MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY
and
CENTER FOR BIOLOGICAL AND COMPUTATIONAL LEARNING
DEPARTMENT OF BRAIN AND COGNITIVE SCIENCES

A.I. Memo No. 1620 December, 1997
C.B.C.L Paper No. 157

On Degeneracy of Linear Reconstruction from Three Views: Linear Line Complex and Applications

Gideon P. Stein

Artificial Intelligence Laboratory
MIT
Cambridge, MA 02139
gideon@ai.mit.edu

Amnon Shashua

Institute of Computer Science
Hebrew University of Jerusalem
Jerusalem 91904, Israel
http://www.cs.huji.ac.il/~ shashua/

This publication can be retrieved by anonymous ftp to publications.ai.mit.edu. The pathname for this publication is: ai-publications/1500-1999/AIM-1620.ps.Z

Abstract

This paper investigates the linear degeneracies of projective structure estimation from point and line features across three views. We show that the rank of the linear system of equations for recovering the trilinear tensor of three views reduces to 23 (instead of 26) in the case when the scene is a Linear Line Complex (set of lines in space intersecting at a common line) and is 21 when the scene is planar. The LLC situation is only linearly degenerate, and we show that one can obtain a unique solution when the admissibility constraints of the tensor are accounted for.

The line configuration described by an LLC, rather than being some obscure case, is in fact quite typical. It includes, as a particular example, the case of a camera moving down a hallway in an office environment or down an urban street. Furthermore, an LLC situation may occur as an artifact such as in direct estimation from spatio-temporal derivatives of image brightness. Therefore, an investigation into degeneracies and their remedy is important also in practice.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology and at the Center for Biological and Computational Learning (CBCL). Support for this research was provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-94-01-0994. G.S. would like acknowledge the financial support from ONR contracts N00014-94-1-0128 and DARPA contracts N00014-94-01-0994, N00014-97-0363 A.S. wishes to thank CBCL for hosting him during the summer of 1997 when this work was performed. A.S. wishes to thank Steve Maybank, Shai Avidan and Nassir Navab for helpful comments and acknowledge the financial support from US-IS Binational Science Foundation 94-00120/2, the European ACTS project AC074 "Vanguard", and from DARPA through ARL Contract DAAL01-97-0101.

1 Introduction

It is known that point and line image features across three perspective views can generally give rise to a linear system of equations for a unique solution for 3D structure and camera motion. The structure and motion parameters are represented by a $3 \times 3 \times 3$ tensor, and the image measurements of matching points and lines provide constraints, trilinear in image coordinates, that as a whole make a linear system of equations for the (unknown) coefficients of the tensor. Finally, the tensor has only 18 degrees of freedom, i.e., the 27 coefficients are subject to non-linear admissibility constraints. In the presence of errors in image measurements one often starts with the Linear solution and improves it further by employing a numerical Gauss-Newton style iterative procedure until a solution that satisfies the admissibility constraints is obtained. (see Appendix for more details).

In this paper we investigate the cases in which the linear solution is degenerate. As it happens, the degeneracy occurs in typical real situations. We show that when the sample of features is taken from a configuration of lines that have a common intersection, known as a Linear Line Complex (LLC), then the rank of the linear system reduces from 26 (in the general case) to 23 — yet, there exists a *unique* solution for the tensor when the non-linear admissibility constraints are accounted for[1]. An LLC includes in particular the case of lines on parallel planes whose degeneracy was observed in [21].

To appreciate the practical importance of investigating LLC configurations, consider a few typical outdoor and indoor scene examples depicted in Fig.1. In Fig.1a the common intersecting line is the edge of the building. All horizontal lines on the two faces of the building meet the edge in the image plane, and the vertical lines meet the edge at infinity. Note also that the vertical line representing the lamp-post also meets the edge of the building (at infinity) thereby included in the LLC configuration. This leaves very few lines (the sidewalk and the oblique line of the lamp-post) not part of the LLC. The common line in Fig.1b is the edge of the book-case leaning on the wall. All vertical lines in the scene meet it at infinity and the horizontal lines of the file cabinet and the arm chair meet it in the image plane. Again, very few lines in the scene do not belong to the LLC (the horizontal line segments of the box near the far end wall and two horizontal line segments attached to the book-case). Next, all lines (except the short line segments of the ceiling light) in the hallway scene in Fig. 1c belong to an LLC whose common line is the vertical edge defined by the meeting of the front wall and the left side wall. Likewise, Fig. 1d is an LLC (with the exception of one short line segment) where the common line is the meeting between the file cabinet and the wall on the right.

Finally, an LLC situation occurs also as an artifact in direct estimation of the Tensor from spatio-temporal derivatives of image brightness [18]. The spatio-temporal derivatives provide an axis of certainty (a one-dimensional uncertainty) for the location of the matching points in views 1,2 relative to points in the reference view 0. The uncertainty axes in views 1,2 are parallel which means that the information gathered from a *general* scene by means of first-order spatio-temporal derivatives is at most comparable to the information gathered from an LLC configuration of discrete matching lines.

Given our main result, an attempt to reconstruct structure and motion from the image line information of the scenes in Fig. 1 using conventional approaches would be at best unstable. The linear system of equations is singular or near singular, and would most likely not serve as a reasonable starting solution for the subsequent Gauss-Newton iterations. Therefore an investigation into degeneracies caused by an LLC and their remedy is important also in practice.

The remainder of the paper is organized as follows. Section 2 contains the main results which include the statement of degeneracy of the linear system forming a null space of dimension 4, and the statement of uniqueness by incorporating the admissibility constraints with a simple constructive algorithm for obtaining a unique solution from an LLC configuration. In Section 3 we discuss the dimension of the null space for a planar object of points, and in Section 4 we verify the theory and the algorithm with experiments with real images. We use a schematized version of the real scene shown in Fig. 1a because it allows for a wider set of experiments. One can accurately find both line and point correspondences and can therefore perform the motion estimation using line correspondences and then verify the results against motion estimates obtained using points.

Notations in general, and tensorial notations in particular, theory and background of the Trilinear Tensor with its contraction and slicing properties and admissibility constraints, are discussed in the Appendix.

2 Linear Line Complex Scene Structure

Consider the tensor $\mathcal{T}_i^{jk}$ applied to the point-line-line configuration:

$$s'_j s''_k (p^i \mathcal{T}_i^{jk}) = 0,$$

where p is a point in image 1 and s', s'' are lines coincident with the matching point p', p'' in image 2 and 3, respectively. Note that $p^i \mathcal{T}_i^{jk}$ is a 3×3 matrix determined by p, which we will denote by B_p, i.e., in matrix notation $s''^\top B_p s' = 0$ for all pairs of lines coincident with p', p''. Assume that there exists a matrix B, independent of p, such that $s''^\top B s' = 0$, then clearly the tensor $\mathcal{T}_i^{jk}$ is not unique: slice the tensor into three matrices $(\mathcal{T}_1^{jk}, \mathcal{T}_2^{jk}, \mathcal{T}_3^{jk})$, then the tensors $(B,0,0), (0,B,0)$ and $(0,0,B)$ (and their linear combinations) all satisfy the constraint $s'_j s''_k p^i \mathcal{T}_i^{jk} = 0$. Hence, such a matrix B does not exist in general. We may, nevertheless, ask *whether there exists a special configuration of points and lines in space for which such a matrix B is valid?* Such a configuration is a Linear Line Complex (LLC):

Theorem 1 *Let $\mathcal{S}$ be a set of lines in 3D which have a common intersecting line L (i.e., $S \wedge L = 0$ for all $S \in \mathcal{S}$). Let $\mathcal{Q}$ be a set of lines in 3D that intersect the*

[1] This in contrast to *critical line configurations* from which a unique solution is not possible, see [9, 3, 10]

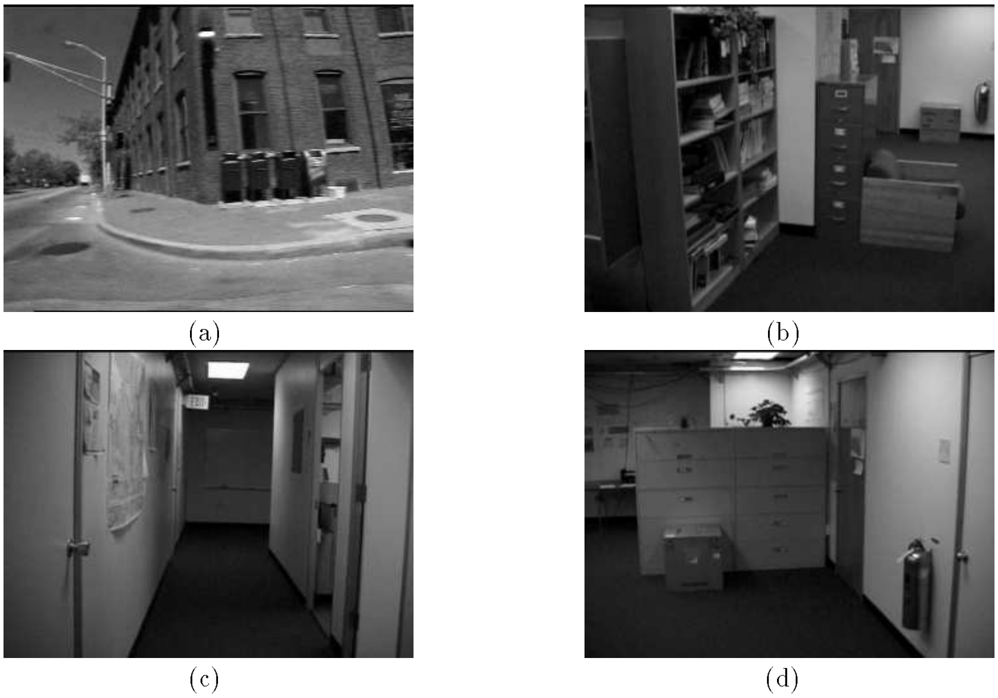


Figure 1: *Typical urban indoor and outdoor scenes. The lines in the images form a Linear Line Complex. See text for more details.*

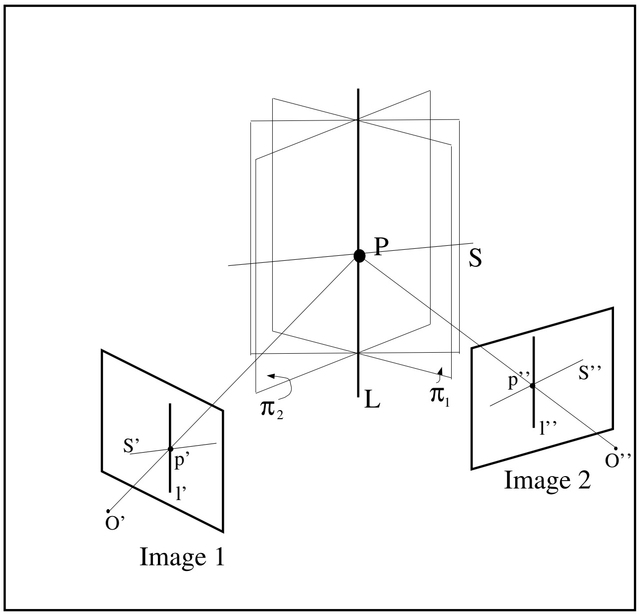


Figure 2: *Figure to accompany theorem 1.*

line joining the two camera centers. Then, there exists a unique matrix B satisfying $s''^{\top} B s' = 0$ for all pairs of projections s', s'' of lines $S \in \mathcal{S}$ onto two distinct views. The matrix B also satisfies $q''^{\top} B q' = 0$ for all pairs of projections q', q'' of lines $Q \in \mathcal{Q}$.

Proof: Let P be the intersection of a line $S \in \mathcal{S}$ with L and denote its projections by p', p'' onto views 1,2 respectively (see Fig. 2). Choose any plane π from the pencil of planes meeting at the line L, and let H_π be the corresponding 2D projective mapping (homography matrix) of points in view 1 to points in view 2 via (projections of) the plane π. Since π contains the line L, then

$$H_\pi p' \cong p''.$$

Let l', l'' be the projections of L, then p' is the intersection of s' and l', thus,

$$H_\pi [l']_x s' \cong p'',$$

where $[l']_x$ denotes the skew-symmetric matrix of cross products, i.e., $l' \times s' = [l']_x s'$. Likewise, s'' is coincident with p'', then

$$s''^{\top} H_\pi [l']_x s' = 0.$$

Denote $B_\pi = H_\pi [l']_x$. We show next that B_π is unique, i.e., independent of the choice of π. Let π_1, π_2 be two distinct planes of the pencil and let H_{π_1}, H_{π_2} be their corresponding homography matrices. It is known that any two homography matrices between two fixed views satisfy,

$$H_{\pi_2} \cong \lambda H_{\pi_1} + e'' n^{\top}$$

where e'' is the projection of the optical center of camera 1 onto the image plane of camera 2 (the epipole), and

n is a free vector. Because π_1, π_2 intersect at L, then, $H_{\pi_1} u \cong H_{\pi_2} u$ for all $u^\top l' = 0$, thus $n \cong l'$ and we have:

$$H_{\pi_2} \cong \lambda H_{\pi_1} + e'' l'^\top,$$

and from which it clearly follows that $B_{\pi_1} = B_{\pi_2}$.

Let D be the intersection of a line $Q \in \mathcal{Q}$ with the plane π and denote its projections by d', d'' onto views 1 and 2. The image line q' passes through the point d' and through the epipole e' and therefore: $q' \cong e' \times d'$. and similarly $q'' \cong e'' \times d''$. We can then write:

$$\begin{aligned} q''^\top B q' &= (e'' \times d'')^\top H [l']_x (e' \times d') \\ &= (e'' \times d'')^\top H (d' \cdot l') e' - (e'' \times d'')^\top H (l' \cdot e') d' \\ &= (d' \cdot l')(e'' \times d'')^\top e'' - (l' \cdot e')(e'' \times d'')^\top d'' \\ &= 0 \end{aligned} \quad (1)$$

where we used the identity:

$$a \times (b \times c) = (c \cdot a) b - (a \cdot b) c \quad (2)$$

and the fact that the homography H maps d' to d'' and e' to e''. $\square$

Corollary 1 *The rank of the estimation matrix of the tensor from image measurements of lines across three views of a Linear Line Complex structure is at most 23.*

Proof: Let the tensor $\mathcal{T}_i^{jk}$ be sliced into three matrices $(\mathcal{T}_1^{jk}, \mathcal{T}_2^{jk}, \mathcal{T}_3^{jk})$, then the tensors $(B, 0, 0), (0, B, 0)$ and $(0, 0, B)$ (and their linear combinations) span the tensors of the form:

$$\mathcal{T}_i^{jk} = \delta_i b^{jk}$$

where δ is a free vector of the family. Then,

$$s'_j s''_k p^i \mathcal{T}_i^{jk} = (p^i \delta_i)(s'_j s''_k b^{jk}) = 0.$$

Since $\delta_i b^{jk}$ does not include the general form of trilinear tensors (eqn. 4), the null space of the estimation matrix includes at least four distinct vectors: the true tensor describing the relative location of the three cameras, and the three 'ghost' tensors $(B, 0, 0), (0, B, 0)$ and $(0, 0, B)$. Thus, the rank is at most $27 - 4 = 23$. $\square$

The ambiguity can be further reduced by incorporating the tensor admissibility constraints (see Appendix) as detailed below.

Theorem 2 *The ambiguity of Tensor estimation from measurements coming from an LLC structure is* **at most** *an 8-fold ambiguity.*

Proof: We assume the correlation matrix slicing of the tensor into the three standard correlation matrices $(\mathcal{T}_1^{jk}, \mathcal{T}_2^{jk}, \mathcal{T}_3^{jk})$ (see Appendix). Let W be the $N \times 27$, $N \geq 27$, estimation matrix for linear estimation of the tensor, i.e., $Wv = 0$ where v is the tensor whose elements are spread as a 27 element vector, and v is spanned by the four-dimensional null space of $W^\top W$. Let v_1, v_2, v_3 be the three 'ghost' tensors corresponding to $(B, 0, 0), (0, B, 0)$ and $(0, 0, B)$, respectively. Let v_0 be the (one dimensional) null space of

$$W^\top W - v_1 v_1^\top - v_2 v_2^\top - v_3 v_3^\top.$$

Since the null space span the admissible tensors, the three standard correlation matrices (T_1, T_2, T_3) of the admissible tensors are spanned by the tensors $v_0, ..., v_3$, i.e.,

$$\begin{aligned} T_1 &= \hat{T}_1 + \alpha_1 B \\ T_2 &= \hat{T}_2 + \alpha_2 B \\ T_3 &= \hat{T}_3 + \alpha_3 B \end{aligned}$$

where $\hat{T}_i$, $i = 1, 2, 3$, are the standard correlation matrices of the tensor v_0, and α_i are scalars. As part of the admissibility constraints (see Appendix), the standard correlation matrices T_i must be of rank 2, thus α_i are generalized eigenvalues of $\hat{T}_i$ and B, and since B is of rank 2, the characteristic equation for each α_i is of second order. Thus, we have at most 8 distinct solutions for T_i. $\square$

Empirical Observation 1 *Only one of the 8 solutions satisfies all the admissibility constraints.*

Explanation: the rank-2 constraint of the standard correlation matrices is closed under linear superposition (see Appendix). Numerical experiments show that only one out of the 8 possible solutions for the generalized eigenvalues α_1, α_2 and α_3 produces standard correlation matrices T_1, T_2, T_3 whose linear superpositions produce rank-2 matrices. $\square$

2.1 Algorithm for recovering structure and motion in the LLC case

1. Using robust estimation techniques determine the line correspondences which belong to the LLC and compute the matrix B (see theorem 1). Here one might use a robust version of the 8 point algorithm [7].
2. From the matrix B create the 3 'ghost' tensors: $v_1 = (B, 0, 0), v_2 = (0, B, 0)$ and $v_3 = (0, 0, B)$.
3. Using the point-line-line correspondences from the 3 views compute W, the $N \times 27, N \geq 27$ estimation matrix for the linear estimation of the tensor.
4. Find v_0, the 4th vector spanning the (row) null space of W orthogonal to v_1, v_2 and v_3 by finding the null space of:
$$W^\top W - v_1 v_1^\top - v_2 v_2^\top - v_3 v_3^\top.$$
In practice take the eigenvector corresponding to the smallest eigenvalue.
5. Find scalars α_i such that the vector:
$$v = v_0 + \sum_{i=1}^{3} \alpha_i v_i$$
is an admissible tensor (see theorem 2). This is done in two stages:
   (a) Let $\hat{T}_i$, T_i , $i = 1, 2, 3$, be the standard correlation matrices of the tensors v_0 and v respectively. Then:
   $$T_I = \hat{T}_i + \alpha_i B$$
   Enforce the constraint that T_i is of rank-2 to find α_i. Since the matrix B is of rank-2 this is quadratic constraint resulting in up to 2 solutions for each α_i for a total of $2^3 = 8$ solutions.
   (b) Prune the number of solutions down to one by enforcing the stronger admissibility constraint that any linear superposition of matrices T_i must be of rank-2. This is done by generating K random sets

of linear coefficients δ_i such that $\sum \delta_i^2 = 1$ and computing the determinant of the linear superposition: $\sum \delta_i T_i$ for each of the 8 possible solutions. The solution that consistently gives $\det(\delta_i T_i) \simeq 0$ is the correct solution.

3 The Case of Planar Configurations

Consider again the point-line-line contraction:

$$p^i s'_j (s''_k \mathcal{T}_i^{jk}) = 0.$$

Denote the matrix $s''_k \mathcal{T}_i^{jk}$ by $E_{s''}$, i.e., in matrix notation we have $s'^{\top} E_{s''} p = 0$. If there exists a matrix E, independent of s'', such that $s'^{\top} E p = 0$ for all lines s' coincident with the matching point p', then clearly the tensor $\mathcal{T}_i^{jk}$ is not unique: slice the tensor into three matrices $(\mathcal{T}_i^{j1}, \mathcal{T}_i^{j2}, \mathcal{T}_i^{j3})$, then the tensors $(E, 0, 0), (0, E, 0)$ and $(0, 0, E)$ (and their linear combinations) all satisfy the constraint $p^i s'_j s''_k \mathcal{T}_i^{jk} = 0$. Hence, such a matrix E does not exist in general. However, if the matching points p, p' are projections of a *coplanar* configuration of points π in space and E is the corresponding homography matrix $Ep \cong p'$, then $s'^{\top} E p = 0$ for all lines s' coincident with p'.

Likewise, let W be the homography matrix due to π, i.e., $Wp \cong p''$, then $s''^{\top} W p = 0$ for all lines s'' coincident with p''. Then, given the slicing of the tensor into three matrices $(\mathcal{T}_i^{1k}, \mathcal{T}_i^{2k}, \mathcal{T}_i^{3k})$, then the tensors $(W, 0, 0), (0, W, 0)$ and $(0, 0, W)$ (and their linear combinations) all satisfy the constraint $p^i s'_j s''_k \mathcal{T}_i^{jk} = 0$.

Therefore, the rank of the null space of the linear system of equations for the tensor is at least 6, since we have just created 6 'ghost' tensors. The question of whether the ghost tensors include the true tensor (in which case the rank of the estimation matrix is 21) or not (rank is 20) is settled below.

Theorem 3 *The rank of the estimation matrix of the tensor from image measurements of three views of a planar configuration of points is at most 21.*

proof: Denote the planar object by π and let E, W be the homography matrices due to π from view 0 to 1, and from view 0 to 2, respectively. The 'ghost' tensors due to E span the tensors of the form:

$$\mathcal{T}_i^{jk} = \delta^k e_i^j$$

where δ is a free vector of the family. Then,

$$p^i s'_j s''_k \mathcal{T}_i^{jk} = (\delta^k s''_k)(p^i s'_j e_i^j) = 0.$$

Likewise, The 'ghost' tensors due to W span the tensors of the form:

$$\mathcal{T}_i^{jk} = \mu^j w_i^k$$

where μ is a free vector of the family. Then,

$$p^i s'_j s''_k \mathcal{T}_i^{jk} = (\mu^j s'_j)(p^i s''_k e_i^k) = 0.$$

The 6 dimensional null space spanned by both families of 'ghost' tensors spans the tensors of the form:

$$\mathcal{T}_i^{jk} = \mu^j w_i^k + \delta^k e_i^j$$

where δ, μ are free vectors of the family. This family includes the true tensor (set $\mu = v'$ and $\delta = v''$). Thus, the rank of the estimation matrix is at most 27-6 = 21.$\square$

Note, that unlike the case of LLC in which the admissibility constraints have reduced the ambiguity to a single solution, here the null space includes admissible tensors (admissibility constraints are satisfied), thus a unique solution is not possible. The ambiguity is also evident by straightforward counting: the tensor is determined by 18 (algebraically independent) parameters, yet two homography matrices (E, W) give rise only to 8 parameters each (because each matrix is up to scale), thus we have 2 parameters missing for uniquely determining the tensor from a planar surface.

4 Experiments

4.1 The experimental procedure

Fig.3a,3b, and 3c show the three input images used for the experiments. The scene is composed of two faces of a cube and another plane on the left which is parallel to the vertical edge of the cube. This is a schematic model of a typical urban scene with an edge of a building such as in Fig 1a.

Corresponding point features were manually extracted. The feature points were saddle points formed by the corners of two black squares which can be found with subpixel accuracy. The point features were grouped into four groups: Points from the left and right faces of the cube form one group each. Points on the planar surface were grouped into two vertical sets of features.

Line features were created by taking pairs of points. If no pair of points has members from more than one group (for example Fig. 3) then we limit ourselves to a Linear Line Complex since all the 3D lines in the scene intersect the edge of the cube. By adding pairs that span two groups we can add lines that do not belong to the LLC. By judiciously choosing pairs we can add lines that are close or far from being part of the LLC (see Fig. 3d). We can also choose pairs of points that define lines passing through the epipoles. This flexibility allows us to verify all the claims in theorem 1.

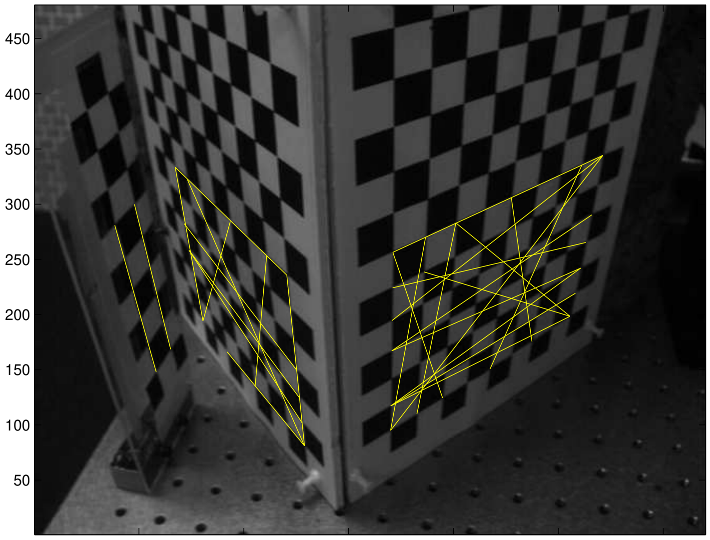
4.1.1 Hardware notes

The images were captured using a Pulnix TM9701 progressive scan camera with a $2/3_{inch}$ CCD and an 8.5_{mm} lens. The image resolution was $640 \times 480_{pixels}$.

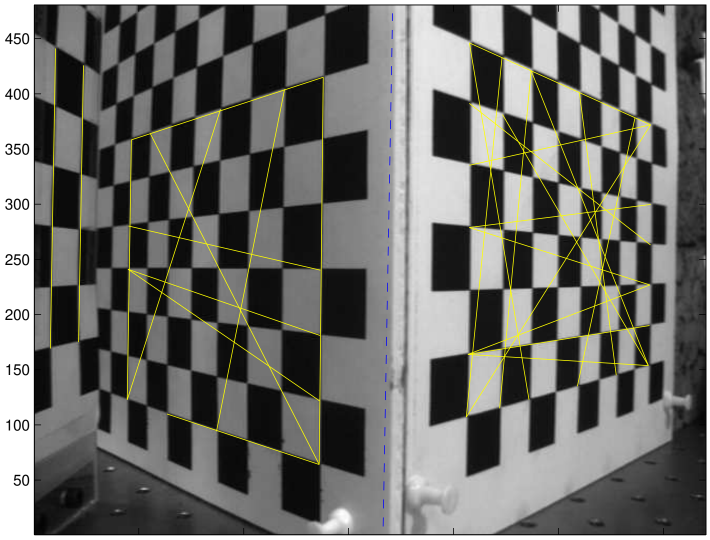
To achieve the results presented here we had to take into account radial lens distortion. Only the first term of radial distortion was used. The radial distortion parameter, $K_1 = 6e-7$ was found using the method described in [19]. We note that that parameter value also minimized the error terms in equation 3.

4.2 Determining the LLC

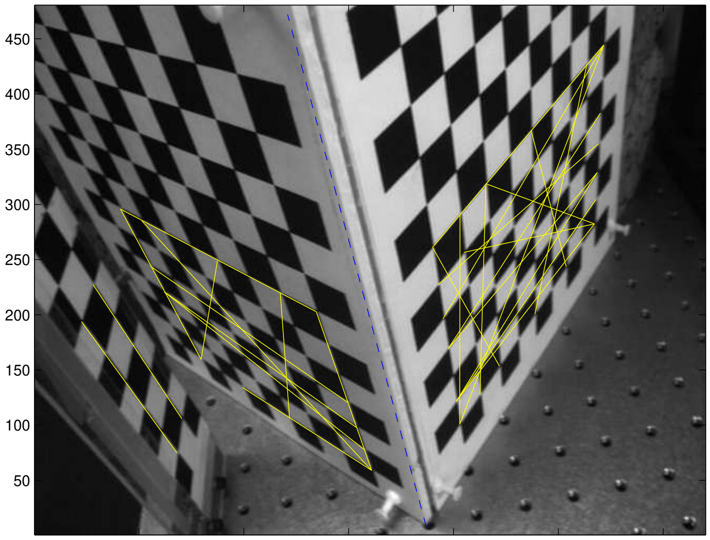
The three input images (Figures 3a,3b, and 3c) will be denoted image 1, 2 and 3 respectively. We chose $N = 28$ pairs of points which defined lines all belonging to the LLC. These are overlaid as white lines in the figures. For each image pair (1,2), (2,3) and (1,3) we used the 8 point algorithm [7] applied to the line correspondences to



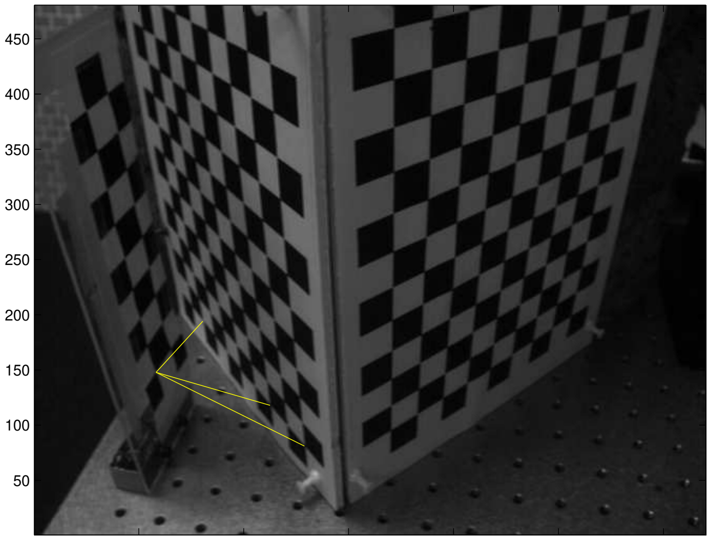
(a)



(b)



(c)



(d)

Figure 3: *The three input images used. All the lines marked are part of a linear line complex. They all intersect the line defined by the edge of the cube. Vertical lines intersect the edge at the point at infinity. The LLC was computed using images (b) and (c). The dashed lines in (b) and (c) are the projection of the common intersection line into the images. The results show it aligns very well with the edge of the true cube. (d) Three line which are not part of the LLC that are used in the experiments.*

Table 1: *Values of the error cost function for estimating the LLC when all the lines belong to the LLC (none) and when we add a line which passes close to or far from the common line of intersection (the edge of the cube).*

Extra Line	E_{23}	E_{12}	E_{13}
None	0.000055	0.00064	0.000079
Close	0.0019	0.0016	0.00054
Middle	0.0065	0.0038	0.0017
Far	0.0119	0.0919	0.0062

compute the matrices B_{12} , B_{23} and B_{13} that minimize:

$$E_{12} = \frac{1}{N}\sum_{i=1}^{N}(s_i B_{12} s_i')^2 \quad (3)$$

$$E_{23} = \frac{1}{N}\sum_{i=1}^{N}(s_i B_{23} s_i'')^2$$

$$E_{13} = \frac{1}{N}\sum_{i=1}^{N}(s_i B_{13} s_i'')^2$$

respectively. The coordinates of the lines s, s', s'' have been scaled as described in [7]. From theorem 1, the left and right null spaces of B_{23} (for example) are the projections of the line L in images 2 and 3. The dashed black line in figures (3b), and (3c) show the lines corresponding to the null spaces overlaid on the input images. They align well with the edge of the cube verifying the theory and showing that the matrix B can be recovered accurately. Similar results were found using the other image pairs.

Fig.3d shows image 1 on which we have overlaid three lines not belonging to the LLC. Table 1 shows the error terms of equations 3 when all the line are from the LLC and when we add one of the lines shown in figure 3d. When the extra line is far away from the common line of intersection the error is large. Even when the line nearly intersects the edge of the cube the error is still significant. Therefore if most of the line come from an LLC robust methods can be used to detect outliers. Other experiments, not reported here, use lines that pass through the epipole to verify the second half of theorem 1.

4.3 Recovering motion and structure

We computed the motion tensor from the three views using four methods. First we used the linear method for a set of 131 point correspondences. Then we used the linear method for a set of 34 non degenerate line correspondences. Next we applied the linear method in a naive way to the set-of 28 line correspondences from an LLC. In other words we ignored the fact that the lines come from an LLC. Finally we estimated the tensor from the 28 degenerate lines using the algorithm described in section 2.1.

4.3.1 Condition of the estimation matrix

Figure 4a (top) shows the four smallest singular values of the estimation matrix W used to compute the tensor from 34 non-degenerate lines. The smallest eigenvalue is considerably smaller than the others indicating that the null space of the matrix is of $rank = 1$ and the problem is well conditioned. Figure 4a (middle) shows the 5 smallest singular values for the estimation matrix computed from 28 lines belonging to an LLC. The 4 smallest singular values are about equal and are considerably smaller than the next smallest. This indicates that the null space of W is of $rank = 4$ as expected from theorem 1. Simply taking the eigenvector corresponding to the smallest eigenvalue would be a mistake.

Figure 4b shows the projection of the vectors v_1, v_2 and v_3 on the eigenvectors of the estimation matrix $W^\top W$. The projections onto the eigenvectors corresponding to the first 23 eigenvalues are close to zero verifying that the vectors v_1, v_2 and v_3 are orthogonal to the first 23 eigenvectors. This verifies part of theorem 2 which states that the vectors v_1, v_2 and v_3 are in the null space of $W^\top W$ (the remaining four eigenvectors).

Following the algorithm described in section 2.1 we compute the eigenvalues and eigenvectors of the matrix

$$W^\top W - v_1 v_1^\top - v_2 v_2^\top - v_3 v_3^\top .$$

Figure 4a (bottom) shows the three smallest eigenvalues. The smallest eigenvalues is significantly smaller than the next smallest value indicating that the null space is now of $rank = 1$.

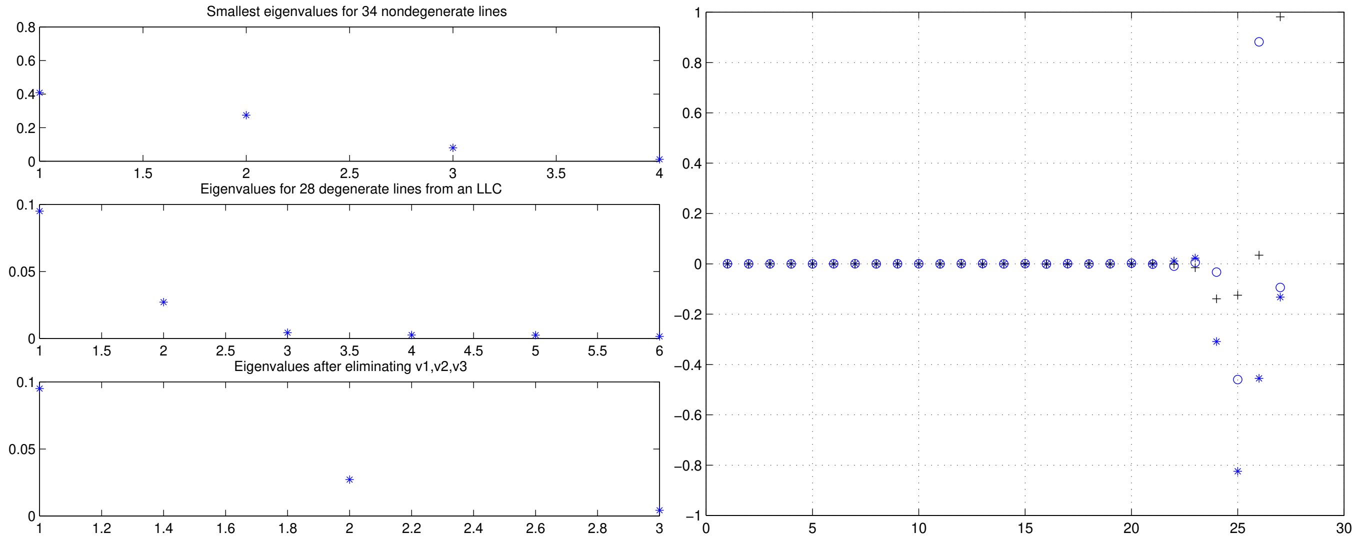
4.3.2 Reprojection of lines using the tensor

After recovering the tensor one can use the tensor to reproject a line given in two images into the third image. In order to test the tensor estimates we used ten additional lines shown in figure 5. Three of the lines lie in the LLC on the left face of the cube. The other 7 do not lie on the LLC.

Figures 5b,c,d show the reprojection results (dashed lines) together with the original lines (solid lines) overlaid on image 1. One can see that if the set of lines used to estimate the tensor all belong to an LLC then other lines in the LLC reproject more or less correctly but the reprojection of lines not in the LLC is incorrect (fig. 5c). On the other hand, reprojection using the tensor computed by taking into account the degeneracy (fig. 5d) gives results as good as if we had a non-degenerate set of lines to estimate the tensor (fig. 5b).

5 Summary

We have shown that linear methods for estimating motion and 3D structure from lines lead to a degenerate set of equations in the case of a Linear Line Complex. The LLC, a configuration of lines that all intersect a common line in $\mathcal{P}^3$, is in fact a common configuration of lines occuring frequently in man-made environments. This degeneracy is due to a bilinear constraint on lines in two views where the constraint equation has a form similar to the epipolar constraint but where lines replace points and the epipoles are replaced by the image of the common line of intersection in the two views. This constraint can be used to determine whether a set of lines belongs to an LLC and enables us to reject a outliers using *least median of squares* or other robust estimation methods.



(a) (b)

Figure 4: *(a) The smallest singular values of the estimation matrix W for a degenerate and non degenerate set of lines. (b) With a degenerate set of lines, the vectors v_1, v_2 and v_3 were projected onto the 27 eigenvectors of $W^\top W$. The values for the first 23 eigenvectors are close to zero verifying that the vectors v_1, v_2 and v_3 are orthogonal to the first 23 eigenvectors and are therefore in the null space of $W^\top W$. (See theorem 2.)*

An LLC is not degenerate for non-linear methods in general. The theoretical analysis leads to a modification of the linear methods that can recover the structure and motion in the LLC case. We have proven that the motion can be recovered up to 8 discrete solutions. Empirical evidence shows that the number of solutions can be reduced further to a single solution one.

We have implemented the algorithm, and experiments with real images verify the theoretical analysis. Although the results using the modified linear algorithm compare favorably with the results obtained using a non-degenerate set of lines, the system at this point is not robust. For example, it requires that lens distortion be taken into account. Further engineering would be involved in making a practical system.

Acknowledgments

A.S. wishes to thank Steve Maybank, Shai Avidan and Nassir Navab for helpful comments and acknowledge the financial support from US-IS Binational Science Foundation 94-00120/2, the European ACTS project AC074 "Vanguard", and from DARPA through ARL Contract DAAL01-97-0101. G.S. would like acknowledge the financial support from ONR contracts N00014-94-1-0128 and DARPA contracts N00014-94-01-0994, N00014-97-0363.

References

[1] S. Avidan and A. Shashua. Tensorial transfer: On the representation of $n > 3$ views of a 3D scene. In *Proceedings of the ARPA Image Understanding Workshop*, Palm Springs, CA, February 1996.

[2] S. Avidan and A. Shashua. View synthesis in tensor space. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Puerto Rico, June 1997.

[3] T. Buchanan. On the critical set for photogrammetric reconstruction using line tokens in p3(c). *Geometriae Dedicata*, 44:223–232, 1992.

[4] O.D. Faugeras and B. Mourrain. On the geometry and algebra of the point and line correspondences between N images. In *Proceedings of the International Conference on Computer Vision*, Cambridge, MA, June 1995.

[5] O.D. Faugeras and T. Papadopoulo. A nonlinear method for estimating the projective geometry of three views. To be published ICCV 98, January 1998.

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

[7] R. Hartley. A linear method for reconstruction from lines and points. In *Proceedings of the International Conference on Computer Vision*, pages 882–887, Cambridge, MA, June 1995.

[8] A. Heyden. Reconstruction from image sequences by means of relative depths. In *Proceedings of the International Conference on Computer Vision*, pages 1058–1063, Cambridge, MA, June 1995.

[9] S.J. Maybank. The critical line concruence for reconstruction from three images. *Applicable Algebra in Engineering Communication and Computing (AAECC)*, 6:89–113, 1995.

[10] N. Navab, O.D. Faugeras, and T. Vieville. The critical sets of lines for camera displacement estimation: A mixed Euclidean-projective and constructive approach. In *Proceedings of the International Conference on Computer Vision*, Berlin, Germany, May 1993.

[11] A. Shashua. Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):779–789, 1995.

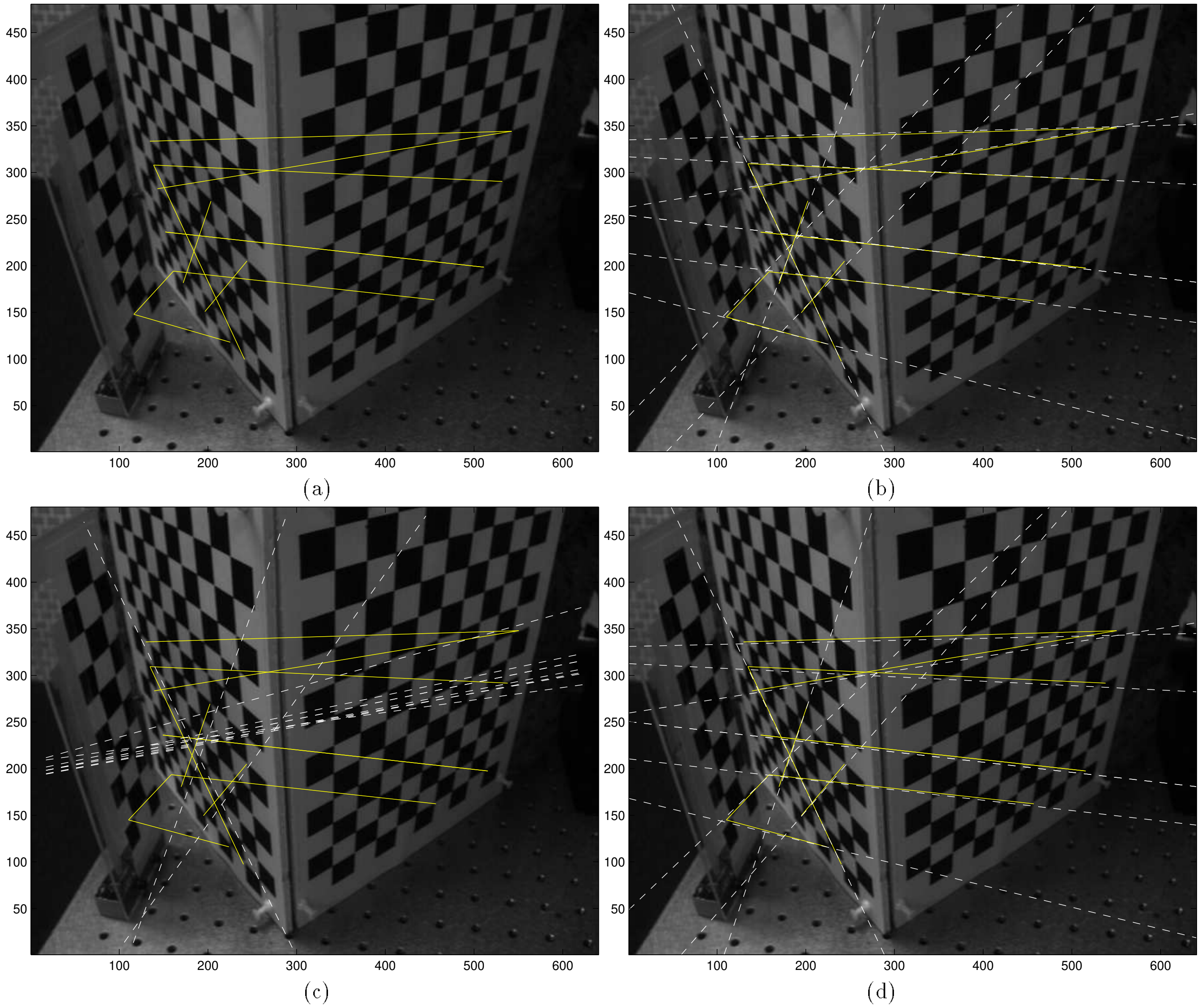


Figure 5: *(a) Ten additional lines are used to test the recovered tensors. Three of the lines lie on the left face of the cube and are therefore the part of the LLC. The other 7 are not. The test lines from images (2) and (3) are reprojected back into image (1) using the recovered tensors. Solid lines are the true locations. Dashed lines are the reprojected lines. (b) The tensor was computed using a set of 34 non degenerate lines. (c) The tensor was computed using a degenerate set of 28 lines using the standard method. Since it does not take into account the degeneracy the tensor fails to correctly reproject the lines. The only lines in (c) that reproject correctly are those that belong to the LLC. (d) Using the new algorithm for the degenerate case the computed tensor correctly reprojects all the lines.*

[12] A. Shashua. Trilinear tensor: The fundamental construct of multiple-view geometry and its applications. Submitted for journal publication, 1997.

[13] A. Shashua and P. Anandan. The generalized trilinear constraints and the uncertainty tensor. In *Proceedings of the ARPA Image Understanding Workshop*, Palm Springs, CA, February 1996.

[14] A. Shashua and S. Avidan. The rank4 constraint in multiple view geometry. In *Proceedings of the European Conference on Computer Vision*, Cambridge, UK, April 1996.

[15] A. Shashua and M. Werman. Trilinearity of three perspective views and its associated tensor. In *Proceedings of the International Conference on Computer Vision*, June 1995.

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

[17] M.E. Spetsakis and J. Aloimonos. A unified theory of structure from motion. In *Proceedings of the ARPA Image Understanding Workshop*, 1990.

[18] G. Stein and A. Shashua. Model based brightness constraints: On direct estimation of structure and motion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Puerto Rico, June 1997.

[19] G. Stein. Lens distortion calibration using point correspondences. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Puerto Rico, June 1997.

[20] B. Triggs. Matching constraints and the joint image. In *Proceedings of the International Conference on Computer Vision*, pages 338–343, Cambridge, MA, June 1995.

[21] 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.

A Mathematical Background and the Trilinear Tensor

Let $\boldsymbol{x}$ be a point in 3D space and its projection in a pair of images be p and p'. Then $p = [I;0]\boldsymbol{x}$ and $p' \cong A\boldsymbol{x}$, where $\cong$ denotes equality up to scale. The left 3×3 minor of A stands for a 2D projective transformation of the chosen plane at infinity and the fourth column of A stands for the epipole (the projection of the center of camera 0 on the image plane of camera 1). In particular, in a calibrated setting the 2D projective transformation is the rotational component of camera motion and the epipole is the translational component of camera motion.

We will occasionally use tensorial notations as described next. We use the covariant-contravariant summation convention: a point is an object whose coordinates are specified with superscripts, i.e., $p^i = (p^1, p^2, ...)$. These are called contravariant vectors. An element in the dual space (representing hyperplanes — lines in $\mathcal{P}^2$), is called a covariant vector and is represented by subscripts, i.e., $s_j = (s_1, s_2, ...)$. Indices repeated in covariant and contravariant forms are summed over, i.e., $p^i s_i = p^1 s_1 + p^2 s_2 + ... + p^n s_n$. This is known as a contraction. An outer-product of two 1-valence tensors (vectors), $a_i b^j$, is a 2-valence tensor (matrix) c_i^j whose i,j entries are $a_i b^j$ — note that in matrix form $C = ba^\top$.

Matching image points across three views will be denoted by p, p', p''; the homogeneous coordinates will be referred to as p^i, p'^j, p''^k, or alternatively as non-homogeneous image coordinates $(x,y), (x',y'), (x'',y'')$ — hence, $p^i = (x, y, 1)$, etc.

Three views, $p = [I;0]\boldsymbol{x}, p' \cong A\boldsymbol{x}$ and $p'' \cong B\boldsymbol{x}$, are known to produce four trilinear forms whose coefficients are arranged in a tensor representing a bilinear function of the camera matrices A, B:

$$\mathcal{T}_i^{jk} = v'^j b_i^k - v''^k a_i^j \tag{4}$$

where $A = [a_i^j, v'^j]$ (a_i^j is the 3×3 left minor and v' is the fourth column of A) and $B = [b_i^k, v''^k]$. The tensor acts on a triplet of matching points in the following way:

$$p^i s_j^\mu r_k^\rho \mathcal{T}_i^{jk} = 0 \tag{5}$$

where s_j^μ are any two lines (s_j^1 and s_j^2) intersecting at p', and r_k^ρ are any two lines intersecting p''. Since the free indices are μ, ρ each in the range 1,2, we have 4 trilinear equations (unique up to linear combinations). If we choose the *standard* form where s^μ (and r^ρ) represent vertical and horizontal scan lines, i.e.,

$$s_j^\mu = \begin{bmatrix} -1 & 0 & x' \\ 0 & -1 & y' \end{bmatrix}$$

then the four trilinear forms, referred to as *trilinearities* [11], have the following explicit form:

$$\begin{aligned}
x''\mathcal{T}_i^{13}p^i - x''x'\mathcal{T}_i^{33}p^i + x'\mathcal{T}_i^{31}p^i - \mathcal{T}_i^{11}p^i &= 0,\\
y''\mathcal{T}_i^{13}p^i - y''x'\mathcal{T}_i^{33}p^i + x'\mathcal{T}_i^{32}p^i - \mathcal{T}_i^{12}p^i &= 0,\\
x''\mathcal{T}_i^{23}p^i - x''y'\mathcal{T}_i^{33}p^i + y'\mathcal{T}_i^{31}p^i - \mathcal{T}_i^{21}p^i &= 0,\\
y''\mathcal{T}_i^{23}p^i - y''y'\mathcal{T}_i^{33}p^i + y'\mathcal{T}_i^{32}p^i - \mathcal{T}_i^{22}p^i &= 0.
\end{aligned}$$

These constraints were first derived in [11]; the tensorial derivation leading to eqns. 4 and 5 was first derived in [13]. The trilinear tensor has been well known in disguise in the context of Euclidean line correspondences and was not identified at the time as a tensor but as a collection of three matrices (a particular contraction of the tensor, correlation contractions, as explained next) [16, 17, 21]. The link between the two and the generalization to projective space was identified later by Hartley [6, 7]. Additional work in this area can be found in [15, 4, 20, 8, 14, 1], and applications in [2, 18].

The tensor has certain contraction properties and can be sliced in three principled ways into matrices with distinct geometric properties. These properties is what makes the tensor distinct from simply being a collection of three matrices and will be briefly discussed next — further details can be found in [12].

A.1 Contraction Properties and Tensor Slices

Consider the matrix arising from the contraction,

$$\delta_k \mathcal{T}_i^{jk} \tag{6}$$

which is a 3×3 matrix, we denote by E, obtained by the linear combination $E = \delta_1 \mathcal{T}_i^{j1} + \delta_2 \mathcal{T}_i^{j2} + \delta_3 \mathcal{T}_i^{j3}$ (which is what is meant by a contraction), and δ_k is an *arbitrary* covariant vector. The matrix E has a general meaning introduced in [15]:

Proposition 1 (Homography Contractions) *The contraction $\delta_k \mathcal{T}_i^{jk}$ for some arbitrary δ_k is a homography matrix from image one onto image two determined by the plane containing the third camera center C'' and the line δ_k in the third image plane. Generally, the rank of E is 3. Likewise, the contraction $\delta_j \mathcal{T}_i^{jk}$ is a homography matrix from image one onto image three.*

For proof see [15]. Clearly, since δ is spanned by three vectors, we can generate up to at most three distinct homography matrices by contractions of the tensor. We define the *Standard Homography Slicing* as the homography contractions associated by selecting δ be $(1,0,0)$ or $(0,1,0)$ or $(0,0,1)$, thus the three standard homography slices between image one and two are $\mathcal{T}_i^{j1}, \mathcal{T}_i^{j2}$ and $\mathcal{T}_i^{j3}$, and we denote them by E_1, E_2, E_3 respectively, and likewise the three standard homography slices between image one and three are $\mathcal{T}_i^{1k}, \mathcal{T}_i^{2k}$ and $\mathcal{T}_i^{3k}$, and we denote them by W_1, W_2, W_3 respectively.

Similarly, consider the contraction

$$\delta^i \mathcal{T}_i^{jk} \tag{7}$$

which is a 3×3 matrix, we denote by T, and where δ^i is an *arbitrary* contravariant vector. The matrix T has a general meaning is well, as detailed below [12]:

Proposition 2 *The contraction $\delta^i \mathcal{T}_i^{jk}$ for some arbitrary δ^i is a rank 2 correlation matrix from image two onto image three, that maps the dual image plane (the space of lines in image two) onto a set of collinear points in image three that form the epipolar line corresponding to the point δ^i in image one. The null space of the correlation matrix is the epipolar line of δ^i in image two. Similarly, the transpose of T is a correlation from image three onto image two with the null space being the epipolar line in image three corresponding to the point δ^i in image one.*

For proof see [12]. We define the *Standard Correlation Slicing* as the correlation contractions associated with selecting δ be $(1,0,0)$ or $(0,1,0)$ or $(0,0,1)$, thus the three standard correlation slices are $\mathcal{T}_1^{jk}, \mathcal{T}_2^{jk}$ and $\mathcal{T}_3^{jk}$, and we denote them by T_1, T_2, T_3, respectively. The three standard correlations date back to the work on structure from motion of lines across three views [16, 21] where these matrices were first introduced.

A.2 Tensor Admissibility Constraints

The 27 coefficients $\mathcal{T}_i^{jk}$ are not independent. One can easily show that the tensor is determined by only 18 parameters; and from the contraction properties discussed above that the constraints among the 27 coefficients, referred to as *admissibility constraints*, are grouped into three classes. Both will be discussed briefly below (further details in [12]).

A.2.1 18 Parameters

The tensor

$$\mathcal{T}_i^{jk} = v'^j b_i^k - v''^k a_i^j$$

is determined by 24 parameters given by the two camera matrices, each has 12 parameters. Two additional parameters drop out because we can scale v' and accordingly b_i^k without changing the tensor, and likewise scale v'' and accordingly a_i^j. An additional parameter drops out because of the global scale factor (tensor is determined up to overall scale). Thus, we readily see there can be at most 21 parameters defining the tensor. We can drop out three more parameters by noticing that the matrices a_i^j and b_i^k belong to a family of homography matrices that leaves the tensor unchanged (uniqueness proof in [11]), as detailed below:

$$\begin{aligned}
\mathcal{T}_i^{jk} &= v'^j b_i^k - v''^k a_i^j && (8)\\
&= v'^j (b_i^k + \alpha_i v''^k) - v''^k (a_i^j + \alpha_i v'^j) && (9)\\
&= \mathcal{T}_i^{jk} + \alpha_i v'^j v''^k - \alpha_i v'^j v''^k && (10)\\
&= \mathcal{T}_i^{jk} && (11)
\end{aligned}$$

hence, we have three free parameters α_i (in geometric terms there is a free choice of reference plane in space). We can select α_i such that the matrix b_i^k will have a vanishing column (this corresponds to selecting a reference plane coplanar with the center of projection of the third view). Therefore, the new matrices a_i^j and b_i^k have only 15 non-vanishing entries, and we have reduced the number of parameters from 21 to 18.

A.2.2 Admissibility Constraints

We may deduce from the Correlation Contractions discussed above the following three groups of constraints that the 27 coefficients must satisfy:

1. Rank($\delta^i \mathcal{T}_i^{jk}$)=2 for all choices of δ. The three standard correlation slices T_1, T_2, T_3 are of rank 2 each and this property is *closed* under all linear combinations.
2. Rank(null(T_1),null(T_2),null(T_3))=2. This follows from the fact the the null space of $\delta^i \mathcal{T}_i^{jk}$ is the epipolar line in the second image corresponding to the point δ in the first image — since all epipolar lines are concurrent, their rank is 2.
3. Rank(null($T_1^\top$),null($T_2^\top$),null($T_3^\top$))=2. These are epipolar lines in third image, thus their rank is 2 as well.

One can easily show that no subset of these constraints is sufficient to describe an admissible tensor of the form of eqn. 4. In practice, in the presence of errors in image measurements one often starts with the Linear solution (that might not satisfy the admissibility constraints) and improves it further by employing a numerical Gauss-Newton style iterative procedure until a solution that satisfies the admissibility constraints is obtained (for example, [5]).