What can be seen in three dimensions with an uncalibrated stereo rig?

Olivier D. Faugeras

INRIA-Sophia, 2004 Route des Lucioles, 06560 Valbonne, France

Abstract. This paper addresses the problem of determining the kind of three-dimensional reconstructions that can be obtained from a binocular stereo rig for which no three-dimensional metric calibration data is available. The only information at our disposal is a set of pixel correspondences between the two retinas which we assume are obtained by some correlation technique or any other means. We show that even in this case some very rich non-metric reconstructions of the environment can nonetheless be obtained.

Specifically we show that if we choose five arbitrary correspondences, then a unique (up to an arbitrary projective transformation) *projective* representation of the environment can be constructed which is *relative* to the five points in three-dimensional space which gave rise to the correspondences.

We then show that if we choose only four arbitrary correspondences, then an *affine* representation of the environment can be constructed. This reconstruction is defined up to an arbitrary affine transformation and is *relative* to the four points in three-dimensional space which gave rise to the correspondences. The reconstructed scene also depends upon three arbitrary parameters and two scenes reconstructed from the same set of correspondences with two different sets of parameter values are related by a projective transformation.

Our results indicate that computer vision may have been slightly overdoing it in trying at all costs to obtain metric information from images. Indeed, our past experience with the computation of such information has shown us that it is difficult to obtain, requiring awkward calibration procedures and special purpose patterns which are difficult if not impossible to use in natural environments with active vision systems. In fact it is not often the case that accurate metric information is necessary for robotics applications for example where *relative* information is usually all what is needed.

1 Introduction

The problem we address in this paper is that of a machine vision system with two cameras, sometimes called a stereo rig, to which no three-dimensional *metric* information has been made available. The only information at hand is contained in the two images. We assume that this machine vision system is capable, by comparing these two images, of establishing correspondences between them. These correspondences can be based on some measures of similitude, perhaps through some correlation-like process. Anyway, we assume that our system has obtained by some means a number of point correspondences. Each such correspondence, noted (m, m') indicates that the two image points m and m' in the two retinas are very likely to be the images of the *same* point out there. It is very doubtful at first sight that such a system can reconstruct anything useful at all. In the machine vision jargon, it does not know either its intrinsic parameters (one set for each camera), nor its extrinsic parameters (relative position and orientation of the cameras).

Surprisingly enough, it turns out that the machine vision system can nonetheless reconstruct some very rich non-metric representations of its environment. These representations are defined up to certain transformations of the environment which we assume to be three-dimensional and euclidean (a realistic assumption which may be criticized by some people). These transformations can be either affine or projective transformations of the surrounding space. This depends essentially on the user (i.e the machine vision system) choice.

This work has been inspired by the work of Jan Koenderink and Andrea van Doorn [4], the work of Gunnar Sparr [9,10], and the work of Roger Mohr and his associates [6,7].

We use the following notations. Vectors and matrixes will be represented in boldface, geometric entities such as points and lines in normal face. For example, m represents a point and $\mathbf{m}$ the vector of the coordinates of the point. The line defined by two points M and N will be denoted by $\langle M, N\rangle$. We will assume that the reader is familiar with elementary projective geometry such as what can be found in [8].

2 The projective case: basic idea

In all the paper we will assume the simple pinhole model for the cameras. In this model, the camera performs a perspective projection from the three-dimensional ambient space considered as a subset of the projective space $\mathcal{P}^3$ to the two-dimensional retinal space considered as a subset of the projective plane $\mathcal{P}^2$. This perspective projection can be represented linearly in projective coordinates. If m is a retinal point represented by the three-dimensional vector $\mathbf{m}$, image of the point M represented by the four-dimensional vector $\mathbf{M}$, the perspective projection is represented by a 3×4 matrix, noted $\tilde{\mathbf{P}}$, such that:

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

Assume now that we are given 5 point matches in two images of a stereo pair. Let A_i, $i = 1, \cdots, 5$ be the corresponding 3D points. We denote their images in the two cameras by a_i, a'_i, $i = 1, 5$. We make three choices of coordinate systems:

in 3D space choose the five (unknown) 3D points as the standard projective basis, i.e $\mathbf{A}_1 = \mathbf{e}_1 = [1, 0, 0, 0]^T, \cdots, \mathbf{A}_5 = \mathbf{e}_5 = [1, 1, 1, 1]^T$.

in the first image choose the four points a_i, $i = 1, \cdots, 4$ as the standard projective basis, i.e, for example $\mathbf{a}_1 = [1, 0, 0]^T$.

in the second image do a similar change of coordinates with the points a'_i, $i = 1, \cdots, 4$.

With those three choices of coordinates, the expressions for the perspective matrixes $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$ for the two cameras are quite simple. Lets us compute it for the first one.

2.1 A simple expression for $\tilde{\mathbf{P}}$

We write that

$$\tilde{\mathbf{P}}\mathbf{A}_i = \rho_i \mathbf{a}_i \quad i = 1, \cdots, 4$$

which implies, thanks to our choice of coordinate systems, that $\tilde{\mathbf{P}}$ has the form:

$$\tilde{\mathbf{P}} = \begin{bmatrix} \rho_1 & 0 & 0 & \rho_4 \\ 0 & \rho_2 & 0 & \rho_4 \\ 0 & 0 & \rho_3 & \rho_4 \end{bmatrix} \tag{1}$$

Let $\mathbf{a}_5 = [\alpha,\ \beta,\ \gamma]^T$, then the relation $\tilde{\mathbf{P}}\mathbf{A}_5 = \rho_5 \mathbf{a}_5$ yields the three equations:

$$\rho_1 + \rho_4 = \rho_5 \alpha \quad \rho_2 + \rho_4 = \rho_5 \beta \quad \rho_3 + \rho_4 = \rho_5 \gamma$$

We now define $\mu = \rho_5$ and $\nu = \rho_4$, matrix $\tilde{\mathbf{P}}$ can be written as a very simple function of the two unknown parameters μ and ν:

$$\tilde{\mathbf{P}} = \mu \tilde{\mathbf{X}} + \nu \tilde{\mathbf{Y}} \tag{2}$$

$$\tilde{\mathbf{X}} = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \\ 0 & 0 & \gamma & 0 \end{bmatrix} \tag{3}$$

$$\tilde{\mathbf{Y}} = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \tag{4}$$

A similar expression holds for $\tilde{\mathbf{P}}'$ which is a function of two unknown parameters μ' and ν':

$$\tilde{\mathbf{P}}' = \mu' \tilde{\mathbf{X}}' + \nu' \tilde{\mathbf{Y}}$$

2.2 Optical centers and epipoles

Equation (2) shows that each perspective matrix depends upon two projective parameters i.e of one parameter. Through the choice of the five points A_i, $i = 1, \cdots, 5$ as the standard coordinate system, we have reduced our stereo system to be a function of only two arbitrary parameters. What have we lost? well, suppose we have another match $(m,\ m')$, it means that we can compute the coordinates of the corresponding three-dimensional point M as a function of two *arbitrary* parameters in the projective coordinate system defined by the five points A_i, $i = 1, \cdots, 5$. Our three-dimensional reconstruction is thus defined up to the projective transformation (unknown) from the absolute coordinate system to the five points A_i and up to the two unknown parameters which we can choose as the ratios $x = \frac{\mu}{\nu}$ and $x' = \frac{\mu'}{\nu'}$. We will show in a moment how to eliminate the dependency upon x and x' by using a few more point matches.

Coordinates of the optical centers and epipoles Let us now compute the coordinates of the optical centers C and C' of the two cameras. We know that the coordinates of C are defined by the equation:

$$\tilde{\mathbf{P}}\mathbf{C} = \mathbf{0}$$

Combining this with the expression (2) for $\tilde{\mathbf{P}}$, we obtain:

$$\mathbf{C} = \left[\frac{1}{\nu - \alpha\mu},\ \frac{1}{\nu - \beta\mu},\ \frac{1}{\nu - \gamma\mu},\ \frac{1}{\nu}\right]^T$$

a set of remarkably simple expressions. Note that the coordinates of C depend only upon the ratio x:

$$\mathbf{C} = \left[\frac{1}{1 - \alpha x},\ \frac{1}{1 - \beta x},\ \frac{1}{1 - \gamma x},\ 1\right]^T$$

Identical expressions are obtained for the coordinates of C' by adding $'$:

$$\mathbf{C}' = \left[\frac{1}{\nu' - \alpha'\mu'},\ \frac{1}{\nu' - \beta'\mu'},\ \frac{1}{\nu' - \gamma'\mu'},\ \frac{1}{\nu'}\right]^T \equiv \left[\frac{1}{1 - \alpha' x'},\ \frac{1}{1 - \beta' x'},\ \frac{1}{1 - \gamma' x'},\ 1\right]^T$$

If we now use the relation $\tilde{\mathbf{P}}'\mathbf{C} = \mathbf{o}'$ to define the epipole o' in the second image, we immediately obtain its coordinates:

$$\mathbf{o}' = [\frac{\mu'\alpha' - \nu'}{\nu - \mu\alpha} + \frac{\nu'}{\nu}, \frac{\mu'\beta' - \nu'}{\nu - \mu\beta} + \frac{\nu'}{\nu}, \frac{\mu'\gamma' - \nu'}{\nu - \mu\gamma} + \frac{\nu'}{\nu}]^T \equiv [\frac{x'\alpha' - x\alpha}{1 - x\alpha}, \frac{x'\beta' - x\beta}{1 - x\beta}, \frac{x'\gamma' - x\gamma}{1 - x\gamma}]^T \tag{5}$$

We note that they depend only on the ratios x and x'. We have similar expressions for the epipole o defined by $\tilde{\mathbf{P}}\mathbf{C}' = \mathbf{o}$:

$$\mathbf{o} = [\frac{\mu\alpha - \nu}{\nu' - \mu'\alpha'} + \frac{\nu}{\nu'}, \frac{\mu\beta - \nu}{\nu' - \mu'\beta'} + \frac{\nu}{\nu'}, \frac{\mu\gamma - \nu}{\nu' - \mu'\gamma'} + \frac{\nu}{\nu'}]^T \equiv [\frac{x\alpha - x'\alpha'}{1 - x'\alpha'}, \frac{x\beta - x'\beta'}{1 - x'\beta'}, \frac{x\gamma - x'\gamma'}{1 - x'\gamma'}]^T \tag{6}$$

Constraints on the coordinates of the epipoles The coordinates of the epipoles are not arbitrary because of the *epipolar transformation*. This transformation is well-known in stereo and motion [3]. It says that the two pencils of epipolar lines are related by a collineation, i.e a linear transformation between projective spaces (here two projective lines). It implies that we have the equalities of two cross-ratios, for example:

$$\begin{aligned}\{\langle o, a_1\rangle, \langle o, a_2\rangle, \langle o, a_3\rangle, \langle o, a_4\rangle\} &= \{\langle o', a_1'\rangle, \langle o', a_2'\rangle, \langle o', a_3'\rangle, \langle o', a_4'\rangle\}\\ \{\langle o, a_1\rangle, \langle o, a_2\rangle, \langle o, a_3\rangle, \langle o, a_5\rangle\} &= \{\langle o', a_1'\rangle, \langle o', a_2'\rangle, \langle o', a_3'\rangle, \langle o', a_5'\rangle\}\end{aligned}$$

As shown in appendix A, we obtain the two relations (12) and (13) between the coordinates of o and o'.

3 Relative reconstruction of points

3.1 Complete determination of $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$

Assume for a moment that we know the epipoles o and o' in the two images (we show in section 3.3 how to estimate their coordinates). This allows us to determine the unknown parameters as follows. Let, for example, U', V' and W' be the projective coordinates of o'. According to equation (5), and after some simple algebraic manipulations, we have:

$$\frac{U'}{W'} = \frac{x\alpha - x'\alpha'}{x\gamma - x'\gamma'} \cdot \frac{x\gamma - 1}{x\alpha - 1} \quad \frac{V'}{W'} = \frac{x\beta - x'\beta'}{x\gamma - x'\gamma'} \cdot \frac{x\gamma - 1}{x\beta - 1}$$

If we think of the pair (x, x') as defining the coordinates of a point in the plane, these equations show that the points which are solutions are at the intersection of two conics. In fact, it is easy to show, using Maple, that there are three points of intersection whose coordinates are very simple:

$$\begin{aligned}&x = 0 \quad x' = 0\\ &x = \tfrac{1}{\gamma} \quad x' = \tfrac{1}{\gamma'}\\ &x = \frac{\mathbf{o}'\cdot(\mathbf{a}_5 \wedge \mathbf{a}_5')}{\mathbf{o}_5'\cdot(\mathbf{a}_5 \wedge \mathbf{a}_5')} \quad x' = x\frac{\gamma U'V'(\beta-\alpha) + \alpha V'W'(\gamma-\beta) + \beta W'U'(\alpha-\gamma)}{\gamma' U'V'(\beta-\alpha) + \alpha' V'W'(\gamma-\beta) + \beta' W'U'(\alpha-\gamma)}\end{aligned}$$

Where

$$\mathbf{o}_5' = [\alpha U', \beta V', \gamma W']^T$$

One of these points has to be a double point where the two conics are tangent. Since it is only the last pair (x, x') which is a function of the epipolar geometry, it is in general the only solution.

Note that since the equations of the two epipoles are related by the two equations described in appendix A, they provide only two independent equations rather than four.

The perspective matrixes $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$ are therefore uniquely defined. For each match (m, m') between two image points, we can then reconstruct the corresponding three-dimensional point M in the projective coordinate system defined by the five points A_i. Remember that those five points are unknown. Thus our reconstruction can be considered as *relative* to those five points and depending upon an arbitrary perspective transformation of the projective space $\mathcal{P}^3$. All this is completely independent of the intrinsic and extrinsic parameters of the cameras.

We have obtained a remarkably simple result:

In the case where at least eight point correspondences have been obtained between two images of an uncalibrated stereo rig, if we arbitrarily choose five of those correspondences and consider that they are the images of five points in general positions (i.e not four of them are coplanar), then it is possible to reconstruct the other three points and any other point arising from a correspondence between the two images in the projective coordinate system defined by the five points. This reconstruction is uniquely defined up to an unknown projective transformation of the environment.

3.2 Reconstructing the points

Given a correspondence (m, m'), we show how to reconstruct the three-dimensional point M in the projective coordinate system defined by the points A_i, $i = 1, \cdots, 5$.

The computation is extremely simple. Let M_∞ be the point of intersection of the optical ray $\langle C, m\rangle$ with the plane of equation $T = 0$. M_∞ satifies the equation $\mathbf{P}\mathbf{M}_\infty = \mathbf{m}$, where $\mathbf{P}$ is the 3×3 left submatrix of matrix $\tilde{\mathbf{P}}$ (note that $\mathbf{M}_\infty$ is a 3×1 vector, the projective representation of M_∞ being $[\mathbf{M}_\infty^T, 0]^T$). The reconstructed point M can then be written as

$$\mathbf{M} = \mu\mathbf{C} + \lambda \begin{bmatrix} \mathbf{M}_\infty \\ 0 \end{bmatrix}$$

where the scalars λ and μ are determined by the equation $\tilde{\mathbf{P}}'\mathbf{M} = \mathbf{m}'$ which says that m' is the image of M. Applying $\tilde{\mathbf{P}}'$ to both sides of the previous equation, we obtain

$$\mathbf{m}' = \mu\mathbf{o}' + \lambda\mathbf{P}'\mathbf{P}^{-1}\mathbf{m}$$

where $\mathbf{P}'$ is the 3×3 left submatrix of matrix $\tilde{\mathbf{P}}'$.

It is shown in appendix B that

$$\mathbf{P}'\mathbf{P}^{-1} = \begin{bmatrix} \frac{\alpha' x' - 1}{\alpha x - 1} & 0 & 0 \\ 0 & \frac{\beta' x' - 1}{\beta x - 1} & 0 \\ 0 & 0 & \frac{\gamma' x' - 1}{\gamma x - 1} \end{bmatrix}$$

Let us note $a = \frac{\alpha' x' - 1}{\alpha x - 1}$, $b = \frac{\beta' x' - 1}{\beta x - 1}$, and $c = \frac{\gamma' x' - 1}{\gamma x - 1}$. λ and μ are then found by solving the system of three linear equations in two unknowns

$$\mathbf{m}' = \mathbf{A} \begin{bmatrix} \mu \\ \lambda \end{bmatrix}$$

3.4 Choosing the five points A_i

[1] As mentioned before, in order for this scheme to work, the three-dimensional points that we choose to form the standard projective basis must be in general position. This means that no four of them can be coplanar. The question therefore arises of whether we can guarantee this *only* from their projections in the two retinas.

The answer is provided by the following observation. Assume that four of these points are coplanar, for example A_1, A_2, A_3, and A_4 as in figure 1. Therefore, the diagonals of the planar quadrilateral intersect at three points B_1, B_2, B_3 in the same plane. Because the perspective projections on the two retinas map lines onto lines, the images of these diagonals are the diagonals of the quadrilaterals a_1, a_2, a_3, a_4 and a'_1, a'_2, a'_3, a'_4 which intersect at b_1, b_2, b_3 and b'_1, b'_2, b'_3, respectively. If the four points A_i are coplanar, then the points b'_j, $j = 1, 2, 3$ lie on the epipolar line of the points b_j, simply because they are the images of the points B_j. Since we know the epipolar geometry of the stereo rig, this can be tested in the two images.

But this is only a necessary condition, what about the reverse? suppose then that b'_1 lies on the epipolar line of b_1. By construction, the line $\langle C, b_1 \rangle$ is a transversal to the two lines $\langle A_1, A_3 \rangle$ and $\langle A_2, A_4 \rangle$: it intersects them in two points C_1 and C_2. Similarly, $\langle C', b'_1 \rangle$ intersects $\langle A_1, A_3 \rangle$ and $\langle A_2, A_4 \rangle$ in C'_1 and C'_2. Because b'_1 lies on the epipolar line of b_1, the two lines $\langle C, b_1 \rangle$ and $\langle C', b'_1 \rangle$ are coplanar (they lie in the same epipolar plane). The discussion is on the four coplanar points C_1, C_2, C'_1, C'_2. Three cases occur:

1. $C_1 \neq C'_1$ and $C_2 \neq C'_2$ implies that $\langle A_1, A_3 \rangle$ and $\langle A_2, A_4 \rangle$ are in the epipolar plane and therefore that the points a_1, a_2, a_3, a_4 and a'_1, a'_2, a'_3, a'_4 are aligned on corresponding epipolar lines.
2. $C_1 \equiv C'_1$ and $C_2 \neq C'_2$ implies that $\langle A_1, A_3 \rangle$ is in the epipolar plane and therefore that the lines $\langle a_1, a_3 \rangle$ and $\langle a'_1, a'_3 \rangle$ are corresponding epipolar lines.
3. The case $C_1 \neq C'_1$ and $C_2 \equiv C'_2$ is similar to the previous one.
4. $C_1 \equiv C'_1$ and $C_2 \equiv C'_2$ implies that the two lines $\langle A_1, A_3 \rangle$ and $\langle A_2, A_4 \rangle$ are coplanar and therefore also the four points A_1, A_2, A_3, A_4 (in that cas we have $C_1 \equiv C'_1 \equiv C_2 \equiv C'_2 \equiv B_1$).

In conclusion, except for the first three "degenerate cases" which can be easily detected, the condition that b'_1 lies on the epipolar line of b_1 is necessary and sufficient for the four points A_1, A_2, A_3, A_4 to be coplanar.

4 Generalization to the affine case

The basic idea also works if instead of choosing five arbitrary points in space, we choose only four, for example A_i, $i = 1, \cdots, 4$. The transformation of space can now be chosen in such a way that it preserves the plane at infinity: it is an affine transformation. Therefore, in the case in which we choose four points instead of five as reference points, the local reconstruction will be up to an affine transformation of the three-dimensional space.

Let us consider again equation 1, change notations slightly to rewrite it as:

$$\tilde{\mathbf{P}} = \begin{bmatrix} p & 0 & 0 & s \\ 0 & q & 0 & s \\ 0 & 0 & r & s \end{bmatrix}$$

[1] This section was suggested to us by Roger Mohr.

where matrix

$$\mathbf{A} = \begin{bmatrix} U' & aX \\ V' & bY \\ W' & cZ \end{bmatrix}$$

is in general of rank 2. We then have

$$\begin{bmatrix} \mu \\ \lambda \end{bmatrix} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{m}' \tag{7}$$

The coordinates of the reconstructed point M are:

$$\mathbf{M} = \mu[\frac{1}{x\alpha - 1}, \frac{1}{x\beta - 1}, \frac{1}{x\gamma - 1}, -1]^T + \lambda[\frac{X}{x\alpha - 1}, \frac{Y}{x\beta - 1}, \frac{Z}{x\gamma - 1}, 0]^T$$

In which we have taken $\mathbf{m} = [X, Y, Z]^T$. We now explain how the epipoles can be determined.

3.3 Determining the epipoles from point matches

The epipoles and the epipolar transformation between the two retinas can be easily determined from the point matches as follows. For a given point m in the first retina, its epipolar line o_m in the second retina is linearly related to its projective representation. If we denote by $\mathbf{F}$ the 3×3 matrix describing the correspondence, we have:

$$\mathbf{o}_m = \mathbf{Fm}$$

where $\mathbf{o}_m$ is the projective representation of the epipolar line o_m. Since the corresponding point m' belongs to the line e_m by definition, we can write:

$$\mathbf{m}'^T\mathbf{Fm} = 0 \tag{8}$$

This equation is reminiscent of the so-called Longuet-Higgins equation in motion analysis [5]. This is not a coincidence.

Equation (8) is linear and homogeneous in the 9 unknown coefficients of matrix $\mathbf{F}$. Thus we know that, in general, if we are given 8 matches we will be able to determine a unique solution for $\mathbf{F}$, defined up to a scale factor. In practice, we are given much more than 8 matches and use a least-squares method. We have shown in [2] that the result is usually fairly insensitive to errors in the coordinates of the pixels m and m' (up to 0.5 pixel error).

Once we have obtained matrix $\mathbf{F}$, the coordinates of the epipole o are obtained by solving

$$\mathbf{Fo} = \mathbf{0} \tag{9}$$

In the noiseless case, matrix $\mathbf{F}$ is of rank 2 (see Appendix B) and there is a unique vector $\mathbf{o}$ (up to a scale factor) which satisfies equation 9. When noise is present, which is the standard case, o is determined by solving the following classical constrained minimization problem

$$\min_{\mathbf{o}} ||\mathbf{Fo}||^2 \text{ subject to } ||\mathbf{o}||^2 = 1$$

which yields $\mathbf{o}$ as the unit norm eigenvector of matrix $\mathbf{E}^T\mathbf{F}$ corresponding to the smallest eigenvalue. We have verified that in practice the estimation of the epipole is very insensitive to pixel noise.

The same processing applies in reverse to the computation of the epipole o'.

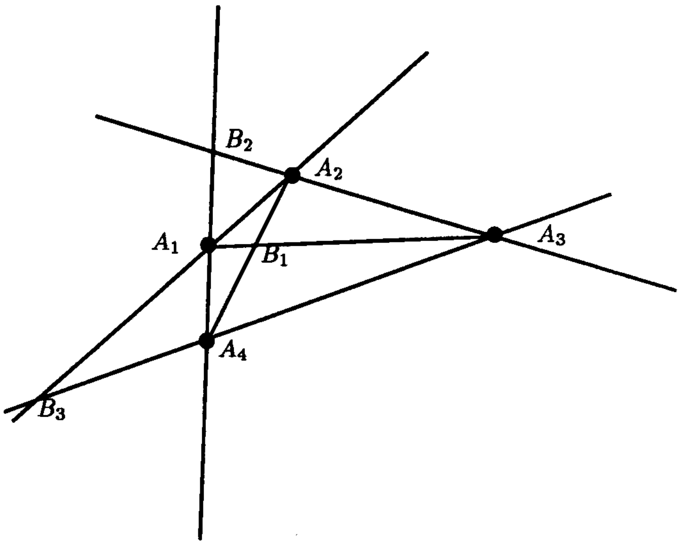

Fig. 1. If the four points A_1, A_2, A_3, A_4 are coplanar, they form a planar quadrilateral whose diagonals intersect at three points B_1, B_2, B_3 in the same plane

with a similar expression with $'$ for $\tilde{\mathbf{P}}'$. Each perspective matrix now depends upon 4 projective parameters, or 3 parameters, making a total of 6. If we assume, like previously, that we have been able to compute the coordinates of the two epipoles, then we can write four equations among these 6 unknowns, leaving two. Here is how it goes.

It is very easy to show that the coordinates of the two optical centers are:

$$\mathbf{C} = [\frac{1}{p}, \frac{1}{q}, \frac{1}{r}, -\frac{1}{s}]^T \quad \mathbf{C}' = [\frac{1}{p'}, \frac{1}{q'}, \frac{1}{r'}, -\frac{1}{s'}]^T$$

from which we obtain the coordinates of the two epipoles:

$$\mathbf{o} = [\frac{p}{p'} - \frac{s}{s'}, \frac{q}{q'} - \frac{s}{s'}, \frac{r}{r'} - \frac{s}{s'}]^T \quad \mathbf{o}' = [\frac{p'}{p} - \frac{s'}{s}, \frac{q'}{q} - \frac{s'}{s}, \frac{r'}{r} - \frac{s'}{s}]^T$$

Let us note

$$x_1 = \frac{p'}{p} \quad x_2 = \frac{q'}{q} \quad x_3 = \frac{r'}{r} \quad x_4 = \frac{s'}{s}$$

We thus have for the second epipole:

$$\frac{x_1 - x_4}{x_3 - x_4} = \frac{U'}{W'} \quad \frac{x_2 - x_4}{x_3 - x_4} = \frac{V'}{W'} \tag{10}$$

and for the second:

$$\frac{x_1 - x_4}{x_3 - x_4} \cdot \frac{x_3}{x_1} = \frac{U}{W} \quad \frac{x_2 - x_4}{x_3 - x_4} \cdot \frac{x_3}{x_2} = \frac{V}{W} \tag{11}$$

The first two equations (10) determine x_1 and x_2 as functions of x_3 by replacing in equations (11):

$$x_1 = \frac{U'W}{W'U} x_3 \quad x_2 = \frac{V'W}{W'V} x_3$$

replacing these values for x_1 and x_2 in equations (10), we obtain a system of two linear equations in two unknowns x_3 and x_4:

$$\begin{cases} x_3U'(W-U)+x_4U(U'-W')=0 \\ x_3V'(W-V)+x_4V(V'-W')=0 \end{cases}$$

Because of equation (12) of appendix A, the discriminant of these equations is equal to 0 and the two equations reduce to one which yields x_4 as a function of x_3:

$$x_4 = \frac{U'(W-U)}{U(W'-U')} \quad x_3 = \frac{V'(W-V)}{V(W'-V')}x_3$$

We can therefore express matrixes $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$ as very simple functions of the four projective parameters p, q, r, s:

$$\tilde{\mathbf{P}}' = \begin{bmatrix} \frac{U'W}{W'U}p & 0 & 0 & \frac{U'(W-U)}{U(W'-U')}s \\ 0 & \frac{V'W}{W'V}q & 0 & \frac{U'(W-U)}{U(W'-U')}s \\ 0 & 0 & r & \frac{U'(W-U)}{U(W'-U')}s \end{bmatrix}$$

There is a detail that changes the form of the matrixes $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$ which is the following. We considered the four points A_i, $i=1,\cdots,4$ as forming an *affine* basis of the space. Therefore, if we want to consider that the last coordinates of points determine the plane at infinity we should take the coordinates of those points to have a 1 as the last coordinate instead of a 0. It can be seen that this is the same as multiplying matrixes $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$ on the right by the matrix

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Similarly, the vectors representing the points of $\mathcal{P}^3$ must be multiplied by $\mathbf{Q}^{-1}$. For example

$$\mathbf{C} = [\frac{1}{p}, \frac{1}{q}, \frac{1}{r}, \frac{1}{p}+\frac{1}{q}+\frac{1}{r}-\frac{1}{s}]^T$$

We have thus obtained another remarkably simple result:

In the case where at least eight point correspondences have been obtained between two images of an uncalibrated stereo rig, if we arbitrarily choose four of these correspondences and consider that they are the images of four points in general positions (i.e not coplanar), then it is possible to reconstruct the other four points and any other point arising from a correspondence between the two images in the affine coordinate system defined by the four points. This reconstruction is uniquely defined up to an unknown affine transformation of the environment.

The main difference with the previous case is that instead of having a unique determination of the two perspective projection matrixes $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$, we have a family of such matrixes parameterized by the point of $\mathcal{P}^3$ of projective coordinates p, q, r, s. Some simple parameter counting will explain why. The stereo rig depends upon $22 = 2 \times 11$ parameters, 11 for each perspective projection matrix. The reconstruction is defined up to an affine transformation, that is $12 = 9 + 3$ parameters, the knowledge of the two epipoles and the epipolar transformation represents $7 = 2 + 2 + 3$ parameters. Therefore we are left with $22 - 12 - 7 = 3$ loose parameters which are the p, q, r, s.

Similarly, in the previous projective case, the reconstruction is defined up to a projective transformation, that is 15 parameters. The knowledge of the epipolar geometry still provides 7 parameters which makes a total of 22. Thus our result that the perspective projection matrixes are uniquely defined in that case.

4.1 Reconstructing the points

Given a pair (m, m') of matched pixels, we want to compute now the coordinates of the reconstructed three-dimensional point M (in the affine coordinate system defined by the four points A_i, $i = 1, \cdots, 4$). Those coordinates will be functions of the parameters p, q, r, s.

The computation is extremely simple and analogous to the one performed in the previous projective case. We write again that the reconstructed point M is expressed as

$$\mathbf{M} = \mu \mathbf{C} + \lambda \begin{bmatrix} \mathbf{M}_\infty \\ 0 \end{bmatrix}$$

The scalars λ and μ are determined as in section 3.2. Let $u_{12} = \frac{U'W}{W'U}$ and $v_{12} = \frac{V'W}{W'V}$. We have

$$\mathbf{P}'\mathbf{P}^{-1} = \begin{bmatrix} u_{12} & 0 & 0 \\ 0 & v_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

λ and μ are given by equation 7 in which matrix

$$\mathbf{A} = \begin{bmatrix} U' & u_{12}X \\ V' & v_{12}Y \\ W' & Z \end{bmatrix}$$

is in general of rank 2. The projective coordinates of the reconstructed point M are then:

$$\mathbf{M} = \mathbf{Q}^{-1}(\mu[\frac{1}{p}, \frac{1}{q}, \frac{1}{r}, -\frac{1}{s}]^T + \lambda[\frac{X}{p}, \frac{Y}{q}, \frac{Z}{r}, 0]^T)$$

4.2 Choosing the parameters p, q, r, s

The parameters p, q, r, s can be chosen arbitrarily. Suppose we reconstruct the same scene with two different sets of parameters p_1, q_1, r_1, s_1 and p_2, q_2, r_2, s_2. Then the relationship between the coordinates of a point M_1 and a point M_2 reconstructed with those two sets from the same image correspondence (m, m') is very simple in projective coordinates:

$$\mathbf{M}_2 = \mathbf{Q}^{-1} \begin{bmatrix} \frac{p_2}{p_1} & 0 & 0 & 0 \\ 0 & \frac{q_2}{q_1} & 0 & 0 \\ 0 & 0 & \frac{r_2}{r_1} & 0 \\ 0 & 0 & 0 & \frac{s_2}{s_1} \end{bmatrix} \mathbf{Q}\mathbf{M}_1 = \begin{bmatrix} \frac{p_2}{p_1} & 0 & 0 & 0 \\ 0 & \frac{q_2}{q_1} & 0 & 0 \\ 0 & 0 & \frac{r_2}{r_1} & 0 \\ \frac{p_2}{p_1} - \frac{s_2}{s_1} & \frac{q_2}{q_1} - \frac{s_2}{s_1} & \frac{r_2}{r_1} - \frac{s_2}{s_1} & \frac{s_2}{s_1} \end{bmatrix} \mathbf{M}_1$$

The two scenes are therefore related by a projective transformation. It may come as a surprise that they are not related by an affine transformation but it is clearly the case that the above transformation preserves the plane at infinity if and only if

$$\frac{p_2}{p_1} = \frac{q_2}{q_1} = \frac{r_2}{r_1} = \frac{s_2}{s_1}$$

If we have more information about the stereo rig, for example if we know that the two optical axis are coplanar, or parallel, then we can reduce the number of free parameters. We have not yet explored experimentally the influence of this choice of parameters on the reconstructed scene and plan to do it in the future.

5 Putting together different viewpoints

An interesting question is whether this approach precludes the building of composite models of a scene by putting together different local models. We and others have been doing this quite successfully over the years in the case of metric reconstructions of the scene [1,11]. Does the loss of the metric information imply that this is not possible anymore? fortunately, the answer to this question is no, we can still do it but in the weaker frameworks we have been dealing with, namely projective and affine reconstructions.

To see this, let us take the case of a scene which has been reconstructed by the affine or projective method from two different viewpoints with a stereo rig. We do not need to assume that it is the *same* stereo rig in both cases, i.e we can have changed the intrinsic and extrinsic parameters between the two views (for example changed the base line and the focal lengths). Note that we do not require the knowledge of these changes.

Suppose then that we have reconstructed a scene S_1 from the first viewpoint using the five points A_i, $i = 1, \ldots, 5$ as the standard projective basis. We know that our reconstruction can be obtained from the *real* scene by applying to it the (unknown) projective transformation that turns the four points A_i which have perfectly well defined coordinates in a coordinate system attached to the environment into the standard projective basis. We could determine these coordinates by going out there with a ruler and measuring distances, but precisely we want to avoid doing this. Let T_1 denote this collineation of $\mathcal{P}^3$.

Similarly, from the second viewpoint, we have reconstructed a scene S_2 using five *other* points B_i, $i = 1, \ldots, 5$ as the standard projective basis. Again, this reconstruction can be obtained from the real scene by applying to it the (unknown) projective transformation that turns the four points B_i into the standard projective basis. Let T_2 denote the corresponding collineation of $\mathcal{P}^3$. Since the collineations of $\mathcal{P}^3$ form a group, S_2 is related to S_1 by the collineation $T_2 T_1^{-1}$. This means that the two reconstructions are related by an *unknown* projective transformation.

Similarly, in the case we have studied before, the scenes were related by an unknown rigid displacement [1,11]. The method we have developed for this case worked in three steps:

1. Look for potential matches between the two reconstructed scenes. These matches are sets of reconstructed tokens (mostly points and lines in the cases we have studied) which can be hypothesized as being reconstructions of the *same* physical tokens because they have the same metric invariants (distances and angles). An example is a set of two lines with the same shortest distance and forming the same angle.
2. Using these groups of tokens with the same metric invariants, look for a *global* rigid displacement from the first scene to the second that maximizes the number of matched tokens.
3. For those tokens which have found a match, *fuse* their geometric representations using the estimated rigid displacement and measures of uncertainty.

The present situation is quite similar if we change the words *metric invariants* into *projective invariants* and *rigid displacement* into *projective transformation*. There is a difference which is due to the fact that the projective group is larger than the euclidean group. the first one depends on 15 independent parameters whereas the second depends upon only 6 (three for rotation and three for translation). This means that we will have to consider larger sets of tokens in order to obtain invariants. For example two lines depend upon 8 parameters in euclidean or projective space, therefore we obtain $8 - 6 = 2$ metric invariants (the shortest distance and the angle previously mentioned) but no projective invariants. In order to obtain some projective invariants, we need to consider sets of four lines for which there is at least one invariant (16-15=1)[2]. Even though this is not theoretically significant, it has obvious consequences on the complexity of the algorithms for finding matches between the two scenes (we go from an $o(n^2)$ complexity to an $o(n^4)$, where n is the number of lines).

We can also consider points, or mixtures of points and lines, or for that matter any combination of geometric entities but this is outside the scope of this paper and we will report on these subjects later.

The affine case can be treated similarly. Remember from section 4 that we choose four arbitrary noncoplanar points A_i, $i = 1, \cdots, 4$ as the standard affine basis and reconstruct the scene locally to these points. The reconstructed scene is related to the real one by a three-parameter family of affine transformations. When we have two reconstructions obtained from two different viewpoints, they are both obtained from the real scene by applying to it two unknown *affine* transformations. These two transformations depend each upon three arbitrary parameters, but they remain affine. This means that the relationship between the two reconstructed scenes is an unknown *affine* transformation[3] and that everything we said about the projective case can be also said in this case, changing projective into affine. In particular, this means that we are working with a smaller group which depends only upon 12 parameters and that the complexity of the matching should be intermediate between the metric and projective cases.

6 Experimental results

This theory has been implemented in Maple and C code. We show the results on the calibration pattern of figure 2. We have been using this pattern over the years to calibrate our stereo rigs and it is fair enough to use it to demonstrate that we will not need it anymore in the forthcoming years.

The pattern is made of two perpendicular planes on which we have painted with great care black and white squares. The two planes define a natural euclidean coordinate frame in which we know quite accurately the coordinates of the vertexes of the squares. The images of these squares are processed to extract the images of these vertexes whose pixel coordinates are then also known accurately. The three sets of coordinates, one set in three dimensions and two sets in two dimensions, one for each image of the stereo rig, are then used to estimate the perspective matrixes $\tilde{\mathbf{P}}_1$ and $\tilde{\mathbf{P}}_2$ from which we can compute the intrinsic parameters of each camera as well as the relative displacement of each of them with respect to the euclidean coordinate system defined by the calibration pattern.

We have used as input to our program the pixel coordinates of the vertexes of the images of the squares as well as the pairs of corresponding points[4]. From these we can estimate the epipolar geometry and perform the kind of local reconstruction which has

[2] In fact there are two which are obtained as follows: given the family of all lines, if we impose that this line intersects a given line, this is one condition, therefore there is in general a finite number of lines which intersect four given lines. This number is in general two and the two invariants are the cross-ratios of the two sets of four points of intersection.

[3] This is true only, according to section 4.2, if the two reconstructions have been performed using the same parameters p, q, r and s.

[4] In practice, these matches are obtained automatically by a program developed by Régis Vaillant which uses some a priori knowledge about the calibration pattern.

been described in this paper. Since it is hard to visualize things in a projective space, we have corrected our reconstruction before displaying it in the following manner.

We have chosen A_1, A_2, A_3 in the first of the two planes, A_4, A_5 in the second, and checked that no four of them were coplanar. We then have reconstructed all the vertexes in the projective frame defined by the five points A_i, $i = 1, \cdots, 5$. We know that this reconstruction is related to the real calibration pattern by the projective transformation that transforms the five points (as defined by their known projective coordinates in the euclidean coordinate system defined by the pattern, just add a 1 as the last coordinate) into the standard projective basis. Since in this case this transformation is known to us by construction, we can use it to test the validity of our projective reconstruction and in particular its sensitivity to noise. In order to do this we simply apply the inverse transformation to all our reconstructed points obtaining their "corrected" coordinates in euclidean space. We can then visualize them using standard display tools and in particular look at them from various viewpoints to check their geometry. This is shown in figure 3 where it can be seen that the quality of the reconstruction is quite good.

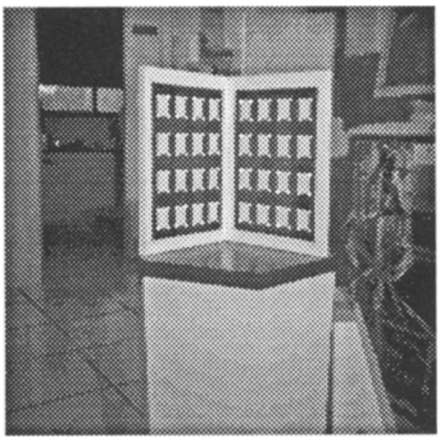

Fig. 2. A grey scale image of the calibration pattern

7 Conclusion

This paper opens the door to quite exciting research. The results we have presented indicate that computer vision may have been slightly overdoing it in trying at all costs to obtain metric information from images. Indeed, our past experience with the computation of such information has shown us that it is difficult to obtain, requiring awkward calibration procedures and special purpose patterns which are difficult if not impossible to use in natural environments with active vision systems. In fact it is not often the case that accurate metric information is necessary for robotics applications for example where *relative* information is usually all what is needed.

In order to make this local reconstruction theory practical, we need to investigate in more detail how the epipolar geometry can be automatically recovered from the environment and how sensitive the results are to errors in this estimation. We have started doing

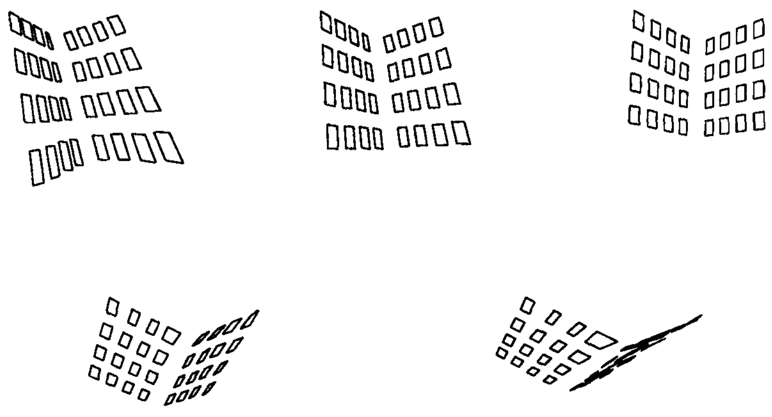

Fig. 3. Several rotated views of the "corrected" reconstructed points (see text)

this and some results are reported in a companion paper [2]. We also need to investigate the sensitivity to errors of the affine and projective invariants which are necessary in order to establish correspondences between local reconstructions obtained from various viewpoints.

Acknowledgements:
I want to thank Théo Papadopoulo and Luc Robert for their thoughtful comments on an early version of this paper as well as for trying to keep me up to date on their latest software packages without which I would never have been able to finish this paper on time.

A Computing some cross-ratios

Let U, V, W de the projective coordinates of the epipole o. The projective representations of the lines $\langle o, a_1\rangle$, $\langle o, a_2\rangle$, $\langle o, a_3\rangle$, $\langle o, a_4\rangle$ are the cross-products $\mathbf{o} \wedge \mathbf{a}_1 \equiv \mathbf{l}_1$, $\mathbf{o} \wedge \mathbf{a}_2 \equiv \mathbf{l}_2$, $\mathbf{o} \wedge \mathbf{a}_3 \equiv \mathbf{l}_3$, $\mathbf{o} \wedge \mathbf{a}_4 \equiv \mathbf{l}_4$.

A simple algebraic computation shows that

$$\mathbf{l}_1 = [0,\, W, -V]^T \qquad \mathbf{l}_2 = [-W,\, 0,\, U]^T$$
$$\mathbf{l}_3 = [V,\, -U,\, 0]^T \; \mathbf{l}_4 = [V - W,\, W - U,\, U - V]^T$$

This shows that, projectively (if $W \neq 0$):

$$\mathbf{l}_3 = U\mathbf{l}_1 + V\mathbf{l}_2 \quad \mathbf{l}_4 = (W - U)\mathbf{l}_1 + (W - V)\mathbf{l}_2$$

The cross-ratio of the four lines is equal to cross-ratio of the four "points" l_1, l_2, l_3, l_4:

$$\{\langle o, a_1\rangle, \langle o, a_2\rangle, \langle o, a_3\rangle, \langle o, a_4\rangle\} = \{0, \infty, \frac{V}{U}, \frac{W - V}{W - U}\} = \frac{V(W - U)}{U(W - V)}$$

Therefore, the projective coordinates of the two epipoles satisfy the first relation:

$$\frac{V(W-U)}{U(W-V)} = \frac{V'(W'-U')}{U'(W'-V')} \tag{12}$$

In order to compute the second pair of cross-ratios, we have to introduce the fifth line $\langle o, a_5 \rangle$, compute its projective representation $\mathbf{l}_5 = \mathbf{o} \wedge \mathbf{a}_5$, and express it as a linear combination of $\mathbf{l}_1$ and $\mathbf{l}_2$. It comes that:

$$\mathbf{l}_5 = (U\gamma - W\alpha)\mathbf{l}_1 + (V\gamma - W\beta)\mathbf{l}_2$$

From which it follows that the second cross-ratio is equal to:

$$\{\langle o, a_1 \rangle, \langle o, a_2 \rangle, \langle o, a_3 \rangle, \langle o, a_5 \rangle\} = \{0, \infty, \tfrac{V}{U}, \tfrac{V\gamma - W\beta}{U\gamma - W\alpha}\} = \frac{V(V\gamma - W\beta)}{U(U\gamma - W\alpha)}$$

Therefore, the projective coordinates of the two epipoles satisfy the second relation:

$$\frac{V(V\gamma - W\beta)}{U(U\gamma - W\alpha)} = \frac{V'(V'\gamma' - W'\beta')}{U'(U'\gamma' - W'\alpha')} \tag{13}$$

B The essential matrix

We relate here the essential matrix $\mathbf{F}$ to the two perspective projection matrixes $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$. Denoting as in the main text by $\mathbf{P}$ and $\mathbf{P}'$ the 3×3 left sub-matrixes of $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$, and by $\mathbf{p}$ and $\mathbf{p}'$ the left 3×1 vectors of these matrixes, we write them as:

$$\tilde{\mathbf{P}} = [\mathbf{P}\ \mathbf{p}] \quad \tilde{\mathbf{P}}' = [\mathbf{P}'\ \mathbf{p}']$$

Knowing this, we can write:

$$\mathbf{C} = \begin{bmatrix} \mathbf{P}^{-1}\mathbf{p} \\ -1 \end{bmatrix} \quad \mathbf{M}_\infty = \mathbf{P}^{-1}\mathbf{m}$$

and we obtain the coordinates of the epipole o' and of the image m'_∞ of M_∞ in the second retina:

$$\mathbf{o}' = \tilde{\mathbf{P}}'\mathbf{C} = \mathbf{P}'\mathbf{P}^{-1}\mathbf{p} - \mathbf{p}' \quad \mathbf{m}'_\infty = \mathbf{P}'\mathbf{P}^{-1}\mathbf{m}$$

The two points o' and m'_∞ define the epipolar line o_m of m, therefore the projective representation of o_m is the cross-product of the projective representations of o' and m'_∞:

$$\mathbf{o}_m = \mathbf{o}' \wedge \mathbf{m}'_\infty = \tilde{\mathbf{o}}'\mathbf{m}'_\infty$$

where we use the notation $\tilde{\mathbf{o}}'$ to denote the 3×3 antisymmetric matrix representing the cross-product with the vector $\mathbf{o}'$.

From what we have seen before, we write:

$$\mathbf{P}'\mathbf{P}^{-1} = \begin{bmatrix} \frac{\alpha' x' - 1}{\alpha x - 1} & 0 & 0 \\ 0 & \frac{\beta' x' - 1}{\beta x - 1} & 0 \\ 0 & 0 & \frac{\gamma' x' - 1}{\gamma x - 1} \end{bmatrix}$$

Thus:

$$\mathbf{P}'\mathbf{P}^{-1}\mathbf{p} - \mathbf{p}' = \begin{bmatrix} \frac{\alpha' x' - 1}{\alpha x - 1} - 1 \\ \frac{\beta' x' - 1}{\beta x - 1} - 1 \\ \frac{\gamma' x' - 1}{\gamma x - 1} - 1 \end{bmatrix} \cong \begin{bmatrix} \frac{\alpha' x' - \alpha x}{\alpha x - 1} \\ \frac{\beta' x' - \beta x}{\beta x - 1} \\ \frac{\gamma' x' - \gamma x}{\gamma x - 1} \end{bmatrix}$$

We can therefore write:

$$\tilde{o}' = \begin{bmatrix} 0 & -\frac{\gamma' x' - \gamma x}{\gamma x - 1} & \frac{\beta' x' - \beta x}{\beta x - 1} \\ \frac{\gamma' x' - \gamma x}{\gamma x - 1} & 0 & -\frac{\alpha' x' - \alpha x}{\alpha x - 1} \\ -\frac{\beta' x' - \beta x}{\beta x - 1} & \frac{\alpha' x' - \alpha x}{\alpha x - 1} & 0 \end{bmatrix}$$

and finally:

$$\mathbf{F} = \begin{bmatrix} 0 & -\frac{\beta' x' - 1}{\beta x - 1} \cdot \frac{\gamma' x' - \gamma x}{\gamma x - 1} & \frac{\gamma' x' - 1}{\gamma x - 1} \cdot \frac{\beta' x' - \beta x}{\beta x - 1} \\ \frac{\alpha' x' - 1}{\alpha x - 1} \cdot \frac{\gamma' x' - \gamma x}{\gamma x - 1} & 0 & -\frac{\gamma' x' - 1}{\gamma x - 1} \cdot \frac{\alpha' x' - \alpha x}{\alpha x - 1} \\ -\frac{\alpha' x' - 1}{\alpha x - 1} \cdot \frac{\beta' x' - \beta x}{\beta x - 1} & \frac{\beta' x' - 1}{\beta x - 1} \cdot \frac{\alpha' x' - \alpha x}{\alpha x - 1} & 0 \end{bmatrix}$$

References

1. Nicholas Ayache and Olivier D. Faugeras. Maintaining Representations of the Environment of a Mobile Robot. *IEEE transactions on Robotics and Automation*, 5(6):804–819, December 1989. also INRIA report 789.
2. Olivier D. Faugeras, Tuan Luong, and Steven Maybank. Camera self-calibration: theory and experiments. In *Proceedings of the 2nd European Conference on Computer Vision*, 1992. Accepted.
3. Olivier D. Faugeras and Steven Maybank. Motion from point matches: multiplicity of solutions. *The International Journal of Computer Vision*, 4(3):225–246, June 1990. also INRIA Tech. Report 1157.
4. Koenderink Jan J. and Andrea J. van Doorn. Affine Structure from Motion. *Journal of the Optical Society of America*, 1992. To appear.
5. H.C. Longuet-Higgins. A Computer Algorithm for Reconstructing a Scene from Two Projections. *Nature*, 293:133–135, 1981.
6. R. Mohr and E. Arbogast. It can be done without camera calibration. *Pattern Recognition Letters*, 12:39–43, 1990.
7. 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.
8. J.G. Semple and G.T. Kneebone. *Algebraic Projective Geometry*. Oxford: Clarendon Press, 1952. Reprinted 1979.
9. Gunnar Sparr. An algebraic-analytic method for reconstruction from image correspondences. In *Proceedings 7th Scandinavian Conference on Image Analysis*, pages 274–281, 1991.
10. Gunnar Sparr. Projective invariants for affine shapes of points configurations. In J.L. Mundy and A. Zisserman, editors, *Proceedings of DARPA-ESPRIT Workshop on Applications of Invariance in Computer Vision*, pages 151–169, 1991.
11. Z. Zhang and O.D. Faugeras. A 3D world model builder with a mobile robot. *International Journal of Robotics Research*, 1992. To appear.

This article was processed using the LTEX macro package with ECCV92 style