New Geometric Insight on the Trifocal Tensor and its Constraints

Andrea Dell'Acqua, Claudio Landi, Augusto Sarti, Stefano Tubaro

Dip. di Elettronica e Informazione – Politecnico di Milano
Piazza Leonardo da Vinci 32, I-20133 Milano – Italy
Email: dellacqu/landi/sarti/tubaro@elet.polimi.it

Abstract

The literature is rich with descriptions of the possible multi-view relations and their use in computer vision. Some of them arise from a purely geometric approach, some others are derived in more of an algebraic fashion. In this paper we show that, with a re-definition of the involved objects and some additional geometric tools, we can come to a more intuitive understanding of some of the most traditionally employed tools in computer vision. In particular, by re-defining camera models, incidence relations and duality we provide an alternative geometric interpretation of the tri-linearities, the trifocal tensor and the related constraints.

1 Introduction

A lot has been said about relations between views and a great deal of effort has been spent in trying to find a self-consistent, evocative and simple notation that would keep all such relations at hand and under control. A complete account of all relations that arise from multi-view projective geometry can be found in [1]. Some authors (e.g. [4]), however, have felt the need of seeking additional algebraic properties to help formulate multi-view constraints and relations in a more complete and unified fashion. An interesting work that follows an entirely algebraic approach is able to describe multi-view relations in the form of rank conditions on a properly defined matrix [6]. One way to retain the effectiveness of algebraic solutions without having to give up the evocative power of geometry is to make a productive use of tools of Geometric Algebra (GA) [5], which is a high-level language that deals directly with geometric objects, and not just with the coordinates of their representation in some reference frame. GA has also been effectively used in computer vision applications [2, 3].

In this paper we propose a unified treatment of the properties of point and line correspondences between n views, with particular emphasis on the three-view case. By following an entirely geometric approach we derive a complete set of constraints that characterise an admissible trifocal tensor and the geometric relationships that they imply, all from a novel perspective. Duality is one of the key concepts used throughout this paper, as it allows us to describe geometric constraints between different objects (points, lines and planes) and their projections, all in the same space. Other central tools are the bundles of lines and planes, which are used to model the correspondence relationships between projections of points and lines. All such concepts are typical tools of GA, therefore we will give a brief summary of them in the next Section.

2 A quick overview on GA tools

Basic GA elements are scalars and vectors in the usual sense. In addition to them, we have higher-grade elements such as bivectors, trivectors, etc., which are the result of a repeated application of the outer product operator. The most general elements of GA are, in fact, combinations of multivectors of various grades, and the basic operators between such elements are the (grade-increasing) outer product ($\wedge$) and the (grade-decreasing) inner product ($\cdot$) [5]. Such operators generalise the vector product and the scalar product, respectively, as they act on subspaces instead of just vectors. The fundamental operator in GA is the geometric product, which is a combination of outer and inner product and provides GA with a rich algebraic structure, as it is invertible [5].

In the GA projective space a point is identified by a 4D homogeneous vector. The bivector $A \wedge B$ represents the line passing through the points A and B, except for when the points concide, in which case the outer product returns the scalar 0. The trivector $A \wedge B \wedge C$ identifies the plane passing through the three points A, B and C, unless the points are aligned, in which case the outer product is the scalar 0. The outer product of more than four vectors is always equal to the scalar 0, as five vectors are always linearly dependent in a 4D space.

Duality is a powerful tool because the GA space is closed with respect to it. The dual of an element is computed by using a special element of the GA space, called *pseudoscalar*. Given an orthonormal basis $\{e_1,\ e_2,\ e_3,\ e_4\}$, the pseudoscalar of a projective GA space is defined as $I_4 = e_1 \wedge e_2 \wedge e_3 \wedge e_4$, and its inverse is equal to itself $I_4^{-1} = I_4$. In general, given 4 vectors X_i, $i = 1, \ldots, 4$, their outer product is $X_1 \wedge X_2 \wedge X_3 \wedge X_4 = \alpha I_4$, where $\alpha = |X_1 X_2 X_3 X_4|$ is the determinant of the matrix whose columns are the coordinates of the 4 vectors. The geometric division of an element A by the pseudoscalar I_4 is called the *dual* of A, and is denoted by $A^* = A I_4^{-1}$. If we think of A as a subspace, then its dual becomes the "complementary" space of A in the 4D space. Another useful operator is the *dual bracket* $[[\cdot]]$, which acts on a set of GA elements and returns the dual of the outer product between them. A commonly adopted notation for this operator is $[[A_1\ A_2\ A_3\ A_4]] = (A_1 \wedge A_2 \wedge A_3 \wedge A_4)^*$. It is through the dual bracket that we define the reciprocal frame (dual frame) associated to a generic vector basis X_i, $i = 1, \ldots, 4$, as the basis $\overline{X}_{i_1} = [[X_{i_2} X_{i_3} X_{i_4}]]$ (where all subscripts are different from each other), which satisfies the relationships $X_i^{(1)} \cdot \overline{X}_j^{(1)} \simeq \delta_{ij}$, where "$\simeq$" denotes proportionality. Since the trivector $X_{i_2} \wedge X_{i_3} \wedge X_{i_4}$ identifies a plane Π_{i_1}, the reciprocal frame for vectors can be thought of as the dual of the plane basis $\overline{X}_i \simeq \Pi_i^*$. The reciprocal frame associated to a bivector (line) basis L_i can be similarly defined [5].

Given two subspaces A and B their "join" $J = A \cup B$ is defined as the smallest subspace containing them both. Conversely, their "meet" $A \vee B$ is the largest common subspace of A and B. The meet operator can be defined in terms of the join as $A \vee B = (AJ^{-1} \wedge BJ^{-1})J$. When the join is between two planes Π_1 and Π_2, the result is a pseudoscalar. In this case their line of intersection is $\Pi_1 \vee \Pi_2 = [[\Pi_1^* \Pi_2^*]] = \Pi_1^* \cdot \Pi_2$. With the help of the meet operator, any line can be dually defined as the intersection of two planes. For three lines of the basis we obtain $L_{i_1} = X_{i_2} \wedge X_{i_3} \simeq \Pi_{i_1} \vee \Pi_4$ (where $i_1, i_2, i_3 = 1, 2, 3$ are all different from each other). Similarly, for the other lines we obtain $X_{i_1} \wedge X_4 \simeq \Pi_{i_2} \vee \Pi_{i_3}$. The dual representation of a point is the intersection of three planes: $X_{i_1} = \Pi_{i_2} \vee \Pi_{i_3} \vee \Pi_4$.

The adopted camera model is here a simple pinhole model, which is completely specified by its optical center and three points on the image plane. Any such set of four independent vectors thus corresponds to a pin-hole camera. A good choice for these vectors is given, for example, by the optical center $C = X_4$, the principal point X_3, and the directions X_1 and X_2 of the $\mathbf{x}$ and $\mathbf{y}$ axes, respectively, which are points at infinity on the image plane. Another way of defining the camera is with four independent planes: the image plane and three more planes, which can be conveniently chosen in such a way that the intersection of any pair of them will give one of the previously-defined axes. The planes just described are the dual of the reciprocal vector frame. The projection of a point on the camera is the intersection between optical ray (line joining optical center and point) and image plane. Given a generic point A, the previously-defined camera basis allows us to express its image projection using the inner product $A^{(\text{cam})} = \sum_{i=1}^{3} (A \cdot \overline{X}_i) X_i$. Similarly, the projection of a line V on the camera can be written as $V^{(\text{cam})} = \sum_{i=1}^{3} (V \cdot \overline{L}_i) L_i$.

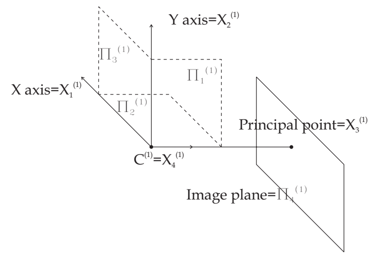


Figure 1: Vector camera frame X_i and plane camera frame Π_j.

3 Bundles of lines and planes

Given a point in 3D space, the corresponding optical rays are bound to be incident and will form a "bundle of lines". As a line is given by the intersection of two planes, this incidence condition can be replaced by the condition that all such planes will form a bundle. As a consequence, the matrix made of the normals to the planes has rank three, meaning that any of its 4x4 minors will have zero determinant. Notice that the fact that a minor has zero determinant corresponds to saying that the planes whose normals are the rows of that minor will form a bundle.

In the GA framework the normal to a plane is the dual $\Phi_j^* = (\Phi_j)\, I_4^{-1}$ of the plane itself. The incidence constraint can thus be expressed as rank $M = 3$, where M is the matrix made of the normals to N pairs of planes that define the N rays. This implies the all the 4×4 minors of M have zero determinant. We recall that this condition on a 4×4 minor corresponds to a "bundling" condition on a subset of 4 of the above planes, which can be written as $[[\Phi_1^*\ \Phi_2^*\ \Phi_3^*\ \Phi_4^*]] = 0$, which corresponds to $(\Phi_1 \vee \Phi_2) \wedge (\Phi_3 \vee \Phi_4) = 0$. In other words, the lines of intersection corresponding to two pairs of plains are incident.

Let us now turn our attention to the projection of a line onto the cameras. The back-projection planes of the line form a pencil of planes, therefore the matrix of the normal vectors has rank 2. This means that $a\Phi_i^* + b\Phi_j^* + c\Phi_k^* = 0$, i.e. any three planes of them are linearly dependent. This implies $(a\Phi_i^* + b\Phi_j^* + c\Phi_k^*) \wedge \Phi_k^* = 0$, and therefore $\Phi_i \vee \Phi_k \simeq \Phi_j \vee \Phi_k \simeq \Phi_i \vee \Phi_j$, which means that the lines of intersection are coincident.

We can now consider the hybrid case of lines and points. We assume that for some of the views we have information on the projection of incident lines, and for the others we have the image locations of of the meeting point of the same lines. This situation can be described by a matrix of the form

$$M = (\phi_{1,1}^*\ \phi_{2,1}^*\ \phi_{1,2}^*\ \phi_{2,2}^* \cdots | \phi_1^*\ \phi_2^* \cdots)^T \qquad (1)$$

which collects the normals to a set of planes: those on the left of the separator define backprojection rays associated to the meeting point, while those on the right are the backprojection planes generated by the lines. Notice that this is a rank-3 matrix, as it describes a bundle of planes. As the choice of the planes associated to the back-projection ray is arbitrary, we can use pairs of planes that encode information about the position and orientation of cameras. Let $p^{(i)} = m_1 X_1^{(i)} + m_2 X_2^{(i)} + m_3 X_3^{(i)} + X_4^{(i)}$ be the projection of the point onto the i-th camera. The two planes defined by the back-projection ray and the first two axes of the camera are

$$\Phi_1^{(i)} = m_2 \Pi_3^{(i)} - m_3 \Pi_2^{(i)}, \Phi_2^{(i)} = -m_1 \Pi_3^{(i)} + m_3 \Pi_1^{(i)}\ . \qquad (2)$$

We could have also used the plane $\Phi_3^{(i)}$ that contains the third axis, therefore there are three possible pairs of planes for any ray. Since these planes form a pencil, any pair of them linearly depends on the other pairs.

In the two-camera case the two incident backprojection rays originate a bundle of four planes. The rank condition on the matrix of the four normals of the planes is thus equivalent to the geometric constraint $(\Phi_1^{(1)} \vee \Phi_2^{(1)}) \wedge (\Phi_1^{(2)} \vee \Phi_2^{(2)}) = 0$, which tells us that the two rays given by the intersection of the two pairs of planes are incident. The same constraint can be expressed using the fundamental matrix. Assuming $F_{i_1 j_1} = [[\Pi_{i_2}^{*(1)} \Pi_{i_3}^{*(1)} \Pi_{j_2}^{*(1)} \Pi_{j_3}^{*(1)}]]$, we derive $\sum m_i^{(1)} F_{ij} m_j^{(2)} = 0$ where $m_i^{(1)}$ and $m_j^{(2)}$ are the coordinates of the projections of the point on the two cameras.

3.1 Feature correspondences in 3 views

Let us now add a third camera and consider the projections of a point onto the three image planes. The bundle constraint on the matrix built with the normals of the 6 planes that define the 3 back-projection rays is given by rank$(M) \leq 3$, where M is defined like in eq. (1). The number of 4x4 minors taken from the 6 lines of M is $\binom{6}{4} = 15$. The minors involving 4 planes associated to only two cameras correspond to the already described bilinear relations; they are M_{1234}, M_{3456}, M_{1256}. There are 12 more minors that can be divided in 3 groups, any of which contains those minors that include both planes from one camera. The constraint on a minor of the first group is, for example, $\det(M_{1235}) = [[\Phi_1^{*(1)} \Phi_2^{*(1)} \Phi_1^{*(2)} \Phi_1^{*(3)}]] = 0$. From the properties of the dual bracket, it follows that $\det(M_{1235}) = 0$ is equivalent to $\sum \epsilon_{j_1 j_2} \epsilon_{k_1 k_2} m_i^{(1)} T_{i j_1 k_1} m_{j_2}^{(2)} m_{k_2}^{(3)} = 0$ where

$$T_{i_1 jk} = [[\Pi_{i_2}^{*(1)} \Pi_{i_3}^{*(1)} \Pi_j^{*(2)} \Pi_k^{*(3)}]],\ i_1 \neq i_2 \neq i_3 \qquad (3)$$

and $j_1 \neq j_2$, $j_1, j_2 \in \{2,3\}$, the same for k_1, k_2 and $\epsilon_{j_1 j_2} = +1$ if $(j_1, j_2) = (2,3)$, $\epsilon_{j_1 j_2} = -1$ for $(j_1, j_2) = (3,2)$. This equation involves the sum of 12 homogeneous third-grade terms, and is referred to as a trilinearity. That condition on the determinant of a minor is thus equivalent to setting a trilinearity to zero.

The matrix M comes from an arbitrary choice of the planes that define the back-projection rays. For any ray there are three possible pairs of planes (see 2), therefore we obtain 9 possible trilinearities for a given pair of planes on the first camera. Since there

Case	Constraint
PPL	$\sum \epsilon_{j_1 j_2} m_i^{(1)} T_{i j_1 k} m_{j_2}^{(2)} l_k^{(3)} = 0$
PLL	$\sum m_i^{(1)} T_{ijk} l_j^{(2)} l_k^{(3)} = 0$
LLL	$l_i^{(1)} = \sum_{j,k} l_j^{(2)} l_k^{(3)} T_{ijk}$

Table 1: Structure of M for hybrid cases. The PPL case has a rank3 constraint with 4 trilinearities. The PLL case, again is rank 3, but the number of trilinearities is down to 1. Finally, the LLL case is rank 2, with 3 trilinearities.

are at most 2 independent planes for a single ray, we derive 4 independent trilinearities per camera, which correspond to those obtained with the chosen matrix M.

In the hybrid case the matrix M is made of $n_\phi = 2 \times n_P + n_L$ planes, where n_L is the number of line projections and n_P is the number of projections of the incidence point. The number of 4x4 minors is then $\binom{n_\phi}{4}$. Table 1 collects some details on M and its constraint.

For the point-point-line (PPL) and line-line-line (LLL) cases, one of the possible constraints is shown in Table 1. The rank for the line-line-line (LLL) case is two because the planes form a pencil.

The above incidence equations include those described in [1]. Our PPL and PLL equations, however, can describe situations where the projection of the point or the line is known onto any of the three cameras, with the same degrees of difficulty. Furthermore, our presentation also shows that the PLL equation is satisfied by the projections of two incident lines, and not just by two projections of the same line. Methods for transferring lines and points from two cameras to a third one are also obtained geometrically in [2] and [4].

4 The trifocal tensor

While the trilinearities depend on the structure of the scene, the trifocal tensor encodes only the spatial relationship between the cameras. As the tensor estimated from image coordinates of lines and points is affected by noise, it does not provide valid geometric relations between cameras. It is thus necessary to determine the constraints that describe the structure of a consistent trifocal tensor in order to project its estimate on the manifold of the admissible trifocal tensors.

The $3 \times 3 \times 3$ elements T_{ijk} of the trifocal tensor [1] can be grouped into three 3×3 "slices" $T_{i\cdot\cdot}$, $i = 1, 2, 3$. If $\mathbf{x}^{(2)}$ is the vector of the image coordinates on the second camera and $\mathbf{x}_\times^{(2)}$ is its skew-symmetric matrix representation (for the computation of the vector product by $\mathbf{x}^{(2)}$), then the incidence equation in the algebraic form of [1] for the PPP case is

$$f \mathbf{x}_\times^{(2)} (\sum x_i^{(1)} T_i) \mathbf{x}_\times^{(3)} = \mathbf{0}_{3\times 3} = \begin{pmatrix} M_{1235} & M_{1236} & M_{123*} \\ M_{1245} & M_{1246} & M_{124*} \\ M_{12*5} & M_{12*6} & M_{12**} \end{pmatrix}$$

where M_{ijkl} is the PPP trilinearity generated by the minor of M corresponding to the four lines i, j, k, l. In addition to the four independent trilinearities that appear on the upper-left 2×2 submatrix of the above eq., there are five more trilinearities that are linearly dependent on them. Such trilinearities are marked with a "$*$" subscript, to indicate that the corresponding line refers to a different matrix M defined using the planes $\Phi_3^{(i)}$ that were not used to build M (see eq. (2)). As we can see, our GA-based derivation automatically produces the minimum possible number of trilinearities, even for the hybrid cases.

4.1 Dual interpretations of the tensor

The elements of the trifocal tensor can be seen as the coordinates of the projections of points and lines that depend only on the internal and the external camera parameters. From the definition of the trifocal tensor it follows that the vector $T_{\cdot jk}$ (a "bar" that runs through all "slices" of the tensor) contains the coordinates of the projection onto the first camera of the line of intersection $\Pi_j^{(2)} \vee \Pi_k^{(3)}$ between a plane on the second camera and a plane on the third camera. If we project $\Pi_j^{(2)} \vee \Pi_k^{(3)}$ onto the first camera we obtain $T_{\cdot jk}^{(1)} = \sum_i T_{ijk} L_i^{(1)}$.

Let us now consider the vectors $T_{ij\cdot}$, which are rows of a slice of the tensor. By expanding the dual bracket we obtain $T_{i_1 jk} = [[\Pi_{i_2}^{*(1)} \Pi_{i_3}^{*(1)} \Pi_j^{*(2)} \Pi_k^{*(3)}]] = -\Pi_k^{*(3)} \cdot (\Pi_j^{(2)} \vee (C^{(1)} \wedge X_{i_1}^{(1)}))$. The second term of the inner product is the intersection point $P = \Pi_j^{(2)} \vee (C^{(1)} \wedge X_{i_1}^{(1)})$ of the axis $C^{(1)} \wedge X_{i_1}^{(1)}$ of the first camera with a plane $\Pi_j^{(2)}$ of the second camera. The vectors $T_{ij\cdot}$ thus express the coordinates of the projection of point P onto the third camera. By projecting P on the first camera we thus obtain $T_{ij\cdot}^{(3)} = \sum_k T_{ijk} X_k^{(3)}$.

By following a similar procedure we find that the vectors $T_{i\cdot k}$ represent the image coordinates of the point $\Pi_k^{(3)} \vee (C^{(1)} \wedge X_i^{(1)})$ on the second camera.

4.2 Tensors for a triplet of cameras

We will now show how to geometrically derive the tensors referred to the second and third camera from the the tensor related to the first camera. In order to obtain this result we use both interpretations of the trifocal tensor in the dual form ("line view" for the known tensor) and its specular form ("point view" for the unknown tensor). Since the tensor contains projective information, we must choose the elements on the two tensors so that they will represent coordinates of projections of lines and points onto the same camera. Let us consider, for example, the tensor $T^{(1)(2)(3)}$ referred to the first camera, and the unknown tensor $T^{(2)(3)(1)}$ referred to the second camera. The vectors $T_{\cdot jk}^{(1)(2)(3)}$ are the projections of the intersection line between the two planes $\Pi_j^{(2)}$ and $\Pi_k^{(3)}$ onto the first camera. The vectors $T_{p_1 q\cdot}^{(2)(3)(1)}$ are the projection on the first camera of the intersection points between the axis of the second camera $\Pi_{p_2}^{(2)} \vee \Pi_{p_3}^{(2)} = C^{(2)} \wedge X_{p_1}^{(2)}$ and the plane of the third camera $\Pi_q^{(3)}$. In order to derive such vectors, we must look at the first tensor and search for two lines that meet at the point $T_{p_1 q\cdot}^{(2)(3)(1)}$. The lines $\Pi_{p_2}^{(2)} \vee \Pi_q^{(3)}$ and $\Pi_{p_3}^{(2)} \vee \Pi_q^{(3)}$ satisfy this constraint because $(\Pi_{p_2}^{(2)} \vee \Pi_q^{(3)}) \vee (\Pi_{p_3}^{(2)} \vee \Pi_q^{(3)}) = \Pi_{p_2}^{(2)} \vee \Pi_{p_3}^{(2)} \vee \Pi_q^{(3)}$. Using these lines and points we obtain a relation between $T^{(1)(2)(3)}$ and $T^{(2)(3)(1)}$ of the form

$$\sum_r T_{p_1 q r}^{(2)(3)(1)} X_r^{(1)} = \sum_i T_{i p_2 q}^{(1)(2)(3)} L_i^{(1)} \vee \sum_i T_{i p_3 q}^{(1)(2)(3)} L_i^{(1)} .$$

This equation allows us to obtain the rows of the slices $T_{p\cdot\cdot}^{(2)(3)(1)}$, but their scaling turns out to be not uniform. It is, however, possible to use the specular approach, which considers the "line view" for the first tensor and "point view" for the unknown tensor. This way we obtain the columns of the slices of the tensor (and even the "bars" that pass through all slices) and obtain a uniformly scaled tensor.

5 Constraints on points

In this Section we present a derivation of the point-constraints that takes advantage of both the "point view" and "line view" of the tensor. Such constraints are collected and illustrated in Tables 2 and 3. We will see that such "tensor views" turn out to be equally powerful tools for the analysis of the properties of the trifocal tensor. Again, we will make an intensive use of the duality (and of the reciprocal frame) [2, 3]. A study of the tensor based on the intersections of planes is reported in [4].

5.1 Constraints on tensor slices (type 0)

From the former analysis it follows that the rows of the slices $T_{i\cdot\cdot}$ represent projections of points on the third camera. The three rows of the i-th slice correspond to the projections of three points $P_j = \Pi_j^{(2)} \vee (C^{(1)} \wedge X_i^{(1)})$, $j = 1, 2, 3$. Such points all lie on the axis of the first camera $C^{(1)} \wedge X_i^{(1)}$. This alignment must hold true for the projections of the points as well. Since the projections coincide with the vectors $T_{ij\cdot}$ and the alignment of three points can be expressed by setting their outer product to zero, it follows that $T_{i1\cdot}^{(3)} \wedge T_{i2\cdot}^{(3)} \wedge T_{i3\cdot}^{(3)} = 0$, where $T_{ij\cdot}^{(3)} = \sum_k T_{ijk} X_k^{(3)}$. The three projections are then linearly dependent, therefore the determinant of the matrices $T_{i\cdot\cdot}$ is zero. An alternative formulation of the same constraint can be obtained through a row decomposition of the slices $T_{i\cdot\cdot}$

$$|T_{i1\cdot} T_{i2\cdot} T_{i3\cdot}| = 0, \; i = 1, 2, 3. \qquad (4)$$

The right null space of the i-th slice of the tensor is the projection on the third camera of the axis $C^{(1)} \wedge X_i^{(1)}$. The same constraints can be derived for projections on the second camera (columns of slices). This time the left null spaces encode projections on the second camera of the axes of the first camera. A similar geometric illustration of these constraints can be found in [4] and [3].

5.2 Incidence constraints (type 1)

The null spaces of the three slices are related by another constraint. The three axes of the first camera $(C^{(1)} \wedge X_1^{(1)})$, $(C^{(1)} \wedge X_2^{(1)})$, $(C^{(1)} \wedge X_3^{(1)})$ are, in fact, incident at the optical center of the first camera, therefore their projections on the third camera meet at the projection of the optical center $C^{(1)}$ on the same camera, which is the epipole $E_{(3)(1)}$. As a consequence, the null spaces of the three matrices $T_{i\cdot\cdot}$ describe incident vectors, therefore the matrix whose rows correspond to these null spaces has zero determinant. Moreover, the null space of the above-defined matrix is the epipole $E_{(3)(1)}$. These geometric relationships link projections of points and lines on the image plane of the third camera, therefore it is possible to work directly in a 2D projective space. Here the element of maximum grade (pseudoscalar) is a trivector. In the case of the third camera, the trivector represents the image plane $\Pi_{\text{im}}^{(3)}$, therefore the dual of a generic image vector V is $[[V]] = V(\Pi_{\text{im}}^{(3)})^{-1} = V \cdot (\Pi_{\text{im}}^{(3)})^{-1}$. The meet of two lines is simply $L_1 \vee L_2 = [[L_1^* L_2^*]]$. As for the image points of Table 2, we can write

$$\begin{aligned} E^{(3)(1)} &= (F_1 \wedge F_2) \vee (G_1 \wedge G_2) \\ &\simeq [[F_1 F_2 G_1]] G_2 - [[F_1 F_2 G_2]] G_1 \ . \end{aligned}$$

Now we can force the epipole $E^{(3)(1)}$ to lie on the line $H_1 \wedge H_2$, therefore we must have $H_1 \wedge H_2 \wedge E^{(3)(1)} = 0$, which can be rewritten as

$$0 \simeq [[F_1 F_2 G_1]]\,[[H_1 H_2 G_2]] - [[F_1 F_2 G_2]]\,[[H_1 H_2 G_1]] \ .$$

On the projective plane, similarly to the 4D projective space, the dual of the outer product of three independent vectors is equal to the determinant of the matrix of such vectors. We can thus derive a constraint on the tensor of the form

$$\begin{aligned} 0 \;=\;& |T_{i_1 j_1 \cdot} T_{i_1 j_2 \cdot} T_{i_2 j_1 \cdot}| \; |T_{i_3 j_1 \cdot} T_{i_3 j_2 \cdot} T_{i_2 j_2 \cdot}| \\ & - |T_{i_1 j_1 \cdot} T_{i_1 j_2 \cdot} T_{i_2 j_2 \cdot}| \; |T_{i_3 j_1 \cdot} T_{i_3 j_2 \cdot} T_{i_2 j_1 \cdot}| \ . \end{aligned} \quad (5)$$

With reference to the “Type-1” figure of Table 2, since there are three points lying on any of the three incident lines, a total of $\binom{3}{2} \times \binom{3}{2} \times \binom{3}{2} = 27$ constraints can be found. Such constraints are not all independent. If the point of intersection of lines $G_1 \wedge G_2$ and $H_1 \wedge H_2$ lies on both lines $F_1 \wedge F_2$ and $F_2 \wedge F_3$, then it follows that the same point lies on the line $F_1 \wedge F_3$ as well. The number of independent constraints is thus $2 \times 2 \times 2 = 8$, any of which involves 18 elements of the tensor. Constraints of the same type could also be derived on the image plane of the second camera, obtaining 8 more constraints. These constraints correspond to the constraints on the null spaces of the tensor slices presented in [4]. Our analysis provides geometric insight for any step of the derivation and produces constraints in a form that is consistent with the other types of constraints.

5.3 Constr. on sum of lines (Type 2, 2b)

Constraints based on points are always obtained from the incidence of three lines, with the exception of eq. (4). With reference to Table 2, we can prove that the lines $L_1 = A \wedge B$, $L_2 = C \wedge D$, and the line resulting from the combination $(A \wedge B) + (C \wedge D)$, are all incident. A point lies on a line when the outer product of the two is zero, therefore this constraint can be written as $0 = (L_1 + L_2) \wedge ((A \wedge B) \vee (C \wedge D))$. Looking at Table 2, we can recognize three constraints of this type for any pair of axes of the first camera projected on the second one. Constraints of the same type can be derived for the third camera too, therefore we have a total of 9 constraints on the second camera plus additional 9 constraints on the third camera. These constraints link 12 elements of the trifocal tensor.

Also these constraints can be expressed in a form that involves only the elements of the tensor. With reference to Table 2 we can write $E^{(2)(1)} = (A \wedge B) \vee (C \wedge D)$, therefore the constraint becomes $(A \wedge D + C \wedge B) \wedge E^{(2)(1)} = 0$. Working in the 2D projective space of the image plane $\Pi_{\text{im}}^{(2)}$ we can derive the general expression (Type-2 constraints):

$$\begin{aligned} 0 \;=\;& \left|T_{i_1 \cdot k_2} T_{i_2 \cdot k_1} T_{i_1 \cdot k_1}\right| \; \left|T_{i_1 \cdot k_2} T_{i_2 \cdot k_1} T_{i_2 \cdot k_2}\right| \\ & - \left|T_{i_2 \cdot k_2} T_{i_1 \cdot k_1} T_{i_2 \cdot k_1}\right| \; \left|T_{i_2 \cdot k_2} T_{i_1 \cdot k_1} T_{i_1 \cdot k_2}\right| \ . \end{aligned} \quad (6)$$

Constraints of the same type can also be derived for the second image plane $\Pi_{\text{im}}^{(2)}$ [4].

The epipole $E^{(2)(1)}$ also coincides with the intersection of the two lines $G \wedge B$ and $H \wedge B$. Working in the image plane $\Pi_{IMM}^{(2)}$ we obtain (Type-2b constraints)

$$\begin{aligned} 0 \;=\;& \left|T_{i_1 \cdot k_2} T_{i_1 \cdot k_3} T_{i_2 \cdot k_1}\right| \; \left|T_{i_2 \cdot k_2} T_{i_2 \cdot k_3} T_{i_1 \cdot k_3}\right| \\ & - \left|T_{i_1 \cdot k_2} T_{i_1 \cdot k_3} T_{i_1 \cdot k_1}\right| \; \left|T_{i_2 \cdot k_2} T_{i_2 \cdot k_3} T_{i_1 \cdot k_2}\right| \ . \end{aligned} \quad (7)$$

The constraint 2b involves the 18 elements of the tensor, as shown in Table 2. With respect to similar derivations developed in more limited cases [3], the geometric procedure that we followed to derive eqs. (5), (6) and (7) is greatly simplified by the fact that we worked on the projective plane instead of the 4D projective space.

6 Constraints on lines

So far we have considered only the point interpretation of the trifocal tensor. Other constraints can be derived from the dual view of the trifocal tensor, based on projections of lines. These new constraints are derived following a dual pattern with respect to the former ones. The objects involved are now lines, instead of points. The incidence of three lines is replaced by the alignment of three points. While the lines were derived before as the outer product of points, the points are now obtained as intersections of lines. Duality is then achieved by specularly replacing the meet operator with the outer product.

The constraints on lines are derived in a consistent form with respect to the previous equations (5), (6) and (7). Some of the constraints on the tensor as a projection of lines can be found in [4] and [3, 2].

6.1 Alignment constraints (type 1)

The vectors $T_{\cdot jk}$ contain the coordinates of epipolar lines on the first camera. We can use this information in order to derive alignment relations between points defined as intersections of coplanar lines. With reference to Table 3, the points $Q_i = (C^{(2)} \wedge X^{(2)}_{j_1}) \vee \Pi^{(3)}_i$, $i = 1, 2, 3$ are the intersections of the axis $(C^{(2)} \wedge X^{(2)}_{j_1})$ of the second camera with the three planes $\Pi^{(3)}_i$ of the third camera, therefore they turn out to be aligned: $Q_1 \wedge Q_2 \wedge Q_3 = 0$. The axis has a dual representation in terms of meet of planes, $C^{(2)} \wedge X^{(2)}_{j_1} = \Pi^{(2)}_{j_2} \vee \Pi^{(2)}_{j_3}$, therefore we can write $Q_i = (\Pi^{(2)}_{j_2} \vee \Pi^{(2)}_{j_3}) \vee \Pi^{(3)}_i$. The same geometric relations must hold true for the projections of the points on a camera, for example the first. Since $T_{\cdot j_2 k}$ contains the coordinates of the projection of the bivector $\Pi^{(2)}_{j_2} \vee \Pi^{(3)}_k$ onto the first camera, the projection of the point Q_k results as $T^{(1)}_{\cdot j_2 k} \vee T^{(1)}_{\cdot j_3 k}$ where $T^{(1)}_{\cdot jk} = \sum_i T_{ijk} L^{(1)}_i$. The alignment constraint between projections can thus be expressed as

$$(T^{(1)}_{\cdot j_1 k_1} \vee T^{(1)}_{\cdot j_2 k_1}) \wedge (T^{(1)}_{\cdot j_1 k_2} \vee T^{(1)}_{\cdot j_2 k_2}) \wedge (T^{(1)}_{\cdot j_1 k_3} \vee T^{(1)}_{\cdot j_2 k_3}) = 0 \,. \quad (8)$$

This constraint involves 6 vectors $T^{(1)}_{\cdot jk}$ that depend on 3 elements of the tensor. The constraint then expresses a relation between 18 elements of the trifocal tensor. The same type of constraint holds true for all the three axes of the second camera, therefore we have 3 constraints like in eq. (8). Constraints of the same form can be applied to the axes $C^{(3)} \wedge X^{(3)}_k$ of the third camera as well, therefore we end up with a total of 6 constraints on the structure of the tensor.

We can derive the constraints as products of determinants by expressing all the relations in the image plane $\Pi^{(1)}_{\text{im}}$. The relation between the points Q_3, Q_2, Q_1 of the Table 3 becomes $0 = Q_3 \wedge Q_2 \wedge Q_1$, therefore

$$[[S^*_1 S^*_2 S^*_3]]\ [[S^*_5 S^*_6 S^*_4]] - [[S^*_1 S^*_2 S^*_4]]\ [[S^*_5 S^*_6 S^*_3]] = 0 \,.$$

Since the dual of a line in a plane coincides with its normal vector, which contains the projective coordinates of the line itself, the constraints from eq. (8) can be written as:

$$\begin{aligned} 0 \ = \ & \left|T_{\cdot j_1 k_1} T_{\cdot j_2 k_1} T_{\cdot j_1 k_2}\right| \ \left|T_{\cdot j_2 k_2} T_{\cdot j_1 k_3} T_{\cdot j_2 k_3}\right| \quad (9) \\ & - \left|T_{\cdot j_1 k_1} T_{\cdot j_2 k_1} T_{\cdot j_2 k_2}\right| \ \left|T_{\cdot j_1 k_2} T_{\cdot j_1 k_3} T_{\cdot j_2 k_3}\right| \,. \end{aligned}$$

Constraints of the same type exist for all the axes of the third camera, therefore we have a total of 6 alignment constraints [3].

6.2 Constr. on sum of points (types 2, 2b)

There are other alignment constraints that express more complex geometric relations. As the Type-2 constraint for points involves a sum of lines, the constraint derived here is obtained by forcing the alignment of three points, one of which is expressed as the sum of two other points.

With reference to Table 3, a linear combination of A and B is bound to lie on the bivector $A \wedge B$. In particular, we can prove that $E+F$ is a combination of A and B, therefore we have three aligned points and the constraint $A \wedge B \wedge (E+F) = 0$. Similarly, it is possible to show that $E - F$ is a combination of C and D and, therefore, it lies on the line $C \wedge D$. The constraint derived from this relation, however, is the same as the one obtained from $E + F$.

This constraint involves four vectors $T_{\cdot jk}$ (four lines) and, therefore, 12 elements of the tensor. For any axis of the second camera $X^{(2)}_4 \wedge X^{(2)}_j$ it is possible to derive 9 constraints plus 9 more for the three axes of the third camera $X^{(3)}_4 \wedge X^{(3)}_k$. Any constraint referred to the third camera, however, is equivalent to one of the constraints referred to the second one, therefore there are 9 independent constraints in total. Working on the image plane $\Pi^{(1)}_{\text{im}}$ we can derive for the axes of the second camera the Type-2 constraint

$$\begin{aligned} 0 \ = \ & \left|T_{\cdot j_1 k_1} T_{\cdot j_2 k_1} T_{\cdot j_1 k_2}\right| \ \left|T_{\cdot j_2 k_2} T_{\cdot j_1 k_2} T_{\cdot j_2 k_1}\right| \quad (10) \\ & - \left|T_{\cdot j_1 k_1} T_{\cdot j_2 k_1} T_{\cdot j_2 k_2}\right| \ \left|T_{\cdot j_1 k_2} T_{\cdot j_1 k_1} T_{\cdot j_2 k_2}\right| \,. \end{aligned}$$

A similar equation can be derived for the axes of the third camera. A geometric derivation of Type-2 constraints can also be found in [3] and, in algebraic form, in [4]. As done for Type-2b line constraints, we can now derive other alignment relations that involve more elements of the tensor. With reference to Table 3, the points A and B are the projections of the intersection points of the axis of the second camera $X^{(2)}_4 \wedge X^{(2)}_1$ with the planes $\Pi^{(3)}_{k_1}$ and $\Pi^{(3)}_{k_2}$. Another point H on the axis can be obtained as the intersection between the axis and the third plane $\Pi^{(3)}_{k_3}$ of the third camera

$$(X^{(2)}_4 \wedge X^{(2)}_{j_3}) \vee \Pi^{(3)}_{k_3} \overset{\text{proj. on the 1st cam.}}{\longrightarrow} T^{(1)}_{\cdot j_1 k_3} \vee T^{(1)}_{\cdot j_2 k_3} = H.$$

The points H, A, B and $E + F$ are thus aligned, therefore we have $H \wedge B \wedge (E+F) = 0$ and $A \wedge H \wedge (E+F) = 0$. Such constraints involve 6 lines (or 6 vectors $T^{(1)}_{\cdot jk}$), which account for 18 elements of the tensor. There are 3 constraints of this type for each axis of the second and of the third camera, therefore there exist a total of 18 alignment constraints.

The algebraic version of this constraint can be derived with a similar procedure to what used for eqs. (9) and (10)

$$0 = \left|T_{\cdot j_1 k_3} T_{\cdot j_2 k_3} T_{\cdot j_1 k_2}\right| \left|T_{\cdot j_2 k_2} T_{\cdot j_1 k_2} T_{\cdot j_2 k_1}\right| - \left|T_{\cdot j_1 k_3} T_{\cdot j_2 k_3} T_{\cdot j_1 k_2}\right| \left|T_{\cdot j_1 k_2} T_{\cdot j_1 k_1} T_{\cdot j_2 k_2}\right| . \quad (11)$$

These constraints are derived with a mixture between geometric and algebraic tools in [3]. In this paper we show that the eqs. (9), (10) and (11) can be geometrically derived by exploiting the duality between the constraints for lines and points.

Tables 2 and 3 collect a summary of the constraints found. Notice that there exist no Type-0 constraints for lines, as there are no incident lines directly encoded in the trifocal tensor.

7 Conclusions

In this paper we showed that by enriching the set of geometry tools at hand, it is possible to derive a wide variety of constraints in an elegant and unified fashion, which otherwise would be quite difficult to devise. We are currently working on a flexible camera tracker that incorporates all such constraints [7].

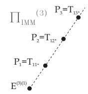
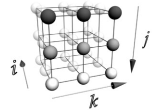
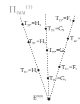
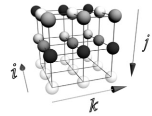
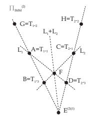
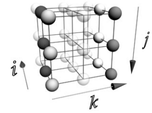
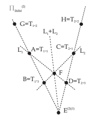
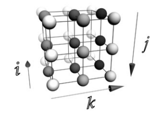
Constr. for points	**Tensor**
 Type 0: 3+3 constr.	 9 elements
 Type 1: 8+8 constr.	 18 elements
 Type 2: 9+9 constr.	 Type 2: 12 elements
 Type 2b: 9+9 constr.	 Type 2b: 18 elements

Table 2: Summary of constraints for points.

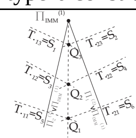
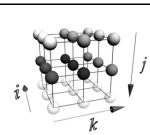
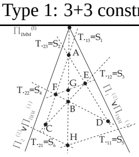
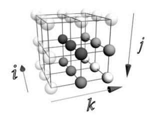
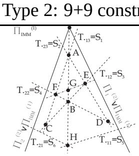
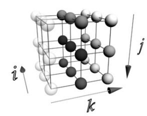
Constr. for lines	**Tensor**
No type-0 constraints	
 Type 1: 3+3 constr.	 Type 1: 18 elements
 Type 2: 9+9 constr.	 Type 2: 12 elements
 Type 2b: 9+9 constr.	 Type 2b: 18 elements

Table 3: Summary of constraints for lines

References

[1] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision.* Cambridge Univ. Press, 2000.

[2] C. B. U. Perwass, *Applications of Geometric Algebra in Computer Vision.*, PhD Thesis, University of Cambridge, 2000.

[3] C.B.U. Perwass, J. Lasenby, "A Geometric Derivation of the Trifocal Tensor and its Constraints". Tech Rep. No.331, Oct. 98, C.U. Eng. Dept., Cambridge University.

[4] O. Faugeras, Q.-T. Luong, *The Geometry of Multiple Images.* The MIT Press, 2001.

[5] D. Hestenes, G. Sobczyk, *Clifford Algebra to Geometric Calculus*, Kluwer Academic Plubishers, 2001.

[6] Y. Ma, R. Vidal, J. Košecká, S. Sastry, "Rank Conditions on the Multiple View Matrix", *Intl. J. of Comp. Vis.*, 2001.

[7] C. Defferara, A. Dell'Acqua, F. Negroni, A. Sarti, S. Tubaro, "New perspectives on camera calibration using geometric algebra". *Intl. Conference on Image Processing*, Sept. 22-25, 2002, Rochester, NY, USA.