NEW PERSPECTIVES ON CAMERA CALIBRATION USING GEOMETRIC ALGEBRA

C. Defferara *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 [2, 3]. 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 [4] (GA) is currently gaining more and more of the interest of researchers in computer vision [5] 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.

2. A NOVEL GA FRAMEWORK

Adopting the same notation used in [5], a generic point p of the projective space $\mathbb{P}^3$ can be written in homogeneous form as $p = a_1e_1 + a_2e_2 + a_3e_3 + e_4$, where e_1, e_2, e_3, e_4 form a base of $\mathbb{P}^3$. The line l passing through a given pair of points p_1 and p_2 can be expressed as a bivector of the form $l = p_1 \wedge p_2$, 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 p_1, p_2 and p_3 can be written as the grade-3 blade $\pi = p_1 \wedge p_2 \wedge p_3$.

Another important issue is to test whether two subspaces are incident. A general condition for the incidence of two subspaces A and B is given in geometric algebra as $A \cdot B^* = 0$, which becomes $A \wedge B = 0$ when the grade of $A \wedge B$ 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 l_1 and l_2 are found to intersect in a point p if and only if $l_1 \wedge l_2 = 0$. This allows us to formulate of the epipolar constraint in quite a straightforward fashion. Let c_1 and c_2 be the centers of the cameras and p_1 and p_2 be the projections (world coordinates) of a point p onto the first and second camera, respectively. The epipolar constraint can be written as

$$(c_1 \wedge p_1) \wedge (c_2 \wedge p_2) = 0 \quad . \tag{1}$$

A simple pin-hole camera model is completely specified by an optical center c, a focal length f and the directions of the camera axes x_1, x_2 and x_3. Under these assumptions, a point of homogeneous image coordinates $m = [m_1, m_2, m_3]^T$, with $m_3 = 1$, turns out to be expressed as $p = m_1x_1 + m_2x_2 + m_3fx_3 + c$ in the world coordinate frame. If we consider two different views of the same point p, of homogeneous coordinates $m = [m_1, m_2, m_3]^T$ and $n = [n_1, n_2, n_3]^T$, eq. (1) can be specialized as follows

$$\begin{aligned}(m_1 (c_1 \wedge x_1) + m_2 (c_1 \wedge x_2) + m_3 f_1 (c_1 \wedge x_3)) \wedge \\ (n_1 (c_2 \wedge y)_1 + n_2 (c_2 \wedge y_2) + n_3 f_2 (c_2 \wedge y_3)) = 0 \, ,\end{aligned} \tag{2}$$

where f_1, c_1, x_i are the parameters of the first camera and f_2, c_2, y_i 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 $m_in_j\varepsilon_{ij}I_4$, where ε_{ij} are unknown scalars, therefore the epipolar constraint takes on the form $\sum_{i,j=1..3} m_in_j\varepsilon_{ij}I_4 = 0$. This expression, after eliminating I_4, can be written in matrix form as

$$m^T En = 0 \tag{3}$$

where E is the 3×3 matrix of the coefficients ε_{ij}, which is the classical formulation of the epipolar constraint where 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 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 [3].

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

$$l = a_1 l_1 + a_2 l_2 + a_3 l_3 + b_1 \widehat{l_1} + b_2 \widehat{l_2} + b_3 \widehat{l_3} \qquad (4)$$

where $l_1 = e_2 \wedge e_3$, $l_2 = e_3 \wedge e_1$, $l_3 = e_1 \wedge e_2$, $\widehat{l_1} = e_4 \wedge e_1$, $\widehat{l_2} = e_4 \wedge e_2$ and $\widehat{l_3} = e_4 \wedge e_3$. 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 $b_1 \widehat{l_1} + b_2 \widehat{l_2} + b_3 \widehat{l_3}$ passing through the origin of the world reference frame ("finite" component);
- a line $a_1 l_1 + a_2 l_2 + a_3 l_3$ 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 a_i and b_i can be obtained by computing the inner product between the line l and the corresponding base bivector, l_i or $\widehat{l_i}$. For example, we have

$$\begin{aligned} l \cdot l_i &= (a_1 l_1 + a_2 l_2 + a_3 l_3 + b_1 \widehat{l_1} + b_2 \widehat{l_2} + b_3 \widehat{l_3}) \cdot l_i \\ &= a_i l_i \cdot l_i = -a_i \quad . \end{aligned} \qquad (5)$$

A camera with center c, and axis directed as x_1, x_2 and x_3 can be represented by the three lines $c \wedge x_1$, $c \wedge x_2$ and $c \wedge x_3$ 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 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. $x_1 = e_1, x_2 = e_2, x_3 = e_3$ and $c_1 = e_4$. With this assumption, we can rework the epipolar constraint (2) to obtain nine equations of the form

$$\varepsilon_{ij} I_4 = (e_4 \wedge e_i) \wedge (c_2 \wedge y_j) \qquad (6)$$

all involving the quadrivector I_4. As we can see, there are three equations for each axis y_j, whose unknowns are both ε_{ij} and the axes $c_2 \wedge y_j$ of the second camera. If we compute the inner product between both sides of eq. (6) and the bivector l_i, we obtain $\varepsilon_{ij} I_4 \cdot l_i = (\widehat{l_i} \wedge (c_2 \wedge y_j)) \cdot l_i$. Using the known equalities $I_4 \cdot l_i = \widehat{l_i}$, and $(A \wedge B) \cdot C = A \cdot (B \cdot C)$, we can write

$$\varepsilon_{ij} \widehat{l_i} = \widehat{l_i} \cdot ((c_2 \wedge y_j) \cdot l_i) \quad . \qquad (7)$$

Notice that the term $(c_2 \wedge y_j) \cdot l_i$ in the right-hand side of eq. (7) is a scalar, therefore we can write $\varepsilon_{ij} = (c_2 \wedge y_j) \cdot l_i$. As shown in eq. (5), the inner product between a bivector l and the base bivector l_i at infinity, returns the relative coefficient a_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 l_i of the camera-2 axis y_j. 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

$$c_2 \wedge y_i \quad = \quad -E_i^T l - R_i^T \widehat{l}, \quad i = 1, 2, 3, \qquad (8)$$

where $E_j = \begin{bmatrix} \varepsilon_{1j} & \varepsilon_{2j} & \varepsilon_{3j} \end{bmatrix}^T$, $j = 1, ..., 3$, are the columns of E; the vectors $R_j = \begin{bmatrix} r_{1j} & r_{2j} & r_{3j} \end{bmatrix}^T$, $j = 1, ..., 3$, collect the unknowns; while l and $\widehat{l}$ are defined as $l = \begin{bmatrix} l_1 & l_2 & l_3 \end{bmatrix}^T$ and $\widehat{l} = \begin{bmatrix} \widehat{l_1} & \widehat{l_2} & \widehat{l_3} \end{bmatrix}^T$. We will now prove that 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 $\pi_\infty = e_1 \wedge e_2 \wedge e_3$, which can be written as $(\pi_\infty \cdot I_4) \cdot l = l \cdot \pi_\infty^*$. We can write

$$\begin{aligned} (a_1 l_1 + a_2 l_2 + a_3 l_3 + b_1 \widehat{l_1} + b_2 \widehat{l_2} + b_3 \widehat{l_3}) \cdot (-e_4) = \\ -b_1 \widehat{l_1} \cdot e_4 - b_2 \widehat{l_2} \cdot e_4 - b_3 \widehat{l_3} \cdot e_4 = \\ b_1 e_1 + b_2 e_2 + b_3 e_3 \quad . \end{aligned}$$

Also, eq. (8) implies that the directions y_1, y_2, y_3 of the camera-2 axes can be written as a function of the directions $x_1 = e_1, x_2 = e_2, x_3 = e_3$ of the camera-1 axes

$$\begin{aligned} y_1 &= -r_{11} e_1 - r_{21} e_2 - r_{31} e_3 \\ y_2 &= -r_{12} e_1 - r_{22} e_2 - r_{32} e_3 \\ y_3 &= -r_{13} e_1 - r_{23} e_2 - r_{33} e_3 \quad . \end{aligned}$$

It is now quite apparent that $R = [\ R_1 \quad R_2 \quad R_3\ ]^T$ is, in fact, the rotation matrix of the second camera.

3. RETRIEVING THE SECOND CAMERA

We now have enough tools to derive an alternative formulation of the self-calibration problem. The essential matrix 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 c_2. This leads to the following pairwise-incident conditions

$$\begin{array}{rcl} (c_2 \wedge y_1) \wedge (c_2 \wedge y_2) & = & 0 \\ (c_2 \wedge y_1) \wedge (c_2 \wedge y_3) & = & 0 \\ (c_2 \wedge y_2) \wedge (c_2 \wedge y_3) & = & 0 \quad , \end{array} \tag{9}$$

which can be rewritten as

$$\left\{ \begin{array}{l} E_1^T R_2 + E_2^T R_1 = 0 \\ E_1^T R_3 + E_3^T R_1 = 0 \\ E_2^T R_3 + E_3^T R_2 = 0 \end{array} \right. . \tag{10}$$

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 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 $Q = a + bl_1 + cl_2 + dl_3$, which has a scalar component a and a bivector component $bl_1 + cl_2 + dl_3$, subjected to the normalization constraint $a^2 + b^2 + c^2 + d^2 = 1$. Incidentally, the bivector component $bl_1 + cl_2 + dl_3$ only involves bivectors at infinity in the projective space $\mathbb{P}^3$. Represent rotations with rotors, the orthonormal constraint on R is automatically satisfied.

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

$$\begin{array}{rcl} R_1 & = & [\ a^2 - d^2 - c^2 + b^2 \quad 2bc + 2ad \quad 2bd - 2ac\ ]^T \\ R_2 & = & [\ 2bc - 2ad \quad -b^2 + a^2 + c^2 - d^2 \quad 2ab + 2cd\ ]^T \\ R_3 & = & [\ 2bd + 2ac \quad -2ab + 2cd \quad -c^2 - b^2 + a^2 + d^2\ ]^T . \end{array}$$

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

$$E_1^T R_1 = 0, \quad E_2^T R_2 = 0, \quad E_3^T R_3 = 0 . \tag{11}$$

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 a_i, $i = 1, \ldots 3$, and the coefficients b_i, $i = 1, \ldots 3$ of the "finite" base bivectors, must satisfy the consistency constraint $a_1 b_1 + a_2 b_2 + a_3 b_3 = 0$. This result can also be proven using classical tools of geometric algebra. This is the additional constraint mentioned in Section 2.1, which reduces the notational redundancy of the bivector representations. Eqs. (10) and (11) can be expressed in terms of $[\ a \quad b \quad c \quad d\ ]^T$. Along with the normalization constraint on quadrivectors we end up with a nonlinear system of seven equations in four unknowns. As 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 (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).

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 $m^T F n = 0$, where 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 E.

As we know, the relationship between the essential matrix E and the fundamental matrix F is

$$E = K_2^T F K_1 \tag{12}$$

where K_1 =diag$(f_1, f_1, 1)$ and 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. (12) can be expanded as

$$E = \left[\begin{array}{ccc} f_{11} & f_{12} & f_{13}/f_2 \\ f_{21} & f_{22} & f_{23}/f_2 \\ f_{31}/f_1 & f_{32}/f_1 & f_{33}/(f_1 f_2) \end{array} \right] .$$

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), (11) 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 $[\ a \quad b \quad c \quad d \quad f_1 \quad f_2\ ]$. 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.

5. 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. Specifically, we compared the focal lengths and rotations estimated through our solution with the ones obtained respectively with the Newsam method [2] and the canonical decomposition of the Essential matrix described in [1]. As we can

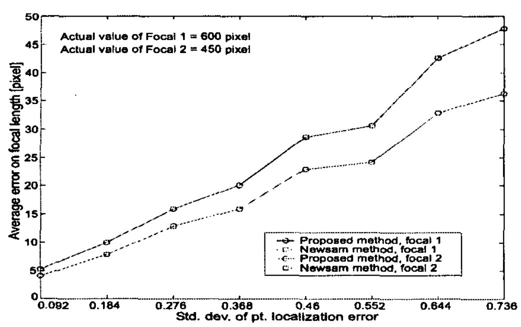

Fig. 1. Average estimation error on (variable) focal length. The proposed method and the Newsam method produce almost exactly the same results for every noise level, in fact only two curves (for the first and the second focal distance) are visible instead of four.

see in Figure 1, the Newsam method and the proposed one produce almost exactly the same results at the various noise levels on the images. However, Figure 2 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 focal lengths and rotations simultaneously, with the result of "distributing" the impact of noise on such quantities more evenly. In addition, it implicitly represents rotations through rotors/quaternions, which are less sensitive to noise than rotation matrices. It is important to stress that this method performs a simultaneous estimation of both intrinsic and extrinsic camera parameters with no ambiguities.

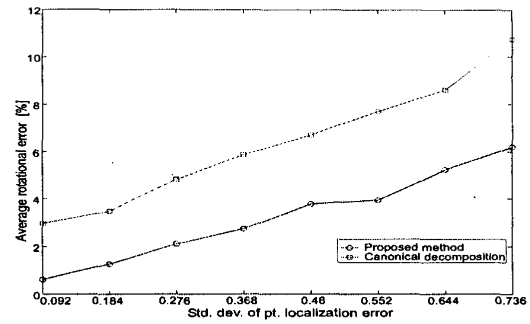

Fig. 2. Average rotational estimation error in the case of cameras with different focal lengths. Rotational distance is measured as the total rotation around the Euler axis.

6. CONCLUSIONS

In this paper we proposed a novel geometric interpretation of essential and rotation matrices in terms of bivectors in geometric algebra. From this parametrization we derived a procedure for computing both focal lengths together with position and orientation of second camera with respect to the first one, without introducing projective ambiguities. The proposed approach seems to open new possibilities in Computer Vision research, as it enables a direct geometric interpretation of incidence relations that involve lines, otherwise not possible in projective geometry.

7. REFERENCES

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

[2] 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.

[3] K. Huang, R. Fossum, Y. Ma, "Generalized Rank Conditions in Multiple View Geometry with Applications to Dynamical Scenes". *ECCV 2002*, Copenhagen, Denmark, May 2002.

[4] W. Baylis, ed., *Clifford (Geometric) Algebras*. Birkhäuser, 1996.

[5] 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.