



Trinocular Geometry Revisited

Jean Ponce, Martial Hebert

► **To cite this version:**

Jean Ponce, Martial Hebert. Trinocular Geometry Revisited. CVPR - IEEE Conference on Computer Vision and Pattern Recognition, Jun 2014, Columbus, Ohio, United States. 2014. <hal-01053676>

HAL Id: hal-01053676

<https://hal.inria.fr/hal-01053676>

Submitted on 1 Aug 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Trinocular Geometry Revisited

Jean Ponce*
Ecole normale supérieure

Martial Hebert
Carnegie-Mellon University

Abstract

When do the visual rays associated with triplets of point correspondences converge, that is, intersect in a common point? Classical models of trinocular geometry based on the fundamental matrices and trifocal tensor associated with the corresponding cameras only provide partial answers to this fundamental question, in large part because of underlying, but seldom explicit, general configuration assumptions. This paper uses elementary tools from projective line geometry to provide necessary and sufficient geometric and analytical conditions for convergence in terms of transversals to triplets of visual rays, without any such assumptions. In turn, this yields a novel and simple minimal parameterization of trinocular geometry for cameras with non-collinear or collinear pinholes.

1. Introduction

The images of points recorded by multiple cameras may only match when the corresponding visual rays *converge*—that is, intersect in a common point (Figure 1, left). For two views, this condition is captured by the bilinear *epipolar constraint* and the corresponding *fundamental matrix* [8, 9]. Three images can be characterized by both the pairwise epipolar constraints associated with any two of the pictures, and a set of *trilinearities* associated with all three views and parameterized by the associated *trifocal tensor* [5, 15, 16, 22]. For cameras with non-collinear pinholes, at least, the rays associated with three image points that satisfy the corresponding epipolar constraints almost always converge: The *only* exception is when the points have been matched incorrectly, *and* all lie in the trifocal plane spanned by the three pinholes (Figure 1, right). Interestingly, Hartley and Zisserman state that the fundamental matrices associated with three cameras with non-collinear pinholes determine the corresponding trifocal tensor [6, Result 14.5], while Faugeras and Mourrain [3] and Ponce *et al.* [12], for example, note that the rays associated with three points only satisfying certain (and different) subsets of the trilinearities *alone* must intersect. These claims contradict

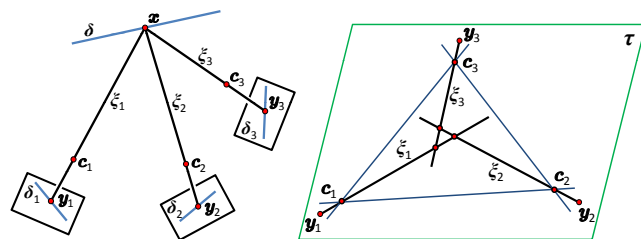


Figure 1. Left: Visual rays associated with three (correct) correspondences. Right: Degenerate epipolar constraints associated with three coplanar, but non-intersecting rays lying in the trifocal plane τ (as in the rest of this presentation, the image planes are omitted for clarity in this part of the figure). See text for details.

each other, since rays that satisfy epipolar constraints do not *always* converge, but they are true under some *general configuration* assumptions, rarely made explicit. It is thus worth clarifying these assumptions, and understanding exactly how much the trifocal constraints add to the epipolar ones for point correspondences. This is the problem addressed in this paper, using elementary projective line geometry. In particular, our analysis shows that exploiting both the epipolar constraints and one or two of the trinocular ones, depending on whether the camera pinholes are collinear or not, always guarantees the convergence of the corresponding visual rays. Our analysis also provides, in both cases, a novel and simple minimal parameterization of trinocular geometry.

1.1. Related Work

Geometric constraints involving multiple perspective views of the same point (Figure 1, left) have been studied in computer vision since the seminal work of Longuet-Higgins, who proposed in 1981 the essential matrix as a bilinear model of epipolar constraints between two calibrated cameras [8]. Its uncalibrated counterpart, the fundamental matrix, was introduced by Luong and Faugeras [9]. The trilinear constraints associated with three views of a straight line were discovered by Spetsakis and Aloimonos [16] and by Weng, Huang and Ahuja [22]. The uncalibrated case was tackled by Shashua [15] and by Hartley [5], who coined the term trifocal tensor. The quadrifocal tensor was introduced by Triggs [20], and Faugeras and Mourrain gave a sim-

*Willow project team. DI/ENS, ENS/CNRS/INRIA UMR 8548.

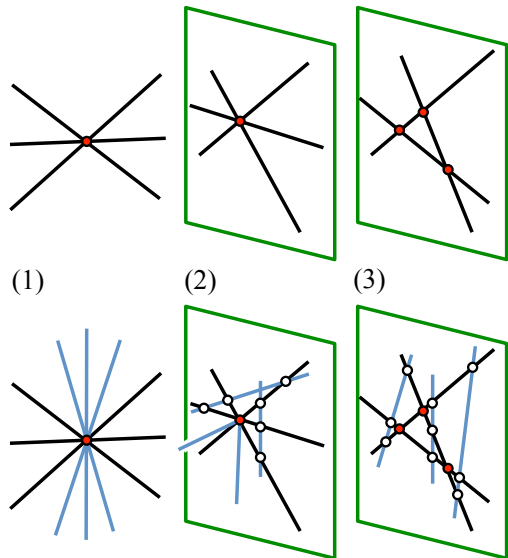


Figure 2. Top: The possible configurations of three pairwise-coplanar distinct lines, classified according to the way they intersect. The three given lines are shown in black; the planes where two of them intersect are shown in green; and the points where two of the lines intersect are shown in red. Bottom: Transversals to the three lines, shown in blue, and forming (1) a line bundle; (2) a degenerate congruence; and (3) a line field.

ple characterization of all multilinear constraints associated with multiple perspective images of a point [3]. The usual formulation of the trilinear constraints associated with three images of the same point are asymmetric, one of the images playing a privileged role. A simple and symmetric formulation based on line geometry was introduced in [12]. A few minimal parameterizations of trinocular geometry are also available [1, 11, 14, 19]. From a historical point of view, it is worth noting that epipolar constraints were already known by photogrameters long before they were (re)discovered by Longuet-Higgins [8], as witnessed by the 1966 Manual of Photogrammetry [17], but that this book does not mention trilinear constraints, although it discusses higher-order trinocular (scale-restraint condition equations).

The direct derivation of trifocal constraints for point correspondences typically amounts to writing that all 4×4 minors of some $k \times 4$ matrix are zero, thus guaranteeing that the three lines intersect [3, 12]. These determinants are then rewritten as linear combinations of *reduced minors* that are bilinear or trilinear functions of the image point coordinates. The whole difficulty lies in selecting an appropriate subset of reduced minors that will always guarantee that the rays intersect. We have already observed that the bilinear epipolar constraints, alone, are not sufficient. We are not aware of any *fixed* set of four trilinearities that, alone, guarantee convergence in all cases. This suggests seeking instead appropriate combinations of bilinear and trilinear constraints, which is the approach taken in this presentation.

1.2. Problem Statement and Proposed Approach

As noted earlier, the goal of this paper is to understand exactly how much the trifocal constraints add to the epipolar ones for point correspondences. Since both types of constraints model incidence relationships among the light rays joining the cameras' pinholes to observed points, we address this problem using the tools of projective geometry [21] in general, and line geometry [13] in particular. As noted earlier, the trifocal tensor was originally invented to characterize the fact that three image lines δ_1 , δ_2 , and δ_3 are the projections of the same scene line δ [15, 16, 22] (Figure 1, left). The trilinearities associated with three image points y_1 , y_2 , and y_3 were then obtained by constructing lines δ_1 , δ_2 , and δ_3 passing through these points, and whose preimage is a line δ passing through the corresponding scene point x . By construction, this line is a *transversal* to the three rays ξ_1 , ξ_2 , and ξ_3 , that is, it intersects them. It is therefore not surprising that much of the presentation will be dedicated to the characterization of the set of transversals to a triplet of lines.

In particular, we have already seen that the fact that three lines intersect pairwise is necessary, but not sufficient for these lines to intersect. We will show in the rest of this presentation that a necessary and sufficient condition for three lines to converge is in fact that they be pairwise coplanar *and* admit a well defined family of transversals. We will also give a simple geometric and analytical characterization of these transversals under various assumptions. When applied to camera systems, it will provide in turn a new and simple minimal parameterization of trinocular geometry.

Contributions

- We give a new *geometric* characterization of triplets of converging lines in terms of transversals to these lines (Proposition 1).
- We provide a novel and simple *analytical* characterization of triplets of converging lines (Lemma 3 and Proposition 2), that does not rely on the assumptions of general configuration implicit in [12].
- We show by applying these results to camera geometry that the three epipolar constraints and one of the trifocal ones (two if the pinholes are collinear) are necessary and sufficient for the corresponding optical rays to converge (Propositions 3 and 4).
- We introduce a new analytical parameterization of epipolar and trifocal constraints, leading to a *minimal* parameterization of trinocular geometry (Propositions 5 and 6).

2. Converging Triplets of Lines

2.1. Geometric Point of View

All lines considered from now on are assumed to be different from each other. A *transversal* to some family of lines is a line intersecting every element of this family. We

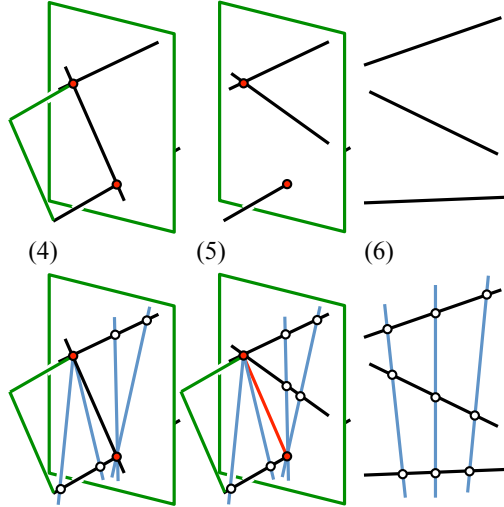


Figure 3. Top: The possible configurations of three distinct, non-pairwise-coplanar lines, classified according to the way they intersect. Bottom: Transversals to the three lines, forming (4) two pencils of lines having one of the input lines (in black) in common (5) two pencils of lines having one line (in red) in common; and (6) a non-degenerate regulus. See text for details.

prove in this section the following main result.

Proposition 1. *A necessary and sufficient condition for three lines to converge is that they be pairwise coplanar, and that they admit a transversal not contained in the planes defined by any two of them.*

To prove Proposition 1, we need two intermediate results. In projective geometry, two straight lines are either skew to each other or coplanar, in which case they intersect in exactly one point. Our first lemma enumerates the possible incidence relationships among three lines.

Lemma 1. *Three distinct lines can be found in exactly six configurations (Figures 2 and 3, top): (1) the three lines are not all coplanar and intersect in exactly one point; (2) they are coplanar and intersect in exactly one point; (3) they are coplanar and intersect pairwise in three different points; (4) exactly two pairs of them are coplanar (or, equivalently, intersect); (5) exactly two of them are coplanar; or (6) they are pairwise skew.*

The proof is by enumeration. Lemma 1 has an immediate, important corollary—that is, when three lines are pairwise coplanar, either they are not coplanar and intersect in one point (case 1); they are coplanar and intersect in one point (case 2); or they are coplanar, and intersect pairwise in three different points (case 3). In particular, epipolar constraints are satisfied for triplets of (incorrect) correspondences associated with images of points in the trifocal plane containing the pinholes of three non-collinear cameras.

To go further, it is useful to introduce a notion of linear (in)dependence among lines. The geometric definition of

independence of lines matches the usual algebraic definition of linear independence, in which, given a coordinate system, a necessary and sufficient for k lines to be linearly dependent is that some nontrivial linear combination of their Plücker coordinate vectors (Section 2.2.1) be the zero vector of \mathbb{R}^6 . Geometrically, the lines linearly dependent on three skew lines form a *regulus* [21]. A regulus is either a *line field*, formed by all lines in a plane; a *line bundle*, formed by all lines passing through some point; the union of all lines belonging to two flat pencils lying in different planes but sharing one line; or a *non-degenerate regulus* formed by one of the two sets of lines ruling a hyperboloid of one sheet or a hyperbolic paraboloid. Linear (in)dependence of four or more lines can be defined recursively. Armed with these definitions, we obtain an important corollary of Lemma 1.

Lemma 2. *Three distinct lines always admit an infinity of transversals, that can be found in exactly six configurations (Figures 2 and 3, bottom): (1) the transversals form a bundle of lines; (2) they form a degenerate congruence consisting of a line field and of a bundle of lines; (3) they form a line field; (4) they form two pencils of lines having one of the input lines in common; (5) they form two pencils of lines having a line passing through the intersection of two of the input lines in common; or (6) they form a non-degenerate regulus, with the three input lines in the same ruling, and the transversals in the other one.*

Lemma 2 should not come as a surprise since the transversals to three given lines satisfy three linear constraints and thus form in general a rank-3 family (the degenerate congruence is a rank-4 exception [21]). Without additional assumptions, not much more can be said in general, since Lemma 2 tells us that any three distinct lines admit an infinity of transversals. When the lines are, in addition, pairwise coplanar, cases 4 to 6 in Lemma 2 are eliminated, and we obtain Proposition 1 as an immediate corollary of this lemma.

2.2. The Analytical Point of View

2.2.1 Preliminaries

To translate the geometric results of the previous section into analytical ones, it is necessary to recall a few basic facts about projective geometry in general, and line geometry in particular. Readers familiar with Plücker coordinates, the join operator, etc., may safely proceed to Section 2.2.2. Given some choice of coordinate system for some two-dimensional projective space \mathbb{P}^2 , points and lines in \mathbb{P}^2 can be identified with their homogeneous coordinate vectors in \mathbb{R}^3 . In addition, if \mathbf{x} and \mathbf{y} are two distinct points on a line ξ in \mathbb{P}^2 , we have $\xi = \mathbf{x} \times \mathbf{y}$. A necessary and sufficient condition for a point \mathbf{x} to lie on a line ξ is $\xi \cdot \mathbf{x} = 0$, and two lines intersect in exactly one point or coincide. A nec-

ecessary and sufficient conditions for three lines to intersect is that they be linearly dependent, or $\text{Det}(\xi_1, \xi_2, \xi_3) = 0$.

In three dimensions, given any choice of coordinate system for a three-dimensional projective space \mathbb{P}^3 , we can identify any line in \mathbb{P}^3 with its Plücker coordinate vector $\xi = (\mathbf{u}; \mathbf{v})$ in \mathbb{R}^6 , where \mathbf{u} and \mathbf{v} are vectors of \mathbb{R}^3 , and we use a semicolon to indicate that the coordinates of \mathbf{u} and \mathbf{v} have been stacked onto each other to form a vector in \mathbb{R}^6 . In addition, if \mathbf{x} and \mathbf{y} are two distinct points on some line $\xi = (\mathbf{u}; \mathbf{v})$ in \mathbb{P}^3 , we have

$$\mathbf{u} = \begin{bmatrix} x_4 y_1 - x_1 y_4 \\ x_4 y_2 - x_2 y_4 \\ x_4 y_3 - x_3 y_4 \end{bmatrix}, \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}. \quad (1)$$

A Plücker coordinate vector is only defined up to scale, and its \mathbf{u} and \mathbf{v} components are by construction orthogonal to each other—this is sometimes known as the *Klein constraint* $\mathbf{u} \cdot \mathbf{v} = 0$. Let us consider the symmetric bilinear form $\mathbb{R}^6 \times \mathbb{R}^6 \rightarrow \mathbb{R}$ associating with two elements $\lambda = (\mathbf{a}; \mathbf{b})$ and $\mu = (\mathbf{c}; \mathbf{d})$ of \mathbb{R}^6 the scalar $(\lambda|\mu) = \mathbf{a} \cdot \mathbf{d} + \mathbf{b} \cdot \mathbf{c}$. A necessary and sufficient for a nonzero vector ξ in \mathbb{R}^6 to represent a line is that $(\xi|\xi) = 0$, and a necessary and sufficient condition for two lines λ and μ to be coplanar (or, equivalently, to intersect) is that $(\lambda|\mu) = 0$.

Let us denote the basis points of some arbitrary projective coordinate system by \mathbf{x}_0 to \mathbf{x}_4 , with coordinates $\mathbf{x}_0 = (0, 0, 0, 1)^T$, $\mathbf{x}_1 = (1, 0, 0, 0)^T$, $\mathbf{x}_2 = (0, 1, 0, 0)^T$, $\mathbf{x}_3 = (0, 0, 1, 0)^T$, and $\mathbf{x}_4 = (1, 1, 1, 1)^T$. Points \mathbf{x}_0 to \mathbf{x}_3 are called the *fundamental points*. The point \mathbf{x}_4 is the *unit point*. Let us also define four *fundamental planes* \mathbf{p}_j ($j = 0, 1, 2, 3$) whose coordinate vectors are the same as those of the fundamental points. The unique line joining two distinct points is called the *join* of these points and it is denoted by $\mathbf{x} \vee \mathbf{y}$. Likewise, the unique plane defined by a line $\xi = (\mathbf{u}; \mathbf{v})$ and some point \mathbf{x} not lying on this line is called the *join* of ξ and \mathbf{x} , and it is denoted by $\xi \vee \mathbf{x}$. Algebraically, we have $\xi \vee \mathbf{x} = [\xi_\vee] \mathbf{x}$, where $[\xi_\vee]$ is the *join matrix* defined by

$$[\xi_\vee] = \begin{bmatrix} [\mathbf{u}_\times] & \mathbf{v} \\ -\mathbf{v}^T & 0 \end{bmatrix}. \quad (2)$$

A necessary and sufficient condition for a point \mathbf{x} to lie on a line ξ is that $\xi \vee \mathbf{x} = 0$.

2.2.2 Back to Transversals

Let us translate some of the geometric incidence constraints derived in the previous section into algebraic ones. We assume that some projective coordinate system is given, and identify points, planes, and lines with their homogeneous coordinate vectors. Let us consider three distinct

lines $\xi_j = (\xi_{1j}, \dots, \xi_{6j})^T$ ($j = 1, 2, 3$) and define

$$D_{ijk} = \begin{vmatrix} \xi_{i1} & \xi_{i2} & \xi_{i3} \\ \xi_{j1} & \xi_{j2} & \xi_{j3} \\ \xi_{k1} & \xi_{k2} & \xi_{k3} \end{vmatrix} \quad (3)$$

to be the 3×3 minor of the 6×3 matrix $[\xi_1, \xi_2, \xi_3]$ corresponding to its rows i , j , and k . A necessary and sufficient condition for this matrix to have rank 2, and thus for the three lines to form a flat pencil (Section 2.1), is that all the minors $T_0 = D_{456}$, $T_1 = D_{234}$, $T_2 = D_{315}$, and $T_3 = D_{126}$ be equal to zero.

Lemma 3. *Given some integer j in $\{0, 1, 2, 3\}$, a necessary and sufficient condition for $\xi_1, \xi_2,$ and ξ_3 to admit a transversal passing through \mathbf{x}_j is that $T_j = 0$.*

Proof. Let us prove the result in the case $j = 0$. The proofs for the other cases are similar. A necessary and sufficient condition for a line $\delta = (\mathbf{u}; \mathbf{v})$ to pass through \mathbf{x}_0 is that $\mathbf{v} = 0$ (this follows from the form of the join matrix). Thus a necessary and sufficient condition for the existence of a line δ passing through \mathbf{x}_0 and intersecting the lines $\xi_j = (\mathbf{u}_j; \mathbf{v}_j)$ is that there exists a vector $\mathbf{u} \neq \mathbf{0}$ such that $(\xi_j|\delta) = \mathbf{v}_j \cdot \mathbf{u} = 0$ for $j = 1, 2, 3$, or, equivalently, that the determinant $T_0 = D_{456} = |\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3|$ be zero. \square

Combining Proposition 2 and Lemma 3 now yields the following important result.

Proposition 2. *A necessary and sufficient condition for three lines $\xi_1, \xi_2,$ and ξ_3 to converge is that $(\xi_i|\xi_j) = 0$ for all $i \neq j$ in $\{1, 2, 3\}$, and that $T_j = 0$ for all j in $\{0, 1, 2, 3\}$.*

Proof. The condition is clearly necessary. To show that is sufficient, note that since the three lines are pairwise coplanar, they either intersect in exactly one point (cases 1 and 2 of Lemma 2), or are all coplanar, intersecting pairwise in three distinct points, with all their transversals in the same plane (case 3). But the latter case is ruled out by Lemma 3 and the condition $T_j = 0$ for $j = 0, 1, 2, 3$ since the fundamental points \mathbf{x}_j are by construction not all coplanar, and at least one of them (and thus the corresponding transversal) does not lie in the plane containing the three lines. \square

3. Converging Triplets of Visual Rays

3.1. Bilinearities or Trilinearities?

Let us now turn our attention from general systems of lines to the visual rays associated with three cameras. As noted earlier, it follows from Lemma 1 that the epipolar constraints alone do not ensure that the corresponding viewing rays intersect (Figure 1, right). On the other hand, the only case where they do not is when the corresponding rays lie in the trifocal plane when the camera pinholes are not

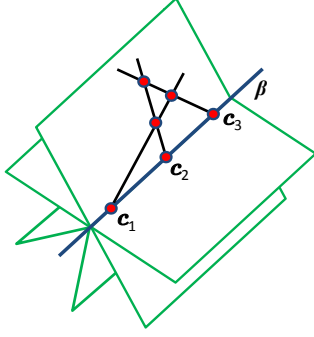


Figure 4. Degenerate epipolar constraints associated with three images when the three pinholes are collinear and the rays are coplanar but don't intersect in a common point.

collinear, or in any plane containing the line joining the three pinholes when they are (Figure 4).

Contrary to the claim of [12, Appendix], the trilinear conditions $T_j = 0$ ($j = 0, 1, 2, 3$) associated with three visual rays do not guarantee, on their own, that the rays intersect: In fact, one can in general construct a two-dimensional family of triplets of non-intersecting visual rays passing through three given non-collinear pinholes and satisfying these constraints. Likewise, although one can show that *some* set of trilinearities can always be chosen to ensure the convergence of the corresponding visual rays, we are not aware of any *fixed* set of trilinearities with the same guarantees, which in turns appears to contradict [3, Sec. 4.2.2] (also the discussion in [7]). This apparent contradiction stems from the fact that both Faugeras and Mourrain [3] and Ponce *et al.* [12] characterize the convergence of visual rays by the vanishing of certain trilinear *reduced minors* of a $k \times 4$ matrix, and have to (implicitly at times) resort to general configuration assumptions to select a representative set of minors. Characterizing the convergence of triplets of lines directly in terms of *both* binocular and trinocular constraints, as in Proposition 2, avoids this difficulty.

3.2. Bilinearities and Trilinearities

By definition, for any choice of projective coordinate system, the four fundamental points x_j ($j = 0, 1, 2, 3$) are not coplanar. When the three pinholes are not collinear, it is thus always possible to choose a projective coordinate system such that one of the fundamental points, say x_0 , does not lie in the trifocal plane, and we obtain the following immediate corollary of Proposition 2.

Proposition 3. *Given three cameras with non-collinear pinholes c_1 , c_2 , and c_3 , and any projective coordinate system such that x_0 does not belong to the trifocal plane, a necessary and sufficient for the three rays $\xi_j = c_j \vee y_j$ ($j = 1, 2, 3$) to converge is that is that $(\xi_i | \xi_j) = 0$ for all $i \neq j$ in $\{1, 2, 3\}$, and $T_0 = 0$.*

When the three pinholes are collinear (but of course dis-

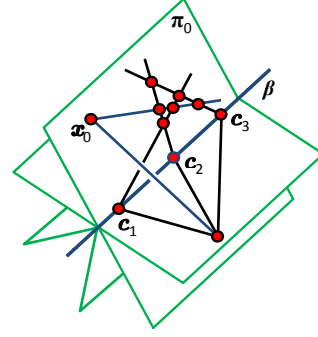


Figure 5. For collinear pinholes, there exists a single scene plane π_0 in the pencil passing through the baseline β that contains x_0 and for which the condition $T_0 = 0$ is ambiguous.

ting), the three cameras admit a single pencil of epipolar planes, and three rays in epipolar correspondence are in fact always coplanar (Figure 4). The trifocal constraints are necessary in this case to ensure that the three lines intersect in exactly one point. Note that, given three cameras with collinear pinholes, one can always choose a projective coordinate system such that the two fundamental points x_0 and x_j (for any j in $\{1, 2, 3\}$) and the baseline joining the three pinholes are not coplanar. The following result characterizes the fact that visual rays intersect in this setting.

Proposition 4. *Given three cameras with collinear pinholes c_1 , c_2 , and c_3 , and any projective coordinate system such that the fundamental points x_0 and x_1 and the baseline β joining the pinholes are not coplanar, a necessary and sufficient condition for the three rays $\xi_j = c_j \vee y_j$ ($j = 1, 2, 3$) to intersect is that is that $(\xi_i | \xi_j) = 0$ for all $i \neq j$ in $\{1, 2, 3\}$, and $T_0 = T_j = 0$ for some $j \neq 0$.*

Proof. The condition is clearly necessary. Because of the epipolar constraints, the three rays must be coplanar, and either intersect in three distinct points with all their transversals in the same plane, intersect in a single point, or coincide with the baseline. Unless the point x_0 lies in the plane π_0 that contains the rays (Figure 5), the first case is ruled out by the condition $T_0 = 0$. If x_0 lies in π_0 , x_j does not (by construction), and the first case is ruled out by $T_j = 0$. \square

3.3. Minimal Parameterizations

3.3.1 Non-Collinear Pinholes

We assume in this section that the three pinholes are not aligned. In this case, we can always choose a projective coordinate system such that the three fundamental points distinct from x_0 are the three camera centers—that is, $c_j = x_j$ for $j = 1, 2, 3$, and x_0 does not lie in the trifocal plane.

With our choice of coordinate system, and the notation $y_j = (y_{1j}, y_{2j}, y_{3j}, y_{4j})^T$, the three epipolar constraints

can be written as

$$\begin{aligned} (\mathbf{x}_1 \vee \mathbf{y}_1 | \mathbf{x}_2 \vee \mathbf{y}_2) = 0 & \quad y_{41}y_{32} = y_{31}y_{42} \\ (\mathbf{x}_1 \vee \mathbf{y}_1 | \mathbf{x}_3 \vee \mathbf{y}_3) = 0 & \iff y_{41}y_{23} = y_{21}y_{43} \\ (\mathbf{x}_2 \vee \mathbf{y}_2 | \mathbf{x}_3 \vee \mathbf{y}_3) = 0 & \quad y_{42}y_{13} = y_{12}y_{43} \end{aligned} \quad (4)$$

Given these constraints, we know from Proposition 2 that a necessary and sufficient conditions for the three visual rays to intersect is that $T_0 = 0$ (the other three trilinearities are trivially satisfied with our choice of coordinate system), which is easily rewritten in our case as

$$y_{21}y_{32}y_{13} = y_{31}y_{12}y_{23}. \quad (5)$$

Note that $y_{4j} = 0$ if and only if \mathbf{y}_j lies in \mathbf{p}_0 , which is also the trifocal plane in our case. As expected, it follows immediately from Eqs. (4) and (5) that, unless $y_{41} = y_{42} = y_{43} = 0$, that is, the observed point lies in the trifocal plane, the epipolar constraints imply the trifocal ones. We now need to translate Eqs. (4-5) to the corresponding equations in *image coordinates*. Let us denote by Π_j ($j = 1, 2, 3$) the 4×3 matrix formed by the coordinate vectors of the basis points for the retinal plane of camera number j . The position of an image point with coordinate vector \mathbf{u}_j in that basis is thus $\mathbf{y}_j = \Pi_j \mathbf{u}_j$. Let us denote by π_{ij}^T the i th row of the matrix Π_j , and use superscripts to index coordinates, i.e., for $k = 1, 2, 3$, π_{ij}^k denotes the k th coordinate of π_{ij} .

Proposition 5. *Given three cameras with non-collinear pinholes and hypothetical point correspondences \mathbf{u}_1 , \mathbf{u}_2 , and \mathbf{u}_3 , a necessary and sufficient condition for the three corresponding rays to converge is that*

$$\begin{aligned} \mathbf{u}_1^T F_{12} \mathbf{u}_2 = 0 & \quad F_{12} = \pi_{41} \pi_{32}^T - \pi_{31} \pi_{42}^T \\ \mathbf{u}_1^T F_{13} \mathbf{u}_3 = 0 & \quad \text{where } F_{13} = \pi_{41} \pi_{23}^T - \pi_{21} \pi_{43}^T, \text{ and} \\ \mathbf{u}_2^T F_{23} \mathbf{u}_3 = 0 & \quad F_{23} = \pi_{42} \pi_{13}^T - \pi_{12} \pi_{43}^T \end{aligned} \quad (6)$$

$$(\pi_{21} \cdot \mathbf{u}_1)(\pi_{32} \cdot \mathbf{u}_2)(\pi_{13} \cdot \mathbf{u}_3) = (\pi_{31} \cdot \mathbf{u}_1)(\pi_{12} \cdot \mathbf{u}_2)(\pi_{23} \cdot \mathbf{u}_3), \quad (7)$$

where the vectors $\pi_1 = (\pi_{21}; \pi_{31}; \pi_{41})$, $\pi_2 = (\pi_{12}; \pi_{32}; \pi_{42})$, and $\pi_3 = (\pi_{13}; \pi_{23}; \pi_{43})$, satisfy the 6 homogeneous constraints

$$\begin{aligned} \pi_{21}^1 = 0, \quad \pi_{32}^2 = 0, \quad \pi_{13}^3 = 0, \\ \pi_{31}^2 = \pi_{41}^3, \quad \pi_{12}^3 = \pi_{42}^1, \quad \pi_{23}^1 = \pi_{43}^2, \end{aligned} \quad (8)$$

and are thus defined by three groups of 7 coefficients, each one uniquely determined up to a separate scale. This is a minimal, 18dof parameterization of trinocular geometry.

Proof. Equations (6) and (7) are obtained immediately by substitution in Eqs. (4) and (5). Together, they provide a 24dof parameterization of the trifocal geometry by the three vectors $\pi_j = (\pi_{1j}; \pi_{2j}; \pi_{3j})$ ($j = 1, 2, 3$), each defined up to scale in \mathbb{R}^9 by 8 independent parameters. Locating the camera pinholes at the fundamental points \mathbf{x}_j ($j = 1, 2, 3$)

freezes 9 of the 15 degrees of freedom of the projective ambiguity of projective structure from motion. It is possible to exploit the remaining 6 degrees of freedom, and to impose the constraints of Eq. (8) on the vectors π_j .

Indeed, the general form of a projective transform Q mapping the three fundamental points \mathbf{x}_j onto themselves has 7 coefficients defined up to scale. Applying such a transform to the matrices Π_j ($j = 1, 2, 3$) defined in some arbitrary projective coordinate system, and writing that the matrices $Q\Pi_j$ must satisfy the constraints of Eq. (8) yields a system of 6 homogeneous equations in the 7 nonzero entries of Q . Note that we can generate many different sets of homogeneous constraints by choosing different sets of entries of the vectors π_1 , π_2 , and π_3 . It can be shown that there is always some choice for which the system defining Q admits a unique solution defined up to scale, and that this solution is nonsingular, thus defining a valid change of coordinates. Together, Eqs. (6), (7) and (8) provide us with a minimal, 18dof parameterization of the trinocular geometry by the three vectors π_1 , π_2 and π_3 now each defined up to scale in \mathbb{R}^9 by only 6 independent parameters. \square

To the best of our knowledge, the minimal parameterization of trinocular geometry proposed by Papadopoulos and Faugeras [11] is the only other one known so far to be one-to-one and parametric (other minimal ones, e.g., [1, 19], impose algebraic constraints). Contrary to [11], our parameterization does not require the use of a computer algebra system to impose rank constraints (see [11] for details). In addition, our parameterization is symmetric, none of the cameras playing a privileged role.

Let us close this section by noting that Eq. (7) has an interesting geometric interpretation: Any point with coordinate vector \mathbf{u}_1 in the first image that matches points with coordinate vectors \mathbf{u}_2 and \mathbf{u}_3 in the other two, must satisfy (7) and thus belong to the ‘‘trinocular line’’ (our terminology):

$$\tau_1 = [(\pi_{32} \cdot \mathbf{u}_2)(\pi_{13} \cdot \mathbf{u}_3)]\pi_{21} - [(\pi_{12} \cdot \mathbf{u}_2)(\pi_{23} \cdot \mathbf{u}_3)]\pi_{31}. \quad (9)$$

This should not come as a surprise since classical trifocal geometry is defined in terms of line correspondences, and Eq. (7) merely expresses the fact that the image point \mathbf{y}_1 lies on the projection τ_1 of the line τ_0 passing through \mathbf{x}_0 that intersects the rays passing through the other two image points, \mathbf{y}_2 and \mathbf{y}_3 . What is less well known is that the lines τ_1 belong to the pencil generated by the lines π_{21} and π_{31} , which intersect at the point $\mathbf{z}_1 = \pi_{21} \times \pi_{31}$ of the first image. The same reasoning applies to the other two images.

3.3.2 Collinear Pinholes

Let us now assume that the three pinholes are collinear (as noted in [10], this case may be important in practice, in aerial photography for example). Let us position the two

pinholes c_1 , c_2 in x_1 and x_2 , and the third pinhole, c_3 , in $x_1 + x_2$. We are free to do this since this amounts to choosing c_1 and c_2 as the fundamental points of the baseline joining the three pinholes, and c_3 as its unit point. From Eq. (4):

$$y_{41}y_{32} = y_{31}y_{42}, \quad y_{41}y_{33} = y_{31}y_{43}, \quad y_{42}y_{33} = y_{32}y_{43}, \quad (10)$$

and write $T_0 = 0$ and $T_3 = 0$ respectively as

$$\begin{aligned} y_{31}y_{32}(y_{23} - y_{13}) + y_{33}(y_{31}y_{12} - y_{21}y_{32}) &= 0, \\ y_{41}y_{42}(y_{23} - y_{13}) + y_{43}(y_{41}y_{12} - y_{21}y_{42}) &= 0. \end{aligned} \quad (11)$$

The other two minors T_1 and T_2 are zero with our choice of coordinate system.

We can rewrite as before Eqs. (10) and (11) in terms of the rows of the matrices Π_j ($j = 1, 2, 3$). Given the special role of $y_{23} - y_{13}$ in Eq. (11), it is convenient to introduce the vector $\omega_3 = \pi_{23} - \pi_{13}$, and we obtain the following characterization of the trinocular geometry.

Proposition 6. *Given three cameras with collinear pinholes and hypothetical point correspondences u_1 , u_2 , and u_3 , a necessary and sufficient condition for the three corresponding rays to converge is that*

$$\begin{aligned} u_1^T F_{12} u_2 &= 0 & F_{12} &= \pi_{41} \pi_{32}^T - \pi_{31} \pi_{42}^T, \\ u_1^T F_{13} u_3 &= 0 & \text{where } F_{13} &= \pi_{41} \pi_{33}^T - \pi_{31} \pi_{43}^T, \\ u_2^T F_{23} u_3 &= 0 & F_{23} &= \pi_{42} \pi_{33}^T - \pi_{32} \pi_{43}^T, \end{aligned} \quad (12)$$

$$\begin{aligned} 0 &= (\pi_{31} \cdot u_1)(\pi_{32} \cdot u_2)(\omega_3 \cdot u_3) + \\ &(\pi_{33} \cdot u_3)[(\pi_{31} \cdot u_1)(\pi_{12} \cdot u_2) - (\pi_{21} \cdot u_1)(\pi_{32} \cdot u_2)], \\ 0 &= (\pi_{41} \cdot u_1)(\pi_{42} \cdot u_2)(\omega_3 \cdot u_3) + \\ &(\pi_{43} \cdot u_3)[(\pi_{41} \cdot u_1)(\pi_{12} \cdot u_2) - (\pi_{21} \cdot u_1)(\pi_{42} \cdot u_2)], \end{aligned} \quad (13)$$

where the vectors $\pi_1 = (\pi_{21}; \pi_{31}; \pi_{41})$, $\pi_2 = (\pi_{12}; \pi_{32}; \pi_{42})$ and $\pi_3 = (\omega_3; \pi_{33}; \pi_{43})$ satisfy the 8 homogeneous constraints

$$\begin{aligned} \pi_{21}^1 &= 0, & \pi_{31}^2 &= 0, & \pi_{12}^1 &= 0, & \pi_{42}^2 &= 0, \\ \pi_{31}^3 &= \pi_{21}^3, & \pi_{32}^3 &= \pi_{42}^3, & \omega_3^1 &= \omega_3^2 &= \omega_3^3, \end{aligned} \quad (14)$$

and are thus defined by three groups of, respectively, 6, 6, and 7 independent coefficients, each uniquely determined up to a separate scale, for a total of 16 independent parameters. This is a minimal, 16dof trinocular parameterization.

Proof. Equations (12) and (13) are obtained immediately by substitution in Eqs. (10) and (11). Together they provide a 24dof parameterization of the trifocal geometry by the three vectors π_j ($j = 1, 2, 3$), each defined up to scale in \mathbb{R}^9 by 8 independent parameters. Locating the camera pinholes in x_1 , x_2 , and $x_1 + x_2$ freezes 7 of the 15 degrees of freedom of the projective ambiguity of projective structure from motion. Similar to the proof of Proposition 5, the remaining 8 degrees of freedom can be used to impose the constraints

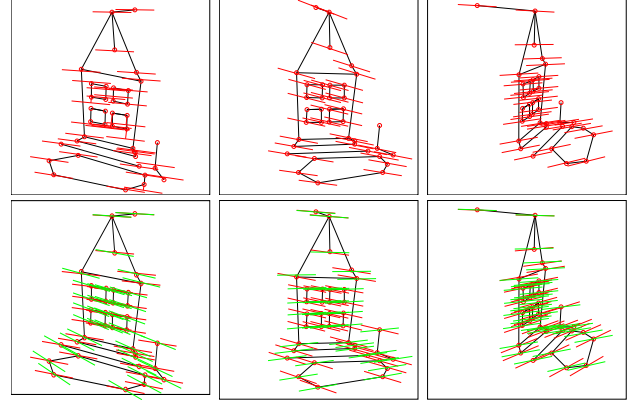


Figure 6. (Top) Example trinocular lines recovered from correspondences in three images; (Bottom) Estimated epipolar lines (two sets per image). Note that the two families of epipolar lines associated with an image typically contain (near) degenerate pairs that can be disambiguated using trilinearity.

of Eq. (14) on the vectors π_j . Together, Eqs. (12), (13) and (14) provide us with a minimal, 16dof parameterization of the trinocular geometry by the three vectors π_1 , π_2 and π_3 now each defined up to scale in \mathbb{R}^9 by only 5, 5, and 6 independent parameters. \square

3.4. Preliminary Implementation

Proposition 5 can be used to estimate the vectors π_j associated with three cameras with non-collinear pinholes from at least six correspondences between three images: Initial values for these vectors are easily obtained from the corresponding projection matrices, estimated from six triplets of matching points using an affine or projective model [2, 18]. The vectors π_j are then refined by minimizing the mean-squared distance between all data points and the corresponding epipolar and trinocular lines. We have constructed a preliminary implementation of this method, and Figure 6 shows an example with 38 correspondences between three images, and the corresponding epipolar and trinocular lines (data courtesy of B. Boufama and R. Mohr). Table 1 shows the average distances between the data points and these lines. The mean distance to epipolar lines is on the order of 1pixel, and comparable to that obtained by classical techniques for estimating the fundamental matrix from pairs of images on the same data [4, Ch. 8]. Our method, on the other hand, is by construction robust to degeneracies with points lying near the trifocal plane. Further experiments and comparisons with other minimal parameterizations of the trinocular geometry [1, 11, 14, 19] are of course needed to truly assess the promise of our approach.

4. Discussion

We have characterized both geometrically and analytically the role of point trilinearity in multi-view geom-

Init.	E12	E13	E23	E21	E31	E32	T1	T2	T3
Aff.	1.0	1.0	0.9	1.0	0.9	0.9	6.3	0.9	0.8
Proj.	2.0	1.6	1.3	1.9	1.5	1.2	7.7	1.7	1.1

Table 1. Quantitative results for the dataset of Figure 6 and affine [18] and projective [2] initializations. Here, “E ij ” refers to the distance between points in image i and the corresponding epipolar lines associated with image j , and “T j ” refers to the distance between points in image j and the corresponding trinocular line associated with the other two images.

etry. Although the nature of our presentation has been mainly theoretical (what are trifocal constraints really?), our analysis has led to a new minimal parameterization of trinocular geometry for both non-collinear and collinear pinholes, and we have presented a preliminary implementation in the non-collinear case. A full-fledged experimental evaluation of this implementation and its extension to the collinear case is next on our agenda.

One may of course wonder whether the fact that four lines intersect in exactly one point can also be characterized geometrically or analytically. Indeed, there exists a quadrifocal tensor expressing the corresponding four-view constraints [20], and it has been shown to be redundant with the epipolar and trifocal constraints. In retrospect, it is geometrically obvious that a necessary and sufficient condition for four lines to intersect in exactly one point is that any two triplets of lines among them also does: This follows immediately from the fact that these triplets have two lines in common, so the point where these two lines intersect is also the point where all four lines intersect.

In other words, there is no need to write any equation to realize that considering four lines together instead of a set of triplets does not add anything to the geometric picture in this case. On the other hand, the natural algebraic constraints to write among four lines is that they be linearly dependent, which is equivalent to writing that all 4×4 minors of the 6×4 matrix formed by their Plücker coordinate vectors be zero. This yields a set of quadrilinear constraints similar to the quadrifocal ones. However, the elements of a rank-3 family of lines do not necessarily intersect in a single point: Instead they form a regulus, in one of the configurations shown in Figure 3, which of course includes bundles. Thus quadrilinearities, *on their own*, are neither necessary (which was already known), nor sufficient, to characterize the fact that the corresponding visual rays intersect. This is intriguing, and perhaps a step toward future work.

Acknowledgments. This work was supported in part by the ERC grant VideoWorld, the Institut Universitaire de France, and ONR MURI N000141010934.

References

[1] N. Canterakis. A minimal set of constraints for the trifocal tensor. In *ECCV*, 2000. 2, 6, 7

[2] S. Carlsson. Duality of reconstruction and positioning in multiple views. In *Proc. Workshop on Representation of Visual Scenes*, 1995. 7, 8

[3] O. Faugeras and B. Mourrain. On the geometry and algebra of the point and line correspondences between n images. Technical Report 2665, INRIA, 1995. 1, 2, 5

[4] D.A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice-Hall, 2003. Second edition. 7

[5] R. Hartley. Lines and points in three views and the trifocal tensor. *IJCV*, 22(2):125–140, 1997. 1

[6] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2000. 1

[7] A. Heyden and K. Astrom. Algebraic properties of multilinear constraints. *Mathematical Methods in the Applied Sciences*, 20(13):1135–1162, 1997. 5

[8] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981. 1, 2

[9] Q.-T. Luong and O.D. Faugeras. The fundamental matrix: theory, algorithms, and stability analysis. *IJCV*, 17(1):43–76, 1996. 1

[10] J.C. McGlone, editor. *Manual of Photogrammetry*. ASPRS, 2004. Fifth Edition. 6

[11] T. Papadopoulo and O. Faugeras. A new characterization of the trifocal tensor. In *ECCV*, 1998. 2, 6, 7

[12] J. Ponce, K. McHenry, T. Papadopoulo, M. Teillaud, and B. Triggs. The absolute quadratic complex and its application to camera self calibration. In *CVPR*, 2005. 1, 2, 5

[13] H. Pottmann and J. Wallner. *Computational Line Geometry*. Springer-Verlag, Berlin, 2001. 2

[14] C. Ressel. A minimal set of constraints and a minimal parameterization for the trifocal tensor. In *ISPRS-Comm. III Symposium, Vol. XXXIV, Part 3A*, 2002. 2, 7

[15] A. Shashua. Algebraic functions for recognition. *PAMI*, 17(8):779–789, 1995. 1, 2

[16] M.E. Spetsakis and Y. Aloimonos. Structure from motion using line correspondences. *IJCV*, 4(3):171–183, 1990. 1, 2

[17] M.M. Thompson, R.C. Eller, W.A. Radlinski, and J.L. Speert, editors. *Manual of Photogrammetry*. ASPRS, 1966. Third Edition. 2

[18] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *IJCV*, 9(2):137–154, 1992. 7, 8

[19] P.H.S. Torr and A. Zisserman. Robust parameterization and computation of the trifocal tensor. *IVC*, 15, 1997. 2, 6, 7

[20] B. Triggs. Matching constraints and the joint image. In *ICCV*, 1995. 1, 8

[21] O. Veblen and J. Young. *Projective Geometry*. The Atheneum Press, 1910. 2, 3

[22] J. Weng, T.S. Huang, and N. Ahuja. Motion and structure from line correspondences: closed-form solution, uniqueness, and optimization. *PAMI*, 14(3):318–336, 1992. 1, 2