Homography Tensors: On Algebraic Entities That Represent Three Views of Static or Moving Planar Points

Amnon Shashua and Lior Wolf

School of Computer Science and Engineering,
The Hebrew University,
Jerusalem 91904, Israel
e-mail: {shashua,lwolf}@cs.huji.ac.il

Abstract

We introduce a $3 \times 3 \times 3$ tensor H^{ijk} and its dual H_{ijk} which represent the 2D projective mapping of points across three projections (views). The tensor H^{ijk} is a generalization of the well known 2D collineation matrix (homography matrix) and it concatenates two homography matrices to represent the joint mapping across three views. The dual tensor H_{ijk} concatenates two dual homography matrices (mappings of line space) and is responsible for representing the mapping associated with **moving** *points along straight-line paths, i.e., H_{ijk} can be recovered from line-of-sight measurements only.*

1 Introduction

In this paper we revisit the fundamental element of projective geometry, the *collineation* (also referred to as homography matrix) between two sets of points on the projective plane undergoing a projective mapping. The role of homography matrices responsible for mapping point sets between two views of a planar object is basic in multiple-view geometry in computer vision. The object stands on its own as a point-transfer vehicle for planar scenes (aerial photographs, for example) and in applications of mosaicing, camera stabilization and tracking [6]; a homography matrix is a standard building block in handling 3D scenes from multiple 2D projections: the "plane+parallax" framework [7,4,5,2] uses a homography matrix for setting up a parallax residual field relative to a planar reference surface, and the trifocal tensor of three views is represented by a "homography-epipole" structure whose slices are homography matrices as well [3,8].

In our work we first consider a 3-view version of a projective mapping represented by a $3 \times 3 \times 3$ contravariant tensor H^{ijk}, referred to as a homography tensor (abbreviated as "Htensor"). The entries of the Htensor are bilinear products of the orginal pair of homography matrices and its 27 coefficients can be recovered linearly (up to scale) from 4 matching points (lines) across the three views. The Htensor can perform directly the image-to-image mapping or alternatively the original piarwise collineations can be linearly recovered from the slices of the tensor. The 27 entries of the tensor satisfy a number of non-linear constraints which are unique to the coupling of three views together — in pairwise projective mappings all constraints are linear.

We next consider the dual Htensor, a covariant form H_{ijk} whose constituents are dual homography matrices (mapping of line space). The dual Htensor has an interesting "twist" in the sense that it applies to projections of **moving** points following straight-line paths on a planar surface. In other words, if $p^i p'^j p''^k H_{ijk} = 0$ then the three optical rays defined by the image points p, p', p'' in views 1,2,3 respectively meet at a *line*. Consequently, the dual tensor opens new applications in which static and moving points live together as equal partners — the process of mapping and estimation of the dual tensor from image measurements need not know in advance what is moving and what is static.

1.1 Background and Notations

We will be working with the projective plane, i.e., the space $\mathcal{P}^2$. Points and lines are represented by triplets of numbers (not all zero) that define a coordinate vector. Consider a collection of planar points $P_1, ..., P_n$ in space living on a plane π viewed from two views. The projections of P_i are p_i, p'_i in views 1,2 respectively. There exists a unique collineation (homography) 3×3 matrix A_π that satisfies the relation $A_\pi p_i \cong p'_i$, $i = 1, ..., n$, and where A_π is uniquely determined by 4 matching pairs from the set of n matching pairs. Moreover, $A_\pi^{-T} s \cong s'$ will map between matching lines s, s' arising from 3D lines living in the plane π. Likewise, $A_\pi^\top s' \cong s$ will map between matching lines from view 2 back to view 1.

It will be most convenient to use tensor notations from now on because the material we will be using in this paper involves coupling together pairs of collineations into a "joint" object. The distinction of when coordinate vectors stand for points or lines matters when using tensor notations. A point is an object whose coordinates are specified with superscripts, i.e., $p^i = (p^1, p^2, p^3)$. These are called contravariant vectors. A line in $\mathcal{P}^2$ is called a covariant vector and is represented by subscripts, i.e., $s_j = (s_1, s_2, s_3)$. Indices repeated in covariant and contravariant forms are summed over, i.e., $p^i s_i = p^1 s_1 + p^2 s_2 + p^3 s_3$. This is known as a contraction. For example, if p is a point incident to a line s in $\mathcal{P}^2$, then $p^i s_i = 0$.

Vectors are also called 1-valence tensors. 2-valence tensors (matrices) have two indices and the transformation they represent depends on the covariant-contravariant positioning of the indices. For example, a_i^j is a mapping from points to points (a collineation, for example), and hyperplanes (lines in $\mathcal{P}^2$) to hyperplanes, because $a_i^j p^i = q^j$ and $a_i^j s_j = r_i$ (in matrix form: $Ap = q$ and $A^\top s = r$); a_{ij} maps points to hyperplanes; and a^{ij} maps hyperplanes to points. When viewed as a matrix the row and column positions are determined accordingly: in a_i^j and a_{ji} the index i runs over the columns and j runs over the rows, thus $b_j^k a_i^j = c_i^k$ is $BA = C$ in matrix form. An outer-product of two 1-valence tensors (vectors), $a_i b^j$, is a 2-valence tensor c_i^j whose i, j entries are $a_i b^j$ — note that in matrix form $C = ba^\top$. A 3-valence tensor has three indices, say H_i^{jk}. The positioning of the indices reveals the geometric nature of the mapping: for example, $p^i s_j H_i^{jk}$ must be a point because the i,j indices drop out in the contraction process and we are left with a contravariant vector (the index k is a superscript). Thus, H_i^{jk} maps a point in the first coordinate frame and a line in the second coordinate frame into a point in the third coordinate frame. A single contraction, say $p^i H_i^{jk}$, of a 3-valence tensor leaves us with a matrix. Note that when p is $(1,0,0)$ or $(0,1,0)$, or $(0,0,1)$ the result is a "slice" of the tensor.

We will make extensive use of the "cross-product tensor" ϵ defined next. The cross product (vector product) operation $c = a \times$b is defined for vectors in $\mathcal{P}^2$. The product operation can also be represented as the product $c = [a]_\times b$ where $[a]_\times$ is called the "skew-symmetric matrix of a". In tensor form we have $\epsilon_{ijk} a^i b^j = c_k$ representing the cross product of two points (contravariant vectors) resulting in the line (covariant vector) c_k. Similarly, $\epsilon^{ijk} a_i b_j = c^k$ represents the point intersection of the to lines a_i and b_j. The tensor ϵ_{ijk} is the anti-symmetric tensor defined such that $\epsilon_{ijk} a^i b^j c^k$ is the determinant of the 3×3 matrix whose columns are the vectors a, b, c. As such, ϵ_{ijk} contains $0, +1, -1$ where the vanishing entries correspond to arrangement of indecis with repetitions (21 such entries), whereas the odd permutations of ijk correspond to -1 entries and the even permutations to $+1$ entries.

In the sequel we will reserve the indices i, j, k to represent the coordinate vectors of images 1,2,3 respectively. We will denote points in images 1,2,3 as p, p', p'' respectively, thus in a tensor equation these points will appear with their corresponding indecis — for instance, $p^i p'^j p''^k H_{ijk} = 0$.

2 Homography Tensor H^{ijk}

Consider some plane π whose features (points or lines) are projected onto three views and let A be the collineation from view 1 to view 2, and B the collineation from view 1 to 3 (we omit the reference to π in our notation). Let P be some point on the plane π and its projections are p, p', p'' in views 1,2,3 respectively. Let q, s, r be *some* line through p, p', p'' respectively. We have $q^\top (A^\top s \times B^\top r) = 0$ because $A^\top s$ is the projection of the line L on π onto view 1, where L projects to s in view 2, and similarly $B^\top r$ is a line in view 1 matching the line r in view 3. These two lines must intersect at p (see Fig. 1). In tensor form we have:

$$q_i s_j r_k (\epsilon^{inu} a_n^j b_u^k) = 0, \tag{1}$$

and we denote the object in parenthesis

$$H^{ijk} = \epsilon^{inu} a_n^j b_u^k \tag{2}$$

as the *Homography Tensor* (in short, Htensor). In the remainder of this section we will investigate the properties and uses of the Htensor along the following lines:

– Number of matching points (lines) required for a unique solution for the Htensor (Proposition 1,2).
– Slicing properties of the Htensor and the means for recovering the original homography matrices A, B from the Htensor (Theorem 1). The nature of these slices provide the source for 11 non-linear constraints among the tensor's coefficients.
– Image-to-image mapping using the Htensor.

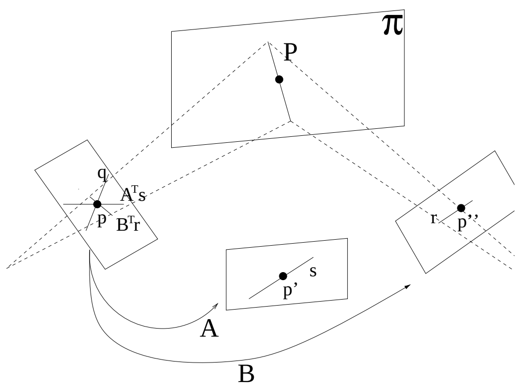

Fig. 1. The lines s, r are mapped by the dual collineations $A^\top, B^\top$ onto view 1, satisfying the relationship $q^\top(A^\top s \times B^\top r) = 0$.

The first result is that 4 matching triplets provide 26 linearly independent constraints on the 27 coefficients of H^{ijk} — hence, a unique solution (up to scale) is provided from image measurements of 4 points.

Proposition 1. *The n'th matching triplet provides* $8-(n-1)$ *linearly independent constraints to the constraints provided from the previous* $n-1$ *matching triplets. Hence, 4 matching triplets provide* $8+7+6+5 = 26$ *linearly independent constraints.*

Proof: The first matching triplet provides 8 linearly independent constraints because a point is spanned by two lines. Take for example the vertical $q_i^1 = (-1, 0, x)$ and the horizontal $q_i^2 = (0, -1, y)$ lines passing through the point p, and similarly the horizontal and vertical lines s_j^1, s_j^2 through the point p' and the lines r_k^1, r_k^2 through p'' and we have the eight constraints:

$$q_i^\mu s_j^\rho r_k^\nu H^{ijk} = 0, \qquad \mu, \rho, \nu = 1, 2.$$

Consider the second matching triplet p_2, p_2', p_2'' and one of the constraints defined by selecting the line q to pass through p_2 and p, the line s to pass through p_2' and p' and the line r to pass through p_2'' and p''. Clearly, these lines are spanned by the lines through p, p', p'', thus the added constraint is linearly spanned by the eight constraints provided by p, p', p''. Hence, the second matching triplet contributes only 7 additional linearly independent constraints to the 8 constraints from the first matching triplet (see Fig. 2). Continue by induction, the n'th point has lines to the previous $n-1$ points, thus n-1 constraints from the 8 possible constraints are already covered by the previous points. ∎

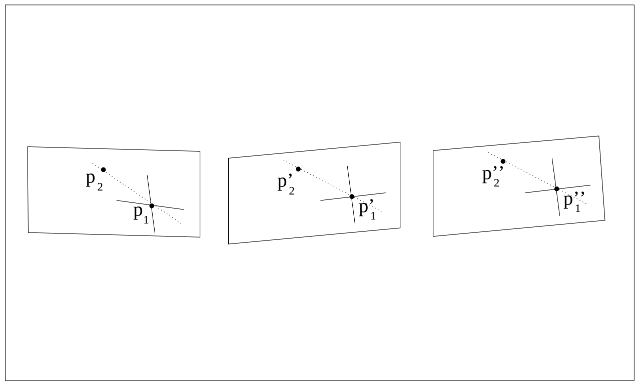

Fig. 2. A triplet of matching points provides 8 constraints. A second triplet provides only 7 additional constraints because the constraint defined by the lines connecting the the two sets of points is already covered by the 8 constraints of the first triplet.

Note that the lines q, s, r in eqn. 1 are *not* matching lines although they pass through matching points. Our next issue is to show that if q, s, r *are* matching lines, then they provide 7 linear constraints, thus 4 matching line triplets provide 28 constraints and a unique solution to H^{ijk}.

Proposition 2. *If q, r, s are matching lines in views 1,2,3 then $q_i s_j H^{ijk}$,$q_i r_k H^{ijk}$ and $s_j r_k H^{ijk}$ are null vectors providing a total of 7 linearly independent constraints on the Htensor. Thus 4 matching lines provide a unique linear solution for H^{ijk}.*

Proof: If q, s, r are matching lines then the rank of the matrix whose columns are $[q, A^\top s, B^\top r]$ is 1. Thus, $q \times A^\top s = 0$,$q \times B^\top r = 0$ and $s^\top A \times B^\top r = 0$. In tensor form, these translate to the following:

$$q_i s_j e_k H^{ijk} = 0 \quad \forall e_k,$$
$$q_i e_j r_k H^{ijk} = 0 \quad \forall e_j,$$
$$e_i s_j r_k H^{ijk} = 0 \quad \forall e_i.$$

Note that $q_i s_j r_k H^{ijk}$ appears three times (once in every row above), thus among the nine constraints arising from the fact that $q_i s_j H^{ijk}$,$q_i r_k H^{ijk}$ and $s_j r_k H^{ijk}$ are null vectors two of the constraints are already accounted for making the total of 7 linearly independent constraints. ∎

So far we have shown that 4 matching points or 4 matching lines across the three views provide a unique solution to the Htensor — just like with collineations: 4 is the number of points or lines that is required for a unique solution. We turn our attention next to single and double contractions of the Htensor — what can be extracted from them and what is their geometric significance.

The double contractions perform mapping operations. For example, $q_i s_j H^{ijk}$ must be a point (contravariant index k is left uncontracted) whose scalar product with the pencil of lines through p'' vanishes — hence, the point must be p''. We have therefore:

$$
\begin{aligned}
q_i s_j H^{ijk} &\cong p''^k \\
q_i r_k H^{ijk} &\cong p'^j \\
s_j r_k H^{ijk} &\cong p^i
\end{aligned}
$$

A single contraction is a correlation mapping of lines to points associated with a Linear Line Complex (LLC) — a set of lines that have a common line intersection called the kernel of the set — which we will derive as follows. Consider some arbitrary covariant vector δ_k and the resulting matrix $\delta_k H^{ijk}$. One can easily verify, by substitution in eqn. 2, that the resulting matrix is $E = A[\mu]_x$ where $\mu = B^\top \delta$. What does the matrix E stand for? Let L be the line on π at the intersection of π and the plane defined by the line δ in view 3 and its center of projection (see Fig. 3). The projections of L in view 1 and 2 are $\mu = B^\top \delta$ and $\eta = A^{-\top}\mu$. Clearly, $E\mu = 0$ and $E^\top \eta = 0$. Consider any other line S intersecting L in space (S is not necessarily on π) projecting onto s, s' in views 1,2 respectively. Then $s'^\top E s = 0$. Taken together, the matrix E maps lines in view 1 onto collinear points (on the line η) in view 2. The set of lines S in 3D whose projections s, s' satisfy $s'^\top E s = 0$ define an LLC whose kernel is the line L whose projections are the null spaces of E and $E^\top$. Moreover, $AE^\top$ is a skew symmetric matrix, thus $AE^\top + EA^\top = 0$ provides 6 linear constraints on the homography matrix A.

By selecting δ to range over the standard basis $(1,0,0), (0,1,0), (0,0,1)$ we obtain three slices of H^{ijk} which we will denote by E_1, E_2, E_3. These slices provide 18 linear constraints for the homography matrix A. Likewise, the three slices $H^{i1k}, H^{i2k}, H^{i3k}$ provide 18 constraints on the homography B and the three slices $H^{1jk}, H^{2jk}, H^{3jk}$ provide 18 constraints on the homography $C = BA^{-1}$ from view 2 to view 3. We summarize these findings in the following theorem:

Theorem 1. *Each of the contractions $\delta_i H^{ijk}, \delta_j H^{ijk}$ and $\delta_k H^{ijk}$ represents a correlation mapping between views $(2,3), (1,3)$ and $(1,2)$ respectively, associated with the LLC whose kernel is the line at the intersection of π and the plane defined by δ of views 1,2,3 respectively and the corresponding center of projection. By setting δ to be $(1,0,0), (0,1,0)$ or $(0,0,1)$ we obtain three different slicings of the tensor: denote the slices of $\delta_i H^{ijk}$ by the matrices G_1, G_2, G_3, the slices of $\delta_j H^{ijk}$ by the matrices W_1, W_2, W_3, and the slices of $\delta_k H^{ijk}$ by the matrices E_1, E_2, E_3. Then these slices provide sufficient (and over-determined) linear*

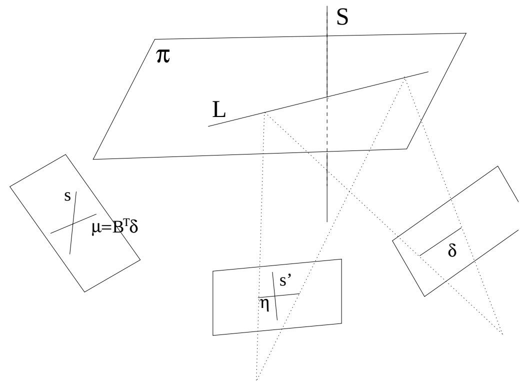

Fig. 3. The contraction $\delta_k H^{ijk}$ is a matrix $E = A[\mu]_x$ where $\mu = B^\top \delta$. The covariant vector δ represents a line in view 3 which together with the center of projection represents a plane whose intersection with π is a line L. The projections of L in views 1,2 are the null spaces of E and $E^\top$ respectively, i.e., $E\mu = 0$ and $E^\top \eta = 0$. Matching lines s, s' in views 1,2 satisfy $s'^\top E s = 0$ if and only if the corresponding 3D lines form a LLC whose kernel is L, i.e., a set of lines S that intersect at L.

constraints for the constituent homography matrices A, B and for $C = BA^{-1}$:

$$CG_i^\top + G_i C^\top = 0, \tag{3}$$
$$BW_i^\top + W_i B^\top = 0, \tag{4}$$
$$AE_i^\top + E_i A^\top = 0, \tag{5}$$

for $i = 1, 2, 3$.

Theorem 3 provides the basis for deriving the "internal consistency" constraints which are 11 non-linear functions on the elements of the tensor that must be satisfied. The details can be found in the full version of this work in [9].

The slicing breakdown can also be useful for performing a direct image-to-image mapping, thus bypassing the need to recover the constituent homography matrices A, B. Consider two slices $\delta_k H^{ijk}$ and $\mu_k H^{ijk}$ for some δ, μ and denote the matrices by E_1, E_2. Let $p' \cong s_1 \times s_2$ for some two lines s_1, s_2. One can verify that:

$$p \cong (E_1^\top s_1 \times E_2^\top s_1) \times (E_1^\top s_2 \times E_2^\top s_2)$$

Thus, given p' and the tensor H^{ijk} one can determine directly the matching point p.

2.1 Concluding Notes

In summary, we have introduced the tensor H^{ijk} representing the joint mapping among three views of a planar surface. The tensor is determined uniquely by 4 matching points or 4 matching lines but in addition lives in a lower dimensional manifold — a fact that places internal non-linear constraints on the

27 entries of the tensor. These internal constraints ensure the group property of collineations arising from a *single* planar surface. In other words, the concatenation of collineations to form a joint mapping could be useful in practice when dealing with sequence of views of a planar surface. Furthermore, we have described in detail the tensor contractions (slices) and their geometric role — notably the role played by the Linear Line Complex configuration.

It is worthwhile noting that the structure of the Htensor bares similarity to the structure of the quadrifocal tensor Q^{ijkl} which is contracted with 4 lines across 4 views: $q_i s_j r_k t_l Q^{ijkl} = 0$ where q, s, r, t are coincident with 4 matching points p, p', p'', p''' across the images. The difference is that Q^{ijkl} applies to a general 3D world whereas the Htensor applies to a coplanar configuration. Proposition 1, for example, stating that the number of linear constraints drop gradually as matching points are introduced is analogous to the gradual drop in independent constraints for the quadrifocal tensor. In fact, the Htensor is a contraction, say $t_l Q^{ijkl}$, a slice, of the quadrifocal tensor where the plane π is determined by the choice of t_l in the example — and as such one can represent the quadrifocal tensor as a sum of three outer-products of epipoles and Htensors. Further details on this issue can be found in a companion paper in these proceedings [10].

In the next section we will explore the dual form H_{ijk} of the Htensor. The dual form turns out to be of particular interest as it applies to *dynamic* point configurations — a feature which opens up new application frontiers as well.

3 The Dual Homography Tensor H_{ijk}

Consider the tensor made up from H^{ijk} but replacing the homographies A, B (in eqn. 2) by their duals A^{-T} and B^{-T}. Denote $A' = A^{-1}$ and $B' = B^{-1}$, then the dual homography tensor is the covariant tensor described below:

$$H_{ijk} = \epsilon_{inu} a'^n_j b'^u_k. \tag{6}$$

Because H_{ijk} is a covariant tensor it applies to 3 points, one in each view. Consider, therefore, the contraction $p^i p'^j p''^k H_{ijk} = 0$. What does that entail on the relationship between p, p', p''? We have:

$$p^i p'^j p''^k H_{ijk} = p^\top (A'p' \times B'p'') = det(p, A'p', B'p'') = 0.$$

In other words, $p^i p'^j p''^k H_{ijk} = 0$ when the rank of the 3×3 matrix $[p, A'p', B'p'']$ is either 1 or 2. The rank is 1 iff the points p, p', p'' match in the usual sense when the three optical rays intersect at a single point in space. The rank is 2, however, when the three optical rays *meet at a line on* π because then $p, A'p'$ and $B'p''$ are collinear points in view 1 (note that A', B' are collineations from view 2 to 1 and view 3 to 1, respectively — see Fig. 4). We thus make the following definitions:

Definition 1. *A triplet of points* p, p', p'' *are said to be* **matching with respect to a static point** *if they are matching in the usual sense of the term,* *i.e., the corresponding optical rays meet at a single point. The triplet are said to be* **matching with respect to a moving point** *if the three optical rays meet at a line.*

We see from the above that the dual Htensor H_{ijk} applies to both static and moving points coming from the planar surface π. The possibility of working with static and moving elements was introduced recently in [1] where it was shown that if a moving point along a general (in 3D) straight path is observed in 5 views, and the camera projection matrices are *known*, then it is possible to set up a linear system for estimating the 3D line. With the dual Htensor H_{ijk}, on the other hand, we have no knowledge of the camera projection matrices, but on the other hand we require that the straight paths the points are taking should all be coplanar (what makes it possible to work with 3 views instead of 5 and not require prior information on camera positions). We have from above the following theorem:

Theorem 2. *The tensor H_{ijk} can be uniquely defined from image measurements associated with moving points only. A triplet of matching points p, p', p'' with respect to a moving point on π contributes one linear constraint $p^i p'^j p''^k H_{ijk} = 0$ on the entries of H_{ijk}.*

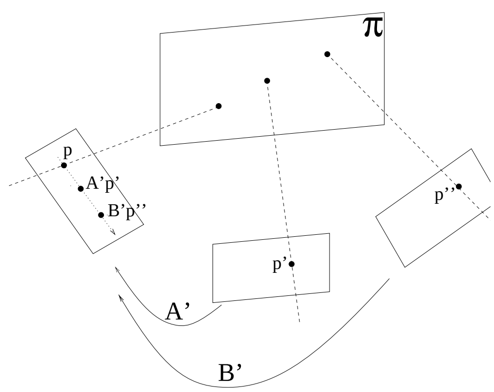

Fig. 4. The dual homography tensor and moving points. The collineations A', B' are from view 2 to 1 and 3 to 1 respectively. If the triplet p, p', p'' are projections of a moving point along a line on π then $p, A'p', B'p''$ are collinear in view 1. Thus, $p^\top(A'p' \times B'p'') = 0$, or $p^i p'^j p''^k H_{ijk} = 0$ where $H_{ijk} = \epsilon_{inu} a'^n_j b'^u_k$.

With 26 matching triplets with respect to moving points on π we can obtain a unique linear solution of the dual tensor. From the principle of duality with Proposition 1 we can state that there could be at most 8 moving points on the first line trajectory on π, at most 7 moving points on the second line, at most 6 points on the third line and at most 5 points on the fourth line. The four trajectory lines should be in general position, i.e., no three of them are concurrent.

If a triplet of points p, p', p'' are *known* to arise from a static point, then by principle of duality with Proposition 2 such a triplet provides 7 constraints on H_{ijk} and thus 4 matching triplets that are *known* to arise from static points (in general position) provide a unique solution for for the dual tensor.

3.1 Mixed Static and Dynamic Points

We have so far applied the principle of duality to assert that the 26 matching triplets with respect to moving points should be arranged along at least 4 line trajectories and that 4 matching triplets arising from static points are sufficient for a solution of the dual Htensor. The dual tensor raises also the possibility of handling a *mixed* situation where some of the matching triplets arise from moving points and some from static points — but without any prior knowledge of what is static and what is dynamic. We call this situation of having a matching triplets without a label of whether they arise from a static or dynamic point as an "unlabeled matching triplet". In this section we will address the following issues:

- In an unlabeled situation what is the maximum number of matching triplets arising from static points that are allowed for a unique solution? We will show that the number is 10, i.e., that among the 26 triplets at least 16 should arise from moving points.
- In case $x \leq 4$ of the matching triplets are labeled as static, how many moving points are required for a unique solution? We will show that we need $16-4x$ triplets arising from moving points.

Theorem 3. *In a situation of unlabeled matching triplets arising from a mixture of static and moving points, let $x \leq 4$ be the number of labeled matching triplets that are known a priori to arise from static points. If $x = 0$, then the matching triplets arising from static points contribute at most 10 linearly independent constraints, therefore the minimal number of matching triplets arising from moving points must be 16. In general, the minimal number of matching triplets arising from moving points is $16 - 4x$ for $x \leq 4$.*

Proof: It is sufficient to prove this theorem for the case where $A = B = I$ (the identity matrix) — because all other cases are transformed into this one by local change of coordinates.

Consider first the case $x = 0$, i.e., all 26 measurements are of the form $p^i p'^j p''^k H_{ijk} = 0$ regardless whether the matching triplet arises from a static or moving point. We wish to show that the dimension of the estimation matrix in case all the measurements arise from static points is 10. Each row of the estimation matrix is some "constraint tensor" G^{ijk} such that $G^{ijk}H_{ijk} = 0$. In the case $A = B = I$, G^{ijk} is a symmetric tensor if the matching triplet arises from a static point (because $p = p' = p''$), i.e., remains the same under permutation of indices — hence contains only 10 different groups of indexes

$$111, 222, 333, 112, 113, 221, 223, 331, 332, 123$$

up to permutations. Therefore, the rank of the estimation matrix from unlabeled static points is at most 10, and in order to solve for the tensor we would have to use at least 16 additional moving points.

Consider the case $x = 1$, i.e., one of the matching triplets contributed 9 constraints of rank 7:

$$\begin{array}{lll} p^i p'^j e_1^k H_{ijk} = 0 & p^i e_1^j p''^k H_{ijk} = 0 & e_1^i p'^j p''^k H_{ijk} = 0 \\ p^i p'^j e_2^k H_{ijk} = 0 & p^i e_2^j p''^k H_{ijk} = 0 & e_2^i p'^j p''^k H_{ijk} = 0 \\ p^i p'^j e_3^k H_{ijk} = 0 & p^i e_3^j p''^k H_{ijk} = 0 & e_3^i p'^j p''^k H_{ijk} = 0, \end{array}$$

where e_1, e_2, e_3 are the standard basis $(1,0,0), (0,1,0), (0,0,1)$. Note that because $A = B = I$, then $p = p' = p''$. Add the three constraints in the first row:

$$G^{ijk} = p^i p^j e_1^k + p^i e_1^j p^k + e_1^k p^j p^k$$

Then, G^{ijk} is a symmetric tensor and thus spanned by the 10-dimensional subspace of the unlabeled static points. Likewise, the constraint tensors resulting from adding the constraint of the second and third row above are also symmetric. Taken together, 3 out of the 7 constraints contributed by a labeled static point are already accounted for by the space of unlabeled static points. Therefore, each labeled static point adds only 4 linearly independent constraints. □

3.2 Contractions of dual Htensor

A double contraction of the tensor performs a point-point to line mapping. For example, $p^i p'^j H_{ijk}$ is a line in view 3 which is the projection of the line on π traced by the moving point onto view 3. In other words, given any two non-matching points p, p' let the line passing through the two intersecting points between the optical rays and π be denoted by L. Then, $p^i p'^j H_{ijk}$ is the projection of L onto view 3, so that any point p'' coincident with the projection will form a matching triplet p, p', p'' associated with a moving point tracing the line L on π.

A single contraction is a correlation mapping points to concurrent lines. Consider, for example, $\delta^k H_{ijk}$ for some contravariant vector (a point in view 3) δ. One can verify by substitution in eqn. 6 that the resulting matrix is $E = [\mu]_x A'$ where $\mu = B'\delta$. Let the matching points of δ in views 1,2 be μ, η respectively. Then, by duality with H^{ijk} we have that $E\eta = 0$ and $E^\top \mu = 0$. Furthermore, Ep' is a line passing through μ and $A'p'$ in view 1. Therefore, E maps the points in view 2 onto concurrent lines that intersect at a fixed point μ, and likewise, $E^\top$ maps points in view 1 onto concurrent lines that intersect at a fixed point η in view 2. Furthermore, $p^\top E p' = 0$ for all pairs of p, p' on matching lines through the fixed points μ, η (see Fig. 5).

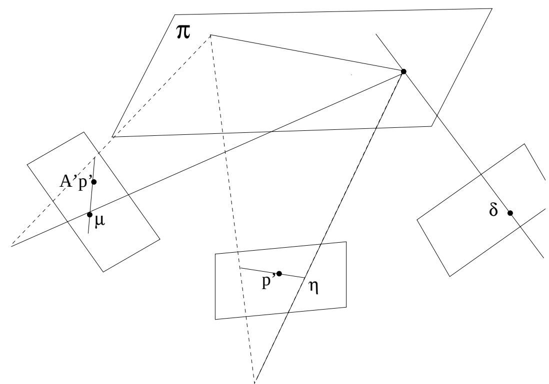

Fig. 5. A single contraction, say $\delta^k H_{ijk}$, is a mapping E between views 1,2 from points to concurrent lines. The null spaces of E and E^T are the matching points μ, η of δ in views 1,2. The image points p' are mapped by E to the lines $A'p' \times \mu$ and the image points p are mapped by $E^\top$ to the lines $A'^{-1}p \times \eta$ in view 2. The bilinear relation $p^\top Ep' = 0$ is satisfied for all pairs of p, p' on matching lines through the fixed points μ, η.

The constituent homography matrices A', B' can be extracted from the slices of H_{ijk} as follows. Let E_1, E_2, E_3 correspond to the three slices of $\delta^k H_{ijk}$ by letting δ be $(1,0,0), (0,1,0)$ and $(0,0,1)$ respectively. Then, $A'^\top E_i + E_i^\top A' = 0$, $i = 1,2,3$. Likewise, B' satisfies such a relation on the slices $\delta^j H_{ijk}$, and the homography $A'^{-1}B$ from the slices $\delta^i H_{ijk}$.

In summary, the dual form of the homography tensor applies to both cases: optical rays meet at a single point (matching points with respect to a static point) and optical rays meet at a line on π (matching points with respect to a moving point). In the case where no distinction can be made to the source of a matching triplet p, p', p'' (static or moving) then we have seen that in a set of at least 26 such matching triplets, 16 of them *must* arise from moving points. In case that a number $x \leq 4$ of these triplets are *known a-priori* to arise from static points, then $16 - 4x$ must arise from moving points. Once the dual tensor is recovered from image measurements it forms a mapping of both moving and static points and in particular can be used to *distinguish* between moving and static points (a triplet p, p', p'' arising from a static point is mapped to null vectors $p^i p'^j H_{ijk}$,$p^i p''^k H_{ijk}$ and $p'^j p''^k H_{ijk}$). The dual Htensor can be useful in practice to handle situations rich in dynamic motion seen from a monocular sequence.

4 Experiments

We conducted tests on the performance of H^{ijk} compared to pairwise homography recovery, and tests on H_{ijk} in order to evaluate the performance on static and moving point configurations. The full details on the experiments of H^{ijk} can be found in [9] which we will briefly summarize its conclusions here. The Htensor is recovered using standard robust estimators for least-squares estimation. The non-linear constraints were not taken into account. The reprojection performance using the recovered Htensor was consistently superior to the performance using a recovered homography matrix between pairs of views. On the other hand, we found out that recovering the constituent homography matrices from the Htensor produced significantly poorer results compared to the reprojection error achieved by the Htensor. Our conclusion is that the recovery of the homography matrix from H^{ijk} requires numerical conditioning which is beyond the scope of this work. It is worthwhile to note that recovering the homography matrix from the skew-symmetric relations with the slices of the tensor is identical to the way one can recover the fundamental matrix from two or more homography matrices. It has been shown empirically (R. Szeliski, private communication) that doing so for the fundamental matrix yields poor results.

In the second experiment, displayed in Fig. 6, we created a scene with mixed static and moving points. The moving points were part of 4 remote controls that were in motion while the camera changed position from one view to the next. The points were tracked along the three views, without knowledge what is static and what is moving. The triplet of matching points were fed into a least-square estimation for H_{ijk}. We then checked the error of reprojection on the static points —these were at sub-pixel level as can be seen in Fig. 6h — and the accuracy of the line trajectory of the moving points. Because the moving points were clustered on only 4 objects (the remote controls), then the accuracy was measured by "eye-balling" the parallelism of the trajectories of all points within a moving object. The lines are closely parallel as can be seen in Fig. 6f. The dual Htensor can also be used to segment the scene into static and moving points — this is shown in Fig. 6e.

5 Summary

Two views of a 3D scene are sufficient for performing reconstruction, yet there exist trifocal and quadrifocal tensors that concatenate 3 and 4 views and display an algebraic added value over 2-view reconstruction. In this paper we have done something similar to the well known Homography matrix — we have shown that there is an added value in investigating a 3-view analogue of the collineation operation. The resulting homography tensor H^{ijk} and its dual H_{ijk} are both intriguing and of practical value. The homography tensor places stronger constraints on the mapping across three views than concatenation of pairwise homography matrices — as evident by the coupling associated with a single linear system, the existence of the non-linear constraints, and the experimental results. This performance is comparable to the sub-space approach under infinitesimal motion recently presented in [12].

The dual Htensor, in our mind, shows promising potential for new application areas and explorations in strucutre from motion. The possibility of handling

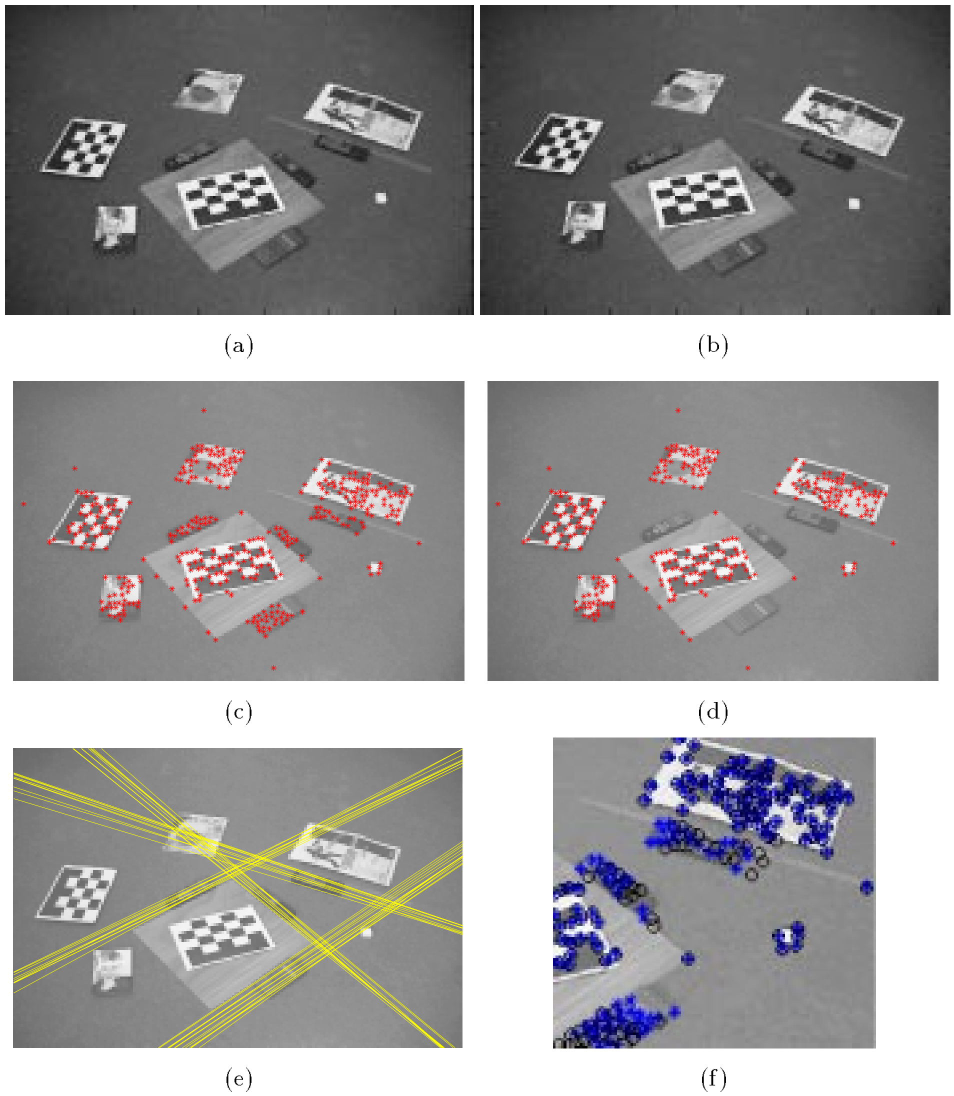

Fig. 6. (a),(b) two of three views of a planar scene with 4 remotes moving along straight lines. (c) The first view with the overlaid tracked points. These points were used for computing the dual homography tensor in a least-squares manner. (d) Segmenting the static from dynamic points using the recovered dual Htensor. Only the static points are shown. (e) The trajectory lines are overlaid on the third image — one can see that the lines of each remote are closely parallel thus providing an indication of accuracy of the dual Htensor. (f) Reprojection results using the Htensor as a point transfer mapping. Note that the static points are aligned with the the transferred points whereas the dynamic points are shifted relative to the transferred points.

static and dynamic points on equal grounds raises a host of new issues in which this paper only begins to address.

The extension of these tensors to higher dimensions is relatively straightforward — in that case the moving points in the dual tensor move along hyperplanes not lines, and the size of the tensor grows exponentially with dimension (and so does the number of constraints for static points). On the other hand, the restriction of the moving points to dimension $k < n - 1$ (in particular, $k = 2$ corresponds to motion along a line) is of more practical interest. For example, the case of $n = 3, k = 2$ has been explored in [11] where the application area is the extension of the classic 3D-to-3D alignment of point clouds to dynamic situations, such as when the structured light pattern attached to a sensor moves along with the sensor while the 3D reconstruction takes place.

References

1. S. Avidan and A. Shashua. Trajectory triangulation of lines: Reconstruction of a 3d point moving along a line from a monocular image sequence. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 1999.
2. A. Criminisi, I. Reid, and A. Zisserman. Duality, rigidity and planar parallax. In *Proceedings of the European Conference on Computer Vision*, Frieburg, Germany, 1998. Springer, LNCS 1407.
3. R.I. Hartley. Lines and points in three views and the trifocal tensor. *International Journal of Computer Vision*, 22(2):125–140, 1997.
4. M. Irani and P. Anandan. Parallax geometry of pairs of points for 3D scene analysis. In *Proceedings of the European Conference on Computer Vision*, LNCS 1064, pages 17–30, Cambridge, UK, April 1996. Springer-Verlag.
5. M. Irani, P. Anandan, and D. Weinshall. From reference frames to reference planes: Multiview parallax geometry and applications. In *Proceedings of the European Conference on Computer Vision*, Frieburg, Germany, 1998. Springer, LNCS 1407.
6. M. Irani, B. Rousso, and S. Peleg. Recovery of ego-motion using image stabilization'. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 454–460, Seattle, Washington, June 1994.
7. A. Shashua and N. Navab. Relative affine structure: Canonical model for 3D from 2D geometry and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):873–883, 1996.
8. 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.
9. A. Shashua and Lior Wolf. Homography tensors: On algebraic entities that represent three views of static or moving planar points. Technical Report CSE99-34, School of CSE, Hebrew University of Jerusalem, December 1999.
10. A. Shashua and Lior Wolf. On the structure and properties of the quadrifocal tensor. In *Proceedings of the European Conference on Computer Vision*, Dublin, Ireland, June 2000.
11. Lior Wolf, A. Shashua, and Y. Wexler. Join tensors: on 3d-to-3d alignment of dynamic sets. In *Proceedings of the International Conference on Pattern Recognition*, Barcelona, Spain, September 2000.
12. L. Zelnik-Manor and M. Irani. Multi-frame alignment of planes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 1999.