NEW PERSPECTIVES ON CAMERA CALIBRATION USING GEOMETRIC ALGEBRA

C. Defferara *A. Dell'Acqua* *F. Negroni* *A. Sarti* *S. Tubaro*

Dip. di Elettronica e Informazione – Politecnico di Milano,
Piazza Leonardo Da Vinci 32, 20133 Milano, Italy

ABSTRACT

In this paper we propose a new approach to the camera self-calibration problem, based on geometric algebra. After a brief introduction on the adopted Clifford algebra framework, we provide new insight on the epipolar constraint as defined in terms of bivectors. On the basis of that, we propose a novel solution for the simultanous determination of the focal lengths of the cameras and the rigid motion between views.

1. INTRODUCTION

Structure from motion (SfM) is often approached in a geometric fashion, by exploiting invariants and constraints of projective geometry [1]. More recently, some effective algebraic solutions based on rank conditions have started to emerge [5, 4]. If teh goal is to devise and implement SfM algorithms that retain the evocative power of geometry, without giving up the effectiveness and the generality of algebraic solutions, we need a mathematical framework where geometry and algebra sinergically co-exist. Geometric (Clifford) algebra (GA) is currently gaining more and more of the interest of researchers in computer vision [2] because it seems to blend such aspects effectively and elegantly.

In this paper we show how geometric algebra can be used to efficently represent the camera geometry and the epipolar constraint, with new insight in its geometric interpretation. Based on that, we propose a novel two-view self-calibration technique. After then, we extend such results to the three-view case and we show what level of further improvement can be achieved.

2. TWO-VIEW ANALYSIS IN THE GA FRAMEWORK

Adopting the same notation used in [2], a generic point $\mathbf{p}$ of the projective space $\mathbb{P}^3$ can be written in homogeneous form as $\mathbf{p} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 + \mathbf{e}_4$, where $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, $\mathbf{e}_4$ form a base of $\mathbb{P}^3$. The line $\mathbf{l}$ passing through a given pair of points $\mathbf{p}_1$ and $\mathbf{p}_2$ can be expressed as a bivector of the form $\mathbf{l} = \mathbf{p}_1 \wedge \mathbf{p}_2$, where the wedge operator denotes the *outer product* between vectors and can be written in terms of the geometric product. Similarly, the plane passing through the three points $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ can be written as the grade-3 blade $\pi = \mathbf{p}_1 \wedge \mathbf{p}_2 \wedge \mathbf{p}_3$.

Another important issue is to test whether two subspaces are incident. A general condition for the incidence of two subspaces $\mathbf{A}$ and $\mathbf{B}$ is given in geometric algebra as $\mathbf{A} \cdot \mathbf{B}^* = 0$, which becomes $\mathbf{A} \wedge \mathbf{B} = 0$ when the grade of $\mathbf{A}\wedge\mathbf{B}$ is smaller or equal to the dimension of the space. This expression becomes very useful when we want to verify the incidence of two lines (bivectors), as the dimension of $\mathbb{P}^3$ is 4. In fact, the two lines $\mathbf{l}_1$ and $\mathbf{l}_2$ are found to intersect in a point $\mathbf{p}$ if and only if $\mathbf{l}_1 \wedge \mathbf{l}_2 = 0$. This allows us to formulate of the epipolar constraint in quite a straightforward fashion. Let $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ be the centers of the cameras and $\mathbf{p}^{(1)}$ and $\mathbf{p}^{(1)}$ be the projections (world coordinates) of a point $\mathbf{p}$ onto the first and second camera, respectively. The epipolar constraint can be written as

$$(\mathbf{c}^{(1)} \wedge \mathbf{p}^{(1)}) \wedge (\mathbf{c}^{(2)} \wedge \mathbf{p}^{(2)}) = 0 \quad . \qquad (1)$$

A simple pin-hole camera model is completely specified by an optical center $\mathbf{c}$, a focal length f and the directions of the camera axes $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$. Under these assumptions, a point of homogeneous image coordinates $\mathbf{m} = [m_1, m_2, m_3]^T$, with $m_3 = 1$, turns out to be expressed as $\mathbf{p} = m_1\mathbf{x}_1 + m_2\mathbf{x}_2 + m_3 f\mathbf{x}_3 + \mathbf{c}$ in the world coordinate frame. If we consider two different views of the same point $\mathbf{p}$, of homogeneous coordinates $\mathbf{m}^{(1)} = \left[m_1^{(1)}, m_2^{(1)}, m_3^{(1)}\right]^T$ and $\mathbf{m}^{(2)} = \left[m_1^{(2)}, m_2^{(2)}, m_3^{(2)}\right]^T$, eq. (1) can be specialized as follows

$$(m_1^{(1)}(\mathbf{c}^{(1)} \wedge \mathbf{x}_1^{(1)}) + m_2^{(1)}(\mathbf{c}^{(1)} \wedge \mathbf{x}_2^{(1)}) + m_3^{(1)} f^{(1)}(\mathbf{c}^{(1)} \wedge \mathbf{x}_3^{(1)}))\wedge$$
$$(m_1^{(2)}(\mathbf{c}^{(2)} \wedge \mathbf{x}_1^{(2)}) + m_2^{(2)}(\mathbf{c}^{(2)} \wedge \mathbf{x}_2^{(2)}) + m_3^{(2)} f^{(2)}(\mathbf{c}^{(2)} \wedge \mathbf{x}_3^{(2)})) = 0\,, \qquad (2)$$

$f^{(1)}$, $\mathbf{c}^{(1)}$, $\mathbf{x}_i^{(1)}$ are the parameters of the first camera and $f^{(2)}$, $\mathbf{c}^{(2)}$, $\mathbf{x}_i^{(2)}$ are those of the second camera. If, for the moment, we assume that $f^{(1)} = f^{(2)} = 1$, then eq. (2) can be expanded as a sum of grade-4 blades of the form $m_i^{(1)} m_i^{(2)} \varepsilon_{ij} \mathbf{I}_4$, where ε_{ij} are unknown scalars, therefore the epipolar constraint takes on the form

$$\sum_{i,j=1..3} m_i^{(1)} m_i^{(2)} \varepsilon_{ij} \mathbf{I}_4 = 0 \quad .$$

This expression, after eliminating $\mathbf{I}_4$, can be written in matrix form as

$$\mathbf{m}^T \mathbf{E} \mathbf{n} = 0 \qquad (3)$$

where $\mathbf{E}$ is the 3×3 matrix of the coefficients ε_{ij}, which is the classical formulation of the epipolar constraint where $\mathbf{E}$ is the well-known essential matrix. More generally, when no assumptions are made on $f^{(1)}$ and $f^{(2)}$, similar considerations hold true and, as we will see later on, eq. (3) becomes the fundamental matrix $\mathbf{F}$.

As we can see, in GA the epipolar constraint is written directly as an incidence relation between lines, which is a something that has no counterpart in projective geometry. In fact, lines have no direct homogeneous representation in projective spaces [1] (they can be represented with Plücker matrices or as an intersection between planes), and are algebraically described through appropriate rank conditions [4].

A line l in GA can be written as the linear combination of the base bivectors as follows

$$\mathbf{l} = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \alpha_3 \mathbf{b}_3 + \beta_1 \hat{\mathbf{b}}_1 + \beta_2 \hat{\mathbf{b}}_2 + \beta_3 \hat{\mathbf{b}}_3 \,, \quad (4)$$

where $\mathbf{b}_1 = \mathbf{e}_2 \wedge \mathbf{e}_3$, $\mathbf{b}_2 = \mathbf{e}_3 \wedge \mathbf{e}_1$, $\mathbf{b}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$, $\widehat{\mathbf{b}}_1 = \mathbf{e}_4 \wedge \mathbf{e}_1$, $\widehat{\mathbf{b}}_2 = \mathbf{e}_4 \wedge \mathbf{e}_2$ and $\widehat{\mathbf{b}}_3 = \mathbf{e}_4 \wedge \mathbf{e}_3$. This notation for the grade-2 base elements emphasizes the fact that base bivectors l_i and $\widehat{l}_i$ are pairwise dual. In fact, a line (bivector) can always be written as the sum of two terms:

- a line $\beta_1 \hat{\mathbf{b}}_1 + \beta_2 \hat{\mathbf{b}}_2 + \beta_3 \hat{\mathbf{b}}_3$ passing through the origin of the world reference frame ("finite" component);
- a line $\alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \alpha_3 \mathbf{b}_3$ on the plane at infinity (component "at infinity").

Notice that this notation for lines is somewhat redundant, as it involves 6 (projective) parameters instead of 5. The extra degree of freedom will be later removed through a consistency constraint on the coefficients.

The coefficients α_i and β_i can be obtained by computing the inner product between the line $\mathbf{l}$ and the corresponding base bivector, $\mathbf{b}_i$ or $\widehat{\mathbf{b}}_i$. For example, we have

$$\begin{aligned} \mathbf{l} \cdot \mathbf{b}_i &= (\alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \alpha_3 \mathbf{b}_3 + \beta_1 \hat{\mathbf{b}}_1 + \beta_2 \hat{\mathbf{b}}_2 + \beta_3 \hat{\mathbf{b}}_3) \cdot \mathbf{b}_i \\ &= \alpha_i \mathbf{b}_i \cdot \mathbf{b}_i = -\alpha_i \quad . \end{aligned} \qquad (5)$$

A camera with center $\mathbf{c}$, and axis directed as $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ can be represented by the three lines $\mathbf{c} \wedge \mathbf{x}_1$, $\mathbf{c} \wedge \mathbf{x}_2$ and $\mathbf{c} \wedge \mathbf{x}_3$ corresponding to its axes. To retrieve position and orientation of a camera we must find the finite and infinite components of these lines.

2.1. Essential matrix

In this Section we will show that the coefficients of the infinite components of the axes of the second camera correspond to the elements ε_{ij} of the essential matrix $\mathbf{E}$. To do so, without loss of generality, we assume that the axes of the world coordinate frame are oriented like the axes of the first camera, and that the origin of the world frame is in the camera's optical center,i. e. $\mathbf{x}_1^{(1)} = \mathbf{e}_1$, $\mathbf{x}_1^{(2)} = \mathbf{e}_2$, $\mathbf{x}_1^{(3)} = \mathbf{e}_3$ and $\mathbf{c}^{(1)} = \mathbf{e}_4$. With this assumption, we can rework the epipolar constraint (2) to obtain nine equations of the form

$$\varepsilon_{ij} \mathbf{I}_4 = (\mathbf{e}_4 \wedge \mathbf{e}_i) \wedge (\mathbf{c}^{(2)} \wedge \mathbf{x}_j^{(2)}) \qquad (6)$$

all involving the quadrivector $\mathbf{I}_4$. As we can see, there are three equations for each axis $\mathbf{x}_j^{(2)}$, whose unknowns are both ε_{ij} and the axes $\mathbf{c}^{(2)} \wedge \mathbf{x}_j^{(2)}$ of the second camera. If we compute the inner product between both sides of eq. (6) and the bivector $\mathbf{b}_i$, we obtain

$$\varepsilon_{ij} \mathbf{I}_4 \cdot \mathbf{b}_i = (\widehat{\mathbf{b}}_i \wedge (\mathbf{c}^{(2)} \wedge \mathbf{x}_j^{(2)}) \cdot \mathbf{b}_i)$$

using the known equalities $\mathbf{I}_4 \cdot \mathbf{b}_i = \widehat{\mathbf{b}}_i$, and $(\mathbf{A} \wedge \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})$, we can write

$$\varepsilon_{ij} \widehat{\mathbf{b}}_i = \widehat{\mathbf{b}}_i \cdot ((\mathbf{c}^{(2)} \wedge \mathbf{x}_j^{(2)}) \cdot \mathbf{b}_i) \quad . \qquad (7)$$

Notice that the term $(\mathbf{c}^{(2)} \wedge \mathbf{x}_j^{(2)}) \cdot \mathbf{b}_i$ in the right-hand side of eq. (7) is a scalar, therefore we can write $\varepsilon_{ij} = (\mathbf{c}^{(2)} \wedge \mathbf{x}_j^{(2)}) \cdot \mathbf{b}_i$. As shown in eq. (5), the inner product between a bivector $\mathbf{l}$ and the base bivector $\mathbf{b}_i$ at infinity, returns the relative coefficient α_i, with a sign change. This shows that the generic element ε_{ij} of the essential matrix is, in fact, the coefficient of the component at infinity $\mathbf{b}_i$ of the camera-2 axis $\mathbf{x}_j^{(2)}$. We can thus conclude that, knowing the essential matrix, we already have the components at infinity of the camera-2 axes.

2.2. Rotation matrix

In order to determine position and orientation of the second camera we still need to compute the coefficients of the base bivectors $\widehat{l}_j$ that pass through the world origin. We will show how this coefficients correspond to the elements of the rotation matrix which brings from the first camera to the second. With this goal in mind, we need a compact notation for the axes of the second camera

$$\mathbf{c}^{(2)} \wedge \mathbf{x}_i^{(2)} = -\mathbf{E}_i^T \mathbf{b} - \mathbf{R}_i^T \widehat{\mathbf{b}}, \quad i = 1, 2, 3, \qquad (8)$$

where $\mathbf{E}_j = \begin{bmatrix} \varepsilon_{1j} & \varepsilon_{2j} & \varepsilon_{3j} \end{bmatrix}^T$, $j = 1, ..., 3$, are the columns of $\mathbf{E}$; the vectors $\mathbf{R}_j = \begin{bmatrix} r_{1j} & r_{2j} & r_{3j} \end{bmatrix}^T$, $j = 1, ..., 3$, collect the unknowns; while $\mathbf{b}$ and $\widehat{\mathbf{b}}$ are defined as $\mathbf{b} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix}^T$ and $\widehat{\mathbf{b}} = \begin{bmatrix} \widehat{\mathbf{b}}_1 & \widehat{\mathbf{b}}_2 & \widehat{\mathbf{b}}_3 \end{bmatrix}^T$. We will now prove that $\mathbf{R}_j$, $j = 1, ..., 3$, are the columns of the rotation matrix of the second camera.

One interesting property of a generic line (4) of the projective space $\mathbb{P}^3$ is that its orientation is given by its intersection with the plane at infinity $\pi_\infty = \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$, which can be written as $(\pi_\infty \cdot \mathbf{I}_4) \cdot \mathbf{l} = \mathbf{l} \cdot \pi^*_\infty$. We can write

$$(\alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \alpha_3 \mathbf{b}_3 + \beta_1 \hat{\mathbf{b}}_1 + \beta_2 \hat{\mathbf{b}}_2 + \beta_3 \hat{\mathbf{b}}_3) \cdot (-\mathbf{e}_4) = \\ -\beta_1 \hat{\mathbf{b}}_1 \cdot \mathbf{e}_4 - \beta_2 \hat{\mathbf{b}}_2 \cdot \mathbf{e}_4 - \beta_3 \hat{\mathbf{b}}_3 \cdot \mathbf{e}_4 = \\ \beta_1 \mathbf{e}_1 + \beta_2 \mathbf{e}_2 + \beta_3 \mathbf{e}_3 \quad .$$

Also, eq. (8) implies that the directions $\mathbf{x}_1^{(2)}$, $\mathbf{x}_2^{(2)}$, $\mathbf{x}_3^{(2)}$ of the camera-2 axes can be written as a function of the directions $\mathbf{x}_1^{(1)} = \mathbf{e}_1$, $\mathbf{x}_2^{(1)} = \mathbf{e}_2$, $\mathbf{x}_3^{(1)} = \mathbf{e}_3$ of the camera-1 axes

$$\begin{aligned} \mathbf{x}_1^{(2)} &= -r_{11}\mathbf{e}_1 - r_{21}\mathbf{e}_2 - r_{31}\mathbf{e}_3 \\ \mathbf{x}_2^{(2)} &= -r_{12}\mathbf{e}_1 - r_{22}\mathbf{e}_2 - r_{32}\mathbf{e}_3 \\ \mathbf{x}_3^{(2)} &= -r_{13}\mathbf{e}_1 - r_{23}\mathbf{e}_2 - r_{33}\mathbf{e}_3 \quad . \end{aligned}$$

It is now quite apparent that matrix of the unknowns $\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 & \mathbf{R}_3 \end{bmatrix}^T$ is, in fact, the rotation matrix of the second camera.

2.3. Retrieving the second camera

We now have enough tools to derive an alternative formulation of the self-calibration problem. The essential matrix $\mathbf{E}$ can, in fact, be computed using a few point-corrispondences between the two views (see [1]), therefore all we need for determining the orientation of the second camera are the coefficients r_{ij} that describe the "finite" component of the camera-2 axes. In order to estimate the coefficients of this component, a set of constraints between the known and unknown parameters needs to be found. First of all, the axes of the second camera must meet in the optical center $\mathbf{c}_2$. This leads to the following pairwise-incident conditions

$$\begin{aligned} (\mathbf{c}^{(2)} \wedge \mathbf{x}_1^{(2)}) \wedge (\mathbf{c}^{(2)} \wedge \mathbf{x}_2^{(2)}) &= 0 \\ (\mathbf{c}^{(2)} \wedge \mathbf{x}_1^{(2)}) \wedge (\mathbf{c}^{(2)} \wedge \mathbf{x}_3^{(2)}) &= 0 \\ (\mathbf{c}^{(2)} \wedge \mathbf{x}_2^{(2)}) \wedge (\mathbf{c}^{(2)} \wedge \mathbf{x}_3^{(2)}) &= 0 \quad , \end{aligned} \tag{9}$$

which can be rewritten as

$$\begin{cases} \mathbf{E}_1^T \mathbf{R}_2 + \mathbf{E}_2^T \mathbf{R}_1 = 0 \\ \mathbf{E}_1^T \mathbf{R}_3 + \mathbf{E}_3^T \mathbf{R}_1 = 0 \\ \mathbf{E}_2^T \mathbf{R}_3 + \mathbf{E}_3^T \mathbf{R}_2 = 0 \end{cases} \quad . \tag{10}$$

Such equations, however, are only meant to imply that the axes will meet pairwise, therefore we also need an additional orthogonality constraint on the axes. This could be done by imposing that $\mathbf{R}$ be an orthonormal matrix with unit determinant. However, it is more convenient to represent rotations with *rotors*, which better exploit the characteristics of geometric algebra and are intrinsecally related to quaternions. In fact, the generic rotor in the the metric space $\mathbb{E}^3$ is expressed as a multivector of the form $\mathbf{Q} = q_0 + q_1\mathbf{b}_1 + q_2\mathbf{b}_2 + q_3\mathbf{b}_3$, which has a scalar component q_0 and a bivector component $q_1\mathbf{b}_1 + q_2\mathbf{b}_2 + q_3\mathbf{b}_3$, subjected to the normalization constraint

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1 \quad . \tag{11}$$

Incidentally, the bivector component $q_1\mathbf{b}_1 + q_2\mathbf{b}_2 + q_3\mathbf{b}_3$ only involves bivectors at infinity in the projective space $\mathbb{P}^3$. Represent rotations with rotors, the orthonormal constraint on $\mathbf{R}$ is automatically satisfied.

Notice however, that it is not difficult to derive the rotation matrix from the rotor's components

$$\begin{aligned} \mathbf{R}_1 &= \begin{bmatrix} q_0^2 - q_3^2 - q_2^2 + q_1^2 & 2q_1q_2 + 2q_0q_3 & 2q_1q_3 - 2q_0q_2 \end{bmatrix}^T \\ \mathbf{R}_2 &= \begin{bmatrix} 2q_1q_2 - 2q_0q_3 & -q_1^2 + q_0^2 + q_2^2 - q_3^2 & 2q_0q_1 + 2q_2q_3 \end{bmatrix}^T \\ \mathbf{R}_3 &= \begin{bmatrix} 2q_1q_3 + 2q_0q_2 & -2q_0q_1 + 2q_2q_3 & -q_2^2 - q_1^2 + q_0^2 + q_3^2 \end{bmatrix}^T . \end{aligned}$$

An additional set of constraints can be derived from the fact that the essential matrix $\mathbf{E}$ can always be written in closed form as $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$, where $\mathbf{t}$ and $\mathbf{R}$ are the translation vector and the rotation matrix of the second camera with respect to the first one, and $[\mathbf{t}]_\times$ is the skew-simmetric matrix form of $\mathbf{t}$ [1]. This implies that each row of $\mathbf{E}$ is bound to be orthogonal to the corresponding row of $\mathbf{R}$, i.e.

$$\mathbf{E}_1^T \mathbf{R}_1 = 0, \quad \mathbf{E}_2^T \mathbf{R}_2 = 0, \quad \mathbf{E}_3^T \mathbf{R}_3 = 0\,. \tag{12}$$

This leads to an interesting property of lines in geometric algebra. In fact, if we write a generic line (4) as the outer product of two of its points in $\mathbb{P}^3$, the coefficients of the bivectors at infinity α_i, $i = 1, ...3$, and the coefficients β_i, $i = 1, ...3$ of the "finite" base bivectors, must satisfy the consistency constraint $\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 = 0$. This result can also be proven using classical tools of geometric algebra. This constraint is equivalent to the Plücker consraint mentioned in Section 2.1, which reduces the notational redundancy of the bivector representations. Eqs. (10) and (12) can be expressed in terms of $\begin{bmatrix} q_0 & q_1 & q_2 & q_3 \end{bmatrix}^T$. Along with the normalization constraint on quadrivectors we end up with a nonlinear system of seven equations in four unknowns. As $\mathbf{E}$ is a rank-2 matrix, only six of these seven equations are, in fact, linearly independent. It is thus possible to compute position and orientation of the second camera by numerically solving the system. This way we end up with two solutions, only one of which corresponds to a camera whose optical axis is oriented consistently with that of the first camera.

2.4. Focal length estimation

In the previous Sections we made the assumption that the focal lengths of the cameras were equal to one. We will now remove this limitation and show how to estimate the unknown focal lengths. Once again, with no loss of generality, we assume the first camera to be placed in the origin of the world reference frame. In this case the epipolar constraint takes on the usual form $\mathbf{m}^T\mathbf{F}\mathbf{n} = 0$, where $\mathbf{F}$ is the fundamental matrix. On the other hand, the coefficients ε_{ij} of the axes of the second camera are still the elements of the essential matrix $\mathbf{E}$.

As we know, the relationship between the essential matrix $\mathbf{E}$ and the fundamental matrix $\mathbf{F}$ is

$$\mathbf{E} = \mathbf{K}^{(2)T}\mathbf{F}\mathbf{K}^{(2)} \qquad (13)$$

where $\mathbf{K}^{(1)}$ =diag$(f^{(1)}, f^{(1)}, 1)$ and $\mathbf{K}^{(2)}$ =diag$(f^{(2)}, f^{(2)}, 1)$ are the matrices of intrinsic parameters (in this case only the focal lengths) of the first and second camera, respectively. Eq. (13) can be expanded as

$$\mathbf{E} = \begin{bmatrix} f_{11} & f_{12} & f_{13}/f^{(2)} \\ f_{21} & f_{22} & f_{23}/f^{(2)} \\ f_{31}/f^{(1)} & f_{32}/f^{(1)} & f_{33}/(f^{(1)}f^{(2)}) \end{bmatrix} \quad . \qquad (14)$$

Similarly to what done in the previous Section, we can use the notation (8) to express the axes of the second camera. The system of equations formed by (10), (12) and the constraint on quaternions is still sufficient to retrieve both orientation and focal lengths of the second camera. In fact we now have seven nonlinear equations (six linearly independent) in the six unknowns $[\ q_0 \quad q_1 \quad q_2 \quad q_3 \quad f^{(1)} \quad f^{(2)}\ ]$. This system is fully constrained and allows us to find both focal lengths, plus position and orientation of the second camera with respect to the first one. Notice that the system has more than one solution, only one of which is correct. This solution can be easily determined as the one such that $f^{(1)} > 0$, $f^{(2)} > 0$, and focal axes consistently oriented.

3. THREE-VIEW ANALYSIS IN THE GA FRAMEWORK

We consider now the case of three cameras in the projective space. Let us assume, for the moment, that all the focal lengths of all three cameras are equal to 1. In order to simplify the derivation of the trifocal constraint, we introduce a new camera parametrization. As the trifocal constraint will be derived from a correspondence condition that involves line features on the image plane, it is more convenient to parametrize the cameras through its camera planes (the planes spanned by its axes).

Given a camera with center $\mathbf{c}$ and axes $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ its camera planes are defined as

$$\begin{aligned} \pi_1 &= \mathbf{c} \wedge \mathbf{x}_2 \wedge \mathbf{x}_3 \\ \pi_2 &= \mathbf{c} \wedge \mathbf{x}_3 \wedge \mathbf{x}_1 \\ \pi_3 &= \mathbf{c} \wedge \mathbf{x}_1 \wedge \mathbf{x}_2 \quad . \end{aligned}$$

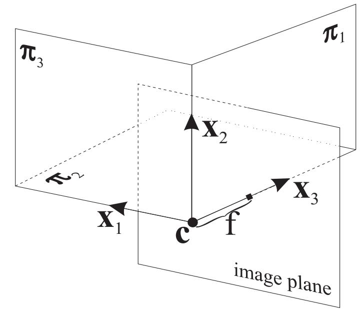

Fig. 1. A camera is completely defined by its axes and center, or by the three camera planes, and its focal lenght.

This second parametrization is completely equivalent to the previous one, as the camera axes and the optical center can be easily obtained as intersections between the camera planes. In GA, the intersection of subspaces is computed with the *meet* operator ($\vee$), therefore the camera axes are

$$\begin{aligned} \mathbf{x}_1 &= \pi_2 \vee \pi_3 \\ \mathbf{x}_2 &= \pi_3 \vee \pi_1 \\ \mathbf{x}_3 &= \pi_1 \vee \pi_2 \quad , \end{aligned}$$

and the optical center is

$$\mathbf{c} = \pi_1 \vee \pi_2 \vee \pi_3 \quad .$$

A camera plane parametrization simplifies the specification of line features on the image plane. In fact, the backprojection $\pi_\mathbf{l}$ of a line $\mathbf{l} = [\ l_1 \quad l_2 \quad l_3\ ]^T$ on the image plane can be written as a linear combination

$$\pi_\mathbf{l} = l_1\pi_1 + l_2\pi_2 + l_3\pi_3 \qquad (15)$$

of the camera planes π_1, $\pi_2$and π_3.

Let us consider a line $\mathbf{L}$ in 3-space, which projects onto the three image lines $\mathbf{l}^{(1)}$, $\mathbf{l}^{(2)}$ and $\mathbf{l}^{(3)}$. The backprojections of such lines

$$\pi_\mathbf{l}^{(i)} = l_1^{(i)}\pi_1^{(i)} + l_2^{(i)}\pi_2^{(i)} + l_3^{(i)}\pi_3^{(i)} = \sum_{r=1}^{3} l_r^{(i)}\pi_r^{(i)} \quad i = 1,2,3 \quad ,$$

are bound to intersect in $\mathbf{L}$.

The line $\mathbf{L}$ in 3-space can also be written in terms of the *meet* operator as

$$\mathbf{L} = \pi_\mathbf{l}^{(2)} \vee \pi_\mathbf{l}^{(3)} \quad . \qquad (16)$$

Considering that the meet operator is distributive with respect to the sum, the expression (16) can reworked as

$$\begin{aligned} \pi_\mathbf{l}^{(2)} \vee \pi_\mathbf{l}^{(3)} &= \left(\sum_{s=1}^{3} l_s^{(2)}\pi_s^{(2)}\right) \vee \left(\sum_{t=1}^{3} l_t^{(3)}\pi_t^{(3)}\right) \\ &= \sum_{s=1}^{3}\sum_{t=1}^{3} l_s^{(2)} l_t^{(3)} (\pi_s^{(2)} \vee \pi_t^{(3)}) \quad . \end{aligned}$$

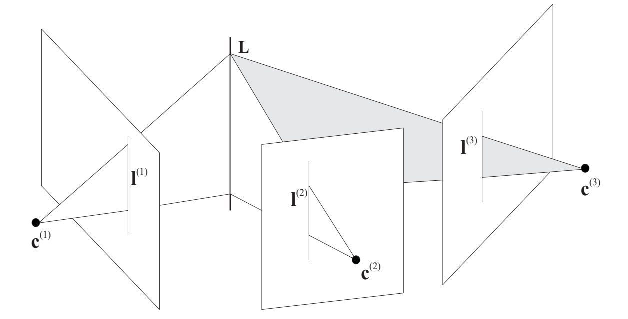

Fig. 2. Three image lines are projections of the same line in the space (i.e. satisfies the trifocal constraint) if and only if the corresponding back-projection planes intersect in a single line in the 3D space.

If the lines $\mathbf{l}^{(1)}$, $\mathbf{l}^{(2)}$ and $\mathbf{l}^{(3)}$ are homologous (corresponding to same line $\mathbf{L}$), then the outer product of the intersection $(\pi_{\mathbf{l}}^{(2)} \vee \pi_{\mathbf{l}}^{(3)})$ and the center of the first camera $\mathbf{c}^{(1)}$ must be equal to the back projection plane $\pi_{\mathbf{l}}^{(1)}$. We can thus write the trifocal constraint as

$$\pi_{\mathbf{l}}^{(1)} = \mathbf{c}^{(1)} \wedge (\pi_{\mathbf{l}}^{(2)} \vee \pi_{\mathbf{l}}^{(3)}) \quad . \tag{17}$$

We can simplify this expression by computing the outer product between $\pi_{\mathbf{l}}^{(1)}$ and the direction $\mathbf{x}_j^{(1)}$, $j = 1, 2, 3$, of each one of the axes of the first camera. Let us first consider $\mathbf{x}_1^{(1)}$:

$$\begin{aligned} \pi_{\mathbf{l}}^{(1)} \wedge \mathbf{x}_1^{(1)} &= \mathbf{c}^{(1)} \wedge (\pi_{\mathbf{l}}^{(2)} \vee \pi_{\mathbf{l}}^{(3)}) \wedge \mathbf{x}_1^{(1)} \\ l_1^{(1)}(\mathbf{c}^{(1)} \wedge \mathbf{x}_2^{(1)} \wedge \mathbf{x}_3^{(1)} \wedge \mathbf{x}_1^{(1)}) &= \mathbf{c}^{(1)} \wedge (\pi_{\mathbf{l}}^{(2)} \vee \pi_{\mathbf{l}}^{(3)}) \wedge \mathbf{x}_1^{(1)} \end{aligned} \tag{18}$$

As we can see from eq. (18), we end up with quadrivectors only. In particular, if the directions $\mathbf{x}_1^{(1)}$, $\mathbf{x}_2^{(1)}$ and $\mathbf{x}_3^{(1)}$ are orthonormal, the term $(\mathbf{c}^{(1)} \wedge \mathbf{x}_2^{(1)} \wedge \mathbf{x}_3^{(1)} \wedge \mathbf{x}_1^{(1)})$ can be proven to be equal to $-\mathbf{I}_4$, therefore eq. (18) becomes

$$\begin{aligned} l_1^{(1)}\mathbf{I}_4 &= \mathbf{c}^{(1)} \wedge \mathbf{x}_1^{(1)} \wedge (\pi_{\mathbf{l}}^{(2)} \vee \pi_{\mathbf{l}}^{(3)}) \\ &= \mathbf{c}^{(1)} \wedge \mathbf{x}_1^{(1)} \wedge \sum_{s=1}^{3}\sum_{t=1}^{3} l_s^{(2)} l_t^{(3)} (\pi_s^{(2)} \vee \pi_t^{(3)}) \\ &= \sum_{s=1}^{3}\sum_{t=1}^{3} l_s^{(2)} l_t^{(3)} (\mathbf{c}^{(1)} \wedge \mathbf{x}_1^{(1)} \wedge (\pi_s^{(2)} \vee \pi_t^{(3)})) \quad . \end{aligned}$$

As all terms $(\mathbf{c}^{(1)} \wedge \mathbf{x}_1^{(1)} \wedge (\pi_s^{(2)} \vee \pi_t^{(3)}))$ are pseudoscalars, they are all proportional to $\mathbf{I}_4$. The substitutions $(\mathbf{c}^{(1)} \wedge \mathbf{x}_1^{(1)} \wedge (\pi_s^{(2)} \vee \pi_t^{(3)})) = \tau_{1st}\mathbf{I}_4$ lead to

$$l_1^{(1)} = \sum_{s=1}^{3}\sum_{t=1}^{3} l_s^{(2)} l_t^{(3)} \tau_{1st} \quad .$$

Similar expressions can be derived for $l_2^{(1)}$ and $l_3^{(1)}$. We have thus re-derived the classic formulation of the trifocal constraint, where the elements τ_{rst} are exactly the elements of the Hartley tensor [1].

3.1. Trifocal tensor

Exactly as seen in Section 2.1, we will now show that the elements of the trifocal tensor can be geometrically interpreted in terms of infinite components of the camera parametrization.

As seen above, the elements of the trifocal tensor can be written in terms of GA operators as

$$\tau_{rst}\mathbf{I}_4 = (\mathbf{c}^{(1)} \wedge \mathbf{x}_r^{(1)}) \wedge (\pi_s^{(2)} \vee \pi_t^{(3)}) \quad . \tag{19}$$

Like in the bifocal case, we assume that the axes of the world coordinate frame are oriented like the axes of the first camera, and that the origin of the world frame is in the camera's optical center. Then eq. (19) becomes

$$\tau_{rst}\mathbf{I}_4 = (\mathbf{e}_4 \wedge \mathbf{e}_r) \wedge (\pi_s^{(2)} \vee \pi_t^{(3)}) \quad .$$

This expression is very similar to (6). In fact, in both cases we have the outer product between an unknown bivector ($(\mathbf{c}^{(2)} \wedge \mathbf{x}_j^{(2)})$ in the bifocal case, $(\pi_s^{(2)} \vee \pi_t^{(3)})$ in the trifocal case) and a base bivector. With a similar procedure to that followed in the bifocal case, we can prove that the elements τ_{rst} *are the coefficients of the component at infinity of the lines* $(\pi_s^{(2)} \vee \pi_t^{(3)})$. These lines are the intersections of the planes of the second camera and the planes of the third camera. Notice that in the bifocal case the unknown lines were the axes of the second camera. Such axes are obtained as mutual intersections between the 3 planes of the second camera.

In the trifocal case we have three planes for the second camera and other three for the third camera, therefore we have nine lines of cross-intersection between planes of the second and third camera. Each one of these lines has three coefficients that characterize the infinite component, therefore we need 27 parameters to specify all infinite components of the unknown lines. This number corresponds to the number of elements of the trifocal tensor. Moreover, the three lines obtained by intersecting one camera plane with the three planes of the other camera are bound to be coplanar (linearly dependent), therefore only the coefficients of six lines turn out to be linearly independent. This confirms the fact that the trifocal tensor has 18 degrees of fredoom.

3.2. Retrieving the cameras

The trifocal tensor can be computed directly from line (or point) correspondances on the images. In order to compute position and orientation of the second and third camera we must find a system of equations to retreive the expression of the planes of the second and third camera from the trifocal tensor. The first camera is assumed to be in canonical position.

The camera planes are trivectors, and therefore linear combination of the base trivectors of $\mathbb{P}^3$ that, in the following, will be indicated with the notation:

$$\begin{aligned} \mathbf{t}_1 &= \mathbf{e}_{2,3,4} \\ \mathbf{t}_2 &= \mathbf{e}_{3,1,4} \\ \mathbf{t}_3 &= \mathbf{e}_{1,2,4} \\ \mathbf{t}_4 &= \mathbf{e}_{3,1,2} \quad . \end{aligned}$$

In the case of trivectors we have that three of them represent planes through the origin ($\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_3$), and are therefore finite planes, while the last base trivector ($\mathbf{t}_4$) corresponds to the plane at infinity. The planes of the second and the third camera can be written as

$$\begin{aligned} \pi_s^{(2)} &= c_{s1}^{(2)}\mathbf{t}_1 + c_{s2}^{(2)}\mathbf{t}_2 + c_{s3}^{(2)}\mathbf{t}_3 + c_{s4}^{(2)}\mathbf{t}_4 \\ \pi_t^{(3)} &= c_{t1}^{(3)}\mathbf{t}_1 + c_{t2}^{(3)}\mathbf{t}_2 + c_{t3}^{(3)}\mathbf{t}_3 + c_{t4}^{(3)}\mathbf{t}_4 \quad , \end{aligned}$$

where $c_{ij}^{(k)}$ are the unknowns that we must retreive in order to derive the position and the orientation of the cameras. Given two of these planes we can directly compute their *meet* as

$$\begin{gathered} (\pi_s^{(2)} \vee \pi_t^{(3)}) = \\ = (c_{s1}^{(2)}c_{t4}^{(3)} - c_{s4}^{(2)}c_{t1}^{(3)})e_{2,3} + (c_{s2}^{(2)}c_{t4}^{(3)} - c_{s4}^{(2)}c_{t2}^{(3)})e_{3,1} + \\ + (c_{s3}^{(2)}c_{t4}^{(3)} - c_{s4}^{(2)}c_{t3}^{(3)})e_{1,2} + (c_{s3}^{(2)}c_{t2}^{(3)} - c_{s2}^{(2)}c_{t3}^{(3)})e_{4,1} + \\ + (c_{s1}^{(2)}c_{t3}^{(3)} - c_{s3}^{(2)}c_{t1}^{(3)})e_{4,2} + (c_{s2}^{(2)}c_{t1}^{(3)} - c_{s1}^{(2)}c_{t2}^{(3)})e_{4,3} \quad . \end{gathered}$$

The elements of the trifocal tensor correspond to the coefficients of the infinite components of the lines $\pi_s^{(2)} \vee \pi_t^{(3)}$, $r, s = 1, 2, 3$. This provides us with a set of 27 equations of the form

$$\tau_{rst} = c_{sr}^{(2)}c_{t4}^{(3)} - c_{s4}^{(2)}c_{tr}^{(3)} \quad , \tag{20}$$

which are 27 nonlinear constraints on the 24 unknowns.

We now need a second set of geometric constraints to impose the orthogonality between the three planes of a same camera. This constraint could easily expressed in terms of the inner product:

$$\begin{aligned} c_{11}^{(j)}c_{21}^{(j)} + c_{12}^{(j)}c_{22}^{(j)} + c_{13}^{(j)}c_{23}^{(j)} &= 0 \\ c_{11}^{(j)}c_{31}^{(j)} + c_{12}^{(j)}c_{32}^{(j)} + c_{13}^{(j)}c_{33}^{(j)} &= 0 \\ c_{31}^{(j)}c_{21}^{(j)} + c_{32}^{(j)}c_{22}^{(j)} + c_{33}^{(j)}c_{23}^{(j)} &= 0 \quad . \end{aligned} \tag{21}$$

By combining eqs. (20) and (21) we obtain a set of nonlinear equations for the computation of position and orientation of the cameras from the trifocal tensor.

The constraints (21) can also be expressed using rotors. In order to do so, let us consider the intersection of a plane π with the plane at infinity π_∞. Using the meet operator we can write

$$\begin{aligned} \pi \vee \pi_\infty &= (c_1\mathbf{t}_1 + c_2\mathbf{t}_2 + c_3\mathbf{t}_3 + c_4\mathbf{t}_4) \vee \pi_\infty \\ &= c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 \quad , \end{aligned}$$

which is a line on the plane at infinity.
The vector $\begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}^T$ formed by the coefficients of the finite components of the plane identifies a line on π_∞. Using the duality principle, this vector can also be interpreted as a point on π_∞. This point represents the direction of all the lines that are orthogonal to π, therefore it characterizes the orientation of π. This is in accordance to what happens with lines, as their component at infinity provides us with information on its orientation. As orthogonal directions imply orthogonal planes, we can impose the orthogonality constraint of the i-th camera planes, directly on the columns of the matrix

$$\mathbf{R}^{(i)} = \begin{bmatrix} c_{11}^{(i)} & c_{21}^{(i)} & c_{31}^{(i)} \\ c_{12}^{(i)} & c_{22}^{(i)} & c_{32}^{(i)} \\ c_{13}^{(i)} & c_{23}^{(i)} & c_{33}^{(i)} \end{bmatrix} \quad .$$

Since rotating camera planes corresponds to rotating camera axes of the same amount, this is the matrix that rotates the first (normalized) camera into the i-th (normalized) camera, $i = 2, 3$. Being a rotation matrix, $\mathbf{R}^{(i)}$ can be represented by a quaternion, exaclty as done in the bifocal case. This reduces the number of unknowns from nine to four thus leading to a more efficient and robust parametrization of the rotation. In addition, the orthogonality constraint on the elements of $\mathbf{R}^{(i)}$ becomes a simple normalization constraint (11) on the equivalent quaternion.

3.3. Calibration in the trifocal case

If we remove the assumption that the focal lengths are equal to one, the equation that describes a back-projected line as a linear combination of camera planes, becomes

$$\pi_{\mathbf{l}} = l_1 f \pi_1 + l_2 f \pi_2 + l_3 \pi_3 \quad . \tag{22}$$

Let $f^{(1)}$, $f^{(2)}$ and $f^{(3)}$ be the focal lenghts of the first, second and third camera, respectively. Following the same steps as in the "normalized" case of the previous Section, we can start from eq. (22) and prove that

$$\begin{aligned} \hat{\tau}_1 &= \begin{bmatrix} (\frac{f^{(2)}f^{(3)}}{f^{(1)}})\tau_{111} & (\frac{f^{(2)}f^{(3)}}{f^{(1)}})\tau_{112} & (\frac{f^{(2)}}{f^{(1)}})\tau_{113} \\ (\frac{f^{(2)}f^{(3)}}{f^{(1)}})\tau_{121} & (\frac{f^{(2)}f^{(3)}}{f^{(1)}})\tau_{122} & (\frac{f^{(2)}}{f^{(1)}})\tau_{123} \\ (\frac{f^{(3)}}{f^{(1)}})\tau_{131} & (\frac{f^{(3)}}{f^{(1)}})\tau_{132} & (\frac{1}{f^{(1)}})\tau_{133} \end{bmatrix} \\ \hat{\tau}_2 &= \begin{bmatrix} (\frac{f^{(2)}f^{(3)}}{f^{(1)}})\tau_{211} & (\frac{f^{(2)}f^{(3)}}{f^{(1)}})\tau_{212} & (\frac{f^{(2)}}{f^{(1)}})\tau_{213} \\ (\frac{f^{(2)}f^{(3)}}{f^{(1)}})\tau_{221} & (\frac{f^{(2)}f^{(3)}}{f^{(1)}})\tau_{222} & (\frac{f^{(2)}}{f^{(1)}})\tau_{223} \\ (\frac{f^{(3)}}{f^{(1)}})\tau_{231} & (\frac{f^{(3)}}{f^{(1)}})\tau_{232} & (\frac{1}{f^{(1)}})\tau_{233} \end{bmatrix} \quad (23) \\ \hat{\tau}_3 &= \begin{bmatrix} (f^{(2)}f^{(3)})\tau_{311} & (f^{(2)}f^{(3)})\tau_{312} & f^{(2)}\tau_{313} \\ (f^{(2)}f^{(3)})\tau_{321} & (f^{(2)}f^{(3)})\tau_{322} & f^{(2)}\tau_{323} \\ f^{(3)}\tau_{331} & f^{(3)}\tau_{332} & \tau_{333} \end{bmatrix} \quad , \end{aligned}$$

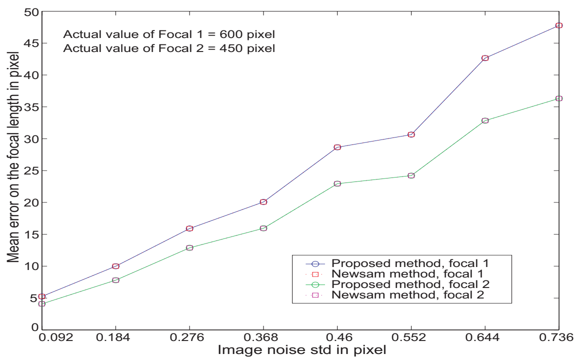

Fig. 3. Average error in focals estimation in the case of cameras with different focal lengths. The proposed methods and the Newsam method produce almost exactly the same results for every noise level, in fact only two lines, one for the first and one for the second focal, are visible instead of four.

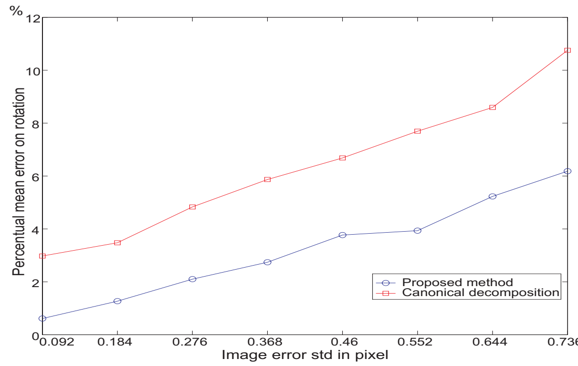

Fig. 4. Average rotational error in the case of cameras with arbitrary focal lengths. Rotational error is measured as the magnitude of the Euler vector that describes the estimated rotation matrix. The proposed method exhibits a more stable behavior, and its accuracy turns out to be better than in the case of canonical decomposition of the essential matrix.

where $\hat{\tau}_i \; i = 1..3$ are "slice" matrices of the trifocal tensor in the case of unknown focal lenghts, while τ_{rst} are the element of the trifocal tensor in the "essential" (normalized) case. The equations (23) play the same role as eq. (14) in the bifocal case. As eq. (20) holds true also in the case of unknown focal lenghts, the variables $f^{(1)}$, $f^{(2)}$ and $f^{(3)}$ can be simply added as unknowns to the systems of equations (20) and (21), which provide enough constraints to retrieve position and orientation of the second and third camera, as well as the focal lenghts.

4. SIMULATION RESULTS

A series of experiments have been conducted on noisy image coordinates of clouds of points with the goal of comparing the proposed solution with existing others. First we compared the focal lengths and rotations that we estimated with our approach in the bifocal case, with the same parameters derived with Newsam's method ([5]) and the canonical decomposition of the Essential matrix (see [1]), respectively. Finally, we compared the performance of our 2-view method with the 3-view method that we proposed in Section 3.3, in order to assess the impact of the third view on the global accuracy.

In figure (3) we can see how Newsam's method and the proposed method produce almost exactly the same results for every noise level on the images. However, figure (4) shows that the rotation retrieved with the proposed method is more accurate than the one obtained using the canonical composition of E. This can be explained considering that our method computes simultaneously both focal lengths and rotation, distributing more efficiently the noise on these quantities, and it employees quaternions to represent rotations, that are intrinsically less sensitive to noise than rotation matrices. The proposed method has also been tested in the case of cameras with fixed focal lengths, and has produced results comparable with the previous case. The experiments confirmed our method's accuracy to be comparable with state-of-the art methods in the literature. Our method, however, enables the estimation of both intrinsic and extrinsic camera parameters simultaneously and with no ambiguities.

In the three-view calibration case, the comparison is done between bifocal estimation and trifocal estimation. The two-view estimation, however, incorporates a preliminary step in which a trifocal tensor is computed in order to make a more robust estimation of the Fundamental matrices. This means that what we measure now is what we gain from computing focal lengths directly from the trifocal tensor. Fig. 5 shows the average relative errors in the estimation of a fixed focal length, obtained when using the two-view and the three-view methods. The three-view calibration method turns out to be more stable and more accurate than the other one, as expected.

5. CONCLUSIONS

In this paper we proposed a novel geometric interpretation of essential and rotation matrices, and of the trifocal tensor in terms of bivectors and trivectors in geometric algebra. From this parametrization we derived two procedures, one for the bifocal case and one for the trifocal case, for computing unknown focal lengths as well as camera positions and orientations, without introducing projective ambiguities.

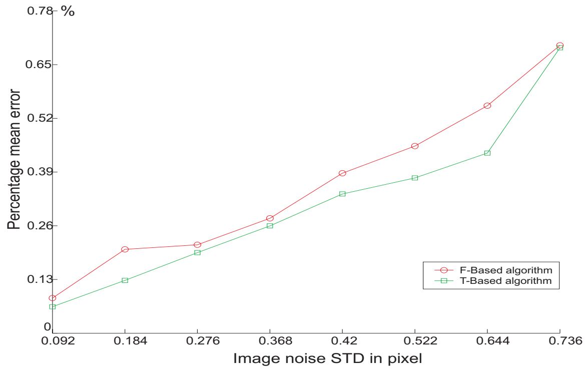

Fig. 5. Average relative error on the focal length estimation in the case of three cameras with fixed focal length. The calibration method based on three views results as being more accurate and stable than in the two-view case for all noise levels.

6. REFERENCES

[1] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge Univ. Press, 2000.

[2] E. Baryo-Corrachano, J. Lasenby, G. Sommer, "Geometric Algebra: A framework for computing point and line correspondances and projective structure using *n* uncalibrated cameras". *ICPR-96*, Vienna, Aug. 1996.

[3] O. Faugeras, Q. Luong, S. Maybank, "Camera self-calibration: theory and experiments". *ECCV-92, LNCS 588, pp. 321-334,* 1992.

[4] Y. Ma, K. Huang, R. Vidal, J. Kosecka, S. Sastry, "Rank conditions on the multiple view matrix". *Kluwer Academic Publishers*, 2001.

[5] G. Newsam, D. Q. Huynh, M. Brooks, H. P. Pan, "Recovering unknown focal lenghts in self-calibration: an essentially linear algorithm and generate configurations". *Intl. Arch. Photogrammetry & Remote Sensing*, Vol. XXXI-B3, pp. 575-80, 1996.