2} \longrightarrow \mathbb{P}^{*2}$ *is represented by the bilinear application* $\boldsymbol{\mathcal{T}}_i$ *such that* $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k) \simeq \tilde{\boldsymbol{\mathcal{P}}}_i(\boldsymbol{\mathcal{P}}_j^{\mathrm{T}}\boldsymbol{l}_j \vartriangle \boldsymbol{\mathcal{P}}_k^{\mathrm{T}}\boldsymbol{l}_k)$. $\boldsymbol{\mathcal{T}}_i$ *has the following properties:*

1. *It is equal to* $\mathbf{0}$ *iff*

   (a) $l_j$ *and* $l_k$ *are epipolar lines with respect to the* $i$*th view, or*

   (b) $l_j$ *and* $l_k$ *are corresponding epipolar lines with respect to the pair* $(j, k)$ *of cameras.*




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

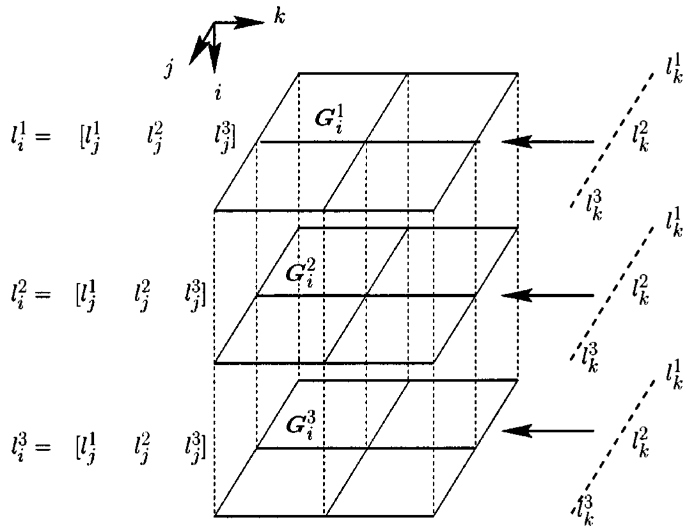


Figure 6. A three-dimensional representation of the trifocal tensor.

2. *Let $l_k$ be an epipolar line with respect to view $i$ and $l_i$ the corresponding epipolar line in view $i$, then for all lines $l_j$ in view $j$ which are not epipolar lines with respect to view $i$: $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k) \simeq \boldsymbol{l}_i$.*

3. *Similarly, let $l_j$ be an epipolar line with respect to view $i$ and $l_i$ the corresponding epipolar line in view $i$, then for all lines $l_k$ in view $k$ which are not epipolar lines with respect to view $i$: $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k) \simeq \boldsymbol{l}_i$.*

4. *If $l_j$ and $l_k$ are non-corresponding epipolar lines with respect to the pair $(j, k)$ of views, then $\boldsymbol{\mathcal{T}}(\boldsymbol{l}_j, \boldsymbol{l}_k) = \boldsymbol{t}_i$, the trifocal line of the $i$th view, if the optical centres are not aligned and $\mathbf{0}$ otherwise.*

*Proof*. We have already proved point 1. In order to prove point 2, we notice that when $l_k$ is an epipolar line with respect to view $i$, the line $L$ is contained in an epipolar plane for the pair $(i, k)$ of cameras. Two cases can happen. If $L$ goes through $C_i$, i.e. if $l_j$ is an epipolar line with respect to view $i$, then $l_i$ is reduced to a point and this is point 1(a) of the theorem. If $L$ does not go through $C_i$, $l_j$ is not an epipolar line with respect to view $i$ and the image of $L$ in view $i$ is independent of its position in the epipolar plane for the pair $(i, k)$, it is the epipolar line $l_i$ corresponding to $l_k$. The proof of point 3 is identical after exchanging the roles of cameras $k$ and $j$.

If $l_j$ and $l_k$ are non-corresponding epipolar lines for the pair $(j, k)$ of views, the two planes $(C_j, l_j)$ and $(C_k, l_k)$ intersect along the line $(C_j, C_k)$. Thus, if $C_i$ is not on that line, its image $l_i$ in view $i$ is indeed the trifocal line $t_i$; see figure 2. ∎

A more pictorial view is shown in figure 6: the tensor is represented as a $3 \times 3$ cube, the three horizontal planes representing the matrices $\boldsymbol{G}_i^n$, $n = 1, 2, 3$. It can be thought of as a black box which takes as its input two lines, $l_j$ and $l_k$ and outputs a third, $l_i$. Hartley (1994, 1997) has shown that the trifocal tensors can be very simply parametrized by the perspective projection matrices $\boldsymbol{\mathcal{P}}_n$, $n = 1, 2, 3$ of the three cameras. This result is summarized in the following proposition:

**Proposition 5.2 (Hartley).** *Let $\mathcal{P}_n$, $n = 1, 2, 3$ be the three perspective projection matrices of three cameras in general viewing position. After a change of coordinates, those matrices can be written, $\mathcal{P}_1 = [\boldsymbol{I}_3\boldsymbol{0}]$, $\mathcal{P}_2 = [\boldsymbol{X}\boldsymbol{e}_{2,1}]$ and $\mathcal{P}_3 = [\boldsymbol{Y}\boldsymbol{e}_{3,1}]$ and the matrices $\boldsymbol{G}_1^n$ can be expressed as*

$$\boldsymbol{G}_1^n = \boldsymbol{e}_{2,1}^{\mathrm{T}}\boldsymbol{Y}^{(n)} - \boldsymbol{X}^{(n)}\boldsymbol{e}_{3,1}^{\mathrm{T}}, \qquad n = 1, 2, 3, \tag{5.7}$$

*where the vectors $\boldsymbol{X}^{(n)}$ and $\boldsymbol{Y}^{(n)}$ are the column vectors of the matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively.*

We use this proposition as a definition:

**Definition 5.3.** *Any tensor of the form (5.7) is a trifocal tensor.*

### (*c*) *A third affine digression*

If the three cameras are affine, i.e. if $\boldsymbol{\Theta}_i \simeq \boldsymbol{\Theta}_j \simeq \boldsymbol{\Theta}_k \simeq \boldsymbol{e}_4$, then we can read from equation (5.6) the form of the matrices $\boldsymbol{G}_i^n$, $n = 1, 2, 3$.

**Proposition 5.4.** *For affine cameras, the trifocal tensor takes the simple form:*

$$\boldsymbol{G}_i^1 = \begin{bmatrix} |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_j\boldsymbol{\Lambda}_k| & 0 \\ |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Lambda}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_i\boldsymbol{\Lambda}_j\boldsymbol{\Lambda}_k| & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{G}_i^2 = \begin{bmatrix} |\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_i\boldsymbol{\Gamma}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_i\boldsymbol{\Gamma}_j\boldsymbol{\Lambda}_k| & 0 \\ |\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Theta}_i\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_j\boldsymbol{\Lambda}_k| & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{G}_i^3 = \begin{bmatrix} |\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\boldsymbol{\Gamma}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\boldsymbol{\Gamma}_j\boldsymbol{\Lambda}_k| & |\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\boldsymbol{\Gamma}_j\boldsymbol{\Theta}_k| \\ |\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\boldsymbol{\Lambda}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\boldsymbol{\Lambda}_j\boldsymbol{\Lambda}_k| & |\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\boldsymbol{\Lambda}_j\boldsymbol{\Theta}_k| \\ |\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_j\boldsymbol{\Gamma}_k| & |\boldsymbol{\Gamma}_i\boldsymbol{\Lambda}_i\boldsymbol{\Theta}_j\boldsymbol{\Lambda}_k| & 0 \end{bmatrix}.$$

### (*d*) *Algebraic and geometric properties of the trifocal tensors*

The matrices $\boldsymbol{G}_i^n$, $n = 1, 2, 3$ have interesting properties which are closely related to the epipolar geometry of the views $j$ and $k$. We start with the following proposition, which was proved, for example, in Hartley (1997). The proof hopefully gives some more geometric insight into what is going on:

**Proposition 5.5 (Hartley).** *The matrices $\boldsymbol{G}_i^n$ are of rank 2 and their nullspaces are the three epipolar lines, noted $l_k^n$ in view $k$ of the three projection rays of camera $i$. These three lines intersect at the epipole $e_{k,i}$. The corresponding lines in view $i$ are represented by $\boldsymbol{e}_n \times \boldsymbol{e}_{i,k}$ and can be obtained as $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k^n)$, $n = 1, 2, 3$ for any $l_j$ not equal to $l_j^n$ (see proposition 5.6).*

*Proof*. The nullspace of $\boldsymbol{G}_i^n$ is the set of lines $l_k^n$ such that $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k^n)$ has a zero in the $n$th coordinate for all lines $l_j$. The corresponding lines $l_i$ such that $\boldsymbol{l}_i = \boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j, \boldsymbol{l}_k^n)$ all go through the point represented by $\boldsymbol{e}_n$, $n = 1, 2, 3$ in the $i$th retinal plane. This is true if and only if $l_k^n$ is the image in the $k$th retinal plane of the projection ray $\boldsymbol{\Lambda}_i \triangle \boldsymbol{\Theta}_i (n = 1)$, $\boldsymbol{\Theta}_i \triangle \boldsymbol{\Gamma}_i (n = 2)$ and $\boldsymbol{\Gamma}_i \triangle \boldsymbol{\Lambda}_i (n = 3)$: $l_k^n$ is an epipolar line with respect to view $i$ and theorem 5.1, point 2, shows that for each $n$ the corresponding line in view $i$ is independent of $l_j$. Moreover, it is represented by $\boldsymbol{e}_n \times \boldsymbol{e}_{i,k}$. ■





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

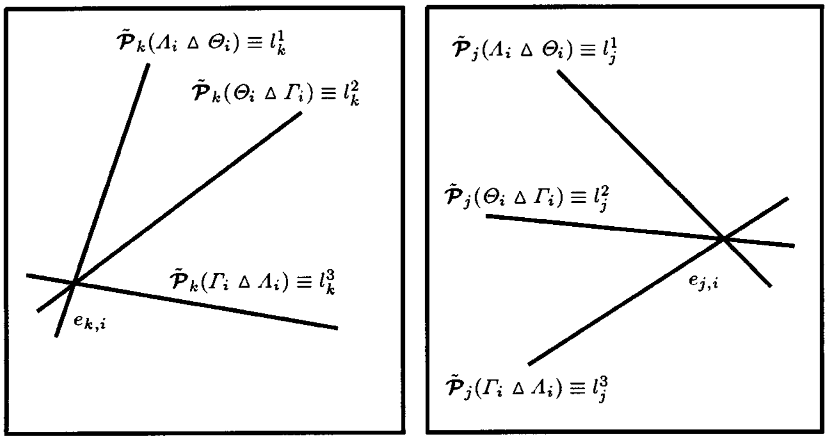


Figure 7. The lines $l_j^n$ (resp. $l_k^n$), $n = 1, 2, 3$ in the nullspaces of the matrices $\boldsymbol{G}_i^{n\mathrm{T}}$ (resp. $\boldsymbol{G}_i^n$) are the images of the three projection rays of camera $i$. Hence, they intersect at the epipole $e_{j,i}$ (resp. $e_{k,i}$). The corresponding epipolar lines in camera $i$ are obtained as $\mathcal{T}_i(l_j^n, l_k)$ (resp. $\mathcal{T}_i(l_j, l_k^n)$) for $l_k \neq l_k^n$ (resp. $l_j \neq l_j^n$).

A similar reasoning applies to the matrices $\boldsymbol{G}_i^{n\mathrm{T}}$:

**Proposition 5.6 (Hartley).** *The nullspaces of the matrices $\boldsymbol{G}_i^{n\mathrm{T}}$ are the three epipolar lines, noted $l_j^n$, $n = 1, 2, 3$, in the $j$th retinal plane of the three projection rays of camera $i$. These three lines intersect at the epipole $e_{j,i}$; see figure 7. The corresponding lines in view $i$ are represented by $\boldsymbol{e}_n \times \boldsymbol{e}_{i,j}$ and can be obtained as $\boldsymbol{\mathcal{T}}_i(\boldsymbol{l}_j^n, \boldsymbol{l}_k), n = 1, 2, 3$ for any $l_k$ not equal to $l_k^n$.*

This provides a geometric interpretation of the matrices $\boldsymbol{G}_i^n$: they represent mappings from the set of lines in view $k$ to the set of points in view $j$ located on the epipolar line $l_j^n$ defined in proposition 5.6. This mapping is geometrically defined by taking the intersection of the plane defined by the optical centre of the $k$th camera and any line of its retinal plane with the $n$th projection ray of the $i$th camera and forming the image of this point in the $j$th camera. This point does not exist when the plane contains the projection ray. The corresponding line in the $k$th retinal plane is the epipolar line $l_k^n$ defined in proposition 5.5. Moreover, the three columns of $\boldsymbol{G}_i^n$ represent three points which all belong to the epipolar line $l_j^n$.

Similarly, the matrices $\boldsymbol{G}_i^{n\mathrm{T}}$ represent mappings from the set of lines in view $j$ to the set of points in view $k$ located on the epipolar line $l_k^n$.

**Remark 5.7.** *It is important to note that the rank of the matrices $\boldsymbol{G}^n$ cannot be less than 2. Consider for example the case $n = 1$. We have seen in proposition 5.5 that the nullspace of $\boldsymbol{G}^1$ is the image of the projection ray $\boldsymbol{\Lambda}_i \triangle \boldsymbol{\Theta}_i$ in view $k$. Under our general viewpoint assumption, this projection ray and the optical centre $C_k$ define a unique plane unless it goes through $C_k$, a situation that can be avoided by a change of coordinates in the retinal plane of the $i$th camera. Therefore there is a unique line in the right nullspace of $\boldsymbol{G}_i^1$ and its rank is equal to 2. Similar reasonings apply to $\boldsymbol{G}_i^2$ and $\boldsymbol{G}_i^3$.*

A question that will turn out to be important later is that of knowing how many distinct lines $l_k^n$ (resp. $l_j^n$) can there be. This is described in the following proposition:

**Proposition 5.8.** *Under the general viewpoint assumption, the rank of the matrices* $[\boldsymbol{l}_k^1\boldsymbol{l}_k^2\boldsymbol{l}_k^3]$ *and* $[\boldsymbol{l}_j^1\boldsymbol{l}_j^2\boldsymbol{l}_j^3]$ *is 2.*

*Proof*. We know from propositions 5.5 and 5.6 that the the ranks are less than or equal to 2 because each triplet of lines intersect at an epipole. In order for the ranks to be equal to 1, we would need to have only one line in either retinal plane. But this would mean that the three planes defined by $C_k$ (resp. $C_j$) and the three projection rays of the $i$th camera are identical which is impossible since $C_i \neq C_k$ (resp. $C_j \neq C_k$) and the three projection rays of the $i$th camera are not coplanar. ∎

Algebraically, this implies that the three determinants $\det(\boldsymbol{G}_i^n)$, $n = 1, 2, 3$ are equal to 0. Another constraint implied by proposition 5.8 is that the $3 \times 3$ determinants formed with the three vectors in the nullspaces of the $\boldsymbol{G}_i^n$, $n = 1, 2, 3$ (resp. of the $\boldsymbol{G}_i^{n\mathrm{T}}$, $n = 1, 2, 3$) are equal to 0. It turns out that the applications $\boldsymbol{\mathcal{T}}_i$, $i = 1, 2, 3$ satisfy other algebraic constraints which are also important in practice.

The question of characterizing exactly the constraints satisfied by the tensors is of great practical importance for the problem of estimating the tensors from triplets of line correspondences (see Faugeras & Papadopoulo 1998). To be more specific, we know that the tensor is equivalent to the knowledge of the three perspective projection matrices and that they depend upon 18 parameters. On the other hand a trifocal tensor depends upon 27 parameters up to scale, i.e. 26 parameters. To be more precise, this means that the set of trifocal tensors is a manifold of dimension 18 in the projective space of dimension 26. There must therefore exist constraints between the coefficients that define the tensor. Our next task is to discover some of those constraints and find subsets of them which characterize the trifocal tensors, i.e. that guarantee that they have the form (5.7).

To simplify a bit the notations, we will assume in the sequel that $i = 1$, $j = 2$, $k = 3$ and will ignore the $i$th index everywhere, e.g. denote $\boldsymbol{\mathcal{T}}_1$ by $\boldsymbol{\mathcal{T}}$.

We have already seen several such constraints when we studied the matrixes $\boldsymbol{G}^n$. Let us summarize those constraints in the following proposition:

**Proposition 5.9.** *Under the general viewpoint assumption, the trifocal tensor* $\boldsymbol{\mathcal{T}}$ *satisfies the three constraints, called the rank constraints:*

$$\operatorname{rank}(\boldsymbol{G}^n) = 2 \Longrightarrow \det(\boldsymbol{G}^n) = 0, \qquad n = 1, 2, 3.$$

*The trifocal tensor* $\boldsymbol{\mathcal{T}}$ *satisfies the two constraints, called the epipolar constraints:*

$$\operatorname{rank}([\boldsymbol{l}_2^1\boldsymbol{l}_2^2\boldsymbol{l}_2^3]) = \operatorname{rank}([\boldsymbol{l}_3^1\boldsymbol{l}_3^2\boldsymbol{l}_3^3]) = 2 \Longrightarrow |\boldsymbol{l}_2^1\boldsymbol{l}_2^2\boldsymbol{l}_2^3| = |\boldsymbol{l}_3^1\boldsymbol{l}_3^2\boldsymbol{l}_3^3| = 0.$$

Those five constraints which are clearly algebraically independent since the rank constraints say nothing about the way the kernels are related constrain the form of the matrices $\boldsymbol{G}^n$.

We now show that the coefficients of $\boldsymbol{\mathcal{T}}$ satisfy nine more algebraic constraints of degree 6 which are defined as follows. Let $\boldsymbol{e}_n$, $n = 1, 2, 3$ be the canonical basis of $\mathbb{R}^3$ and let us consider the four lines $\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2}, \boldsymbol{e}_{k_3})$, $\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2}, \boldsymbol{e}_{k_3})$, $\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2}, \boldsymbol{e}_{l_3})$ and $\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2}, \boldsymbol{e}_{l_3})$ where the indexes $k_2$ and $l_2$ (resp. $k_3$ and $l_3$) are different. For example, if $k_2 = k_3 = 1$ and $l_2 = l_3 = 2$, the four lines are the images in camera 1 of the four 3D lines $\boldsymbol{\Gamma}_2 \vartriangle \boldsymbol{\Gamma}_3$, $\boldsymbol{\Lambda}_2 \vartriangle \boldsymbol{\Gamma}_3$, $\boldsymbol{\Gamma}_2 \vartriangle \boldsymbol{\Lambda}_3$ and $\boldsymbol{\Lambda}_2 \vartriangle \boldsymbol{\Lambda}_3$.

These four lines can be chosen in nine different ways satisfy an algebraic constraint which is detailed in the following theorem which is proved in Faugeras & Mourrain (1995*b*).

**Theorem 5.10.** *The trifocal tensor $\boldsymbol{\mathcal{T}}$ satisfies the 9 algebraic constraints of degree 6, called the vertical constraints:*

$$|\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2},\boldsymbol{e}_{k_3})\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2},\boldsymbol{e}_{l_3})\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2},\boldsymbol{e}_{l_3})||\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2},\boldsymbol{e}_{k_3})\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2},\boldsymbol{e}_{k_3})\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2},\boldsymbol{e}_{l_3})| \\ - |\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2},\boldsymbol{e}_{k_3})\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2},\boldsymbol{e}_{l_3})\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2},\boldsymbol{e}_{l_3})||\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2},\boldsymbol{e}_{k_3})\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2},\boldsymbol{e}_{k_3})\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2},\boldsymbol{e}_{l_3})| = 0. \quad (5.8)$$

The reader can convince himself that if he takes any general set of lines, then equation (5.8) is in general not satisfied. For instance, let $\boldsymbol{l}_i = \boldsymbol{e}_i$, $i = 1, 2, 3, 4$. It is readily verified that the left-hand side of (5.8) is equal to $-2$.

Referring to figure 6, what theorem 5.10 says is that if we take four vertical columns of the trifocal cube (shown as dashed lines in the figure) arranged in such a way that they form a prism with a square basis, then the expression (5.8) is equal to 0. This is the reason why we call these constraints the *vertical* constraints in the sequel. Representing each line as $\boldsymbol{\mathcal{T}}_{\cdot k_2 k_3}$, etc., we rewrite equation (5.8) as

$$|\boldsymbol{\mathcal{T}}_{\cdot k_2k_3}\boldsymbol{\mathcal{T}}_{\cdot k_2l_3}\boldsymbol{\mathcal{T}}_{\cdot l_2l_3}||\boldsymbol{\mathcal{T}}_{\cdot k_2k_3}\boldsymbol{\mathcal{T}}_{\cdot l_2k_3}\boldsymbol{\mathcal{T}}_{\cdot l_2l_3}| \\ - |\boldsymbol{\mathcal{T}}_{\cdot l_2k_3}\boldsymbol{\mathcal{T}}_{\cdot k_2l_3}\boldsymbol{\mathcal{T}}_{\cdot l_2l_3}||\boldsymbol{\mathcal{T}}_{\cdot k_2k_3}\boldsymbol{\mathcal{T}}_{\cdot l_2k_3}\boldsymbol{\mathcal{T}}_{\cdot k_2l_3}| = 0. \quad (5.9)$$

It turns out that the same kind of relations hold for the other two principal directions of the cube (shown as solid lines of different widths in the same figure):

**Theorem 5.11.** *The trifocal tensor $\boldsymbol{\mathcal{T}}$ satisfies also the nine algebraic constraints, called the row constraints:*

$$|\boldsymbol{\mathcal{T}}_{k_1\cdot k_3}\boldsymbol{\mathcal{T}}_{k_1\cdot l_3}\boldsymbol{\mathcal{T}}_{l_1\cdot l_3}||\boldsymbol{\mathcal{T}}_{k_1\cdot k_3}\boldsymbol{\mathcal{T}}_{l_1\cdot k_3}\boldsymbol{\mathcal{T}}_{l_1\cdot l_3}| \\ - |\boldsymbol{\mathcal{T}}_{l_1\cdot k_3}\boldsymbol{\mathcal{T}}_{k_1\cdot l_3}\boldsymbol{\mathcal{T}}_{l_1\cdot l_3}||\boldsymbol{\mathcal{T}}_{k_1\cdot k_3}\boldsymbol{\mathcal{T}}_{l_1\cdot k_3}\boldsymbol{\mathcal{T}}_{k_1\cdot l_3}| = 0, \quad (5.10)$$

*and the nine algebraic constraints, called the column constraints:*

$$|\boldsymbol{\mathcal{T}}_{k_1k_2\cdot}\boldsymbol{\mathcal{T}}_{k_1l_2\cdot}\boldsymbol{\mathcal{T}}_{l_1l_2\cdot}||\boldsymbol{\mathcal{T}}_{k_1k_2\cdot}\boldsymbol{\mathcal{T}}_{l_1k_2\cdot}\boldsymbol{\mathcal{T}}_{l_1l_2\cdot}| \\ - |\boldsymbol{\mathcal{T}}_{l_1k_2\cdot}\boldsymbol{\mathcal{T}}_{k_1l_2\cdot}\boldsymbol{\mathcal{T}}_{l_1l_2\cdot}||\boldsymbol{\mathcal{T}}_{k_1k_2\cdot}\boldsymbol{\mathcal{T}}_{l_1k_2\cdot}\boldsymbol{\mathcal{T}}_{k_1l_2\cdot}| = 0. \quad (5.11)$$

*Proof*. We do the proof for the first set of constraints which concern the columns of the matrices $\boldsymbol{G}^n$. The proof is analogous for the other set concerning the rows.

The three columns of $\boldsymbol{G}^n$ represent three points $G_k^n, k = 1, 2, 3$ of the epipolar line $l_2^n$ (see the discussion after proposition 5.6). To be concrete, let us consider the first two columns of $\boldsymbol{G}^1$ and $\boldsymbol{G}^2$, the proof is similar for the other combinations. We consider the two sets of points defined by $a_2\boldsymbol{G}_1^1 + b_2\boldsymbol{G}_2^1$ and $a_2\boldsymbol{G}_1^2 + b_2\boldsymbol{G}_2^2$. These two sets are in projective correspondence, the collineation being the identity. It is known that the line joining two corresponding points envelops a conic. It is easily shown that the determinant of the matrix defining this conic is equal to:

$$|\boldsymbol{\mathcal{T}}_{1\cdot 1}\boldsymbol{\mathcal{T}}_{1\cdot 2}\boldsymbol{\mathcal{T}}_{2\cdot 2}||\boldsymbol{\mathcal{T}}_{1\cdot 1}\boldsymbol{\mathcal{T}}_{2\cdot 1}\boldsymbol{\mathcal{T}}_{2\cdot 2}| - |\boldsymbol{\mathcal{T}}_{2\cdot 1}\boldsymbol{\mathcal{T}}_{1\cdot 2}\boldsymbol{\mathcal{T}}_{2\cdot 2}||\boldsymbol{\mathcal{T}}_{1\cdot 1}\boldsymbol{\mathcal{T}}_{2\cdot 1}\boldsymbol{\mathcal{T}}_{1\cdot 2}|.$$

In order to show that this expression is equal to 0, we show that the conic is degenerate, containing two points. This result is readily obtained from a geometric interpretation of what is going on.

The point $a_2\boldsymbol{G}_1^1 + b_2\boldsymbol{G}_2^1$ is the image by $\boldsymbol{G}^1$ of the line $a_2\boldsymbol{e}_1 + b_2\boldsymbol{e}_2$, i.e. the first set of points is the image by $\boldsymbol{G}^1$ of the pencil of lines going through the point $e_3$. Using again the geometric interpretation of $\boldsymbol{G}^1$, we realize that those points are the images in the second image of the points of intersection of the first projection ray $\boldsymbol{\Lambda} \triangle \boldsymbol{\Theta}$ of the first camera with the pencil of planes going through the third projection ray





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

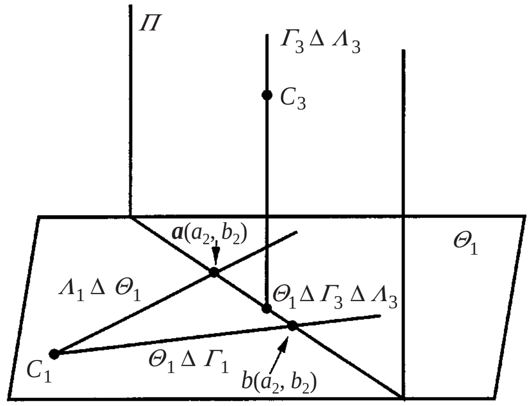


Figure 8. A plane of the pencil of axis $\boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$ intersects the plane $\boldsymbol{\Gamma}$ along a line going through the point $\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$. The points $\boldsymbol{a}(a_2, b_2)$ and $\boldsymbol{b}(a_2, b_2)$ are the images by $\boldsymbol{G}_1$ of $a_2\boldsymbol{G}_1^1 + b_2\boldsymbol{G}_2^1$ and $a_2\boldsymbol{G}_1^2 + b_2\boldsymbol{G}_2^2$, respectively.

$\boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$ of the third camera. Similarly, the second set of points is the image of the points of intersection of the second projection ray $\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma}$ of the first camera with the pencil of planes going through the third projection ray $\boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$ of the third camera.

The lines joining two corresponding points of those two sets are thus the images of the lines joining the two points of intersection of a plane containing the third projection ray $\boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$ of the third camera with the first and the second projection rays, $\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}$ and $\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma}$, of the first camera. This line lies in the third projection plane $\boldsymbol{\Theta}$ of the first camera and in the plane $\boldsymbol{\Pi}$ of the pencil. Therefore it goes through the point of intersection of the third projection plane $\boldsymbol{\Theta}$ of the first camera and the third projection ray $\boldsymbol{\Theta}' \vartriangle \boldsymbol{\Gamma}_3$ of the third camera; see figure 8. In image 2, all the lines going through two corresponding points go through the image of that point. A special case occurs when the plane $\boldsymbol{\Pi}$ goes through the first optical centre, the two points are identical to the epipole $e_{2,1}$ and the line joining them is not defined. Therefore the conic is reduced to the two points $e_{2,1}$ and the point of intersection of the two lines $(G_1^1, G_1^2)$ and $(G_2^1, G_2^2)$. This point is the image in the second camera of the point of intersection $\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$ of the first projection plane, $\boldsymbol{\Gamma}$, of the first camera with the third projection ray, $\boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$, of the third camera. ■

The theorem draws our attention to three sets of three points, i.e. three triangles, which have some very interesting properties. The triangle that came up in the proof is the one whose vertexes are the images of the points $\boldsymbol{A}_1 = \boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}_3 \vartriangle \boldsymbol{\Theta}_3$, $\boldsymbol{B}_1 = \boldsymbol{\Gamma} \vartriangle \boldsymbol{\Theta}_3 \vartriangle \boldsymbol{\Gamma}_3$ and $\boldsymbol{D}_1 = \boldsymbol{\Gamma} \vartriangle \boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$. The other two triangles are those whose vertexes are the images of the points $\boldsymbol{A}_2 = \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Lambda}_3 \vartriangle \boldsymbol{\Theta}_3$, $\boldsymbol{B}_2 = \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}_3 \vartriangle \boldsymbol{\Gamma}_3$, $\boldsymbol{D}_2 = \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$ on one hand, and $\boldsymbol{A}_3 = \boldsymbol{\Theta} \vartriangle \boldsymbol{\Lambda}_3 \vartriangle \boldsymbol{\Theta}_3$, $\boldsymbol{B}_3 = \boldsymbol{\Theta} \vartriangle \boldsymbol{\Theta}_3 \vartriangle \boldsymbol{\Gamma}_3$, $\boldsymbol{D}_3 = \boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma}_3 \vartriangle \boldsymbol{\Lambda}_3$ on the other.

Note that the three sets of vertexes $A_1$, $A_2$, $A_3$, $B_1$, $B_2$, $B_3$ and $D_1$, $D_2$, $D_3$, are aligned on the three projection rays of the third camera and therefore their images are also aligned, the three lines $l_{21} = (a_1, a_2, a_3)$, $l_{22} = (b_1, b_2, b_3)$ and $l_{23} = (d_1, d_2, d_3)$ converging to the epipole $e_{2,3}$; see figure 9. The corresponding epipolar lines in image 3 are represented by $\boldsymbol{l}_{3i} = \boldsymbol{e}_{3,2} \times \boldsymbol{e}_i$, $i = 1, 2, 3$, respectively. Note that all points $a_i$, $b_i$, $d_i$, $i = 1, 2, 3$ can be expressed as simple functions of the columns of the matrixes

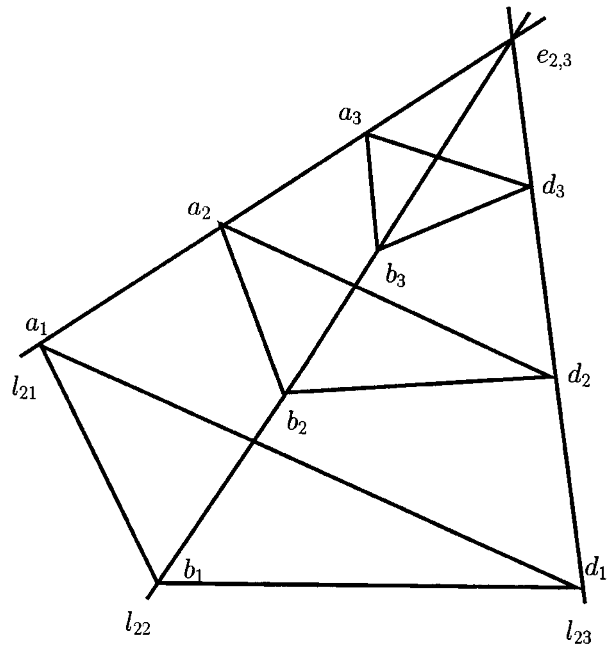


Figure 9. The three triangles have corresponding vertexes aligned on epipolar lines for the pair $(2, 3)$ of images.

$\boldsymbol{G}^n$. For example,

$$\boldsymbol{a}_1 = (\boldsymbol{G}_1^1 \times \boldsymbol{G}_1^2) \times (\boldsymbol{G}_2^1 \times \boldsymbol{G}_2^2).$$

The same is true of the constraints on the rows of the matrices $\boldsymbol{G}^n$. More specifically the constraints (5.11) introduce nine other points $(H_i, K_i, M_i)$, $i = 1, 2, 3$ with $\boldsymbol{H}_1 = \boldsymbol{\Gamma} \triangle \boldsymbol{\Lambda}' \triangle \boldsymbol{\Theta}'$, $\boldsymbol{K}_1 = \boldsymbol{\Gamma} \triangle \boldsymbol{\Theta}' \triangle \boldsymbol{\Gamma}'$, $\boldsymbol{M}_1 = \boldsymbol{\Gamma} \triangle \boldsymbol{\Gamma}' \triangle \boldsymbol{\Lambda}'$, $\boldsymbol{H}_2 = \boldsymbol{\Lambda} \triangle \boldsymbol{\Lambda}' \triangle \boldsymbol{\Theta}'$, $\boldsymbol{K}_2 = \boldsymbol{\Lambda} \triangle \boldsymbol{\Theta}' \triangle \boldsymbol{\Gamma}'$, $\boldsymbol{M}_2 = \boldsymbol{\Lambda} \triangle \boldsymbol{\Gamma}' \triangle \boldsymbol{\Lambda}'$, and $\boldsymbol{H}_3 = \boldsymbol{\Theta} \triangle \boldsymbol{\Lambda}' \triangle \boldsymbol{\Theta}'$, $\boldsymbol{K}_3 = \boldsymbol{\Theta} \triangle \boldsymbol{\Theta}' \triangle \boldsymbol{\Gamma}'$, $\boldsymbol{M}_3 = \boldsymbol{\Theta} \triangle \boldsymbol{\Gamma}' \triangle \boldsymbol{\Lambda}'$. The three sets of points $H_1$, $H_2$, $H_3$, $K_1$, $K_2$, $K_3$ and $M_1$, $M_2$, $M_3$ are aligned on the three projection rays of the second camera and therefore their images are also aligned, the three lines $l'_{31} = (h_1, h_2, h_3)$, $l'_{32} = (k_1, k_2, k_3)$ and $l'_{33} = (m_1, m_2, m_3)$ converging to the epipole $e_{3,2}$. The corresponding epipolar lines, $l'_{2i}$, $i = 1, 2, 3$ in image 3 are represented by $\boldsymbol{e}_{2,3} \times \boldsymbol{e}_i$, $i = 1, 2, 3$, respectively.

Note that this yields a way of recovering the fundamental matrix $\boldsymbol{F}_{23}$, since we obtain the two epipoles $e_{2,3}$ and $e_{3,2}$ and three pairs of corresponding epipolar lines, in fact six pairs. We will not address further here the problem of recovering the epipolar geometry of the three views, let us simply mention the fact that the fundamental matrices which are recovered from the trifocal tensor are *compatible* in the sense that they satisfy the constraints (5.1). There is a further set of constraints that are satisfied by any trifocal tensor and are also of interest. They are described in the next proposition:

**Proposition 5.12.** *The trifocal tensor $\boldsymbol{\mathcal{T}}$ satisfies the 10 algebraic constraints, called the extended rank constraints:*

$$\operatorname{rank}\left(\sum_{n=1}^{3} \lambda_n \boldsymbol{G}^n\right) \leqslant 2 \qquad \forall \lambda_n, \quad n = 1, 2, 3.$$

*Proof.* The proof can be done either algebraically or geometrically. The algebraic proof simply uses the parametrization (5.7) and verifies that the constraints described in proposition 5.13 are satisfied. In the geometric proof one notices that for fixed values (not all zero) of the $\lambda_n$s, and for a given line $l_3$ in view 3, the point which is the image in view 2 of line $l_3$ by $\sum_{n=1}^{3}\lambda_n \boldsymbol{G}^n$ is the image of the point defined by

$$\lambda_1 \boldsymbol{\mathcal{P}}_3^{\mathrm{T}} l_3 \vartriangle (\boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta}) + \lambda_2 \boldsymbol{\mathcal{P}}_3^{\mathrm{T}} l_3 \vartriangle (\boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma}) + \lambda_3 \boldsymbol{\mathcal{P}}_3^{\mathrm{T}} l_3 \vartriangle (\boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}).$$

This expression can be rewritten as

$$\boldsymbol{\mathcal{P}}_3^{\mathrm{T}} l_3 \vartriangle (\lambda_1 \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} + \lambda_2 \boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} + \lambda_3 \boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}). \tag{5.12}$$

The line $\lambda_1 \boldsymbol{\Lambda} \vartriangle \boldsymbol{\Theta} + \lambda_2 \boldsymbol{\Theta} \vartriangle \boldsymbol{\Gamma} + \lambda_3 \boldsymbol{\Gamma} \vartriangle \boldsymbol{\Lambda}$ is an optical ray of the first camera (proposition 3.3), and when $l_3$ varies in view 3, the point defined by (5.12) is well defined except when $l_3$ is the image of that line in view 3. In that case the meet in (5.12) is zero and the image of that line is in the nullspace of $\sum_{n=1}^{3}\lambda_n \boldsymbol{G}^n$. ∎

Note that the proposition 5.12 is equivalent to the vanishing of the 10 coefficients of the homogeneous polynomial of degree 3 in the three variables $\lambda_n$, $n = 1, 2, 3$ equal to $\det(\sum_{n=1}^{3}\lambda_n \boldsymbol{G}^n)$. The coefficients of the terms $\lambda_n^3$, $n = 1, 2, 3$ are the determinants $\det(\boldsymbol{G}^n)$, $n = 1, 2, 3$. Therefore the extended rank constraints contain the rank constraints.

To be complete, we give the expressions of the seven extended rank constraints which are different from the three rank constraints:

**Proposition 5.13.** *The seven extended rank constraints are given by*

$$\lambda_1^2\lambda_2 \quad |\boldsymbol{G}_1^1\boldsymbol{G}_2^1\boldsymbol{G}_3^2| + |\boldsymbol{G}_1^1\boldsymbol{G}_2^2\boldsymbol{G}_3^1| + |\boldsymbol{G}_1^2\boldsymbol{G}_2^1\boldsymbol{G}_3^1| = 0, \tag{5.13}$$

$$\lambda_1^2\lambda_3 \quad |\boldsymbol{G}_1^1\boldsymbol{G}_2^1\boldsymbol{G}_3^3| + |\boldsymbol{G}_1^1\boldsymbol{G}_2^3\boldsymbol{G}_3^1| + |\boldsymbol{G}_1^3\boldsymbol{G}_2^1\boldsymbol{G}_3^1| = 0, \tag{5.14}$$

$$\lambda_2^2\lambda_1 \quad |\boldsymbol{G}_1^2\boldsymbol{G}_2^2\boldsymbol{G}_3^1| + |\boldsymbol{G}_1^2\boldsymbol{G}_2^1\boldsymbol{G}_3^2| + |\boldsymbol{G}_1^1\boldsymbol{G}_2^2\boldsymbol{G}_3^2| = 0, \tag{5.15}$$

$$\lambda_2^2\lambda_3 \quad |\boldsymbol{G}_1^2\boldsymbol{G}_2^2\boldsymbol{G}_3^3| + |\boldsymbol{G}_1^2\boldsymbol{G}_2^3\boldsymbol{G}_3^2| + |\boldsymbol{G}_1^3\boldsymbol{G}_2^2\boldsymbol{G}_3^2| = 0, \tag{5.16}$$

$$\lambda_3^2\lambda_1 \quad |\boldsymbol{G}_1^3\boldsymbol{G}_2^3\boldsymbol{G}_3^1| + |\boldsymbol{G}_1^3\boldsymbol{G}_2^1\boldsymbol{G}_3^3| + |\boldsymbol{G}_1^1\boldsymbol{G}_2^3\boldsymbol{G}_3^3| = 0, \tag{5.17}$$

$$\lambda_3^2\lambda_2 \quad |\boldsymbol{G}_1^3\boldsymbol{G}_2^3\boldsymbol{G}_3^2| + |\boldsymbol{G}_1^3\boldsymbol{G}_2^2\boldsymbol{G}_3^3| + |\boldsymbol{G}_1^2\boldsymbol{G}_2^3\boldsymbol{G}_3^3| = 0, \tag{5.18}$$

$$\begin{aligned}\lambda_1\lambda_2\lambda_3 \quad & |\boldsymbol{G}_1^1\boldsymbol{G}_2^2\boldsymbol{G}_3^3| + |\boldsymbol{G}_1^1\boldsymbol{G}_2^3\boldsymbol{G}_3^2| + |\boldsymbol{G}_1^2\boldsymbol{G}_2^1\boldsymbol{G}_3^3| \\ & + |\boldsymbol{G}_1^2\boldsymbol{G}_2^3\boldsymbol{G}_3^1| + |\boldsymbol{G}_1^3\boldsymbol{G}_2^1\boldsymbol{G}_3^2| + |\boldsymbol{G}_1^3\boldsymbol{G}_2^2\boldsymbol{G}_3^1| = 0.\end{aligned} \tag{5.19}$$

### (*e*) *Constraints that characterize the tensor*

We now show two results which are related to the question of finding subsets of constraints which are sufficient to characterize the trifocal tensors. These subsets are the implicit equations of the manifold of the trifocal tensors. The first result is given in the following theorem:

**Theorem 5.14.** *Let $\boldsymbol{\mathcal{T}}$ be a bilinear mapping from $\mathbb{P}^{*2} \times \mathbb{P}^{*2}$ to $\mathbb{P}^{*2}$ which satisfies the 14 rank, epipolar and vertical constraints. Then this mapping is a trifocal tensor, i.e. it satisfies definition 5.3. Those 14 algebraic equations are a set of implicit equations of the manifold of trifocal tensors.*

The second result is that the 10 extended constraints and the epipolar constraints characterize the trifocal tensors:

**Theorem 5.15.** *Let $\mathcal{T}$ be a bilinear mapping from $\mathbb{P}^{*2} \times \mathbb{P}^{*2}$ to $\mathbb{P}^{*2}$ which satisfies the 12 extended rank and epipolar constraints. Then this mapping is a trifocal tensor, i.e. it satisfies definition 5.3. Those 12 algebraic equations are another set of implicit equations of the manifold of trifocal tensors.*

The proof of those theorems will take us some time. We start with a proposition we will use to prove that the three rank constraints and the two epipolar constraints are not sufficient to characterize the set of trifocal tensors:

**Proposition 5.16.** *If a tensor $\mathcal{T}$ satisfies the three rank constraints and the two epipolar constraints, then its matrices $\boldsymbol{G}^n$, $n = 1, 2, 3$ can be written:*

$$\boldsymbol{G}^n = a_n \boldsymbol{X}^{(n)} \boldsymbol{Y}^{(n)\mathrm{T}} + \boldsymbol{X}^{(n)} \boldsymbol{e}_{3,1}^{\mathrm{T}} + \boldsymbol{e}_{2,1} \boldsymbol{Y}^{(n)\mathrm{T}}, \tag{5.20}$$

*where $e_{2,1}$ (resp. $e_{3,1}$) is a fixed point of image 2 (resp. of image 3), the three vectors $\boldsymbol{X}^{(n)}$ represent three points of image 2, and the three vectors $\boldsymbol{Y}^{(n)}$ represent three points of image 3.*

*Proof*. The rank constraints allow us to write:

$$\boldsymbol{G}^n = \boldsymbol{X}_1^{(n)} \boldsymbol{Y}_1^{(n)\mathrm{T}} + \boldsymbol{X}_2^{(n)} \boldsymbol{Y}_2^{(n)\mathrm{T}}, \tag{5.21}$$

where the six vectors $\boldsymbol{X}_1^{(n)}$, $\boldsymbol{X}_2^{(n)}$, $n = 1, 2, 3$ represent six points of the second image and the six vectors $\boldsymbol{Y}_1^{(n)}$, $\boldsymbol{Y}_2^{(n)}$, $n = 1, 2, 3$ represent six points of the third image. The right nullspace of $\boldsymbol{G}^n$ is simply the cross-product $\boldsymbol{X}_1^{(n)} \times \boldsymbol{X}_2^{(n)}$, the left nullspace being $\boldsymbol{Y}_1^{(n)} \times \boldsymbol{Y}_2^{(n)}$. Those two sets of three nullspaces are of rank 2 (proposition 5.9). Let us consider the first set. We can write the corresponding matrix as

$$[\boldsymbol{X}_1^{(1)} \times \boldsymbol{X}_2^{(1)} \boldsymbol{X}_1^{(2)} \times \boldsymbol{X}_2^{(2)} \boldsymbol{X}_1^{(3)} \times \boldsymbol{X}_2^{(3)}] = \boldsymbol{Z}_1 \boldsymbol{T}_1^{\mathrm{T}} + \boldsymbol{Z}_2 \boldsymbol{T}_2^{\mathrm{T}}.$$

With obvious notations, we have in particular

$$\boldsymbol{X}_1^{(1)} \times \boldsymbol{X}_2^{(1)} = T_{11} \boldsymbol{Z}_1 + T_{21} \boldsymbol{Z}_2.$$

Let us now interpret this equation geometrically: the line represented by the vector $\boldsymbol{X}_1^{(1)} \times \boldsymbol{X}_2^{(1)}$, i.e. the line going through the points $X_1^{(1)}$ and $X_2^{(1)}$ belongs to the pencil of lines defined by the two lines represented by the vectors $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$. Therefore it goes through their point of intersection represented by the cross-product $\boldsymbol{Z}_1 \times \boldsymbol{Z}_2$ and we write $\boldsymbol{X}_2^{(1)}$ as a linear combination of $\boldsymbol{X}_1^{(1)}$ and $\boldsymbol{Z}_1 \times \boldsymbol{Z}_2$:

$$\boldsymbol{X}_2^{(1)} = \alpha_1 \boldsymbol{X}_1^{(1)} + \beta_1 \boldsymbol{Z}_1 \times \boldsymbol{Z}_2.$$

We write $\boldsymbol{e}_{2,1}$ for $\boldsymbol{Z}_1 \times \boldsymbol{Z}_2$ and note that our reasoning is valid for $X_1^{(n)}$ and $X_2^{(n)}$:

$$\boldsymbol{X}_2^{(n)} = \alpha_n \boldsymbol{X}_1^{(n)} + \beta_n \boldsymbol{e}_{2,1}, \qquad n = 1, 2, 3.$$

The same exact reasoning can be applied to the pairs $\boldsymbol{Y}_1^{(n)}$, $\boldsymbol{Y}_2^{(n)}$, $n = 1, 2, 3$ yielding the expression:

$$\boldsymbol{Y}_1^{(n)} = \gamma_n \boldsymbol{Y}_2^{(n)} + \delta_n \boldsymbol{e}_{3,1}.$$

We have exchanged the roles of $\boldsymbol{Y}_1^{(n)}$ and $\boldsymbol{Y}_2^{(n)}$ for reasons of symmetry in the final expression of $\boldsymbol{G}_n$. Replacing $\boldsymbol{X}_2^{(n)}$ and $\boldsymbol{Y}_1^{(n)}$ by their values in the definition (5.21) of the matrix $\boldsymbol{G}_n$, we obtain

$$\boldsymbol{G}^n = (\alpha_n + \gamma_n) \boldsymbol{X}_1^{(n)} \boldsymbol{Y}_2^{(n)\mathrm{T}} + \delta_n \boldsymbol{X}_1^{(n)} \boldsymbol{e}_{3,1}^{\mathrm{T}} + \beta_n \boldsymbol{e}_{2,1} \boldsymbol{Y}_2^{(n)\mathrm{T}}.$$

We can absorb the coefficients $\delta_n$ in $\boldsymbol{X}_1^{(n)}$, the coefficients $\beta_n$ in $\boldsymbol{Y}_2^{(n)}$ and we obtain the announced relation. ∎





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

The next proposition is a proof of theorem 5.14 that the 14 rank and epipolar constraints characterize the set of trifocal tensors:

**Proposition 5.17.** *Let $\boldsymbol{\mathcal{T}}$ be a bilinear mapping from $\mathbb{P}^{*2} \times \mathbb{P}^{*2}$ to $\mathbb{P}^{*2}$ which satisfies the 14 rank, epipolar and vertical constraints. Then its matrices $\boldsymbol{G}^n$ take the form:*

$$\boldsymbol{G}^n = \boldsymbol{e}_{2,1}\boldsymbol{Y}^{(n)\mathrm{T}} + \boldsymbol{X}^{(n)}\boldsymbol{e}_{3,1}^{\mathrm{T}}. \tag{5.22}$$

*Proof*. In order to show this, we show that the nine vertical constraints imply that $\boldsymbol{\mathcal{T}}(\boldsymbol{l}_{21}, \boldsymbol{l}_{31}) = \boldsymbol{0}$ for all pair of epipolar lines $(l_{21}, l_{31})$, i.e. for all pairs of lines such that $l_{21}$ contains the point $e_{2,1}$ and $l_{31}$ contains the point $e_{3,1}$ defined in (5.8). Indeed, this implies that $a_n(\boldsymbol{l}_{21}^{\mathrm{T}}\boldsymbol{X}^{(n)}) \cdot (\boldsymbol{Y}^{(n)\mathrm{T}}\boldsymbol{l}_{31}) = 0$ for all pairs of epipolar lines $(l_{21}, l_{31})$ which implies $a_n = 0$ unless either $X^{(n)}$ is identical to $e_{2,1}$ or $Y^{(n)}$ is identical to $e_{3,1}$ which contradicts the hypothesis that the rank of $\boldsymbol{G}^n$ is two.

In order to show this it is sufficient to show that each of the nine constraints implies that $\boldsymbol{\mathcal{T}}(\boldsymbol{l}_{21i}, \boldsymbol{l}_{31j}) = \boldsymbol{0}$, $i, j = 1, 2, 3$ where $l_{21i}$ (resp. $l_{31j}$) is an epipolar line for the pair $(1, 2)$ (resp. the pair $(1, 3)$) of cameras, going the $i$th (resp. the $j$th) point of the canonical basis. This is sufficient because we can assume that, for example, $e_{2,1}$ does not belong to the line represented by $\boldsymbol{e}_3$. In that case, any epipolar line $l_{21}$ can be represented as a linear combination of $\boldsymbol{l}_{211}$ and $\boldsymbol{l}_{212}$:

$$\boldsymbol{l}_{21} = \alpha_2\boldsymbol{l}_{211} + \beta_2\boldsymbol{l}_{212}.$$

Similarly, any epipolar line $l_{31}$ can be represented as a linear combination of $\boldsymbol{l}_{311}$ and $\boldsymbol{l}_{312}$, given that $e_{3,1}$ does not belong to the line represented by $\boldsymbol{e}_3$:

$$\boldsymbol{l}_{31} = \alpha_3\boldsymbol{l}_{311} + \beta_3\boldsymbol{l}_{312}.$$

The bilinearity of $\boldsymbol{\mathcal{T}}$ allows us to conclude that $\boldsymbol{\mathcal{T}}(\boldsymbol{l}_{21}, \boldsymbol{l}_{31}) = \boldsymbol{0}$.

To simplify the notations we define

$$\boldsymbol{\lambda}_1 = \boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2}, \boldsymbol{e}_{k_3}), \qquad \boldsymbol{\lambda}_2 = \boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2}, \boldsymbol{e}_{k_3}),$$
$$\boldsymbol{\lambda}_3 = \boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2}, \boldsymbol{e}_{l_3}), \qquad \boldsymbol{\lambda}_4 = \boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2}, \boldsymbol{e}_{l_3}).$$

To help the reader follow the proof, we encourage him or her to take the example $k_2 = k_3 = 1$ and $l_2 = l_3 = 2$. If the tensor $\boldsymbol{\mathcal{T}}$ were a trifocal tensor, the four lines $\lambda_1$, $\lambda_2$, $\lambda_3$, $\lambda_4$ would be the images of the 3D lines $\boldsymbol{\Gamma}' \triangle \boldsymbol{\Gamma}_3$, $\boldsymbol{\Lambda}' \triangle \boldsymbol{\Gamma}_3$, $\boldsymbol{\Gamma}' \triangle \boldsymbol{\Lambda}_3$, $\boldsymbol{\Lambda}' \triangle \boldsymbol{\Lambda}_3$, respectively.

We now consider the two lines $d_1$ and $d_2$ in image 1 which are defined as follows. $d_1$ goes through the point of intersection of $\lambda_1$ and $\lambda_2$ (the image of the point $\boldsymbol{\Gamma}' \triangle \boldsymbol{\Lambda}' \triangle \boldsymbol{\Gamma}_3$) and the point of intersection of the lines $\lambda_3$ and $\lambda_4$ (the image of the point $\boldsymbol{\Gamma}' \triangle \boldsymbol{\Lambda}' \triangle \boldsymbol{\Lambda}_3$). In our example, $d_1$ is the image of the projection ray $\boldsymbol{\Gamma}' \triangle \boldsymbol{\Lambda}'$. $d_2$, on the other hand, goes through the point of intersection of $\lambda_2$ and $\lambda_4$ (the image of the point $\boldsymbol{\Lambda}' \triangle \boldsymbol{\Gamma}_3 \triangle \boldsymbol{\Lambda}_3$) and the point of intersection of $\lambda_1$ and $\lambda_3$ (the image of the point $\boldsymbol{\Gamma}' \triangle \boldsymbol{\Gamma}_3 \triangle \boldsymbol{\Lambda}_3$). In our example, $d_2$ is the image of the projection ray $\boldsymbol{\Gamma}_3 \triangle \boldsymbol{\Lambda}_3$. Using elementary geometry, it is easy to find

$$\boldsymbol{d}_1 = |\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{\lambda}_1 - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{\lambda}_2,$$
$$\boldsymbol{d}_2 = |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{\lambda}_2 + |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{\lambda}_4.$$

According to the definition of the lines $\lambda_1$, $\lambda_2$, $\lambda_3$, $\lambda_4$, $d_1$ is the image by $\boldsymbol{\mathcal{T}}$ of the two lines $(|\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_2} - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{l_2}, \boldsymbol{e}_{k_3})$ and $d_2$ the image by $\boldsymbol{\mathcal{T}}$ of the two lines $(\boldsymbol{e}_{l_2}, |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_3} + |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{e}_{l_3})$.





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

We now proceed to show that

$$\boldsymbol{\mathcal{T}}(|\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_2} - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{l_2}, |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_3} + |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{e}_{l_3}) = \mathbf{0}.$$

Using the bilinearity of $\boldsymbol{\mathcal{T}}$, we have

$$\begin{aligned}\boldsymbol{\mathcal{T}}(&|\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_2} - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{l_2}, |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_3} + |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{e}_{l_3}) \\ &= |\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4||\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2}, \boldsymbol{e}_{k_3}) + |\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4||\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{k_2}, \boldsymbol{e}_{l_3}) \\ &\quad - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4||\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2}, \boldsymbol{e}_{k_3}) - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4||\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{\mathcal{T}}(\boldsymbol{e}_{l_2}, \boldsymbol{e}_{l_3}).\end{aligned}$$

We now use the constraint (5.8):

$$|\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4||\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_4| - |\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4||\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|,$$

to replace the coefficient of the second term by $|\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4||\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_4|$. The coefficient $|\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|$ is a factor and we have

$$\begin{aligned}\boldsymbol{\mathcal{T}}(&|\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_2} - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{l_2}, |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_3} + |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{e}_{l_3}) \\ &= |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|(|\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{\lambda}_1 - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{\lambda}_2 + |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_4|\boldsymbol{\lambda}_3 - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{\lambda}_4).\end{aligned}$$

The second factor is seen to be equal to 0 because of Cramer's relation.

We therefore have two sets of three lines, one in image 2 noted $l_{21i}, i = 1, 2, 3$, one in image 3 noted $l_{31j}$, $j = 1, 2, 3$, corresponding to the choices of $k_2$, $l_2$, $k_3$, $l_3$ and such that $\boldsymbol{\mathcal{T}}(\boldsymbol{l}_{21i}, \mathbf{1}_{31j}) = \mathbf{0}$, $i, j = 1, 2, 3$. For example, one of the lines $l_{21i}$ is represented by $|\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_2} - |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{l_2}$ and one of the lines $l_{31j}$ is represented by $|\boldsymbol{\lambda}_1\boldsymbol{\lambda}_3\boldsymbol{\lambda}_4|\boldsymbol{e}_{k_3} + |\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2\boldsymbol{\lambda}_3|\boldsymbol{e}_{l_3}$.

Let us see what this means in terms of the linear applications defined by the matrices $\boldsymbol{G}^n$. Consider the first line in image 3, $l_{311}$ its image by $\boldsymbol{G}^n$, $n = 1, 2, 3$ is a point on the line $l_2^n$, $n = 1, 2, 3$. According to what we have just proved, those three points are also on the three lines $l_{21i}$, $i = 1, 2, 3$; see figure 10. This is only possible if (*a*) the three lines $l_{21i}$, $i = 1, 2, 3$ are identical, which they are not in general, or if (*b*) the three points are identical and the three lines go through that point. The second possibility is the correct one and implies that (*a*) the three points are identical with the point of intersection, $e_{2,1}$, of the three lines $l_2^n$, $n = 1, 2, 3$ and (*b*) that the three lines $l_{21i}$, $i = 1, 2, 3$ go through $e_{2,1}$. A similar reasoning shows that the three lines $l_{31j}$, $j = 1, 2, 3$ go through the epipole $e_{3,1}$.

This completes the proof of the proposition and of theorem 5.14. ■

An intriguing question is whether there are other sets of constraints that imply this parametrization, or in other words does there exist simpler implicit parametrization of the manifold of trifocal tensors? One answer is contained in theorem 5.15. Before we prove it we prove two interesting results, the first one is unrelated, the second is

**Proposition 5.18.** *Any bilinear mapping* $\boldsymbol{\mathcal{T}}$ *which satisfies the 14 rank, epipolar and vertical constraints also satisfies the 18 row and columns constraints.*

*Proof*. The proof consists in noticing that if $\boldsymbol{\mathcal{T}}$ satisfies the rank, epipolar and vertical constraints, according to proposition 5.17, it satisfies definition 5.3 and therefore, according to theorem 5.11, it satisfies the row and column constraints. ■

The reader may wonder about the 10 extended rank constraints. Are they sufficient to characterize the trilinear tensor? The following proposition answers this question negatively.

**Proposition 5.19.** *The 10 extended rank constraints do not imply the epipolar constraints.*

*Proof*. The proof consists in exhibiting a counterexample. The reader can verify that the tensor $\boldsymbol{\mathcal{T}}$ defined by

$$\boldsymbol{G}^1 = \begin{bmatrix} 0 & 0 & 0 \\ -1 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{G}^2 = \begin{bmatrix} -1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \boldsymbol{G}^3 = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

satisfies the 10 extended rank constraints and that the corresponding three left nullspaces are the canonic lines represented by $\boldsymbol{e}_n$, $n = 1, 2, 3$ which do not satisfy one of the epipolar constraints. ∎

Before we prove theorem 5.15 we prove the following proposition:

**Proposition 5.20.** *The three rank constraints and the two epipolar constraints do not characterize the set of trifocal tensors.*

*Proof*. Indeed, proposition 5.16 gives us a parametrization of the matrices $\boldsymbol{G}^n$ in that case. It can be verified that for such a paramerization, the vertical constraints are not satisfied. Assume now that the rank and epipolar constraints imply that the tensor is a trifocal tensor, then, according to proposition 5.10, it satisfies the vertical constraints, a contradiction. ∎

We are now ready to prove theorem 5.15:

*Proof*. The proof consists in showing that any bilinear application $\boldsymbol{\mathcal{T}}$ that satisfies the five rank and epipolar constraints, i.e. whose matrices $\boldsymbol{G}^n$ can be written as in (5.20) and the remaining seven extended rank constraints (5.13)–(5.19) can be written as in (5.22), i.e. is such that $a_n = 0$, $n = 1, 2, 3$.

If we use the parametrization (5.20) and evaluate the constraints (5.13)–(5.18), we find

$$-a_2|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}|, \tag{5.23}$$

$$-a_3|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(3)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(3)}|, \tag{5.24}$$

$$-a_1|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}|, \tag{5.25}$$

$$-a_3|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}|, \tag{5.26}$$

$$-a_1|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(3)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(3)}|, \tag{5.27}$$

$$-a_2|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}|. \tag{5.28}$$

In those formulas, our attention is drawn to determinants of the form $|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(i)}\boldsymbol{X}^{(j)}|$, $i \neq j$ (type 2) and $|\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(i)}\boldsymbol{Y}^{(j)}|$, $i \neq j$ (type 3). The nullity of a determinant of the first type implies that the epipole $e_{2,1}$ (resp. $e_{3,1}$) is on the line defined by the two points $X^{(i)}$, $X^{(j)}$ (resp. $Y^{(i)}$, $Y^{(j)}$), if the corresponding points are distinct.

If all determinants are non-zero, the constraints (5.23)–(5.28) imply that all $a_n$s are zero. Things are slightly more complicated if some of the determinants are equal to 0.

We prove that if the matrices $\boldsymbol{G}^n$ are of rank 2, no more than one of the three determinants of each of the two types can equal 0. We consider several cases. The

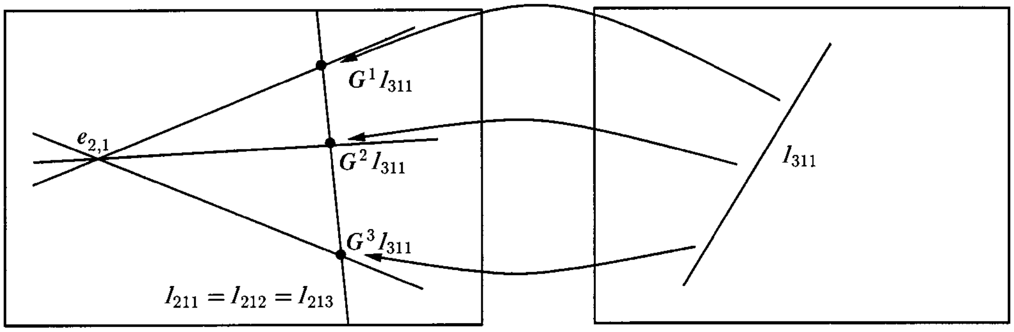


Figure 10. The three lines $l_{21i}$, $i = 1, 2, 3$ are identical, see the proof of proposition 5.17.

first case is when all points of one type are different. Suppose first that the three points represented by the three vectors $\boldsymbol{X}^{(n)}$ are not aligned. Then, having two of the determinants of type 2 equal to 0 implies that the point $e_{2,1}$ is identical to one of the points $X^{(n)}$ since it is at the intersection of two of the lines they define. But, according to equation (5.20), this implies that the corresponding matrix $\boldsymbol{G}^n$ is of rank 1, contradicting the hypothesis that this rank is 2. Similarly, if the three points $X^{(n)}$ are aligned, if one determinant is equal to 0, the epipole $e_{2,1}$ belongs to the line $(X^{(1)}, X^{(2)}, X^{(3)})$ which means that the three epipolar lines $l_2^1, l_2^2, l_2^3$ are identical contradicting the hypothesis that they form a matrix of rank 2. Therefore, in this case, all three determinants are non-null.

The second case is when two of the points are equal, e.g. $\boldsymbol{X}^{(1)} \simeq \boldsymbol{X}^{(2)}$. The third point must then be different, otherwise we would only have one epipolar line contradicting the rank 2 assumption on those epipolar lines, and, if it is different, the epipole $e_{2,1}$ must not be on the line defined by the two points for the same reason. Therefore in this case also at most one of the determinants is equal to 0.

Having at most one determinant of type 2 and one of type 3 equal to 0 implies that at least two of the $a_n$ are 0. This is seen by inspecting the constraints (5.23)–(5.28). If we now express the seventh constraint:

$$
\begin{aligned}
&a_1a_2a_3|\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}||\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}| \\
&\quad - (|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(3)}| + |\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}||\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(3)}|)a_1 \\
&\quad + (|\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}||\boldsymbol{e}_{2,1}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}| + |\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}||\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}|)a_2 \\
&\quad - (|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(3)}| + |\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}||\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(3)}|)a_3 \\
&\quad + (|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}||\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}| + |\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}||\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}|)a_1a_2 \\
&\quad + (|\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}||\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}| + |\boldsymbol{e}_{2,1}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}||\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}|)a_2a_3 \\
&\quad - (|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(3)}||\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}\boldsymbol{Y}^{(3)}| + |\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}\boldsymbol{X}^{(3)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(3)}|)a_1a_3,
\end{aligned}
$$

we find that it is equal to the third $a_n$ multiplied by two of the non-zero determinants, implying that the third $a_n$ is null and completing the proof.

Let us give a few examples of the various cases. Let us assume first that

$$|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}| = |\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}| = 0.$$

We find that the constraints (5.27), (5.28) and (5.26) imply $a_1 = a_2 = a_3 = 0$. The second situation occurs if we assume, for example, $|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}| = |\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(3)}| =$

0. We find that the constraints (5.28) and (5.26) imply $a_2 = a_3 = 0$. The constraint (5.19) takes then the form:

$$-|\boldsymbol{e}_{2,1}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(3)}||\boldsymbol{e}_{3,1}\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}|a_1,$$

and implies $a_1 = 0$. ■

Note that from a practical standpoint, theorem 5.15 provides a simple set of sufficient constraints than theorem 5.14. The 10 extended constraints are of degree 3 in the elements of $\boldsymbol{\mathcal{T}}$ whereas the nine vertical constraints are of degree 6 as are the two epipolar constraints.

This situation is more or less similar to the one with the $E$-matrix (Longuet-Higgins 1981). It has been shown in several places (e.g. Faugeras 1993, propositions 7.2 and 7.3) that the set of real $E$-matrices is characterized either by the two equations:

$$\det(\boldsymbol{E}) = 0, \qquad \tfrac{1}{2}\,\mathrm{Tr}^2(\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}) - \mathrm{Tr}((\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}})^2) = 0,$$

or by the nine equations:

$$\tfrac{1}{2}\,\mathrm{Tr}(\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}})\boldsymbol{E} - \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{E} = \boldsymbol{0}.$$

In a somewhat analogous way, the set of trifocal tensors is characterized either by the 14 rank, epipolar and vertical constraints (theorem 5.14) or by the 12 extended rank and epipolar constraints (theorem 5.15).

## 6. Conclusion

We have shown a variety of applications of the Grassmann–Cayley or double algebra to the problem of modelling systems of up to three pinhole cameras. We have analysed in detail the algebraic constraints satisfied by the trilinear tensors which characterize the geometry of three views. In particular, we have isolated two subsets of those constraints that are sufficient to guarantee that a tensor that satisfies them arises from the geometry of three cameras. Each of those subsets is a set of implicit equations for the manifold of trifocal tensors. We have shown elsewhere (Faugeras & Papadopoulo 1998) how to use some of those equations to parametrize the tensors and estimate them from line correspondences in three views.

This work was partly supported by the EEC under the reactive LTR project 21914-CUMULI.

## References

Avidan, S. & Shashua, A. 1996 Tensorial transfer: representation of $n > 3$ views of 3D scenes. In *Proc. ARPA Image Understanding Workshop.*

Barnabei, M., Brini, A. & Rota, G.-C. 1985 On the exterior calculus of invariant theory. *J. Algebra* **96**, 120–160.

Carlsson, S. 1994 Multiple image invariance using the double algebra. In *Applications of invariance in computer vision* (ed. J. L. Mundy, A. Zissermann & D. Forsyth), pp. 145–164. Lecture Notes in Computer Science, vol. 825. Springer.

Doubilet, P., Rota, G.-C. & Stein, J. 1974 On the foundations of combinatorial theory. IX. Combinatorial methods in invariant theory. *Stud. Appl. Math.* **53**, 185–216.

Eklundh, J.-O. (ed.) 1994 Lecture Notes in Computer Science, vol. 800–801. Stockholm, Sweden: Springer.

Faugeras, O. 1992 What can be seen in three dimensions with an uncalibrated stereo rig. In *Proc. 2nd European Conf. on Computer Vision* (ed. G. Sandini), pp. 563–578. Lecture Notes in Computer Science, vol. 588. Santa Margherita Ligure, Italy: Springer.

Faugeras, O. 1993 *Three-dimensional computer vision: a geometric viewpoint*. MIT Press.

Faugeras, O. & Mourrain, B. 1995*a* On the geometry and algebra of the point and line correspondences between $n$ images. In *Proc. Int. Conf. on Computer Vision*, pp. 951–956. Boston, MA: IEEE Computer Society Press.

Faugeras, O. & Mourrain, B. 1995*b* On the geometry and algebra of the point and line correspondences between $n$ images. Technical Report 2665, INRIA.

Faugeras, O. & Papadopoulo, T. 1998 A nonlinear method for estimating the projective geometry of three views. In *Proc. 6th Int. Conf. on Computer Vision*, pp. 477–484. Bombay, India: IEEE Computer Society Press.

Hartley, R. 1994 Lines and points in three views—an integrated approach. In *Proc. ARPA Image Understanding Workshop*. Defense Advanced Research Projects Agency, Morgan Kaufmann Publishers.

Hartley, R. 1995 In defence of the 8-point algorithm. *Proc. Int. Conf. on Computer Vision*, pp. 1064–1070. Boston, MA: IEEE Computer Society Press.

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

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

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

Luong, Q.-T. & Viéville, T. 1994 Canonic representations for the geometries of multiple projective views. In *Proc. European Conf. on Computer Vision*, pp. 589–599.

Shashua, A. 1994 Trilinearity in visual recognition by alignment. In *ECCV 1994*, pp. 479–484.

Shashua, A. 1995 Algebraic functions for recognition. *IEEE Trans. Pattern Analysis Machine Intell.* **17**, 779–789.

Shashua, A. & Werman, M. 1995 Trilinearity of three perspective views and its associated tensor. In *Proc. Int. Conf. on Computer Vision*, pp. 920–925.

Spetsakis, M. E. & Aloimonos, J. 1990*a* Structure from motion using line correspondences. *Int. J. Computer Vision* **4**, 171–183.

Spetsakis, M. E. & Aloimonos, Y. 1990*b* A unified theory of structure from motion. In *Proc. DARPA IU Workshop*, pp. 271–283.

Torr, P. & Zissermann, A. 1997 Performance characterization of fundamental matrix estimation under image degradation. *Machine Vision Applic.* **9**, 321–333.

Weng, J., Huang, T. & Ahuja, N. 1992 Motion and structure from line correspondences: closed-form solution, uniqueness and optimization. *IEEE Trans. Pattern Analysis Machine Intell.* **14**.

Zhang, Z., Deriche, R., Faugeras, O. & Luong, Q.-T. 1995 A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artif. Intell. J.* **78**, 87–119.


## *Discussion*

W. Triggs (*INRIA, France*). This is a very nice approach to the trifocal tensor. However, although we have the rank 2 constraint on each matrix $G_n$, we also have the rank 2 constraint on any linear combination of the three matrices. This therefore gives 10 constraints none of which contain the epipoles, which are third order and are therefore simpler to use than those Dr Faugeras gave.




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

O. Faugeras. Theorems 5.14 and 5.15 in the paper give two sets of algebraic relations that are sufficient for a tensor to be a trifocal tensor. The second set is somewhat simpler than the first. The difficulty is to prove sufficiency. This is done in the paper.

W. Triggs. How does one deal with quadrics in the Grassmann–Cayley algebra?

O. Faugeras. The Grassmann–Cayley algebra is a way to deal algebraically with such operations as sums and intersections of vector subspaces of a vector space. Thus, quadrics apparently do not fall immediately in that framework. But since they are ruled surfaces, the Grassmann–Cayley algebra can in fact be very elegantly used to describe families of lines of which quadrics are a special case.

T. Vieville (*INRIA, France*). How automatically can we derive the minimal parametrization of $T$? Can one program an automatic derivation of the parametrization? In other words, can stupid people also use the Grassmann–Cayley algebra to do vision?

O. Faugeras. Yes, of course, one can compute the parametrization automatically just as one can for the fundamental matrix.

J. Lasenby (*Department of Engineering, University of Cambridge, UK*). Is it easy to use the Grassmann–Cayley algebra and Dr Faugeras's system to extract all the epipoles from a given set of 27 numbers which make up some trilinear tensor?

O. Faugeras. Two of the three fundamental matrices are obtained in a straightforward fashion from the left and right nullspaces of the matrices $\boldsymbol{G}^n$, $n = 1, 2, 3$ (propositions 5.5 and 5.6). Recovering the third fundamental matrix can be done as explained at the end of theorem 5.11.

J. Lasenby. Is it possible for to expand the Grassmann–Cayley algebra to cover things other than projective geometry?

O. Faugeras. Dr Lasenby probably knows more about this than I do! Yes, we can of course look at including the rigid motion group but here I was focusing on the projective geometry approach which we have found to be quite useful. The way to include the group of rigid motions is to go to Clifford algebras which are 'deformations' of exterior algebras. The best known example are the quaternions and the dual quaternions which can be efficiently used to represent the set of rigid displacements of a three-dimensional Euclidean affine space.

J. Lasenby. Are the reconstructions all based on a nonlinear estimation of $T$?

O. Faugeras. Yes of course. But we have developed in my group a large software system which can use $F$ or $T$ to perform 3D reconstruction and can switch between the two. The estimation of $T$ is nonlinear as a linear estimation is highly non-robust to noise since it neglects the nonlinear constraints.

R. I. Hartley (*GE Corporate Research and Development, Niskayuna, NY, USA*). Once the inverse problem satisfies these constraints, it is then a trifocal tensor. Is that a different concept from just saying that the set of constraints is a complete set of constraints, a generator for the complete ideal?

O. Faugeras. No, it is the same concept. The problem that I have been addressing is that of finding a good compromise between the number of generators of the ideal and their simplicity (i.e. their degree). From that standpoint the result of theorem 5.15 is better than the result of theorem 5.14.

R. I. Hartley. Wouldn't the thing then be something like just a complete polynomial modulo this ideal?

O. Faugeras. Well, the ideal we have not been able to compute, actually, if that is the question. This means precisely that we have not been able to compute a Gröbner basis. Because, simply there are too many variables and so on and so forth. Even Macaulay (a symbolic algebra package) couldn't do it for us.

R. I. Hartley. What I'm saying is, if you had the ideal, then essentially you have the trifocal tensor, you have something which is therefore almost by definition the trifocal tensor. I'm being very vague.

O. Faugeras. As I said before, we did not know the ideal, i.e. we did not have either a set of generators or a Gröbner basis. One of the contributions of the paper is to give, in theorems 5.14 and 5.15, two different sets of generators.

W. Triggs. Locally, at least the construction that I had with the 10 cubic constraints; if you look at the Jacobian of that, it's guaranteed that you remove eight of the degrees of freedom, so you get down to the 18 that we know the trifocal tensor has. So, *locally* you know that those constraints are sufficient, at least locally and generically. That does not mean that there may not be some other tensor somewhere else in the space which satisfies those constraints but which are not the trifocal tensor.

O. Faugeras. As proved in the paper in proposition 5.19, the 10 cubic contraints are not sufficient (they generate a larger ideal). You need to add something, for example the two epipolar constraints (theorem 5.15).

It seems that the manifold of trifocal tensors has at least two components, the one for the generic configuration when the three optical centres are not aligned, and another one, smaller, when they are aligned. This is a conjecture.