From santanuc@cse.iitd.ernet.in Wed Nov 11 12:57:16 1998 Date: Wed, 11 Nov 1998 16:42:35 +0530 (IST) From: Santanu Chaudhury ¡santanuc@cse.iitd.ernet.in¿ To: jl@eng.cam.ac.uk Subject: psfig.tex

Geometric Techniques for the Computation of Projective Invariants using n Uncalibrated Cameras

Eduardo Bayro-Corrochano
Computer Science Institute
Christian Albrechts University
Kiel, Germany 24105
edb@informatik.uni-kiel.de

Joan Lasenby
Department of Engineering
Cambridge University
Cambridge, UK CB2 1PZ
jl@eng.cam.ac.uk

Abstract

One goal of computer vision is to automatically recognize objects in real-world scenes. Given the large range of possible viewing conditions, it is often desirable to look for geometric properties of the object which remain invariant under changes in observation parameters. The main contribution of this paper lies in outlining the formation of 3D projective invariants using image points and lines and the trifocal tensor. The experimental analysis indicates that the invariants computed using the trifocal tensor perform better than those formed from bilinearities.

1 Computer Vision using Geometric Algebra

This section aims to outline the basic geometric algebra tools required for the treatment of problems in computer vision. This includes the formulation of projective geometry, the algebra of incidence and the geometry of an image sequence in terms of a mathematical language called *geometric algebra* (GA). The algebra is based on the manipulation of geometric objects – for more detail the reader is referred to [4, 5]. Throughout the paper the summation convention will be used (i.e. repeated indices are summed over) unless explicitly stated otherwise.

1.1 The geometric algebras of the camera and 3D visual space

$\mathcal{G}_n$, the GA of n-D, is a graded linear space in which we have, in addition to vector addition and scalar multiplication, a non- commutative product which is associative and distributive over addition – the *geometric* of *Clifford* product. The symmetric part of the geometric product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, written as $\boldsymbol{ab}$, is the inner (or scalar) product, $\boldsymbol{a}\cdot\boldsymbol{b}$, producing a scalar. The antisymmetric part is the outer (or wedge) product, $\boldsymbol{a}\wedge\boldsymbol{b}$, producing a *bivector* (oriented area). Similarly, $\boldsymbol{a}\wedge\boldsymbol{b}\wedge\boldsymbol{c}$ will produce a *trivector* or oriented volume. The highest grade element in an n-D space wil be an n-vector and this will be unique up to scale and is called the *pseudoscalar* for the space.

In 3-D Euclidean space we have three basis vectors $\{\sigma_i\}$ $i = 1, 2, 3$ which can be extended to produce a basis for the whole GA having $2^3 = 8$ elements given by:

$$\underbrace{1}_{scalar}, \underbrace{\{\sigma_1, \sigma_2, \sigma_3\}}_{vectors}, \underbrace{\{\sigma_1\sigma_2, \sigma_2\sigma_3, \sigma_3\sigma_1\}}_{bivectors}, \underbrace{\{\sigma_1\sigma_2\sigma_3\} \equiv I}_{trivector}. \tag{1}$$

Each basis vector squares to $+1$. It can then easily be verified that the pseudoscalar, $\sigma_1\sigma_2\sigma_3$, squares to -1 and commutes with all elements of the 3-D space. The unit pseudoscalar, I, has a crucial role when discussing duality.

In order to make concrete computations, we require a basis and metric for the 4D projective space (R^4 or P^3) and a means of moving between projective and 3D Euclidean space, $\mathcal{E}^3$. This is achieved using the basis $\{\gamma_i\}$, $i = 1, 2, 3, 4$, which extends to give a basis for the whole GA;

$$\underbrace{1}_{scalar}, \underbrace{\gamma_k}_{4\ vectors}, \underbrace{\gamma_2\gamma_3, \gamma_3\gamma_1, \gamma_1\gamma_2, \gamma_1\gamma_4, \gamma_2\gamma_4, \gamma_3\gamma_4}_{6\ bivectors}, \underbrace{I\gamma_k}_{4\ trivectors}, \underbrace{I}_{pseudoscalar} \tag{2}$$

where $\gamma_4^2 = +1$, $\gamma_k^2 = -1$ for $k = 1, 2, 3$. The pseudoscalar is $I = \gamma_1\gamma_2\gamma_3\gamma_4$ with $I^2 = -1$. The fourth basis vector γ_4 can be seen as a selected direction and is used to project vectors from projective to Euclidean space via an operation called the *projective split*. The role and use of the projective split for a variety of problems involving the algebra of incidence can be found in [6].

1.2 Algebra in projective space

Consider three non-collinear points, P_1, P_2, P_3, represented by vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ in $\mathcal{E}^3$ and by vectors $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ in R^4. The line L_{12} joining points P_1 and P_2 can be expressed in R^4 by the following bivector,

$$L_{12} = \mathbf{X}_1 \wedge \mathbf{X}_2. \tag{3}$$

Any point P, represented in R^4 by $\mathbf{X}$, on the line through P_1 and P_2, will satisfy

$$\mathbf{X} \wedge L_{12} = \mathbf{X} \wedge \mathbf{X}_1 \wedge \mathbf{X}_2 = 0. \tag{4}$$

Similarly, the plane Φ_{123} passing through points P_1, P_2, P_3 is expressed by the following trivector in R^4

$$\Phi_{123} = \mathbf{X}_1 \wedge \mathbf{X}_2 \wedge \mathbf{X}_3. \tag{5}$$

1.2.1 Intersection of a line and a plane

Consider a line $A = \mathbf{X}_1 \wedge \mathbf{X}_2$ intersecting a plane $\Phi = \mathbf{Y}_1 \wedge \mathbf{Y}_2 \wedge \mathbf{Y}_3$ – all vectors are in R^4. The intersection point is expressible using the *meet* operation

$$A \vee \Phi = (\mathbf{X}_1 \wedge \mathbf{X}_2) \vee (\mathbf{Y}_1 \wedge \mathbf{Y}_2 \wedge \mathbf{Y}_3). \tag{6}$$

In GA the meet can be written using the inner product. If U and V are two multivectors then their meet is given by ,[2],

$$U \vee V = (UI^{-1}) \cdot V. \tag{7}$$

After some algebraic manipulations [2] the expression in equation (6) reduces to

$$A \vee \Phi = [\mathbf{X}_1\mathbf{X}_2\mathbf{Y}_2\mathbf{Y}_3]\mathbf{Y}_1 + [\mathbf{X}_1\mathbf{X}_2\mathbf{Y}_3\mathbf{Y}_1]\mathbf{Y}_2 + [\mathbf{X}_1\mathbf{X}_2\mathbf{Y}_1\mathbf{Y}_2]\mathbf{Y}_3 \tag{8}$$

giving the intersection point (vector in R^4).

1.2.2 Intersection of two planes

We now consider the intersection of two planes $\Phi_1 = \mathbf{X}_1 \wedge \mathbf{X}_2 \wedge \mathbf{X}_3$ and $\Phi_2 = \mathbf{Y}_1 \wedge \mathbf{Y}_2 \wedge \mathbf{Y}_3$. The meet of Φ_1 and Φ_2 is given by

$$\Phi_1 \vee \Phi_2 = (\mathbf{X}_1 \wedge \mathbf{X}_2 \wedge \mathbf{X}_3) \vee (\mathbf{Y}_1 \wedge \mathbf{Y}_2 \wedge \mathbf{Y}_3) \tag{9}$$

which after similar algebraic manipulations reads

$$\Phi_1 \vee \Phi_2 = [\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{Y}_1](\mathbf{Y}_2 \wedge \mathbf{Y}_3) + \\ +[\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{Y}_2](\mathbf{Y}_3 \wedge \mathbf{Y}_1) + [\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{Y}_3](\mathbf{Y}_1 \wedge \mathbf{Y}_2), \tag{10}$$

producing a line of intersection (bivector in R^4).

1.2.3 Intersection of two lines

Two lines will intersect at a point only if they are coplanar, this will mean that their representations in R^4, $A = \mathbf{X}_1 \wedge \mathbf{X}_2$, and $B = \mathbf{Y}_1 \wedge \mathbf{Y}_2$ will satisfy

$$A \wedge B = 0. \tag{11}$$

i.e. Any one vector is expressible as a linear combination of the other three vectors. We therefore need to work only in a 2D Euclidean space, $\mathcal{E}^2$, which has an associated 3D projective counterpart, R^3. It can be shown [2] that the intersection point is given by

$$A \vee B = [\mathbf{X}_1\mathbf{X}_2\mathbf{Y}_1]\mathbf{Y}_2 - [\mathbf{X}_1\mathbf{X}_2\mathbf{Y}_2]\mathbf{Y}_1 \tag{12}$$

where the bracket $[\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3]$ in R^3 is understood to mean $(\mathbf{A}_1 \wedge \mathbf{A}_2 \wedge \mathbf{A}_3)I_3{}^{-1}$ with I^3 the pseudoscalar for R^3,

1.3 Geometry of 1, 2 and 3 Views

Consider the monocular case shown in figure 1. Here the image plane is defined by three points $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$, with the optical centre given by $\mathbf{A}_0$ (all vectors in R^4). We are then able to define a bivector basis, $\{L_i^A\}$, $i = 1, 2, 3$, spanning all lines in the image plane $\Phi_A = \mathbf{A}_1 \wedge \mathbf{A}_2 \wedge \mathbf{A}_3$. The *optical planes*, ϕ_i are then given by $\phi_i = \mathbf{A}_0 \wedge L_i^A$. Any plane through $\mathbf{A}_0$ and the image plane can then be expanded in terms of the ϕ_i.

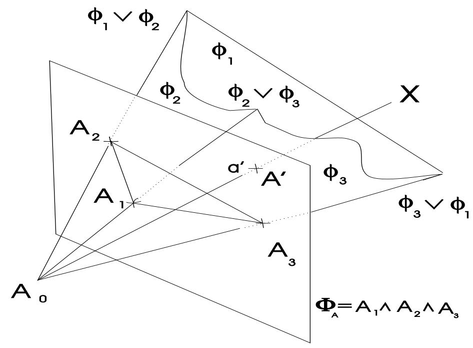

Figure 1: Monocular geometry

In later sections we will compute invariants in 4D and then 'project down' onto 3D – it is therefore necessary to provide an outline of two and three-view geometry in GA. Figure 2 shows that a world point $\mathbf{X}$ projects onto points $\mathbf{A}'$ and $\mathbf{B}'$ in the two image planes. The so called epipoles $\boldsymbol{E}_{AB}$ and $\boldsymbol{E}_{BA}$ correspond to the intersections of the line joining the optical centres with the image planes. Since the points $\mathbf{A}_0, \mathbf{B}_0, \mathbf{A}', \mathbf{B}'$ are coplanar, we can formulate the bilinear constraint as the outer product of these four vectors which must therefore vanish: $\mathbf{A}_0 \wedge \mathbf{B}_0 \wedge \mathbf{A}' \wedge \mathbf{B}'$

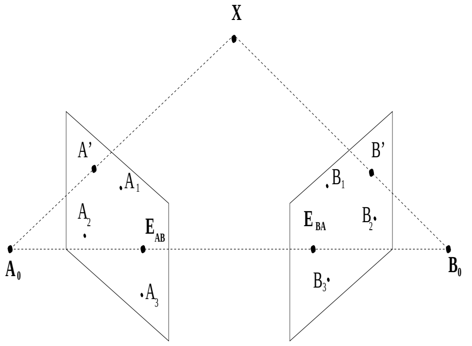

Figure 2: Binocular geometry

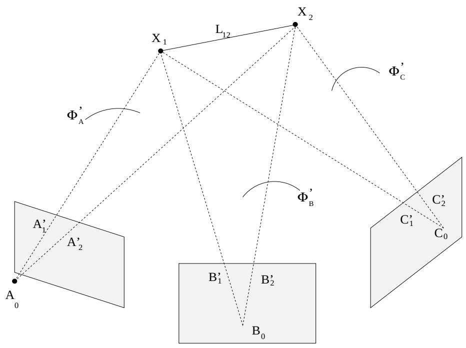

Figure 3: Trinocular geometry

$= 0$. Now, if we let $\mathbf{A}' = \alpha_i \mathbf{A}_i$ and $\mathbf{B}' = \beta_j \mathbf{B}_j$, then this equation can be written as

$$\alpha_i \beta_j \{\mathbf{A}_0 \wedge \mathbf{B}_0 \wedge \mathbf{A}_i \wedge \mathbf{B}_j\} = 0. \tag{13}$$

Defining $\tilde{F}_{ij} = \{\mathbf{A}_0 \wedge \mathbf{B}_0 \wedge \mathbf{A}_i \wedge \mathbf{B}_j\} I^{-1}$ gives us

$$\tilde{F}_{ij} \alpha_i \beta_j = 0 \tag{14}$$

which corresponds to the well-known relationship between the components of the fundamental matrix, F, and the image coordinates. This suggests that F can be seen as a linear function mapping two vectors onto a scalar $F(\mathbf{A}, \mathbf{B}) = \{\mathbf{A}_0 \wedge \mathbf{B}_0 \wedge \mathbf{A} \wedge \mathbf{B}\} I^{-1}$ so that $\tilde{F}_{ij} = F(\mathbf{A}_i, \mathbf{B}_j)$.

In the case of trinocular geometry the so-called trilinear constraint captures the geometric relationships of points and lines between three cameras. Figure 3 shows three image planes Φ_A,Φ_B and Φ_C with bases $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{C}_i$ where i=1,2,3 and optical centres $\mathbf{A}_0, \mathbf{B}_0, \mathbf{C}_0$, and the intersections of two world points $\mathbf{X}_i$ at points $\mathbf{B}'_i, \mathbf{C}'_i$, $i = 1, 2$. The line joining the world points is $L_{12} = \mathbf{X}_1 \wedge \mathbf{X}_2$ and the projected lines are denoted by L'_A, L'_B and L'_C. We firstly define three planes: $\Phi'_A = \mathbf{A}_0 \wedge \mathbf{A}'_1 \wedge \mathbf{A}'_2$, $\Phi'_B = \mathbf{B}_0 \wedge \mathbf{B}'_1 \wedge \mathbf{B}'_2$, and $\Phi'_C = \mathbf{C}_0 \wedge \mathbf{C}'_1 \wedge \mathbf{C}'_2$. It is clear that L_{12} is formed by intersecting Φ'_B and Φ'_C, i.e. $L_{12} = \Phi'_B \vee \Phi'_C = (\mathbf{B}_0 \wedge L'_B) \vee (\mathbf{C}_0 \wedge L'_C)$. If $L_1 = \mathbf{A}_0 \wedge \mathbf{A}'_1$ and $L_2 = \mathbf{A}_0 \wedge \mathbf{A}'_2$, then we can easily see that L_1 and L_2 intersect with L_{12} at $\mathbf{X}_1$ and $\mathbf{X}_2$ respectively. We therefore have $L_1 \wedge L_{12} = 0$ and $L_2 \wedge L_{12} = 0$ which can then be written as

$$(\mathbf{A}_0 \wedge \mathbf{A}'_i) \wedge \{(\mathbf{B}_0 \wedge L'_B) \vee (\mathbf{C}_0 \wedge L'_C)\} = 0 \tag{15}$$

for $i = 1, 2$. This suggests that we should define a linear function T which maps a point and two lines onto a scalar:

$$T(\mathbf{A}, L_B, L_C) = (\mathbf{A}_0 \wedge \mathbf{A}) \wedge \{(\mathbf{B}_0 \wedge L_B) \vee (\mathbf{C}_0 \wedge L_C)\}. \tag{16}$$

Now, using the vector basis of the planes Φ_B and Φ_C it is also possible to define their line basis as follows: $L_1^B = \mathbf{B}_2 \wedge \mathbf{B}_3$, $L_2^B = \mathbf{B}_3 \wedge \mathbf{B}_1$, $L_3^B = \mathbf{B}_1 \wedge \mathbf{B}_2$ etc. So that we can write

$$\mathbf{A} = \alpha_i \mathbf{A}_i, \quad L'_B = l_j^B L_j^B, \quad L'_C = l_k^C L_k^C. \tag{17}$$

If we define the components of a tensor as $\tilde{T}_{ijk} = T(\mathbf{A}_i, L_j^B, L_k^C)$, then if $\mathbf{A}, L'_B, L'_C$ are all derived from projections of the same two world points, equation (15) tells us that we can write

$$\tilde{T}_{ijk} \alpha_i l_j^B l_k^C = 0. \tag{18}$$

This is the trilinear constraint arrived at in [7] using camera matrices. In contrast, here this constraint was produced from purely geometric considerations. Using $\tilde{T}_{ijk}$ we can also relate three projected lines. Consider a projected line in the image plane Φ_A

$$L'_A = \mathbf{A}'_1 \wedge \mathbf{A}'_2 = (\mathbf{A}_0 \wedge L_{12}) \vee \Phi_A. \tag{19}$$

Considering L_{12} as the meet of the planes $\Phi'_B \vee \Phi'_C$ and using the expansions of L'_A, L'_B, L'_C given in equation (17), we can rewrite this equation as

$$l_m^A L_m^A = \Big(\mathbf{A}_0 \wedge l_j^B l_k^C \{(\mathbf{B}_0 \wedge L_j^B) \vee (\mathbf{C}_0 \wedge L_k^C)\}\Big) \vee \Phi_A. \tag{20}$$

Using some standard results, [2], we are able to expand this equation as follows

$$l_i^A L_i^A = [(\mathbf{A}_0 \wedge \mathbf{A}_i) \wedge l_j^B l_k^C \{(\mathbf{B}_0 \wedge L_j^B) \vee (\mathbf{C}_0 \wedge L_k^C)\}] L_i^A \tag{21}$$

where there is no summation over i on the LHS. When we equate coefficients this gives

$$l_i^A = \tilde{T}_{ijk} l_j^B l_k^C \tag{22}$$

which is the familiar expression relating projected lines in the three views.

2 Projective Invariants from Two Uncalibrated Cameras

Given 6 general 3D points whose R^4 and $\mathcal{E}^3$ representations are $\{\mathbf{X}_i, \boldsymbol{x}_i\}$ $i = 1,..,6$, a number of projective invariants can be formed, an example of which is

$$Inv = \frac{[\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{X}_4][\mathbf{X}_4\mathbf{X}_5\mathbf{X}_2\mathbf{X}_6]}{[\mathbf{X}_1\mathbf{X}_2\mathbf{X}_4\mathbf{X}_5][\mathbf{X}_3\mathbf{X}_4\mathbf{X}_2\mathbf{X}_6]}. \quad (23)$$

where $[\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{X}_4] = (\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{X}_4)I^{-1}$ and is equivalent to the definition of the bracket in the Grassmann-Cayley algebra [1]. Let $\mathbf{X}_i$ project down to $\mathbf{A}'_i$ and $\mathbf{B}'_i$ in the two image planes A and B, and $\mathbf{A}'_{ijmn} = (\mathbf{A}'_i \wedge \mathbf{A}'_j) \vee (\mathbf{A}'_m \wedge \mathbf{A}'_n)$, and similarly for $\mathbf{B}'_{ijmn}$. It is shown in [2] that the bracket of 4 world points (in R^4) can be written as

$$S_{1234} = [\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{X}_4] \equiv [\mathbf{A}_0\mathbf{B}_0\mathbf{A}'_{1234}\mathbf{B}'_{1234}]. \quad (24)$$

If we expand this bracket using $\mathbf{A}'_i = \alpha_{ij}\mathbf{A}_j$ and $\mathbf{B}'_i = \beta_{ij}\mathbf{B}_j$) and use our matrix $\tilde{\boldsymbol{F}}$ given in the previous section

$$\tilde{F}_{ij} = [\mathbf{A}_0\mathbf{B}_0\mathbf{A}_i\mathbf{B}_j] \quad (25)$$

we can write

$$Inv = \frac{(\boldsymbol{\alpha}^T{}_{1234}\tilde{\boldsymbol{F}}\boldsymbol{\beta}_{1234})(\boldsymbol{\alpha}^T{}_{4526}\tilde{\boldsymbol{F}}\boldsymbol{\beta}_{4526})}{(\boldsymbol{\alpha}^T{}_{1245}\tilde{\boldsymbol{F}}\boldsymbol{\beta}_{1245})(\boldsymbol{\alpha}^T{}_{3426}\tilde{\boldsymbol{F}}\boldsymbol{\beta}_{3426})} \quad (26)$$

where $\boldsymbol{\alpha}_{1234} = (\alpha_{1234,1}, \alpha_{1234,2}, \alpha_{1234,3})$ and $\boldsymbol{\beta}_{1234} = (\beta_{1234,1}, \beta_{1234,2}, \beta_{1234,3})$ etc.

We now look at what happens when we attempt to express Inv in terms of what we actually observe – the 3D image coordinates and the fundamental matrix calculated *from* these image coordinates. Let us define a matrix $\boldsymbol{F}$ by

$$\tilde{F}_{kl} = (\mathbf{A}_k \cdot \gamma_4)(\mathbf{B}_l \cdot \gamma_4)F_{kl}. \quad (27)$$

If $\boldsymbol{a}'_i = \delta_{ij}\boldsymbol{a}_j$ and $\boldsymbol{b}'_i = \epsilon_{ij}\boldsymbol{b}_j$ ($\boldsymbol{a}'_i$ and $\boldsymbol{b}'_i$ are the $\mathcal{E}^3$ representations of $\mathbf{A}'_i$ and $\mathbf{B}'_i$), then it can be shown [2] that the following relationships hold

$$\alpha_{ij} = \frac{\mathbf{A}'_i \cdot \gamma_4}{\mathbf{A}_j \cdot \gamma_4}\delta_{ij} \quad \text{and} \quad \beta_{ij} = \frac{\mathbf{B}'_i \cdot \gamma_4}{\mathbf{B}_j \cdot \gamma_4}\epsilon_{ij}. \quad (28)$$

Thus we are able to write

$$\alpha_{ik}\tilde{F}_{kl}\beta_{il} = (\mathbf{A}'_i \cdot \gamma_4)(\mathbf{B}'_i \cdot \gamma_4)\delta_{ik}F_{kl}\epsilon_{il}. \quad (29)$$

Now let us look again at Inv. According to the above we can write Inv as

$$Inv = \frac{(\boldsymbol{\delta}^T{}_{1234}\boldsymbol{F}\boldsymbol{\epsilon}_{1234})(\boldsymbol{\delta}^T{}_{4526}\boldsymbol{F}\boldsymbol{\epsilon}_{4526})\phi_{1234}\phi_{4526}}{(\boldsymbol{\delta}^T{}_{1245}\boldsymbol{F}\boldsymbol{\epsilon}_{1245})(\boldsymbol{\delta}^T{}_{3426}\boldsymbol{F}\boldsymbol{\epsilon}_{3426})\phi_{1245}\phi_{3426}} \quad (30)$$

where $\phi_{pqrs} = (\mathbf{A}'_{pqrs} \cdot \gamma_4)(\mathbf{B}'_{pqrs} \cdot \gamma_4)$. We see therefore that the ratio of the terms $\boldsymbol{\delta}^T\boldsymbol{F}\boldsymbol{\epsilon}$ which resembles the expression for the invariant in R^4, but uses only the observed coordinates and the estimated fundamental matrix, will not be an invariant. Instead, we need to include the factors ϕ_{1234} etc., which do not cancel. Since $\boldsymbol{a}'_3$, $\boldsymbol{a}'_4$ and $\boldsymbol{a}'_{1234}$ are collinear we can write $\boldsymbol{a}'_{1234} = \mu_{1234}\boldsymbol{a}'_4 + (1 - \mu_{1234})\boldsymbol{a}'_3$. Then, by expressing $\mathbf{A}'_{1234}$ as the intersection of the line joining $\mathbf{A}'_1$ and $\mathbf{A}'_2$ with the plane through $\mathbf{A}_0, \mathbf{A}'_3, \mathbf{A}'_4$ it is relatively easy to show [2] that we can write

$$\frac{(\mathbf{A}'_{1234} \cdot \gamma_4)(\mathbf{A}'_{4526} \cdot \gamma_4)}{(\mathbf{A}'_{3426} \cdot \gamma_4)(\mathbf{A}'_{1245} \cdot \gamma_4)} = \frac{\mu_{1245}(\mu_{3426} - 1)}{\mu_{4526}(\mu_{1234} - 1)}. \quad (31)$$

The values of μ are readily obtainable from the images. The factors $\mathbf{B}'_{pqrs} \cdot \gamma_4$ are found in a similar way so that if $\boldsymbol{b}'_{1234} = \lambda_{1234}\boldsymbol{b}'_4 + (1 - \lambda_{1234})\boldsymbol{b}'_3$ etc., the overall expression for the invariant becomes

$$Inv = \frac{(\boldsymbol{\delta}^T{}_{1234}\boldsymbol{F}\boldsymbol{\epsilon}_{1234})(\boldsymbol{\delta}^T{}_{4526}\boldsymbol{F}\boldsymbol{\epsilon}_{4526})}{(\boldsymbol{\delta}^T{}_{1245}\boldsymbol{F}\boldsymbol{\epsilon}_{1245})(\boldsymbol{\delta}^T{}_{3426}\boldsymbol{F}\boldsymbol{\epsilon}_{3426})}\eta_1\eta_2$$

$$\eta_1 = \frac{\mu_{1245}(\mu_{3426} - 1)}{\mu_{4526}(\mu_{1234} - 1)} \quad \eta_2 = \frac{\lambda_{1245}(\lambda_{3426} - 1)}{\lambda_{4526}(\lambda_{1234} - 1)}. \quad (32)$$

To conclude: given the coordinates of a set of 6 corresponding points in the two image planes (where these 6 points are projections from arbitrary non-coplanar world points) we can form 3D projective invariants provided we have some estimate of $\boldsymbol{F}$. See [2] for a more detailed discussion of this issue.

3 Projective Invariants from Three Uncalibrated Cameras

The technique used to form the 3D projective invariants for two views can be straightforwardly extended to give expressions for invariants from three views. Consider the scenario shown in figure 3. Suppose we have four world points, $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4\}$ (or two lines $\mathbf{X}_1 \wedge \mathbf{X}_2$ and $\mathbf{X}_3 \wedge \mathbf{X}_4$) projected into three camera planes, where we use the same notation as in Section 2. As before, we can write $\mathbf{X}_1 \wedge \mathbf{X}_2 = (\mathbf{A}_0 \wedge L^A{}_{12}) \vee (\mathbf{B}_0 \wedge L^B{}_{12})$ and $\mathbf{X}_3 \wedge \mathbf{X}_4 = (\mathbf{A}_0 \wedge L^A{}_{34}) \vee (\mathbf{C}_0 \wedge L^C{}_{34})$ (where $L^A_{12} = \mathbf{A}'_1 \wedge \mathbf{A}'_2$ etc.). Once again, we can combine the above expressions to give an equation for the 4-vector $\mathbf{X}_1 \wedge \mathbf{X}_2 \wedge \mathbf{X}_3 \wedge \mathbf{X}_4$:

$$\begin{aligned} &\mathbf{X}_1 \wedge \mathbf{X}_2 \wedge \mathbf{X}_3 \wedge \mathbf{X}_4 = \\ &= [(\mathbf{A}_0 \wedge L^A{}_{12}) \vee (\mathbf{B}_0 \wedge L^B{}_{12})] \wedge [(\mathbf{A}_0 \wedge L^A{}_{34}) \vee (\mathbf{C}_0 \wedge L^C{}_{34})] \\ &= (\mathbf{A}_0 \wedge \mathbf{A}'_{1234}) \wedge [(\mathbf{B}_0 \wedge L^B{}_{12}) \vee (\mathbf{C}_0 \wedge L^C{}_{34})]. \end{aligned} \quad (33)$$

The third line of equation (33) follows from a standard rearrangement result (see [2]). Now, in terms of line and point coordinates we have $L_{12}^B = l_{12,j}^B L_j^B$, $L_{34}^C = l_{34,j}^C L_j^C$ and $\mathbf{A}'_{1234} = \alpha_{1234,i}\mathbf{A}_i$. It has been shown in Section 2 that the components of the trilinear tensor (which plays the role of the fundamental matrix for 3 views), can be written in geometric algebra as

$$\tilde{T}_{ijk} = (\mathbf{A}_0 \wedge \mathbf{A}_i) \wedge [(\mathbf{B}_0 \wedge L_j^B) \vee (\mathbf{C}_0 \wedge L_k^C)] \tag{34}$$

so that equation (33) now reduces to

$$\mathbf{X}_1 \wedge \mathbf{X}_2 \wedge \mathbf{X}_3 \wedge \mathbf{X}_4 = \tilde{T}_{ijk}\alpha_{1234,i} l^B{}_{12,j} l^C{}_{34,k}. \tag{35}$$

The invariant Inv can then be expressed as

$$Inv = \frac{(\tilde{T}_{ijk}\alpha_{1234,i} l^B{}_{12,j} l^C{}_{34,k})(\tilde{T}_{mnp}\alpha_{4526,m} l^B{}_{26,n} l^C{}_{45,p})}{(\tilde{T}_{qrs}\alpha_{1245,q} l^B{}_{12,r} l^C{}_{45,s})(\tilde{T}_{tuv}\alpha_{3426,t} l^B{}_{26,u} l^C{}_{34,v})}. \tag{36}$$

We therefore have an expression for invariants in three views which is a direct extension of that found in the two-view case. Now, as before, when forming the invariants from *observed* quantities, some correction factors will be necessary since equation (36) is given in terms of R^4 quantities.

It is straightforward to show [2] that it is only the point coordinates which require scaling. If we now take $l_{ij,k}^A$ to represent the *observed* line coordinates of the line joining $\mathbf{A}'_i$ and $\mathbf{A}'_j$ in image plane A (and similarly for $l_{ij,k}^B$ etc.), and use $\boldsymbol{a}'_{1234} = \delta_{1234,i}\boldsymbol{a}_i$ the invariant is given by

$$Inv = \frac{[T_{ijk}\delta_{1234}^A l_{12}^B l^C{}_{34}][T_{ijk}\delta_{4526}^A l_{45}^B l^C{}_{26}](\mathbf{A}'_{1234}\cdot\gamma_4)(\mathbf{A}'_{4526}\cdot\gamma_4)}{[T_{ijk}\delta_{1245}^A l_{12}^B l^C{}_{45}][T_{ijk}\delta_{3426}^A l_{34}^B l^C{}_{26}](\mathbf{A}'_{1245}\cdot\gamma_4)(\mathbf{A}'_{3426}\cdot\gamma_4)}. \tag{37}$$

Here we have defined $\tilde{T}_{ijk} = (\mathbf{A}_i\cdot\gamma_4)(\mathbf{B}'_{j1}\cdot\gamma_4)(\mathbf{B}'_{j2}\cdot\gamma_4)(\mathbf{C}'_{k1}\cdot\gamma_4)(\mathbf{C}'_{k2}\cdot\gamma_4)T_{ijk}$ for $j1, j2 = 1, 2$ if $j = 3$ etc., and similarly for $k1, k2$. The terms containing the $\mathbf{A}'$s represent the *correction factor* and are given (following the result in the previous section) by

$$\frac{\mu_{1245}(\mu_{3426} - 1)}{\mu_{4526}(\mu_{1234} - 1)}. \tag{38}$$

In equation (37) the quantities are all *observed* quantities or entities we form from *observed* quantities. Tests in MAPLE show that the above expression is indeed invariant.

4 Experiments

In this section we present computations on both synthetic and real data. In each case the estimates of the fundamental matrix and the trifocal tensor were made using simple linear methods (no constraints were enforced) – this is adequate for the purposes of this paper.

Four different sets of six points $S_i = \{\boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, \boldsymbol{X}_{i3}, \boldsymbol{X}_{i4}, \boldsymbol{X}_{i5}, \boldsymbol{X}_{i6}\}$, i=1,..,4, were considered in the simulation and these were viewed from several different camera positions. The three possible invariants, $\{I_{1,i}, I_{2,i}, I_{3,i}\}$ (here the invariants are the three linearly independent versions of Inv which arise from permuting the order of the points), were computed for each set. These invariants of each set were represented as 3D vectors, $\mathbf{v}_i = [I_{1,i}, I_{2,i}, I_{3,i}]^T$. The comparison of the invariants was done using Euclidean distances of the vectors $d(\boldsymbol{v}_i, \boldsymbol{v}_j) = \left[1 - \left|\frac{\boldsymbol{v}_i\cdot\boldsymbol{v}_j}{||\boldsymbol{v}_i||||\boldsymbol{v}_j||}\right|\right]^{\frac{1}{2}}$. For any $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ the distance $d(\mathbf{v}_i, \boldsymbol{v}_j)$ lies between 0 and 1 and it does not vary when $\boldsymbol{v}_i$ or $\boldsymbol{v}_j$ is multiplied by a nonzero constant – this follows Hartley's analysis given in [3].

Figure 4 shows two sets of tables. The (i, j)th entry in the top left-hand box shows $d(\boldsymbol{v}_i, \boldsymbol{v}_j)$ for invariants formed from two views when gaussian noise of $\sigma = 0.005$ was added to the image points. The boxes below this show the same thing for increasing σ. The right-hand column shows the equivalent for invariants formed from three views using the expression given in Section 3. Clearly, we would like the diagonal elements to be as close as possible to zero since the invariants should be the same in all views in the zero noise case. The off-diagonal elements give some indication of the usefulness of the invarinats in distinguishing between sets of points (we would like these to be as close to 1 as possible).

We can see that the performance of the invariants based on trilinearities is rather better than those based on bilinearities.

In the case of real images we use a sequence of images taken by a moving robot equipped with a binocular head. Figure 5 shows an example of images taken with the left and right eyes. Image pairs, one from the left sequence and one from the right sequence were taken to form invariants using F. For the formation of invariants using T, two from the left and one from the right sequence were used. 38 points were semi-automatically taken and 6 sets of 6 general points were selected. The vector of invariants for each set was formed using both F and T and the set of distances found are shown in figure 6. We again see that computing the invariants from 3 views is more robust and useful than computing them from 2 views – one would expect this from a theoretical viewpoint.

5 Conclusions

Here we have reviewed the computation of general 3D projective point invariants from two views and given a novel expression for computing these invari-

Invariants using F

0.000	0.590	0.670	0.460
	0	0.515	0.68
		0.59	0
			0.69

0.063	0.650	0.750	0.643
	0.67	0.78	0.687
		0.86	0.145
			0.531

0.148	0.600	0.920	0.724
	0.60	0.96	0.755
		0.71	0.97
			0.596

0.900	0.838	0.690	0.960
	0.276	0.693	0.527
		0.98	0.59
			0.663

Invariants using T

0.000	0.590	0.310	0.630
	0	0.63	0.338
		0.134	0.67
			0.29

0.044	0.590	0.326	0.640
	0	0.63	0.376
		0.192	0.67
			0.389

0.031	0.100	0.352	0.660
	0.031	0.337	0.67
		0.31	0.67
			0.518

0.000	0.640	0.452	0.700
	0.063	0.77	0.545
		0.321	0.63
			0.643

Figure 4: *The distance matrices show the performance of the invariants with increasing Gaussian noise σ (from top to bottom): 0.005, 0.015, 0.025 and 0.04.*

ants from three views. Using geometric algebra, the formation of these invariant expressions is straightforward. Experimental evidence confirms that the invariants formed from the trifocal tensor exhibit better performance than those formed from the fundamental matrix.

References

[1] Carlsson, S. 1994. The double algebra: an effective tool for computing invariants in computer vision. *Applications of Invariance in Computer Vision*, Lecture Notes in Computer Science 825; Proceedings of 2nd joint Europe-US Workshop, Azores, October 1993. Eds. Mundy, Zisserman and Forsyth. Springer-Verlag.

[2] Lasenby, J. and Bayro-Corrochano, E. Computing Invariants in Computer Vision using Geometric Algebra. Cambridge University Engineering Department Technical Report, CUED/F-INENG/TR.244. 1998

[3] Hartley R.I. 1994. Projective Reconstruction and Invariants from Multiple Images *IEEE Trans.PAMI*, Vol 16, No.10, 1036-1041.

[4] Hestenes, D. and Sobczyk, G. 1984. Clifford Algebra to Geometric Calculus: A unified language for mathematics and physics. *D. Reidel*, Dordrecht.

[5] Hestenes, D. and Ziegler, R. 1991. Projective Geometry with Clifford Algebra. *Acta Applicandae Mathematicae*, 23: 25–63.

[6] Lasenby J. and Bayro-Corrochano E. 1997. Computing 3D Projective Invariants from points and lines. In G. Sommer, K. Daniilidis, and J. Pauli, editors, Computer Analysis of Images and Patterns, 7th Int. Conf., CAIP'97, Kiel, Springer-Verlag, Sept., pp. 82-89.

[7] Shashua, A. and Werman, M. 1995. Trilinearity of three perspective views and its associated tensor. In International Conference on Computer Vision, pp. 920-925, Cambridge, MA , 1995.

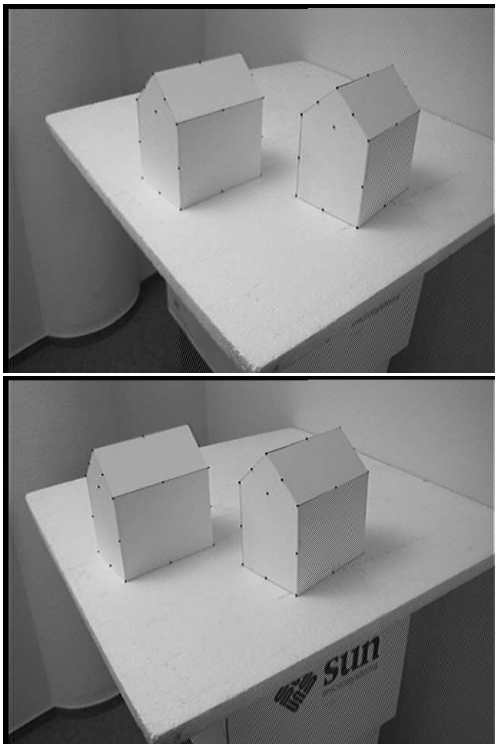

Figure 5: *Examples of left and right images from an image sequence taken during navigation by the binocular head of a mobile robot.*

using F

0.04	0.79	0.646	0.130	0.679	0.89
	0.023	0.2535	0.278	0.268	0.89
		0.0167	0.723	0.606	0.862
			0.039	0.808	0.91
				0.039	0.808
					0.039

using T

0.021	0.779	0.346	0.930	0.759	0.81
	0.016	0.305	0.378	0.780	0.823
		0.003	0.83	0.678	0.97
			0.02	0.908	0.811
				0.008	0.791
					0.01

Figure 6: *The distance matrices show the performance of the computed invariants using bilinearities (top) and trilinearities (bottom) for the real image sequence.*