What can two images tell us about a third one — **Source link** ↗

Olivier Faugeras, Luc Robert

Institutions: French Institute for Research in Computer Science and Automation

Related papers:

- Three-dimensional computer vision: a geometric viewpoint

- A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry

- Algebraic functions for recognition

- Multiple view geometry in computer vision

- Trilinearity in visual recognition by alignment

What can two images tell us about a third one

Olivier Faugeras, Luc Robert

What can two images tell us about a third one?

Olivier FAUGERAS

Luc ROBERT

N° 2018

Juillet 1993

———— PROGRAMME 4 ————

Robotique,
image
et vision

*Rapport
de recherche*

1993

What can two images tell us about a third one?

Olivier FAUGERAS
Luc ROBERT

Abstract: This paper discusses the problem of predicting image features in an image from image features in two other images and the epipolar geometry between the three images. We adopt the most general camera model of perpective projection and show that *a point* can be predicted in the third image as a *bilinear* function of its images in the first two cameras, that *the tangents* to three corresponding curves are related by a *trilinear* function, and that *the curvature* of a curve in the third image is a *linear* function of the curvatures at the corresponding points in the other two images. Our analysis relies heavily on the use of the fundamental matrix which has been recently introduced [7] and on the properties of a special plane which we call the *trifocal* plane.

We thus answer completely the following question: given two views of an object, what would a third view look like? the question and its answer bear upon several areas of computer vision, stereo, motion analysis, and model-based object recognition. Our answer is quite general since it assumes the general perspective projection model for image formation and requires only the knowledge of the epipolar geometry for the triple of views. We show that in the special case of orthographic projection our results for points reduce to those of Ullman and Basri [24]. We demonstrate on synthetic as well as on real data the applicability of our theory.

(Résumé : tsvp)

Que nous apprennent deux images sur une troisième?

Résumé : Dans cet article, nous étudions la prédiction des éléments d'une image à partir des primitives de deux autres images et de la géométrie épipolaire entre les trois images. Nous adoptons le modèle de caméra le plus général : celui de la projection perspective. Nous démontrons que l'image d'un point dans la troisième caméra est une une fonction *bilinéaire* de ses images dans les deux premières caméras, que *les tangentes* à trois courbes en correspondance sont reliées par une fonction *trilinéaire*, et que *la courbure* d'une courbe dans la troisième image est une fonction *linéaire* des courbures aux points correspondants dans les deux autres images. Notre analyse utilise de manière extensive la matrice fondamentale qui a été introduite récemment [7], et les propriétés d'un plan particulier que nous appelons le plan *trifocal*.

Ainsi, nous répondons complètement à la question suivante : étant donné deux vues d'un objet, à quoi ressemblerait une troisième vue de celui-ci ? La question et sa réponse embrassent divers domaines de la vision par ordinateur : la stéréo, l'analyse du mouvement, et la reconnaissance d'objets à base de modèles. Notre solution est très générale puisqu'elle repose sur le modèle général de projection perspective, et suppose seulement que la géométrie épipolaire du triplet d'images est connue. Nous montrons que dans le cas particulier de la projection orthographique, nos résultats sur les points se réduisent à ceux de Ullman et Basri [24]. Nous démontrons sur des données synthétiques et réelles l'applicabilité de notre théorie.

Contents

1   Introduction

This paper is about the problem of predicting a third view of an object or a scene from two other views. Stated in these terms this is a fairly general problem which has been studied mostly from three different but complementary viewpoints.

First, people developing stereo systems have recognized quite early the advantages of using a third camera to reduce the ambiguity of the correspondence problem, either using active illumination [25] or using standard passive systems [9, 13, 17, 18, 1, 20, 11]. Most of them have used the idea of using the epipolar geometry for predicting the position and, for a few of them, the orientation of an image token in one image given a correspondence between image tokens in the other two images.

Second, people interested in the recognition and pose estimation of three-dimensional objects have recognized more recently that the variety of views of the same rigid object under rigid transformations can be expressed as the combinations of a small number of views [24]. The connection between the two problems has only been realized even more recently [3] and in a rather incomplete way even though the central role played by the epipolar geometry of the set of cameras has been acknowledged.

Third, people interested in photogrammetry and model-based vision as well, have developed so-called *transfer* methods which find for one or more image points in a given image set, the corresponding points in some new image set. If the camera geometries are known, transfer is done in a straightforward fashion by 3-D reconstruction and reprojection. If the camera geometries are unknown, this can still be done by methods based on using projective invariants [2].

What emerges from these three different viewpoints as some sort of unifying factor is the central role played by projective geometry in image formation. If we accept to think in terms of projective geometry while keeping in mind its close connection with the more familiar affine and euclidean geometries, then many of these questions find simple and elegant answers [6, 10]. This is the viewpoint which has recently allowed to arrive at a deep understanding of what is essential in the geometry of sets of cameras. It is not surprising that it can be described very simply in terms of projective geometry [14, 15, 5, 8, 7]. We will make heavy use of elementary projective geometry in this paper. The reader who is not familiar with the subject is referred to the nice appendix of [16] and then to the introductory textbook [21].

One particular important idea that is beginning to find its way in the computer vision and robotics communities is the idea that many useful vision tasks can be achieved without camera calibration for systems using only one camera or with a weak form of camera calibration consisting only of the epipolar geometry of the cameras, for systems involving several cameras [5, 8, 23, 22]. This paper takes this idea a step further and shows that it can solve elegantly most part of the above three problems in the following sense. Assume that we have a system of three cameras that is looking at objects containing lines (straight or not). These lines are supposed to be rigidly attached to the objects, in other words we do not consider here the problem of extremal boundaries. Given two images of one of these lines, can we predict its image in a third view, assuming only that we know the epipolar geometry between the views? we answer this question in the affirmative in a sense that is

general on three counts. First, as said before, we assume only that the epipolar geometry between the views is known, second, we assume that the cameras perform a full perspective projection of the 3-D world on their retinas and third, we develop prediction operators for points, tangents, and curvatures of the lines which go beyond what has previously been published.

The paper starts by giving some background on the geometry of camera systems in section 2. We then develop in section 3 the three predictors for points, straight lines or tangents and curvatures. Section 4 shows that in the orthographic case the predictor for points reduces, as expected, to the linear one used in [24]. In section 5, we validate our theory on synthetic and real data and conclude in section 6.

2 Background

In the article we use boldface letters for vectors and matrixes. When we will need to distinguish between a projective quantity and an affine or euclidean one, we will add a ˜ on top. For example, a pixel m with image coordinates u, v is represented by the vector $\mathbf{m} = [u, v]^T$ if we think of it as an affine or euclidean point and by the vector $\tilde{\mathbf{m}} = [u, v, 1]^T = [\mathbf{m}^T, 1]^T$ if we think of it as a projective point.

2.1 The geometry of a single camera

We use the general pinhole camera model in which the camera performs a full perspective projection from the three-dimensional space to the retinal plane $\mathcal{R}$ through the optical center C. In more detail, we can choose the coordinate system (C, x, y, z) for the three-dimensional space and (c, u, v) for the retinal plane as indicated in figure 1. The z-axis is perpendicular to the retinal plane, it is the optical axis of the camera. The coordinate system (C, x, y, z) is called the standard coordinate system of the camera. From this figure it should be clear that the relationship between image coordinates and 3D space coordinates can be written as

$$-\frac{f}{z} = \frac{u}{x} = \frac{v}{y} \tag{1}$$

which can be rewritten linearly as

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{2}$$

where we define

$$u = U/S \quad v = V/S \quad \text{if} \ \ S \neq 0 \tag{3}$$

We interpret U, V, S as the projective coordinates of a point in the retina. If $S = 0$, i.e., if $z = 0$, the 3D point is in the focal plane $\mathcal{F}$ of the camera. Thus the coordinates u and v are not defined, and the corresponding point is at infinity. The points such that $S = 0$ are
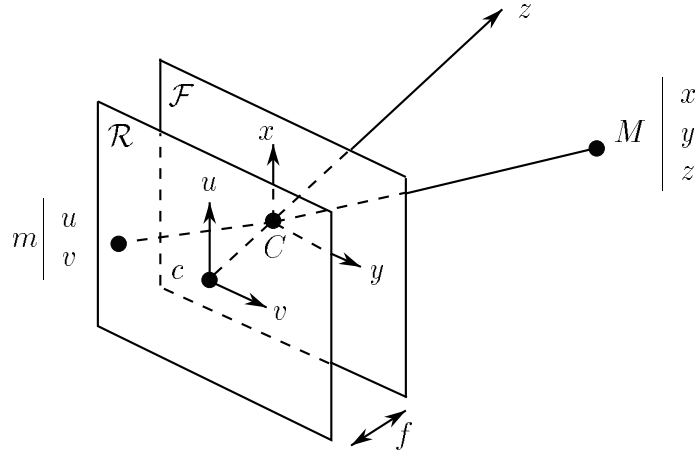
Figure 1: The focal plane (x, y) is parallel to the retinal plane (u, v) and at a distance f from it; f is the focal length.

called points at infinity of the retinal plane. $S = 0$ is the equation of the line at infinity of the retinal plane and this line is the "image" of the focal plane.

Note that equation (2) is projective, i.e., it is defined up to a scale factor, and we can rewrite it by using the projective coordinates (X, Y, Z, T) of M:

$$
\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}
\tag{4}
$$

The above formula expresses the fact that the relationship between image and space coordinates is linear in projective coordinates and can be written in matrix form as

$$
\tilde{\mathbf{m}} = \mathbf{P}\tilde{\mathbf{M}}
\tag{5}
$$

where $\tilde{\mathbf{m}} = [U, V, S]^T$ and $\tilde{\mathbf{M}} = [X, Y, Z, T]^T$. A camera can be considered as a system that performs a linear projective transformation from the projective space $\mathcal{P}^3$ into the projective plane $\mathcal{P}^2$. This is one of the many examples where the use of projective geometry makes things simpler: instead of dealing with the nonlinear equations (1), we can use the linear relation (5) and use the power of linear algebra.

It is easy to convince ourselves that relation (5) still holds true with different matrices $\mathbf{P}$ and for any choice of the 3-D and retinal plane coordinate systems. We sometimes refer to the 3-D coordinate system as to the world coordinate system. More precisely, if we perform a change of projective coordinates in the retinal plane defined by a 3×3 matrix $\mathbf{A}$ and a change of projective coordinates in the three-dimensional space defined by the 4×4 matrix $\mathbf{D}$, the new perspective projection matrix is given by

$$
\mathbf{P}_{\text{new}} = \mathbf{A}^{-1}\mathbf{P}\mathbf{D}
\tag{6}
$$

As an application of this, let us consider the translation of $-f$ along the z-axis of the standard coordinate system (figure 1). This translation brings the focal plane onto the retinal plane. In the standard coordinate system the matrix $\mathbf{P}$ is given by equation (2). Therefore the new perspective projection matrix, according to equation (6), is given by

$$\mathbf{P} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & -f \end{bmatrix}$$

Since $\mathbf{P}$ is defined up to a scale factor, this is the same as

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{f} & 1 \end{bmatrix}$$

If we now let f go to infinity, the matrix becomes

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

With the previous notations, this defines a transformation $U = x$, $V = y$, $S = 1$, and therefore $u = x$ and $v = y$. This is known as an orthographic projection parallel to the z-axis. It appears as the limit of the general perspective projection when the focal length f becomes large with respect to the distance z of the camera to the object.

2.2   The geometry of two cameras

In the previous section, we have discussed the geometry of one camera. We are now going to introduce a second camera and study the geometric properties of a set of two cameras. The main new geometric property is known in computer vision as the epipolar constraint and can readily be understood by looking at figure 2.

Let C (resp. C') be the optical center of the first camera (resp. the second). The line $\langle C, C' \rangle$ projects to a point e (resp. $e^{'}$) in the first retinal plane $\mathcal{R}$ (resp. in the second retinal plane $\mathcal{R}'$). The points e, e' are the epipoles. The lines through e in the first image and the lines through e' in the second image are the epipolar lines. The epipolar constraint is well-known in stereovision: for each point m in the first retina, its corresponding point m' lies on its epipolar line l'_m. If the relative camera geometry is known then, given a pixel m, its epipolar line l'_m can be computed, and its correspondent m' has only to be searched along l'_m rather than in the whole image.

Let us enrich this idea and consider the one parameter family of planes going through $\langle C, C' \rangle$ as shown in figure 3. This family is a pencil of planes. Its intersection with each retinal plane is a family of lines going through the points e and e'. These two families are pencils of lines called the pencils of epipolar lines. Moreover, as shown for example in [7], the epipolar planes induce a natural correspondence between the two pencils of epipolar lines: two epipolar lines correspond to each other if and only if they belong to the same epipolar
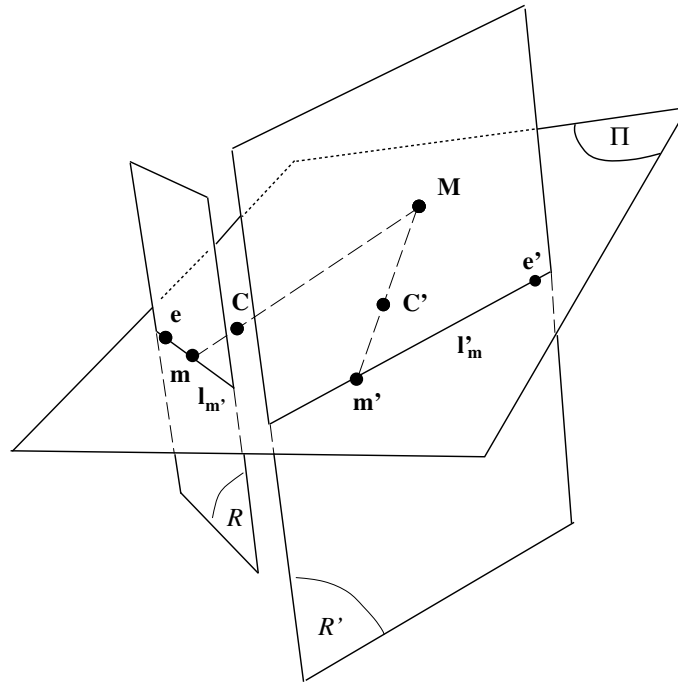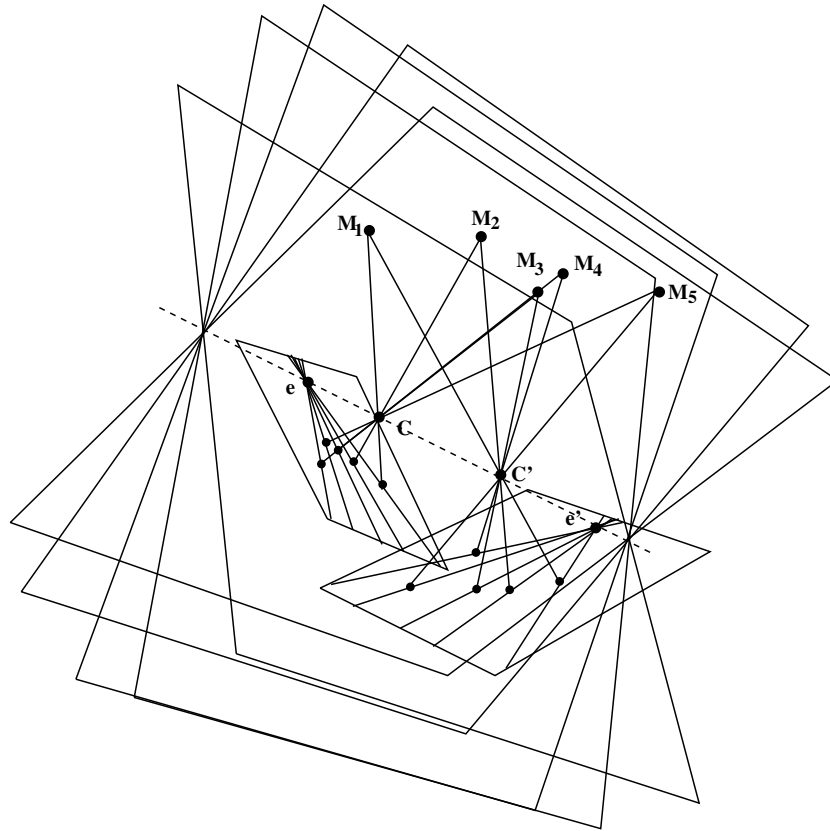
Figure 2: The epipolar geometry.

Figure 3: The epipolar pencils.

plane. This relation is clearly one to one and, in fact projective linear: it is a homography, called the epipolar homography, and can be represented with three parameters [7].

From these considerations it can be shown ([6] for example) that the correspondence between a point m and its epipolar line l'_m is quite simple, i.e. linear, in projective coordinates. It can therefore be written

$$\mathbf{l}_m = \mathbf{F}\tilde{\mathbf{m}} \tag{8}$$

The 3×3 matrix $\mathbf{F}$ is called the *fundamental matrix* [7, 12]. It is of rank 2. A vector $\mathbf{e}$ in its null space is a projective representation of the epipole e. The fact that m' belongs to l_m can be written very simply as follows:

$$\tilde{\mathbf{m}}'^T \mathbf{F} \tilde{\mathbf{m}} = 0$$

This relation shows that the epipolar line $l_{m'}$ of m' is represented by $\mathbf{F}^T \tilde{\mathbf{m}}'$. As a consequence the second epipole e' is represented by a vector of the null space of matrix $\mathbf{F}^T$. The fundamental matrix $\mathbf{F}$ contains all the information about the epipolar correspondence between the two cameras: the two epipoles, as mentioned earlier, and the three parameters of the epipolar homography. A stereo rig is said to be *weakly calibrated* if only the fundamental matrix is known. Since it depends only upon seven parameters, this is clearly less information than the knowledge of the two perspective projection matrixes $\mathbf{P}$ and $\mathbf{P}'$ which depend upon 22 parameters. Nonetheless, the information contained in the F-matrix is sufficient for recovering affine or projective information of the three-dimensional environment as shown in [5, 8].

2.3   The geometry of three cameras

If we now add a third camera, the geometry becomes even richer. We denote the cameras by $1, 2, 3$ and consider the three fundamental matrices $\mathbf{F}_{ij}$, $i = 1, 2, 3$, $j = i + 1$ with the obvious convention on the indexes. If m_i is a pixel in image i, its epipolar line in image j is represented by $\mathbf{F}_{ij}\tilde{\mathbf{m}}_i$. Note that we have $\mathbf{F}_{ij} = \mathbf{F}_{ji}^T$. As shown in figure 4, it is natural to consider the plane defined by the three optical centers C_1, C_2 and C_3 which we call the *trifocal* plane. It is a special plane which is simultaneously an epipolar plane for the three pairs of camera. It intersects each retinal plane along a line $d_i, i = 1, 2, 3$ which contains the epipoles e_{ii+1} and e_{ii+2} of the ith camera with respect to the $(i + 1)st$ and the $(i + 2)nd$. Note that, because of the epipolar geometry, we have

$$\mathbf{F}_{ii+1}\tilde{\mathbf{e}}_{ii+2} = \mathbf{d}_{i+1} = \tilde{\mathbf{e}}_{i+1,i} \wedge \tilde{\mathbf{e}}_{i+1,i+2} \tag{9}$$

This equation simply says that, for example, the epipolar line of e_{13} in image 2 is the line of intersection of the epipolar plane defined by $\langle C_1, C_2 \rangle$ and e_{13}. From figure 4, it is apparent that this plane is the trifocal plane and therefore that this line of intersection is d_2. This explains the first equality in (9). Since e_{21} and e_{23} are on d_2, this explains the second equality in (9).. To be consistent with the two cameras case, we will say that a system of three cameras is weakly calibrated if the three fundamental matrixes are known but not more. Note that because of (9), the three fundamental matrixes $\mathbf{F}_{12}, \mathbf{F}_{23}$ and $\mathbf{F}_{31}$ are not independent.
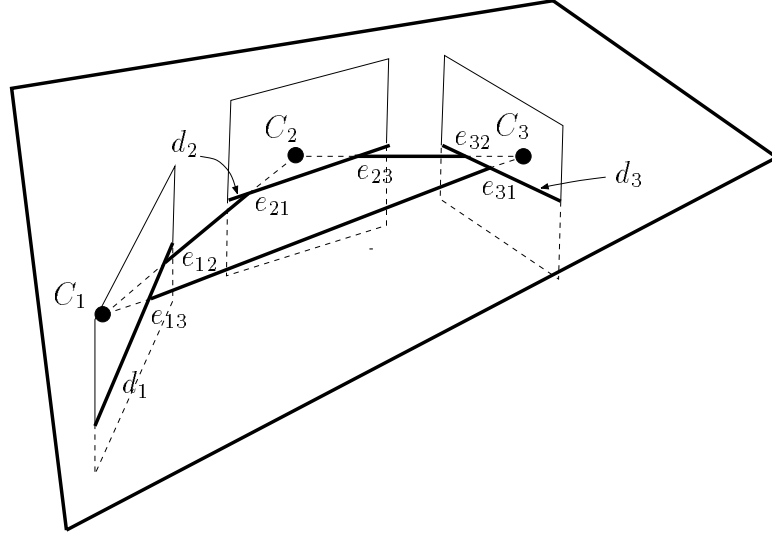
Figure 4: The trifocal plane is defined by the three optical centers C_1 C_2 and C_3.

3   Prediction

We now consider two views of a static scene obtained from two different viewpoints. The question we now ask is the following: can we predict how the scene would look like from a third viewpoint? Note that this is the same question as the following: given two views of an object, can we predict a third view of that same object? We will completely answer these questions under the general assumption that the three views are *weakly* calibrated.

3.1   Prediction of points

We assume that we are given two corresponding pixels m_1 and m_2 in images 1 and 2. The question is where should pixel m_3 be in image 3? The answer is quite simple, pixel m_3 must belong to the epipolar line of m_1 in the third image, represented by $\mathbf{F}_{13}\tilde{\mathbf{m}}_1$ and to the epipolar line of pixel m_2 in the third image, represented by $\mathbf{F}_{23}\tilde{\mathbf{m}}_2$ as well. Therefore, m_3 belongs to the intersection of these two lines and we can write:

$$\tilde{\mathbf{m}}_3 = \mathbf{F}_{13}\tilde{\mathbf{m}}_1 \wedge \mathbf{F}_{23}\tilde{\mathbf{m}}_2 \tag{10}$$

This equation predicts the position of the pixel in the third image given the positions of the corresponding pixels in the other two images. This formula is *bilinear* in m_1 and m_2. This means that the projective coordinates of m_3 are linear functions of the projective coordinates of m_1 and m_2 independently (in particular, they are *not* linear in the coordinates of the vector $[\tilde{\mathbf{m}}_1^T, \tilde{\mathbf{m}}_2^T]^T$). This is true if we consider projective coordinates. If we insist on using affine coordinates, then we have to divide by the third component of the right hand side of (10) (when it is not equal to zero). We see that the affine coordinates of m_3 are rational functions of the affine coordinates of m_1 and m_2 with numerators and denominators which are bilinear functions of these coordinates.

3.2   Prediction of lines

We assume now that we are given two corresponding lines l_1 and l_2 in images 1 and 2. The question is where should line l_3 be in image 3? The answer is again quite simple. Let m_1, m'_1 be two points of l_1. They define two points m_2, m'_2 of l_2 as the intersections of the epipolar line of m_1 represented by $\mathbf{F}_{12}\tilde{\mathbf{m}}_1$ and of m'_1 represented by $\mathbf{F}_{12}\tilde{\mathbf{m}}'_1$ with l_2. Therefore we can write:

$$\tilde{\mathbf{m}}_2 = \mathbf{F}_{12}\tilde{\mathbf{m}}_1 \wedge \mathbf{l}_2 \quad \tilde{\mathbf{m}}'_2 = \mathbf{F}_{12}\tilde{\mathbf{m}}'_1 \wedge \mathbf{l}_2 \tag{11}$$

Therefore, the line l_3 is defined by the two points m_3 and m'_3, intersections of the epipolar lines of m_1 and m'_1 and m_2 and m'_2 in the third image. Therefore we can write:

$$\mathbf{l}_3 = \tilde{\mathbf{m}}_3 \wedge \tilde{\mathbf{m}}'_3 = (\mathbf{F}_{13}\tilde{\mathbf{m}}_1 \wedge \mathbf{F}_{23}\tilde{\mathbf{m}}_2) \wedge (\mathbf{F}_{13}\tilde{\mathbf{m}}'_1 \wedge \mathbf{F}_{23}\tilde{\mathbf{m}}'_2)$$

Expanding the cross-products and using some standard properties of determinants, we obtain:

$$\mathbf{l}_3 = \mid \mathbf{F}_{13}\tilde{\mathbf{m}}_1, \mathbf{F}_{23}\tilde{\mathbf{m}}_2, \mathbf{F}_{23}\tilde{\mathbf{m}}'_2 \mid \mathbf{F}_{13}\tilde{\mathbf{m}}'_1 - \mid \mathbf{F}_{23}\tilde{\mathbf{m}}_2, \mathbf{F}_{13}\tilde{\mathbf{m}}'_1, \mathbf{F}_{13}\tilde{\mathbf{m}}_1 \mid \mathbf{F}_{23}\tilde{\mathbf{m}}'_2$$

We notice that $\mathbf{F}_{23}\tilde{\mathbf{m}}_2 \wedge \mathbf{F}_{23}\tilde{\mathbf{m}}'_2$ and $\mathbf{F}_{13}\tilde{\mathbf{m}}'_1 \wedge \mathbf{F}_{13}\tilde{\mathbf{m}}_1$ represent the points of intersection of the epipolar lines in the third image of points m_2, m'_2 and m_1, m'_1. These are known to intersect at the epipoles e_{32} and e_{31}, represented by the vectors in the nullspaces of matrices $\mathbf{F}_{32}$ and $\mathbf{F}_{31}$. This allows us to write:

$$\mathbf{l}_3 = (\tilde{\mathbf{m}}_1^T\mathbf{F}_{31}\tilde{\mathbf{e}}_{32})\mathbf{F}_{13}\tilde{\mathbf{m}}'_1 - (\tilde{\mathbf{m}}_2^T\mathbf{F}_{32}\tilde{\mathbf{e}}_{31})\mathbf{F}_{23}\tilde{\mathbf{m}}'_2$$

an equation where we can replace $\tilde{\mathbf{m}}_2$ and $\tilde{\mathbf{m}}'_2$ by their values obtained from equations (11).

This is the general case of lines. In the case where we know that the lines are the images of the tangent to a three-dimensional curve, there is a simpler way of performing prediction which will be useful for the next section in which we describe the prediction of curvature.

Let us assume that we are given the three image curves (c_i), $i = 1, 2, 3$ of a three-dimensional curve and three corresponding points m_i on these three curves. We can write three epipolar constraints

$$\tilde{\mathbf{m}}_j^T\mathbf{F}_{ij}\tilde{\mathbf{m}}_i = 0 \quad i = 1, 2, 3 \quad j = 2, 3, 1$$

In this equation it is understood that

$$\tilde{\mathbf{m}}_i = \left[\begin{array}{c} \mathbf{m}_i \\ 1 \end{array} \right]$$

Let us differentiate each of these three equations with respect to the arclengths of the curves (c_i) and (c_j), we obtain the following three relations:

$$\begin{array}{lll} \tilde{\mathbf{m}}_2^T\mathbf{F}_{12}\tilde{\mathbf{t}}_1 ds_1 & +\tilde{\mathbf{t}}_2^T\mathbf{F}_{12}\tilde{\mathbf{m}}_1 ds_2 & = 0 \\ \tilde{\mathbf{m}}_3^T\mathbf{F}_{23}\tilde{\mathbf{t}}_2 ds_2 & +\tilde{\mathbf{t}}_3^T\mathbf{F}_{23}\tilde{\mathbf{m}}_2 ds_3 & = 0 \\ \tilde{\mathbf{t}}_1^T\mathbf{F}_{31}\tilde{\mathbf{m}}_3 ds_1 & +\tilde{\mathbf{m}}_1^T\mathbf{F}_{31}\tilde{\mathbf{t}}_3 ds_3 & = 0 \end{array} \tag{12}$$

in which the vectors $\tilde{\mathbf{t}}_i$, $i = 1, 2, 3$ are equal to

$$\tilde{\mathbf{t}}_i = \left[\begin{array}{c} \mathbf{t}_i \\ 0 \end{array} \right]$$

where the vectors $\mathbf{t}_i$, $i = 1, 2, 3$ are unit tangent vectors to the corresponding curves. Since we know that there exists a non-zero solution to this system of three linear homogeneous equations in the differentials of the arclengths, the corresponding determinant must be equal to 0:

$$\begin{vmatrix} \tilde{\mathbf{m}}_2^T \mathbf{F}_{12} \tilde{\mathbf{t}}_1 & \tilde{\mathbf{t}}_2^T \mathbf{F}_{12} \tilde{\mathbf{m}}_1 & 0 \\ 0 & \tilde{\mathbf{m}}_3^T \mathbf{F}_{23} \tilde{\mathbf{t}}_2 & \tilde{\mathbf{t}}_3^T \mathbf{F}_{23} \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{t}}_1^T \mathbf{F}_{31} \tilde{\mathbf{m}}_3 & 0 & \tilde{\mathbf{m}}_1^T \mathbf{F}_{31} \tilde{\mathbf{t}}_3 \end{vmatrix}$$

Expanding this determinant yields the nice symmetric formula:

$$\tilde{\mathbf{m}}_1^T \mathbf{F}_{31} \tilde{\mathbf{t}}_3 \; \tilde{\mathbf{m}}_2^T \mathbf{F}_{12} \tilde{\mathbf{t}}_1 \; \tilde{\mathbf{m}}_3^T \mathbf{F}_{23} \tilde{\mathbf{t}}_2 \; + \; \tilde{\mathbf{t}}_1^T \mathbf{F}_{31} \tilde{\mathbf{m}}_3 \; \tilde{\mathbf{t}}_2^T \mathbf{F}_{12} \tilde{\mathbf{m}}_1 \; \tilde{\mathbf{t}}_3^T \mathbf{F}_{23} \tilde{\mathbf{m}}_2 \; = \; 0 \qquad (13)$$

which is linear in the coordinates of the unit vector $\mathbf{t}_3$ tangent to (c_3). This equation allows us to compute the direction of the tangent to the curve (c_3) from:

1. The points m_1 and m_2 in the first two images.

2. The unit tangent vectors $\mathbf{t}_1$ and $\mathbf{t}_2$ to the two image curves (c_1) and (c_2).

3. The corresponding point m_3 in the third image.

In fact the formula is *trilinear* in the coordinates of the three tangent vectors $\mathbf{t}_1, \mathbf{t}_2$ and $\mathbf{t}_3$.

3.3   Prediction of curvatures

We can also obtain a prediction formula for the curvature as follows. We differentiate equation (13) with respect to the three arclengths and use the standard Frenet formula $\frac{d\mathbf{t}}{ds} = \kappa \mathbf{n}$, where κ is the curvature of and $\mathbf{n}$ the unit normal to the curve. It should be clear that this yields an expression of the form:

$$(a_1 \kappa_1 + b_1) ds_1 + (a_2 \kappa_2 + b_2) ds_2 + (a_3 \kappa_3 + b_3) ds_3 = 0$$

where the coefficients a_i and b_i are functions of the three points m_i, the three tangents $\mathbf{t}_i$ and the three normals $\mathbf{n}_i$.

If we now use the relations (12), to eliminate ds_2 and ds_3 for example, we obtain an expression of the form:

$$(A_1 \kappa_1 + A_2 \kappa_2 + A_3 \kappa_3 + B) ds_1 = 0$$

implies:

$$A_1 \kappa_1 + A_2 \kappa_2 + A_3 \kappa_3 + B = 0 \qquad (14)$$

This is again a nice symmetric formula that allows to compute the curvature in the third image as an affine function of the curvatures in the other two. It turns out that the constant term is 0, as shown next. Thus we have the following *linear* formula between the curvatures at three corresponding points:

$$A_1 \kappa_1 + A_2 \kappa_2 + A_3 \kappa_3 = 0 \qquad (15)$$

The complete computation is performed in appendix B.

3.3.1 The constant term in the curvature prediction equation is 0

We prove this result in three steps. We first choose special coordinate systems in which the relationships between three-dimensional and image entities can be expressed very simply. Second, we express the tangent prediction equation (13) in these coordinate systems and, third, express the constant term of equation (14) in the same coordinate systems and show that it is identically zero.

Choosing coordinate systems In image i, $i = 1, 2, 3$, we consider the two epipoles e_{ii+1}, e_{ii+2}. Let also $m_i, i = 1, 2, 3$ be three corresponding points, the images of a three-dimensional point M, and C_i, $i = 1, 2, 3$ the three optical centers. If M, C_1, C_2, C_3 are in general positions (i.e. no three of them are aligned, which we will assume in what follows), we can take these four points and the point C given by:

$$\tilde{\mathbf{C}} = \tilde{\mathbf{m}} + \sum_{i=1}^{3} \tilde{\mathbf{C}}_i$$

as a projective basis of the three-dimensional space considered as embedded in the projective space $\mathcal{P}^3$. We thus choose their projective coordinates to be

$$\tilde{\mathbf{C}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \tilde{\mathbf{C}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \tilde{\mathbf{C}}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \tilde{\mathbf{m}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \tilde{\mathbf{C}} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

We then choose three special image coordinate systems, one in each image. For each i, $i = 1, 2, 3$, the point m_i is chosen to be the origin of coordinates $[0, 0, 1]^T$ and the two epipoles e_{ii+1}, e_{ii+2} are chosen to have coordinates $[1, 0, 0]^T$ and $[0, 1, 0]^T$. The point $c_i, i = 1, 2, 3$, image of the point C in the i-th image has coordinates $[1, 1, 1]^T$. The corresponding situation is shown in figure 6 where the points c_i do not appear.

Let now $\tilde{\mathbf{P}}_i$, $i = 1, 2, 3$ be the three 3×4 perspective projection matrixes relating a three-dimensional point Q to its three images q_i in projective coordinates. In other words

$$\tilde{\mathbf{q}}_i = \rho \tilde{\mathbf{P}}_i \tilde{\mathbf{Q}}$$

for some non-zero scalar ρ. In particular, we have

$$\tilde{\mathbf{e}}_{ij} = \rho_j \tilde{\mathbf{P}}_i \tilde{\mathbf{C}}_j, \ j = i + 1, i + 2$$

and

$$\tilde{\mathbf{P}}_i \tilde{\mathbf{C}}_i = \mathbf{0}, \ i = 1, 2, 3$$

Both equations are simply translations of the well-known relations between epipoles and optical centers. From them we can deduce the analytical form of the matrixes $\tilde{\mathbf{P}}_i$:

$$\tilde{\mathbf{P}}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \tilde{\mathbf{P}}_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \tilde{\mathbf{P}}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
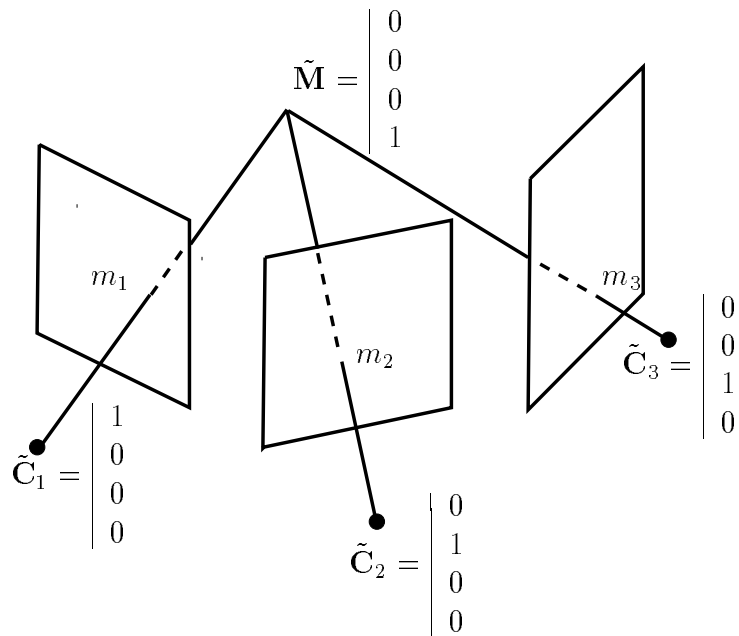
$$\tilde{\mathbf{M}} = \begin{vmatrix} 0 \\ 0 \\ 0 \\ 1 \end{vmatrix}$$

$$\tilde{\mathbf{C}}_3 = \begin{vmatrix} 0 \\ 0 \\ 1 \\ 0 \end{vmatrix}$$

$$\tilde{\mathbf{C}}_1 = \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

$$\tilde{\mathbf{C}}_2 = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 0 \end{vmatrix}$$

m_1 m_2 m_3

Figure 5: The choice of a special 3-D coordinate system: the origin is M and the three optical centers C_1, C_2 and C_3 are in the plane at infinity.

- $m_1 = [0, 0, 1]^T$

- $e_{13} = [0, 1, 0]^T$
- $e_{12} = [1, 0, 0]^T$

- $e_{21} = [0, 1, 0]^T$

- $e_{23} = [1, 0, 0]^T$
- $m_2 = [0, 0, 1]^T$

- $e_{31} = [1, 0, 0]^T$
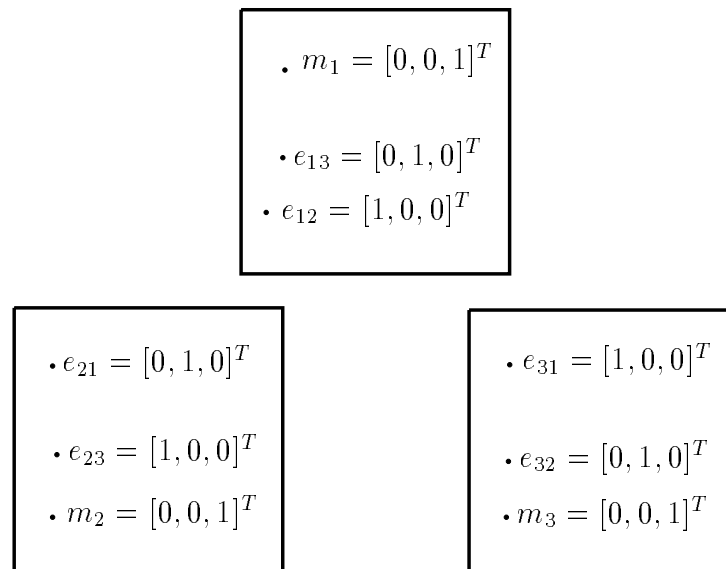
- $e_{32} = [0, 1, 0]^T$
- $m_3 = [0, 0, 1]^T$

Figure 6: This choice of image coordinates greatly simplifies the computations.

To determine the F-matrixes $\mathbf{F}_{ii+1}, i \neq j$ we use the four correspondences:

$$e_{ii+1} \to 0 \qquad e_{ii+2} \to \langle e_{i+1i}, e_{i+1,i+2} \rangle$$
$$c_i \to \langle e_{i+1i}, c_{i+1} \rangle \qquad m_i \to \langle e_{i+1i}, m_{i+1} \rangle$$

Note that the second correspondence comes from equation (9). They yield immediatly:

$$\mathbf{F}_{12} = \mathbf{F}_{23} = \mathbf{F}_{31} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{16}$$

We will use these special coordinate systems to express equation (13) and the term B in equation (14).

This is allright even though the corresponding changes of coordinates in the retinal planes are projective and we are considering *metric* quantities. Indeed, equation (13) is homogeneous of degree one in $\tilde{\mathbf{m}}_i$ and $\tilde{\mathbf{t}}_i$, $i = 1, 2, 3$ and therefore the fact that it is zero is a projective invariant. Similarly, as shown later, the term B in equation (14) is homogeneous of degree 2 in $\tilde{\mathbf{t}}_1$ and of degree 1 in the other terms.

Relation between the tangents Let us write $\tilde{\mathbf{t}}_i = [x_i, y_i, z_i]^T$, $i = 1, 2, 3$ and rewrite equation (13) with the previous expressions for the F-matrixes. We readily obtain:

$$y_1 y_2 y_3 - x_1 x_2 x_3 = 0 \tag{17}$$

which we will use in the next section.

The constant term of the curvature equation The constant term of equation (14) can be shown to be equal to:

$$\tilde{\mathbf{t}}_1^T \mathbf{F}_{31} \tilde{\mathbf{t}}_3 \; \tilde{\mathbf{m}}_2^T \mathbf{F}_{12} \tilde{\mathbf{t}}_1 \; \tilde{\mathbf{m}}_3^T \mathbf{F}_{23} \tilde{\mathbf{t}}_2 + \tilde{\mathbf{t}}_1^T \mathbf{F}_{31} \tilde{\mathbf{m}}_3 \; \tilde{\mathbf{t}}_2^T \mathbf{F}_{12} \tilde{\mathbf{t}}_1 \; \tilde{\mathbf{t}}_3^T \mathbf{F}_{23} \tilde{\mathbf{m}}_2 - \tilde{\mathbf{t}}_1^T \mathbf{F}_{31} \tilde{\mathbf{m}}_3 \; \tilde{\mathbf{m}}_2^T \mathbf{F}_{12} \tilde{\mathbf{t}}_1 \; \tilde{\mathbf{t}}_3^T \mathbf{F}_{23} \tilde{\mathbf{t}}_2$$

As we mentioned before this term is homogeneous of degree 2 in $\tilde{\mathbf{t}}_1$ and of degree 1 in the other terms. Therefore if it takes the value 0 in some coordinate systems, it will remain equal to 0 after any projective change of coordinates.

Expressing all this in function of x_i, y_i, z_i, $i = 1, 2, 3$, we find:

$$z_1 (y_1 y_2 y_3 - x_1 x_2 x_3)$$

and, according to equation (17), this is equal to 0.

4 The orthographic case

The case of an orthographic projection can be easily deduced from the analysis done in the previous section. Indeed, in the orthographic case, the three centers of projection C_i, $i = 1, 2, 3$ are at infinity which means in the plane at infinity. This is precisely what we have achieved by choosing the coordinates of C_1, C_2 and C_3 in the plane at infinity in figure

6). Note that matrix $\tilde{\mathbf{P}}_3$ is the same as in equation (7) which describes an orthographic projection. As a consequence of this, the epipoles are also at infinity in each retinal plane since they lie on the intersections of these retinal planes with the trifocal plane which is the plane at infinity in that case.

Let m_1, m_2 and m_3 be three point correspondences. Because the epipoles are at infinity in this case, it is possible, by performing *affine* changes of coordinate systems in the three retinal planes to arrange things in such a way that the coordinates of the epipoles and those of the points m_i, $i = 1, 2, 3$ are as in figure 6. Indeed, an affine transformation is a projective transformation which preserves the line at infinity. It is defined by three point correspondences. Its matrix $\mathcal{A}$ can be written as:

$$\mathcal{A} = \left[\begin{array}{cc} \mathbf{A} & \mathbf{b} \\ \mathbf{0}_2^T & 1 \end{array} \right] \tag{18}$$

where $\mathcal{A}$ is a nonsingular 2×2 matrix and $\mathbf{b}$ a 2×1 vector. Note that we use the projective notations even though we are dealing with the affine case. The reason for this is that the prediction equation (10) is projective. (18) defines an affine transformation because of its third row which ensures that it preserves the line at infinity.

The three point correspondences we have are

$$\tilde{\mathbf{m}}_i \leftrightarrow \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right] \quad \tilde{\mathbf{e}}_{ii+1} \leftrightarrow \left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right] \quad \tilde{\mathbf{e}}_{ii+2} \leftrightarrow \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right] \quad \tilde{\mathbf{c}}_i \leftrightarrow \left[\begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right]$$

Even though it looks as if we have four point correspondences, the fact is that we have only three since, the points e_{ii+1} and e_{ii+2} being at infinity, they represent directions and therefore "count" only for half points. The point m_i is the origin of the new affine coordinate system and the points at infinity e_{ii+1} and e_{ii+2} define the directions of its axes.

Let us denote by $\mathbf{F}^0$ the fundamental matrix defined by equation (16). We thus have

$$\mathbf{F}_{ij} = \mathcal{A}_j^T \mathbf{F}^0 \mathcal{A}_i \tag{19}$$

Let q_1, q_2 and q_3 be three corresponding pixels, the prediction equation (10) can be written:

$$\tilde{\mathbf{q}}_3 = \mathcal{A}_3^T \mathbf{F}^{0T} \mathcal{A}_1 \tilde{\mathbf{q}}_1 \wedge \mathcal{A}_3^T \mathbf{F}^0 \mathcal{A}_2 \tilde{\mathbf{q}}_2$$

Using a result proved in appendix A, this is the same as:

$$\mathbf{q}_3 = \mathcal{A}_3^{T*} (\mathbf{F}^{0T} \mathcal{A}_1 \mathbf{q}_1 \wedge \mathbf{F}^0 \mathcal{A}_2 \mathbf{q}_2) \tag{20}$$

where $\mathcal{A}_3^{T*}$ is the matrix of cofactors of matrix $\mathcal{A}_3^T$. Because of (18) and (16, we have:

$$\mathbf{F}^{0T} \mathcal{A}_1 = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \\ \alpha_1 & \beta_1 & \gamma_1 \end{array} \right] \quad \mathbf{F}^0 \mathcal{A}_2 = \left[\begin{array}{ccc} 0 & 0 & -1 \\ 0 & 0 & 0 \\ \alpha_2 & \beta_2 & \gamma_2 \end{array} \right]$$

where the α_i, β_i and γ_i depend upon the matrixes $\mathcal{A}_i$, $i = 1, 2$.

Taking now $\mathbf{q}_i = [x_i, y_i, 1]^T$, $i = 1, 2, 3$, we find that

$$\mathbf{F}^{0T}\mathcal{A}_1\mathbf{q}_1 \wedge \mathbf{F}^0\mathcal{A}_2\mathbf{q}_2 = \left[\begin{array}{c} \alpha_2 x_2 + \beta_2 y_2 + \gamma_2 \\ -(\alpha_1 x_1 + \beta_1 y_1 + \gamma_1) \\ 1 \end{array} \right]$$

Since $\mathcal{A}_3^{T*} = det(\mathcal{A}_3)\mathcal{A}_3^{-1}$, equation (20) shows that x_3 and y_3, the affine coordinates of q_3 are linear combinations of the coordinates of q_1 and q_2. This is the result used by Ullman and Basri [24] in their work on object recognition.

5 Experiments

5.1 Prediction on real images

A set of images has been acquired with a weakly calibrated stereo rig. Some point correspondences have been extracted by hand in two images. For each point, the differential properties of the corresponding edge have been computed. These are shown in the two upper images of figure 7 in the following way:

- a cross shows the predicted point,

- a line segment passing through this point indicates the tangent,

- a portion of circle tangent to this line segment, whose radius is the inverse of the predicted curvature, shows the curvature.

The prediction scheme has then been applied to the points, the tangents and the curvatures. The result of prediction in shown in the lower (third) image.

Figure 7 demonstrates the validity of the prediction scheme: indeed, both points, tangents and curvatures are correctly predicted. Furthermore, we can see that the process is accurate, in particular for predicting the point and the tangent. Predicted curvatures seem correct, but the figure does not precisely show how accurate the predicted curvatures are. For this reason, we ran another experiment, which is described in the next section.

5.2 Prediction of the equation of a conic given two points, two tangents, and one curvature

Representation of conics   Let us consider a conic (γ), of equation:

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

Since the equation is homogeneous in a, b, c, d, e, and f, the conic has 5 degrees of freedom. In other words, we should be able to compute the equation of (γ) if we know, for instance,

- two points that belong to (γ),

- the tangents to (γ) at these two points,

Figure 7: Point correspondences in the two top images are used to predict what happens in the third image. Points are represented by crosses, tangents by line segments, and curvatures by portions of osculating circles.

- the curvature of (γ) at one of them.

We will not give the details of the computation, which is quite simple. However, it is important to mention that two points and two tangents do not define a conic, but a one parameter family of conics. If we do not want to introduce another point or tangent line for defining the conic, then we can, for example, use the curvature information.

Planar conics in space   Let us now consider a planar conic (Γ) in space. Its image on camera i is a conic (γ_i), of equation:

$$a_i x^2 + 2b_i xy + c_i y^2 + d_i x + e_i y + f_i = 0$$

Let us assume that we know the equations of (γ_1) and (γ_2). Then, we can compute the equation of (γ_3) as soon as we know two point correspondences on curves $(\gamma_1),(\gamma_2)$ and that we can compute tangents and curvatures at those points (only one curvature is necessary). Indeed, we apply the prediction scheme to these five entities, and determine the conic in the third image defined by the predicted data.

We ran this process on a set of real images, in which one can find some ellipses, images of three-dimensional circles. The parameters of the ellipses are determined in images 1 and 2 by means of a simple fitting technique applied on to the edges of the ellipses [4]. Then, we generate some random point correspondences between the ellipses of images 1 and 2, which are used to predict the shape of the corresponding ellipse in the third image (figure 8). As we can see in the figure, the ellipse predicted with the pair of points (1,0) is better than the one derived from the pair (1,2). The reason is that the epipolar line in image 1 or 2 is almost tangent to the ellipse at point 2, so we are closer to a degenerate configuration.

5.3   Application to stereo matching

As we mentioned earlier, the prediction scheme is a way of eliminating ambiguity in the stereovision matching process. Most trinocular stereo algorithms use the prediction scheme on points. Only a few of them used the geometric property of the tangent [13, 1, 26], formulated as a three-dimensional constraint, thus assuming strong calibration of the cameras.

A trinocular stereovision algorithm was recently developped, based on the prediction scheme described in this article [19]. Here, we show the results provided by this algorithm on triplet of office images. For each trinocular correspondence of edges, we represent with thin curves the two edges used for prediction, and with a thick curve the predicted one. The algorithm switches images according to the orientation of the edges, so that the epipolar lines in the two images used for prediction intersect the edges as accurately as possible.

6   Conclusion

We have shown in this paper that simple ideas of projective geometry could shed light on a class of questions in computer vision and give very simple theoretical answers to these questions. The class of questions is summarized in the title of the paper: what can two
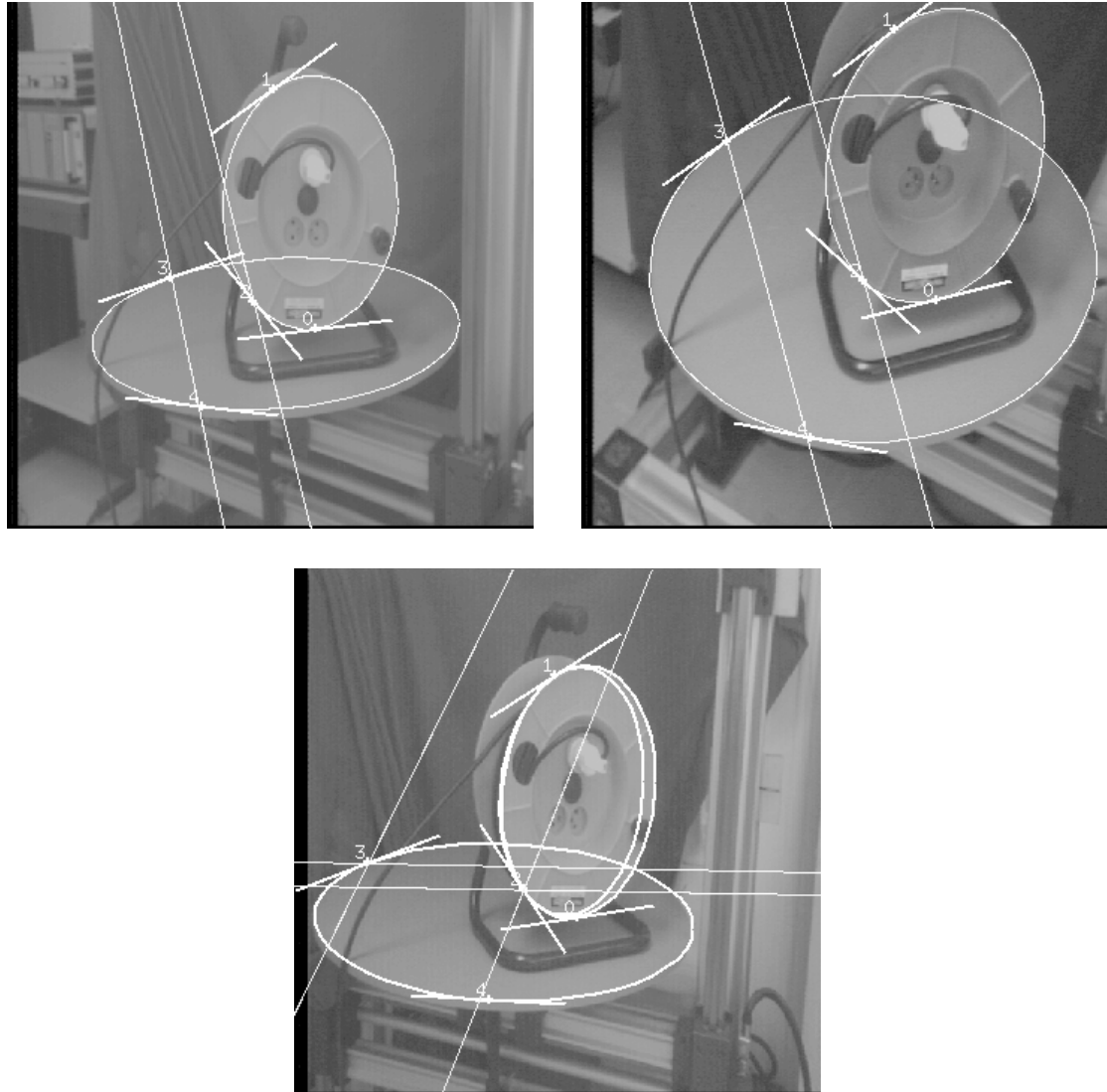
Figure 8: Prediction of ellipses in the third image: The points and tangents shown in the two top images are used to predict the ones of the third image. In this image, we show the ellipses predicted from pairs of points (1,2), (1,0) and (3,4). In the three images, we show the epipolar lines issued from points 2 and 3 in images 1 and 2.

Figure 9: Application of the prediction scheme to the problem of stereo matching.

images tell us about a third one? the answer we have given bears upon several areas of computer vision, stereo, motion analysis, and model-based object recognition.

As shown in the experimental part of the paper these ideas from projective geometry not only produce theoretical answers to the questions but also lead to robust algorithms that have been used on real images. We have shown results of these algorithms for the prediction of features from two images to a third but clearly they could also be used as a building block of a model-based recognition system.

A   Proof of a result

Let $\mathbf{A}$ be a nonsingular 3×3 matrix. We have the following relation:

$$\mathbf{A}\mathbf{x} \wedge \mathbf{A}\mathbf{y} = \mathbf{A}^*(\mathbf{x} \wedge \mathbf{y})$$

for all vectors $\mathbf{x}$ and $\mathbf{y}$. Matrix $\mathbf{A}^*$ is the matrix of cofactors of matrix $\mathbf{A}$.

The proof is straightforward and involves the three column vectors $\mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_3$ of $\mathbf{A}$. We first observe that

$$\mathbf{A}\mathbf{x} = \sum_{i=1}^{3} x_i \mathbf{a}_i$$

Next we expand

$$\mathbf{A}\mathbf{x} \wedge \mathbf{A}\mathbf{y} = (x_1 y_2 - x_2 y_1)\mathbf{a}_1 \wedge \mathbf{a}_2 + (x_2 y_3 - x_3 y_2)\mathbf{a}_2 \wedge \mathbf{a}_3 + (x_3 y_1 - x_1 y_3)\mathbf{a}_3 \wedge \mathbf{a}_1$$

which is equal to

$$[\mathbf{a}_2 \wedge \mathbf{a}_3 \quad \mathbf{a}_3 \wedge \mathbf{a}_1 \quad \mathbf{a}_1 \wedge \mathbf{a}_2](\mathbf{x} \wedge \mathbf{y})$$

we recognize the matrix $[\mathbf{a}_2 \wedge \mathbf{a}_3 \quad \mathbf{a}_3 \wedge \mathbf{a}_1 \quad \mathbf{a}_1 \wedge \mathbf{a}_2]$ as the matrix $\mathbf{A}^*$ of cofactors of matrix $\mathbf{A}$.

B   The curvature equation

Let us first introduce some new and more concise notations. If $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ represent two symbols chosen among $\tilde{\mathbf{m}}, \tilde{\mathbf{t}}, \tilde{\mathbf{n}}$, We denote by $\boldsymbol{\alpha}_j \boldsymbol{\beta}_i$ the product $\boldsymbol{\alpha}_j \mathbf{F}_{ij} \boldsymbol{\beta}_i$ (for instance, $\tilde{\mathbf{m}}_2 \mathbf{F}_{12} \tilde{\mathbf{t}}_1$ is denoted by $\tilde{\mathbf{m}}_2 \tilde{\mathbf{t}}_1$). The fact that $\mathbf{F}_{ji} = \mathbf{F}_{ij}^T$ implies that $\boldsymbol{\alpha}_j \boldsymbol{\beta}_i = \boldsymbol{\beta}_i \boldsymbol{\alpha}_j$.

Using these notations, (12) can be rewritten as

$$\forall i, j \quad \tilde{\mathbf{t}}_i \tilde{\mathbf{m}}_j ds_i + \tilde{\mathbf{m}}_i \tilde{\mathbf{t}}_j ds_j = 0$$

and (13) becomes

$$\tilde{\mathbf{m}}_1 \tilde{\mathbf{t}}_3 \ \tilde{\mathbf{m}}_2 \tilde{\mathbf{t}}_1 \ \tilde{\mathbf{m}}_3 \tilde{\mathbf{t}}_2 \ + \ \tilde{\mathbf{t}}_1 \tilde{\mathbf{m}}_3 \ \tilde{\mathbf{t}}_2 \tilde{\mathbf{m}}_1 \ \tilde{\mathbf{t}}_3 \tilde{\mathbf{m}}_2 \ = \ 0$$

Differentiating equation 13 yields

$$\kappa_1 \quad (\tilde{m}_1\tilde{t}_3 \ \tilde{m}_2\tilde{n}_1 \ \tilde{m}_3\tilde{t}_2 \ + \ \tilde{n}_1\tilde{m}_3 \ \tilde{t}_2\tilde{m}_1 \ \tilde{t}_3\tilde{m}_2) \ ds_1 \ +$$
$$\kappa_2 \quad (\tilde{m}_1\tilde{t}_3 \ \tilde{m}_2\tilde{t}_1 \ \tilde{m}_3\tilde{n}_2 \ + \ \tilde{t}_1\tilde{m}_3 \ \tilde{n}_2\tilde{m}_1 \ \tilde{t}_3\tilde{m}_2) \ ds_2 \ +$$
$$\kappa_3 \quad (\tilde{m}_1\tilde{n}_3 \ \tilde{m}_2\tilde{t}_1 \ \tilde{m}_3\tilde{t}_2 \ + \ \tilde{t}_1\tilde{m}_3 \ \tilde{t}_2\tilde{m}_1 \ \tilde{n}_3\tilde{m}_2) \ ds_3 \ +$$
$$\quad (\tilde{t}_1\tilde{t}_3 \ \tilde{m}_2\tilde{t}_1 \ \tilde{m}_3\tilde{t}_2 \ + \ \tilde{t}_1\tilde{m}_3 \ \tilde{t}_2\tilde{t}_1 \ \tilde{t}_3\tilde{m}_2) \ ds_1 \ +$$
$$\quad (\tilde{m}_1\tilde{t}_3 \ \tilde{t}_2\tilde{t}_1 \ \tilde{m}_3\tilde{t}_2 \ + \ \tilde{t}_1\tilde{m}_3 \ \tilde{t}_2\tilde{m}_1 \ \tilde{t}_3\tilde{t}_2) \ ds_2 \ +$$
$$\quad (\tilde{m}_1\tilde{t}_3 \ \tilde{m}_2\tilde{t}_1 \ \tilde{t}_3\tilde{t}_2 \ + \ \tilde{t}_1\tilde{m}_3 \ \tilde{t}_2\tilde{m}_1 \ \tilde{t}_3\tilde{m}_2) \ ds_3 \ = \ 0$$

After factorizing ds_1 and performing some simple algebraic manipulation we find:

$$\kappa_1 \quad (\tilde{m}_1\tilde{t}_3 \ \tilde{m}_2\tilde{n}_1 \ \tilde{m}_3\tilde{t}_2 \ + \ \tilde{n}_1\tilde{m}_3 \ \tilde{t}_2\tilde{m}_1 \ \tilde{t}_3\tilde{m}_2) \ \tilde{m}_1\tilde{t}_2 \ \tilde{m}_1\tilde{t}_3 \ -$$
$$\kappa_2 \quad (\tilde{m}_1\tilde{t}_3 \ \tilde{m}_2\tilde{t}_1 \ \tilde{m}_3\tilde{n}_2 \ + \ \tilde{t}_1\tilde{m}_3 \ \tilde{n}_2\tilde{m}_1 \ \tilde{t}_3\tilde{m}_2) \ \tilde{m}_2\tilde{t}_1 \ \tilde{m}_1\tilde{t}_3 \ -$$
$$\kappa_3 \quad (\tilde{m}_1\tilde{n}_3 \ \tilde{m}_2\tilde{t}_1 \ \tilde{m}_3\tilde{t}_2 \ + \ \tilde{t}_1\tilde{m}_3 \ \tilde{t}_2\tilde{m}_1 \ \tilde{n}_3\tilde{m}_2) \ \tilde{m}_1\tilde{t}_2 \ \tilde{m}_3\tilde{t}_1 \ +$$
$$2 \quad (\tilde{t}_1\tilde{t}_3 \ \tilde{m}_2\tilde{t}_1 \ \tilde{m}_3\tilde{t}_2 \ + \ \tilde{t}_1\tilde{m}_3 \ \tilde{t}_2\tilde{t}_1 \ \tilde{t}_3\tilde{m}_2 \ - \ \tilde{t}_1\tilde{m}_3 \ \tilde{m}_2\tilde{t}_1 \ \tilde{t}_3\tilde{t}_2) \ \tilde{m}_1\tilde{t}_2 \ \tilde{m}_1\tilde{t}_3 \ = \ 0$$

The last term is zero according to the proof given in paragraph 3.3.1. In this linear equation in the variables $\kappa_1, \kappa_2, \kappa_3$, the coefficient of κ_i is A_i introduced in paragraph 3.3.

References

[1] N. Ayache and F. Lustman. Fast and reliable passive trinocular stereovision. In *Proceedings ICCV '87, London*, pages 422–427. IEEE, June 1987.

[2] Eamon B. Barrett, Michael H. Brill, Nils N. Haag, and Paul M. Payton. Invariant Linear Methods in Photogrammetry and Model-Matching. In Joseph L. Mundy and Andrew Zimmerman, editors, *Geometric Invariance in Computer Vision*, chapter 14. MIT Press, 1992.

[3] Ronen Basri. On the Uniqueness of Correspondence under Orthographic and Perspective Projections. In *Proc. of the Image Understanding Workshop*, pages 875–884, 1993.

[4] R. Deriche. Using Canny's Criteria to Derive an Optimal Edge Detector Recursively Implemented. In *The International Journal of Computer Vision*, volume 2, pages 15–20, April 1987.

[5] Olivier D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig. In Giulio Sandini, editor, *Proceedings of the 2nd European Conference on Computer Vision*, pages 563–578. Springer-Verlag, Lecture Notes in Computer Science 588, May 1992.

[6] Olivier D. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993. To appear.

[7] Olivier D. Faugeras, Tuan Luong, and Steven Maybank. Camera self-calibration: theory and experiments. In Giulio Sandini, editor, *Proceedings of the 2nd European Conference on Computer Vision*, pages 321–334. Springer-Verlag, Lecture Notes in Computer Science 588, 1992.

[8] Richard Hartley, Rajiv Gupta, and Tom Chang. Stereo from Uncalibrated Cameras. In *Proceedings of CVPR92, Champaign, Illinois*, pages 761–764, June 1992.

[9] M. Ito and A. Ishii. Three-View Stereo Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:524–532, 1986.

[10] Kenichi Kanatani. Computational Projective Geometry. *CVGIP: Image Understanding*, 54(3):333–348, November 1991.

[11] Yoshifumi Kitamura and Masahiko Yachida. Three-dimensional data acquisition by trinocular vision. *Advanced Robotics*, 4(1):29–42, 1990. Robotics Society of Japan.

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

[13] V.J. Milenkovic and T. Kanade. Trinocular Vision Using Photometric and Edge Orientation Constraints. In *Proceedings of DARPA Image Understanding Workshop*, pages 163–175, , December 1985.

[14] R. Mohr and E. Arbogast. It can be done without camera calibration. *Pattern Recognition Letters*, 12:39–43, 1990.

[15] Roger Mohr, Luce Morin, and Enrico Grosso. Relative positioning with poorly calibrated cameras. In J.L. Mundy and A. Zisserman, editors, *Proceedings of DARPA-ESPRIT Workshop on Applications of Invariance in Computer Vision*, pages 7–46, 1991.

[16] Joseph L. Mundy and Andrew Zimmerman, editors. *Geometric Invariance in Computer Vision*. MIT Press, 1992.

[17] M. Pietikainen and D. Harwood. Depth from three-camera stereo. In *Proc. International Conference on Computer Vision and Pattern Recognition*, pages 2–8. IEEE, 1986. Miami Beach, Florida.

[18] M. Pietikainen and D. Harwood. Progress in trinocular stereo. In *Proceedings NATO Advanced Workshop on Real-time Object and Environment Measurement and classification, Maratea, Italy*, August 31 - September 3 1987.

[19] L. Robert. *Perception Stéréoscopique de Courbes et de Surfaces Tridimensionnelles. Applications à la Robotique Mobile*. PhD thesis, Ecole Polytechnique, Mars 1993.

[20] Luc Robert and Olivier D. Faugeras. Curve-Based Stereo: Figural Continuity And Curvature. In *CVPR91*, pages 57–62. IEEE, June 1991. Maui, Hawai.

[21] J.G. Semple and G.T. Kneebone. *Algebraic Projective Geometry*. Oxford: Clarendon Press, 1952. Reprinted 1979.

[22] A. Shashua. On Geometric and Algebraic Aspects of 3D Affine and Projective Structures from Perspective 2D Views. Technical Report A.I. Memo No. 1405, MIT, July 1993.

[23] Amnon Shashua. Projective Depth: A Geometric Invariant for 3D Reconstruction From Two Perspective\Orthographic Views and For Visual Recognition. In *Proc. Fourth International Conference on Computer Vision*, pages 583–590, 1993.

[24] Shimon Ullman and Ronen Basri. Recognition by Linear Combinations of Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(10):992–1006, 1991.

[25] M. Yachida. 3D Data Acquisition by Multiple Views. In O.D. Faugeras and G. Giralt, editors, *Robotics Research: the Third International Symposium*, pages 11–18. MIT Press, Cambridge, Mass., 1986.

[26] D. Yang and J. Illingworth. Line Based Trinocular Stereo. In *British Machine Vision Conference*, pages 327–336, 1992. Leeds, UK.

INRIA