Duality, rigidity and planar parallax

A. Criminisi, I. Reid and A. Zisserman

Department of Engineering Science, Oxford University
Parks Road, Oxford, OX1 3PJ, UK
tel: +1865 273148
fax: +1865 273908
e-mail: [criminis,ian,az]@robots.ox.ac.uk

Abstract. We investigate the geometry of two views of seven points, four of which are coplanar, and the geometry of three views of six points, four of which are coplanar. We prove that the two are dual, and that the fundamental geometric constraints in each case are encapsulated by a planar homology. The work unifies a number of previously diverse results related to planar parallax, duality and planar homologies.
In addition, we make a number of practical contributions, including formulae for computing the distance of the cameras from a distinguished world plane and formulae for structure computations. We show that the trifocal tensor is obtained uniquely from three views of six points, four of which are coplanar, and give a simple interpretation of the trifocal geometry.
We give examples of these computations on real images.

1 Introduction

In recent work Carlsson and Weinshall *et al.* [2, 3, 18, 19] have demonstrated the fundamental duality of the 3D reconstruction problem. They show that for points and camera in *general position*, the problem of computing camera positions from n points in m views is mathematically equivalent to the problem of reconstructing $m + 4$ points in $n - 4$ views.

In this paper we investigate the case where the points are not in general position, but where four of the space points are coplanar (which we refer to as the plane + points configuration). We show that in this case there exists an additional dual relationship which is described by a *planar homology* [13, 17], which encapsulates the fundamental geometric constraints which can be obtained. A summary of the duality results contrasted with the general position cases is shown in table 1.

The plane + points configuration has received significant attention in the past, not least because it arises frequently in everyday scenes. A useful and popular approach to the problem decomposes the image motion into a planar homographic transfer plus a residual image parallax vector [7, 8, 12]. This decomposition has the advantage that it partially factors out dependence on the camera relative rotation and internal parameters. Furthermore it can be shown

m views	n pts	general position	coplanar
2	7	$3n + 7 = 28$ d.o.f. $2mn = 28$ constraints F determined up to a 3-fold ambiguity no further constraints	$3n - 1 + 7 = 27$ d.o.f. $2mn = 28$ constraints F determined uniquely motion constraint (one) in addition $\exists$ *homology*: maps between views, *vertex* is epipole, i.e. intersection of plane and camera baseline, *axis* is intersection of plane with plane containing remaining three points.
3	6	$3n + 18 = 36$ d.o.f. $2mn = 36$ constraints T determined up to a 3-fold ambiguity no further constraints	$3n - 1 + 18 = 35$ d.o.f. $2mn = 36$ constraints T determined uniquely structure constraint (one) in addition $\exists$ *homology*: maps between points, *vertex* is intersection of plane and line joining the remaining two points, *axis* is intersection of plane and plane containing the camera centres.

Table 1. Camera/point duality results for (i) points in general position and (ii) four points lying on a distinguished plane. The fundamental matrix F has 7 degrees of freedom (d.o.f.) and the trifocal tensor T has 18 d.o.f.

[7] that the relative structure of points, and the rigidity of a scene, can be determined directly from the image measurements – i.e. the parallax vectors – without needing to compute the epipolar geometry. We show that these constraints, and other equivalent and dual ones, are consequences of the planar homology.

In fact, the work here unifies a number of previously diverse results related to planar parallax [7, 8, 12], duality [2, 3, 18, 19] and planar homologies [17]. In addition to this theoretical contribution, we make a number of practical contributions, including formulae for computing the distance of the cameras from a distinguished world plane, formulae for structure computations, and we derive the trifocal tensor [6, 15, 16] in the plane + points case, showing that it is obtained uniquely.

The remainder of the paper is organised as follows. We begin with a discussion of background material – notation, parallax geometry and, planar homologies.

We then turn to the geometry of two views, seven points, four of which are coplanar. We show that there exists a homology on the plane relating the two views and derive necessary conditions for the homology directly in terms of the parallax measurements. In section 4 we show the duality of the geometry of three views, six points (four coplanar) to the two view, seven point case, and hence obtain analogous necessary conditions. We also derive the trifocal tensor and show that it is over-constrained. In section 5 we derive expressions for the height of the cameras from the distinguished plane and the structure of points in terms of affine invariants, and give examples of various applications. We conclude with a discussion and directions for future study in section 6.

2 Background

2.1 Notation

We denote 3D points in general position by upper case bold symbols (e.g. $\mathbf{P}$) and image positions and vectors by lower case bold symbols (e.g. $\mathbf{p}$, $\boldsymbol{\mu}_p$) and scalars by lower case normal symbols (e.g. d, h_p). Matrices are denoted by typewriter style capitals (e.g. $\mathtt{A}$, $\mathtt{B}$).

The area of a triangle on a plane with vertices $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$ is denoted $A_{\mathbf{pqr}}$, and can be determined via the formula $A_{\mathbf{pqr}} = \frac{1}{2}|\mathbf{pqr}|$ where the points $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$ are represented as homogeneous 3-vectors with last component equal to one.

Numbered subscripts are used to distinguish different views, with the first camera centre given by $\mathbf{O}_1$, the second by $\mathbf{O}_2$ and the third by $\mathbf{O}_3$. The projection of an image point onto the distinguished world plane from the i^{th} view is denoted $\mathbf{p}_i$.

2.2 Planar parallax

The underlying parallax geometry is shown in figure 1. The distinguished world plane induces a homography between the views meaning that the images of points on the plane can be transferred via the homography between views 1 and 2. The homography can be determined from a minimum of four correspondences in the two views of points (or lines) on the distinguished plane [11].

The parallax vector in the first view is the vector joining the image of a world point $\mathbf{P}$ with the transferred location of $\mathbf{P}$'s image in the other view (i.e. the image of $\mathbf{p}_2$). Furthermore, since the three planes (distinguished world plane and two image planes) are equivalent up to a plane projectivity, we can also measure parallax in the second view, or – if we know the image to world plane homographies – on the distinguished world plane. In fact it is particularly elegant to work with the world plane. In this case all dependence on the rotational and internal parameters of the cameras is removed (aggregated into the image plane to world plane homographies) leaving only a dependence on the camera centres.

Since the clarity of the underlying geometry is greatly increased, we depict all relevant points and vectors on the world plane in all of our figures. However

the computations in general do *not* require the image to world homographies to be known.

The parallax vector is directed towards (or away from) the epipole, so two such vectors are sufficient to compute its position, and the full epipolar geometry follows [1, 9, 10]. The magnitude of the parallax vector is related to the distance of the world point and cameras from the world plane. Although others have described this function in detail [7, 8, 12], we re-derive the relationship in section 5.

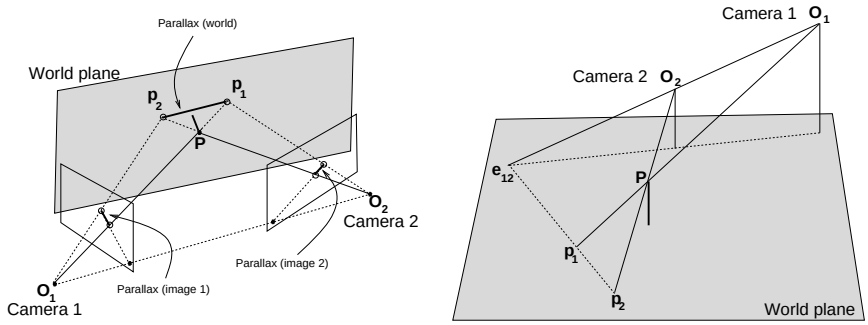

Fig. 1. Parallax geometry: (a) general configuration; (b) viewed on the distinguished plane. The parallax vector $< \mathbf{p}_1, \mathbf{p}_2 >$ passes through the epipole $\mathbf{e}_{12}$.

2.3 Planar homologies

A *planar homology* is a plane projective transformation with five degrees of freedom, having a line of fixed points, called the *axis* and a distinct fixed point not on the axis known as the *vertex* (figure 2). Algebraically, such a transformation has one distinct eigenvalue, with corresponding eigenvector being the vertex, and two repeated eigenvalues, whose corresponding eigenvectors span the axis. Planar homologies arise naturally in an image when two planes related by a perspectivity in 3-space are imaged [17].

If two triangles on a plane are related such that the lines joining their corresponding vertices are concurrent, then they are said to be in a *Desargues configuration*, and Desargues' Theorem states that the intersections of their corresponding sides are collinear [13] (figures 4, 6). Such triangles are related by a planar homology, with the common point of intersection being the vertex of the transformation, and the axis being the line containing the intersections of corresponding sides. Conversely, any triple of points in correspondence under a homology must be in a Desargues configuration.

The projective transformation representing the homology can be parametrized directly in terms of the 3-vector representing the axis $\mathbf{a}$, the 3-vector representing

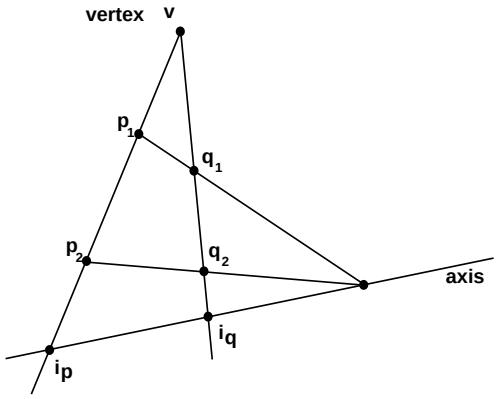

Fig. 2. A planar homology is defined by a vertex and an axis. Its characteristic invariant is given by the cross-ratio $< \mathbf{v}, \mathbf{p_1}, \mathbf{p_2}, \mathbf{i}_p >$ where $\mathbf{p_1}$ and $\mathbf{p_2}$ are *any* pair of corresponding points and $\mathbf{i}_p$ is the intersection of the line through $\mathbf{p_1}$ and $\mathbf{p_2}$ and the axis. The point $\mathbf{p_1}$ is projected onto the point $\mathbf{p_2}$ under the homology, and similarly for $\mathbf{q}_1$ and $\mathbf{q}_2$.

the vertex $\mathbf{v}$, and the characteristic cross-ratio μ as:

$$\mathtt{H} = \mathtt{I} + (\mu - 1)\frac{\mathbf{v}\mathbf{a}^\top}{\mathbf{v}.\mathbf{a}}$$

Having five degrees of freedom (the scales of $\mathbf{v}$ and $\mathbf{a}$ have no effect), a homology can be determined by 2.5 point correspondences. Three point correspondences therefore, provide an additional constraint. In the next section we derive the link between homologies and the structure and motion.

3 Geometry of two views

We consider the case of imaging seven points, four of which are coplanar from two distinct viewpoints. Each of the three points $\mathbf{P}$,$\mathbf{Q}$ and $\mathbf{R}$ not on the plane gives rise to a parallax vector, which is depicted on the world plane in figure 3.

The plane $\mathbf{PQR}$ intersects the world plane in a line, and the camera baseline intersects the world plane in a point. It can be seen by inspection of figures 3 and 4 that the geometry under consideration (seven points, two views) leads directly to a Desargues configuration in which the epipole is the vertex of the homology and the intersection of plane $\mathbf{PQR}$ with the world plane is the axis of the homology. The two triangles in the Desargues configuration are the two images of the space triangle $\mathbf{PQR}$. This key observation underpins the results which follow.

As stated in the previous section, a homology has five degrees of freedom, and therefore three point correspondences over-determine the homology. The extra constraint available can be used as a test for the rigidity of the scene and is equivalent to the epipolar constraint.

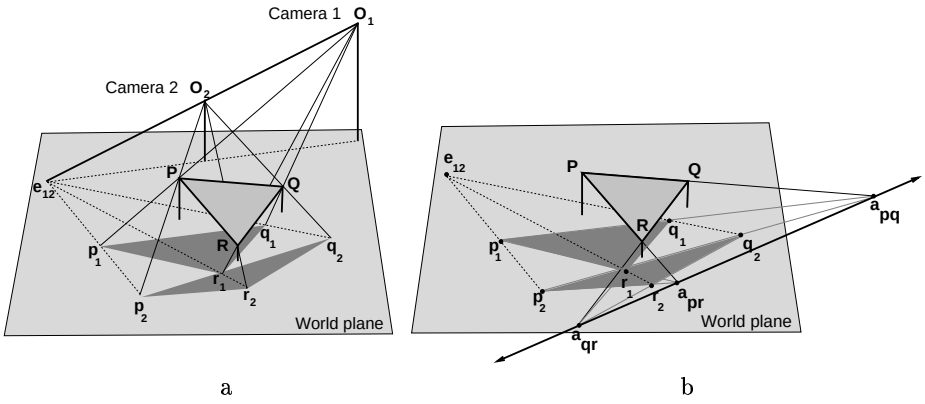

Fig. 3. (a) The geometry of three points in two views. The triangle $\mathbf{p}_i\mathbf{q}_i\mathbf{r}_i$ is the "shadow" of $\mathbf{PQR}$ under the camera $\mathbf{O}_i$; (b) The axis of the homology is given by the intersection of the plane $\mathbf{PQR}$ with the world plane, and the vertex (epipole e_{12}) by the intersection of the baseline with the world plane.

Clearly the constraint can be tested geometrically by using point correspondences either to construct the intersections of corresponding sides and testing their collinearity, or testing the concurrence of the parallax vectors. Alternatively an algebraic test could, for example, compute the epipole using two point correspondences, use the epipole plus the three point correspondences to solve for a general homography, then test the homography to determine if it is a homology.

The former has the disadvantage of requiring the construction of features which may be far removed from the measured image features themselves, while the latter gives little insight into the underlying geometry.

Below we derive novel bilinear and trilinear constraints which are necessary conditions on the homology. We refer to these as *motion constraints* and they are equivalent to the epipolar constraint, but have the advantage that the computations involve only those features which can be measured directly, namely the parallax vectors.

3.1 Motion constraints

Here we give necessary conditions for the homology (which are therefore necessary for scene rigidity in two views) in the form of an identity involving only areas computable from the parallax vectors. Two such conditions and their symmetric forms can be determined. The first is derived from the collinearity of the points $\mathbf{a}_{qr}$, $\mathbf{a}_{pr}$ and $\mathbf{a}_{pq}$ and leads to a constraint which is trilinear in the areas. The second is derived from the collinearity of the epipole $\mathbf{e}_{12}$ and corresponding points and is bilinear in the areas. The results are summarised in table 2.

The areas $A_{\mathbf{pqr}}$ can be computed either in the image or on the distinguished plane. The latter requires knowledge of the world to image homography, the former only the homography between images.

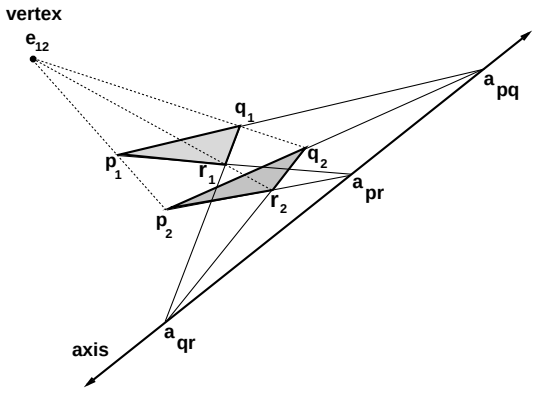

Fig. 4. Three points in two views relative to a known plane leads directly to a Desargues configuration on the plane.

Violation of any of (1) – (5) is a clear indication that there has been non-rigid motion between the views. However if any (or all) of the points $\mathbf{P}, \mathbf{Q}, \mathbf{R}$ moves in its own epipolar planes then the equations are still satisfied and non-rigidity is not detected.

MOTION CONSTRAINTS		
T_1	$A_{p_1p_2r_1}A_{q_1q_2p_1}A_{q_1r_1r_2} = A_{p_1p_2q_1}A_{q_1q_2r_1}A_{p_1r_1r_2}$	(1)
T_2	$A_{p_2p_1r_2}A_{q_2q_1p_2}A_{q_2r_2r_1} = A_{p_2p_1q_2}A_{q_2q_1r_2}A_{p_2r_2r_1}$	(2)
B_1	$A_{r_2p_1p_2}A_{r_1q_1q_2} = A_{r_1p_1p_2}A_{r_2q_1q_2}$	(3)
B_2	$A_{p_2r_1r_2}A_{p_1q_1q_2} = A_{p_1r_1r_2}A_{p_2q_1q_2}$	(4)
B_3	$A_{q_2p_1p_2}A_{q_1r_1r_2} = A_{q_1p_1p_2}A_{q_2r_1r_2}$	(5)

Table 2. Two view bilinear (B_i) and trilinear (T_i) motion constraints equivalent to the epipolar constraint.

4 Geometry of three views

We now consider the geometry of six points, four of which are coplanar, in three views. This is the situation addressed by Irani and Anandan [7]. The geometry is shown in figure 5.

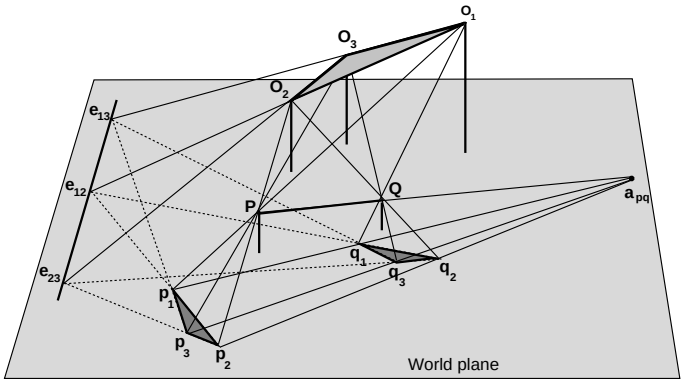

Fig. 5. The geometry of three views with two points off the plane. The three epipoles are collinear, lying on the line which is the intersection of the plane $\mathbf{O}_1\mathbf{O}_2\mathbf{O}_3$ with the world plane. The point $\mathbf{a}_{pq}$ is the intersection of the line $\mathbf{PQ}$ with the world plane.

We begin by demonstrating the duality of this case to the two view case in section 3, and obtain a *structural constraint* directly from the measured image features. We then derive the trifocal tensor for the three view, six point (four coplanar) case. Since the trifocal tensor is over-constrained by six points, four of which are coplanar, we also obtain another form of the structure constraint.

4.1 Duality

It is clear by inspection of figure 5 that the three view geometry is dual to that of figure 3 in which we have directly exchanged points off the plane for camera positions. The vertex of the homology is given by the intersection of the line $\mathbf{PQ}$ with the world plane, and the axis by the intersection of the trifocal plane containing the three camera centres $\mathbf{O}_1$, $\mathbf{O}_2$, $\mathbf{O}_3$ with the world plane (figure 6).

Having established the duality of the two situations, we are now in a position to invoke duality in order to prove further results. We make the substitutions:

2 views	$\mathbf{e}_{12}$	$\mathbf{p}_1$	$\mathbf{p}_2$	$\mathbf{r}_1$	$\mathbf{r}_2$	$\mathbf{q}_1$	$\mathbf{q}_2$	$\mathbf{a}_{pr}$	$\mathbf{a}_{pq}$	$\mathbf{a}_{qr}$
3 views	$\mathbf{a}_{pq}$	$\mathbf{p}_1$	$\mathbf{q}_1$	$\mathbf{p}_2$	$\mathbf{q}_2$	$\mathbf{p}_3$	$\mathbf{q}_3$	$\mathbf{e}_{12}$	$\mathbf{e}_{13}$	$\mathbf{e}_{23}$

and the dual trilinear and bilinear constraints given in table 3 follow. Note that the bilinear constraints (8) – (10) are exactly the constraints given by Irani and Anandan [7]. The trilinear constraints are new.

4.2 The trifocal tensor

In this section it is shown that the trifocal tensor is uniquely determined from three views of six points, four of which are coplanar. We begin with a familiar

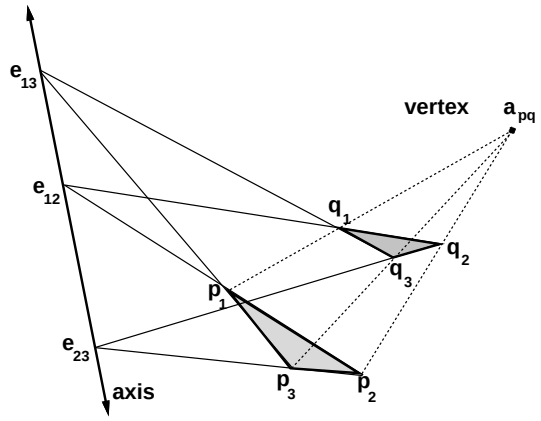

Fig. 6. The planar geometry in the three view two point case is also clearly a Desargues configuration and point correspondences $\mathbf{p_1} \rightarrow \mathbf{q_1}$, $\mathbf{p_2} \rightarrow \mathbf{q_2}$ and $\mathbf{p_3} \rightarrow \mathbf{q_3}$ are related by a homology. This situation is clearly dual to that in figure 4.

STRUCTURAL CONSTRAINTS		
T_1	$A_{\mathbf{p_1q_1p_2}} A_{\mathbf{p_3q_3p_1}} A_{\mathbf{p_3p_2q_2}} = A_{\mathbf{p_1q_1p_3}} A_{\mathbf{p_3q_3p_2}} A_{\mathbf{p_1p_2q_2}}$	(6)
T_2	$A_{\mathbf{q_1p_1q_2}} A_{\mathbf{q_3p_3q_1}} A_{\mathbf{q_3q_2p_2}} = A_{\mathbf{q_1p_1q_3}} A_{\mathbf{q_3p_3q_2}} A_{\mathbf{q_1q_2p_2}}$	(7)
B_1	$A_{\mathbf{q_2p_1q_1}} A_{\mathbf{p_2p_3q_3}} = A_{\mathbf{p_2p_1q_1}} A_{\mathbf{q_2p_3q_3}}$	(8)
B_2	$A_{\mathbf{q_1p_2q_2}} A_{\mathbf{p_1p_3q_3}} = A_{\mathbf{p_1p_2q_2}} A_{\mathbf{q_1p_3q_3}}$	(9)
B_3	$A_{\mathbf{q_3p_1q_1}} A_{\mathbf{p_3p_2q_2}} = A_{\mathbf{p_3p_1q_1}} A_{\mathbf{q_3p_2q_2}}$	(10)

Table 3. Three view bilinear (B_i) and trilinear (T_i) structure constraints.

form of the trifocal tensor (after [5]) in which we consider the image projection matrices and image point locations. We then show how the form of the tensor is simplified when we consider all geometric objects (lines, points, etc) projected onto the distinguished plane.

Image form: In order to discriminate between points (and epipoles) in images, as opposed to their projections onto the distinguished plane, we use primes to indicate the image in which the points appear; i.e. the images of $\mathbf{P}$ in views 1, 2 and 3 are, respectively, $\mathbf{p}$, $\mathbf{p}'$ and $\mathbf{p}''$. Suppose the homographies induced by the plane of the points are $\mathtt{A}$ and $\mathtt{B}$, so that $\mathbf{p}' = \mathtt{A}\mathbf{p}$ and $\mathbf{p}'' = \mathtt{B}\mathbf{p}$ for images of points on the plane. These homographies are computed from the images of the four coplanar points.

The images of the first camera centre in the second and third images, denoted $\mathbf{e}'$ and $\mathbf{e}''$ respectively, are the epipoles. They are determined using parallax vectors, as described in section 3, so that $\mathtt{F}_{12} = [\mathbf{e}']_\times \mathtt{A}$, and $\mathtt{F}_{13} = [\mathbf{e}'']_\times \mathtt{B}$. It can be shown that the three camera projection matrices can be chosen as

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}], \quad \mathtt{P}' = [\mathtt{A} \mid \mathbf{e}'], \quad \mathtt{P}'' = [\mathtt{B} \mid \lambda \mathbf{e}''] \tag{11}$$

up to a homography of 3-space, where λ is an unknown scalar. This unknown scalar is determined by line transfer.

The line through the (non-coplanar) points $\mathbf{P}$, $\mathbf{Q}$, is imaged as $\mathbf{l} = \mathbf{p} \times \mathbf{q}$, $\mathbf{l}' = \mathbf{p}' \times \mathbf{q}'$, $\mathbf{l}'' = \mathbf{p}'' \times \mathbf{q}''$ in the first, second and third views respectively. It is then straightforward to show that lines transfer as follows (we could alternatively consider point transfer):

$$\mathbf{l} = \lambda(\mathbf{e}''.\mathbf{l}'')\mathtt{A}^\top \mathbf{l}' - (\mathbf{e}'.\mathbf{l}')\mathtt{B}^\top \mathbf{l}'' \tag{12}$$

The scalar λ is the only unknown in this equation. It is determined by taking the vector product with $\mathbf{l}$.

$$\lambda(\mathbf{e}''.\mathbf{l}'')\mathbf{l} \times (\mathtt{A}^\top \mathbf{l}') = (\mathbf{e}'.\mathbf{l}')\mathbf{l} \times (\mathtt{B}^\top \mathbf{l}'') \tag{13}$$

This provides two equations in the one unknown λ and so we can solve uniquely for the trifocal tensor and obtain one further constraint, namely the rigidity condition that the imaged intersection of the line through $\mathbf{P}, \mathbf{Q}$ is the same when computed from views one and two ($\mathbf{l} \times (\mathtt{A}^\top \mathbf{l}')$) as from views one and three ($\mathbf{l} \times (\mathtt{B}^\top \mathbf{l}'')$). This is yet another form of the constraints (6) – (10). The scale factor lambda is obtained by normalising both sides of (13):

$$\lambda = \frac{||(\mathbf{e}'.\mathbf{l}')\mathbf{l} \times (\mathtt{B}^\top \mathbf{l}'')||}{||(\mathbf{e}''.\mathbf{l}'')\mathbf{l} \times (\mathtt{A}^\top \mathbf{l}')||} \tag{14}$$

Distinguished plane form: On the distinguished plane $\mathtt{A} = \mathtt{B} = \mathtt{I}$, so the equivalent of (12) for point transfer is

$$\mathbf{r_3} = \lambda \mathbf{e}_{13}(\mathbf{l}_2.\mathbf{r}_1) - (\mathbf{e}_{12}.\mathbf{l}_2)\mathbf{r}_1 \tag{15}$$

where $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$ are the distinguished plane images of a general 3D point $\mathbf{R}$, and $\mathbf{l}_2$ is any line through $\mathbf{r}_2$. This equation depends only on the positions of the epipoles on the distinguished plane, with all dependence on camera internals and relative rotations having been factored out into the image to plane homographies.

Additionally the projection matrices have the very simple form

$$\mathtt{P}_1 = [\mathtt{I} \mid \mathbf{0}], \quad \mathtt{P}_2 = [\mathtt{I} \mid \mathbf{e}_{12}], \quad \mathtt{P}_3 = [\mathtt{I} \mid \lambda \mathbf{e}_{13}] \tag{16}$$

Hence, representing a general 3D point as $\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 \\ \rho \end{bmatrix}$, we determine the distingished plane images to be:

$$\mathbf{r}_1 = \mathtt{P}_1\mathbf{R}, \quad , \mathbf{r}_2 = \mathtt{P}_2\mathbf{R} = \mathbf{r}_1 + \rho \mathbf{e}_{12}, \quad \mathbf{r}_3 = \mathtt{P}_3\mathbf{R} = \mathbf{r}_1 + \rho\lambda \mathbf{e}_{13} \tag{17}$$

We now give an interpretation of ρ and λ on the distinguished plane (see figure 7).

The ratio λ depends only on the camera centres, not on the points, and can be determined as $\lambda = d(\mathbf{e}_{12}, \mathbf{e}_{23})/d(\mathbf{e}_{13}, \mathbf{e}_{23})$ where $d()$ is the distance between the points on the distinguished plane. The parameter ρ is the relative affine invariant of Shashua [14], and is related to the point depth. On the distinguished plane it is obtained as $\rho = d(\mathbf{r}_2, \mathbf{r}_1)/d(\mathbf{r}_2, \mathbf{e}_{12})$.

So point transfer using the trifocal tensor simply involves computing the ratio ρ from $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{e}_{12}$ and employing λ to define the transferred point $\mathbf{r}_3$ on the line between $\mathbf{e}_{13}$ and $\mathbf{r}_1$ as $\mathbf{r}_3 = \mathbf{r}_1 + \rho\lambda\mathbf{e}_{13}$ in (17). This is identical to the point transfer of (15), as can be seen by considering similar triangles in figure 7. In the case that the three camera centres are collinear there is no degeneracy in point/line transfer. The ratio λ is still defined and can be obtained using the distinguished plane equivalent of (14).

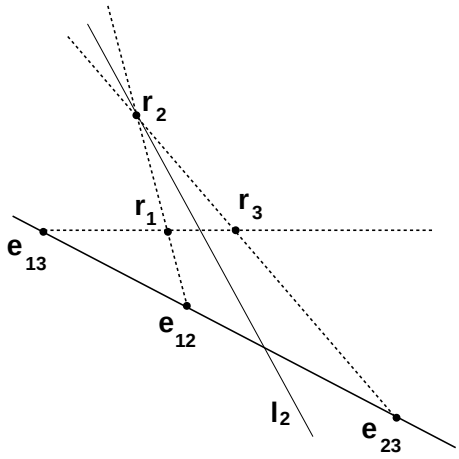

Fig. 7. Point transfer: the ratios of the distances between $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{e}_{12}$ and the three epipoles define the transfer of the point $\mathbf{r}_1$ to $\mathbf{r}_3$.

5 Structural computations and applications

In this section we discuss a number of useful structural computations which can be achieved using ratios of areas. We require *affine* measurements on the world plane, which can be obtained either from four world plane points known up to an affinity (and hence the image to world plane homographies), or from the inter-image homography and vanishing line of the world plane in each image. In either case we obtain results for the scene structure without resorting first to computing the epipolar geometry. A significant novel aspect of the formulae

given in sections 5.1 and 5.2 is that the vanishing point for the direction of measurement need not be known.

The results are derived for the two view, seven point case. However because of the fundamental duality proved in section 4.1, they are equally valid (with appropriate symbol substitutions) in the three view, six point case. For example (22) can be used to compute the height of a third point given two other known heights in the two view, seven point case; dually, in the three view, six point case, it can be used to obtain the height of a third camera given the other two camera heights.

We begin by re-deriving the basic parallax relationship for the case where the parallax is measured on the distinguished world plane (18).

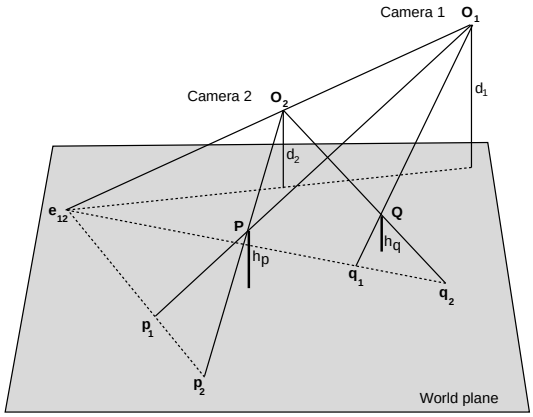

Fig. 8. Parallax geometry of two points.

Considering figure 8 and writing $\mathbf{p}_1 = \mathbf{O}_1 + \frac{d_1}{d_1 - h}(\mathbf{P} - \mathbf{O}_1)$, $\mathbf{p}_2 = \mathbf{O}_2 + \frac{d_2}{d_2 - h}(\mathbf{P} - \mathbf{O}_2)$ and $\mathbf{e}_{12} = \mathbf{O}_2 + \frac{d_2}{d_2 - h}(\mathbf{O}_1 - \mathbf{O}_2)$. Then eliminating $\mathbf{O}_1$ and $\mathbf{O}_2$ yields

$$\boldsymbol{\mu}_p = \frac{h_p}{d_2 - h_p}\frac{\Delta_d}{d_1}(\mathbf{p}_1 - \mathbf{e}_{12}) \tag{18}$$

where $\boldsymbol{\mu}_p = \mathbf{p_2} - \mathbf{p_1}$ is the planar parallax vector and $\Delta_d = d_1 - d_2$ is the component of camera translation towards the plane.

Let γ be the ratio of the distance of a point to the plane and the point to the first camera (measured in the same direction), i.e $\gamma_p = \frac{h_p}{d_1 - h_p}$ and $\gamma_q = \frac{h_q}{d_1 - h_q}$ then combining the basic parallax equation (18) for two points $\mathbf{P}$ and $\mathbf{Q}$ gives

$$\gamma_q \boldsymbol{\mu}_p - \gamma_p \boldsymbol{\mu}_q = \gamma_p \gamma_q \frac{\Delta_d}{d_2}(\mathbf{p_2} - \mathbf{q_2}) \tag{19}$$

Finally, taking the cross product of both sides of the equation with $\mathbf{p_2} - \mathbf{q_2}$ and taking magnitudes yields an expression for $\frac{\gamma_q}{\gamma_p}$ as a ratio of areas (see figure 9)

of the form

$$\frac{\gamma_q}{\gamma_p} = \frac{|\boldsymbol{\mu}_q \times (\mathbf{p_2} - \mathbf{q_2})|}{|\boldsymbol{\mu}_p \times (\mathbf{p_2} - \mathbf{q_2})|} = \frac{A_{\mathbf{q_1 q_2 p_2}}}{A_{\mathbf{p_1 p_2 q_2}}} \tag{20}$$

This ratio is computable solely from the parallax measurements, and is clearly affine invariant, being a ratio of areas. Our derivation is equivalent to Irani and Anandan's construction [7], but note that in our formulation we have used only affine constructs (no perpendicularity has been assumed and the formulae are homogeneous).

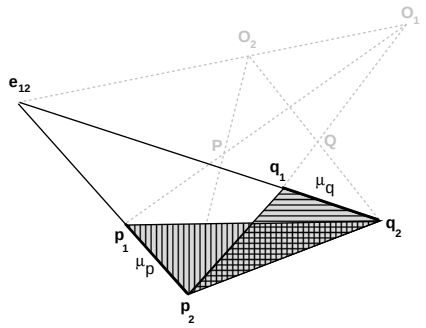

Fig. 9. The relative structure $\frac{\gamma_q}{\gamma_p}$ can be expressed as a ratio of areas.

5.1 Camera distance to world plane

Here we show that: given the parallax vectors of two world points $\mathbf{P}$ and $\mathbf{Q}$, the Euclidean distances of these points from the world plane (measured in the same but arbitrary direction), h_p and h_q, and affine measurements on the world plane, then we can determine the Euclidean distance of either camera to the world plane (measured in the same direction as h_p and h_q).

We can rearrange (20) to give an expression for the distance of the first camera from the plane (a similar formula can be derived for the second camera):

$$d_1 = \frac{h_p h_q (A_{\mathbf{p_1 p_2 q_2}} - A_{\mathbf{q_1 q_2 p_2}})}{h_q A_{\mathbf{p_1 p_2 q_2}} - h_p A_{\mathbf{q_1 q_2 p_2}}} \tag{21}$$

In figure 11 we have used the formulae to compute the heights of the camera above the floor as 566cm and 586cm (left and right views respectively).

5.2 Measuring the structure of other points

Here we show that: given the parallax vectors of three world points $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$, the Euclidean distances of two of these points from the world plane (measured in

the same but arbitrary direction), h_p and h_q ($h_p \neq h_q$), and affine measurements on the world plane, then we can determine the Euclidean distance of the third point from the plane (measured in the same direction as h_p and h_q).

From (20) we have that $\gamma_r = \gamma_p \frac{A_{\mathbf{r}_1\mathbf{r}_2\mathbf{p}_1}}{A_{\mathbf{p}_1\mathbf{p}_2\mathbf{r}_1}}$ and $\gamma_r = \gamma_q \frac{A_{\mathbf{r}_1\mathbf{r}_2\mathbf{q}_1}}{A_{\mathbf{q}_1\mathbf{q}_2\mathbf{r}_1}}$. After eliminating the camera distance d between these equations we obtain an expression for h_r:

$$h_r = \frac{|\mathtt{M}_1|}{|\mathtt{M}_2|}, \quad \mathtt{M}_1 = h_p h_q \begin{bmatrix} A_{\mathbf{r}_1\mathbf{r}_2\mathbf{q}_1} & A_{\mathbf{q}_1\mathbf{q}_2\mathbf{r}_1} \\ A_{\mathbf{r}_1\mathbf{r}_2\mathbf{p}_1} & A_{\mathbf{p}_1\mathbf{p}_2\mathbf{r}_1} \end{bmatrix}, \quad \mathtt{M}_2 = \begin{bmatrix} A_{\mathbf{r}_1\mathbf{r}_2\mathbf{q}_1} & A_{\mathbf{q}_1\mathbf{q}_2\mathbf{r}_1} & A_{\mathbf{q}_1\mathbf{q}_2\mathbf{r}_1} \\ h_p & 0 & -h_q \\ A_{\mathbf{p}_1\mathbf{p}_2\mathbf{r}_1} & A_{\mathbf{p}_1\mathbf{p}_2\mathbf{r}_1} & A_{\mathbf{r}_1\mathbf{r}_2\mathbf{p}_1} \end{bmatrix} \tag{22}$$

Note that if h_p and h_q are equal then the formulae above degenerate, however this situation can be avoided in practice. The degeneracy is understood in terms of the geometry as follows: in general we obtain projective structure (since $\mathtt{F}$ is determined uniquely). If in addition $h_p \neq h_q$, the line $\mathbf{PQ}$ intersects $\boldsymbol{\pi}_\infty$ in a point which can be identified in both images (details omitted). This point and the vanishing line of the world plane, $\mathbf{l}_\infty$, determine $\boldsymbol{\pi}_\infty$, hence we can obtain affine structure. When $h_p = h_q$ then the line $\mathbf{PQ}$ intersects $\boldsymbol{\pi}_\infty$ on $\mathbf{l}_\infty$ and so no additional information about $\boldsymbol{\pi}_\infty$ is obtained; it is determined only up to a one parameter family (the pencil of planes with $\mathbf{l}_\infty$ as its axis).

Figure 10 shows an example in which point heights have been estimated. The floor tiling and the perpendicular heights of two other points were measured by hand with a tape measure. A second example is shown in Figure 11. As before, we have used the patterned floor tiling to compute the image to world plane homography, and measured the heights of two points with a tape measure as reference heights.

6 Discussion

We have considered the geometry of two views of seven points, four of which are coplanar, and the geometry of three views of six points, four of which are coplanar. These configurations were shown to be dual to one another and that the fundamental geometric constraints are captured by a planar homology relating the images of points across views. Consequently a number of previously diverse results related to planar parallax, duality and planar homologies have been unified.

The constraints derived from the homology are easily computed from directly measured image features, and have been tested on real imagery. Formulae for determining the height of the cameras from the world plane, and the heights of points from the plane have been developed and tested on real imagery. We are currently evaluating the accuracy of the method. In particular we are investigating how uncertainty and errors in image measurements propagate to 3D measurements [4]. A subject of further investigation will be the sensitivity of the structure and motion constraints.

a b

Tape measure	191.5	165.0	135.0	109.1	35.0	22.9
Image parallax	191.8	164.8	134.8	108.9	35.2	22.1

c

Fig. 10. Estimating the heights of points from two views of the lecturn in The Queen's College chapel. The heights of the reference points, shown black on white, were measured by hand ($\pm 0.5cm$) to be 150cm (top edge of lecturn), and 35cm (height of foot stool). In (c), all heights are given in centimetres. The error between the computed and measured heights is always less than one centimetre.

a b

Fig. 11. Estimating the heights of points from two views of The Queen's College dining hall. The heights of the reference points, shown black on white, were measured by hand ($\pm 0.5cm$) to be 76cm (table top), and 230cm (fireplace). The step was computed to be 11.4cm high and measured by hand as 11.0cm.

Acknowledgements The work has been supported by Technology Foresight Grant AMVIR, Esprit Project IMPROOFS and by an EPSRC Advanced Reareach Fellowship. We are grateful to A. Fitzgibbon and F. Schaffalitzky for assistance and discussions.

References

1. P. A. Beardsley, D. Sinclair, and A. Zisserman. Ego-motion from six points. Insight meeting, Catholic University Leuven, Feb 1992.
2. Carlsson S. Duality of reconstruction and positioning from projective views. In *ICCV Workshop on Representation of Visual Scenes, Boston*, 1995.
3. Carlsson S. and Weinshall D. Dual computation of projective shape and camera positions from multiple images. *Int'l J. of Computer Vision*, 1998. in Press.
4. A. Criminisi, I. Reid, and A. Zisserman. A plane measuring device. In *Proc. BMVC*, UK, September 1997.
5. R. I. Hartley. Lines and points in three views – a unified approach. In *ARPA Image Understanding Workshop, Monterrey*, 1994.
6. R. I. Hartley. A linear method for reconstruction from lines and points. In *Proc. ICCV*, pages 882–887, 1995.
7. M. Irani and P. Anandan. Parallax geometry of pairs of points for 3d scene analysis. In B. Buxton and R. Cipolla, editors, *Proc. ECCV*, pages 17–30. Springer, 1996.
8. R. Kumar, P. Anandan, M. Irani, J. Bergen, and K. Hanna. Representation of scenes from collections of images. In *ICCV Workshop on the Representation of Visual Scenes*, 1995.
9. Q.-T. Luong and T. Viéville. Canonic representations for the geometries of multiple projective views. In *Proc. ECCV*, pages 589–599, May 1994.
10. R. Mohr. Projective geometry and computer vision. In Chen, Pau, and Wang, editors, *Handbook of Pattern Recognition and Computer Vision*. 1992.
11. J. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision*. MIT Press, 1992.
12. H. S. Sawhney. Simplifying motion and structure analysis using planar parallax and image warping. In *Proc. CVPR*, 1994.
13. J. Semple and G. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1979.
14. A. Shashua. On geometric and algebraic aspects of 3d affine and projective structures from perspective 2d views. In J. Mundy, A. Zisserman, and D. Forsyth, editors, *Applications of Invariance in Computer Vision LNCS 825*, pages 127–143. Springer-Verlag, 1994.
15. A. Shashua. Trilinearity in visual recognition by alignment. In *Proc. ECCV*, volume 1, pages 479–484, May 1994.
16. M. E. Spetsakis and J. Aloimonos. Structure from motion using line correspondences. *Intl. J. of Computer Vision*, 4(3):171–183, 1990.
17. Van Gool L., Proesmans M., and Zisserman A. Grouping and invariants using planar homologies. In Mohr, R. and Chengke, W., editors, *Europe-China workshop on Geometrical Modelling and Invariants for Computer Vision*, pages 182–189. Xidan University Press, Xi'an, China, 1995.
18. D. Weinshall, M. Werman, and A. Shashua. Duality of multi-point and multi-frame geometry: Fundamental shape matrices and tensors. In B. Buxton and R. Cipolla, editors, *Proc. ECCV*, pages 217–227. Springer–Verlag, 1996.
19. Werman M., Weinshall D., and Shashua A. Shape tensors for efficient and learnable indexing. In *ICCV Workshop on Representation of Visual Scenes, Boston*, 1995.