The Rank 4 Constraint in Multiple (≥ 3) View Geometry

Amnon Shashua[1] *and Shai Avidan*[2]

[1] Technion — Israel Institute of Technology, CS department, Haifa 32000, Israel
[2] Hebrew University, Institute of Computer Science, Jerusalem 91904, Israel

Abstract. It has been established that certain trilinear froms of three perspective views give rise to a tensor of 27 intrinsic coefficients [8]. Further investigations have shown the existence of quadlinear forms across four views with the negative result that further views would not add any new constraints [3, 12, 5]. We show in this paper first general results on any number of views. Rather than seeking new constraints (which we know now is not possible) we seek connections across trilinear tensors of triplets of views. Two main results are shown: (i) trilinear tensors across $m > 3$ views are embedded in a low dimensional linear subspace, (ii) given two views, all the induced homography matrices are embedded in a four-dimensional linear subspace. The two results, separately and combined, offer new possibilities of handling the consistency across multiple views in a linear manner (via factorization), some of which are further detailed in this paper.

Keywords: Trilinearity, 3D recovery from 2D views, Matching Constraints, Projective Structure, Algebraic and Projective Geometry.

1 Introduction

The algebraic and geometric relations across multiple perspective views is a recent and growing interest which is relevant to a number of topics including (i) issues of 3D reconstruction from 2D data, (ii) representations of visual scenes from video data, (iii) image synthesis and animation, and (iv) visual recognition and indexing.

Typical to these topics is the question about the limitations and possibilities of going from two-dimensional (2D) measurements of point matches (correspondences) across two or more views to properties of the three-dimensional (3D) object or scene. Since the relationship between the 3D world and the 2D image space combines together 3D shape parameters, camera viewing parameters and 2D image measurements, the question of limitations and possibilities, in its widest scope, is about (i) 2D constraints across multiple views (matching constraints), (ii) characterizations of the space of all images of a particular object (indexing functions). In other words, one seeks to best represent, in terms of efficiency, compactness, flexibility and scope of use, two kinds of manifolds: (i) the manifold of image and viewing parameters (invariance to shape), and (ii) the manifold of image and object parameters (invariance to viewing parameters).

It has been established that certain trilinear forms of three perspective views give rise to a tensor of 27 intrinsic coefficients [8]. Further work on the properties of the "trilinear tensor" with relevancy to 3D reconstruction was described in [10, 4]. Other investigations have established the existence of quadlinear forms (with total of 81 coefficients) across four views with the negative result that further views would not add any new constraints [3, 12, 5]. Also, adopting the representation put forward by [3], dual trilinear tensors were established by [2, 13].

In this paper we extend the investigation to any number $m > 3$ views. There are two motivations to this work. First, is the practical aspect — if any additional view over the fourth view is redundant, what is the best and most efficient way of capturing that redundancy (in a linear manner)? Second, the existence of the quad-linearities is somewhat unsettling because the number of coefficients has risen from 27 to 81, whereas a view adds only 12 parameters. In other words, the quad-linearities may be too redundant a representation of the constraints over four views.

Our line of approach is to investigate the space of *all* trilinear tensors and to look for rank deficiencies in that space. Any finding of that sort is extremely useful because it readily allows a statistical way of putting together many views, simply by means of factorization. Moreover, a finding of that nature promises progress on the task of novel-view synthesis from model images ("image-based rendering") because a rank deficiency implies that trilinear tensors are related together by linear combinations — which is a necessary property for synthesizing tensors from a number of model tensors.

Two main results are shown: (i) trilinear tensors across $m > 3$ views are embedded in a low dimensional linear subspace, (ii) given two views, all the induced homography matrices are embedded in a four-dimensional linear subspace. The two results, separately and combined, offer new possibilities of handling multiple views in a linear manner (via factorization), some of which are further detailed in this paper.

2 Preliminaries About the Trilinear Tensor

Let P be a point in 3D projective space projecting onto p, p', p'' three views ψ, ψ', ψ'' represented by the two dimensional projective space. The relationship between the 3D and the 2D spaces is represented by the 3×4 matrices, $[I, 0]$, $[A, v']$ and $[B, v'']$, i.e.,

$$p = [I, 0]P$$
$$p' \cong [A, v']P$$
$$p'' \cong [B, v'']P$$

We may adopt the convention that $p = (x, y, 1)^\top$, $p' = (x', y', 1)^\top$ and $p'' = (x'', y'', 1)^\top$, and therefore $P = [x, y, 1, \rho]$. The coordinates $(x, y), (x'y'), (x'', y'')$

are matching points (with respect to some arbitrary image origin — say the geometric center of each image plane). The 3×3 matrices A and B are 2D collineations (homography matrices) from ψ to ψ' and ψ'', respectively, induced by the plane $\rho = 0$. The vectors v' and v'' are the epipolar points (the projection of the first camera center onto views ψ' and ψ'', respectively). The trilinear tensor is an array of 27 entries:

$$\alpha_i^{jk} = v'^k b_i^j - v''^j a_i^k. \qquad i, j, k = 1, 2, 3 \tag{1}$$

where superscripts denote contravariant indices (representing points in the 2D plane, like v') and subscripts denote covariant indices (representing lines in the 2D plane, like the rows of A). Thus, a_i^k is the element of the k'th row and i'th column of A, and v'^k is the k'th element of v'. The tensor α_i^{jk} forms the set of coefficients of certain trilinear forms that vanish on any corresponding triplet p, p', p'' and which have the following form: let s_k^l be the matrix,

$$s = \begin{bmatrix} 1 & 0 & -x' \\ 0 & 1 & -y' \end{bmatrix}$$

and, similarly, let r_j^m be the matrix,

$$r = \begin{bmatrix} 1 & 0 & -x'' \\ 0 & 1 & -y'' \end{bmatrix}$$

Then, the tensorial equations are:

$$s_k^l r_j^m p^i \alpha_i^{jk} = 0, \tag{2}$$

with the standard summation convention that an index that appears as a subscript and superscript is summed over (known as a contraction). For further details on this derivation, see Appendix A. Hence, we have four trilinear equations (note that $l, m = 1, 2$). In more explicit form, these functions (referred to as "trilinearities") are:

$$\begin{aligned}
x''\boldsymbol{\alpha}_i^{13}p^i - x''x'\boldsymbol{\alpha}_i^{33}p^i + x'\boldsymbol{\alpha}_i^{31}p^i - \boldsymbol{\alpha}_i^{11}p^i &= 0, \\
y''\boldsymbol{\alpha}_i^{13}p^i - y''x'\boldsymbol{\alpha}_i^{33}p^i + x'\boldsymbol{\alpha}_i^{32}p^i - \boldsymbol{\alpha}_i^{12}p^i &= 0, \\
x''\boldsymbol{\alpha}_i^{23}p^i - x''y'\boldsymbol{\alpha}_i^{33}p^i + y'\boldsymbol{\alpha}_i^{31}p^i - \boldsymbol{\alpha}_i^{21}p^i &= 0, \\
y''\boldsymbol{\alpha}_i^{23}p^i - y''y'\boldsymbol{\alpha}_i^{33}p^i + y'\boldsymbol{\alpha}_i^{32}p^i - \boldsymbol{\alpha}_i^{22}p^i &= 0.
\end{aligned}$$

Since every corresponding triplet p, p', p'' contributes four linearly independent equations, then seven corresponding points across the three views uniquely determine (up to scale) the tensor α_i^{jk}. More details and applications can be found in [8, 9]. Also worth noting is that these trilinear equations are an extension of the three equations derived by [11] under the context of unifying line and point geometry.

Furthermore, for any arbitrary (covariant) vector $\rho_j = (\rho_1, \rho_2, \rho_3)$, the matrix $\rho_j \alpha_i^{jk}$ (recall the summation convention, i.e., $\rho_j \alpha_i^{jk} = \rho_1 \alpha_i^{1k} + \rho_2 \alpha_i^{2k} +$

$\rho_3\alpha_i^{3k}$, which is a matrix) is not just any matrix, it is a 2D *homography* (a 2D collineation) from image 1 to image 2 via some plane, whose parameters are determined by ρ_j (the vector ρ_j is in direction of the normal to the plane in a coordinate system whose origin is the first camera center and its axes are aligned with the axes of the third camera coordinate system). Therefore, if we take ρ_j to be $(1,0,0),(0,1,0)$ and $(0,0,1)$, we obtain three homography matrices, which we will denote by E_1, E_2, E_3, respectively. In other words, the elements of the tensor α_i^{jk} are *rearranged* to comprise three matrices E_1, E_2, E_3 ($E_1 = \alpha_i^{1k}$, for example). For example, the "fundamental" matrix F between ψ and ψ' can be linearly determined from the tensor by: ${E_j}^\top F + F^\top E_j = 0$, which yields 18 linear equations of rank 8 for F. More details can be found in [10].

3 Tensors and Rank 12

Consider the following arrangement: we are given views $\psi_1, \psi_2, ...\psi_{m+2}$, $m \geq 1$. For each (ordered) triplet of views there exists a unique trilinear tensor. Consider m triplets of views each containing ψ_1, ψ_2, i.e., the triplets $<\psi_1, \psi_2, \psi_i>$, $i = 3, ..., m+2$. Consider each of the tensors as a vector of 27 components and concatenate all these vectors as columns of a $27 \times m$ matrix. The question is what is the rank of this matrix when $m \geq 27$? Clearly, if the rank is smaller than 27 we obtain a line of attack on the task of putting together many views. The motivation for considering this arrangement is that a view adds only 12 parameters (up to scale). It may be the case that the redundancy of representing an additional view with 27 numbers (a column vector in the $27 \times m$ matrix), instead of 12, comes to bear only at a non-linear level — in which case it will not affect the rank of the system above. Therefore, a rank deficiency implies an important property of a collection of tensors.

We arrange each tensor as a 27 column vector as follows:

$$\begin{pmatrix} E_1 = \alpha_i^{1k} \\ E_2 = \alpha_i^{2k} \\ E_3 = \alpha_i^{3k} \end{pmatrix}$$

where E_j (a 3×3 matrix) is arranged column-wise as a 9 vector. To simplify notation of indices, let $[B_i, v^{(i)}]$ denote the camera transformation matrices for view $\psi_3, ...,$ and $[A, v']$ the camera transformation matrix for view ψ_2. In the next formula, $A, B_i, v', v^{(i)}$ all appear as column vectors: A is a arranged as a 9-vector column wise, i.e., $(a_1^1, a_2^1, a_3^1, a_1^2, ..., a_3^3)$, and likewise $B_i^\top$. It is simply a matter of observation to verify that the following holds:

$$\begin{bmatrix} \quad \end{bmatrix}_{27\times m} = \begin{bmatrix} v' & 0 & \cdot & \cdot & 0 \\ 0 & v' & \cdot & \cdot & 0 \\ \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & v' \end{bmatrix}_{27\times 9} \begin{bmatrix} B_1^\top & \cdots & B_m^\top \end{bmatrix}_{9\times m} - \begin{bmatrix} A & 0 & 0 \\ & & \\ 0 & A & 0 \\ & & \\ 0 & 0 & A \end{bmatrix}_{27\times 3} \begin{bmatrix} v^{(1)} & \cdots & v^{(m)} \end{bmatrix}_{3\times m}$$

$$= \begin{bmatrix} v' & 0 & \cdot & \cdot & 0 & A & 0 & 0 \\ 0 & v' & \cdot & \cdot & 0 & & & \\ \cdot & 0 & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & A & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & & & \\ 0 & 0 & \cdot & \cdot & v' & 0 & 0 & A \end{bmatrix}_{27\times 12} \begin{bmatrix} B_1^\top & \cdots & B_m^\top \\ & & \\ v^{(1)} & \cdots & v^{(m)} \end{bmatrix}_{12\times m}$$

Thus, we have proved the following theorem:

Theorem 1 (Rank 12) *All trilinear tensors live in a manifold of $\mathcal{P}^{26}$. The space of all trilinear tensors with two of the views fixed, is a 12'th dimensional linear sub-space of $\mathcal{R}^{27}$.*

From the factorization principle above we see that each additional view adds, linearly, only 12 parameters — as expected. Moreover, these 12 parameters constitute the camera transformation matrix associated with the new view. Next we show that the linear subspace of all tensors with two views fixed is *closed*, i.e., any linear combination of such tensors produces an admissible tensor describing the configuration of the two fixed cameras and some third camera position.

Theorem 2 *The linear subspace containing tensors of views $<\psi_1, \psi_2, \psi>$, for all views ψ, is closed.*

The proof is detailed in Appendix B. In particular, consider the two tensors $<\psi_1, \psi_2, \psi_i>$ and $<\psi_1, \psi_2, \psi_j>$ as two points on the (non-linear) manifold of *all* tensors; the inifinite line passing through the two points *lives* inside the manifold — which implies that the manifold is decomposed into ruled surfaces. This property is the basis for novel-view synthesis which we will touch upon later in the paper.

4 Collineations and Rank 4

Consider a similar exercise done with homography matrices between two fixed views. Given some plane in space projecting onto views ψ and ψ', the corresponding image points are mapped to each other by a collineation (homography matrix), $Ap \cong p'$ for all matching pairs p, p'. Since the homography matrix A depends on the orientation and location of the planar object, we obtain a *family* of homography matrices when we consider all possible planes. It is also known, that given a homography matrix A of *some* plane, then all other homography matrices can be described by,

$$\lambda A + v' n^\top.$$

Consider homography matrices $A_1, A_2, ..., A_k$ each as a column vector in a $9 \times k$ matrix. We ask again, what is the rank of the system? It would be convenient if it were 4, because each additional homography matrix represents a plane, and a plane is determined by 4 parameters. Let $A_i = \lambda_i A + v' n_i{}^\top$. The following can be verified by inspection:

$$\begin{aligned}
\left[\;\right]_{9\times k} &= \begin{bmatrix} \\ \lambda_1 A \cdots \lambda_k A \\ \\ \end{bmatrix}_{9\times k} + \begin{bmatrix} v' & 0 & 0 \\ 0 & v' & 0 \\ 0 & 0 & v' \end{bmatrix}_{9\times 3} \begin{bmatrix} n_1 \cdots n_k \end{bmatrix}_{3\times k} \\
&= \begin{bmatrix} & v' & 0 & 0 \\ A & 0 & v' & 0 \\ & 0 & 0 & v' \end{bmatrix}_{9\times 4} \begin{bmatrix} \lambda_1 \cdots \lambda_k \\ n_1 \cdots n_k \end{bmatrix}_{4\times k}
\end{aligned}$$

We have thus proven the following result:

Theorem 3 (Collineations, Rank 4) *The space of all homography matrices between two fixed views is embedded in a 4 dimensional linear subspace of* $\mathcal{P}^8$.

5 Tensors and Rank 4

We recall from Section 2 that the tensor α_i^{jk} can be contracted into three homography matrices, associated with three distinct planes, between ψ_1 and ψ_2. Hence, consider the same situation as before where we have the tensors of the triplets $< \psi_1, \psi_2, \psi_i >$, $i = 3, ..., m+2$. But now, instead of arranging each tensor as a 27 column vector, we arrange it in a 9×3 block, where each column is the homography α_i^{jk}, $j = 1, 2, 3$. We obtain a $9 \times 3m$ matrix,

$$M = \begin{bmatrix} \\ \alpha_i^{1k} \; \alpha_i^{2k} \; \alpha_i^{3k} \; \cdots \\ \\ \end{bmatrix}_{9\times 3m}$$

From Theorem 3 we know that its rank must be 4. Therefore, we have the following result:

Theorem 4 (Tensors and Rank 4) *The space of all trilinear tensors with two of the views fixed can be decomposed into three separate linear subspaces, each of dimension 4, of* $\mathcal{R}^{27}$.

Four columns of M span all tensors $< \psi_1, \psi_2, \psi >$, for any view ψ. Instead of being spanned as a 27 vector it is spanned three times each as a 9 vector. Thus, the new tensor is determined by 12 coefficients (of the linear combinations of the 4 columns). As a consequence, each additional tensor would require only 6 matching points, instead of 7:

Corollary 1 (Tensor-and-a-third) *A tensor of views* $< \psi_1, \psi_2, \psi_3 >$ *and "third" of the tensor* $< \psi_1, \psi_2, \psi_4 >$*, linearly span, with 12 coefficients, all tensors* $< \psi_1, \psi_2, \psi >$ *(over all views* ψ*). Each such tensor can be recovered using 6 matching points with* ψ_1 *and* ψ_2.

We can use instead a single tensor to linearly span all tensors $< \psi_1, \psi_2, \psi >$ (for all views ψ) by recovering the fundamental matrix F from the tensor (see Section 2). The homography matrix $[v']_x F$ (see [6]), where $[v']_x$ is the skew-symmetric matrix associated with vector products, can replace the "third" tensor of above (note that $[v']_x F$ is of rank 2, therefore is not linearly spanned by the three homography matrices provided from the tensor — unless the three camera centers are collinear). We have therefore the following result:

Corollary 2 (Tensor + F) *The tensor of views* $< \psi_1, \psi_2, \psi_3 >$ *and the epipolar constraint (matrix* F *and epipole* v'*) together linearly span, with 12 coefficients, all other tensors* $< \psi_1, \psi_2, \psi >$ *(running over all views* ψ*). Each additional view* ψ *contributes linearly 12 parameters and its tensor with* ψ_1, ψ_2 *can be determined linearly using 6 matching points.*

6 Applications and Experimental Results

The results presented in this paper have three areas of application (that have been identified so far). First, is the obvious application of enforcing numerical consistency across two or more tensors. This can done as follows.

Assume we are given views $\psi_1, \psi_2, ...\psi_{m+2}$ and consider the tensors of the triplets $< \psi_1, \psi_2, \psi_i >$, $i = 3, ..., m+2$. Arrange the tensors into a $9 \times 3m$ matrix M as described in the previous section. Perform an SVD and keep only the 4 largest singular values. We have thus obtained a new matrix $\hat{M}$ that enforces the rank 4 constraint, which in turn, enforces the consistency across all the tensors. Separate the tensors of interest from $\hat{M}$ (each tensor still occupies a 9×3 block of $\hat{M}$).

Another variant on this application is to recover the four principle components of M, as follows. Given M as above, perform a principle component analysis and obtain the four principle components $A_1, ..., A_4$ (each is a 9-vector). These vectors encode the geometry between ψ_1, ψ_2 alone (represent four homography

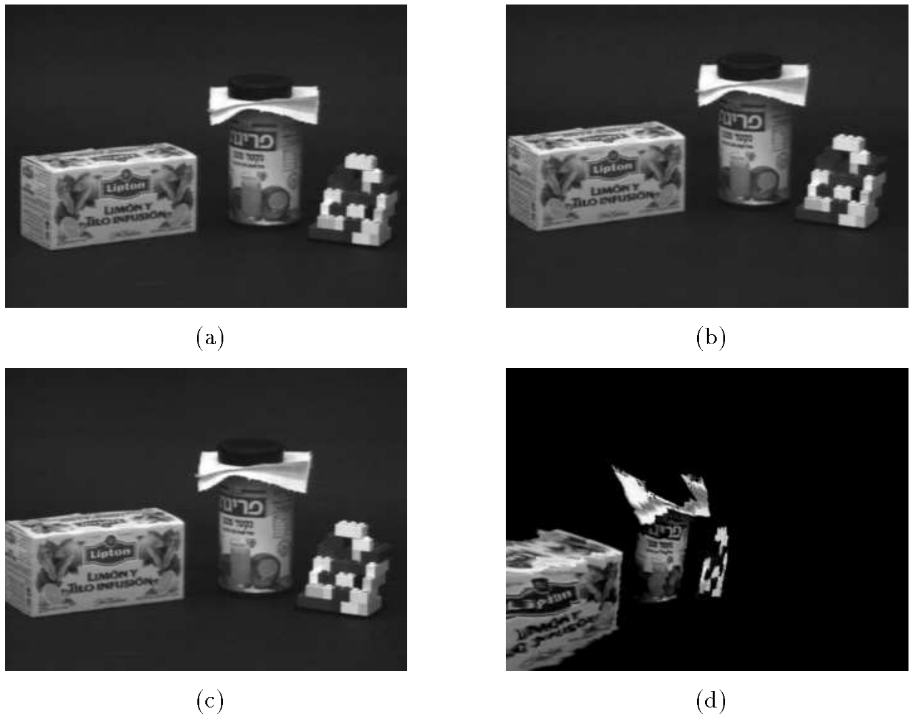

Fig. 1. An example of image synthesis using optic-flow and a tensor. (a) (b) and (c) are the original three images. (d) is a synthetic image created by the tensor of the three model images, the user-specified virtual camera motion, and the (dense) correspondence between images (a) and (b). Note that the virtual view is significantly outside the viewing cone of the original model views. For more details see [1].

matrices). Hence, we can use them to recover F from the (over-determined) linear system of 24 equations $A_j{}^{\top} F + F^{\top} A_j = 0$, $j = 1, ..., 4$.

The second area of application is in novel-view synthesis (image-based rendering). Theorem 4 can be rewritten as follows:

$$\begin{pmatrix} \hat{E}_1 \\ \hat{E}_2 \\ \hat{E}_3 \end{pmatrix} = \begin{bmatrix} l_{11} & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & l_{34} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ A \end{bmatrix} \tag{3}$$

where E_1, E_2, E_3 are the three homography matrices comprising the tensor of views $< \psi_1, \psi_2, \psi_3 >$, A is some arbitrary homography matrix from ψ_1 to ψ_2 not spanned by E_1, E_2, E_3; the matrix L is a 3×4 matrix comprising the 12 coefficients necessary to span a new tensor of the views $< \psi_1, \psi_2, \psi >$, for some view ψ, arranged as $\hat{E}_1, \hat{E}_2, \hat{E}_3$ as a 3×9 matrix, i.e., each matrix $\hat{E}_j$ is a 9-vector arranged column-wise. Furthermore, the matrix L is a camera transformation matrix from 3D space to the new camera position that produces the view ψ. This result implies that in principle we could be in a position to *synthesize* new tensors on demand (by specifying L) and from those new tensors to reproject views ψ_1, ψ_2 and create views ψ. This idea was taken further in [1] where the possibility of creating synthetic movies of a 3D scene was demonstrated. Fig. 1 shows three model views of a scene, and a new synthesized image of the scene created by synthesizing a new tensor from the tensor of the three model views and then reprojecting the two model views on the top row onto the new viewing position using the new synthesized tensor. Note that the synthesized virtual view is significantly outside the viewing cone of the three model views, and that no 3D model of the scene was created in the process. For more details see [1].

The third application using the "rank 4" results is in the area of image stabilization. In [7] it is shown that Theorem 3 can be used as a building block for recovering the small-angle approximation of the rotational component of camera motion *directly* and in closed-form from a tensor of three views. The significance of this result is that the translational component of camera motion need not be recovered in the process, and that the process is fairly robust in practice.

7 Summary

This paper has presented a new approach for investigating the inter-relationship among a collection of four or more views. The approach is based on searching for rank deficiencies in the space of all trilinear tensors.

We have shown that families of trilinear tensors are embedded in a low dimensional linear subspace of tensor space (the manifold where all tensors live). First, this result enables a factorization approach to enforce consistency among many views (via consistency among the tensors). We showed, for instance, that one can use the factorization principle to obtain the fundamental matrix of two views from any number of views. Secondly, the theorems and their corollaries provide a tight bound on the contribution of additional views over three views.

We view these results as forming "building blocks" for future applications, not necessarily as forming an application on their own. We have pointed out two areas of application where these results have already proven fruitful — the first is in the area of "image-based rendering" where one is interested in synthesizing novel views of a 3D scene without necesserily creating a 3D model of the scene, and the second application is the area of video sequence stabilization.

A Deriving the Trilinear Tensor

The trilinear equations were first derived in [8] together with the equation of the tensor. The derivation presented here is more compact and more details can be found in [9].

The camera transformation between images ψ and ψ' is represented by $p' \cong [A, v']P$ where $P = (x, y, 1, \rho)^\top$. Let s_k^l be the matrix,

$$s = \begin{bmatrix} 1 & 0 & -x' \\ 0 & 1 & -y' \end{bmatrix}$$

It can be verified by inspection that $p' \cong [A, v']P$ can be represented by the following two equations:

$$\rho s_k^l v'^k + p^i s_k^l a_i^k = 0, \tag{4}$$

with the standard summation convention that an index that appears as a subscript and superscript is summed over (known as a contraction). Note that we have two equations because $l = 1, 2$ is a free index.

Similarly, the camera transformation between views ψ and ψ'' is $p'' \cong [B, v'']P$. Likewise, let r_j^m be the matrix,

$$r = \begin{bmatrix} 1 & 0 & -x'' \\ 0 & 1 & -y'' \end{bmatrix}$$

And likewise,

$$\rho r_j^m v''^j + p^i r_j^m b_i^j = 0, \tag{5}$$

Note that k and j are dummy indices (are summed over) in equations 4 and 5, respectively. We used different dummy indices because now we are about to eliminate ρ and combine the two equations together. Likewise, l, m are free indices, therefore in the combination they must be separate indices. We eliminate ρ and after some rearrangement and grouping we obtain:

$$s_k^l r_j^m p^i (v'^k b_i^j - v''^j a_i^k) = 0,$$

and the term in parenthesis is the trilinear tensor.

B The linear subspace of tensors is closed under linear combinations

Theorem 2: *The linear subspace containing tensors of views $< \psi_1, \psi_2, \psi >$, for all views ψ, is closed.*

Proof. We need to show that for any views ψ_3, ψ_4, the linear combination of tensors $< \psi_1, \psi_2, \psi_3 >$ and $< \psi_1, \psi_2, \psi_4 >$ produces a tensor $< \psi_1, \psi_2, \psi >$ for some view ψ. Let $[B_3; v^{(3)}]$ and $[B_4; v^{(4)}]$ be the camera transformation matrices associated with views ψ_3 and ψ_4, respectively. From Theorem 1 it is clear that the linear combination will produce a tensor $< \psi_1, \psi_2, \psi >$ where the camera transformation matrix associated with view ψ is $[aB_3 + bB_4; av^{(3)} + bv^{(4)}]$ where a, b are the coefficients of the linear combination. The subtle point in this argument is that although the homography matrix $aB_3 + bB_4$ does not correspond to the same plane (the plane $\rho = 0$, see Section 2) associated with B_3 and B_4, it can be transformed into such a homography matrix by correspondingly transforming the projective representation of the 3D scene. In other words, the projective representation of the scene can undergo a projective transformation (which effectively translates the scene along the optical axes of the first view) which is interchangeable with the camera motion from view to view. The interchangeability point is effectively contained in the arguments of [3] about the geometric content of each trilinearity. $\square$

References

1. S. Avidan and A. Shashua. Novel view synthesis in tensor space. CS report, CIS-9602, Technion, January 1996. Also http://www.cs.huji.ac.il/ avidan/render/render.html.
2. S. Carlsson. Duality of reconstruction and positioning from projective views. In *Proceedings of the workshop on Scene Representations*, Cambridge, MA., June 1995.
3. 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.
4. 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.
5. 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.
6. 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.
7. B. Rousso, S. Avidan, A. Shashua, and S. Peleg. Robust recovery of camera rotation from three frames. In *Proceedings of the ARPA Image Understanding Workshop*, Palm Springs, CA, February 1996.
8. A. Shashua. Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):779–789, 1995.
9. 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.
10. A. Shashua and M. Werman. On the trilinear tensor of three perspective views and its underlying geometry. In *Proceedings of the International Conference on Computer Vision*, June 1995.
11. M.E. Spetsakis and J. Aloimonos. A unified theory of structure from motion. In *Proceedings of the ARPA Image Understanding Workshop*, 1990.
12. B. Triggs. Matching constraints and the joint image. In *Proceedings of the International Conference on Computer Vision*, pages 338–343, Cambridge, MA, June 1995.
13. D. Weinshall, M. Werman, and A. Shashua. Shape tensors for efficient and learnable indexing. In *Proceedings of the workshop on Scene Representations*, Cambridge, MA., June 1995.