Canonic representations for the geometries of multiple projective views

Q.-T. Luong*
EECS, Cory Hall 211-215
University of California
Berkeley, CA 94720

T. Viéville
I.N.R.IA.
2004, route de Lucioles
06902 Sophia-Antipolis, France

qtluong@robotics.eecs.berkeley.edu

Technical Report UCB/CSD-93-772

October 7, 1993

Abstract

We show how a special decomposition of a set of two or three general projection matrices, called *canonic* enables us to build geometric descriptions for a system of cameras which are invariant with respect to a given group of transformations. These representations are minimal and capture completely the properties of each level of description considered: Euclidean (in the context of calibration, and in the context of structure from motion, which we distinguish clearly), affine, and projective, that we also relate to each other. In the last case, a new decomposition of the well-known *fundamental matrix* is obtained. Dependencies, which appear when three or more views are available, are studied in the context of the canonic decomposition, and new composition formulas are established. The theory is illustrated by examples with real images.

Keywords
3D vision, perspective projection, invariants, motion, self-calibration

*The Californian Department of Transportation is acknowledged for support through a PATH project grant MOU94 to QTL.

Contents

1 Introduction and background

Motivations This paper is about an unified framework to account for the Euclidean, affine, and projective geometries of two, three, or more cameras. Three dimensional problems involving several views such as model-based recognition, stereovision or motion and structure from motion analysis have traditionnally been studied under the assumption that the cameras are calibrated. The idea that several classical vision tasks could be performed without full calibration of the cameras, but only using some geometric information which can be obtained from mere point correspondences between uncalibrated images, has generated during the last few years an active research area, whose framework has been projective geometry (see the collective volume [23]). However, the constraints provided by projective geometry have sometimes proven quite weak for some applications. More recently, affine geometry has been found to provide an interesting framework (see for instance [1], where no less than six papers about affine structure can be found), borrowing some nice characteristics from both Euclidean geometry and projective geometry.

However, one can remark that the representations adopted in the literature are of very disparate nature, and that often they are not even minimal. The relationships between different levels of representation has not been investigated thoroughly, which is a consequence of the fact that as the mathematical language used was quite different, comparisons were difficult. Another important point which has not yet received much attention is the problem of dealing with multiple viewpoints to build a coherent representation in the case of uncalibrated cameras. Thus a unified representation is needed, to account in a single framework for the different geometric levels of representation, in the case of two, three, or more views. The principal aim of this paper is to describe such a framework, the *canonic decomposition*. In this section, some background material is presented. Section 2 first introduces representations for each level of description considered, and then gives the canonic decomposition for two views. It is extended to the case of three and more views in Section 3. Section 4. discuss the relations between levels of representation, and the problem of their recovery from image measurements. The paper ends with some examples, presented in Section 5.

The projective model The camera model which we consider is the pinhole model. In this model, the camera performs a perspective projection of an object point M onto a pixel m in the retinal plane $\mathcal{R}$ through the optical center C. The main property of this camera model is thus that *the relationship between the world coordinates and the pixel coordinates is linear projective.* This property is independent of the choice of the coordinate systems in the retinal plane or in the three-dimensional space. The consequence is that the relationship between 2-D pixel coordinates 3-D and any world coordinates can be described by a 3×4 matrix $\tilde{\mathbf{P}}$, called projection matrix, which maps points from $\mathcal{P}^3$ to $\mathcal{P}^2$:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{[\mathbf{P} \;\; \mathbf{p}]}_{\tilde{\mathbf{P}}} \begin{bmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \\ \mathcal{X}_3 \\ \mathcal{X}_4 \end{bmatrix} \tag{1}$$

where the retinal projective coordinates x_1, x_2, x_3 are related to usual pixel coordinates by $(u, v) = (x_1/x_3, x_2/x_3)$ and the projective world coordinates $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{X}_3$, $\mathcal{X}_4$ are related to usual affine world coordinates by $(X, Y, Z) = (\mathcal{X}_1/\mathcal{X}_4, \mathcal{X}_2/\mathcal{X}_4, \mathcal{X}_3/\mathcal{X}_4)$. The points for which $\mathcal{X}_4 = 0$ cannot be related to affine space, and are called *points at infinity.* Although this model has been known for decades, its main characteristic is emphasized in the recent appellation *projective camera* [23].

Note that since we assume a full perspective model, there is an optical center at finite distance. Such a point $\mathbf{C}$ has to verify the matrix equation:

$$\tilde{\mathbf{P}} \begin{bmatrix} \mathbf{C} \\ 1 \end{bmatrix} = \mathbf{0}$$

It is easy to see that it is uniquely defined if, and only if the 3×3 submatrix $\mathbf{P}$ is invertible, which is an assumption that we will use all the way through the paper. This is in opposition with another class of simplified models ranging from orthographic, weak perspective, to the affine camera [23], the most general form of these models, which is studied extensively in [28].

The goal of this paper is to exploit equation (1) to its fullest extend by deriving algebraic consequences (with geometric interpretations) of this equation in the case where several viewpoints are available. The generality of the approach comes from the fact that only projection matrices are manipulated in the paper, thus the results found do not depend on the different primitives one may be interested in, or the algorithms used for the estimation. In the case of a calibrated system, we are just left with the classical description in terms of rotations and translations. In the case of an uncalibrated system of cameras, we show that a complete description for the geometry of two views is given by fundamental matrices, which are described next.

Fundamental matrices as projective view invariants When considering two projective views, the main geometric property is known in computer vision as the epipolar constraint. It can readily be understood by looking at the left part of figure 1. Let $\mathbf{C}$ (resp. $\mathbf{C}'$) be the optical center of the first camera (resp. the second). The line $\langle \mathbf{C}, \mathbf{C}' \rangle$ projects to a point $\mathbf{e}$ (resp. $\mathbf{e}'$) in the first retinal plane $\mathcal{R}$ (resp. in the second retinal plane $\mathcal{R}'$). The points $\mathbf{e}$, $\mathbf{e}'$ are the epipoles. The lines through $\mathbf{e}$ in the first image and the lines through $\mathbf{e}'$ in the second image are the epipolar lines. The epipolar constraint is well-known in stereovision: for each point $\mathbf{m}$ in the first retina, its corresponding point $\mathbf{m}'$ lies on its epipolar line $\mathbf{l}'_m$, projection of $\langle \mathbf{C}, \mathbf{M} \rangle$ in the second retina.

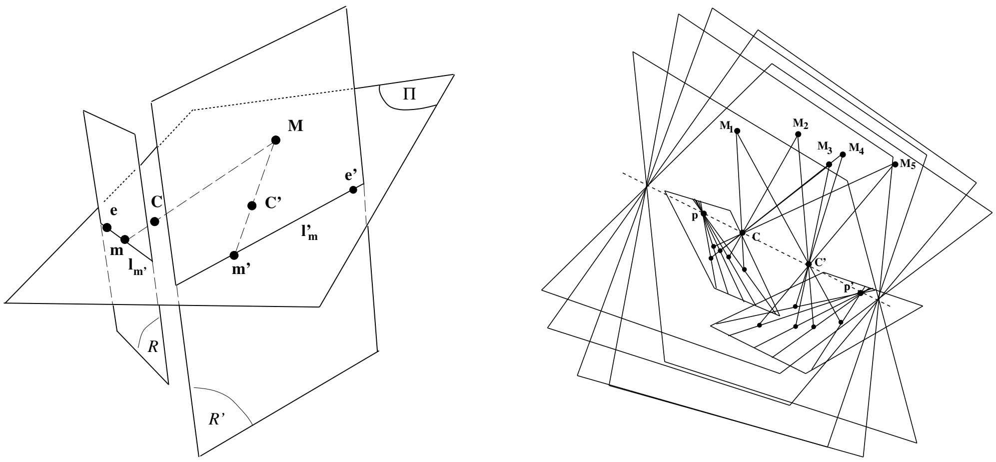

Figure 1: The epipolar geometry and epipolar pencils

Let us consider the one parameter family of planes going through $\langle \mathbf{C}, \mathbf{C}' \rangle$. This family is a pencil of planes, shown at the right of figure 1. Let Π be any plane containing $\langle \mathbf{C}, \mathbf{C}' \rangle$. Then Π projects to an epipolar line l in the first image and to an epipolar line l' in the second image. The correspondences $\Pi \barwedge l$ and $\Pi \barwedge l'$ are homographies between the two pencils of epipolar lines and the pencil of planes containing $\langle \mathbf{C}, \mathbf{C}' \rangle$. It follows that the correspondence $l \barwedge l'$ is a homography, called the *epipolar transformation*.

An algebraic formulation of these properties has been introduced in [6] thanks to the key notion of *fundamental matrix*, or F-matrix. It can be shown only from the hypothesis (1) that the relationship between the projective retinal coordinates of a point $\mathbf{m}$ and the projective coordinates of the corresponding epipolar line $\mathbf{l}'_m$ is linear. The *fundamental matrix* describes this correspondence:

$$\begin{bmatrix} l'_1 \\ l'_2 \\ l'_3 \end{bmatrix} = \mathbf{l}'_m = \mathbf{F}\mathbf{m} = \mathbf{F} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix}$$

The epipolar constraint has then a very simple expression: since the point $\mathbf{m}'$ corresponding to $\mathbf{m}$ belongs to the line $\mathbf{l}'_m$ by definition, it follows that

$$l'_1 x'_1 + l'_2 x'_2 + l'_3 x'_3 = \mathbf{m}'^T \mathbf{F}\mathbf{m} = 0 \tag{2}$$

The epipolar transformation is characterized by the 2×2 projective coordinates of the epipoles $\mathbf{e}$ and $\mathbf{e}'$ (which are defined respectively by $\mathbf{F}\mathbf{e} = 0$ and $\mathbf{F}^T\mathbf{e}' = 0$), and by the 3 coefficients of the homography between the two pencils of epipolar lines. It follows that the epipolar transformation, like the fundamental matrix, depends on seven independent parameters, which represent the only generic information relating two uncalibrated views. Unless further hypotheses are made, there is no way to extract other geometric parameters from correspondences, since one has, in this case, to assume that the transformation between the two retinal plane is a general projective transformation, whereas the fundamental matrix contains the only geometric quantities which are invariant by any projective transformation. Thus the fundamental matrix can be described as an *invariant of views*, which means that it is a function of the projection matrices which is left invariant by any projective transformation of the 3D space $\mathcal{P}^3$.

We want to extend this idea to show that the fundamental matrix and the classical description in terms of intrinsic parameters, rotations, translations, can be represented in a single framework, which includes also a new intermediate representation. We will see that the fundamental matrix represents indeed the minimal information (two views, no additional hypotheses), in a hierarchy of representations obtained by making further assumptions and adding views. Since the observation of planes play a critical role in the sequel, we recall now another important result.

Homographies generated by a plane Let $\mathbf{M}_i$ be space points which happen to lie in the same plane Π and $\mathbf{m}_i$ be their images by a projective linear relation from $\mathcal{P}^3$ to $\mathcal{P}^2$. Its restriction to Π is a projective linear relation between points of $\mathcal{P}^2$, which is an homography h. This relation is invertible, in the generic case where the plane does not contain the optical center. If two images of the points $\mathbf{M}_i$ lying in a plane, $\mathbf{m}_i$ and $\mathbf{m}'_i$ are available, we can consider the relation $h' \circ h^{-1}$ between these two images. It is thus an homography, which means there is a 3×3 invertible matrix $\mathbf{H}$, such that the following projective relation holds for each i:

$$\mathbf{m}'_i = \mathbf{H}\mathbf{m}_i \tag{3}$$

2 A global framework to analyze the geometries of two views

In this section, we introduce a canonic representation of a system of two projection matrices which enables to group Euclidean, affine, and projective invariance properties in a single framework. To lay the ground for such a representation, we first detail the characteristic elements of representation at each level and give a geometric interpretation for them. In particular, a detailed discussion of affine invariants, and new decomposition of the fundamental matrix are given with algebraic and geometric descriptions.

2.1 The Euclidean parameters and the absolute conic

Using the QR theorem, it can be seen that the projection matrix can be decomposed uniquely in the following way:

$$\tilde{\mathbf{P}} = \lambda_w \underbrace{\begin{bmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{R}_w & \mathbf{T}_w \\ \mathbf{0}_3^T & 1 \end{bmatrix}}_{\mathbf{D}_w} \tag{4}$$

where $\mathbf{A}$ is a 3×3 matrix describing the change of retinal coordinate system[1], whose five entries are called *intrinsic parameters*, and $\mathbf{D}_w$ is a 4×4 matrix describing the change of world coordinate system (the pose of the camera) called *extrinsic parameters*. It can be seen that the 5 intrinsic parameters and the 6 pose parameters together account for the 11 parameters of $\tilde{\mathbf{P}}$, which is a 3×4 matrix defined up to a scale factor.

The Euclidean structure of $\mathcal{P}^3$ is characterized[2] by the the absolute conic Ω which lies in the plane at infinity Π_∞ ($\mathcal{X}_4 = 0$) and has equation:

$$\mathcal{X}_1^2 + \mathcal{X}_2^2 + \mathcal{X}_3^2 = 0 \tag{5}$$

A transformation of $\mathcal{P}^3$ leaves Ω invariant if, and only if it is a similarity, which is a rigid displacement multiplied by a scale factor [7, 5]. There is an interesting and important relationship between the camera intrinsic parameters and the absolute conic, already used in [7] and in [22]. Since the absolute conic is invariant under these transformations, its image ω by the camera, which is also a conic with only complex points, does not depend on the pose of the camera. Therefore, its equation in the retinal coordinate system does not depend on the extrinsic parameters and depends only on the intrinsic parameters. Its matrix is $\mathbf{B} = \mathbf{A}^{-1}\mathbf{A}^{-1T}$, whereas its dual conic (the set of its tangents) has matrix $\mathbf{K} = \mathbf{B}^*$, the adjoint matrix of $\mathbf{B}$, defined as $\mathbf{B}^* = \det(\mathbf{B})\mathbf{B}^{-1T}$ [6]. The matrix $\mathbf{K}$ is called the Kruppa matrix, and is a better description of camera calibration than the intrinsic parameters, since it does not depend on the choice of a particular model. There is a one-to-one correspondence between these matrices of Kruppa coefficients and the intrinsic parameters [6, 37]. It can be shown [5] that ω determines the angle between optical rays, which is in coherence with the fact that the similarities conserve angles.

In the case of two views, the extrinsic parameters of the stereo system are classically described by a rotation $\mathbf{R}$ and a translation $\mathbf{T}$, such that if $\mathbf{M}$ (resp. $\mathbf{M}'$) represents the coordinates of a point in the first (resp. second) camera coordinate system, then $\mathbf{M}' = \mathbf{RM} + \mathbf{T}$. When a calibration object and its associated coordinate system are known, the projection matrices can be fully recovered by model-based calibration [33, 9, 34], and the transformation between the two camera coordinate systems is described as a rigid displacement $\mathbf{R}, \mathbf{T}$ which leaves absolute distances invariant. However, in the structure from motion paradigm where the data used is only image measurements, there is an ambiguity between the amount of displacement, represented by $||\mathbf{T}||$, and the depth of objects, thus only the direction of translation can be determined, and we have to consider a global scale factor, and only relative distances. The transformation between the two camera coordinate systems has then to be described by a similarity. It can be noted that the observation of just one line segment of known length would be sufficient to eliminate the scale indetermination.

2.2 Affine structure and the plane at infinity

The projective space $\mathcal{P}^3$ can be described as the union of the usual affine space (points $[\mathbf{M}, M_4]^T$ with $M_4 \neq 0$) and the plane at infinity Π_∞ (points $[\mathbf{M}, 0]^T$). Affine transformations are transformations of $\mathcal{P}^3$ which conserve parallelism. Since in projective geometry the direction $\mathbf{d}$ of a

[1]To obtain a unique decomposition (4), we have to restrict the form of $\mathbf{A}$. Different formulations are possible, as long as the unicity of decomposition is preserved. Other models with direct physical interpretation can be found for example in [9, 5]. The one adopted here for simplicity appeared in [32] and is also discussed in [38].

[2]The idea first appeared in the work of Caylay, and has been introduced in the computer vision literature by [7].

line $\mathbf{l} = [\mathbf{d}, d_4]^T$ can be represented by its intersection with the plane at infinity Π_∞, $[\mathbf{d}, 0]^T$, the conservation of parallelism by a general transformation $\mathcal{A}$ of $\mathcal{P}^3$ is equivalent to the fact that $\mathcal{A}$ leaves the plane at infinity Π_∞ invariant. Expressing this invariance leads to the fact that the last row of the matrix of $\mathcal{A}$ has to be $[0, 0, 0, \mu]$, with $\mu \neq 0$. Since $\mathcal{A}$ is defined only up to a scale factor, we can take $\mu = 1$, and then the transformation $\mathcal{A}$ is fully described by its first 3×4 submatrix $[\mathcal{A}_1, \mathbf{a}]$, which is consistent with the classic definition of an affine transformation of the affine space $\mathbf{M}' = \mathcal{A}_1\mathbf{M} + \mathbf{a}$.

Vanishing points are images of the points at infinity of $\mathcal{P}^3$. Parallel lines of $\mathcal{P}^3$ have the same direction, hence the same point of infinity, thus their projection is a set of lines of $\mathcal{R}$ which contains the image of this vanishing point. If the projection matrix is $\tilde{\mathbf{P}} = [\mathbf{P}, \mathbf{p}]$, then, since the image of each point at infinity $[\mathbf{d}, 0]^T$ is $\mathbf{v} = \mathbf{Pd}$, $\mathbf{P}$ can be considered as the homography between the plane at infinity Π_∞ and the retinal plane $\mathcal{R}$. If a second projection matrix $\tilde{\mathbf{P}}' = [\mathbf{P}', \mathbf{p}']$ is considered, then the transformation:

$$\mathbf{H}_\infty = \mathbf{P}'\mathbf{P}^{-1} \tag{6}$$

is an homography from the first image to the second image, which maps vanishing points to vanishing points, as already remarked by [24]. Introducing the intrinsic parameters into the classical equation for the motion of planes [35, 3] yields the following expression for a general homography:

$$\mathbf{H} = \mathbf{A}'(\mathbf{R} + \frac{1}{d}\mathbf{t}\mathbf{n}^T)\mathbf{A}^{-1} \tag{7}$$

where $\mathbf{n}$ is the normal vector of the plane and d the distance of the plane to the origin. This expression has for limit (6) when $d \to \infty$, which shows that $\mathbf{H}_\infty$ is indeed the limit of the homographies defined by finite planes when they move towards infinity. We will call the matrix $\mathbf{H}_\infty$ *infinity homography*. It allows us to determine whether two lines of $\mathcal{P}^3$ are parallel or not by just checking if their intersection in the first image is mapped to their intersection in the second image by $\mathbf{H}_\infty$. That fact is equivalent to the belonging of the intersection of the two lines of $\mathcal{P}^3$ to Π_∞. Thus the knowledge of $\mathbf{H}_\infty$ determines of parallelism of lines of $\mathcal{P}^3$, which is in coherence with the fact that affine transformations preserve parallelism.

The matrix $\mathbf{H}_\infty$ is proportional to $\mathbf{Q} = \mathbf{A}'\mathbf{R}\mathbf{A}^{-1}$, and thus does not depend on the translational component of the displacement, a propriety used by [2, 39] to obtain the rotational component of the displacement between two cameras once the intrinsic parameters are known. The other component of the representation, $\mathbf{e}'$ depends only on the translational component, since it is proportional to $\mathbf{s} = \mathbf{A}'\mathbf{T}$. These two quantities together define the uncalibrated motion, called *Qs-representation* in [37] which appears to be the simplest generalization of what is well known in the calibrated case, with identical laws of composition as it will be seen in section 3. Unlike $\mathbf{H}_\infty$ and $\mathbf{e}'$, which are basically projective quantities defined up to a scale factor[3], the norm of $\mathbf{Q}$ and $\mathbf{s}$ is completely fixed. Thus the full Qs representation, function of 12 parameters, is a superset of the affine representation $\mathbf{H}_\infty$, $\mathbf{e}'$, but depending on the computational context, one or two unknown scale factors are to be taken into account.

We will see in the next section that when the homography $\mathbf{H}_\infty$ is known, the knowledge of the position of an epipole is sufficient to determine completely the epipolar geometry. Reciprocally, if the fundamental matrix is known, then a consequence of the relation between F-matrices and H-matrices (8) is that $\mathbf{H}_\infty$ is specified by three parameters, which is coherent with the fact already mentioned in [4, 26, 29] that affine structure has still three degrees of freedom given only the epipolar geometry. Without an "affine calibration" [24] consisting of an identification of the plane at infinity, one would have to chose an arbitrary plane to play its role. While some affine properties would then be conserved, the special properties of Π_∞ would be lost. It can be noted that in the special case of an affine camera [28] , affine shape reconstruction is possible from point correspondences, as first discovered by [13] for the case of orthographic projection. This is because an affine camera performs an affine transformation between $\mathcal{P}^3$ and the retina,

[3] They are defined from the projection matrices, which are themselves projective quantities

thus conserving affine properties such as parallelism. In the case of a projective camera, some limited knowledge about the scene such as identification of parallel lines, vanishing points, or horizon, has to be used. This is intermediate between the full calibration required for the Euclidean representation, and the absence of any requirement other than correspondences, for the projective "weak calibration".

2.3 A new representation for projective structure

Only a little has to be said about projective transformations, since the only geometric property which is left invariant is incidence. It can be shown that this entails the invariance of cross-ratio, which will not be used in this paper. Thus this section is only devoted to the detailed study of a new representation for fundamental matrices.

Factorizations of the fundamental matrix It has been shown in [18] that if we have two projective views of a scene, the homography matrix $\mathbf{H}$ mapping the image of a point of a given 3D plane in the first view to its image in the second view, and the fundamental matrix are related by the following system of equations:

$$\mathbf{H}^T\mathbf{F} + \mathbf{F}^T\mathbf{H} = \mathbf{0} \tag{8}$$

We are going to show that this condition is equivalent to the fact that:

$$\mathbf{F} = [\mathrm{e}']_\times \mathbf{H} \tag{9}$$

If this condition is satisfied, We will say then that the matrix $\mathbf{H}$ is *compatible* with the fundamental matrix $\mathbf{F}$. It is straightforward to verify that (8) is satisfied if the substitution of (9) is made in that equation. Now let suppose that $\mathbf{F}$ and $\mathbf{H}$ satisfy (8), and consider a pair of corresponding points $\mathbf{m}$ and $\mathbf{m}'$. The equation (8) is equivalent to the fact that $\mathbf{F}^T\mathbf{H}$ is an antisymmetric matrix, thus:

$$(\mathbf{Fm})^T\mathbf{Hm} = 0 \tag{10}$$

Since the fundamental matrix maps points to corresponding epipolar lines, $\mathbf{Fm} = \mathbf{e}' \times \mathbf{m}'$. Using this relation, we see that (10) is equivalent to: $\mathbf{m}'^T[\mathrm{e}']_\times\mathbf{Hm} = 0$. If we identify this equation with the epipolar constraint: $\mathbf{m}'^T\mathbf{Fm} = 0$, we obtain the expression (9). Note that the proof does not depend on the fact that $\mathbf{H}$ is an invertible (or homography) matrix. In the case where this matrix is singular, it also defines a mapping from the first plane to the second plane, but this mapping is not invertible. We also obtain as an easy consequence of (8) that if $\mathbf{H}$ is compatible with $\mathbf{F}$, then:

$$\mathbf{He} = \lambda\mathbf{e}' \quad \text{or} \quad \mathbf{He} = 0 \tag{11}$$

The first case corresponds to the homography matrices, defined by a plane in general position. The second case corresponds to the degenerated case where the plane contains the optical center $\mathbf{C}'$, thus yielding a non-invertible correspondence.

It is obvious that the decomposition (9) is not unique, since $\mathbf{H}$ can be any matrix defining a correspondence compatible with $\mathbf{F}$. More precisely, since the matrix equation (8) includes six homogeneous equations, such matrices form a 3 dimensional set, each of them being defined by three corresponding points, as shown in [27]. This is coherent with the fact that a plane is defined by three points. If two matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ satisfy (9), then $[\mathrm{e}']_\times\mathbf{M} = 0$, where $\mathbf{M} = \lambda\mathbf{H}_1 - \mathbf{H}_2$. This implies that $\mathbf{M} = \mathbf{e}'\mathbf{r}^T$ for a certain vector $\mathbf{r}$. Thus we find (like [10]) that any two matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ satisfying the decomposition (9) are related by:

$$\lambda\mathbf{H}_2 = \mathbf{H}_1 + \mathbf{e}'\mathbf{r}^T \tag{12}$$

The S-matrix We are going to define a special matrix $\mathbf{S}$ compatible with $\mathbf{F}$, and *which is only function of* $\mathbf{F}$. Using the relation:

$$\|\mathbf{v}\|^2 \mathbf{I}_3 = \mathbf{v}\mathbf{v}^T - [\mathbf{v}]_\times^2 \tag{13}$$

it is seen that:

$$\mathbf{F} = \underbrace{\frac{\mathbf{e}'}{\|\mathbf{e}'\|^2}\mathbf{e}'^T\mathbf{F}}_{\mathbf{0}} + [\mathbf{e}']_\times \underbrace{(-\frac{[\mathbf{e}']_\times}{\|\mathbf{e}'\|^2}\mathbf{F})}_{\mathbf{S}} \tag{14}$$

This relation shows that $\mathbf{S}$ is determined by $\mathbf{F}$ (since $\mathbf{e}'$ is determined by $\mathbf{F}$), and that $\mathbf{F}$ is determined by $\mathbf{S}$ and $\mathbf{e}'$. An analogy can be noted with the decomposition of the essential matrix $\mathbf{E} = [\mathbf{T}]_\times \mathbf{R}$ as the product of an antisymmetric matrix and a rotation matrix $\mathbf{R}$, since the fundamental matrix is decomposed as the product of an antisymmetric matrix and the singular special matrix $\mathbf{S}$, which we call *epipolar projection* matrix, for geometric reasons which will be explained later.

An alternative way of defining $\mathbf{S}$ is to add an additional constraint to the condition (9) to ensure the unicity:

$$\mathbf{F} = [\mathbf{e}']_\times \mathbf{S} \qquad \mathbf{S}^T\mathbf{e}' = 0 \tag{15}$$

Let $\mathbf{H}$ be an homography matrix compatible with $\mathbf{F}$. By transposing (12) we get: $\mathbf{H}^T = \mathbf{S}^T + \mathbf{r}\mathbf{e}'^T$. Since $\mathbf{S}^T\mathbf{e}' = 0$, we obtain $\mathbf{r} = \mathbf{H}^T\mathbf{e}'/\|\mathbf{e}'\|^2$. By using the relation (13) a substitution of this value back to (12) yields $\mathbf{S} = -\frac{[\mathbf{e}']_\times^2}{\|\mathbf{e}'\|^2}\mathbf{H}$. Since $\mathbf{H}$ is compatible with $\mathbf{F}$ the relation (9) applies, and the expression given in (14), independent of $\mathbf{H}$, is again found for $\mathbf{S}$.

Once the fundamental matrix is known, equation (12) can be used to characterize any plane Π by the vector $\mathbf{r}_\Pi$ such that:

$$\mathbf{H}_\Pi = \mathbf{S} + \mathbf{e}'\mathbf{r}_\Pi^T \tag{16}$$

Let use write again equation (12) between the matrix $\mathbf{S}$ and the homography $\mathbf{H}_\infty$ with the value of $\mathbf{r}$ found previously:

$$\mathbf{H}_\infty = \mathbf{S} + \mathbf{e}' \underbrace{\mathbf{e}'^T\mathbf{H}_\infty/\|\mathbf{e}'\|^2}_{\mathbf{r}_\infty^T} \tag{17}$$

The vector $\mathbf{r}_\infty$ which appears in this relation is a complete characterization of the plane at infinity, once the fundamental matrix is known. By combining (7), (16) and (17), we see that $\mathbf{r}_\Pi = \mathbf{r}_\infty + \mathbf{A}^T\mathbf{n}/d$, and thus we can interpret the vector $\mathbf{r}_\Pi$ as the projective characterization of the plane Π. An affine characterization would be the vector $\vec{\nu} = \mathbf{A}^T\mathbf{n}/d$ [38], whereas an Euclidean characterization would be $\mathbf{n}/d$, the normal vector of the plane divided by the distance of the plane to the origin.

Relation with projection matrices The two projection matrices are noted $\tilde{\mathbf{P}} = [\mathbf{P}, \mathbf{p}]$ and $\tilde{\mathbf{P}}' = [\mathbf{P}', \mathbf{p}']$. The epipole in the second image is the projection of the optical center of the first camera into the second camera, thus:

$$\mathbf{e}' = \tilde{\mathbf{P}}' \begin{bmatrix} -\mathbf{P}^{-1}\mathbf{p} \\ 1 \end{bmatrix} = \mathbf{p}' - \mathbf{P}'\mathbf{P}^{-1}\mathbf{p} \tag{18}$$

As shown in section 2.2, $\mathbf{P}'\mathbf{P}^{-1}$ defines an homography matrix compatible with $\mathbf{F}$, noted $\mathbf{H}_\infty$, which gives an expression of $\mathbf{S}$ from the projection matrices, using (14). A geometric interpretation is that the epipolar line of a point $\mathbf{m}$ of the first retina is defined by the epipole $\mathbf{e}'$, and the image by the second camera of the point of infinity of the optical ray $\langle \mathbf{C}, \mathbf{m} \rangle$. We have seen that this image is $\mathbf{H}_\infty \mathbf{m}$, and thus we obtain as expected:

$$\mathbf{F} = [\mathbf{p}' - \mathbf{P}'\mathbf{P}^{-1}\mathbf{p}]_\times \mathbf{P}'\mathbf{P}^{-1} \; , \; \mathbf{S} = [\mathbf{p}' - \mathbf{P}'\mathbf{P}^{-1}\mathbf{p}]_\times^2 \mathbf{P}'\mathbf{P}^{-1}/\|\mathbf{p}' - \mathbf{P}'\mathbf{P}^{-1}\mathbf{p}\|^2 \tag{19}$$

To obtain the second equality, we applied the property (proved for example in [8]):

$$\forall \mathbf{M} \in \mathcal{GL}_3, \forall \mathbf{x}, \mathbf{y} \in \mathcal{R}^3, \mathbf{M}^*[\mathbf{x}]_\times \mathbf{y} = [\mathbf{M}\mathbf{x}]_\times \mathbf{M}\mathbf{y} \quad (20)$$

Let us now show that $\mathbf{S}$ relates the projection matrices through two relations which will be used in section 2.5:

$$\mathbf{P}' = \mathbf{S}\mathbf{P} + \mathbf{e}'\mathcal{L}^T \quad (21)$$

$$\mathbf{p}' = \mathbf{S}\mathbf{p} + \mathbf{e}'\nu \quad (22)$$

First, multiplying relation (17) by $\mathbf{P}$ yields equation (21), with $\mathcal{L} = \mathbf{P}^T\mathbf{r}_\infty = \mathbf{P}'^T\mathbf{e}'/||\mathbf{e}'||^2$. Using the expression (14) and expending the double crossproduct leads to:

$$\mathbf{S}\mathbf{p} = 1.\mathbf{H}_\infty\mathbf{p} - (\frac{\mathbf{e}'^T\mathbf{H}_\infty\mathbf{p}}{||\mathbf{e}'||^2})\mathbf{e}'$$

Since from (18), $\mathbf{H}_\infty\mathbf{p} = \mathbf{p}' - \mathbf{e}'$, we obtain equation (22), with $\nu = (1 + \mathbf{e}'^T\mathbf{H}_\infty\mathbf{p}/||\mathbf{e}'||^2) = \mathbf{p}'^T\mathbf{e}'/||\mathbf{e}'||^2$. Although the quantities are projective, we have kept the scale factors, in order to make their dependencies explicit, which is important since we often have to *add* these quantities in this paper. Note that the use of the projection matrices enable us to determine all the scale factors which are undetermined in a projective context, and to obtain for example:

$$\mathbf{H}\mathbf{e} = -\mathbf{e}' \qquad \mathbf{F}' = (\det \mathbf{H}_\infty)\mathbf{F}^T \quad (23)$$

instead of the mere projective relations (11) and $\mathbf{F}' \sim \mathbf{F}^T$. A technique which has been found useful is to express all the quantities as functions of elements of the projection matrices, which ensures that the scale factors are coherent, and to account for the fact that these scale factors are not always observable in a latter stage.

2.4 A geometric interpretation

In this section, we drop the scale factors, since we are not seeking algebraic relations. Let us first establish a result which will be used in this section: a plane $\mathbf{\Pi}$ of direction $\mathbf{d}$ containing the optical center $\mathbf{C}$ of a camera with projection matrix $\tilde{\mathbf{P}} = [\mathbf{P}, \mathbf{p}]$, is projected on the line $\mathbf{P}^{-1T}\mathbf{d}$ by this camera. By writing the equivalence of the fact that $\mathbf{\Pi}$ contains $\mathbf{M}$: $\mathbf{\Pi}^T\mathbf{M} = 0$ and of the fact that its image $\mathbf{l}_\Pi$ contains the projection of $\mathbf{M}$: $\mathbf{l}_\Pi^T\mathbf{P}\mathbf{M} = 0$, we obtain two equations. They are compatible because $\mathbf{C}$ belongs to $\mathbf{\Pi}$: $\mathbf{\Pi}^T\mathbf{C} = 0$, and eventually yield: $\mathbf{l}_\Pi = \mathbf{P}^{-1T}\mathbf{d}$.

Although the matrix $\mathbf{S}$ is singular, the fact that it is compatible with the fundamental matrix allows to interpret this matrix as a correspondence between the two retinas induced by a plane. Some computations that are detailed in Appendix B, show that this plane has projective equation:

$$\mathbf{\Pi}_{e'} = \begin{bmatrix} \mathbf{P}'^T\mathbf{e}' \\ \mathbf{p}'^T\mathbf{e}' \end{bmatrix}$$

It can be easily verified that this plane contains the optical center of the second camera:

$$\mathbf{C}' = \begin{bmatrix} \mathbf{P}'^{-1}\mathbf{p}' \\ -1 \end{bmatrix}$$

The correspondence from the first image to the second image is thus a mapping to the line which is the image of $\mathbf{\Pi}_{e'}$ in the second retina, and this mapping is a projection and is not invertible. Applying the result proved at the beginning of this section, we obtain that the matrix $\mathbf{S}$ is the correspondence defined by the place $\mathbf{\Pi}_{e'}$ which contains the optical center of the second camera, and whose image in the second camera is the line $\langle\mathbf{e}'\rangle$. It can also be verified that in the canonic decomposition proposed in (2.5) the homography $\mathcal{H}$ from the first view to the second view maps the plane $\mathbf{\Pi}_{e'}$ to the plane at infinity $\mathbf{\Pi}_\infty$. Let see that this is coherent with definition (14). The matrix $\mathbf{F}$ maps points to lines, the matrix $[\mathbf{e}']_\times$ either maps lines to points or points to lines, thus from (14) the matrix $\mathbf{S}$ maps points to points. More specifically, a point $\mathbf{m}$ is mapped to $\mathbf{m}'_1 = \mathbf{e}' \times \mathbf{Fm}$, which is the intersection of the epipolar line of $\mathbf{m}$ with the line $\langle \mathbf{e}' \rangle$. We can note that this point is always defined as soon as $\mathbf{m} \neq \mathbf{e}$ since the distinctive property of the line $\langle \mathbf{e}' \rangle$ is that it does not contain the point $\mathbf{e}'$, as we have always $\mathbf{e}'^T\mathbf{e}' = \|\mathbf{e}'\|^2 \neq 0$. The interpretation of (14) is that the epipolar line $\mathbf{l}'_m = \mathbf{Fm}$ is defined by joining the epipole $\mathbf{e}'$ and the point $\mathbf{m}'_1 = \mathbf{Sm}$, (intersection of the epipolar line and the epipole) thus the transformation $\mathbf{S}$ and the epipole $\mathbf{e}'$ define completely the epipolar geometry.

Let consider $\mathbf{S}^T = \mathbf{F}^T[\mathbf{e}']_\times$, and an epipolar line $\mathbf{l}'$ of the second retina. Then $\mathbf{m}'_1 = \mathbf{e}' \times \mathbf{l}'$ is a point of $\mathbf{l}'$ which is always defined if $\mathbf{l}' \neq \langle \mathbf{e}' \rangle$. Then $\mathbf{l} = \mathbf{F}^T\mathbf{m}'_1$ is the epipolar line of $\mathbf{m}'_1$, but since $\mathbf{m}'_1$ is a point of $\mathbf{l}'$, $\mathbf{l}$ and $\mathbf{l}'$ are corresponding epipolar lines. Thus the matrix $\mathbf{S}^T$ maps epipolar lines to corresponding epipolar lines (a property which is also true of any homography matrix as mentioned by [10, 30]). From relations (15,14), and the fact that $\mathbf{F}' = \mathbf{F}^T$, we obtain:

$$\mathbf{S}' = [\mathbf{e}]_\times \mathbf{S}^T [\mathbf{e}']_\times$$

which can be interpreted as follows: a point $\mathbf{m}'$ is mapped to an epipolar line $\mathbf{l}'$ passing through it, which is then mapped into the corresponding epipolar line $\mathbf{l}$. The final result is the intersection of this line with the line $\langle \mathbf{e} \rangle$ which is what we expected.

The epipole $\mathbf{e}'$ depends on two independent parameters, since it is defined only up to a scale factor. The transformation $\mathbf{S}$ is a linear projection of a projective plane (the first retina) on a projective line $\langle \mathbf{e}' \rangle$, thus it is defined by a 2×3 matrix defined up to a scale factor. Since the line $\langle \mathbf{e}' \rangle$ is also defined by the same parameters than the epipole, we see that the knowledge of the linear projection (5 parameters) and the epipole (2 parameters) define completely the 3×3 matrices $\mathbf{S}$ and $\mathbf{F}$, which is consistent with the result that the fundamental matrix depends on 7 parameters [6]. We have thus exhibited a new decomposition of this matrix in two subsets of independent parameters, which has a sound geometric interpretation in terms of epipolar mappings.

To determine the infinity homography $\mathbf{H}_\infty$ three more parameters are needed. These parameters are for example the coordinates of the vector $\mathbf{r}_\infty = \mathbf{H}_\infty^T\mathbf{e}'/\|\mathbf{e}'\|^2$, which has been shown to characterize the infinity homography once the epipolar geometry is known. The direction of this vector defines the line in the first retina which is the image by $\mathbf{H}_\infty^T$ of the special line $\langle \mathbf{e}' \rangle$ of the second retina. This line is the projection of the plane parallel to $\mathbf{\Pi}_{e'}$ and containing the optical center $\mathbf{C}$ of the first camera. Note that a general property of the dual homography $\mathbf{H}_\infty^T$ is that two lines $\mathbf{l}$ and $\mathbf{l}'$, respective projections of the planes $\mathbf{\Pi}$ and $\mathbf{\Pi}'$ in each retina, are in correspondence by $\mathbf{H}_\infty^T$ if, and only if the planes $\mathbf{\Pi}$ and $\mathbf{\Pi}'$ are parallel. This comes from the relation between planes and lines established at the beginning of the section. Taking the direction of $\mathbf{r}_\infty$ drops one parameter, since we have lost the information corresponding to the norm of this vector. Thus knowing the epipolar geometry and the line $\langle \mathbf{H}_\infty^T\mathbf{e}' \rangle$ is not sufficient to determine $\mathbf{H}_\infty^T$, and we need another piece of geometric information, for instance the line $\langle \mathbf{H}_\infty^{-1T}\mathbf{e} \rangle$ in the second retina. The lines $\langle \mathbf{e} \rangle$ and $\langle \mathbf{H}_\infty^T\mathbf{e}' \rangle$ are never identical, since we have $\mathbf{e}^T\mathbf{H}_\infty^T\mathbf{e}' = -\mathbf{e}'^T\mathbf{e}' \neq 0$, and thus these two line correspondences, together with the epipolar projection $\mathbf{S}$ are sufficient to characterize the infinity homography.

2.5 The canonic decomposition for two views

If two projective views are considered, the most complete description is given through the two projection matrices $\tilde{\mathbf{P}} = [\mathbf{P}, \mathbf{p}]$ and $\tilde{\mathbf{P}}' = [\mathbf{P}', \mathbf{p}']$. Since each matrix is defined up to a scale factor, this representation is not unique and the total number of parameters is 22. However, a total determination of these matrices cannot be done except in the case where a calibration object and its associated coordinate system are known. This total determination is not necessary: for example, in the Euclidean case, the choice of a particular world coordinate system is arbitrary, which means that the representation is defined up to a displacement. One is generally interested only in descriptions of the geometric relationship between the two images

EUCLIDEAN (calibration)	**displacements** *preserve* **angles**, **distances** $\Vert\mathcal{D}(\mathbf{i})\Vert = \Vert\mathbf{i}\Vert$			
$\mathcal{SO}_3$	$\mathcal{D} = \begin{bmatrix} \mathcal{R} & \mathcal{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix}$	$\mathcal{R}$: rotation matrix $\mathcal{T}$: translation vector		6
invariant description	$\mathbf{A}$,$\mathbf{A}'$: intrinsic parameters of cameras			5+5
	$\mathbf{R}$: rotation from camera 1 to camera 2			3
	$\mathbf{T}$: translation from camera 1 to camera 2			3
canonic decomposition	$\begin{cases} \tilde{\mathbf{P}} = \mathbf{A}[\mathbf{I}_3, 0]\ \mathcal{D} \\ \tilde{\mathbf{P}}' = \mathbf{A}'[\mathbf{R}, \mathbf{T}]\ \mathcal{D} \end{cases}$	$\begin{array}{c} \mathbf{R} = \mathbf{R}'_w \mathbf{R}_w^T \\ \mathbf{T} = \mathbf{T}'_w - \mathbf{R}'_w \mathbf{R}_w^T \mathbf{T}_w \\ \hline \mathcal{R} = \mathbf{R}_w \\ \mathcal{T} = \mathbf{T}_w \end{array}$	where $\tilde{\mathbf{P}} = \mathbf{A}[\mathbf{R}_w, \mathbf{T}_w]$ $\tilde{\mathbf{P}}' = \mathbf{A}'[\mathbf{R}'_w, \mathbf{T}'_w]$	
EUCLIDEAN (motion)	**similarities** *preserve* **angles**, **relative distances** $\mathcal{S}(\Omega) = \Omega$			
$\lambda\mathcal{SO}_3$	$\mathcal{S} = \begin{bmatrix} \mathcal{R} & \mathcal{T} \\ \mathbf{0}_3^T & \lambda \end{bmatrix}$	$\mathcal{R}$: rotation matrix $\mathcal{T}$: translation vector λ: non-null scalar		7
invariant description	$\mathbf{A}$,$\mathbf{A}'$: intrinsic parameters of cameras			5+5
	$\mathbf{R}$: rotation from camera 1 to camera 2			3
	$\mathbf{t}$: direction of translation from camera 1 to camera 2			2
canonic decomposition	$\begin{cases} \tilde{\mathbf{P}} = \mathbf{A}[\mathbf{I}_3, 0]\ \mathcal{S} \\ \tilde{\mathbf{P}}' = \mathbf{A}'[\mathbf{R}, \mathbf{t}]\ \mathcal{S} \end{cases}$	$\begin{array}{c} \mathbf{R} = \mathbf{R}'_w \mathbf{R}_w^T \\ \mathbf{t} = \mathbf{T}/\Vert\mathbf{T}\Vert \\ \hline \mathcal{R} = \mathbf{R}_w \\ \mathcal{T} = \mathbf{T}_w \\ \lambda = \Vert\mathbf{T}\Vert \end{array}$	where $\mathbf{T} = \mathbf{t}'_w - \mathbf{R}'_w \mathbf{R}_w^T \mathbf{T}_w$	
AFFINE	**affine transformations** *preserve* **parallelism**, **center of mass** $\mathcal{A}(\Pi_\infty) = \Pi_\infty$			
$\mathcal{GA}_3$	$\mathcal{A} = \begin{bmatrix} \mathcal{M} & \mathcal{V} \\ \mathbf{0}_3^T & \mu \end{bmatrix}$	$\mathcal{M}$: non-singular 3×3 matrix $\mathcal{V}$: 3D vector μ: non-null scalar	$\mathcal{A}$ defined up to a global scale factor	12
invariant description	$\mathbf{H}_\infty$: infinity homography from image 1 to image 2			8
	$\mathbf{e}'_N$: normalized epipole in image 2			2
canonic decomposition	$\begin{cases} \tilde{\mathbf{P}} = [\mathbf{I}_3, 0]\ \mathcal{A} \\ \tilde{\mathbf{P}}' = [\mathbf{H}_\infty, \mathbf{e}'_N]\ \mathcal{A} \end{cases}$	$\begin{array}{c} \mathbf{H}_\infty = \mathbf{P}'\mathbf{P}^{-1} \\ \mathbf{e}'_N = \mathbf{e}/\Vert\mathbf{e}'\Vert \\ \hline \mathcal{M} = \mathbf{P} \\ \mathcal{V} = \mathbf{p} \\ \mu = \Vert\mathbf{e}'\Vert \end{array}$	where $\mathbf{H}_\infty \sim \mathbf{A}'\mathbf{R}\mathbf{A}^{-1}$ $\mathbf{e}' = \mathbf{p}' - \mathbf{H}_\infty\mathbf{p} \sim \mathbf{A}'\mathbf{T}$	
PROJECTIVE	**homographies** *preserve* **collinearity**, **cross-ratio** $\mathcal{H}(\mathcal{P}^3) = \mathcal{P}^3$			
$\mathcal{GL}_4$	$\mathcal{H} = \begin{bmatrix} \mathcal{M} & \mathcal{V} \\ \mathcal{L}_N^T & \nu_N \end{bmatrix}$	$\mathcal{M}$: 3×3 matrix $\mathcal{V}$, $\mathcal{L}_N$: 3D vectors ν_N: scalar	$\mathcal{H}$ non-singular defined up to a global scale factor	15
invariant description	$\mathbf{S}$: epipolar projection from image 1 to image 2			5
	$\mathbf{e}'_N$: normalized epipole in image 2			2
canonic decomposition	$\begin{cases} \tilde{\mathbf{P}} = [\mathbf{I}_3, 0]\ \mathcal{H} \\ \tilde{\mathbf{P}}' = [\mathbf{S}, \mathbf{e}'_N]\ \mathcal{H} \end{cases}$	$\begin{array}{c} \mathbf{S} = -[\mathbf{e}'_N]_\times{}^2\mathbf{H}_\infty \\ \mathbf{e}'_N = \mathbf{e}/\Vert\mathbf{e}'\Vert \\ \hline \mathcal{M} = \mathbf{P} \\ \mathcal{V} = \mathbf{p} \\ \mathcal{L}_N = \mathbf{P}'^T\mathbf{e}'_N \\ \nu_N = \mathbf{p}'^T\mathbf{e}'_N \end{array}$	where $\mathbf{H}_\infty = \mathbf{P}'\mathbf{P}^{-1}$ $\mathbf{e}' = \mathbf{p}' - \mathbf{H}_\infty\mathbf{p}$	

Table 1: The geometries of two views: canonic representation

that are invariant by some group $\mathcal{G}$ of transformation of the projective space $\mathcal{P}^3$, which will be refered to as *descriptions of level* $\mathcal{G}$ or $\mathcal{G}$-invariant descriptions. The properties which can be recovered from these descriptions are those which are left invariant by the transformations of $\mathcal{G}$.

The basic idea is very simple and powerful, and allows to derive in a purely algebraic way invariant descriptions of views knowing only the properties of the elements of $\mathcal{G}$. Note that by the term view invariant, we do not mean a function attached to a set of observed geometric entities, but rather a function attached to the set of two cameras, which describes the geometric relations between these cameras, in a way which is invariant with respect to a transformation of $\mathcal{G}$ applied to the object space $\mathcal{P}^3$. Further details about this derivation are given in Appendix A for the affine case.

> *The matrices $\mathcal{I}$, $\mathcal{I}'$, expressed as functions of a pair of generic projection matrices $\tilde{\mathbf{P}}$, $\tilde{\mathbf{P}}'$, such that there is a unique decomposition, called canonic:*
>
> $$\tilde{\mathbf{P}} = \mathcal{I}\mathcal{T} \qquad \tilde{\mathbf{P}}' = \mathcal{I}'\mathcal{T} \tag{24}$$
>
> *with $\mathcal{T}$ being an element of $\mathcal{G}$, provide a a complete description of the geometric properties of two projective views which are left invariant by the group of transformation $\mathcal{G}$.*

If the pair $\tilde{\mathbf{P}}$, $\tilde{\mathbf{P}}'$ is transformed by any member $\mathcal{T}_1$ of $\mathcal{G}$, then a decomposition of the same form still holds where $\mathcal{T}$ is just replaced by $\mathcal{T}_2 = \mathcal{T}\mathcal{T}_1$. The equivalence of these representations comes from the group properties of $\mathcal{G}$. It follows that the relation $\Re(\mathcal{G})$ defined by:

$$\{(\tilde{\mathbf{P}}_1, \tilde{\mathbf{P}}'_1) \; \Re(\mathcal{G}) \; (\tilde{\mathbf{P}}_2, \tilde{\mathbf{P}}'_2)\} \quad \Leftrightarrow \quad \{\exists \mathcal{T} \in \mathcal{G} \;,\; \tilde{\mathbf{P}}_2 = \tilde{\mathbf{P}}_1\mathcal{T} \;\wedge\; \tilde{\mathbf{P}}'_2 = \tilde{\mathbf{P}}'_1\mathcal{T}\}$$

is an equivalence relation over the set of pairs of projection matrices. The canonic decomposition is an explicit decomposition over a canonic representant of each equivalence class of $\Re(\mathcal{G})$. Let us list some consequences of this construction:

- The sum of the number of parameters in the representation $\mathcal{I}$, $\mathcal{I}'$ and in the generic transformation $\mathcal{T}$ has to be 22.
- Every quantity which depends only on the projection matrices and is invariant with respect to $\mathcal{G}$ is also function of $\mathcal{I}'$ and $\mathcal{I}'$.
- The quantities which appear in matrix $\mathcal{T}$ are not measurable from two views using the representation of level $\mathcal{G}$. But they may be expressed using representations of the previous level, instead. Example of such quantities include the scale factor λ, which corresponds to the norm of the translation, and the vector $[\mathcal{L}^T, \nu]$ which corresponds to the reciprocal image of the plane at infinity by the homography $\mathcal{H}$.
- The decomposition provides a tool for building explicitly a pair of projection matrices $\tilde{\mathbf{P}}, \tilde{\mathbf{P}}'$ from the invariants obtained with respect to $\mathcal{G}$, which captures all the properties of a pair of views up to a transformation of $\mathcal{G}$. For example, if $\mathcal{G}$ is the projective group, the general invariant of two views is the fundamental matrix. Now given a particular fundamental matrix $\mathbf{F}$, the decomposition formulas (24) gives two projection matrices whose fundamental matrix is $\mathbf{F}$. Further, if we reconstruct 3D points using these projection matrices, then we obtain a reconstruction which is coherent up to a projective transformation of $\mathcal{P}^3$, as done already in [4, 11].

We consider as group of transformations $\mathcal{G}$ the group of the displacements $\mathcal{SO}_3$, the group of similarities $\lambda\mathcal{SO}_3$, which are the product of a multiplication by a scalar and a displacement, the group of affine transformations $\mathcal{GA}_3$ and the group of homographies of $\mathcal{P}^3$, $\mathcal{GL}_4$. Note that $\mathcal{SO}_3$ is not relevant in the context of analysis from views, since the scale factor ambiguity can not be resolved in this framework, because of the well-known *depth-speed ambiguity.* The previous sections have laid the groundfor the results which are summerized in table 1, in which we mention:

- the characteristic properties and generic decomposition of a member of each of these group of transformations,
- a canonic decomposition of the form (24) of two projection matrices. The quantities above the horizontal line are the elements of the invariant description, the quantities under that line are non-measurable,
- indication of links with the previous level,
- the number of parameters, whose sum is exactly 22.

The verification of each decomposition is straightforward, except for the Euclidean decomposition of the projection matrices, already discussed in section 2.1, and for the projective case, where it depends on the equations (21,22) already established. It should be noted that the invariants $\mathbf{e}'$, $\mathbf{H}_\infty$, $\mathbf{S}$ are projective, thus defined only up to a scale factor, as well as the matrices $\mathcal{A}$ and $\mathcal{H}$. It can be verified that this reflects coherently the fact that the projection matrices $\tilde{\mathbf{P}}, \tilde{\mathbf{P}}'$ are also projective quantities, as follow:

$\tilde{\mathbf{P}}$	$\tilde{\mathbf{P}}'$	$\mathbf{e}'$,	$\mathbf{H}_\infty$	$\mathbf{S}$	$\mathcal{A}$	$\mathcal{H}$
$\lambda\tilde{\mathbf{P}}$	$\tilde{\mathbf{P}}'$	$\mathbf{e}'$,	$\lambda^{-1}\mathbf{H}_\infty$	$\lambda^{-1}\mathbf{S}$	$\lambda\mathcal{A}$	$\lambda\mathcal{H}$
$\tilde{\mathbf{P}}$	$\mu\tilde{\mathbf{P}}'$	$\mu\mathbf{e}'$,	$\mu\mathbf{H}_\infty$	$\mu\mathbf{S}$	$\mu\mathcal{A}$	$\mu\mathcal{H}$

It is important to notice that reciprocally, multiplying any of the elements $\mathbf{e}'$, $\mathbf{H}_\infty$, $\mathbf{S}$ in the canonic decomposition will only multiply the projection matrices $\tilde{\mathbf{P}}, \tilde{\mathbf{P}}'$ by a scale factor, which is not significant. Using the projective epipole $\mathbf{e}'$ as an invariant would have been perfectly adequate in this two-view analysis, because the norm of this quantity is not constrained in any way by two mere views. However, we will see in the next section that it is constrained if three views are considered, and thus, in order to be coherent with the sequel, we have taken a scale-invariant representation for the epipole, the normalized epipole $\mathbf{e}'_N = \mathbf{e}/||\mathbf{e}'||$. The vectors $\mathbf{r}_\infty$, $\mathcal{L}$ and the scalar ν have to be scaled accordingly, resulting in the quantities:

$$\mathbf{r}_{\infty N} = \mathbf{H}^T\mathbf{e}'_N \qquad \mathbf{L}_N = \mathbf{P}'^T\mathbf{e}'_N \qquad \nu_N = \mathbf{p}'^T\mathbf{e}'_N \tag{25}$$

Although the canonic decomposition given in table 1 are somewhat arbitrary, we can note that they are based on the idea of minimizing the change of coordinate system in the first camera, and thus are very natural. The fact that in the affine and projective formulation, the first invariant matrix of the canonic decomposition is the identity means that it is possible to work directly from pixel coordinates, no retinal coordinate change being necessary. Another thing worth noting is that the epipole $\mathbf{e}'$ is part of the representation for the affine and projective formulation, which confirm the importance of computing this quantity accurately, a fact already stressed in [16] where algorithmic issues were considered.

3 The geometries of three projective views

3.1 Some composition relations

We turn now to the case of three views, and will use the subscripts 1, 2, 3, to designate them. A transformation noted $\mathbf{M}_{ij}$ will always map quantities of view i to quantities of view j. A quantity present in view i in relation with view j will be noted $\mathbf{v}_{ij}$. We first describe the relations between the descriptions relating each of the three pairs of views.

Euclidean representation The knowledge of the intrinsic parameters ensures that in the case where the translations are completely determined, the composition relations are obviously obtained as composition of displacements. If the displacements $\mathcal{D}_{12}$ and $\mathcal{D}_{23}$ are known only up to a scale factor, (which means that their translational component is: $\mathbf{t}_{12} = \mathbf{T}_{12}/||\mathbf{T}_{12}||$ and $\mathbf{t}_{23} = \mathbf{T}_{23}/||\mathbf{T}_{23}||$), it is still possible to determine the rotation:

$$\mathbf{R}_{13} = \mathbf{R}_{23}\mathbf{R}_{12} \tag{26}$$

but not the direction of the translation of $\mathcal{D}_{23}\mathcal{D}_{12}$, since there are two additional unknowns in the relation:

$$\alpha_1 \mathbf{t}_{13} = \mathbf{R}_{23}\mathbf{t}_{12} + \alpha_2 \mathbf{t}_{23} \tag{27}$$

which are $\alpha_1 = ||\mathbf{T}_{13}||/||\mathbf{T}_{12}||$ and $\alpha_2 = ||\mathbf{T}_{23}||/||\mathbf{T}_{12}||$. But if $\mathcal{D}_{12}$, $\mathcal{D}_{23}$ and the ratio α_2 are known, then it is possible to determine $\mathbf{t}_{13}$ and the other ratio. This means that $\mathcal{D}_{12}$ and $\mathcal{D}_{23}$ (resp $\mathcal{D}_{13}$) being known, the knowledge of $\mathcal{D}_{13}$ (resp. $\mathcal{D}_{23}$) is equivalent to that of α_2 (resp α_1). When we have three views, there is only a global scale indetermination, but the ratio of the local scale factors is completely determined. This was the idea behind a self-calibration approach to analyze the motion from three views [19]. Thus, the description is complete if, in addition to the direction of translations, the ratio of their norm is known.

Affine representation Let us first point out to a general result for homographies generated by a plane $\mathbf{\Pi}$ between three images:

$$\mathbf{H}_{13} = \mathbf{H}_{23}\mathbf{H}_{12} \tag{28}$$

This result is obvious for the infinite homography considering the equation (6), but it is also easy to see that it is true for any homography, by decomposing each homography $\mathbf{H}_{ij}$ as a product of the homographies between retina $\mathcal{R}_j$ and plane $\mathbf{\Pi}$, and between plane $\mathbf{\Pi}$ and retina $\mathcal{R}_i$. If the epipoles are determined by the formula (18), then we have also easily, using this formula, and equation (6):

$$\mathbf{e}_{31} = \mathbf{H}_{\infty 23}\mathbf{e}_{21} + \mathbf{e}_{32} \tag{29}$$

Thus it can be seen that the formulas for composition of affine[4] descriptions have the same form than the one for composition of Euclidean description. In this sense, they can be thought as "uncalibrated motion", as already noted in section 2.2.

The reader may wonder why these projective quantities can be treated as Euclidean values. The reason is that when starting from the formulas (6) and (18), scale factors coherent with the three projection matrices are automatically obtained. But if not starting from the projection matrices, the scale factors are undetermined, as expected. That means that we have to normalize the description, if we are to use more than two views, as it was done in the case of the similarity-invariant description, as opposed to the displacement-invariant one. Thus in the description, we have to replace $\mathbf{e}_{ij}$ by the normalized epipole $\mathbf{e}_{Nij} = \mathbf{e}_{ij}/||\mathbf{e}_{ij}||$. We then see that the reasoning introduced in the preceding paragraph still holds. Thus a complete description for the affine parameters of three views includes the descriptions between views 1 and 2, 2 and 3, and the ratio $\beta_2 = ||\mathbf{e}_{21}||/||\mathbf{e}_{32}||$, these quantities being defined by (18). It can be noted that β_2 can be computed from α_2 and the elements of the similarity invariant description.

Projective representation The reason why the similarity-invariant description yields more complicated composition relations than the displacement-invariant description is that we have in the first case an indetermination in the description, which is represented by the last element of the canonic similarity matrix, a non-measurable quantity. Only the ratios of these quantities can be obtained, using the three descriptions together. We have just seen that the behavior of the affine description is the same than the one of the similarity-invariant in this respect. Now the projective-invariant description in turn has more non-measurable variables, which are the last row of this matrix, thus the composition relations are even more complicated. If we start from the projection matrices, the following relationswhich are proved in Appendix B hold:

$$\mathbf{e}_{31} = \mathbf{S}_{23}\mathbf{e}_{21} + \mathbf{e}_{32}(\mathbf{q}_2^T\mathbf{e}_{21}) \qquad \mathbf{S}_{13} = \mathbf{S}_{23}\mathbf{S}_{12} + \mathbf{e}_{32}(\mathbf{q}_2^T\mathbf{S}_{12} + \frac{\mathbf{e}_{12}^T}{||\mathbf{e}_{12}||^2}) + \mathbf{e}_{31}\mathbf{q}_1^T \tag{30}$$

[4]Note that (29) does not hold with general homographies.

where $\mathbf{q}_2 = \mathbf{r}_{\infty 23} - \mathbf{r}_{\infty 21}$, and $\mathbf{q}_1 = \mathbf{r}_{\infty 12} - \mathbf{r}_{\infty 13}$, the vectors $\mathbf{r}_{\infty ij}$ being the non-measurable quantities defined in (17). Now introducing the scale factors $\beta_1 = \|\mathbf{e}_{31}\|/\|\mathbf{e}_{21}\|$, $\beta_2 = \|\mathbf{e}_{32}\|/\|\mathbf{e}_{21}\|$, we can write these relations:

$$\begin{aligned} \beta_1 \mathbf{e}_{N31} &= \mathbf{S}_{23}\mathbf{e}_{N21} + \beta_2 \mathbf{e}_{N32}(\mathbf{q}_{N2}^T \mathbf{e}_{N21}) && (31) \\ \mathbf{S}_{13} &= \mathbf{S}_{23}\mathbf{S}_{12} + \beta_2 \mathbf{e}_{N32}(\mathbf{q}_{N2}^T \mathbf{S}_{12} + \gamma_1 \mathbf{e}_{N12}^T) + \beta_1 \mathbf{e}_{N31}\mathbf{q}_{N1}^T && (32) \end{aligned}$$

where $\gamma_1 = \|\mathbf{e}_{21}\|/\|\mathbf{e}_{12}\|$, and:

$$\begin{aligned} \mathbf{q}_{N1} = \; & \|\mathbf{e}_{21}\|\mathbf{q}_1 = \; \mathbf{r}_{\infty N12} - \frac{1}{\beta_1}\mathbf{r}_{\infty N13} \\ \mathbf{q}_{N2} = \; & \|\mathbf{e}_{21}\|\mathbf{q}_2 = \; \frac{1}{\beta_2}\mathbf{r}_{\infty N23} - \gamma_1 \mathbf{r}_{\infty N21} \end{aligned}$$

all the normalized vectors (denoted with a subscript N) being defined in (17). Note that the quantities γ_1 and $\mathbf{e}_{12}$ appearing in these equations are not part of the description 1-2, however they can be derived from it although there is no simple closed-form formula, by noting that by definition $\mathbf{S}_{12}\mathbf{e}_{12} = 0$, which gives $\mathbf{e}_{N12}$. The ratio γ_1 is computed by assigning an identical value to an identical component in the kernels of $\mathbf{S}_{12}^T$ and of $\mathbf{S}_{12}$.

By analogy with the previous geometries where only a ratio of the non-measurable quantities (the norms of translations or epipoles) was available, we see that only a difference of non-measurable vectors is accessible. The following relations are verified with easy algebra by substracting the canonic description equations (21) and (22) between 1 and 3:

$$\begin{aligned} \mathbf{P}_3 &= (\mathbf{S}_{13} - \mathbf{e}_{31}\mathbf{q}_1^T)\mathbf{P}_1 + \mathbf{e}_{31}\mathcal{L}_{12}^T = && (\mathbf{S}_{13} - \beta_1 \mathbf{e}_{N31}\mathbf{q}_{N1}^T)\mathbf{P}_1 + \beta_1 \mathbf{e}_{N31}\mathcal{L}_{N12}^T && (33) \\ \mathbf{p}_3 &= (\mathbf{S}_{13} - \mathbf{e}_{31}\mathbf{q}_1^T)\mathbf{p}_1 + \mathbf{e}_{31}\nu_{12} = && (\mathbf{S}_{13} - \beta_1 \mathbf{e}_{N31}\mathbf{q}_{N1}^T)\mathbf{p}_1 + \beta_1 \mathbf{e}_{N31}\nu_{N12} && (34) \end{aligned}$$

Let us give an interpretation of these results. From equations (31) and (32), it can be seen that the quantities $\beta_1 \mathbf{e}_{N31}$ and $\mathbf{S}_{13} - \beta_1 \mathbf{e}_{N31}\mathbf{n}_{N1}^T$ can be computed from the elements of the invariant descriptions 1-2 and 2-3, the 3D vector $\mathbf{q}_{N2}$, and the ratio β_2. Equivalently, these quantities could be obtained from the description 1-3, the 3D vector $\mathbf{q}_{N1}$, and the ratio β_1, which gives a complete description together with 1-2. This means that the description depends on 7+7+3+1=18 parameters. This result proves that the three fundamental matrices $\mathbf{F}_{12}$, $\mathbf{F}_{23}$, and $\mathbf{F}_{13}$ are not independent, but are linked by three equations. This may be an explanation why the purely projective self-calibration algorithms based on the Kruppa equations presented in [6, 17] do not perform as well as the formulation based on the global minimization of epipolar constraints with respect to Euclidean parameters [15, 17]: the latter formulation allows to take into account the composition constraints and thus to proceed with a minimal parameterization, whereas the former misses these constraints. A geometric argument enables to see what the three constraints are. Let us consider $\mathbf{F}_{31}\mathbf{e}_{32}$, the epipolar line of $\mathbf{e}_{32}$ in the first image. This line is the projection of the line $\langle \mathbf{C}_2, \mathbf{C}_3 \rangle$ in the first retina, through the optical center $\mathbf{C}_1$. Thus it is the intersection of the *trifocal plane* [8, 37] $\langle \mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3 \rangle$ with the first retina. Now the epipoles $\mathbf{e}_{12}$ and $\mathbf{e}_{13}$ also belong to this plane and to the first retina, thus the epipolar line $\mathbf{e}_{12} \times \mathbf{e}_{13}$ is identical to $\mathbf{F}_{31}\mathbf{e}_{32}$. By a circular permutation of indices follow the three equations:

$$\mathbf{F}_{31}\mathbf{e}_{32} = \mathbf{e}_{12} \times \mathbf{e}_{13} \qquad \mathbf{F}_{12}\mathbf{e}_{13} = \mathbf{e}_{23} \times \mathbf{e}_{21} \qquad \mathbf{F}_{23}\mathbf{e}_{21} = \mathbf{e}_{31} \times \mathbf{e}_{32} \tag{35}$$

3.2 The canonic decomposition for three (and more) views

Three projective views are now considered, and their most complete description is given through the three projection matrices $\tilde{\mathbf{P}}_i = [\mathbf{P}_i, \mathbf{p}_i]$, $i = 1, 2, 3$, totalizing 33 independent parameters. The canonic decomposition for three views is defined as the unique representation:

$$\tilde{\mathbf{P}}_1 = \mathcal{I}_1 \mathcal{T} \qquad \tilde{\mathbf{P}}_2 = \mathcal{I}_2 \mathcal{T} \qquad \tilde{\mathbf{P}}_3 = \mathcal{I}_3 \mathcal{T} \tag{36}$$

EUCLIDEAN (calibration)	$\mathcal{D} \in \mathcal{SO}_3$		**displacement**	6
	$\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3$		intrinsic parameters	5+5+5
invariant descriptions	$\mathbf{R}_{12},\ \mathbf{R}_{23}$	$\mathbf{R}_{12},\ \mathbf{R}_{13}$	rotations	3+3
	$\mathbf{T}_{12},\ \mathbf{T}_{23}$	$\mathbf{T}_{12},\ \mathbf{T}_{13}$	translations	3+3
canonic decomposition	$\begin{cases} \tilde{\mathbf{P}}_1 = \mathbf{A}_1[\mathbf{I}_3, 0]\mathcal{D} \\ \tilde{\mathbf{P}}_2 = \mathbf{A}_2[\mathbf{R}_{12}, \mathbf{T}_{12}]\mathcal{D} \\ \tilde{\mathbf{P}}_3 = \begin{array}{l} \mathbf{A}_3[\mathbf{R}_{13}\quad, \mathbf{T}_{13}]\mathcal{D} \\ \mathbf{A}_3[\mathbf{R}_{23}\mathbf{R}_{12}, \mathbf{R}_{23}\mathbf{T}_{12} + \mathbf{T}_{23}]\mathcal{D} \end{array} \end{cases}$			
EUCLIDEAN (motion)	$\mathcal{S} \in \lambda\mathcal{SO}_3$		**similarity**	7
	$\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3$		intrinsic parameters	5+5+5
invariant descriptions	$\mathbf{R}_{12},\ \mathbf{R}_{23}$	$\mathbf{R}_{12},\ \mathbf{R}_{13}$	rotations	3+3
	$\mathbf{t}_{12},\ \mathbf{t}_{23}$	$\mathbf{t}_{12},\ \mathbf{t}_{13}$	directions of translations	2+2
	α_1	α_2	ratios of translation norms	1
canonic decomposition	$\begin{cases} \tilde{\mathbf{P}}_1 = \mathbf{A}_1[\mathbf{I}_3, 0]\mathcal{S} \\ \tilde{\mathbf{P}}_2 = \mathbf{A}_2[\mathbf{R}_{12}, \mathbf{t}_{12}]\mathcal{S} \\ \tilde{\mathbf{P}}_3 = \begin{array}{l} \mathbf{A}_3[\mathbf{R}_{13}\quad, \alpha_1\mathbf{t}_{13}]\mathcal{S} \\ \mathbf{A}_3[\mathbf{R}_{23}\mathbf{R}_{12}, \mathbf{R}_{23}\mathbf{t}_{12} + \alpha_2\mathbf{t}_{23}]\mathcal{S} \end{array} \end{cases}$			$\alpha_1 = \Vert\mathbf{T}_{13}\Vert/\Vert\mathbf{T}_{12}\Vert$ $\alpha_2 = \Vert\mathbf{T}_{23}\Vert/\Vert\mathbf{T}_{12}\Vert$
AFFINE	$\mathcal{A} \in \mathcal{GA}_3$		**affine transformation**	12
	$\mathbf{H}_{\infty 12},\mathbf{H}_{\infty 23}$	$\mathbf{H}_{\infty 12},\mathbf{H}_{\infty 13}$	infinity homographies	8+8
invariant descriptions	$\mathbf{e}_{N21},\mathbf{e}_{N32}$	$\mathbf{e}_{N21},\mathbf{e}_{N31}$	normalized epipoles	2+2
	β_1	β_2	ratios of epipole norms	1
canonic decomposition	$\begin{cases} \tilde{\mathbf{P}}_1 = [\mathbf{I}_3, 0]\mathcal{A} \\ \tilde{\mathbf{P}}_2 = [\mathbf{H}_{\infty 12}, \mathbf{e}_{N21}]\mathcal{A} \\ \tilde{\mathbf{P}}_3 = \begin{array}{l} [\mathbf{H}_{\infty 13}\quad, \beta_1\mathbf{e}_{31N}]\mathcal{A} \\ [\mathbf{H}_{\infty 23}\mathbf{H}_{\infty 12}, \mathbf{H}_{\infty 23}\mathbf{e}_{N21} + \beta_2\mathbf{e}_{N32}]\mathcal{A} \end{array} \end{cases}$			$\beta_1 = \Vert\mathbf{e}_{31}\Vert/\Vert\mathbf{e}_{21}\Vert = \alpha_1\Vert\mathbf{A}_3\mathbf{t}_{13}\Vert/\Vert\mathbf{A}_2\mathbf{t}_{12}\Vert$ $\beta_2 = \Vert\mathbf{e}_{32}\Vert/\Vert\mathbf{e}_{21}\Vert = \alpha_2\Vert\mathbf{A}_3\mathbf{t}_{23}\Vert/\Vert\mathbf{A}_2\mathbf{t}_{12}\Vert$
PROJECTIVE	$\mathcal{H} \in \mathcal{GL}_4$		**homography**	15
	$\mathbf{S}_{12},\mathbf{S}_{23}$	$\mathbf{S}_{12},\mathbf{S}_{13}$	epipolar projections	5+5
invariant descriptions	$\mathbf{e}_{N21},\mathbf{e}_{N32}$	$\mathbf{e}_{N21},\mathbf{e}_{N31}$	normalized epipoles	2+2
	$\mathbf{q}_{N1}$	$\mathbf{q}_{N2}$	differences of r_∞-vectors	3
	β_1	β_2	ratios of epipole norms	1
canonic decomposition	$\begin{cases} \tilde{\mathbf{P}}_1 = [\mathbf{I}_3, 0]\mathcal{H} \\ \tilde{\mathbf{P}}_2 = [\mathbf{S}_{12}, \mathbf{e}_{N21}]\mathcal{H} \\ \tilde{\mathbf{P}}_3 = \begin{array}{l} [\mathbf{S}_{13} - \beta_1\mathbf{e}_{31N}\mathbf{q}_{N1}^T, \beta_1\mathbf{e}_{31N}]\mathcal{H} \\ [\mathbf{S}_{23}\mathbf{S}_{12} + \beta_2\mathbf{e}_{N32}(\mathbf{q}_{N2}^T\mathbf{S}_{12} + \gamma_1\mathbf{e}_{N12}^T)\ , \\ \qquad \mathbf{S}_{23}\mathbf{e}_{N21} + \beta_2\mathbf{e}_{N32}(\mathbf{q}_{N2}^T\mathbf{e}_{N21})]\mathcal{H} \end{array} \end{cases}$			$\mathbf{q}_{N1} = \mathbf{H}_{\infty 12}^T\mathbf{e}_{N21} - \frac{1}{\beta_1}\mathbf{H}_{\infty 13}^T\mathbf{e}_{N31}$ $\mathbf{q}_{N2} = \frac{1}{\beta_2}\mathbf{H}_{\infty 23}^T\mathbf{e}_{N32} - \gamma_1\mathbf{H}_{\infty 21}^T\mathbf{e}_{N12}$ $\beta_1 = \Vert\mathbf{e}_{31}\Vert/\Vert\mathbf{e}_{21}\Vert$ $\beta_2 = \Vert\mathbf{e}_{32}\Vert/\Vert\mathbf{e}_{21}\Vert$ *with* $\gamma_1 = \Vert\mathbf{e}_{21}\Vert/\Vert\mathbf{e}_{12}\Vert$

Table 2: The geometries of three views: canonic representation

where $\mathcal{I}_1$ and $\mathcal{I}_2$ have the same form than in the canonic decomposition for two views (24). Of course, the form of $\mathcal{I}_3$ is expected to be in general different from the form of $\mathcal{I}_2$. Let us list the consequences of this construction:

- The two-view canonic decomposition and its properties are extended: construction of an invariant description with respect to a given group of transformations, determination of the exact number of parameters in the representation and of the nature of the measurable and the non-measurable quantities. Explicit formulas are obtained, to build three projection matrices which are captures all the properties of a triple of views up to a transformation of $\mathcal{G}$, thus allowing to perform trinocular reconstruction up to a certain group of transformations.
- There are two descriptions for the invariants of three views, one build upon the pair of descriptions 1-2, 1-3, the other upon the pair of descriptions 1-2, 2-3. They are more than the two descriptions for two views, including some additional parameters, which cannot be determined from them. Rather, these parameters are functions of descriptions of the previous level.
- The equivalence of the two forms of the alternative descriptions for three views gives the dependency of the composed description 1-3 (resp. 2-3) over the descriptions for 1-2 and 2-3 (resp. 1-3), and the additional parameters.
- The additional parameters can be determined from the knowledge of the *three* descriptions 1-2, 2-3, and 1-3. It means that knowing all the triples of descriptions for two views are equivalent to a description for three views. The formulas are given explicitly in the next section. From the count of parameters, it is seen that the triple of descriptions for two views is not a minimal representation.

In order to make a purely algebraic derivation for the last points, we notice that simultaneously to (36), the following canonic representations of two views must hold:

$$\tilde{\mathbf{P}}_1 = \mathcal{I}'_1 \mathcal{T}' \quad \tilde{\mathbf{P}}_3 = \mathcal{I}'_2 \mathcal{T}' \tag{37}$$

$$\tilde{\mathbf{P}}_2 = \mathcal{I}''_1 \mathcal{T}'' \quad \tilde{\mathbf{P}}_3 = \mathcal{I}''_2 \mathcal{T}'' \tag{38}$$

where all the quantities with the same subscript have to be of the same form. In Appendix A, we give as an example the derivation of the composition properties for the affine case using this approach. The groups of transformations considered are the same than in section (2.5), and we have not repeated their properties, nor have we repeated the definitions of the elements of the canonic representation for two views, since these elements are already found in table 1. We have summarized in table 2 the results specific to the canonic decomposition of a triple of projection matrices:

- the nature of the two equivalent invariant description, the quantities above the horizontal line being the elements of the invariant description for two views, the quantities under that line being the additional parameters, which are measurable from three views but not from two pairs of invariant description for two views,
- the two alternative expressions for $\tilde{\mathbf{P}}_3$, as a function of the description 1-2, 1-3, or of the descriptions 1-2, 2-3,
- the definition of the additional elements, as a function of the description of previous level,
- the number of parameters, whose sum is exactly 33.

One advantage of the previous formalism is that the generalization of the canonic decomposition to the case of N views is straightforward, since to build a description for N views is just a matter of considering the triplet of the first and second views and the Nth view, or alternatively any triple of views including the Nth view. Thus the elements of the description are exactly the same than for three views, and can be summarized in the table 3 where it can be verified that the total number of parameters is $11N$:

EUCLIDEAN (calibration)	displacement	6
	intrinsic parameters	5N
	rotations	3 (N-1)
	translations	3 (N-1)
EUCLIDEAN (motion)	similarity	7
	rotations	3 (N-1)
	directions of translations	2 (N-1)
	ratio of translation norms	N-2
AFFINE	affine transformation	12
	infinity homographies	8 (N-1)
	normalized epipoles	2 (N-1)
	ratios of epipole norms	N-2
PROJECTIVE	homography	15
	epipolar projections	5 (N-1)
	normalized epipoles	2 (N-1)
	differences of r_∞ vectors	3 (N-2)
	ratios of epipole norms	N-2

Table 3: The geometries of N views: canonic representation

3.3 From triples of two views to three views

In this section, we give details about the computation of the three views invariant description which has been described previously from when triples of two views invariant descriptions. Scale factors indeterminations are discussed.

Euclidean representation First, there is no ambiguity for the rotational part. For the translational part, if we take by definition $\mathbf{t}_{ij} = \mathbf{T}_{ij}/||\mathbf{T}_{ij}||$, as previously done, then there is no sign ambiguity. But in the structure from motion paradigm, the sign information contained in $\mathbf{T}_{ij}$ is also lost in the direction of translation $\mathbf{t}_{ij}$, a fact we take into account by writing that the recovered quantities are: $\mathbf{t}_{12} = \varepsilon_{12}\mathbf{T}_{12}/||\mathbf{T}_{12}||$ and $\mathbf{t}_{23} = \varepsilon_{23}\mathbf{T}_{23}/||\mathbf{T}_{23}||$. In the relation (27), the definition of the scale factors is thus: $\alpha_1 = \varepsilon_{12}\varepsilon_{13}||\mathbf{T}_{13}||/||\mathbf{T}_{12}||$ and $\alpha_2 = \varepsilon_{12}\varepsilon_{23}||\mathbf{T}_{23}||/||\mathbf{T}_{12}||$.

However, the relation (27) constraints $\mathbf{t}_{13}$ to lie in the plane $\langle \mathbf{R}_{23}\mathbf{t}_{12}, \mathbf{t}_{23}\rangle$, and thus if $\mathbf{t}_{12}$, $\mathbf{t}_{23}$, and $\mathbf{t}_{13}$ are all known, the ratio α_2 can be computed by expressing the proportionality constraint:

$$\mathbf{t}_{13} \times (\mathbf{R}_{23}\mathbf{t}_{12} + \alpha_2\mathbf{t}_{23}) = 0 \tag{39}$$

Once α_2 is determined, α_1 can also be computed. We obtain:

$$\alpha_2 = -\frac{\mathbf{t}_{13} \times \mathbf{R}_{23}\mathbf{t}_{12}}{\mathbf{t}_{13} \times \mathbf{t}_{23}} \qquad \alpha_1 = \frac{\mathbf{R}_{23}\mathbf{t}_{12} + \alpha_2\mathbf{t}_{23}}{\mathbf{t}_{13}} = -\frac{\mathbf{t}_{23} \times \mathbf{R}_{13}\mathbf{R}_{21}^{-1}\mathbf{t}_{12}}{\mathbf{t}_{13} \times \mathbf{t}_{23}} \qquad \text{for } i = 1,2,3 \tag{40}$$

where we have used the usual division symbol for the the vector division of two proportional vectors: if $\mathbf{v}_1 = \lambda\mathbf{v}_2$, then $\mathbf{v}_1/\mathbf{v}_2 = \lambda$. Please note that it would be incorrect to consider only the quotient of the norms, since the signs would be lost.

Affine representation Using the fact that the composition laws are similar, and applying the same technique, it is found that: $\beta_1\mathbf{e}_{N31} = \mathbf{H}_{\infty 23}\mathbf{e}_{N21} + \beta_2\mathbf{e}_{N32}$ with:

$$\beta_2 = -\frac{\mathbf{e}_{N31} \times \mathbf{H}_{\infty 23}\mathbf{e}_{N21}}{\mathbf{e}_{N31} \times \mathbf{e}_{N32}} \qquad \beta_1 = \frac{\mathbf{e}_{N32} \times \mathbf{H}_{\infty 13}\mathbf{H}_{\infty 12}^{-1}\mathbf{e}_{N21}}{\mathbf{e}_{N31} \times \mathbf{e}_{N32}} \qquad \text{for } i = 1,2,3 \tag{41}$$

In addition to the indetermination on the epipole, we have to take into account the fact that in the affine description for two views, the infinity homography is defined only up to a scale factor.

It means that we know only the matrices $\lambda_{12}\mathbf{H}_{\infty 12}$, $\lambda_{23}\mathbf{H}_{\infty 23}$, $\lambda_{13}\mathbf{H}_{\infty 13}$. Expliciting also the dependencies on the scale factors which are in (41), the invariant descriptions become:

$$\left\{\begin{array}{l}\mathcal{I}_1 = [\mathbf{I}_3, 0] \\ \mathcal{I}_2 = [\lambda_{12}\mathbf{H}_{\infty 12}, \mathbf{e}_{N21}] \\ \mathcal{I}_3 = [\lambda_{13}\mathbf{H}_{\infty 13}, \lambda_{13}\lambda_{12}^{-1}\beta_1\mathbf{e}_{31N}]\end{array}\right. \qquad \left\{\begin{array}{l}\mathcal{I}_1 = [\mathbf{I}_3, 0] \\ \mathcal{I}_2 = [\lambda_{12}\mathbf{H}_{\infty 12}, \mathbf{e}_{N21}] \\ \mathcal{I}_3 = [\lambda_{12}\lambda_{23}\mathbf{H}_{\infty 23}\mathbf{H}_{\infty 12}, \lambda_{23}\mathbf{H}_{\infty 23}\mathbf{e}_{N21} + \lambda_{23}\beta_2\mathbf{e}_{N32}]\end{array}\right.$$

Let us remark that the pairs of projection matrices which yield the same $\mathbf{H}_\infty$,$\mathbf{e}'$ description than $\{[\mathbf{P}, \mathbf{p}]\ ,\ [\mathbf{P}', \mathbf{p}']\}$ are of the form: $\{[\lambda\mu\mathbf{P}, \lambda\mathbf{p}]\ ,\ [\lambda'\mu\mathbf{P}', \lambda'\mathbf{p}']\}$. Then we see that both the invariant descriptions are not affected by a change of the unknown scale factors λ_{12}, λ_{23}, λ_{13}. Note that this result depends on the way the ratios β_i are computed. If, for example, the ratio β_1 was computed from $\mathbf{H}_{\infty 23}$ instead of $\mathbf{H}_{\infty 13}\mathbf{H}_{\infty 12}^{-1}$, the result would not hold.

Projective representation Let first deal with the first canonic decomposition, and compute β_1 and $\mathbf{q}_{N1}$ from the pairs of descriptions 1-2, 1-3, and 2-3. The following pairs of projection matrices have to yield the same projective view invariants:

$$\left\{\begin{array}{l}[\mathbf{I}_3, 0] \\ [\mathbf{S}_{32}, \mathbf{e}_{N23}]\end{array}\right. \qquad \left\{\begin{array}{l}[\frac{1}{\beta_1}\mathbf{S}_{13} - \mathbf{e}_{N31}\mathbf{q}_{N1}^T, \mathbf{e}_{N31}] \\ [\mathbf{S}_{12}, \mathbf{e}_{N21}]\end{array}\right.$$

Writing that the fundamental matrices and epipoles obtained from the two pairs are proportional yield eventually the two equations:

$$\begin{array}{l}([\mathbf{e}_{N23}]_\times\mathbf{S}_{32})(\mathbf{S}_{13} - \mathbf{e}_{N31}\beta_1\mathbf{q}_{N1}^T) \sim ([\mathbf{e}_{N23}]_\times\mathbf{S}_{12}) \\ \mathbf{e}_{N23} \times (\mathbf{e}_{N21} - \beta_1\mathbf{S}_{12}(\mathbf{S}_{13} - \mathbf{e}_{N31}\beta_1\mathbf{q}_{N1}^T)^{-1}\mathbf{e}_{N31}) = 0\end{array}$$

The solution is obtained by:

$$\beta_1(\mathbf{q}_{N1})_k = \frac{([\mathbf{e}_{N23}]_\times\mathbf{S}_{32}\mathbf{S}_{13})_k \times ([\mathbf{e}_{N23}]_\times\mathbf{S}_{12})_k}{([\mathbf{e}_{N23}]_\times\mathbf{S}_{32}\mathbf{e}_{N31}) \times ([\mathbf{e}_{N23}]_\times\mathbf{S}_{12})_k}\ ,\quad \beta_1 = \frac{\mathbf{e}_{N23} \times \mathbf{e}_{N21}}{\mathbf{e}_{N23} \times \mathbf{S}_{12}(\mathbf{S}_{13} - \mathbf{e}_{N31}(\beta_1\mathbf{q}_{N1})^T)^{-1}\mathbf{e}_{N31}} \tag{42}$$

where the notation $(\mathbf{M})_k$ (resp. $(\mathbf{v})_k$) designates the kth column vector (resp. component) of a matrix (resp. vector). From the formulas given above, it is easy to see that the solution is not affected by a scaling of the matrices $\mathbf{S}_{ij}$, since β_1 is proportional to λ_{13}^{-1} and $\mathbf{q}_{N1}$ to λ_{13}^2.

The second canonic decomposition is dealt with in the same way, by expressing the projective equivalence of the pairs of projection matrices:

$$\left\{\begin{array}{l}[\mathbf{I}_3, 0] \\ [\mathbf{S}_{31}, \mathbf{e}_{N13}]\end{array}\right. \qquad \left\{\begin{array}{l}[\mathbf{S}_{23}\mathbf{S}_{12} + \beta_2\mathbf{e}_{N32}(\mathbf{q}_{N2}^T\mathbf{S}_{12} + \gamma_1\mathbf{e}_{N12}^T)\ , \mathbf{S}_{23}\mathbf{e}_{N21} + \beta_2\mathbf{e}_{N32}(\mathbf{q}_{N2}^T\mathbf{e}_{N21})] \\ [\mathbf{I}_3, 0]\end{array}\right.$$

In writing the proportionality of the fundamental matrices, the term with γ_1 cancels because of (35), and we are left with the equation:

$$[\mathbf{e}_{N13}]_\times\mathbf{S}_{31}(\mathbf{S}_{23}\mathbf{S}_{12} + \mathbf{e}_{N32}(\beta_2\mathbf{q}_{N2}^T)\mathbf{S}_{12})\ \sim\ [\mathbf{e}_{N13}]_\times$$

which can be solved similarly for $\beta_2\mathbf{q}_{N2}$. Then β_2 is determined by solving:

$$(\mathbf{S}_{23}\mathbf{S}_{12} + \mathbf{e}_{N32}((\beta_2\mathbf{q}_{N2})^T\mathbf{S}_{12} + \beta_2\gamma_1\mathbf{e}_{N12}^T))\mathbf{e}_{N13} \times (\mathbf{S}_{23}\mathbf{e}_{N21} + \mathbf{e}_{N32}((\beta_2\mathbf{q}_{N2})^T\mathbf{e}_{N21})) = 0$$

We have just given a sketch of the solution, and have not explicited it, because in practice, one would prefer to deal with the first equivalent canonic decomposition, which is less complicated. It can be noted that the simpler formulas[5]:

$$\beta_2 = \frac{\mathbf{e}_{N31} \times \mathbf{S}_{13}\mathbf{e}_{N12}}{\gamma_1\mathbf{e}_{N31} \times \mathbf{e}_{N32}} \qquad \beta_2\mathbf{q}_{N2} = \gamma_1\frac{\mathbf{S}_{21}^T\mathbf{S}_{13}^T\mathbf{e}_{N32} + \mathbf{e}_{N32}^T\mathbf{e}_{N31}\mathbf{S}_{23}^T\mathbf{e}_{N31}}{1 - (\mathbf{e}_{N32}^T\mathbf{e}_{N31})^2}$$

do not give consistent solutions when the scale of the matrices $\mathbf{S}_{ij}$ is changed.

[5]The first one is obtained by equalizing the two first 3×3 submatrix of the two forms of the canonic decomposition, then taking the cross-product with $\mathbf{e}_{31}$ and the dot-product with $\mathbf{e}_{N12}$, the second is from [37].

4 Some consequences of the representations

4.1 Relations between the levels of representation

From tables 1 and 2 we remark that each invariant description of a given level is formulated in terms of descriptions of the previous level. This heritage of descriptors is quite natural, except for the link between the similarity-invariant representations and the affine representations where a different change of retinal coordinates occur. It can be seen that the simplest representation is the affine one.

Projective representation and affine representation Since the infinite homography $\mathbf{H}_\infty$ included in the affine representation is compatible with the fundamental matrix, and since the epipole $\mathbf{e}'$ is also part of the representation, it can be seen that all the results of section 2.3 apply entirely. In particular, we have:

$$\mathbf{F} = [\mathbf{e}']_\times \mathbf{H}_\infty \qquad \mathbf{S} = -\frac{[\mathbf{e}']_\times^2}{\|\mathbf{e}'\|^2}\mathbf{H}_\infty$$

The knowledge of the infinity homography does not constraint the projective representation more than does the knowledge of any homography. However, the infinity homography, unlike the other homographies, is guaranteed to be compatible with the fundamental matrix.

Let us examine the case of three views. From the two fundamental matrices $\mathbf{F}_{12}$ and $\mathbf{F}_{23}$ alone, there is no way to obtain the fundamental matrix $\mathbf{F}_{13}$, since four additional parameters are needed. The canonic representation shows also that adding the homography matrices $\mathbf{H}_{\infty 12}$ and $\mathbf{H}_{\infty 23}$ is not sufficient, but that the scale ratio β_2 is also needed. In order to cope with the problem of scale factors, let us work with the non-normalized quantities, and establish the interesting relation:

$$\begin{aligned} \mathbf{F}_{13} &= \mathbf{H}^*_{\infty 23}\mathbf{F}_{12} + \mathbf{F}_{23}\mathbf{H}_{\infty 12} = (\det \mathbf{H}_{\infty 23})(\mathbf{H}^T_{\infty 32}\mathbf{F}_{12} + \mathbf{F}^T_{32}\mathbf{H}_{\infty 12}) \\ &= (\det \mathbf{H}_{\infty 23})(\mathbf{H}^T_{\infty 32}([\mathbf{e}_{21}]_\times - [\mathbf{e}_{23}]_\times)\mathbf{H}_{\infty 12}) \end{aligned} \tag{43}$$

Applying the property (20), and then (29) and (28) yield the first equality of (43):

$$\mathbf{H}^*_{\infty 23}\mathbf{F}_{12} = \mathbf{H}^*_{\infty 23}[\mathbf{e}_{21}]_\times \mathbf{H}_{\infty 12} = [\mathbf{e}_{31} - \mathbf{e}_{32}]_\times \mathbf{H}_{\infty 23}\mathbf{H}_{\infty 12} = \underbrace{[\mathbf{e}_{31}]_\times \mathbf{H}_{\infty 13}}_{\mathbf{F}_{13}} - \underbrace{[\mathbf{e}_{32}]_\times \mathbf{H}_{\infty 23}}_{\mathbf{F}_{23}}\mathbf{H}_{\infty 12}$$

To obtain the second equality, we also use (23). The last member is obtained by substitution of the decompositions of $\mathbf{F}_{12}$ and $\mathbf{F}_{32}$. It allows to write the following expression using normalized quantities:

$$\mathbf{F}_{13} \sim \mathbf{H}^T_{\infty 32}(\gamma_2[\mathbf{e}_{N21}]_\times - \beta_2[\mathbf{e}_{N23}]_\times)\mathbf{H}_{\infty 12} \tag{44}$$

where $\gamma_2 = \|\mathbf{e}_{32}\|/\|\mathbf{e}_{23}\|$ is obtained in the same way than the ratio γ_1 appearing in (32).

It can be noted that the system of equations obtained by writing (44) between the three images can not determine the infinity homography matrices from the knowledge of the three fundamental matrices, because there are 21 parameters in the affine representation, versus only 18 in the projective representation. However, it can be seen that this system determines the infinity homography matrices as soon as one of them is known. By substitution of $\mathbf{H}_{\infty 32} = \mathbf{S}_{32} + \mathbf{e}_{N23}\mathbf{r}^T_{\infty 32}$ into (44), and some easy algebra, the following formula is obtained:

$$(\mathbf{r}_{\infty 32})_k = \frac{(\mathbf{H}^T_{\infty 12}(\beta_2[\mathbf{e}_{N23}]_\times - \gamma_2[\mathbf{e}_{N21}]_\times)\mathbf{S}_{32})_k \times (\mathbf{F}_{31})_k}{\gamma_2(\mathbf{H}^T_{\infty 12}[\mathbf{e}_{N21}]_\times \mathbf{e}_{N23}) \times (\mathbf{F}_{31})_k} \tag{45}$$

where the notations are the same than in (42). The additional knowledge needed correspond for example to one of the three vectors $\mathbf{r}_\infty$ defined in (17), which identify the plane at infinity. This means that is this quantity is identified anywhere in an image sequence, it can be *propagated* along the whole sequence.

Euclidean representation and affine representation It is easy to see that the relation:

$$\mathbf{H}_\infty = \mathbf{A}'\mathbf{R}\mathbf{A}^{-1} \tag{46}$$

together with the relation between the fundamental and the essential matrix [6]:

$$\mathbf{F} = \mathbf{A}'^{-1T}[\mathbf{t}]_\times \mathbf{R}\mathbf{A}^{-1} \tag{47}$$

which entails $\mathbf{e}' = \mathbf{A}'\mathbf{t}$ allows to determine directly the motion parameters, the rotation $\mathbf{R}$ and the direction of translation $\mathbf{t}$ from the affine representation, if the intrinsic parameters are determined.

We examine now the relation with the intrinsic parameters. In the case of self-calibration in the projective framework, equation (47) was used to constraint the intrinsic parameters from the fundamental matrix by means of the rigidity constraint expressing that the matrix:

$$\mathbf{E} = \mathbf{A}'^{-1T}\mathbf{F}\mathbf{A}^{-1} \tag{48}$$

is an essential matrix [14] (product of an antisymmetric matrix by a rotation matrix, for several characterizations see [12, 21]), thus eliminating $\mathbf{R}$ and $\mathbf{t}$ [17]. Now by analogy, if we eliminate $\mathbf{R}$ from (46), by expressing that it is a rotation matrix, we will get equations relating the intrinsic parameters and the infinity homography matrix. This can be done much more simply than in the projective framework. The fact that $\mathbf{R}$ is a rotation matrix is equivalent to:

$$\mathbf{R}\mathbf{R}^T = \mathbf{I}_3 \tag{49}$$

Substituting $\mathbf{R} = \mathbf{A}'^{-1}\mathbf{H}_\infty\mathbf{A}$ obtained from equation (46) into (49), yield:

$$\mathbf{K}' = \mathbf{H}_\infty\mathbf{K}\mathbf{H}_\infty^T \tag{50}$$

where the matrices $\mathbf{K} = \mathbf{A}\mathbf{A}^T$ and $\mathbf{K}' = \mathbf{A}'\mathbf{A}'^T$ represent the dual of the image of the absolute conic in each camera coordinate system, as mentioned in section 2.1. Each of these matrices is symmetric and defined only up to a scale factor, thus they depend on five independent parameters. It can be seen that relation (50) allows us to *update* camera calibration through a sequence of images where they do not remain constant, once that the initial camera parameters are known.

We have obtained from (50) five constraints on the intrinsic parameters, which are linear, whereas from the projective invariants, only two quadratic constraints were obtained [22, 6]. These last constraints are in fact implied by the former ones.

Let us suppose that $\mathbf{A}$ and $\mathbf{A}'$ satisfy the constraint (50), then $\mathbf{H}_\infty\mathbf{A} = \mathbf{A}'\mathbf{R}$. The matrix $\mathbf{H}_\infty$ has also to be compatible with the epipolar geometry, thus, $\mathbf{F} = [\mathbf{e}']_\times\mathbf{H}_\infty$. By substitution of the second, and then the first relation into (48), we obtain: $\mathbf{E} = \mathbf{A}'^T[\mathbf{e}']_\times\mathbf{A}'\mathbf{R}$ which is the product of an antisymmetric matrix and a rotation matrix. Thus we have shown that the rigidity constraint (and in particular the Kruppa equations of [22, 6]) are indeed implied by (50), and do not yield additional equations.

If we have three views, then the relations (50) yield the following system of equations:

$$\left\{\begin{array}{l} \mathbf{K}_2 = \mathbf{H}_{\infty 12}\mathbf{K}_1\mathbf{H}_{\infty 12}^T \\ \mathbf{K}_3 = \mathbf{H}_{\infty 23}\mathbf{K}_2\mathbf{H}_{\infty 23}^T \\ \mathbf{K}_1 = \mathbf{H}_{\infty 31}\mathbf{K}_3\mathbf{H}_{\infty 31}^T \end{array}\right. \tag{51}$$

Since the last equation is a consequence of the two first ones, the system provides with at most 10 independent equations. This is coherent with a simple count of the number of parameters. We have seen that in the case of structure from motion, the Euclidean representation has 26 independent parameters, whereas the affine representation has 21 independent parameters. Thus the knowledge of 5 more variables is needed to determine the Euclidean representation from the affine representation. These parameters are for instance the set of intrinsic parameters of one of the three cameras, or, equivalently, the equation of the image of the absolute conic in one of the cameras.

4.2 Computing from the images

We discuss now which parts of the representations previously described can be recovered from measurements made in the images, and in particular we consider the self-calibration problem.

Uncalibrated cameras When working with uncalibrated cameras, the only information that is accessible are pixel coordinates. Let us suppose that we observe corresponding points in two images. They are the image of same point $\tilde{\mathbf{M}}$ of $\mathcal{P}^3$:

$$\kappa \mathbf{m} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} \qquad \kappa' \mathbf{m}' = \tilde{\mathbf{P}}'\tilde{\mathbf{M}}$$

From the canonic decomposition, we see that this can also be written:

$$\kappa \mathbf{m} = [\mathbf{I}_3, 0]\tilde{\mathbf{M}}_1 \qquad \kappa' \mathbf{m}' = [\mathbf{P}_c, \mathbf{p}_c]\tilde{\mathbf{M}}_1$$

where $[\mathbf{P}_c \mathbf{p}_c]$ are one of the canonic forms of the second projection matrix, either the projective or the affine one. The substitution of the first relation into the second yields:

$$\kappa' \mathbf{m}' = \kappa \mathbf{P}_c \mathbf{m} + \mathbf{p}_c \tag{52}$$

There are several consequences of this relation. The first one is that since the quantities κ and κ' are unknown, the only equation relating $\mathbf{m}$ and $\mathbf{m}'$ is obtained by writing that the vectors $\mathbf{m}'$, $\mathbf{P}_c\mathbf{m}$, and $\mathbf{p}_c$ are coplanar. This is a generalization of the Longuet-Higgins condition [14], as first pointed out to in [6]. By writing that the mixed product of the three vectors is zero, we obtain: $\mathbf{m}'^T[\mathbf{p}_c]_\times \mathbf{P}_c \mathbf{m} = 0$, which is the epipolar constraint, with fundamental matrix $\mathbf{F} = [\mathbf{p}_c]_\times \mathbf{P}_c$. Thus we find again that this fundamental matrix is the only information which can be obtained from image correspondences alone. We can notice that the projective decomposition and the affine decomposition yield the same fundamental matrix. A second consequence is that one can see equation (52) as a generalization of the relation: $Z'\mathbf{m}' = Z\mathbf{R}\mathbf{m} + \mathbf{t}$ which holds in the calibrated case, provided a normalized representation is taken for $\mathbf{m}$ and $\mathbf{m}'$. Thus, we can interpret the quantities κ and κ' as affine or projective depths, and use equation (52) to recover 3D structure up to an affine or projective transformation. Only the elements of the representation $\mathbf{P}_c$ and $\mathbf{p}_c$ are needed, as shown in [37]. An alternative interpretation of similar quantities can be found in [29].

Affine calibration To recover the three parameters which define the affine representation, in addition to the seven parameters representing the projective representation, some additional knowledge is needed. Several approximations and heuristics can be used. Drawing upon the concepts of section 2.2, there are two ways to compute $\mathbf{H}_\infty$:

- Identifying parallel directions: if images of three set of lines which are parallel in 3D space can be identified (a task addressed for example in [20, 25]) then we obtain three vanishing points $\mathbf{v}_i$ and $\mathbf{v}'_i$ in each image, and the infinity homography matrix is the solution of the system $\{\mathbf{H}_\infty \mathbf{v}_i = \lambda_i \mathbf{v}'_i\}, \mathbf{H}_\infty \mathbf{e} = \lambda \mathbf{e}'$, as mentioned by [24].
- Identifying infinity: the homography $\mathbf{H}_\infty$ is the limit of any homography induced by a plane, when the distance of the plane to the optical centers increases arbitrarily. Thus observing corresponding points which are at the horizon, or at remote distances provide an approximative way to compute $\mathbf{H}_\infty$, used already in [36].

If we write equation (52) in the affine case: $\kappa'_a \mathbf{m}' = \kappa_a \mathbf{H}_\infty \mathbf{m} + \mathbf{e}'$, we also see, since $\mathbf{e}' = \mathbf{A}'\mathbf{T}$ that:

- When the translational component of the motion can be neglected, the average homography (ie the homography computed from all the point correspondences) is an approximation of the infinity homography.

Recently, the following result has been proved and used in [38]:

- The average homography computed from a random set of points is almost equal to the homography of the infinity homography

The case of a single moving camera It has been shown [22, 6] that Euclidean representations can be recovered from three displacements of a single uncalibrated camera, which means a camera whose intrinsic parameters remain constant. However, it is not possible to use such a constraint to recover *directly* intermediate affine representations, since this information cannot be expressed within this representation.

If the two cameras are the same, then from $\mathbf{H}_\infty = \mathbf{A}\mathbf{R}\mathbf{A}^{-1}$ it is seen that the infinity homography matrix is similar to a rotation matrix. Considering that this matrix is recovered up to a scale factor, its eigenvalues λ_i are $\alpha, \alpha e^{i\theta}, \alpha e^{-i\theta}$. They have thus to satisfy two constraints, which are for example $||\lambda_1|| = ||\lambda_2||$ and $||\lambda_3||^3 = \lambda_1\lambda_2\lambda_3$. They result in two algebraic constraints on the entries of the infinity homography matrix which would be interesting to investigate further. We can conclude that in this case the matrix $\mathbf{H}_\infty$ depends only on six parameters. Note that these constraints should be incorporated in any algorithm to estimate the matrix $\mathbf{H}_\infty$ for a single camera. A consequence is that the natural approach to solve the linear system of equations $\mathbf{K} = \mathbf{H}_\infty \mathbf{K} \mathbf{H}_\infty^T$ for the Kruppa coefficients would fail, since this system would not provide with five independent equations. Let us make this point more precise. The equation (50) becomes:

$$\mathbf{K} = (\mathbf{A}\mathbf{R}\mathbf{A}^{-1})\mathbf{K}(\mathbf{A}\mathbf{R}\mathbf{A}^{-1})^T$$

Using the fact that $\mathbf{R}\mathbf{R}^T = \mathbf{I}_3$, this equation can be transformed into:

$$\mathbf{R}\underbrace{\mathbf{A}^{-1}\mathbf{K}\mathbf{A}^{-1T}}_{\mathbf{X}} = \underbrace{\mathbf{A}^{-1}\mathbf{K}\mathbf{A}^{-1T}}_{\mathbf{X}}\mathbf{R} \tag{53}$$

The matrix $\mathbf{K}$ and the matrix $\mathbf{X}$ are obtained from each other in a unique way. Using $\mathbf{X}$ as unknown, we see that (53) is a commutator equation, but with the additional condition that $\mathbf{X}$ has to be symmetric. All the solutions to this matrix equation $\mathbf{R}\mathbf{X} = \mathbf{X}\mathbf{R}$ can be obtained by using a rationnal parameterization of $\mathbf{R}$ by quaternions[6] which accounts for the orthogonal structure of the matrix $\mathbf{R}$, and writing explicitly the linear system of equations in the entries of $\mathbf{X}$. Such an approach, which is not detailed here, enables to prove that the rank of the system if 4. Let us instead give a less analytical solution. Obviously, the identity matrix $\mathbf{I}_3$ is a solution. If $\mathbf{u}$ is the rotation axis of the rotation $\mathbf{R}$, we have:

$$(\mathbf{R}\mathbf{u})\mathbf{u}^T = \mathbf{u}\mathbf{u}^T = \mathbf{u}(\mathbf{R}^T\mathbf{u})^T = \mathbf{u}\mathbf{u}^T\mathbf{R}$$

thus the symmetrical matrix $\mathbf{u}\mathbf{u}^T$ is the other solution, and we can conclude that the general solution is $\mathbf{X} = \lambda\mathbf{u}\mathbf{u}^T + \mu\mathbf{I}_3$. This means that by solving the equation (50) for the Kruppa matrix, we recover: $\mathbf{K} = \lambda\mathbf{A}\mathbf{u}(\mathbf{A}\mathbf{u})^T + \mu\mathbf{A}\mathbf{A}^T$, and thus there is an indetermination on $\mathbf{K}$ in the direction $\mathbf{A}\mathbf{u}(\mathbf{A}\mathbf{u})^T$. Thus two displacements with non-parallel rotation axes, or further constraints on the parameters, such as a restricted model, are necessary to recover $\mathbf{K}$ unambiguously.

As soon as three views are available, the knowledge of only one infinity homography matrix allows to obtain all the infinity homography matrices, and thus the Kruppa coefficients can be solved using the system of equations (51). Although the solution is more simple, since it involves only linear equations, we see that no more information is obtained than by proceeding only from the more general projective framework, where no affine calibration is necessary. Thus the best way to use affine information seems to be the updating scheme, where a camera is first self-calibrated, and then affine information is used to track its calibration parameters as they evolve over an image sequence.

5 Examples

Two examples are provided to illustrate the theory. The associated numerical values are given in appendix C. They involve real images, and tasks that we shall hopefully be able to perform automatically in a near future. Since our goal here is not to address stability or robustness

[6] See for example [40, 5]

issues, we have computed the fundamental matrices and the infinity homography matrices using calibration or self-calibration methods, in order to obtain very coherent values for these matrices. Robust methods to obtain the fundamental matrix from point correspondences exist already[16, 31], whereas practical methods to estimate the infinity homography are currently being developed [38].

5.1 Computation of the canonic representation for three views and reconstruction

Three images of the calibration grid are used. We start from the three fundamental matrices $\mathbf{F}_{12}$, $\mathbf{F}_{32}$, $\mathbf{F}_{13}$. The normalized epipoles $\mathbf{e}_{Nij}$ are obtained by solving $\mathbf{F}_{ij}\mathbf{e}_{Nij} = 0$, and this yield the epipolar projection matrices $\mathbf{S}_{ij} = [\mathbf{e}_{Nij}]_\times \mathbf{F}_{ij}$. Applying the formulas (42), we obtain β_1 and $\mathbf{q}_{N1}$, and then compute the invariant descriptions $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_3$ from table 2. General projective transformations can yield very strange deformations[7] which may make the shape of an object totally impossible to recognize, thus we first apply a carefully chosen projective transformation of $\mathcal{P}^3$ to the three matrices above, in order to obtain three projection matrices $\tilde{\mathbf{P}}_i$. The epipolar geometry computed from the matrices $\tilde{\mathbf{P}}_i$ is shown figure 2. It can be seen that it is perfectly coherent. Thus, the classical *trinocular* algorithms can be used for matching and reconstruction. This is a progress over previous methods [4, 11] where only two cameras could be considered at the same time, since it is well known that trinocular methods are more precise, robust, and efficient. A 3D reconstruction is shown figure 3. It can be verified on the stereogram that collinarity and coplanarity are preserved. The two last views, taken under orthographic projections along the normal of each of the two planes, show clearly that parallelism is not preserved.

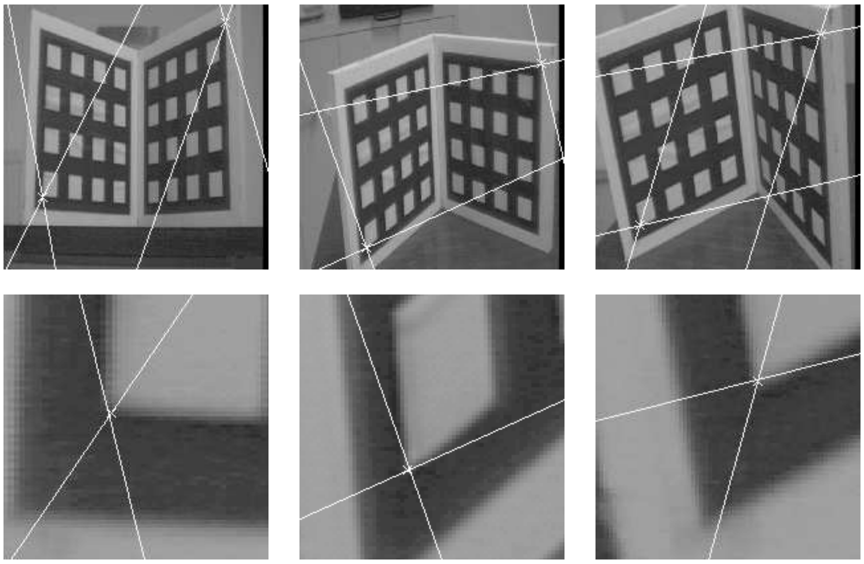

Figure 2: Epipolar lines from the projection matrices obtained using only fundamental matrices. Control points are at the two opposite corners of the grid.

If, in addition to the fundamental matrices, the knowledge of *one* infinity homography matrix is added, then the *three* infinity homography matrices can be computed using (45). Applying (41), the value of β_1 is obtained, and then affine invariant descriptions are given by table 2. Unlike the projective reconstruction, the affine reconstruction limits deformations, since parallelism has to be preserved, and thus we can use directly the invariant description $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$

[7] For instance, the inversion can map finite points to points at infinity.

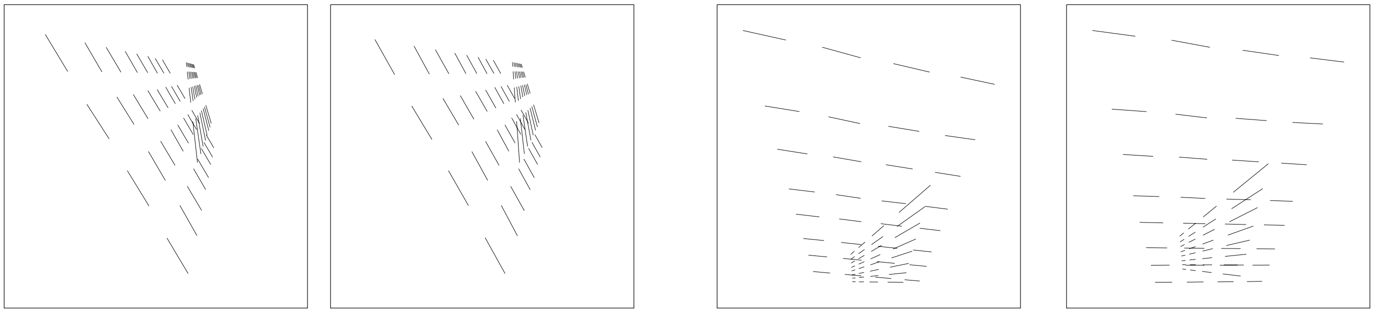

Figure 3: Projective 3D reconstruction. Left: stereogram. Right: two orthographic projections

as projection matrices. Although the reconstructed points in the stereogram shown in figure 4 might seem all coplanar, they are not. It is just that the angle between the two planes in this reconstruction is very small. Applying a chosen affine transformation of $\mathcal{P}^3$ to the three matrices above makes the difference between the planes appear, as seen in the two last reprojected views. These views, taken under orthographic projection along the normal of each of the two planes, illustrate clearly that parallelism is preserved, but that distances and angles are not. This is to be compared with an Euclidean reconstruction, obtained from three views, given figure 5. Although some information is lost, the affine reconstruction remains attractive.

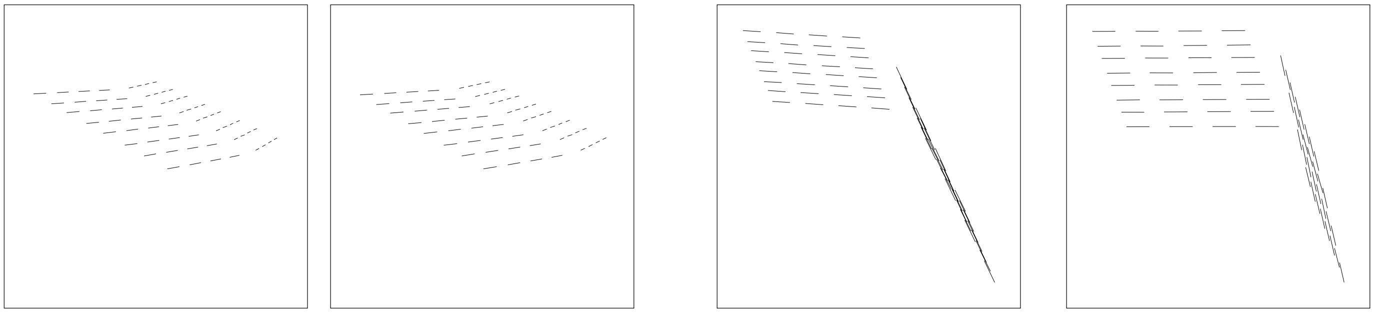

Figure 4: Affine 3D reconstruction. Left: stereogram. Right: two orthographic projections

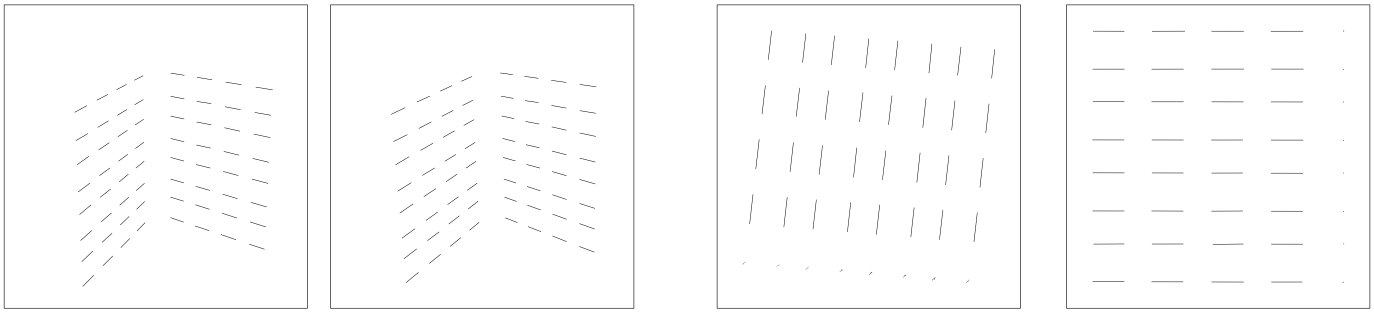

Figure 5: Euclidean 3D reconstruction. Left: stereogram. Right: two orthographic projections

5.2 Using affine information to perform self-calibration

Three images of an indoor scene taken by a zooming camera are used. The focal has been changed from 8mm to 12mm (read on the lens barrel). The infinity homography matrix between the two first views is:

$$\mathbf{H}_{\infty 12} = \begin{bmatrix} 1.72134 & -.172135 & -218.805 \\ .758041 & 1.59733 & -620.388 \\ .00109138 & .000592603 & 1.00000 \end{bmatrix}$$

The three eigenvalues of $\mathbf{H}_{\infty 12}$ have the same norm, 1.57575, thus we conclude that the intrinsic parameters remain constant between image 1 and image 2. The system of equations (50), with $\mathbf{K} = \mathbf{K}'$, is linear. Let us use, in conformity with previous work [6], the notation:

$$\mathbf{K} = \begin{bmatrix} -\delta_{23} & \delta_3 & \delta_2 \\ \delta_3 & -\delta_{13} & \delta_1 \\ \delta_2 & \delta_1 & -\delta_{12} \end{bmatrix}$$

Normalizing by $\delta_{12} = -1$, we obtain a 5×5 matrix for the system. The singular values of this matrix are: 957, 462, .19, .00055, .82 10^{-11}. The rank of this matrix is thus 4, as expected, and the solutions found are:

$$\begin{aligned} \delta_1 &= 776.394 - .00115042t \\ \delta_2 &= -118.969 + .000817533t \\ \delta_3 &= 400000 - .747472t \end{aligned} \qquad \begin{aligned} \delta_{13} &= -731756 + .353407t \\ \delta_{23} &= -40381.4 - .562483t \end{aligned}$$

A further assumption is needed to obtain a unique solution. Let us suppose that the two axes of the retinal coordinate system are orthogonal, which means that $\gamma = 0$ in equation (4), then we have the additional constraint [6]: $-\delta_3\delta_{12} = \delta_1\delta_2$, which yields the two solutions: $t = 448873$, $t = .116628\ 10^7$. The first solution leads to:

$$\delta_1 = 260,\ \ \delta_2 = 248,\ \ \delta_3 = 64480,\ \ \delta_{13} = -573121, \delta_{23} = -292865$$

and then to the intrinsic parameters (formulas are in [6]):

$$\alpha_u = 481,\ \ \alpha_v = 711,\ \ u_0 = 248,\ \ v_0 = 260$$

The second solution is to be discarded, because it leads to:

$$\delta_1 = -565.319,\ \ \delta_2 = 834.501,\ \ \delta_3 = -471759,\ \ \delta_{13} = -319585,\ \ \delta_{23} = -696393$$

Thus we see that $\delta_2^2 = \delta_{23}\delta_{12}$ and $\delta_1^2 = \delta_{13}\delta_{12}$, which corresponds to a degenerate case where $\mathbf{K}$ represents a line, and $\alpha_u = \alpha_v = 0$.

Now, having the three fundamental matrices between the images, we are able to obtain, using (45), the other infinity homography matrices from $\mathbf{H}_{\infty 12}$. In particular, we have:

$$\mathbf{H}_{\infty 23} = \begin{bmatrix} 1.07992 & .302292 & -35.2662 \\ -.905154 & .959146 & 413.566 \\ -.0000364269 & -.000355526 & 1.00000 \end{bmatrix}$$

The norms of the eigenvalues are $1.21, 1.21, 1.0$, thus we conclude that the intrinsic parameters have changed between image 2 and image 3. Applying formula (50) gives immediately the new matrix $\mathbf{K}'$, from which the following parameters are computed:

$$\alpha_u = 642,\ \ \alpha_v = 950,\ \ u_0 = 248,\ \ v_0 = 263$$

It can be verified that they correspond to the variation of focal length described previously. From the intrinsic parameters and the fundamental matrices, we can obtain the essential matrices, and thus the motions, up to a scale factor. Applying (40) as done in [19] it is thus possible to recover three projection matrices which are equivalent to the initial ones, up to a global scale factor and a change of coordinate system. Some epipolar lines obtained from these new projection matrices are show figure 6. Note that if we have a fourth view, the same method will allow us to obtain a fourth projection matrix with Euclidean information, and so on.

This example has illustrated how knowing only one infinity homography matrix, enables to perform (partial) self-calibration from two views in the case of constant intrinsic parameters, but also to deal with the case of variable intrinsic parameters, by propagating affine, and then Euclidean information.

Figure 6: Three images taken by a zooming camera, with epipolar lines

6 Conclusion

This paper lays the ground for further studies about problems involving 3D information, multiple viewpoints, and uncalibrated cameras. It confirms the interest of the affine representation, which turns out to yield simple and powerful descriptions.

We have described the *canonic decomposition*, an idea to account in a single framework for the different geometric levels of representation, in the case of two views, three views, or more. The approach is very general, since it involves only reasoning about the projection matrices. We first presented new descriptions for the affine and projective geometries of two views, which are respectively the *infinity homography matrix* and the *epipolar projection matrix*, which have been described from both an algebraic and geometric viewpoint. Then, a coherent hierarchy of representations has been studied. In particular, we have exhibited minimal and complete representations for each level of description, and showed clearly which elements of representation change and which ones are conserved across two different levels. These representations are description of the geometry of the cameras which are invariant with respect to a given group of transformations. In the case of three views, new representations and their associated composition formulas have been established. They allow to deal with the case of multiple viewpoints while working with uncalibrated cameras. We have also investigated the relationships which occur between the different levels of representation, and some computational consequences, which have began to be explored [37, 38] in order to recover efficiently the invariant descriptions studied in this paper from various primitives extracted in uncalibrated images.

Acknowledgements

We would like to thank L. Robert for proofreading the paper and interesting remarks.

A Algebraic derivation of the affine invariants

In this section, we illustrate the principle of the derivation of invariants thanks to the simplest case, the affine one. One choice has to be made, the form of the invariant for the first camera, which we make to be the simplest possible. Then the case of two views gives the nature of the invariant as function of projection matrices, and the case of three views gives also in addition the composition relations. All is done just by applying the definition of the canonic decomposition.

Two views In the affine case, looking for a matrices $\mathbf{X}$, $\mathbf{Y}$, vectors $\mathbf{x}$,$\mathbf{y}$ and a scalar μ such that:

$$\left\{ \begin{array}{l} \tilde{\mathbf{P}} = [\mathbf{I}_3, 0]\, \mathcal{A} \\ \tilde{\mathbf{P}}' = [\mathbf{X}, \mathbf{x}]\, \mathcal{A} \end{array} \right. \quad \text{with} \quad \mathcal{A} = \left[\begin{array}{cc} \mathbf{Y} & \mathbf{y} \\ \mathbf{0}_3^T & \mu \end{array} \right] \tag{54}$$

gives the equations:

$$\mathbf{P} = \mathbf{Y} \ , \ \mathbf{p} = \mathbf{y} \ , \ \mathbf{P}' = \mathbf{X}\mathbf{Y} \ , \ \mathbf{p}' = \mathbf{X}\mathbf{y} + \mu\mathbf{x}$$

which immediately yield the representation $\mathbf{X} = \mathbf{P}'\mathbf{P}^{-1}$, $\mu\mathbf{x} = \mathbf{P}'\mathbf{P}^{-1}\mathbf{p} - \mathbf{p}'$. One arbitrary choice for $||\mathbf{x}||$ has thus to be made, and we take the value 1, thus $\mu = ||\mathbf{x}||$.

Three views Starting from the canonic representation found for two views, we look for matrices $\mathbf{X}$,$\mathbf{Y}'$,$\mathbf{Y}''$, vectors $\mathbf{x}$,$\mathbf{y}'$,$\mathbf{y}''$, and scalars μ', μ'', such that:

$$\left\{\begin{array}{l} \tilde{\mathbf{P}}_1 = [\mathbf{I}_3, \mathbf{0}]\ \mathcal{A} \\ \tilde{\mathbf{P}}_2 = [\mathbf{H}_{12}, \mathbf{e}_{N21}]\ \mathcal{A} \\ \tilde{\mathbf{P}}_3 = [\mathbf{X}, \mathbf{x}]\ \mathcal{A} \end{array}\right. \qquad \text{with} \qquad \mathcal{A} = \left[\begin{array}{cc} \mathbf{Y} & \mathbf{y} \\ \mathbf{0}_3^T & \mu \end{array}\right] \tag{55}$$

$$\left\{\begin{array}{l} \tilde{\mathbf{P}}_2 = [\mathbf{I}_3, \mathbf{0}]\ \mathcal{A}' \\ \tilde{\mathbf{P}}_3 = [\mathbf{H}_{23}, \mathbf{e}_{N32}]\ \mathcal{A}' \end{array}\right. \qquad \text{with} \qquad \mathcal{A}' = \left[\begin{array}{cc} \mathbf{Y}' & \mathbf{y}' \\ \mathbf{0}_3^T & \mu' \end{array}\right] \tag{56}$$

$$\left\{\begin{array}{l} \tilde{\mathbf{P}}_1 = [\mathbf{I}_3, \mathbf{0}]\ \mathcal{A}'' \\ \tilde{\mathbf{P}}_3 = [\mathbf{H}_{13}, \mathbf{e}_{N31}]\ \mathcal{A}'' \end{array}\right. \qquad \text{with} \qquad \mathcal{A}'' = \left[\begin{array}{cc} \mathbf{Y}'' & \mathbf{y}'' \\ \mathbf{0}_3^T & \mu'' \end{array}\right] \tag{57}$$

By equating expressions in (55) and in (56), we obtain:

$$\begin{array}{c} \mathbf{Y}' = \mathbf{H}_{12}\mathbf{Y}\ ,\ \mathbf{y}' = \mathbf{H}_{12}\mathbf{y} + \mu\mathbf{e}_{N21}\ ,\ \mu' = ||\mathbf{e}_{32}|| \\ \mathbf{X} = \mathbf{H}_{23}\mathbf{H}_{12}\ ,\ \mathbf{x} = \mathbf{H}_{23}\mathbf{e}_{N21} + (\mu'/\mu)\mathbf{e}_{N32} \end{array} \tag{58}$$

where $\mu = ||\mathbf{e}_{21}||$, cannot be eliminated from the equations, and by equating expressions in (55) and in (57), we obtain:

$$\mathbf{Y}'' = \mathbf{Y}\ ,\ \mathbf{y}'' = \mathbf{y}\ ,\ \mu' = ||\mathbf{e}_{31}||\ ,\ \mathbf{X} = \mathbf{H}_{13}\ ,\ \mathbf{x} = (\mu''/\mu)\mathbf{e}_{N31} \tag{59}$$

We thus obtain the two alternative representations. By equating the values of $\mathbf{X}$ and $\mathbf{x}$ found in (58) with those found in (59), the composition relations are obtained.

B Proofs of some formulas

B.1 The epipolar projection

We show that the matrix $\mathbf{S}$ describes the correspondence from image 1 to image 2 generated by the plane $\mathbf{\Pi}_{e'}$.

First let compute the intersection of the optical ray of a point $\mathbf{m}$ in the camera defined by the projection matrix $\tilde{\mathbf{P}} = [\mathbf{P}, \mathbf{p}]$ with a given plane $\mathbf{\Pi}$. Two points of this optical ray are the optical center $\mathbf{C}$ and the point of infinity representing the direction $\mathbf{P}^{-1}\mathbf{m}$ of this line, thus a point of this optical ray can be written:

$$\mathbf{M} = \mathbf{C} + \lambda \left[\begin{array}{c} \mathbf{P}^{-1}\mathbf{m} \\ 0 \end{array}\right] = \left[\begin{array}{c} \mathbf{P}^{-1}(\mathbf{p} + \lambda\mathbf{m}) \\ -1 \end{array}\right]$$

By writing that $\mathbf{\Pi}^T\mathbf{M} = 0$, a value for λ is found as a function of $\tilde{\mathbf{P}}$, $\mathbf{\Pi}$, and $\mathbf{m}$, and by back-substitution in the previous relation, we obtain, after some algebra (the double cross-product formula is used):

$$\mathbf{M}_\Pi = \left[\begin{array}{c} \mathbf{P}^{-1}([\mathbf{P}^{-1T}\mathbf{d}]_\times[\mathbf{p}]_\times + \delta\mathbf{I}_3)\mathbf{m} \\ -\mathbf{d}^T\mathbf{P}^{-1}\mathbf{m} \end{array}\right] \qquad \text{where} \quad \mathbf{\Pi} = \left[\begin{array}{c} \mathbf{d} \\ \delta \end{array}\right]$$

Now let define $\mathbf{\Pi}_{e'}$ by $\mathbf{d} = \mathbf{P}^{'T}\mathbf{e}'$ and $\delta = \mathbf{p}^{'T}\mathbf{e}'$. The correspondence to $\mathbf{m}$ through plane $\mathbf{\Pi}_{e'}$ is $\mathbf{m}' = \tilde{\mathbf{P}}'\mathbf{M}_\Pi$, which can be written as:

$$\mathbf{m}' = \mathbf{H}_\infty([\mathbf{H}_\infty^T\mathbf{e}']_\times[\mathbf{p}]_\times + \mathbf{p}^{'T}\mathbf{e}')\mathbf{m} - \mathbf{p}'(\mathbf{e}^{'T}\mathbf{H}_\infty\mathbf{m})$$

Expending the double cross-product and using $\mathbf{H}_\infty\mathbf{p} = \mathbf{p}' - \mathbf{e}'$ yield:

$$\mathbf{H}_\infty[\mathbf{H}_\infty^T\mathbf{e}']_\times[\mathbf{p}]_\times\mathbf{m} = (\mathbf{e}^{'T}\mathbf{H}_\infty\mathbf{m})(\mathbf{p}' - \mathbf{e}') - (\mathbf{e}^{'T}(\mathbf{p}' - \mathbf{e}'))\mathbf{H}_\infty\mathbf{m}$$

Several terms cancels and it follows:

$$\mathbf{m}' = -(\mathbf{e}^{'T}\mathbf{H}_\infty\mathbf{m})\mathbf{e}' + (\mathbf{e}^{'T}\mathbf{e}')\mathbf{H}_\infty\mathbf{m} = -[\mathbf{e}']_\times[\mathbf{e}']_\times\mathbf{H}_\infty\mathbf{m}$$

B.2 Composition of projective representations

The third epipole Starting from the definition of $\mathbf{S}_{23}$, then using (35) and expending the double crossproduct yield:

$$\mathbf{S}_{23}\mathbf{e}_{21} = -\frac{[\mathbf{e}_{32}]_\times}{||\mathbf{e}_{32}||^2}\mathbf{F}_{23}\mathbf{e}_{21} = -\frac{\mathbf{e}_{32}}{||\mathbf{e}_{32}||^2} \times (\mathbf{e}_{32} \times \mathbf{e}_{31}) = \mathbf{e}_{31} - \frac{(\mathbf{e}_{32}^T\mathbf{e}_{31})}{||\mathbf{e}_{32}||^2}\mathbf{e}_{32} \tag{60}$$

Starting from the definition of the vectors $\mathbf{r}_\infty$ and using (29) and (23) yield:

$$(\mathbf{r}_{\infty 23} - \mathbf{r}_{\infty 21})^T \mathbf{e}_{21} = -\frac{\mathbf{e}_{12}^T}{||\mathbf{e}_{12}||^2}\mathbf{H}_{\infty 21}\mathbf{e}_{21} + \frac{\mathbf{e}_{32}^T}{||\mathbf{e}_{32}||^2}\mathbf{H}_{\infty 23}\mathbf{e}_{21} = 1 + \frac{\mathbf{e}_{32}^T}{||\mathbf{e}_{32}||^2}(\mathbf{e}_{31} - \mathbf{e}_{32}) = \frac{(\mathbf{e}_{32}^T\mathbf{e}_{31})}{||\mathbf{e}_{32}||^2}$$

Thus:

$$\mathbf{S}_{23}\mathbf{e}_{21} + ((\mathbf{r}_{\infty 23} - \mathbf{r}_{\infty 21})^T\mathbf{e}_{21})\mathbf{e}_{32} = \mathbf{e}_{31}$$

The third epipolar projection matrix The substitution of the decompositions (17) into (28) yields:

$$\mathbf{S}_{13} - \mathbf{S}_{23}\mathbf{S}_{12} = \mathbf{S}_{23}\mathbf{e}_{21}\mathbf{r}_{\infty 12}^T + \mathbf{e}_{32}\mathbf{r}_{\infty 23}^T\mathbf{S}_{12} + \mathbf{e}_{32}\mathbf{r}_{\infty 23}^T\mathbf{e}_{21}\mathbf{r}_{\infty 12}^T - \mathbf{e}_{31}\mathbf{r}_{\infty 13}^T$$

The first term of the right side is transformed using (60). The fourth term is transformed using $\mathbf{r}_{\infty 23}^T\mathbf{e}_{21} = \mathbf{r}_{\infty 23}^T(\mathbf{e}_{31} - \mathbf{e}_{32}) = \frac{\mathbf{e}_{32}^T\mathbf{e}_{31}}{||\mathbf{e}_{32}||^2} - 1$ This yields after simplifications:

$$\mathbf{S}_{13} - \mathbf{S}_{23}\mathbf{S}_{12} = \mathbf{e}_{32}\mathbf{r}_{\infty 23}^T\mathbf{S}_{12} + (\mathbf{e}_{31} - \mathbf{e}_{32})\mathbf{r}_{\infty 12}^T - \mathbf{e}_{31}\mathbf{r}_{\infty 13}^T$$

By expressing $\mathbf{S}_{12}$ and $\mathbf{r}_{\infty 21}$ as functions of infinite homographies and epipoles and using (13) and then (23), we obtain: $\mathbf{r}_{\infty 12} = \mathbf{S}_{12}^T\mathbf{r}_{\infty 21} - \frac{\mathbf{e}_{12}^T}{||\mathbf{e}_{12}||^2}$. The substitution of this value yields:

$$\mathbf{S}_{13} - \mathbf{S}_{23}\mathbf{S}_{12} = \mathbf{e}_{32}((\mathbf{r}_{\infty 23}^T - \mathbf{r}_{\infty 21}^T)\mathbf{S}_{12} + \frac{\mathbf{e}_{12}^T}{||\mathbf{e}_{12}||^2}) + \mathbf{e}_{31}(\underbrace{\mathbf{S}_{12}^T\mathbf{r}_{\infty 21} - \frac{\mathbf{e}_{12}}{||\mathbf{e}_{12}||^2}}_{\mathbf{r}_{\infty 12}} - \mathbf{r}_{\infty 13})^T$$

C Numerical values

We provide numerical values corresponding to the examples of Section 5.

C.1 Example 1

$$\mathbf{F}_{12} = \begin{bmatrix} -.1395\ 10^{-5} & .1853\ 10^{-5} & .001981 \\ -.8986\ 10^{-5} & .1374\ 10^{-5} & .01201 \\ .001424 & -.01269 & 1 \end{bmatrix},\ \mathbf{F}_{32} = \begin{bmatrix} -.8549\ 10^{-5} & .7565\ 10^{-5} & .01596 \\ -.1561\ 10^{-5} & .2783\ 10^{-5} & .004520 \\ -.01474 & -.01013 & 1 \end{bmatrix}$$

$$\mathbf{F}_{13} = \begin{bmatrix} .5978\ 10^{-5} & .4767\ 10^{-5} & -.02712 \\ .4380\ 10^{-5} & .2787\ 10^{-5} & .005137 \\ .02105 & -.007449 & 1 \end{bmatrix}$$

$$\beta_1 = 2.5651\ ,\ \mathbf{q}_{N1} = [-.00759, .008898, -.03468]^T$$

$$\mathcal{P}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},\ \mathcal{P}_2 = \begin{bmatrix} 6911.48 & -1363.61 & 2373.28 & 7873.86 \\ -1200.83 & 343.822 & -8656.55 & -1381.41 \\ .803724 & -.161854 & 99.3223 & 1.000000 \end{bmatrix}$$

$$\mathcal{P}_3 = \begin{bmatrix} -635.855 & 126.875 & -1835.90 & -694.050 \\ -3351.10 & 640.177 & -1440.08 & -3859.31 \\ .877744 & -.158361 & -42.4041 & 1.00000 \end{bmatrix}$$

$$\mathbf{H}_{\infty 12} = \begin{bmatrix} .7091 & -.004109 & 397.0 \\ -.01186 & 1.004 & -148.7 \\ -.0006201 & .0001081 & 1.000 \end{bmatrix}, \ \mathbf{H}_{\infty 32} = \begin{bmatrix} 1.034 & .3276 & 163.0 \\ -.2321 & 1.177 & -796.9 \\ -.0005753 & .0005679 & 1.000 \end{bmatrix}$$

$$\mathbf{H}_{\infty 13} = \begin{bmatrix} .8853 & -.1793 & 140.9 \\ -.007023 & .7471 & 549.2 \\ -.0002499 & -.0003943 & 1.0000 \end{bmatrix}$$

$$\mathcal{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \ \mathcal{A}_2 = \begin{bmatrix} -5668.34 & 32.85 & -.31736\ 10^7 & 7873.86 \\ 94.778 & -8023.47 & .11885\ 10^7 & -1381.41 \\ 4.9569 & -.864224 & -7994.12 & 1.000 \end{bmatrix}$$

$$\mathcal{A}_3 = \begin{bmatrix} 541.50 & -514.85 & 404452. & -694.05 \\ -20.1608 & 2144.7 & .15767\ 10^7 & -3859.3 \\ -.717456 & -1.1320 & 2870.78 & 1.000 \end{bmatrix}$$

C.2 Example 2

$$\mathbf{F}_{12} = \begin{bmatrix} .00030194 & .00022904 & .20447 \\ -.00018455 & -.000039449 & -.070503 \\ -.34804 & .17286 & 1.000 \end{bmatrix} \ \mathbf{F}_{23} = \begin{bmatrix} .000049813 & -.000052961 & -.022319 \\ .000059537 & .000017487 & -.0043212 \\ -.0026382 & .0021470 & 1.00 \end{bmatrix}$$

$$\mathbf{F}_{13} = \begin{bmatrix} .43369\ 10^{-5} & -.000014001 & .0024273 \\ .000012721 & .19193\ 10^{-5} & -.0068886 \\ -.0053471 & .0050099 & 1.00 \end{bmatrix}$$

References

[1] *Proc. British Machine Vision Conf.*, Guildford, UK, Sept 1993.

[2] B. Caprile and V. Torre. Using Vanishing Points for Camera Calibration. *International Journal of Computer Vision*, 4:127–140, 1990.

[3] O. D. Faugeras and F. Lustman. Let us suppose that the world is piecewise planar. In O. D. Faugeras and G. Giralt, editors, *Robotics Research, The Third International Symposium*, pages 33–40. MIT Press, 1986.

[4] O.D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig. In *Proc. European Conference on Computer Vision*, pages 563–578, 1992.

[5] O.D. Faugeras. *Three-dimensional computer vision: a geometric viewpoint.* MIT Press, 1993. To appear.

[6] O.D. Faugeras, Q.-T. Luong, and S.J. Maybank. Camera self-calibration: theory and experiments. In *Proc. European Conference on Computer Vision*, pages 321–334, Santa-Margerita, Italy, 1992.

[7] O.D. Faugeras and S.J. Maybank. Motion from point matches: multiplicity of solutions. *The International Journal of Computer Vision*, 4(3):225–246, 1990. also INRIA Tech. Report 1157.

[8] O.D. Faugeras and L. Robert. What can two images tell us about a third one ? Technical report, INRIA, July 1993.

[9] O.D. Faugeras and G. Toscani. The calibration problem for stereo. In *Proceedings of CVPR'86*, pages 15–20, 1986.

[10] R. Hartley and R. Gupta. Computing matched-epipolar projections. In *Proc. International Conference on Computer Vision and Pattern Recognition*, pages 549–555, New-York, 1993.

[11] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. of the conf. on Computer Vision and Pattern Recognition*, pages 761–764, Urbana, 1992.

[12] T.S. Huang and O.D. Faugeras. Some properties of the E-matrix in two view motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:1310–1312, 1989.

[13] J. J. Koenderink and A. J. van Doorn. Affine Structure from Motion. *Journal of the Optical Society of America A*, 8(2):377–385, 1992.

[14] H.C. Longuet-Higgins. A Computer Algorithm for Reconstructing a Scene from Two Projections. *Nature*, 293:133–135, 1981.

[15] Q.-T. Luong. *Matrice fondamentale et calibration visuelle sur l'environnement: vers une plus grande autonomie des systèmes robotiques.* PhD thesis, Universite de Paris-Sud., Dec. 1992.

[16] Q.-T. Luong, R. Deriche, O.D. Faugeras, and T. Papadopoulo. On determining the Fundamental matrix: analysis of different methods and experimental results. Technical Report RR-1894, INRIA, 1993.

[17] Q.-T. Luong and O.D. Faugeras. Calibration, motion, and structure, from fundamental matrices and point correspondences. In preparation, 1993.

[18] Q.-T. Luong and O.D. Faugeras. Determining the Fundamental matrix with planes: unstability and new algorithms. In *Proc. Conference on Computer Vision and Pattern Recognition*, pages 489–494, New-York, 1993.

[19] Q.-T. Luong and O.D. Faugeras. Self-calibration of a stereo rig from unknown camera motions and point correspondences. In A. Grun and T.S. Huang, editors, *Calibration and orientation of cameras in computer vision.* Springer-Verlag, 1993. To appear. Also presented at XVII ISPRS, Washington, and INRIA Tech Report.

[20] M.J. Magee and J.K. Aggarwal. Determining vanishing points in perspective images. *Computer Vision, Graphics and Image Processing*, 26:256–267, 1984.

[21] S.J. Maybank. Properties of essential matrices. *International journal of imaging systems and technology*, 2:380–384, 1990.

[22] S.J. Maybank and O.D. Faugeras. A Theory of Self-Calibration of a Moving Camera. *The International Journal of Computer Vision*, 8(2):123–151, 1992.

[23] J. L. Mundy and A. Zisserman, editors. *Geometric invariance in computer vision.* MIT Press, 1992.

[24] L. Quan. Affine stereo calibration for relative affine shape reconstruction. In *Proc. British Machine Vision Conf.*, 1993.

[25] L. Quan and R. Mohr. Determining perspective structures using hierarchical Hough transform. *Pattern Recognition Letters*, 9(4):279–286, 1989.

[26] L. Quan and R. Mohr. Affine shape representation from motion through reference points. *Journal of mathematical imaging and vision*, 1:145–151, 1992.

[27] L. Robert and O.D. Faugeras. Relative 3D Positioning and 3D Convex Hull Computation from a Weakly Calibrated Stereo Pair. In *Proc. International Conference on Computer Vision*, pages 540–543, Berlin, Germany, May 1993.

[28] L.S. Shapiro, A. Zisserman, and M. Brady. Motion from point matches using affine epipolar geometry. Technical Report OUEL 1994/93, Oxford University., June 1993.

[29] A. Shashua. On geometric and algebraic aspects of 3d affine and projective structures from perspective 2d views. Technical Report 1405, MIT AI Memo, July 1993.

[30] A. Shashua. Projective depth: a geometric invariant for 3d reconstruction from two perspective/orthographic views and for visual recognition. In *Proc. International Conference on Computer Vision*, pages 583–590, Berlin, Germany, 1993.

[31] P.H.S. Torr and D.W. Murray. Outlier detection and motion segmentation. Technical Report OUEL 1987/93, Oxford University, September 1993.

[32] G. Toscani. *Systèmes de Calibration et Perception du Mouvement en Vision Articielle.* PhD thesis, Université de Paris-Sud Centre d'Orsay, December 1987.

[33] R.Y. Tsai. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. In *Proceedings CVPR '86, Miami Beach, Florida*, pages 364–374. IEEE, June 1986.

[34] R.Y. Tsai. Synopsis of Recent Progress on Camera Calibration for 3D Machine Vision. In Oussama Khatib, John J. Craig, and Tomás Lozano-Pérez, editors, *The Robotics Review*, pages 147–159. MIT Press, 1989.

[35] R.Y. Tsai and T.S. Huang. Estimating Three-dimensional motion parameters of a rigid planar patch, II: singular value decomposition. *IEEE Transactions on Acoustic, Speech and Signal Processing*, 30, 1982.

[36] T. Viéville, P. Facao, and E. Clergue. Computation of ego-motion and structure from visual and inertial sensors using the vertical cue. In *Proc. International Conference on Computer Vision*, pages 591–598, Berlin, Germany, 1993.

[37] T. Viéville, Q.T. Luong, and O.D. Faugeras. Motion of points and lines in the uncalibrated case. Technical report, INRIA, Sept 1993. In press.

[38] T. Viéville, C. Zeller, and O.D. Faugeras. Using collineations to perform active vision. In preparation, 1993.

[39] L. Wang and W. Tsai. Computing cameras parameters using vanishing line information from a rectangular parallepiped. *Machine Vision and Applications*, 3:129–141, 1990.

[40] Z. Zhang and O.D. Faugeras. *3D dynamic scene analysis.* Springer-Verlag, 1992.