Invariants of Six Points and Projective Reconstruction From Three Uncalibrated Images

Long Quan

***Abstract*— There are three projective invariants of a set of six points in general position in space. It is well known that these invariants cannot be recovered from one image, however an invariant relationship does exist between space invariants and image invariants. This invariant relationship is first derived for a single image. Then this invariant relationship is used to derive the space invariants, when multiple images are available.**

This paper establishes that the minimum number of images for computing these invariants is three, and the computation of invariants of six points from three images can have as many as three solutions. Algorithms are presented for computing these invariants in closed form.

The accuracy and stability with respect to image noise, selection of the triplets of images and distance between viewing positions are studied both through real and simulated images.

Applications of these invariants are also presented. Both the results of Faugeras [1] and Hartley *et al.* [2] for projective reconstruction and Sturm's method [3] for epipolar geometry determination from two uncalibrated images with at least seven points are extended to the case of three uncalibrated images with only six points.

***Index Items*—This invariant, projective reconstruction, epipolar geometry, uncalibrated images, projective geometry, self-calibration.**

I. Introduction

GEOMETRIC invariants are playing a more and more important role in machine vision applications. A lot of work for recognition and shape description using invariants has already been reported, for instance *cf.* the collection book [4]and [5], [6], [1], [7], [8], [9], [10], [2], [11], [12]. Most of the invariants [13], [14] are derived for planar objects using geometric entities such as points, lines and conics, since in this case, there exists a plane projective transformation between object and image space. Plane projective geometry provides an ideal mathematical tool for describing this. As for general geometric configurations in space, it has been shown that it is not possible to estimate invariants from a single image [15], except for some constraint geometric configurations [16]. Therefore, one (*cf.* [17], [1], [2], [11], [18]) basically deals with space projective invariants from two images, provided that the epipolar geometry, or the fundamental matrix [19], [20] of the two images is determined *a priori.*

Manuscript received November 8, 1993; revised July 19,1994. Recommended for acceptance by Dr. Daniel P. Huttenlocher.

This work is partly supported by European Esprit BRA projects Viva which is gratefully acknowledged. Recommended for acceptance by Dr. Daniel P. Huttenlocher. The author is with LIFIA-CNRS-INRIA, 46, Av. Felix Viallet, 38031 Grenoble, France.

IEEE Log Number P95009.

This paper is concerned with the computation of the invariants of sets of six points in space from three images taken with uncalibrated cameras, assuming that the correspondences between image points are known. The main new results obtained in this paper are that the minimum number of images for computing the invariants of six points is three; and the computation of invariants from three images can have as many as three solutions. All solutions are given in closed form. As a consequence of these results, both Sturm's [3] method for epipolar geometry determination and projective reconstruction of Faugeras [1] and Hartley *et al.* [2] developed for two uncalibrated images is extended to the case of three uncalibrated images. Interestingly, in both two-camera and three-camera cases, the maximum number of solutions is three from the minimal data, i.e. six points for three images and seven points for two images. Also, with the help of such results, it may be established that the method of camera self-calibration proposed by Maybank and Faugeras [21] can be achieved with only six points of three images taken by the camera. Part of this work was also presented in [22].

Barrett *et al.* [17] have considered invariants from multiple images in order to solve the "transfer problem" in photogrammetry. Faugeras [1] and Hartley *et al.* [2] have shown that a set of points in 3D can be reconstructed up to a collineation from the point correspondences of two images taken with uncalibrated cameras. Similar results are presented by Mohr *et al.* in [6] for general multiple uncalibrated images using numerical minimization techniques. This permits the computation of projective invariants for sets of points visible from two or more images. In [6], [11], it is mentioned that it might be possible to get invariants from three images, however no explicit solutions are obtained. For other sets of geometric entities, invariants are obtained from two images and some experimental results are presented in [9], [11],[18]. However all these approaches using only two images are essentially based on the *a priori* determination of the epipolar geometry of the two images. The epipolar geometry may be algebraically determined, up to three solutions, with a minimum of seven points by Sturm's method [3],[19] or other equivalent algebraic methods based on the matrix representation of the epipolar geometry [23], [19]. However, Sturm's method is numerically unstable [20]. When more than eight points are available, numerical minimization methods are used to determine it. Therefore, to compute the invariants of six points, the correspondences of more than six points are needed for two images.

In comparison with related work, the method proposed in this paper does not need to first estimate the epipolar geometry between cameras, only the correspondences of six points are necessary, however a third image is needed to get algebraic solutions.

The development of this work is largely inspired by the old mathematical work on invariants of Coble [24] and the more recent work of Faugeras [1]. We assume that readers are familiar with elementary projective geometry and invariant theory, which can be found in [25], [24], [26], [27].

The paper is organized as follows. In Section II, the invariants of the set of six points in space are briefly reviewed. Then in Section III we will show that although invariants of a general set of points can not be obtained from one image, there exists a simple invariant relationship between the invariants of the set of six points in space and the invariants of their projected points in an image. Later in Section IV, this invariant relationship for one image will be applied to three images to compute invariants in space. Applications of the computation of invariants for projective reconstruction, epipolar geometry determination and self-calibration are discussed in Section V. Experimental results for the stability of these invariants and projective reconstruction are presented in Section VI. Future directions are discussed in Section VII.

II. Review of Invariants of Six Points in Space

Given a set of points or other kind of geometric configurations, the number of invariants is, roughly speaking, the difference between the dimension of the configuration and the dimension of the transformation group that acts on the configuration, if the dimension of the isotropy group of the configuration is null (*cf.* [9], [4]). For a set of six points in $\mathcal{P}^3$, there are $3 \times 6 - (16 - 1) = 3$ absolute invariants under the action of general linear group $GL(3)$ in $\mathcal{P}^3$. Since each point in $\mathcal{P}^3$ has three degrees of freedom, the dimension of a set of six points is $3 \times 6 = 18$. The transformation group $GL(3)$ is represented by a 4×4 matrix up to a scaling factor, its dimension is $4 \times 4 - 1 = 15$.

These three invariants can be formed and interpreted differently. For instance, following invariant theory, invariants can be expressed by linear combinations of products of the determinants of 4×4 matrix whose columns are the homogeneous coordinates of the points. This is mainly a domain of symmetric functions of the coordinates of the points. The invariants formed this way can be symmetric, therefore independent of the order in which the points are taken. Sometimes, half symmetric functions are also used. Coble [24] was interested in studying the invariants of associated sets of points between spaces of different dimensions. Using the six half symmetrical Jourbet's functions, he investigated the complete system of the (relative) invariants of six points of different dimension. We will not go further into his work, as it is rather complicated and involves concepts beyond the scope of this paper.

Another way is to consider twisted cubics [25], since there is a unique twisted cubic that passes through the given six general points in $\mathcal{P}^3$. Any four points of a twisted cubic define a cross ratio, therefore a subordinate one-dimensional projective geometry is induced on any twisted cubic. So the three invariants of the set of six points can be taken as the three independent cross ratios of the sets of four points on the twisted cubic. Or more algebraically, we can consider the invariants of cubic forms. Cubic forms in $\mathcal{P}^3$ have five algebraically independent relative invariants of degree 8, 16, 24, 32, 40 (*cf.* [27]). Three absolute invariants can be derived from these five relative invariants. However the expressions for these invariants are very complicated polynomials.

One of the simplest ways to consider these three invariants is by considering the three non-homogeneous projective coordinates of any sixth point with respect to a projective basis defined by any five of them. In the following, we will consider these invariants, since it seems to us that this is the simplest way to deal with them, and also because this methodology has been successfully used in [7], [1], [28], [8], [18], particularly by Faugeras in [1]. Obviously, these non homogeneous projective coordinates admit direct cross-ratio interpretation [25].

III. Invariant Relationship of Six Points from One Image

A. Canonical representation of six points in space

Given any six points $\{P_i,\ i = 1,\ldots,6\}$ in $\mathcal{P}^3$, in view of the fundamental theorem for projective space, any five of them, no three of them collinear and no four of them coplanar, can be given preassigned coordinates, thus we can assign them the canonical projective coordinates as follows:

$$(1,0,0,0)^T,\ (0,1,0,0)^T,\ (0,0,1,0)^T,\ (0,0,0,1)^T \text{ and } (1,1,1,1)^T.$$

This uniquely determines a space collineation $A_{4\times4}$, $det(A_{4\times4}) \neq 0$, which transforms the original five points into this canonical basis. And for the sixth point, it is transformed into its projective coordinates $(X, Y, Z, T)^T$ by $A_{4\times4}$. Therefore,

$X : Y : Z : T$ gives the three independent absolute invariants of six points.

B. Canonical representation of six image points

The projections of these six points onto an image $\{p_i, i = 1, \ldots, 6\}$ are usually given in non-homogeneous coordinates

$$(x_i, y_i)^T,\ i = 1, \ldots, 6$$

take any four of them, no three of them collinear, and assign them the canonical projective coordinates in $\mathcal{P}^2$

$$(1, 0, 0)^T,\ (0, 1, 0)^T,\ (0, 0, 1)^T, \text{ and } (1, 1, 1)^T.$$

A plane collineation $A_{3\times3}$, $det(A_{3\times3}) \neq 0$, can be uniquely determined. $A_{3\times3}$ transforms the fifth and sixth points into $(u_5, v_5, w_5)^T$ and $(u_6, v_6, w_6)^T$. Therefore,

$u_5 : v_5 : w_5$ and $u_6 : v_6 : w_6$ give the four independent absolute invariants of six image points.

The algebraic determination of $A_{3\times3}$ and $A_{4\times4}$ can be found in the Appendix.

C. Projection between space and image plane

If we assume a perspective projection as the camera model, then object space may be considered as embedded in $\mathcal{P}^3$ and image space embedded in $\mathcal{P}^2$. The camera performs the projection from $\mathcal{P}^3$ upon $\mathcal{P}^2$, and this projection can be represented by a 3×4 matrix $C_{3\times 4}$ of rank three whose kernel is the projection center. The relation between the points P_i in $\mathcal{P}^3$ and p_i in $\mathcal{P}^2$ can be written as

$$\lambda_i p_i = C_{3\times 4} P_i,$$

where p_i and P_i are in homogeneous coordinates. I. e.,

$$\lambda_i \begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ T_i \end{pmatrix}.$$

This can be rewritten in ratio form hiding the scaling factor λ_i,

$$\begin{aligned} u_i : v_i : w_i \quad = \quad & (c_{11}X_i + c_{12}Y_i + c_{13}Z_i + c_{14}T_i) : \\ & (c_{21}X_i + c_{22}Y_i + c_{23}Z_i + c_{24}T_i) : \\ & (c_{31}X_i + c_{32}Y_i + c_{33}Z_i + c_{34}T_i) \end{aligned} \qquad (1)$$

For each point, as u_i, v_i and w_i can not all be zero, two independent equations can always be derived from (1). These equations express nothing else than the collinearity of the space points and their corresponding image points. This is a projective property which is preserved by any projective transformation.

The calibration process consists of the determination of all the parameters c_{ij} of the projection matrix $C_{3\times 4}$. As we are assuming that we are working with uncalibrated images, it is meant that $C_{3\times 4}$ is totally unknown.

D. Elimination of camera parameters c_{ij}

For a set of *six* points in image and in space, when the correspondences $p_i \leftrightarrow P_i$, for $i = 1, \ldots, 6$, are given as

$$\begin{array}{ccc} (1,0,0)^T & \leftrightarrow & (1,0,0,0)^T, \\ (0,1,0)^T & \leftrightarrow & (0,1,0,0)^T, \\ (0,0,1)^T & \leftrightarrow & (0,0,1,0)^T, \\ (1,1,1)^T & \leftrightarrow & (0,0,0,1)^T, \\ (u_5, v_5, w_5)^T & \leftrightarrow & (1,1,1,1)^T, \\ (u_6, v_6, w_6)^T & \leftrightarrow & (X, Y\ Z, T)^T, \end{array}$$

this leads to $12 = 2 \times 6$ equations from (1). All entries c_{ij} of $C_{3\times 4}$ are unknowns, as we assumed that the camera was uncalibrated. Since $C_{3\times 4}$ is defined up to a scaling factor, it counts for $3 \times 4 - 1 = 11$ unknowns. So there still remains one $(1 = 12 - 11)$ independent equation after eliminating all unknown camera parameters c_{ij}.

Substituting all canonical projective coordinates of the *six* points in image and in space into equations (1). Then eliminating c_{ij}, we obtain the following homogeneous equation between X, Y, Z, T and $\{(u_i, v_i, w_i)$, for $i = 5, 6\}$,

$$\begin{array}{llllll} w_6(u_5 - v_5)XY & + & v_6(w_5 - u_5)XZ & + & & \\ u_5(v_6 - w_6)XT & + & u_6(v_5 - w_5)YZ & + & & (2) \\ v_5(w_6 - u_6)YT & + & w_5(u_6 - v_6)ZT & = 0. & & \end{array}$$

E. Invariant interpretation of the equation

The above equation (2) will be arranged and interpreted as an invariant relationship between the relative invariants of $\mathcal{P}^3$ and those of $\mathcal{P}^2$ as follows.

If i_j and I_j denote respectively

$$\begin{aligned} i_1 &= w_6(u_5 - v_5), \\ i_2 &= v_6(w_5 - u_5), \\ i_3 &= u_5(v_6 - w_6), \\ i_4 &= u_6(v_5 - w_5), \\ i_5 &= v_5(w_6 - u_6), \\ i_6 &= w_5(u_6 - v_6), \end{aligned}$$

and

$$\begin{aligned} &I_1 = XY,\ I_2 = XZ,\ I_3 = XT, \\ &I_4 = YZ,\ I_5 = YT,\ I_6 = ZT, \end{aligned}$$

then i_j and I_j can be interpreted as respectively the relative invariants of six points in $\mathcal{P}^2$ (image) and those of six points in $\mathcal{P}^3$ (space).

$\{i_j, j = 1, \ldots, 6\}$ are defined up to a common multiplier, so they are only relative invariants of the six points of the image. The five ratios $i_1 : i_2 : i_3 : i_4 : i_5 : i_6$ are (absolute) projective invariants. Since (by the arguments of Section II) for the set of six points in $\mathcal{P}^2$, there are $4 = 2 \times 6 - 8$ independent projective invariants, therefore the relative invariants $\{i_j, j = 1, \ldots, 6\}$ are not independent and are subject to one $(1 = 5 - 4)$ additional constraint which is

$$i_1 + i_2 + i_3 + i_4 + i_5 + i_6 = 0.$$

This can be checked from the above definition of i_j.

$\{I_j, j = 1, \ldots, 6\}$ are relative invariants of $\mathcal{P}^3$, only the five ratios $I_1 : I_2 : I_3 : I_4 : I_5 : I_6$ are projective invariants in space. Since there are only three independent absolute invariants, I_j are subject to two additional constraints which are, by inspection of the definition of I_j,

$$\frac{I_1}{I_2} = \frac{I_5}{I_6} \text{ and } \frac{I_2}{I_3} = \frac{I_4}{I_5}.$$

An independent set of three absolute invariants in space could be taken to be

$$\alpha \equiv \frac{I_2}{I_6} = \frac{X}{T},\ \beta \equiv \frac{I_4}{I_6} = \frac{Y}{T} \text{ and } \gamma \equiv \frac{I_2}{I_3} = \frac{Z}{T}$$

The set of invariants $\{\alpha, \beta, \gamma\}$ is equivalent to $X : Y : Z : T$.

Then the invariant relation (2) is simply expressed as a bilinear homogeneous relation,

$$i_1 I_1 + i_2 I_2 + i_3 I_3 + i_4 I_4 + i_5 I_5 + i_6 I_6 = 0. \qquad (3)$$

This relationship is of course independent of any camera parameters, therefore it can be used for any uncalibrated images. This clearly shows that the 3D invariants can not be computed from one single image, since we have only one invariant relationship for three independent invariants. This invariant relationship can be used in two different ways:

- To verify whether a given set of points is present in a model base. That means I_j can be computed from the model base and i_j from the image, then the invariant relation can be used to check for the presence of the model in the image.
- To compute 3D invariants from more than one image, this leads to the following section.

IV. Computation of the Invariants of Six Points from Three Images

In this section, we will focus on how to find the absolute invariants of six points in space, i.e. $\{\alpha, \beta, \gamma\}$ or equivalently $X : Y : Z : T$ from more than one image. To do this, it is assumed that the six point correspondences through images have already been established.

So from a set of six point in one image, the homogeneous invariant relation (3) can be written in X, Y, Z, T as

$$i_1 XY + i_2 XZ + i_3 XT + i_4 YZ + i_5 YT + i_6 ZT = 0,$$

with coefficients i_j computed from image points. As there are only three independent invariants for a set of six points, we might hope to solve for these three invariants if three images of the set of six points are available. The three homogeneous quadratic equations in X, Y, Z, T from three images can be written as follows.

$$\mathcal{F}_1 \equiv i_1^{(1)} XY + i_2^{(1)} XZ + i_3^{(1)} XT + i_4^{(1)} YZ + i_5^{(1)} YT + i_6^{(1)} ZT = 0, \qquad (3)$$

$$\mathcal{F}_2 \equiv i_1^{(2)} XY + i_2^{(2)} XZ + i_3^{(2)} XT + i_4^{(2)} YZ + i_5^{(2)} YT + i_6^{(2)} ZT = 0 \qquad (5)$$

$$\mathcal{F}_3 \equiv i_1^{(3)} XY + i_2^{(3)} XZ + i_3^{(3)} XT + i_4^{(3)} YZ + i_5^{(3)} YT + i_6^{(3)} ZT = 0 \qquad (6)$$

here the superscripts (j), $j = 1, 2, 3$, distinguish the invariant quantities of different images.

A. Maximum number of possible solutions

Each equation represents a quadratic surface which has rank 3. These quadratic forms have no X^2, Y^2, Z^2, T^2 terms. That means that the quadratic surface goes through the vertices of the tetrahedron of reference whose coordinates are $(0, 0, 0, 1)^T$, $(0, 0, 1, 0)^T$, $(0, 1, 0, 0)^T$, $(1, 0, 0, 0)^T$. This is easily verified by substituting these points in the equations. In addition, as the coefficients of the quadratic form $i^{(j)}$ are subject to the following relations,

$$i_1^{(j)} + i_2^{(j)} + i_3^{(j)} + i_4^{(j)} + i_5^{(j)} + i_6^{(j)} = 0, \text{ for } j = 1, 2, 3,$$

so all the equations (4), (5), and (6) necessarily pass through the unit point $(1, 1, 1, 1)^T$.

According to Bezout's Theorem, three quadratic surfaces must meet in $8 = 2 \times 2 \times 2$ points. Since they already pass through the five known points, so only $3 = 8 - 5$ common points remain. Therefore,

the maximum number of solutions for $X : Y : Z : T$ is three.

B. Find X/T by solving a cubic equation

Now, let us try to explicit these three solutions. For instance, if we want to solve for X/T, the two resultants $\mathcal{G}_1$ and $\mathcal{G}_2$ obtained by eliminating Z between $\mathcal{F}_1$ and $\mathcal{F}_3$ and between $\mathcal{F}_2$ and $\mathcal{F}_3$ are homogeneous polynomials in X, Y and T of degree three:

$$\mathcal{G}_1 \equiv e_1^{(1)} X^2 Y + e_2^{(1)} XY^2 + e_3^{(1)} XYT + e_4^{(1)} X^2 T + e_5^{(1)} XT^2 + e_6^{(1)} Y^2 T + e_7^{(1)} YT^2 = 0, \qquad (7)$$

$$\mathcal{G}_2 \equiv e_1^{(2)} X^2 Y + e_2^{(2)} XY^2 + e_3^{(2)} XYT + e_4^{(2)} X^2 T + e_5^{(2)} XT^2 + e_6^{(2)} Y^2 T + e_7^{(2)} YT^2 = 0, \qquad (8)$$

whose coefficients are still subject to

$$e_1^{(j)} + e_2^{(j)} + e_3^{(j)} + e_4^{(j)} + e_5^{(j)} + e_6^{(j)} + e_7^{(j)} = 0 \text{ for } j = 1, 2,$$

where

$$\begin{aligned}
e_1^{(j)} &= i_4^{(j+1)} i_5^{(j)} - i_4^{(j)} i_5^{(j+1)},\\
e_2^{(j)} &= i_4^{(j+1)} i_6^{(j)} - i_4^{(j)} i_6^{(j+1)},\\
e_4^{(j)} &= i_1^{(j+1)} i_5^{(j)} - i_1^{(j)} i_5^{(j+1)},\\
e_5^{(j)} &= i_1^{(j+1)} i_3^{(j)} - i_1^{(j)} i_3^{(j+1)},\\
e_6^{(j)} &= i_2^{(j+1)} i_6^{(j)} - i_2^{(j)} i_6^{(j+1)},\\
e_7^{(j)} &= i_2^{(j+1)} i_3^{(j)} - i_2^{(j)} i_3^{(j+1)}.
\end{aligned}$$

This is due to the fact that the point $(1, 1, 1)^T$ is still on the two cubic curves. The resultant obtained by eliminating Y between $\mathcal{G}_1$ and $\mathcal{G}_2$ is a homogeneous polynomial in X, T of degree eight, which can be factorized as,

$$XT(X - T)(b_1X^2 + b_2XT + b_3T^2)\ (a_1X^3x + a_2X^2T + a_3XT^2 + a_4T^3).$$

It is evident that the linear factors lead to trivial solutions. The solution $X = 0$ corresponds to the common points $(0, 0, 0, 1)^T$, $(0, 0, 1, 0)^T$ and $(0, 1, 0, 0)^T$. The solution $T = 0$ corresponds to the common points $(1, 0, 0, 0)^T$, $(0, 1, 0, 0)^T$ and $(0, 0, 1, 0)^T$. The solution $X = T$ corresponds to the common point $(1, 1, 1, 1)^T$.

Now, let us have a close-look at the coefficients of the quadratic factor $b_1X^2 + b_2XT + b_3T^2$,

$$\begin{aligned}
b_1 &= i_4^{(2)} i_5^{(2)},\\
b_2 &= i_3^{(2)} i_4^{(2)} + i_2^{(2)} i_5^{(2)} - i_1^{(2)} i_6^{(2)},\\
b_3 &= i_2^{(2)} i_3^{(2)}
\end{aligned}$$

which depend only on the invariant quantities of the points of the second image, so the zeros of the quadratic factor are the parasite solutions introduced by elimination using resultants, they are not the zeros of $\{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3\}$. Thus, the only nontrivial solutions for X/T are those of the cubic equation,

$$\mathcal{C} \equiv a_1 X^3 + a_2 X^2T + a_3 XT^2 + a_4 T^3 = 0. \tag{9}$$

The implicit expressions for a_i (quite long) can be easily obtained with Maple. The cubic equation may be solved algebraically by Cardano's formula, either for X/T or T/X. The choice between solving X/T or T/X depends on the estimation of the moduli of the roots. According to Cauchy's Theorem, the moduli of the roots of cubic equations are bounded by the following,

$$\left(1 + \frac{max\{|a_2|, |a_4|\}}{|a_1|}\right)^{-1} < \left|\frac{X}{T}\right| < 1 + \frac{max\{|a_1|, |a_3|\}}{|a_4|}.$$

If the modulus is large enough, that means T is near to 0, in this case it would be better to solve for its reciprocal T/X instead of X/T.

C. Find $Y : Z : T$ linearly for a given X/T

$Y : Z : T$ should be solved for each given $X : T$. As the equations are symmetric in X, Y, Z and T, we could obtain other similar cubic equations for the other two independent absolute invariants, for instance, Y/T and Z/T or any other ratios if necessary. However, doing this would lead to much more than three solution sets which is not desirable. We expect to obtain unique $Y : Z : T$ for each given $X : T$.

By eliminating Y^2 between $\mathcal{G}_1$ and $\mathcal{G}_2$, we obtain the following homogeneous polynomial

$$\begin{aligned}
\mathcal{H} \equiv\ & (c_1X^3 + c_2X^2T + c_3XT^2 + c_4T^3)Y\\
& + X(d_1X^2 + d_2XT + d_3T^2)T,
\end{aligned} \tag{10}$$

where

$$\begin{aligned}
c_1 &= e_2^{(2)} e_1^{(1)} - e_1^{(2)} e_2^{(1)},\\
c_2 &= e_6^{(2)} e_1^{(1)} - e_3^{(2)} e_2^{(1)} + e_2^{(2)} e_3^{(1)} - e_1^{(2)} e_6^{(1)},\\
c_3 &= e_6^{(2)} e_3^{(1)} - e_7^{(2)} e_2^{(1)} - e_3^{(2)} e_6^{(1)} + e_2^{(2)} e_7^{(1)},\\
c_4 &= e_6^{(2)} e_7^{(1)} - e_7^{(2)} e_6^{(1)},
\end{aligned}$$

and

$$\begin{aligned}
d_1 &= e_2^{(2)} e_4^{(1)} - e_4^{(2)} e_2^{(1)},\\
d_2 &= e_2^{(2)} e_5^{(1)} + e_6^{(2)} e_4^{(1)} - e_5^{(2)} e_2^{(1)} - e_4^{(2)} e_6^{(1)},\\
d_3 &= -e_5^{(2)} e_6^{(1)} + e_6^{(2)} e_5^{(1)}.
\end{aligned}$$

Any zero of $\mathcal{H}$ is also a zero of $\mathcal{G}_1$ and $\mathcal{G}_2$, as $\mathcal{H}$ is a linear combination of $\mathcal{G}_1$ and $\mathcal{G}_2$ with polynomial coefficients, and $\mathcal{H}$ belongs to the ideal generated by $\mathcal{G}_1$ and $\mathcal{G}_2$.

$\mathcal{H}$ can be considered as a linear homogeneous polynomial in Y and T for a given X/T. Therefore a unique $Y : T$ is guaranteed from this equation for each given $\alpha = X : T$,

$$Y : T = -\alpha(d_1\alpha^2 + d_2\alpha + d_3) : (c_1\alpha^3 + c_2\alpha^2 + c_3\alpha + c_4).$$

For a given $X : Y : T$, say $\alpha = X/T$ and $\beta = Y/T$, $Z : T$ is uniquely determined by one of the equations (4), (5), or (6), for instance taking (4), we have

$$Z : T = -\left(i_1^{(1)}\alpha\beta + i_3^{(1)}\alpha + i_5^{(1)}\beta\right) : \left(i_2^{(1)}\alpha + i_4^{(1)}\beta + i_6^{(1)}\right).$$

In conclusion, only at most three solutions exist for $X : Y : Z : T$, hence we establish the following:

Given a set of six point correspondences in three images taken by uncalibrated cameras, there are at most three solutions for the three invariants associated to the set of six points in space, such solutions can be found in closed form.

D. Summary of computation

The computation of the invariants of six points can be summarized in the following simple algorithm.

1) Compute the plane collineation $A_{3\times 3}$ (*cf.* Appendix) such that 4 of the six points of each image go to the canonical projective basis. Apply this collineation to the remaining two points to obtain

$$\left\{u_i^{(j)},\ v_i^{(j)},\ w_i^{(j)},\ \text{for } i = 5,\ 6 \text{ and } j = 1,\ 2,\ 3\right\};$$

2) Compute the relative invariants of each image

$$\left\{i_k^{(j)},\ \text{for } j = 1,\ 2,\ 3 \text{ and } k = 1, \ldots 6\right\}$$

from the above transformed image points;

3) Solve the cubic equation (9) $\mathcal{C} = 0$ for $\alpha = X/T$;

4) For each α, solve the linear equation (10) $\mathcal{H} = 0$ for $\beta = Y/T$;

5) For each α and β, solve the linear equation (4) $\mathcal{F}_1 = 0$ for $\gamma = Z/T$.

E. Remark

We can note that instead of solving $\{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3\}$ for $X : Y : Z : T$, we can equivalently solve

$$\begin{cases} i_1^{(1)} I_1 + i_2^{(1)} I_2 + i_3^{(1)} I_3 + i_4^{(1)} I_4 + i_5^{(1)} I_5 + i_6^{(1)} I_6 = 0 \\ i_1^{(2)} I_1 + i_2^{(2)} I_2 + i_3^{(2)} I_3 + i_4^{(2)} I_4 + i_5^{(2)} I_5 + i_6^{(2)} I_6 = 0 \\ i_1^{(3)} I_1 + i_2^{(3)} I_2 + i_3^{(3)} I_3 + i_4^{(3)} I_4 + i_5^{(3)} I_5 + i_6^{(3)} I_6 = 0 \\ I_1 I_6 - I_2 I_5 = 0 \\ I_1 I_6 - I_3 I_4 = 0 \end{cases}$$

for the ratios $I_1 : I_2 : I_3 : I_4 : I_5 : I_6$.

They meet in $4 = 1 \times 1 \times 1 \times 2 \times 2$ intersection points in $\mathcal{P}^5$. As the sum of the coefficients is zero, they meet in the common unit point $(1, 1, 1, 1, 1, 1)^T$, so only three solutions remain. The same results, cubic equation, can be obtained by a little more symbolic computation.

We also note that when five images are available, a linear solution for the ratios $I_1 : I_2 : I_3 : I_4 : I_5 : I_6$ is possible while ignoring the quadratic constraints $I_1 I_6 = I_2 I_5 = I_3 I_4$.

V. Applications of the Invariants

In this section, we are exploring some interesting applications that may be provided by the computed invariants of six points. The most important is undoubtedly the complete determination of the projection matrices of the three cameras. This gives the most complete projectively invariant description of the three cameras. It follows that projective reconstruction as has been achieved by Faugeras [1] and Hartley *et al.* [2] and the determination of epipolar transformation by Sturm's method [3] are extended to the case of three uncalibrated cameras with fewer points.

A. Complete determination of projection matrices $C_{3\times4}^{(j)}$

In the case of two uncalibrated cameras, Faugeras [1] used epipolar geometry to determine the projection matrix up to projective transformations. Here, in the case of three uncalibrated images, as the projective coordinates of the sixth point is algebraically determined. It means that six points both in $\mathcal{P}^2$ and $\mathcal{P}^3$ are known up to respectively a plane and a space collineation. It turns out that the projection matrix

$$C_{3\times4}^{(j)}$$

for each camera can be completely determined.

For each camera, the correspondences of six points between $\mathcal{P}^2$ and $\mathcal{P}^3$ result in $2 \times 6 = 12$ linear homogeneous equations in the $3 \times 4 = 12$ entries c_{ij} that count only for 11 unknowns due to a common scaling factor. One of the 12 equations has been used to derive the invariant relationship (3) (*cf.* Section III-D), which was then used to compute the invariants. Therefore, there remain only $12 - 1 = 11$ independent linear equations which matches the 11 unknowns of the projection matrix. After some simple algebra, the projection matrix for the *j-th* camera can be written as follows,

$$C_{3\times4}^{(j)} = \begin{pmatrix} c_{11}^{(j)} & 0 & 0 & 1 \\ 0 & c_{22}^{(j)} & 0 & 1 \\ 0 & 0 & c_{33}^{(j)} & 1 \end{pmatrix}$$

where

$$c_{11}^{(j)} = u_5^{(j)} k^{(j)} - 1,$$
$$c_{22}^{(j)} = v_5^{(j)} k^{(j)} - 1,$$
$$c_{33}^{(j)} = w_5^{(j)} k^{(j)} - 1$$

and $k^{(j)}$ can be taken as either

$$\frac{\left(u_6^{(j)} - w_6^{(j)}\right) + w_6^{(j)}\alpha - u_6^{(j)}\gamma}{u_5^{(j)} w_6^{(j)}\alpha - u_6^{(j)} w_5^{(j)}\gamma},$$

or

$$\frac{\left(u_6^{(j)} - v_6^{(j)}\right) + v_6^{(j)}\alpha - u_6^{(j)}\beta}{u_5^{(j)} v_6^{(j)}\alpha - u_6^{(j)} v_5^{(j)}\beta},$$

which are equivalent due to the invariant relationship (3).

B. Projective reconstruction

The complete determination of projection matrices implies that given any point correspondence in images, the corresponding point in space can be reconstructed in $\mathcal{P}^3$ up to a collineation. In fact, projective reconstruction is equivalent to the computation of the projective invariants of the sets of points.

Given the point correspondences

$$\left(u_i^{(j)}, v_i^{(j)}, w_i^{(j)}\right)^T$$

and the projection matrices $C_{3\times4}^{(j)}$, for $j = 1, 2, 3$, we can write two linear homogeneous equations for the unknowns X_i, Y_i, Z_i T_i of each point for each image as follows,

$$\begin{cases} w_i^{(j)} c_{11}^{(j)} X_i - u_i^{(j)} c_{33}^{(j)} Z_i + \left(w_i^{(j)} - u_i^{(j)}\right) T_i = 0, \\ w_i^{(j)} c_{22}^{(j)} Y_i - v_i^{(j)} c_{33}^{(j)} Z_i + \left(w_i^{(j)} - u_i^{(j)}\right) T_i = 0. \end{cases}$$

With three images $j = 1, 2, 3$, a linear system of six homogeneous equations is obtained. An algebraic solution can be found by solving any three of them. A numerical solution can also be found by solving this over-determined linear system. Therefore we establish the following result which can be considered as an extension to the recent work of Faugeras [1] and Hartley *et al.* [2] for reconstruction from two uncalibrated images.

Given a set of $n \geq 6$ point correspondences in three images taken by uncalibrated cameras, the set of points can be reconstructed in $\mathcal{P}^3$ up to a collineation.

The reconstruction may have as many as three different solutions from the minimal data of six points.

Notice that this result reduces the number of points needed to have a projective reconstruction to the strict minimum of six.

C. Determination of the epipolar geometry

Another consequence of the complete determination of the projection matrices is that the epipolar geometry of any pair of the three cameras can be computed with only *six* point correspondences.

Given a set of six point correspondences in three images taken by uncalibrated cameras, the epipolar geometry of any pair of cameras can be determined, up to at most three solutions.

This can be compared with Sturm's method [29], [3], reintroduced into computer vision in [30], [21], [19], which is of great importance for two uncalibrated cameras. Based on projective geometry, Sturm's method yields at most three solutions in general for the epipolar geometry, which are three of the nine points of intersection of two cubic plane curves.

Note that although the two approaches developed respectively for the two-camera and the three-camera cases lead to equivalent results, the parametrizations used in two cases are different. Sturm directly used the epipolar transformation as the parameters for the two uncalibrated cameras, while we used the projective structure of the point set in space as parameters for the three uncalibrated cameras. The underlying equations of the parametrization for the three-camera case are simpler, hence a better numerical stability might be expected. This is to be confirmed by the experimental results presented in Section VI.

For extracting the epipolar geometry from any given pair of projection matrices, one can refer, for instance, to [1] for more details. Here we are content with the following results.

- The epipole in the *j-th* image with respect to the *i-th* camera e_{ji} is given by

$$e_{ji} = C^{(j)}_{3\times 4} ker\left(C^{(i)}_{3\times 4}\right) = \left(1-h_1,\ 1-h_2,\ 1-h_3\right)^T,$$

where

$$h_1 = \frac{c^{(j)}_{11}}{c^{(i)}_{11}},\ h_2 = \frac{c^{(j)}_{22}}{c^{(i)}_{22}},\ h_3 = \frac{c^{(j)}_{33}}{c^{(i)}_{33}}.$$

- The epipolar geometry between *i-th* and *j-th* cameras, represented by the fundamental matrix F_{ij} [19] (the generalization of the essential matrix [23]) is given by

$$F_{ji} = \left[e_{ji}\right] C^{(j)}_{3\times 3}\left(C^{(i)}_{3\times}\right)^{-1} = \begin{pmatrix} 0 & -h_2(h_3-1) & h_3(h_2-1) \\ h_1(h_3-1) & 0 & -h_3(h_1-1) \\ -h_1(h_2-1) & h_2(h_1-1) & 0 \end{pmatrix},$$

where $[e_{ji}]$ is the anti-symmetric matrix associated to e_{ji}, and $C_{3\times 3}$ is the left 3×3 sub-matrix of the 3×4 matrix $C_{3\times 4}$.

D. Camera self-calibration

The above results can also be applied to the camera self-calibration proposed by Maybank and Faugeras [21]. According to [21],[19],[20], from point correspondences of three uncalibrated images and the associated epipolar geometries, one can calibrate and therefore obtain a Euclidean reconstruction (only up to a global scaling factor), if we suppose that all three images are taken by the same uncalibrated camera.

Using Sturm's method or other equivalent algebraic methods for estimating the epipolar geometries for each pair of three cameras, one needs a minimum number of seven points to do that. While using the method presented above which considers the three cameras as a whole, the epipolar geometries of any pair of three cameras can be computed with only six points. Therefore, we can make the following statement.

A camera can be self-calibrated with a minimum of six point correspondences in three images.

VI. Experimental Results

The theoretical results presented above for the computation of the invariants of six points have been implemented. The accuracy and stability of the invariants with respect to various factors such as pixel errors, selection of the triplet of images and positions of the camera are studied both for simulated and real images. The computation should be accurate and stable in order for the invariants to be useful for vision applications. The three invariants used for experiments are $\alpha = X : T$, $\beta = Y : T$ and $\gamma = Z : T$. The application for projective reconstruction is also presented.

A. The experiment with simulated images

A.1. Simulation set-up

We will first simulate a number of images in which a set of six points in space is projected. The simulation is set up as follows.

- First, a real camera is calibrated (*cf.* [31], [6], [32]). This is done by placing a known object of about 50 cm^3 in front of the camera. The camera is then moved through nine different viewing positions, spaced at 10 degree intervals around the calibration object. The intervals between the first three views are slightly bigger than the second three views, and the second three views bigger than the third three ones. This creates nine calibration matrices which are subsequently used in the simulation.

- With these nine calibration matrices as our perfect cameras, we project several sets of six known points into these nine synthetic images.
- Finally the projected positions of the points in the images are perturbed by varying levels of noise of uniform distribution.
- As the positions of the six simulated points in space are known in advance, their invariants are computed by the method given in Section II. These invariants are compared with those calculated from the simulated image data.

In this way, the realism of simulation is preserved, and the image noise can be quantitatively controlled as well in order to observe its influence.

The results that will be presented in the following tables are obtained from the set of six points whose (Euclidean) coordinates are $(2.0,\ 0.0,\ 12.0)^T$ $(0.0,\ 6.0,\ 0.0)^T$ $(12.0, 0.0, 14.0)^T$ $(0.0,\ 6.0,\ 6.0)^T$ $(-1.5,\ 19.5,\ 0.0)^T$ and $(0.0,\ 12.0,\ 12.0)^T$ (the unit is in *cm*). The other sets of six points are also simulated with the same calibration matrices and they have similar numerical behavior and will not be presented in the tables.

A.2 Stability w.r.t. the selection of the triplets of images

Table I shows the effect of choosing different triplets of synthetic images on the stability of the three invariants. All images points are perturbed by a uniform noise of ± 1.5 pixels. In the case where the solution is unique, the computed value is followed by a * in the tables. When multiple solutions occur, the one that is the closest to the ground truth value is selected.

Another experimentation was carried out to test the stability with respect to the change of one of the three images. Four images, four, five, six, and seven are taken, then all triplets are selected among the four. Here pixel error is set to ± 1.5.

TABLE I

TABLE OF INVARIANTS COMPUTED FROM DIFFERENT TRIPLETS OF SYNTHETIC IMAGES. THE POINTS OF IMAGES ARE NOISED BY ± 1.5 PIXEL ERROR, THE INVARIANTS α, β AND γ ARE COMPUTED FOR EACH DIFFERENT IMAGE TRIPLET. d_i ARE THE DIFFERENCES BETWEEN EACH COMPUTED INVARIANT AND ITS MEAN VALUE. σ IS THE STANDARD DEVIATION AND $\sigma/\overline{m}$ THE RATIO OF THE STANDARD DEVIATION TO THE MEAN VALUE.

Image Triplet	α = 0.526421	d_1	β=1.880620	d_2	γ = 0.745762	d_3
1, 2, 3	0.528183	0.0083	1.877117	-0.018	0.760771	0.0050
4, 5, 6	0.516476*	-0.0034	1.899056*	0.0035	0.747555*	-0.0082
1, 3, 5	0.524387	0.0045	1.889926	-0.0056	0.760802*	0.0051
5, 7, 9	0.514523*	-0.0053	1.899576*	0.0041	0.739840*	-0.016
1, 4, 8	0.520458	-0.00053	1.900477	0.0050	0.763048	0.0073
2, 5, 8	0.516762	-0.0032	1.900178	0.0047	0.751696	-0.0040
1, 5, 9	0.518681	-0.0012	1.902235	0.0067	0.766406	0.011
mean $\overline{m}$	0.519924	–	1.895509	–	0.755731	–
σ	0.0048	–	0.0090	–	0.0096	–
$\sigma/\overline{m}$	0.0092	–	0.0047	–	0.013	–
Ground truth	0.526421	–	1.880620	–	0.745762	–

In the. case of multiple solutions, all three possible solutions are given for each invariant, the solution that corresponds to the real solution is marked in bold font in Table II. From this table we note that sometimes only one real solution is possible, in this case, it is the unique correct solution. In the case of existence of multiple solutions, the real solution is always present independent of the selection of the triplet and remains numerically stable.

TABLE II.

TABLE OF INVARIANTS COMPUTED FROM DIFFERENT TRIPLETS OF SYNTHETIC IMAGES. WITH ± 1.5 PIXEL NOISE FOR IMAGE POINTS, ALL THREE SOLUTIONS OF $\{\alpha, \beta, \gamma\}$, IF THERE ARE, ARE SHOWN IN ORDER TO ILLUSTRATE THEIR STABILITY

Image Triplet	α = 0.526421	β = 1.880620	γ = 0.745762
	-3.779942	0.424841	-1.588062
4, 6, 7	**0.515372**	**1.899496**	**0.744176**
	15.791148	–0.023215	–8.050262
4, 5, 6	**0.516476**	**1.899056**	**0.747555**
	–2.006199	0.438052	–1.051868
5, 6, 7	**0.515080**	**1.899593**	**0.742203**
	12.874219	–0.414817	–013.899986
4,5,7	**0.516014**	**1.899551**	**0.746564**

A.3. Stability w.r.t. the pixel errors

For the given triplets of images $\{1,\ 2,\ 3\}$, $\{4,\ 5,\ 6\}$ and $\{7, 8, 9\}$, different pixel noise (uniformly distributed) is added to illustrate the influence of the pixel errors.

Tables III, IV, and V show that with noise levels up to ± 5.5 pixels, the computed invariants remain numerically stable, so that degradation with the increasing pixel noise is graceful.

Tables III, IV, and V also show that the same invariant computed with completely different triplets of images remains stable.

From these empirical results, we note that the uniqueness of solutions depends partly on the positions of the cameras, for the first and the third triplet, the solutions are always multiple. For the second one, they are always unique up to ± 3.5 pixel error.

TABLE III.
IMAGES {1, 2, 3}. α, β AND γ ARE COMPUTED WITH PIXEL ERRORS OF DIFFERENT LEVELS. d_i IS THE DIFFERENCE BETWEEN THE COMPUTED VALUE AND ITS MEAN VALUE

Noise	$\alpha = 0.526421$	$d1$	$\beta = 1.880620$	$d2$	$\gamma = 0.745762$	$d3$
± 0.5	0.527054	0.00063	1.879408	-0.0012	0.751121	0.0054
± 1.5	0.528183	0.0018	1.877117	-0.0035	0.760771	0.015
± 2.5	0.529135	0.0027	1.875013	-0.0056	0.769214	0.023
± 3.5	0.529938	0.0035	1.873026	-0.0076	0.776658	0.031
± 4.5	0.530609	0.0042	1.871119	-0.0095	0.783268	0.038
± 5.5	0.531163	0.0047	1.869260	-0.011	0.789171	0.043

TABLE IV.
TABLE OF INVARIANTS COMPUTED FROM THE TRIPLET OF SYNTHETIC IMAGES {4, 5, 6}. α, β AND γ ARE COMPUTED WITH PIXEL ERRORS OF DIFFERENT LEVELS. d_i IS THE DIFFERENCE BETWEEN THE COMPUTED VALUE AND ITS MEAN VALUE

Noise	$\alpha = 0.526421$	$d1$	$\beta = 1.880620$	$d2$	$\gamma = 0.745762$	$d3$
± 0.5	0.523211*	-0.0032	1.886847*	0.0062	0.746559*	0.00080
± 1.5	0.516476*	-0.0099	1.899056*	0.018	0.747555*	0.0018
± 2.5	0.509556*	-0.017	1.911075*	0.030	0.748065*	0.0023
± 3.5	0.502432*	-0.023	1.922747*	0.042	0.748102*	0.0023
± 4.5	0.495112	-0.031	1.934272	0.054	0.747731	0.0020
± 5.5	0.487509	-0.039	1.945445	0.065	0.746875	0.0011

TABLE V.
TABLE OF INVARIANTS COMPUTED FROM THE TRIPLET OF SYNTHETIC IMAGES {7, 8, 9}. α , β, AND γ IS COMPUTED WITH PIXEL ERRORS OF DIFFERENT LEVELS. D_i IS THE DIFFERENCE BETWEEN THE COMPUTED VALUE AND ITS MEAN VALUE. NOTE THAT WITH THE SMALLEST PIXEL ERROR ± 0.5, THE RESULTS ARE NOT THE BEST ONES

Noise	α=0.526421	d_1	β=1.880620	d_2	γ=0.745762	d_3
± 0.5	0.532012	0.0056	1.921869	0.041	0.710370	-0.035
± 1.5	0.512018	-0.014	1.858706	-0.022	0.743526	-0.0022
± 2.5	0.523644	-0.0028	1.845212	-0.035	0.727474	-0.018
± 3.5	0.496563	-0.030	1.930408	0.050	0.718309	-0.027
± 4.5	0.502496	-0.024	1.966565	0.086	0.721738	-0.024
± 5.5	0.501878	0.025	1.964701	0.084	0.717939	-0.028

A.4. Stability w.r.t. the distance between cameras

In Table I, the different triplets are ordered in increasing camera distance order, globally, the accuracy of the computed invariants is slightly improved by increasing camera distance.

Although the camera distance decreases slightly from {1, 2, 3} to {4, 5, 6 and {7, 8, 9}, comparing Tables III, IV, and V, the difference of the accuracy of the computed invariants is not significant, this is due to the fact that the camera distance for those triplets is already favorable.

In the next section, camera distance influence will be discussed with real image sequences

B. The experiment with real images

We first experimented over a sequence of 24 real images. The object spans about 30 *cm* in space and is placed in about two *m* from the camera. The camera was turned manually around the object, and images are taken every one or two degrees. The first 13 images and the last 11 images have slightly different lighting condition. One gradient image of the sequences is displayed in Fig. 1. It is important to note that the images we used are only a small portion, approximately 1/4 of the original 512×512 images of the sequence.

- Points are automatically tracked over the sequence and the location of the points are optimized by a nonlinear subpixel corner detector [33]. The automatic tracker may have two or three points missing for some images, the missing points are added by hand. The six points used for experimentation are the visible corners of the two cubic objects, marked by white circles in Fig. 1. Concerning the precision of the image points, although corner detection is itself subpixel but the precision of points location can not be considered as subpixel since we can never guarantee that in different images a corner comes from the same physical point in space.
- The dimensions of cubic objects in space are measured by hand to serve as ground truth for comparison with the computed invariants. However these ground truth should be considered relative, since errors in measuring 3D objects are inevitable. Measurement error was estimated as ±2 *mm*.

TABLE VI.

TABLE OF INVARIANTS α, β AND γ COMPUTED FROM THE REAL IMAGE SEQUENCE WITH DIFFERENT TRIPLETS OF IMAGES. IN THE SECOND COLUMN, *CAMERA DIST.* MEANS THE DISTANCE BETWEEN CAMERA POSITIONS. d_i IS THE DIFFERENCE BETWEEN THE COMPUTED VALUE AND THE MEAN VALUE. σ IS THE COMPUTED STANDARD DEVIATION, $\sigma / \overline{m}$ THE RATIO OF THE STANDARD DEVIATION TO THE MEAN VALUE

Ima. Triplet	Camera Dist.	α	d_1	β	d_2	γ	d_3
1, 12, 23	12	3.055932*	-0.032	1.322068*	-0.014	0.044376*	-0.014
2, 13, 24	12	3.040515	0.048	1.3321695	-0.0043	0.051376	-0.0065
1, 7, 13	5	2.981206	-0.11	1.295518	-0.041	0.015540	-0.0424
14, 19, 24	5	3.205979	0.12	1.352756	0.016	0.056934	-0.00097
1, 5, 9	4	3.085201*	-0.0032	1.336204*	-0.00022	0.064528*	0.0066
13, 17, 21	4	3.592866	0.50	1.370506	0.0034	0.053793	-0.0041
14, 18, 22	4	2.976934	-0.11	1.328929	-0.0074	0.049507	-0.0084
2, 5, 8	3	3.006701*	-0.082	1.359429*	0.023	0.117986*	0.060
3, 6, 9	3	3.111329*	0.023	1.353453*	0.017	0.084712*	0.027
14, 17, 20	3	3.417479	0.33	1.366791	0.030	0.059527	0.0016
16, 19, 2	3	2.497743*	-0.59	1.282863*	-0.054	0.038662	-0.019
1, 3, 5	2	3.539165	0.45	1.376995	0.041	1.140780	1.08
15, 17, 19	2	6.754482	3.7	1.509805*	0.17	0.099840	0.042
mean $\overline{m}$	–	3.088353	–	1.336426	–	0.057904	–
s.d. σ	–	0.28	–	0.028	-	0.026	-
$\sigma / \overline{m}$	-	0.091	-	0.021	-	0.45	-
ground truth	-	3.155082	-	1.357450	-	0.073426	-

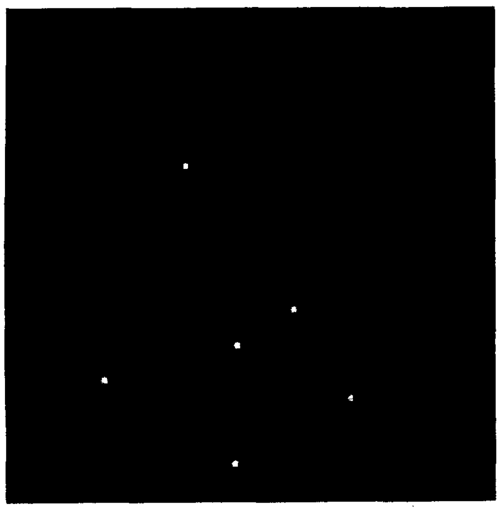


Fig. 1 . One gradient image of the sequence. The six points marked by a white circle are used for the experiment.

The results in Table VI show that the computed invariants from different triplet of images are stable and usable, except for γ in the third row. We note that the computation of γ is systematically worse than that of α and β, this is due to its small modulus, although both have almost the same magnitude of absolute error, about 0.04, the relative error is much more important for γ than α. This is also due to error accumulation, since γ is found using α and β. One possible remedial measure is to change the order of the computation, or recompute them from the corresponding original cubic equations.

In this table, we also note the influence of the camera distances on the stability of the invariants. Since each pair of neighboring images differs by a rotation of roughly one degree, the camera distance in Table VI can roughly be estimated as the rotation angle in degrees with a radius of two meters. We can see in the table that, globally the results degrade by decreasing the distance between camera positions. However the results from the largest distance do not give the best ones, this can be explained by the fact that one side of the objects in these extreme positions are almost tangent to the view line, so the corresponding corners are not well defined in the images, therefore they can not be precisely located by the corner detector. The pixel errors of the points in the extreme images are much more important than in other images. When the distance between cameras becomes quite small, for instance 2, corresponding to the last two rows of the table, the results are considerably degraded, especially for the triplet of the last row. When the camera distance is 1, the results are too bad to be useful, the results are not presented in the table.

Another experiment is performed on a sequence of images of a wooden house. Three views covering about a 45^{o} rotation of the camera around the wooden house are taken. The corners marked in Fig. 2 are tracked for the three images using the same procedure as for the previous image sequence. The five reference points are those numbered 2, 5, 8, 10 and 11 in Fig. 2. Then with any other point, a set of six points is formed. In Table VII, the computed invariants are compared with that computed by transforming the estimated Euclidean coordinates (obtained by the method described in [32]) using the space collineation $A_{4\times4}$ (*cf.* Appendix). In the case of unique real solution, the computed values are followed by a *. In case of multiple solutions, the one that is the closest to the known value is selected by hand.

C. The experiment on projective reconstruction

As we have seen in Section 5 that we can use the invariants to perform projective reconstruction of a point set from three

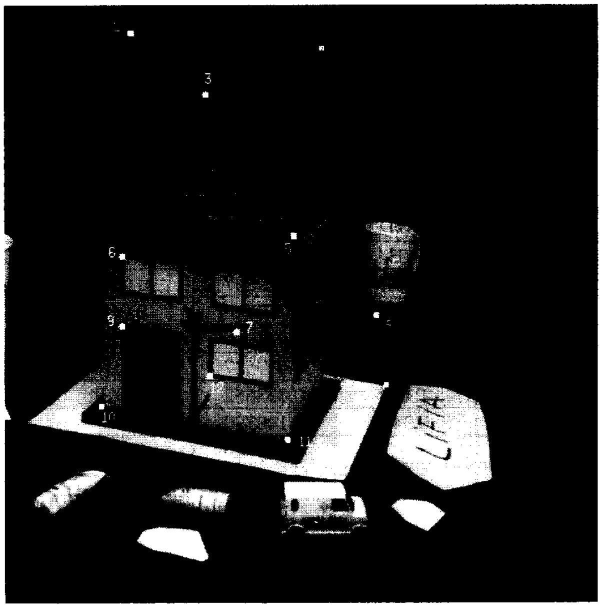


Figure 2. One of the wooden house image sequence. The points used for the experiment are marked by white circles.

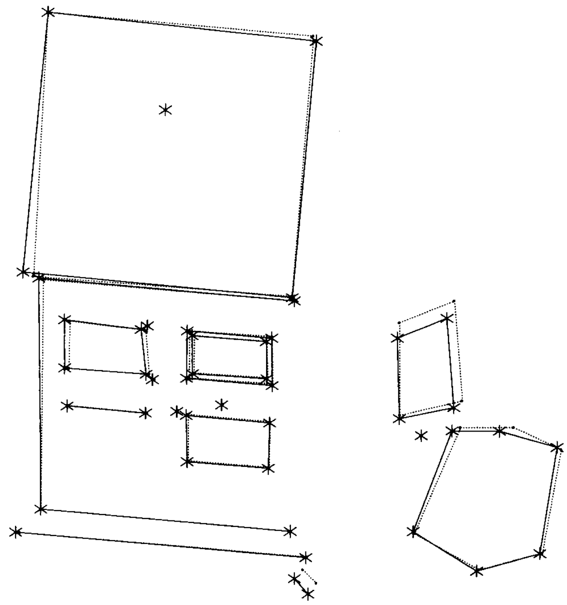


Fig. 3. The rectified projective reconstruction from three uncalibrated images: the points are displayed as stars and the line segments as solid lines. Euclidean reconstruction from five images: the points are dots and the line segments dashed lines. The five reference points used were {2,5,8,10,11} (cf. Fig. 2).

uncalibrated images. We experimented on real images of the wooden house (see Fig. 2). All tracked 46 points of three images are used.

Although the theoretical multiplicity of the projective reconstruction is at most three from the minimal data. With more than the minimum number of points in practice, a unique solution is always obtained. In this experiment, the set of five reference points is {2,5,8,10,11} (*cf.* Fig. 2). The sixth point used to compute the projection matrices is the point numbered 1 in Fig. 2, that is the set of six points {1, 2, 5, 8, 10, 11}. Since the invariants of this set of points are uniquely determined (see the first row of Table VII), the projection matrices are uniquely computed from these invariants, then any tracked points are reconstructed from the projection matrices. The projective reconstruction is then transformed into its Euclidean representation (see Fig. 3) by applying $A_{4\times 4}^{-1}$ (see Appendix) with the known reference points in order to compare the result with direct Euclidean reconstruction from five images [32].

Note that during all reconstruction process, we need to solve only one cubic equation in one variable, all other computational efforts are linear operations

VII. Conclusion

In the first part of the paper, it is shown that the three invariants of a set of six points in general position in space can be algebraically computed from three images of these points. No further information about camera calibration and epipolar geometry is needed. It is shown that the maximum number of solutions is three which are all algebraically explicited in closed form.

Compared with the computation of the invariants from two images, which is essentially based on Sturm's method using seven points, the method proposed in this paper needs only six points of the three images.

The result obtained in this paper has some important consequences. It allows to extend the results of Faugeras [1] and Hartley *et al.* [2] for projective reconstruction and Sturm's method [3] for determining the epipolar geometry from two uncalibrated images to the case of three uncalibrated images with the minimal data of six points instead of seven. It is also shown that a camera could be self-calibrated with only six points of three images.

In the second part of the paper, the stability of the invariants is studied via both simulated and real images with respect to different factors. The computed invariants are numerically stable both with respect to pixel noise and to selection of triplet of images if the distance between cameras is not too small. This confirms that our parametrization has a better numerical behavior than Sturm's. We have also successfully applied the method to perform projective reconstruction of set of points.

We can remark that the way we computed the canonical projective coordinates is not symmetric in the six points, it is therefore order-dependent. The study of order-independent invariants has no theoretical problems, however practically it will produce extremely long and complicated algebraic expression which will compromise numerical stability.

TABLE 7.
TABLE OF INVARIANTS COMPUTED FROM THE THREE IMAGES OF THE WOODEN HOUSE FOR THE DIFFERENT SETS OF SIX POINTS. THE FIRST COLUMN PRESENTS THE SETS OF NUMBERS OF SIX POINTS. IN THE SECOND COLUMN, α', β' AND γ' ARE TRANSFORMED EUCLIDEAN COORDINATES AS GROUND TRUTH. IN THE LAST COLUMN, α, β AND γ ARE SET OF SIX POINTS KNOWN (α', β', γ') COMPUTED (α, β, γ)

Set of 6 Points	known α, β and γ	computed (α, β and γ)
{1,2,5,8,10,11}	(0.594573, 0.944785, 0.157327)	(0.581565*, 0.925869*, 0.143597*)
{2,3,5,8,10,11}	(0.716795, 0.740667, 0.64426)	(0.714164, 0.739041, 0.662835)
{2,4,7,8,10,11}	(1.68619, 12.7413, 0.970638)	(1.714522, 12.729020, 0.995141)
{2,5,6,8,10,11}	(0.681678, 0.241914, 1.7126)	(0.742775*, 0.325263*, 1.822018*)
{2,5,7,8,10,11}	(-0.22391, 0.225495, 0.858623)	(-0.230072, 0.228533, 0.880303)
{2,5,8,9,10,11}	(0.634079, 0.344556, 3.09316)	(0.705443, 0.448216, 3.303431)
{2,5,8,10,11,12}	(-0.816493, 0.264887, 2.45476)	(-0.849691, 0.233549, 2.523767)

APPENDIX

Plane collineation $A_{3\times3}^{-1}$

The plane collineation matrix $A_{3\times3}^{-1}$, which transforms the canonical projective basis in $\mathcal{P}^2$,

$$e_1 = (1,0,0)^T, \quad e_2 = (0,1,0)^T,$$
$$e_3 = (0,0,1)^T, \quad e_4 = (1,1,1)^T,$$

into four given points: $\{p_i = (x_i, y_i, 1)^T, I = 1, \ldots, 4\}$,

$$e_i \rightarrow \mu_i p_i = A_{3\times3}^{-1} e_i,$$

is given by

$$\begin{pmatrix} x_1\Delta_{423} & x_2\Delta_{143} & x_3\Delta_{124} \\ y_1\Delta_{423} & y_2\Delta_{143} & y_3\Delta_{124} \\ \Delta_{423} & \Delta_{143} & \Delta_{124} \end{pmatrix}.$$

where

$$\Delta_{ijk} = det\begin{pmatrix} x_i & y_i & 1 \\ x_j & y_j & 1 \\ x_k & y_k & 1 \end{pmatrix}$$

Space collineation $A_{4\times4}^{-1}$

The space collineation matrix $A_{4\times4}^{-1}$, which transforms the canonical projective basis,

$$e_1 = (1,0,0,0)^T, \quad e_2 = (0,1,0,0)^T, \quad e_3 = (0,0,1,0)^T,$$
$$e_4 = (0,0,0,1)^T, \quad e_5 = (1,1,1,1)^T,$$

into five given points, $\{p_i = (x_i, y_i, z_i, 1)^T, I = 1, \ldots, 5\}$:

$$e_i \rightarrow \mu_i p_i = A_{4\times4}^{-1} e_i,$$

is given by

$$\begin{pmatrix} x_1\Delta_{5234} & x_2\Delta_{1534} & x_3\Delta_{1254} & x_4\Delta_{1235} \\ y_1\Delta_{5234} & y_2\Delta_{1534} & y_3\Delta_{1254} & y_4\Delta_{1235} \\ z_1\Delta_{5234} & z_2\Delta_{1534} & z_3\Delta_{1254} & z_4\Delta_{1235} \\ \Delta_{5234} & \Delta_{1534} & \Delta_{1254} & \Delta_{1235} \end{pmatrix}.$$

where

$$\Delta_{ijkl=det}\begin{pmatrix} x_i & y_i & z_i & 1 \\ x_j & y_j & z_j & 1 \\ x_k & y_k & z_k & 1 \\ x_l & y_l & z_l & 1 \end{pmatrix}$$

Then $A_{3\times3}$ and $A_{4\times4}$ can then be obtained from its inverse either analytically by computing cofactors or numerically by *LU* decomposition.

ACKNOWLEDGEMENT

We would also like to thank B. Boufama, P. Gros, R. Horaud, R. Mohr, D. Sinclair, D. Wang and C. Zhao for providing interesting discussion and help.

REFERENCES

[1] O. Faugeras, "What can be seen in three dimensions with an uncalibrated stereo rig?" *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy,* G. Sandini, Ed. May 1992, pp. 563-578, Springer-Verlag.

[2] R. Hartley, R. Gupta, and T. Chang, "Stereo from uncalibrated cameras," *Proceedings of the Conference on Computer Vision and Pattern Recognition, Urbana-Champaign, Illinois, USA,* 1992, pp. 761-764.

[3] R. Sturm, "Das Problem der Projektivität und seine Anwendung auf die Flächen zweiten Grades," *Math. Ann.,* vol. 1, pp. 533-574, 1869.

[4] J.L. Mundy and A. Zisserman, Eds., *Geometric Invariance in Computer Vision*}, MIT Press, Cambridge, Massachusetts, USA, 1992.

[5] J.J. Koenderink and A. J. van Doorn, "Affine structure from motion," Tech. Rep., Utrecht University, Utrecht, The Netherlands, October 1989.

[6] R. Mohr, L. Quan, F. Veillon, and B. Boufama, "Relative 3D reconstruction using multiples uncalibrated images," Tech. Rep. RT 84-I-IMAG LIFIA 12, LIFIA-IRIMAG, 1992.

[7] R. Mohr, L. Morin, C. Inglebert, and L. Quan, "Geometric solutions to some 3D vision problems," *Integration and Control in Real Time Active Vision,* J.L. Crowley, E. Granum, and R. Storer, Eds., ESPRIT BRA Series. Springer-Verlag, 1991.

[8] L. Quan and R. Mohr, "Affine shape representation from motion

through reference points," *Journal of Mathematical Imaging and Vision*, vol. 1, pp. 145-151, 1992, also in IEEE Workshop on Visual Motion, New Jersey, pages 249-254, 1991.

[9] P. Gros and L. Quan, "Projective Invariants for Vision," Tech. Rep. RT 90 IMAG - 15 LIFIA, LIFIA-IRIMAG, Grenoble, France, December 1992.

[10] G. Sparr, "An algebraic/analytic method for reconstruction from image correspondance," *Proceedings of the 7th Scandinavian Conference on Image Analysis, Aalborg, Denmark*, 1991, pp. 274-281.

[11] R. Hartley, "Invariants of points seen in multiple images," Tech. Rep., G.E. CRD, Schenectady, 1992.

[12] A. Shashua, "Projective Structure from Two Uncalibrated Images : Structure from Motion and Recognition," Tech. Rep. A.I. Memo No. 1363, Massachusetts Institute of Technology, September 1992.

[13] D. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell, "Invariant descriptors for 3D object recognition and pose," *IEEE* TRANSACTIONS ON PAMI, vol. 13, no. 10, pp. 971-991, October 1991.

[14] I. Weiss, "Projective invariants of shapes," *Proceedings of DARPA Image Understanding Workshop, Cambridge, Massachusetts, USA*, 1988, pp. 1125-1134.

[15] J.B. Burns, R. Weiss, and E.M. Riseman, "View variation of point set and line segment features," *Proceedings of DARPA Image Understanding Workshop, Pittsburgh, Pennsylvania, USA*, 1990, pp. 650-659.

[16] C.A. Rothwell, D.A. Forsyth, A. Zisserman, and J.L. Mundy, "Extracting projective structure from single perspective views of 3D point sets", *Proceedings of the 4th International Conference on Computer Vision, Berlin, Germany*, May 1993, pp. 573-582.

[17] E.B. Barrett, M.H. Brill, N.N. Haag, and P.M. Payton, "Invariant linear methods in photogrammetry and model-matching," *Geometric Invariance in Computer Vision*, J. Mundy and A. Zisserman, Eds., pp. 277-292. The MIT press, 1992.

[18] P. Gros and L. Quan, "3D projective invariants from two images," *Geometric Methods in Computer Vision II, SPIE's 1993 International Symposium on Optical Instrumentation and Applied Science*, July 1993, pp. 75-86.

[19] O.D. Faugeras, Q.T. Luong, and S.J. Maybank, "Camera Self-Calibration: Theory and Experiments," *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, G. Sandini, Ed. May 1992, pp. 321-334, Springer-Verlag.

[20] Q.T. Luong, *Matrice Fondamentale et Autocalibration en Vision par Ordinateur*}, Thèse de doctorat, Université de Paris-Sud, Orsay, France, December 1992.

[21] S.J. Maybank and O.D. Faugeras, "A theory of self calibration of a moving camera," *International Journal of Computer Vision*, vol. 8, no. 2, pp. 123-151, 1992.

[22] L. Quan, "Invariants of 6 points from 3 uncalibrated images," *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden*, J.O. Eklundh, Ed. 1994, vol. II, pp. 459-470, Springer-Verlag.

[23] H.C. Longuet-Higgins, "A computer program for reconstructing a scene from two projections," *Nature*, vol. 293, pp. 133-135. XX, September 1981.

[24] A.B. Coble, *Algebraic Geometry and Theta Functions*, American Mathematical Society, 1961.

[25] J.G. Semple and G.T. Kneebone, *Algebraic Projective Geometry*, Oxford Science Publication, 1952.

[26] J.H. Grace and A. Young, *The Algebra of Invariants*, Chelsea Publishing Company, New York, 1903.

[27] J. Dixmier, "Quelques aspects de la théorie des invariants," *Gazette des Mathématiciens*, vol. 43, pp. 39-64, January 1990.

[28] R. Mohr, F. Veillon, and L. Quan, "Relative 3D reconstruction using multiple uncalibrated images," *Proceedings of the Conference on Computer Vision and Pattern Recognition, New York, USA*, June 1993, pp. 543-548.

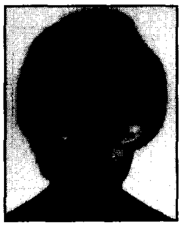
[29] O. Hesse, "Die cubische Gleichung, von welcher die Lösung des Problems der Homographie von M. Chasles abh\"a}ngt," *J. reine angew. Math.*, no. 62, pp. 188-192, 1863.

[30] O. Faugeras and S. Maybank, "Motion from point matches: Multiplicity of solutions," *International Journal of Computer Vision*, vol. 3, no. 4, pp. 225-246, 1990.

[31] D.C. Brown, "Close-range camera calibration," *Photogrammetric Engineering*, vol. 37, no. 8, pp. 855-866, 1971.

[32] B. Boufama, R. Mohr, and F. Veillon, "Euclidian constraints for uncalibrated reconstruction," *Proceedings of the 4th International Conference on Computer Vision, Berlin, Germany*, May 1993, pp. 466-470.

[33] R. Deriche and G. Giraudon, "Accurate corner detection: an analytical study," *Proceedings of the 3rd International Conference on Computer Vision, Osaka, Japan*, 1990.



Long Quan was born in Taiyuan, China, on June 14, 1964. He received the B.Sc. degree in telecommunication from the Northern Jiao-Tong Univerity, Beijing, China in 1984, the M.Sc. degree in computer science from the University of Nancy I in 1986, and Ph.D. degree in computer science from INPL (Institut National Polytechnique de Lorraine), Nancy, France, in 1989.

He joined LIFIA (Laboratoire d'Informatique Fondamentale et d'Intelligence Artificielle), Grenoble, France, as Research Scientist of CNRS (Centre National de Recherche Scientifique) in 1990. His main research interests are currently on applications of invariant theory and projective geometry for computer vision.