Trilinearity of Three Perspective Views and its Associated Tensor

Amnon Shashua*

Hebrew University of Jerusalem
Institute of Computer Science
91904 Jerusalem, Israel
e-mail: samm@cs.huji.ac.il

Michael Werman

Hebrew University of Jerusalem
Institute of Computer Science
91904 Jerusalem, Israel
e-mail: werman@cs.huji.ac.il

Abstract

It has been established that certain trilinear froms of three perspective views give rise to a tensor of 27 intrinsic coefficients [11]. We show in this paper that a permutation of the the trilinear coefficients produces three homography matrices (projective transformations of planes) of three distinct intrinsic planes, respectively. This, in turn, yields the result that 3D invariants are recovered directly — simply by appropriate arrangement of the tensor's coefficients. On a secondary level, we show new relations between fundamental matrix, epipoles, Euclidean structure and the trilinear tensor. On the practical side, the new results extend the existing envelope of methods of 3D recovery from 2D views — for example, new linear methods that cut through the epipolar geometry, and new methods for computing epipolar geometry using redundancy available across many views.

1 Introduction

Given that three-dimensional (3D) objects in the world are modeled by point sets, then their projection onto a number of distinct image planes produces point sets that are related by correspondences. The geometric and algebraic relationship between a 3D point set and the correspondences across the images, created by the pin-hole camera model, is by now well understood when the number of images is two. In the case of two cameras, the correspondence constraint is a bilinear function of image coordinates whose coefficients can be arranged in a matrix, known as the fundamental matrix [3, 1]. The bilinear constraint describes the epipolar geometry ("coplanarity constraint" known in photogrammetric circles), and once it is known the 3D structure of the scene can be recovered (linearly) up to a projective transformation [2, 6, 10].

The case of three cameras is a recent topic of investigation. It has been shown in [11] that the correspondence constraint is expressed by certain trilinear forms (henceforth "trilinearities") having all together 27 distinct coefficients (henceforth "trilinear tensor"). The interest in the trilinearities is threefold. On one hand, the trilinear tensor can be recovered linearly from 7 corresponding points across the three views, whereas the fundamental matrix requires 8 points. Second, concatenation of epipolar geometries across three views fails in cases where trilinearities do not. Third, the trilinearities use all three views together, rather than in pairs, implying that the current class of 3D reconstruction methods that are based on "two-views-at-time" can be extended to "three-views-at-a-time", thereby gaining additional numerical stability.

This paper investigates the relationship between the trilinear tensor and geometric invariants of the 3D world, intrinsic structures of two views (fundamental matrix and epipoles), and show the existence of new intrinsic structures (beside the tensor itself) associated with three views. We show that there are three generic ways to permute the tensor into sets of the three matrices. One of these permutations was recently shown by Hartley [5] to yield, somewhat unexpectedly, the three matrices associated with lines over three views [7, 15, 13]. The other two permutations produce another remarkable result: the sets of matrices turn out to be homography matrices (projective transformations of planes) of distinct *intrinsic* planes. This result has far reaching implications, because it implies that 3D invariants can be recovered directly — simply by appropriate arrangement of the tensor's coefficients.

2 Preliminaries

We consider object space to be three dimensional projective space $\mathcal{P}^3$, and image space to be two dimensional projective space $\mathcal{P}^2$. An object is modeled as a set of points in $\mathcal{P}^3$ and its views (images) are denoted by $\psi_i \subset \mathcal{P}^2$. The projection centers will be denoted by O, O', O'' for views ψ_1, ψ_2, ψ_3, respectively (we consider at most three views in this paper). The symbol $\cong$ denotes equality up to a scale. The skew-symmetric matrix associated with vector product $u \times w$ is denoted by $[u]_x$, i.e., $[u]_x w = u \times w$ for any vector w.

Image coordinates are noted by $p = (x, y, 1)^\top, p' = (x', y', 1)^\top$ and $p'' = (x'', y'', 1)^\top$ for views ψ_1, ψ_2, ψ_3, respectively. Let M denote the transformation from the (unknown) camera coordinate frame to the adopted coordinization of ψ_1, i.e., $zM^{-1}p$ is the location of the object point in the first camera coor-

*Starting from fall 1995, A. Shashua will be at the Computer Science Department of the Technion, Haifa 32000, Israel.

dinate frame. In other words, M represents what is often referred to as the interior orientation of the camera. Similarly, M' and M'' denote the internal parameters of the cameras associated with views ψ_2 and ψ_3, respectively. Let R and T stand for the rotational and translational components (relative orientation) from the first to the second camera coordinate frames, hence we have:

$$p' \cong M'RM^{-1}p + \frac{1}{z}M'T$$

Similarly, R'' and T'' are the rotational and translational components from the first to the third camera coordinate frames. The vector $v' \cong M'T$ is known as the epipole (located on ψ_2) which is the point at the intersection of the line OO' with the image plane of the second camera. Similarly, $v'' \cong M''T''$ is the epipole on ψ_3.

2.1 2-view Geometry via the Homography Matrix: Review

The intrinsic structures associated with two views, say between ψ_1 and ψ_2, can be conveniently described through a 2D projective transformation (homography matrix), as follows. Let π be some arbitrary plane in $\mathcal{P}^3$ parameterized by normal vector n and distance to origin d with respect to the first camera frame, then the homography matrix A_π is described below [4, 14, 12]:

$$A_\pi \cong M'(R + \frac{1}{d}Tn^\top)M^{-1}. \quad (1)$$

The transformation A_π registers the points of π across the two images, i.e., if p, p' are projections of a point $P \in \pi$, then $A_\pi p \cong p'$. We have the following properties: (i) $A_\pi v \cong v'$, where $v \cong MR^\top T$ is the corresponding epipole on ψ_1, and (ii) $p'^\top [v']_x A_\pi p = 0$ for any corresponding pair of image points. The first property is that the homography matrix aligns the epipoles, regardless of π. This property provides a simple method for recovering the homography matrix from corresponding pairs, once the epipoles v, v' are known: take three corresponding pairs p_i, p'_i, $i = 1,2,3$, (their corresponding object points determine the plane π), then the relation $A_\pi p_i \cong p'_i$, and $A_\pi v \cong v'$ provides a linear system of equations to solve for the homography matrix [10, 9]. The second property is the epipolar geometry: $F = [v']_x A_\pi \cong [v']_x M'RM^{-1}$ is known as the "fundamental" matrix [2], which we readily see can be recovered linearly from eight corresponding image pairs. Note that F is an intrinsic structure of two views because it does not depend on the choice of π. It also readily follows that for any π we have $A_\pi{}^\top F + F^\top A_\pi = 0$. For notation purposes, we denote F'' to be the fundamental matrix between ψ_1, ψ_3, i.e., $p''^\top F'' p = 0$.

One can distinguish two intrinsic fiducial planes: the first is the plane at infinity whose homography matrix is $A_\infty = M'RM^{-1}$ arises from setting $d = \infty$ in Eqn. 1 (note that A_∞ cannot be recovered from image measurements alone without prior information, unlike A_π). The second plane, denoted by π_s, is coplanar with the center O' (thus the corresponding homography matrix is of rank 2, and $A_{\pi_s} v = 0$) and perpendicular to the line OO', thus in addition $A_{\pi_s}{}^\top v' = 0$. It then follows [8] that $A_{\pi_s} \cong [v']F$. We will refer to the plane π_s as the "singular intrinsic plane".

The coordinization of object space and its views is based on the (3×4) projection matrix $[I, 0]$ onto the first view ψ_1, and the projection matrix $[A_\pi, v']$ onto the second view ψ_2 (note that π is arbitrary). Thus the projective coordinates of a point P projecting onto p, p' is $[x, y, 1, k]$ where k satisfies the equation $p' \cong A_\pi p + kv'$. The scalar k remains fixed between ψ_1 and any other view ψ_i if the same fiducial plane π is used [12]. In other words, k is invariant to the choice of of the interior and relative orientation parameters (the projection matrix) from object space to ψ_i. If two fiducial planes π_1 and π_2 are used whose homography matrices are recovered up to a mutual scale, then a projective invariant κ can be defined by the equation $p' \cong A_{\pi_1} p + \kappa A_{\pi_2} p$. The scalar κ is invariant to the choice of the interior and relative orientation parameters of both views [10], provided that π_1 and π_2 are fiducials (i.e., used for all pairs of views).

2.2 Trilinearity: Review

The material presented in this paper relies critically on the existence of certain trilinear forms of three views with intrinsic coefficients [11]. Let $A_\pi : \psi_1 \mapsto \psi_2$ and $B_\pi : \psi_1 \mapsto \psi_3$ be homography matrices of some arbitrary plane π, recovered together with v', v'' up to a common scale. Then, there exists a tensor of 27 intrinsic (i.e., do not depend on the choice of π) entries:

$$\alpha_{ijk} = v'_i b_{jk} - v''_j a_{ik}. \qquad i,j,k = 1,2,3 \quad (2)$$

We use the standard notations for tensors, as follows. We identify vectors and matrices by fixing some of the indices while varying others. For example, α_{ijk} is a set of scalars, $\alpha_{ij\cdot}$ is a set of 9 vectors (k varies while i, j remain fixed); $\alpha_{i\cdot\cdot}$ is a set of 3 matrices ($\alpha_{1\cdot\cdot}, \alpha_{2\cdot\cdot}$ and $\alpha_{3\cdot\cdot}$), and so forth.

The tensor α forms the set of coefficients of certain trilinear forms that vanish on any corresponding triplet p, p', p'' (i.e., functions of views that are invariant to object structure). There are nine functions divided into sets of four linearly independent functions[1]. For example, the following four forms are linearly independent invariant functions of three views:

$$x''\boldsymbol{\alpha}_{13\cdot}^\top p - x''x'\boldsymbol{\alpha}_{33\cdot}^\top p + x'\boldsymbol{\alpha}_{31\cdot}^\top p - \boldsymbol{\alpha}_{11\cdot}^\top p = 0,$$
$$y''\boldsymbol{\alpha}_{13\cdot}^\top p - y''x'\boldsymbol{\alpha}_{33\cdot}^\top p + x'\boldsymbol{\alpha}_{32\cdot}^\top p - \boldsymbol{\alpha}_{12\cdot}^\top p = 0,$$
$$x''\boldsymbol{\alpha}_{23\cdot}^\top p - x''y'\boldsymbol{\alpha}_{33\cdot}^\top p + y'\boldsymbol{\alpha}_{31\cdot}^\top p - \boldsymbol{\alpha}_{21\cdot}^\top p = 0,$$
$$y''\boldsymbol{\alpha}_{23\cdot}^\top p - y''y'\boldsymbol{\alpha}_{33\cdot}^\top p + y'\boldsymbol{\alpha}_{32\cdot}^\top p - \boldsymbol{\alpha}_{22\cdot}^\top p = 0.$$

Since every corresponding triplet p, p', p'' contributes four linearly independent equations, then

[1] If we change the order of views, then we have a three-fold scaling: 27 functions, divided into sets of 12 linearly independent functions and 81 coefficients.

seven corresponding points across the three views uniquely determine (up to scale) the tensor α. More details and applications can be found in [11].

3 Homography Matrices via Permutation $\alpha_{\cdot j \cdot}$

The permutations of the tensor obtained by $\alpha_{\cdot j \cdot}$ leads to our main result (Theorem 1) which is remarkably simple:

The three matrices $\alpha_{\cdot j \cdot}$ are three homography matrices of three distinct and intrinsic planes.

Theorem 1 *The matrices E_j corresponding to $\alpha_{\cdot j \cdot}$, i.e., $E_1 = \alpha_{\cdot 1 \cdot}$, and so forth, are homography matrices of three distinct planes:*

$$E_j = \lambda v''_j M'(R + \frac{1}{v''_j} T {u_j}^\top) M^{-1}, \qquad (3)$$

where u_j are the row vectors of $U = M''R''$ and λ is some fixed scale factor.

Proof. From the definition of the tensor in Eqn. 2, we write E_j in terms of the components A, B, v', v'' as follows:

$$E_j = v' {b_j}^\top - v''_j A = M'T {b_j}^\top - v''_j A,$$

where b_j are the row vectors of B, and v' is scaled such that $v' = M'T$. Let v'' be scaled such that $v'' = M''T''$. Since A, B, v', v'' are defined up to a common scale, then the matrices A and B share a common scale λ:

$$A = \lambda M'(R + \frac{1}{d} T n^\top) M^{-1}$$
$$B = \lambda M''(R'' + \frac{1}{d} T'' n^\top) M^{-1},$$

where n and d are the parameters of some arbitrary plane π. By substituting the expressions of A and B in E_j we obtain:

$$E_j = \lambda M'T(u_j + v''_j \frac{n}{d})^\top M^{-1} - \lambda v''_j M'(R + \frac{1}{d} T n^\top) M^{-1},$$

where u_j are the row vectors of $U = M''R''$. By collecting terms we see that the contribution of π drops out and we are left with Eqn. 3. ∎

We see that the arrangement of the tensor into $\alpha_{\cdot j \cdot}$ gives rise directly to homography matrices of some planes in space. These planes are not singular, i.e., they are not coplanar with any of the camera centers, and in case the third camera is calibrated (i.e., $M'' = I$), the planes are perpendicular to the main axes of the third camera coordinate frame, i.e., E_1 is the homography associated with the plane whose normal is the $x - axis$ of the third camera frame and whose distance to the origin O is $1/v''_1$, and so forth. We refer to the matrices E_1, E_2, E_3 as the *intrinsic homography matrices.*

A direct consequence of this result is that we can obtain three projective invariants of 3D space *directly* from the tensor, as follows:

Corollary 1 *The matrices $E_j = \alpha_{\cdot j \cdot}$ give rise to three projective invariants k_1, k_2, k_3 per corresponding pairs p, p' (in images ψ_1, ψ_2 respectively). The invariants satisfy the following equations:*

$$p' \cong E_1 p + k_1 E_2 p$$
$$p' \cong E_1 p + k_2 E_3 p$$
$$p' \cong E_2 p + k_3 E_3 p$$

The scalars k_1, k_2, k_3 are invariant to the camera transformations $\mathcal{P}^3 \mapsto \mathcal{P}^2$ associated with views ψ_1, ψ_2. The third view ψ_3 determines the projective reference frame.

Proof. It was shown in [10] that two homographies (of two distinct planes) recovered up to a common scale provide an invariant (referred to as "projective-depth") which corresponds to a cross-ratio across the line connecting the space point P (projects onto p, p'), the unit basis point P_o and the two intersection points with the the planes associated with the two homographies (note that two planes and a point not coplanar with the planes provide a projective basis for 3D space). By Theorem 1, E_1, E_2, E_3 are homographies of distinct planes, and are determined up to a common scale (because the tensor coefficients are recovered up to a common scale). ∎

We have two outcomes. First, projective invariants of the scene are recovered directly from the tensor, without the need for additional steps of calculations. Second, and quite intriguing, is that those invariants are recovered without "fiducial" points.

For clarity we elaborate briefly on the latter outcome. Computing projective invariants of the scene requires the identification of five fiducial (reference/basis) points, and in the case of two views, can be achieved (linearly) by observing the projections of those five points and the epipoles [2, 10]. For example, the projections of points 1,2,3 and the epipoles determine the homography A_{π_1} of the plane spanned by points 1,2,3. The projections of points 2,3,4 (and the epipoles) determine the homography A_{π_2} of the plane spanned by points 2,3,4. The remaining fifth fiducial point is then used to achieve a common scale for both homographies, and the result is the projective invariance described by the relation [10]:

$$p' \cong A_{\pi_1} p + k A_{\pi_2} p$$

Geometrically, the invariant scalar k is the cross ratio of the fifth fiducial point P_o, the point of interest P (projecting onto p, p') and the intersections of the line connecting P_o and P with the two planes π_1, π_2.

As it turns out, with three views the two planes are provided "for free" without the need to select four fiducial points (and the epipoles). The fifth fiducial point is not required either, because the homographies are already determined up to a common scale. To complete the geometric picture we may describe the location of the fifth fiducial point P_o relative to the planes associated with E_j, as follows. The point P_o is associated with $k=1$, i.e., $p'_o \cong (E_1+E_2)p_o$. The matrix E_1+E_2 is a homography of a plane π_{12}, hence P_o is coplanar with π_{12}. For similar reasons P_o is coplanar with π_{13} and π_{23}, thus P_o is at the intersection of the three planes associated with the homographies E_1+E_2, E_1+E_3 and E_2+E_3.

What we may gather from this is that the third view provides the projective reference frame, i.e., we are replacing fiducial points with a "fiducial view". We summarize this feature of the trilinear tensor below:

The recovery of 3D projective invariants can be accomplished directly from the trilinear tensor (without recovering epipoles or fundamental matrices), nor is it required to identify fiducial points for a projective reference frame.

The finding that $\alpha_{\cdot j \cdot}$, $j=1,2,3$, are homography matrices entails several properties, generally straightforward, such as the ability to compute the fundamental matrix from two homography matrices, and other examples that will be detailed later. Those properties are unrelated to the fact that $\alpha_{\cdot j \cdot}$ have come from the trilinear tensor. However, in addition, the three homography matrices are given up to a *common* scale. From this finding alone we get the curious result describing a direct connection between the matrices $\alpha_{\cdot j \cdot}$ and the epipole $v'' \in \psi_3$:

Theorem 2 *The eigenvalues of $E_j^{-1}E_i$, $i,j = 1,2,3 (i \neq j)$, have multiplicity 2. The multiple eigenvalue is equal to $\frac{v''_i}{v''_j}$.*

Proof. From Eqn. 3 we easily find that

$$E_i - \frac{v''_i}{v''_j}E_j = \lambda v'(u_i - \frac{v''_i}{v''_j}u_j)^\top M^{-1}. \qquad (4)$$

Thus, $E_i - \frac{v''_i}{v''_j}E_j$ is a rank 1 matrix, which in turn proves that $\mu = v''_i/v''_j$ is one of the generalized eigenvalues of $(E_i - \mu E_j)x = 0$, where x is the corresponding generalized eigenvector. Also, since we know there exists μ that makes $E_i - \mu E_j$ rank 1, then μ must be of multiplicity 2. Finally, the generalized eigenvalues of $(E_i - \mu E_j)x = 0$ are the eigenvalues of $(E_j^{-1}E_i - \mu I)x = 0$, i.e., the eigenvalues of $E_j^{-1}E_i$. □

The practical implications of Thm. 2 are, first, that the epipole $v'' \in \psi_3$ can be recovered in closed-form from the tensor matrices (there is a closed-form solution for finding eigenvalues of 3×3 matrices). The components v''_1/v''_3 and v''_2/v''_3 are simply the eigenvalue with multiplicity 2 of the matrices $E_3^{-1}E_1$ and $E_3^{-1}E_2$, respectively. Second, the distribution of the eigenvalues of $E_j^{-1}E_i$ provides an indication of accuracy on finding the epipole v'', i.e., in an error-free case two of the eigenvalues should be equal to v''_i/v''_j.

Next we describe a practical result that follows from the fact that $\alpha_{\cdot j \cdot}$ are homography matrices. In general, given any two homography matrices we have a sufficient (and redundant) set of linear equations for solving for the fundamental matrix:

Corollary 2 *The (over-determined) system of linear equations resulting from:*

$$F^\top E_j + {E_j}^\top F = 0 \qquad j = 1,2,3$$

provides a unique solution for F (up to scale). The solution is unique when at least two of the E_js are used in the system, i.e., $j = j_1, j_2$, $j_1 \neq j_2$, $j_1, j_2 = 1,2,3$.

Proof. Follows directly from the finding that E_j, $j = 1,2,3$, are homography matrices (i.e., $F^T E_j$ is a skew-symmetric matrix). □

The importance of the last corollary is that it provides an algorithm for recovering the fundamental matrix F, associated with the first two views, from the tensor. Moreover, F can be obtained in a least-squares sense because we have a redundant set of equations (18 homogeneous equations for 9 unknowns). In fact, we can use this result and extend the envelope of algorithms for computing the fundamental matrix from two views (+ many point matches) to any number of views $m \geq 3$, as follows.

Consider views $\psi_1, ..., \psi_m$, $m \geq 3$ with many (≥ 7) points matches across them. Assume we would like to compute F between views ψ_1, ψ_m. Let α^{1mj} denote the tensor of views ψ_1, ψ_m, ψ_j (in that order), where $j = 2, ..., m-1$. Each of these tensors can be recovered in a least-squares manner using all the available point matches across the appropriate three views. From Corollary 2 we have that each tensor contributes 18 linear equations for F. Hence, the more views we have the larger system of equations we have for F for a least-squares solution. Non-linear components may be added to this basic scheme by enforcing the rank constraint of F, and other constraints like in [1].

We end this section with two results on the internal structure of the trilinear tensor.

Corollary 3 *The intrinsic homography matrices E_1, E_2, E_3 are generically full rank. In the case the three camera centers O, O', O'' are collinear (i.e., $R''R^\top T = T''$) the intrinsic homography matrices have rank 2 (the corresponding fiducial planes are coplanar with O'). Finally, E_j has rank 1 iff $v''_j = 0$.*

Proof. Eqn. 3 is of the general form depicted by Eqn. 1, rewritten below:

$$A_\pi \cong M'RM^{-1} + \frac{1}{d}v'n^\top M^{-1}. \qquad (5)$$

The right hand terms of Eqn. 5 represent a sum of two matrices, the first of rank 3 and the second of rank 1. Their sum is a matrix of deficient rank if and only if the scalar $\mu = 1/d$ is a generalized eigenvalue of the two matrices. There are two eigenvalues: one with multiplicity 2 making the sum a rank 1 matrix (the case when $v''_j = 0$), and the second eigenvalue making the sum a rank 2 matrix (homography matrix associated with the plane coplanar with O'). Since rank 2 homography matrices necessarily satisfy the constraint $A_\pi v = 0$, it can be readily calculated that the corresponding eigenvalue must equal:

$$\mu = \frac{1}{(M''R''M^{-1})_j{}^\top v}.$$

Therefore, if $M''R''M^{-1}v = v''$, then the rank of all three intrinsic homography matrices is 2. Since, $v = MR^\top T$ and $v'' = M''T''$, the condition reduces to $R''R^\top T = T''$, which in geometric terms means that the three camera centers are collinear. □

Corollary 4 *Let $(E_j)_m$ denote the mth column of matrix E_j. Then the vectors $(E_1)_m, (E_2)_m, (E_3)_m$ span a 2D space for $m = 1, 2, 3$.*

Proof. Follows immediately from the observation that $F^T E_j$ is a skew-symmetric matrix. □

4 The Permutation $\alpha_{i\cdot\cdot}$

The arrangement of the tensor by $\alpha_{i\cdot\cdot}$ produces similar properties as the arrangement $\alpha_{\cdot j \cdot}$, with the difference that the homography matrices apply from ψ_1 to ψ_3 (instead of ψ_2). In other words, if we denote by W_1, W_2, W_3 the three matrices corresponding to $\alpha_{i\cdot\cdot}$, then W_i are homographies from ψ_1 to ψ_3 corresponding to planes associated with the coordinate frame of the second view ψ_2:

$$W_i = \lambda v'_i M''(R'' + \frac{1}{v'_i} T'' g_j{}^\top) M^{-1},$$

where $G = M'R$. The rest follow readily, and for example: $[v'']_x W_i \cong F''$ (recall that F'' is the fundamental matrix between views ψ_1 and ψ_3, and $v'' \in \psi_3$ is the epipole with the first view ψ_1); $F''^\top W_i + W_i{}^\top F'' = 0$. The second view ψ_2 is the fiducial view for reconstruction, rather than the third view.

5 The Permutation $\alpha_{\cdot\cdot k}$

The arrangement $\alpha_{\cdot\cdot k}$ was addressed in full by Hartley in [5]. To complete the picture of the tensor decomposition into sets of matrices we summarize the highlights of this arrangement.

Unlike the previous two arrangements, the matrices $T_k = \alpha_{\cdot\cdot k}$ are not homography matrices, and they are of rank 2 (instead of 3). The epipoles v', v'' can be found as the common perpendiculars to the left (respectively, right) null-spaces of T_k. The matrix:

$$B_{\pi_s} = \frac{1}{v'^\top v'}[T_1{}^\top v', T_2{}^\top v', T_3{}^\top v'].$$

is a homography matrix from ψ_1 onto ψ_3 associated with the plane π_s coplanar with O'. The homography matrix A_{π_s} from ψ_1 onto ψ_2 associated with the plane π_s has the form:

$$A_{\pi_s} = \frac{1}{v''^\top v''}[v'']_x^2 [T_1 v'', T_2 v'', T_3 v''].$$

Note that A_{π_s} is of rank 2 and $A_{\pi_s}{}^\top v' = 0$.

6 Euclidean Results

In the case the cameras are calibrated, the epipole v' (v'') is the translational components of camera motion from ψ_1 to ψ_2 (ψ_3). We show next that given the epipoles we can obtain a closed form solution for the rotational component of camera motion.

Theorem 3 *Given the epipoles v', v'', then in the case the cameras are internally calibrated (M=M'=M"=I) the rotational component of camera motion R'' can be recovered (up to reflection) in closed-form from $\alpha_{\cdot j \cdot}$, and similarly R can be recovered from $\alpha_{i\cdot\cdot}$.*

Proof. Let h_1, h_2, h_3 be the vectors:

$$h_1 = \lambda u_1 - \frac{v''_1}{v''_3}\lambda u_3$$

$$h_2 = \lambda u_2 - \frac{v''_2}{v''_3}\lambda u_3 \quad (6)$$

$$h_3 = \lambda u_1 - \frac{v''_1}{v''_2}\lambda u_2 \quad (7)$$

Given v, v'', the vectors h_i are observable following Eqn. 4, and u_i are the row vectors of R''. Our task is to recover λ and u_1, u_2, u_3 from h_1, h_2, h_3.

We note that since R'' is orthonormal, then $\lambda^2 = h_1{}^\top h_3$. Let $u_1 = u_2 \times u_3$, $u_2 = u_3 \times u_1$ and $u_3 = u_1 \times u_2$, then

$$h_1 \times h_3 = -\lambda(\frac{v''_1}{v''_2}\lambda u_3 + \frac{v''_1}{v''_3}\lambda u_2 + \frac{v''^2_1}{v''_2 v''_3}\lambda u_1).$$

By replacing λu_1, and λu_2 with their expression from Eqn. 6 and 7, respectively, we obtain:

$$u_3 = \frac{-1}{h_1{}^\top h_3(\frac{v''_1}{v''_2} + \frac{v''^3_1}{v''^2_3 v''_2} + \frac{v''_1 v''_2}{v''^2_3})}(h_1 \times h_3 + \frac{\lambda v''^2_1}{v''_2 v''_3} h_1 + \frac{\lambda v''_1}{v''_3} h_2).$$

Thus, we have two solutions for u_3 (because of two solutions for λ), and from there two solutions for R''. □

7 Discussion and Future Research

The paper derives connections between the trilinear forms across three views and invariants of 3D space. The major result is the existence of a permutation of the trilinear coefficients (the tensor) that produces three intrinsic homography matrices (Theorem 1). The corresponding planes of these homography matrices are associated with the camera coordinate frame of the third view and provide a reference basis for reconstruction of invariants. This result also provides a complete decoupling between homography matrices and epipoles (unlike 2-view geometry, and 3-view line geometry).

The material presented in this paper contains a host of practical consequences — of which we mentioned only a few. For example, the envelope of techniques for recovering the fundamental matrix is now extended to multiple views (Corollary 2); projective reconstruction and invariants are recovered without fiducial points; and most importantly, the introduction of the third view has a potentially stabilizing effect on numerical algorithms for 3D calculations from 2D images.

Continuing on the practical motivation, three views admit three distinct fundamental matrices (pairs of views out of three), thus all together we have 27 homogeneous equations (from 9 points), or 24 inhomogeneous equations (from 8 points). However, these equations are decomposed into three independent sets, thus in practice we use only two views at a time. The trilinear coefficients comprising the tensor are determined also by 27 equations but with two important differences: (i) 7 points are needed, (ii) the set of equations is not decomposable linearly, thus all three views are used together. In other words, with the introduction of the tensor there is a better use of the information available in three views. This argument implies a clear superiority, regardless of statistical methods that may be used on top of the results presented here.

Acknowledgments

This research was supported by ARPA through the U.S. Office of Naval Research under Grant N00014-93-1-1202, R&T Project Code 4424341-01.

References

[1] R. Deriche, Z. Zhang, Q.T. Luong, and O.D. Faugeras. Robust recovery of the epipolar geometry for an uncalibrated stereo rig. In *Proceedings of the European Conference on Computer Vision*, pages 567–576, Stockholm, Sweden, May 1994. Springer-Verlag, LNCS 800.

[2] O.D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proceedings of the European Conference on Computer Vision*, pages 563–578, Santa Margherita Ligure, Italy, June 1992.

[3] O.D. Faugeras, Q.T. Luong, and S.J. Maybank. Camera self calibration: Theory and experiments. In *Proceedings of the European Conference on Computer Vision*, pages 321–334, Santa Margherita Ligure, Italy, June 1992.

[4] O.D. Faugeras and F. Lustman. Let us suppose that the world is piecewise planar. In O. D. Faugeras and Georges Giralt, editors, *International Symposium on Robotics Research*, pages 33–40. MIT Press, Cambridge, MA, 1986.

[5] R. Hartley. Lines and points in three views — a unified approach. In *Proceedings Image Understanding Workshop*, Monterey, CA, November 1994.

[6] R. Hartley. Projective reconstruction and invariants from multiple images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(10):1036–1040, 1994.

[7] R. Hartley. Projective reconstruction from line correspondences. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WI, June 1994. IEEE Computer Society Press.

[8] Q.T. Luong and T. Vieville. Canonic representations for the geometries of multiple projective views. In *Proceedings of the European Conference on Computer Vision*, pages 589–599, Stockholm, Sweden, May 1994. Springer Verlag, LNCS 800.

[9] L. Robert and O.D. Faugeras. Relative 3D positioning and 3D convex hull computation from a weakly calibrated stereo pair. In *Proceedings of the International Conference on Computer Vision*, pages 540–544, Berlin, Germany, May 1993.

[10] A. Shashua. Projective structure from uncalibrated images: structure from motion and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):778–790, 1994.

[11] A. Shashua. Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995. in press.

[12] A. Shashua and N. Navab. Relative affine structure: Theory and application to 3D reconstruction from perspective views. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 483–489, Seattle, Washington, 1994.

[13] M.E. Spetsakis and J. Aloimonos. Structure from motion using line correspondences. *International Journal of Computer Vision*, 4(3):171–183, 1990.

[14] R. Tsai and T.S. Huang. Estimating three-dimensional motion parameters of a rigid planar patch, II: singular value decomposition. *IEEE Trans. on Acoustic, Speech and Signal Processing*, 30, 1982.

[15] J. Weng, T.S. Huang, and N. Ahuja. Motion and structure from line correspondences: Closed form solution, uniqueness and optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(3), 1992.