About the correspondences of points between N images

O. Faugeras B. Mourrain
INRIA
2004 route des Lucioles,
BP 93, 06902 Sophia-Antipolis,
FRANCE
faugeras,mourrain@sophia.inria.fr

Abstract

We analyze the correspondence of points between an arbitrary number of images, from an algebraic and geometric point of view. We use the formalism of the Grassmann-Cayley algebra as the simplest way to make both geometric and algebraic statements in a very synthetic and effective way (i.e. allowing actual computation if needed). We propose a systematic way to describe the algebraic relations which are satisfied by the coordinates of the images of a 3-D point.

They are of three types: bilinear relations arising when we consider pairs of images among the N and which are the well-known epipolar constraints, trilinear relations arising when we consider triples of images among the N, and quadrilinear relations arising when we consider four-tuples of images among the N. Moreover, we show how two trilinear relations imply the bilinear ones (i.e. the epipolar constraints). We also show how these trilinear constraints can be used to predict the image coordinates of a point in a third image, given the coordinates of the images in the other two images, even in cases where the prediction by the epipolar constraints fails (points in the trifocal plane, or optical centers aligned).

Finally, we show that the quadrilinear relations are in the ideal generated by the bilinearities and trilinearities, and do not bring in any new information. This completes the algebraic description of correspondence between any number of cameras.

Keywords: Invariants, Geometry, Geometry of N cameras, Grassmann-Cayley algebra, Plücker relations, multiple cameras stereo.

1 Introduction

Understanding the geometry of the correspondences between image primitives that arise from the perspective projection of three-dimensional objects is fundamental for such applications as three-dimensional reconstruction from multiple views, for example stereo and motion, object recognition, image synthesis, image coding. Recent theoretical efforts directed toward the development of such an understanding have demonstrated the importance of projective geometry as the language allowing the simplest description of the underlying phenomena.

In the case of two cameras, the theory is almost complete, the main fact being that correspondences between points in two images are completely described by the epipolar geometry which can itself be summarized algebraically in a 3×3 matrix of rank 2, the fundamental matrix [10, 11, 12].

The case of three images or more has not been studied as extensively as the case of two. Faugeras and Robert [8] have shown that the knowledge of the three fundamental matrixes of the three pairs of images could be used to predict correspondences in a third image from correspondences in the other two by a simple use of the epipolar geometry. We show in the main body of this article that this method may fail in some cases (see also [18]).

In this paper, we propose a systematic way of deriving the necessary and sufficient conditions for points in each retinal plane to be in correspondence. By this approach, we recover all the relations that appeared in previous works (see [9], [13, 14], [16]). More precisely, the set of N-tuples of corresponding points form an algebraic variety that we describe completely by the ideal of polynomial functions that vanish on it. We show that the generators of this ideal are precisely the relations introduced before and that for more than 4 cameras, this ideal is generated by the bilinearities and trilinearities.

We complete this presentation by showing that trilinearities are more powerful that bilinearities in some degenerate cases, closely connected to the previous ideal. This is an extension of the work described in [6, 7].

2 The model

We consider N cameras C_i in the 3-D space. They are classically described by projections on a plane (the retinal plane) from a point O_i (the center of projection). In order to have good geometrical behaviors, we consider the 3-D space as embedded in a three-dimensional projective space that we note $\mathbb{P}^3$. A point M of $\mathbb{P}^3$ is represented by a 4-dimensional vector $\mathbf{M} = (M_1 : M_2 : M_3 : M_4)$ (defined up to a non-zero scalar).

Similarly, we embed the retinal plane in a projective space $\mathbb{P}^2$ of dimension 2 and the images m_i of the point M in the camera C_i are represented by a 3-dimensional vector $\mathbf{m}_i = (x_i : y_i : z_i)$. For each camera C_i, we have the following map

$$\begin{array}{rcl} \mathbf{C}_i : \ \mathbb{P}^3 & \rightarrow & \mathbb{P}^2 \\ \mathbf{M} & \mapsto & \mathbf{m}_i \equiv \mathbf{C}_i.\mathbf{M} \end{array}$$

The i-th perspective projection matrix, noted $\mathbf{C}_i$, is a 3×4 matrix whose 3 row vectors will be denoted by the digits $3 \times (\mathbf{i} - \mathbf{1}) + \mathbf{j}$, $i = 1, \cdots, N$, $j = 1, 2, 3$. Thus, the perspective projection equation of the first camera can be written as:

$$\mathbf{m}_1 \equiv \mathbf{C}_1 \mathbf{M} = \begin{bmatrix} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \end{bmatrix} \mathbf{M},$$

or

$$(x_1 : y_1 : z_1) \equiv (\mathbf{1} \cdot \mathbf{M} : \mathbf{2} \cdot \mathbf{M} : \mathbf{3} \cdot \mathbf{M}) \qquad (1)$$

where for example $\mathbf{1} \cdot \mathbf{M}$ represents the usual inner product of the row vector $\mathbf{1}$ with the column vector $\mathbf{M}$. By convention, bold letters indicate vectors or matrixes.

Note that the projection is not defined at the center of projection: $\mathbf{C}_i.\mathbf{O}_i = 0$.

The equation (1) is equivalent to the three equations

$$\begin{array}{l} x_1 \mathbf{2} \cdot \mathbf{M} - y_1 \mathbf{1} \cdot \mathbf{M} = 0 \\ y_1 \mathbf{3} \cdot \mathbf{M} - z_1 \mathbf{2} \cdot \mathbf{M} = 0 \\ z_1 \mathbf{1} \cdot \mathbf{M} - x_1 \mathbf{3} \cdot \mathbf{M} = 0 \end{array} \qquad (2)$$

We note that a linear combination of these equations vanishes, so that only two of them are useful. In the following, we will take the first two equations of this type for each camera. Stacking them up into a matrix, we obtain

$$\begin{bmatrix} x_1 \, \mathbf{2} - y_1 \, \mathbf{1} \\ y_1 \, \mathbf{3} - z_1 \, \mathbf{2} \\ x_2 \, \mathbf{5} - y_2 \, \mathbf{4} \\ y_2 \, \mathbf{6} - z_2 \, \mathbf{5} \\ \vdots \end{bmatrix} \mathbf{M} = \mathbf{0}. \qquad (3)$$

Let us denote by R_i the i$^{\text{th}}$ row of this matrix. As $\mathbf{M}$ is a non-zero vector, any 4 rows of this matrix are linearly dependent. Consequently

$$det(R_{i_1}, R_{i_2}, R_{i_3}, R_{i_4}) = 0$$

for all $1 \leq i_1 < i_2 < i_3 < i_4 \leq 3\,N$. Conversely if this matrix is of rank less than 3, then we can construct a point $M \in \mathbb{P}^3$ satisfying the relations (3) and (1). In other words, *there exist a point $M \in \mathbb{P}^3$ such that m_i ($1 \leq i \leq N$) is the image of M in the i$^{\text{th}}$ camera if and only if the matrix (3) is of rank less than* 3.

The 4×4 minors of this matrix fall into three classes

- those involving only two cameras, with two rows from each camera.
- those involving three cameras, with two rows from the first camera, one row from the second and one row from the third,
- those involving four cameras, with one row arising from each camera.

In this paper, we are going to describe carefully these relations. But first we need to introduce some mathematical notations.

3 Mathematical background

We denote by $\mathbb{E} = \mathbb{E}^4$ the vector spaces of dimension 4 over the reals $\mathbb{R}$. The rows of the projection matrices belong to this space. They represent planes in $\mathbb{P}^3$. If $\mathbf{U} = (U_1, U_2, U_3, U_4)$ is such a vector, the plane associated to it, has the equation $U_1 \, M_1 + U_2 \, M_2 + U_3 \, M_3 + U_4 \, M_4 = 0$. A non-zero multiple of $\mathbf{U}$ represents the same plane so that the set of planes $\hat{\mathbb{P}}^3$ has naturally a structure of projective space of dimension 3. It is called the dual space of $\mathbb{P}^3$.

In order to be able to manipulate subspaces of $\mathbb{P}^3$, we need to introduce the **exterior algebra** $\wedge\mathbb{E}$ of $\mathbb{E}$. This algebra has an anticommutative product (noted $\wedge$) which satisfied $e_i \wedge (\lambda \, e_j + \mu \, e_k) = \lambda \, e_i \wedge e_j + \mu \, e_i \wedge e_k$ for all $\lambda, \mu \in \mathbb{R}$. Let $e_1, \ldots, e_4$ be the canonical basis of $\mathbb{E}$. Then $\wedge\mathbb{E}$ is a vector space of basis $(e_{i_1} \wedge \cdots \wedge e_{i_k})$ with $1 \leq i_1 < \cdots < i_k \leq 4$.

For any vectors $u_1 = (u_{1,1}, \ldots, u_{4,1})$, $\ldots$, $u_m = (u_{1,m}, \ldots, u_{4,m})$ of $\mathbb{E}$, the coordinates of $u_1 \wedge \cdots \wedge u_m$ in the basis $(e_{i_1} \wedge \cdots \wedge e_{i_m})_{1 \leq i_1 < \cdots < i_m \leq 4}$ of $\wedge^m \mathbb{E}$ are the corresponding determinants $|u_{i_k,j}|_{1 \leq i,k \leq m}$. So $u_1 \wedge \cdots \wedge u_m = 0$ if and only if the vectors $\{u_1, \ldots, u_m\}$ are linearly dependent.

We associate naturally to an element $u_1 \wedge \cdots \wedge u_m$ the subspace of $\mathbb{P}^3$ corresponding to the intersection of the hyperplane $u_1, \ldots u_m$. For instance $u_1 \wedge u_2$ is the line of intersection of the two planes u_1, u_2 and

$u_1 \wedge u_2 \wedge u_3$ represents a point of $\mathbb{P}^3$. Conversely any linear subspace L of $\mathbb{P}^3$ can be represented by a unique element $u_1 \wedge \cdots \wedge u_m$ (up to a non-zero scalar). We just take any basis $u_1, \ldots u_m$ of L. The elements which represent linear subspace of $\mathbb{P}^3$ form a subvariety of $\mathbb{P}(\wedge\mathbb{E})$ called the *Grassmannian.*

We illustrate these notions in our context:

- $x_1\,\mathbf{2} - y_1\,\mathbf{1}, y_1\,\mathbf{3} - z_1\,\mathbf{2}$ *are planes which contain the optical center* O_1 *(because* $\mathbf{C}_1.\mathbf{O}_1 = 0$*) and the point* $\mathbf{M}$ *by construction.*
- $(x_1\,\mathbf{2} - y_1\,\mathbf{1}) \wedge (y_1\,\mathbf{3} - z_1\,\mathbf{2}) = y_1(z_1\mathbf{1} \wedge \mathbf{2} - y_1\mathbf{1} \wedge \mathbf{3} + x_1\mathbf{2} \wedge \mathbf{3})$, *is the* **optical ray** (O_1, M).
- $\mathbf{1} \wedge \mathbf{2} \wedge \mathbf{3}$ *is the intersection of the planes defined by the rows of* C_1. *It is the optical center* O_1 *of the first camera.*

We check immediately that for any vectors $\mathbf{U}_1, \ldots, \mathbf{U}_4$ of $\mathbb{E}$, we have the properties $(\mathbf{U}_1 \cdot \mathbf{U}_2 \wedge \mathbf{U}_3 \wedge \mathbf{U}_4) = [\mathbf{U}_1, \ldots, \mathbf{U}_4]$, where $[\mathbf{U}_1, \ldots, \mathbf{U}_4]$ denotes the 4×4 determinant of these vectors. In particular two lines $(\mathbf{a} \wedge \mathbf{b})$, $(\mathbf{c} \wedge \mathbf{d})$ intersect if and only if $[\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}] = 0$ or equivalently $(\mathbf{a} \wedge \mathbf{b}) \wedge (\mathbf{c} \wedge \mathbf{d}) = 0$. As $\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} \wedge \mathbf{d} = [\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}]\, e_1 \wedge e_2 \wedge e_3 \wedge e_4$, with a slight abuse of notations we may identify $\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} \wedge \mathbf{d}$ with $[\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}]$.

A natural approach when dealing with intrinsic quantities, is to consider the determinants of points as variables $[j_1, \ldots, j_4]$. Therefore we have to manipulate polynomials in these variables. Especially, we want to be able to check when such a polynomial is zero or not. This is not straightforward, for there exist relations between these variables. A fundamental result in invariant theory says that these relations are generated by the **Plücker relations**. This algebra of determinants fall in the category of **algebra with straightening laws**, for which there is an algorithmic way to normalize any element. Moreover, the monomials which are normalized can be described in terms of a partial order on the variables (see for instance [4], [5], or [15]). In this paper, the computations and tests of equality to zero are based on this normalization procedure that we have implemented in Maple.

More details about invariant theory, exterior calculus and the Grassmann-Cayley algebra can be found in the classical book [17], or in the quite accessible book by Sturmfels [15], and in the more advanced article by Barnabei et al. [1]. A good introduction, targeted at computer vision researchers, can also be found in [2].

4 Bilinear constraints

In this section, we consider the 4×4 minors where two rows R_i arise from two cameras. For instance, if we take $det(R_1, R_2, R_3, R_4)$, we obtain

$$y_1 y_2 F_{1,2}(\mathbf{m}_1, \mathbf{m}_2) = y_1 y_2 \;\left(\mathbf{m}_2^T \mathbf{F}_{1,2} \mathbf{m}_1\right) = 0$$

where

$$\begin{aligned} F_{1,2}(\mathbf{m}_1, \mathbf{m}_2) &= (x_1\mathbf{2} \wedge \mathbf{3} + y_1\mathbf{3} \wedge \mathbf{1} + z_1\mathbf{1} \wedge \mathbf{2}) \\ &\wedge \quad (x_2\mathbf{5} \wedge \mathbf{6} + y_2\mathbf{6} \wedge \mathbf{4} + z_2\mathbf{4} \wedge \mathbf{5}) \end{aligned}$$

Geometrically, we are just saying that the two optical rays (O_1, m_1), (O_2, m_2) intersect.

Expanding this product, we obtain a bilinear form in m_1, m_2 whose matrix is given by

$$\mathbf{F}_{1,2} = \begin{bmatrix} [2,3,5,6] & -[1,3,5,6] & [1,2,5,6] \\ -[2,3,4,6] & [1,3,4,6] & -[1,2,4,6] \\ [2,3,4,5] & -[1,3,4,5] & [1,2,4,5] \end{bmatrix}.$$

This is *the fundamental matrix* between cameras 1 and 2.

The epipoles $e_{1,2}, e_{2,1}$ can also be described very easily in terms of determinants. For instance $e_{1,2}$ is the image of $O_2 = \mathbf{4} \wedge \mathbf{5} \wedge \mathbf{6}$ in the first camera:

$$\begin{aligned} \mathbf{e}_{1,2} &= ((\mathbf{1} \cdot \mathbf{O}_2) : (\mathbf{2} \cdot \mathbf{O}_2) : (\mathbf{3} \cdot \mathbf{O}_2)) \\ &= ([1,4,5,6] : [2,4,5,6] : [3,4,5,6]). \end{aligned}$$

Similarly

$$\mathbf{e}_{2,1} = ([1,2,3,4] : [1,2,3,5] : [1,2,3,6]).$$

Using Plücker relations, we check that $\mathbf{F}_{1,2}.\mathbf{e}_{1,2} = \mathbf{e}_{2,1}^T \mathbf{F}_{1,2} = 0$.

Though this case is now well understood, it is worthwhile going through this analysis because it turns out that the way we approach it is rather different from the usual way and can be extended in a straightforward fashion to the trilinear and quadrilinear cases.

5 Trilinear constraints

In this section, we consider relations where two rows R_i are arising from the first camera, one from the second and one from the third. As these trilinearities are not yet very widely used or understood, we will spend a large section on them.

One can for instance take $det(R_1, R_2, R_3, R_5) = y_1 T_{1,2,3,5}(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3)$ where

$$\begin{aligned} &T_{1,2,3,5}(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3) \\ &= (x_1\mathbf{2} \wedge \mathbf{3} + y_1\mathbf{3} \wedge \mathbf{1} + z_1\mathbf{1} \wedge \mathbf{2}) \wedge (x_2\mathbf{5} - y_2\mathbf{4}) \\ &\qquad \wedge (x_3\mathbf{8} - y_3\mathbf{7}) \\ &= x_1\,x_2\,x_3\,[2,3,5,8] - x_1\,x_2\,y_3\,[2,3,5,7] \\ &\quad + x_1\,y_2\,x_3\,[2,4,3,8] - x_1\,y_2\,y_3\,[2,4,3,7] \\ &\quad - y_1\,x_2\,x_3\,[1,3,5,8] + y_1\,x_2\,y_3\,[1,3,5,7] \\ &\quad - y_1\,x_3\,y_2\,[1,4,3,8] + y_1\,y_2\,y_3\,[1,4,3,7] \\ &\quad + z_1\,x_2\,x_3\,[1,2,5,8] - z_1\,x_2\,y_3\,[1,2,5,7] \\ &\quad - z_1\,y_2\,x_3\,[1,2,4,8] + z_1\,y_2\,y_3\,[1,2,4,7]. \end{aligned}$$

Geometrically, we are saying that the optical ray (O_1, m_1), the plane through the line $\mathbf{4} \wedge \mathbf{5}$ and the point m_2, the plane through the line $\mathbf{7} \wedge \mathbf{8}$ and the point m_3 intersect.

We can construct 12 such relations with 3 cameras:

Group I

$T_{1,2,3,5} = (x_1 \mathbf{2} \wedge \mathbf{3} + y_1 \mathbf{3} \wedge \mathbf{1} + z_1 \mathbf{1} \wedge \mathbf{2}) \wedge (x_2 \mathbf{5} - y_2 \mathbf{4}) \wedge (x_3 \mathbf{8} - y_3 \mathbf{7})$

$T_{1,2,4,5} = (x_1 \mathbf{2} \wedge \mathbf{3} + y_1 \mathbf{3} \wedge \mathbf{1} + z_1 \mathbf{1} \wedge \mathbf{2}) \wedge (y_2 \mathbf{6} - z_2 \mathbf{5}) \wedge (x_3 \mathbf{8} - y_3 \mathbf{7})$

$T_{1,2,3,6} = (x_1 \mathbf{2} \wedge \mathbf{3} + y_1 \mathbf{3} \wedge \mathbf{1} + z_1 \mathbf{1} \wedge \mathbf{2}) \wedge (x_2 \mathbf{5} - y_2 \mathbf{4}) \wedge (y_3 \mathbf{9} - z_3 \mathbf{8})$

$T_{1,2,4,6} = (x_1 \mathbf{2} \wedge \mathbf{3} + y_1 \mathbf{3} \wedge \mathbf{1} + z_1 \mathbf{1} \wedge \mathbf{2}) \wedge (y_2 \mathbf{6} - z_2 \mathbf{5}) \wedge (y_3 \mathbf{9} - z_3 \mathbf{8})$

Group II

$T_{1,3,4,5} = (x_1 \mathbf{2} - y_1 \mathbf{1}) \wedge (x_2 \mathbf{4} \wedge \mathbf{5} + y_2 \mathbf{5} \wedge \mathbf{6} + z_2 \mathbf{6} \wedge \mathbf{4}) \wedge (x_3 \mathbf{8} - y_3 \mathbf{7})$

$T_{2,3,4,5} = (y_1 \mathbf{3} - z_1 \mathbf{2}) \wedge (x_2 \mathbf{4} \wedge \mathbf{5} + y_2 \mathbf{5} \wedge \mathbf{6} + z_2 \mathbf{6} \wedge \mathbf{4}) \wedge (x_3 \mathbf{8} - y_3 \mathbf{7})$

$T_{1,3,4,6} = (x_1 \mathbf{2} - y_1 \mathbf{1}) \wedge (x_2 \mathbf{4} \wedge \mathbf{5} + y_2 \mathbf{5} \wedge \mathbf{6} + z_2 \mathbf{6} \wedge \mathbf{4}) \wedge (y_3 \mathbf{9} - z_3 \mathbf{8})$

$T_{2,3,4,6} = (y_1 \mathbf{3} - z_1 \mathbf{2}) \wedge (x_2 \mathbf{4} \wedge \mathbf{5} + y_2 \mathbf{5} \wedge \mathbf{6} + z_2 \mathbf{6} \wedge \mathbf{4}) \wedge (y_3 \mathbf{9} - z_3 \mathbf{8})$

Group III

$T_{1,3,5,6} = (x_1 \mathbf{2} - y_1 \mathbf{1}) \wedge (x_2 \mathbf{5} - y_2 \mathbf{4}) \wedge (x_3 \mathbf{7} \wedge \mathbf{8} + y_3 \mathbf{8} \wedge \mathbf{9} + z_3 \mathbf{9} \wedge \mathbf{7})$

$T_{1,4,5,6} = (x_1 \mathbf{2} - y_1 \mathbf{1}) \wedge (y_2 \mathbf{6} - z_2 \mathbf{5}) \wedge (x_3 \mathbf{7} \wedge \mathbf{8} + y_3 \mathbf{8} \wedge \mathbf{9} + z_3 \mathbf{9} \wedge \mathbf{7})$

$T_{2,3,5,6} = (y_1 \mathbf{3} - z_1 \mathbf{2}) \wedge (x_2 \mathbf{5} - y_2 \mathbf{4}) \wedge (x_3 \mathbf{7} \wedge \mathbf{8} + y_3 \mathbf{8} \wedge \mathbf{9} + z_3 \mathbf{9} \wedge \mathbf{7})$

$T_{2,4,5,6} = (y_1 \mathbf{3} - z_1 \mathbf{2}) \wedge (y_2 \mathbf{6} - z_2 \mathbf{5}) \wedge (x_3 \mathbf{7} \wedge \mathbf{8} + y_3 \mathbf{8} \wedge \mathbf{9} + z_3 \mathbf{9} \wedge \mathbf{7})$

5.1 Reconstruction from trilinearities

Knowing these trilinearities, one can wonder if the geometry of the 3 cameras can be recovered. By this we mean their fundamental matrices and relative positions if they are calibrated. We are going to see that in fact two trilinearities are enough to recover the 3 bilinearities.

By definition

$$T_{1235}(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3) = (R_1 \wedge R_2)(\mathbf{m}_1) \wedge R_3(\mathbf{m}_2) \wedge R_5(\mathbf{m}_3)$$

with $R_5(\mathbf{m}_3) = x_3 \mathbf{8} - y_3 \mathbf{7}$. So by *Cramer's rule*

$$\begin{aligned} R_5(\mathbf{e}_{31}) &= [1,2,3,7]\,\mathbf{8} - [1,2,3,8]\,\mathbf{7} \\ &= [1,2,7,8]\,\mathbf{1} - [1,3,7,8]\,\mathbf{2} + [2,3,7,8]\,\mathbf{3} \end{aligned}$$

Substituting this relation in $T_{1235}(\mathbf{m}_1, \mathbf{m}_2, \mathbf{e}_{31})$ yields

$$\begin{aligned} & (x_1[1,2,7,8] + y_1[1,3,7,8] + z_1[2,3,7,8]) \\ \times\ & (\mathbf{1} \wedge \mathbf{2} \wedge \mathbf{3} \wedge (x_2 \mathbf{5} - y_2 \mathbf{4})) \\ =\ & (x_1[1,2,7,8] + y_1[1,3,7,8] + z_1[2,3,7,8]) \\ & \times (x_2[1,2,3,5] - y_2[1,2,3,4]) \\ =\ & (\mathbf{F}_{1,3}.\mathbf{m}_1)_3\,(\mathbf{m}_2 \wedge \mathbf{e}_{2,1})_{1,2} \end{aligned}$$

We denote the third coordinate of $\mathbf{F}_{1,3}.\mathbf{m}_1$ by $(\mathbf{F}_{1,3}.\mathbf{m}_1)_3$, and by $(\mathbf{m}_2 \wedge \mathbf{e}_{2,1})_{1,2}$ the coefficient of $e_1 \wedge e_2$ in its expansion in the canonical basis.

So we see that $T(\mathbf{m}_1, \mathbf{m}_2, \mathbf{e}_{3,1})$ *is a product of two linear forms.* We can use this property to recover the epipoles and the fundamental matrices.

Decomposing T_{1235} as a polynomial in x_1, y_1, z_1 and x_2, y_2 with coefficients depending on m_3, yields a matrix

$$\begin{array}{cc} & \begin{array}{cc} x_2 & z_2 \end{array} \\ \begin{array}{c} x_1 \\ y_1 \\ z_1 \end{array} & \left[\begin{array}{cc} c_{11}(x_3,y_3) & c_{12}(x_3,y_3) \\ c_{21}(x_3,y_3) & c_{22}(x_3,y_3) \\ c_{31}(x_3,y_3) & c_{32}(x_3,y_3) \end{array} \right]. \end{array}$$

The bilinear form in m_1, m_2 is a product of two linear forms if and only if this matrix is of rank 1. Indeed, if this matrix is of rank 1, the first linear form corresponds (up to a scalar) to the first (or the second) column of this matrix.

This fact can be used to determine the first two coordinates of the epipole $e_{3,1}$ (up to a scalar) as follows:

1. Compute the 2×2 minors of this matrix :

   $$\begin{aligned} & d_{1,2} = [1,2,3,8]\,[3,4,5,8]\,{x_3}^2 \\ & + (-2\,[1,2,3,7]\,[3,4,5,8] + [1,2,3,5]\,[3,4,7,8] \\ & -\ [1,2,3,4]\,[3,5,7,8])\,x_3\,y_3 \\ & +[1,2,3,7]\,[3,4,5,7]\,{y_3}^2 \\ & d_{1,3} = -[1,2,3,8]\,[2,4,5,8]\,{x_3}^2 \\ & + (2\,[1,2,3,7]\,[2,4,5,8] - [1,2,3,5]\,[2,4,7,8] \\ & +[1,2,3,4]\,[2,5,7,8])\,x_3\,y_3 \\ & -[1,2,3,7]\,[2,4,5,7]\,{y_3}^2 \\ & d_{2,3} = [1,2,3,8]\,[1,4,5,8]\,{x_3}^2 \\ & + (-2\,[1,2,3,7]\,[1,4,5,8] + [1,2,3,5]\,[1,4,7,8] \\ & -[1,2,3,4]\,[1,5,7,8])\,x_3\,y_3 \\ & +[1,2,3,7]\,[1,4,5,7]\,{y_3}^2 \end{aligned}$$

2. Take their Greater Common Divisor which is a polynomial of degree 1 in x_3, y_3. It yields the first two coordinates of $e_{3,1}$.

3. Use the factorization of $T_{1,2,3,5}(\mathbf{m}_1, \mathbf{m}_2, \mathbf{e}_{31})$ to recover the first two coordinates of $e_{2,1}$ and the last row of $\mathbf{F}_{1,3}$.

4. Apply the same process with $\mathbf{m}_2 = \mathbf{e}_{2,1}$ in order to get the last row of $\mathbf{F}_{1,2}$.

5. Do it again for $T_{1,2,4,6}$ in order to obtain the last two coordinates of $e_{2,1}$, the last two coordinates of $e_{3,1}$, the first row of $\mathbf{F}_{1,3}$, and $\mathbf{F}_{1,2}$.

To sum up, if we know $T_{1,2,3,5}$ and $T_{1,2,4,6}$, we also know the epipoles $e_{2,1}, e_{3,1}$.

From the relations $\mathbf{e}_{2,1}^T\mathbf{F}_{12} = 0, \mathbf{e}_{3,1}^T\mathbf{F}_{13} = 0$, and with the first and last rows of $\mathbf{F}_{1,2}$ and $\mathbf{F}_{1,3}$, we deduce the second row of these matrices (the above relations gives us the second coordinate with respect to the first and last one). So we can also recover the matrices $\mathbf{F}_{1,2}$ and $\mathbf{F}_{3,1}$.

In order to determine $\mathbf{F}_{2,3}$, we proceed as follows:

1. Consider a point m_2 and its epipolar line in the first image, represented by $\mathbf{F}_{2,1} \cdot \mathbf{m}_2$.
2. Choose any point m_1 different from $e_{1,2}$ on that line and use the two trilinearities $T_{1,2,3,5}$ and $T_{1,2,4,6}$ to predict m_3 in the third image.
3. The point m_3 must be on the epipolar line of m_2 in the third retinal plane, represented by $\mathbf{F}_{2,3} \cdot \mathbf{m}_2$.
4. When m_1 moves along the previous epipolar line, m_3 moves on a locus which contains the epipolar line of m_2. This locus is a conic which splits into two lines,
   - the line of equation $y_3 = 0$ and
   - the line of equation $\mathbf{m}_3^T\mathbf{F}_{2,3}\mathbf{m}_2$.

Algebraically, we can take the determinant of the following linear system in m_1:

$$\begin{cases} F_{1,2}(m_1, m_2) = 0, \\ T_{1,2,3,5}(m_1, m_2, m_3) = 0, \\ T_{1,2,4,6}(m_1, m_2, m_3) = 0 \end{cases}$$

It is an equation of degree 2 in m_3 which factors as

$$\begin{aligned} &(y_2 e_{2,1}[3] - z_2 e_{2,1}[2])(y_2 e_{2,1}[1] - x_2 e_{2,1}[2]) e_{1,2}[3] \\ &\times y_3 \times (\mathbf{m}_3^T \mathbf{F}_{2,3} \mathbf{m}_2) = 0 \end{aligned}$$

($e_{i,j}[k]$ is the k$^{\text{th}}$ coordinate of $e_{i,j}$). One of its factor is precisely the bilinear form we seek.

Hence, we can also compute the last fundamental matrix $\mathbf{F}_{2,3}$ from $T_{1,2,3,5}$ and $T_{1,2,4,6}$.

Consequently the geometry of the 3 *cameras can be recovered completely from those two trilinearities.* By symmetry, this also holds for the other groups.

5.2 The algebraic variety of points in correspondence

We are now going to give a precise description of the triples of points in correspondence for 3 cameras. Those triples of points (m_1, m_2, m_3) that correspond to the images of a point M in $\mathbb{P}^3$ form an algebraic variety that we are going to describe. Our presentation is illustrated by explicit symbolic computations (ie. Gröbner bases computations) on special configurations of cameras. Proofs of the underlying theoretical results will be presented elsewhere.

Let call V_3 this subset of $\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2$. In order to describe this variety, we give the ideal I_3 of the functions that vanish on V_3. We already know some of them, i.e. the bilinearities between two cameras and the trilinearities of the previous section. These functions are polynomial functions on $\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2$, which are homogeneous with respect to m_1, m_2, m_3. We denote by $\mathbb{R}[\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2]$ this set of functions.

The ideal I_3 is the kernel of the following map:

$$\begin{gathered} 0 \rightarrow I_3 \rightarrow \mathbb{R}[\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2] \xrightarrow{\sigma} \mathbb{R}[\mathbb{P}^3] \rightarrow 0 \\ f(m_1, m_2, m_3) \mapsto f(\mathbf{C}_1.\mathbf{M}, \mathbf{C}_2.\mathbf{M}, \mathbf{C}_3.\mathbf{M}) \end{gathered}$$

which substitutes $(\mathbf{1}.\mathbf{M})$ for the variable x_1, $(\mathbf{2}.\mathbf{M})$ for y_1, ... If f is homogeneous in $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$, we obtain a homogeneous polynomial in the coordinates of $\mathbf{M}$, belonging to $\mathbb{R}[\mathbb{P}^3]$.

In order to compute the kernel of this map, we proceed as follows. We introduce new variables (of homogeneity) a, b, c and work in the ring $R = \mathbb{R}[x_i, y_i, z_i, M_1, M_2, M_3, M_4, a, b, c]$. Let us note $\tilde{I_3}$ the ideal in R, generated by $x_1 - a\,(\mathbf{1}.\mathbf{M}), y_1 - a\,(\mathbf{2}\cdot\mathbf{M}), z_1 - a\,(\mathbf{3}\cdot\mathbf{M}), x_2 - b\,(\mathbf{4}\cdot\mathbf{M}), \ldots, z_3 - c\,(\mathbf{9}\cdot\mathbf{M})$. Then the polynomials in I_3 are precisely the elements of $S \cap \tilde{I_3}$ where $S = \mathbb{R}[x_i, y_i, z_i]$. The substitution σ can naturally be extended to R (fixing on $a, \ldots, M_4$), so that $\forall r \in \tilde{I_3}$, $\sigma(r) = 0$. We note that a polynomial f is in I_3 if and only if each of its homogeneous components are in I_3. Therefore, let us consider a polynomial $f \in S$ which is homogeneous of degree d_1 in m_1 (resp. d_2 in m_2 , d_3 in m_3). As we have

$$\begin{aligned} f(m_1, m_2, m_3) &= f(a\,(\mathbf{C}_1\,\mathbf{M}), b\,(\mathbf{C}_2\,\mathbf{M}), a\,(\mathbf{C}_1\,\mathbf{M})) + r \\ &= a^{d_1} b^{d_2} c^{d_3} \sigma(f) + r \end{aligned}$$

with $r \in \tilde{I_3}$, it is easy to check that $\sigma(f) = 0$ if and only if $f \in \tilde{I_3}$. So in order to get $I_3 = S \cap \tilde{I_3}$, we compute the Gröbner basis of $\tilde{I_3}$ with a product order on two blocks of variables: $[x_1, \ldots, z_3]$, $[a, \ldots, M_4]$, which "eliminates" the last set of variables. The polynomial of S in this Gröbner basis form a Gröbner basis of the ideal I_3 (see [3]). This computation has been repeated for several random configurations of cameras (we take random matrices of projection), using the computer algebra systems Macaulay and Maple. It yields an ideal which is generated by the 3 bilinearities and one trilinearity (for instance $T_{1,2,3,5}$). More generally we can prove the following results:

The ideal of functions that vanish on triples of points (m_1, m_2, m_3) *in correspondence is generated by*

- *the three bilinearities* $F_{1,2}, F_{1,3}, F_{2,3}$,
- *any of the trilinearities (ie.* $T_{1,2,3,5}$*).*

Computing the dimension of V_3 yields 3 which is also the dimension of $\mathbb{P}^3$.

One consequence of this result is that any other trilinearity that vanishes on the images is a linear combination of $x_3F_{1,2}, y_3F_{1,2}, z_3F_{1,2}, x_2F_{1,3}, \ldots, z_1F_{2,3}, T_{1,2,3,5}$. This is true, in particular of the trilinearities reported by Hartley [9] and Shashua [13]. We can in fact show that the trilinearities introduced by Hartley, noted $THS_{i,l}$, are precisely (up to a sign) some of our previous trilinearities $T_{i,a,b,c}$.

Let J be the ideal of S generated by the bilinear polynomials $F_{1,2}, F_{1,3}, F_{2,3}$. Let $f(m_1, m_2, m_3) \in I_3$ be a homogeneous polynomial of degree d_1 in m_1 (resp. d_2 in m_2 , d_3 in m_3). We can also rewrite it as $(\mathbf{L}\cdot\mathbf{m}_1)$ where $\mathbf{L}$ is 3-dimensional vector, whose coefficients are homogeneous of degree $d_1 - 1$ in m_1, d_2 in m_2 , d_3 in m_3. Similarly, we denote $F_{1,2}(\mathbf{m}_1, \mathbf{m}_2) = (\mathbf{L}_2 \cdot \mathbf{m}_1)$, $F_{1,3}(\mathbf{m}_1, \mathbf{m}_3) = (\mathbf{L}_3 \cdot \mathbf{m}_1)$, where $\mathbf{L}_2$ (resp. $\mathbf{L}_3$) is a 3-dimensional vector linear in m_2 (resp. m_3). As $f \in I_3$, the system

$$\begin{cases} (\mathbf{L} \cdot \mathbf{m}_1) = 0 \\ (\mathbf{L}_2 \cdot \mathbf{m}_1) = 0 \\ (\mathbf{L}_3 \cdot \mathbf{m}_1) = 0 \end{cases}$$

has a solution in V_3. Consequently, the determinant $D = \det(\mathbf{L}, \mathbf{L}_2, \mathbf{L}_3)$ vanishes in V_3. Let $\mathbf{U}$ be any 3-dimensional vector. By Cramer's rule, we have

$$\det(\mathbf{L}, \mathbf{L}_2, \mathbf{L}_3)\,(\mathbf{U} \cdot \mathbf{m}_1) - \det(\mathbf{L}, \mathbf{L}_2, \mathbf{U})\,(\mathbf{L}_3 \cdot \mathbf{m}_1) + \det(\mathbf{L}, \mathbf{L}_3, \mathbf{U})\,(\mathbf{L}_2 \cdot \mathbf{m}_1) + \det(\mathbf{L}_2, \mathbf{L}_3, \mathbf{U})\,(\mathbf{L} \cdot \mathbf{m}_1) = 0$$

Let us denote by Δ the ideal generated by the coefficients of $\mathbf{U}$ in $\det(\mathbf{L}_2, \mathbf{L}_3, \mathbf{U})$. According to the previous relation, we have

$$\Delta \cdot f \in (F_{1,2}, F_{1,3}, D).$$

If $d_1 = 1$, then D depends only on $(\mathbf{m}_2, \mathbf{m}_3)$ and vanishes for points in correspondence. Hence, it must be divisible by $F_{2,3}$. By induction, we can prove that

$$\Delta^{d_1} \cdot f \in (F_{1,2}, F_{1,3}, F_{2,3})$$

Consequently, if we are outside the variety defined by Δ (ie. if $\mathbf{F}_{1,2} \cdot \mathbf{m}_2$ and $\mathbf{F}_{1,3} \cdot \mathbf{m}_3$ are not linearly dependent) then $(m_1, m_2, m_3) \in V_3$ if and only if $F_{1,2}(m_1, m_2) = F_{1,3}(m_1, m_3) = F_{2,3}(m_2, m_3) = 0$. If the cameras are generic (optical centers not collinear), then the epipolar lines $\mathbf{F}_{1,2} \cdot \mathbf{m}_2$ and $\mathbf{F}_{1,3} \cdot \mathbf{m}_3$ are linearly dependent if and only if M is in the trifocal plane. Apart from this degenerate case, V_3 can be defined by the bilinearities $F_{1,2}, F_{1,3}, F_{2,3}$.

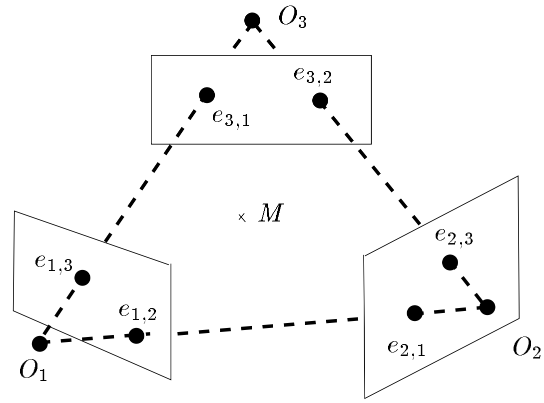
However, the trilinearities are necessary to describe completely the variety V_3, precisely in the degenerate cases.

5.3 Applications

If the trilinearities were only useful to recover the epipolar geometry, then one may wonder why bother. We show now how they are in fact more powerful than the fundamental matrixes:

M in the trifocal plane

In this case, the images m_1, m_2, m_3 are the trifocal lines $(e_{i,i+1}, e_{i,i+2})$, $i = 1, 2, 3$. Consequently, the epipolar m_1 and m_2 are identical in the third camera. They cannot be used to predict m_3.

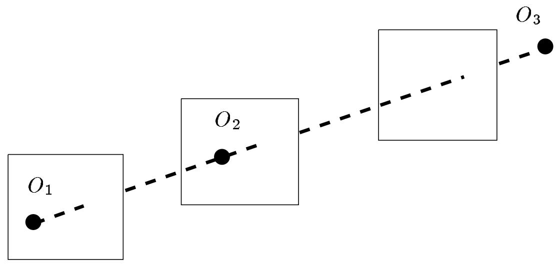


However, substituting $\mathbf{m}_1 = a_1\,\mathbf{e}_{1,2} + b_1\,\mathbf{e}_{1,3}$, $\mathbf{m}_2 = a_2\,\mathbf{e}_{2,1} + b_2\,\mathbf{e}_{2,3}$, $\mathbf{m}_3 = a_3\,\mathbf{e}_{3,1} + b_3\,\mathbf{e}_{3,2}$ in the trilinearity $T_{1,2,3,5}$ yields a non-zero trilinear polynomial in $(a_1, b_1), \ldots, (a_3, b_3)$. This polynomial can be used to construct the point m_3 knowing m_1, m_2.

In this case, we are in an analog situation than in the previous sections, except that we projection a point of the trifocal plane on lines in the retinal planes (it corresponds to a map from $\mathbb{P}^2$ to $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$). By symbolic computation, we can show that the ideal defining V_3 when M is in the trifocal plane is generated by one equation (the corresponding variety is of codimension 1), which arises from a trilinearity of $\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2$. Therefore, any trilinearity is "equivalent" on this degenerate configuration to the trilinearity $T_{1,2,3,5}$, for example. This explains why the ideal I_3 is generated by the bilinearities and only *one* trilinearity.

The optical centers aligned

In this configuration, the epipoles $e_{i,i-1}$ and $e_{i,i+1}$ are identical and the epipolar lines of two corresponding points m_1 and m_2 in the third image are always identical. The corresponding point m_3 is thus undefined on this epipolar line.



This configuration is equivalent to a configuration where the planes $\mathbf{1}, \mathbf{4}, \mathbf{7}$ are linearly dependent. Just as in the previous case, substituting $\mathbf{7} = u\,\mathbf{1} + v\,\mathbf{4}$, we check that the trilinear relations are not identically 0 and can be used to compute the position of m_3 with m_1 and m_3.

6 Quadrilinear constraints

In this section, we consider 4 cameras and relations where one row R_i is arising from each camera. This yields 16 quadrilinear relations. One of them built from the first, third, fifth, and seventh rows, which we note $Q_{1,3,5,7}$, in agreement with our notation for trilinearities, is given by the following expression:

$$\begin{aligned} & x_1\,x_2\,x_3\,x_4\,[2,5,8,11] - x_1\,x_2\,x_3\,y_4\,[2,5,8,10] \\ & - x_1\,x_2\,y_3\,x_4\,[2,5,7,11] + x_1\,x_2\,y_3\,y_4\,[2,5,7,10] \\ & - x_1\,y_2\,x_3\,x_4\,[2,4,8,11] + x_1\,y_2\,x_3\,y_4\,[2,4,8,10] \\ & + x_1\,y_2\,y_3\,x_4\,[2,4,7,11] - x_1\,y_2\,y_3\,y_4\,[2,4,7,10] \\ & - y_1\,x_2\,x_3\,x_4\,[1,5,8,11] + y_1\,x_2\,x_3\,y_4\,[1,5,8,10] \\ & + y_1\,x_2\,y_3\,x_4\,[1,5,7,11] - y_1\,x_2\,y_3\,y_4\,[1,5,7,10] \\ & + y_1\,y_2\,x_3\,x_4\,[1,4,8,11] - y_1\,y_2\,x_3\,y_4\,[1,4,8,10] \\ & - y_1\,y_2\,y_3\,x_4\,[1,4,7,11] + y_1\,y_2\,y_3\,y_4\,[1,4,7,10] \end{aligned}$$

Other examples of *quadrilinear* relations have been reported by Triggs [16]. The geometrical interpretation of this computation is as follows: $x_1\mathbf{2} - y_1\mathbf{1}$, ... represent lines in the images, which are the intersections of the retinal planes with planes through the optical centers of the cameras in $\mathbb{P}^3$. The vanishing of the quadrilinear polynomial is just the condition that these four planes have a common point in $\mathbb{P}^3$.

Here again, we can consider the quadruples of points $(m_1, \ldots, m_4)$ which correspond to the image of a same point $M \in \mathbb{P}^3$. This set of quadruples V_4 is an algebraic variety that we describe by ideal the I_4 of functions in $\mathbb{R}[x_i, y_i, z_i]$ that vanish on V_4. An explicit computation (done for a random configuration of cameras) shows that *the ideal I_4 defining V_4 is generated by*

- *the 6 bilinearities corresponding to pairs of cameras,*
- *the 4 trilinearities corresponding to triples of cameras.*

A surprising fact is that no quadrilinearity is needed for the description of V_4. In other words, the quadrilinearities are linear combinations of quadrilinearities obtain from the bilinearities and the trilinearities. *If we have 4 cameras or more, no more information will be available than if we consider any subset of 3 cameras among them*

The variety V_4 is also of dimension 3, as expected.

7 Conclusions

We have shown that the correspondences between the images of a single 3-D point in N cameras can be described by three types of relations between the coordinates of the image points. These relations fall into three classes of which only the first two are sufficient since all elements in the third one are algebraically dependent of elements in the first two. The coefficients of these relations are intrinsic quantities, independent of the referential. They have been shown to be 4×4 determinants of the row vectors of the perspective projection matrixes of the cameras.

We have shown how two trilinear relations allow us to recover the fundamental matrices and that the trilinear relations are useful in some degenerate cases of practical importance where the bilinear relations cannot be used for prediction.

For 3 cameras, we focus on the algebraic variety V_3 of points $(m_1, m_2, \ldots)$ in correspondence and give an explicit description of the ideal I_3 defining V_3. More precisely, we show that this ideal is generated by the 3 bilinearities and one trilinearity $T_{1,2,3,5}$. This trilinearity is necessary to handle correctly the degenerated cases. This approach also applies to 4 cameras. However, the ideal defining the quadruples in correspondence is defined by the bilinearities and the trilinearities. No quadrilinearity is required. This shows that no more information will be available if we consider 4 cameras or more.

This has been achieved by symbolic computations in projective geometry using the Grassmann-Cayley algebra formalism, showing the advantage of using such tools for problems in Computer Vision.

Acknowledgments

The authors gratefully acknowledge the contributions of Didier Bondyfalat and Stéphane Laveau.

References

[1] M. Barnabei, A. Brini, and G.C. Rota. On the exterior calculus of invariant theory. *J. of Algebra*, 96:p 120–160, 1985.

[2] Stefan Carlsson. Multiple image invariance using the double algebra. In Joseph L. Mundy, Andrew Zissermann, and David Forsyth, editors, *Applications of Invariance in Computer Vision*, volume 825 of *Lecture Notes in Computer Science.* Springer-Verlag, 1994.

[3] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra.* Undergraduate Texts in Mathematics. Springer, 1992.

[4] C. DeConcini, D. Eisenbud, and C. Procesi. Hodge algebras. *Astérisque*, 91, 1982.

[5] P. Doubilet, G. C. Rota, and J. Stein. Foundations of Combinatorics IX: Combinatorial methods in Invariant Theory. *Studies in Appl. Math.*, 53:185–216, 1974.

[6] Olivier Faugeras and Bernard Mourrain. On the geometry and algebra of the point and line correspondences between n images. In *Proceedings of the ICCV95*, Boston, MA, June 1995. IEEE Computer Society Press. To appear.

[7] Olivier Faugeras and Bernard Mourrain. On the geometry and algebra of the point and line correspondences between n images. In Roger Mohr and Chengke Wu, editors, *Proceedings of Europe-China Workshop on Geometrical Modeling and Invariants for Computer Vision*, pages 102–109. Xidan University Press, 1995.

[8] Olivier Faugeras and Luc Robert. What can two images tell us about a third one? *The International Journal of Computer Vision*, 1994. To appear. Also INRIA Technical report 2018.

[9] Richard Hartley. Lines and points in three views-an integrated approach. In *Proc. Image Understanding Workshop, Monterey, California*, November 1994.

[10] Quang-Tuan Luong. *Matrice Fondamentale et Calibration Visuelle sur l'Environnement-Vers une plus grande autonomie des systèmes robotiques.* PhD thesis, Université de Paris-Sud, Centre d'Orsay, December 1992.

[11] Quang-Tuan Luong, Rachid Deriche, Olivier Faugeras, and Théodore Papadopoulo. On Determining the Fundamental Matrix: Analysis of Different Methods and Experimental Results. Technical Report 1894, INRIA, 1993.

[12] Quang-Tuan Luong and Olivier Faugeras. The Fundamental matrix: theory, algorithms, and stability analysis. *The International Journal of Computer Vision*, 1994. To appear.

[13] Amnon Shashua. Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994. In press.

[14] Amnon Shashua and Michael Werman. Fundamental tensor: On the geometry of three perspective views. Technical report, The Hebrew University of Jerusalem, September 1994.

[15] Bernd Sturmfels. *Algorithms in Invariant Theory.* Springer-Verlag Wien New york, 1993.

[16] Bill Triggs. The Geometry of Projective Reconstruction I: Matching Constraints and the Joint Image. Technical report, LIFIA, INRIA Rhône Alpes, November 1994.

[17] H. Weyl. *The Classical Groups, their invariants and representations.* Princeton University Press, 1939.

[18] Andrew Zisserman and Steven Maybank. A Case Against Epipolar Geometry. In Joseph L. Mundy, Andrew Zissermann, and David Forsyth, editors, *Applications of Invariance in Computer Vision*, volume 825 of *Lecture Notes in Computer Science.* Springer-Verlag, 1994.