A PRIMARY DECOMPOSITION IN COMPUTER VISION

JUSTIN KATZ AND IRENA SWANSON

1. Introduction

2. The Multilinear Forms

A projective pinhole camera is modeled by rank 3 projective transformation $A : \mathbb{P}^3 \setminus \{f_A\} \to \mathbb{P}^2$ where f_A is the generator of the kernel of A. For a derivation of the camera matrix see [5]. Given a family of m rank-3 projective transformations $A_i : \mathbb{P}^3 - \{f_{A_i}\} \to \mathbb{P}^2$, the m-view scenario is modeled by the rational map

$$\Phi_A : \mathbb{P}^3 - \{f_1, ..., f_m\} \quad \to \quad \underbrace{\mathbb{P}^2 \times \cdots \times \mathbb{P}^2}_{\text{m times}}.$$

This model is imminently dissatisfactory because one imagines the multiview scenario as a projection of three dimensional projective space onto a family of two dimensional affine hyperplanes embedded in the ambient projective 3-space. By instead mapping into the product of (unembedded) 2-spaces, one would worry that the geometry that constrains the points on the camera planes is lost. For example, one thinks the epipolar geometry of the 2-view scenario is induced by the pencil of lines passing through the epipole, as obtained by intersecting the pencil of planes passing through the epipolar line with the image plane. Fortunately, the elemetary geometry can be recovered by algebraicizing [6, 5, 4] as follows.

Let $X \in \mathbb{P}^3$ and $A_1, \ldots, A_m$ be 3×4 camera matrices such that $A_i X = x_i \in \mathbb{P}^2$ for each $i \in \{1, .., m\}$. Corespondingly there exist representing $\vec{X} \in \mathbb{R}^4 \setminus \{0\}$, $\vec{x}_i \in \mathbb{R}^3 \setminus \{0\}$ and $\lambda_i \in \mathbb{R} \setminus \{0\}$ such that

$$A_i \vec{X} = \lambda_i \vec{x}_i.$$

We can compile this data into the equation

$$Mu = 0,$$

where

$$M = \begin{bmatrix} A_1 & \vec{x}_1 & 0_{3\times1} & \cdots & 0_{3\times1} \\ A_2 & 0_{3\times1} & \vec{x}_2 & \cdots & 0_{3\times1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_m & 0_{3\times1} & 0_{3\times1} & \cdots & \vec{x}_m \end{bmatrix} \qquad u = \begin{bmatrix} \vec{X} \\ \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}.$$

We call M the *master matrix of the m-view scenario.*

Now because u is nonzero, M has a nontrivial kernel, and the strict inequality

$$\operatorname{rank} M < m + 4$$

holds. Consequently, the determinant of any $(m+4) \times (m+4)$ submatrix of M will be zero. This suggests that if one thinks of the $\vec{x}_i$ as indeterminates in the polynomial ring $R := \mathbb{R}[x_1, ..., x_m, y_1, .., y_m, z_1, ..., z_m]$, then insisting that all $(m+4) \times (m+4)$ minors of M be zero will impose genuine polynomial constraints on the coordinates in the m-fold product of projective 2-spaces. There are three distinct types of choices of $(m+4)\times(m+4)$ minors, distinguished by how many rows are taken from how many cameras.

To construct the first type, pick two cameras A_i, A_j, and consider the $(m+4)\times(m+4)$ submatrix of M

$$D = \begin{bmatrix} A_1^{q_1} & \vec{x}_1^{q_1} & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_i & 0_{3\times1} & \cdots & \vec{x}_i & \cdots & 0_{3\times1} & \cdots & 0_{3\times1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_j & 0_{3\times1} & \cdots & 0_{3\times1} & \cdots & \vec{x}_j & \cdots & 0_{3\times1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_m^{q_m} & 0 & \cdots & 0 & \cdots & 0 & \cdots & \vec{x}_m^{q_m} \end{bmatrix}$$

where each $q_k \in \{1,2,3\}$ is picked arbitrarily.Taking the determinant of D, expanding along the columns containing single indeterminates, we discover

$$\det D = \pm\left(\prod_{\substack{k=1\\k\neq i,j}}^{m} \vec{x}_k^{q_k}\right)\det\begin{bmatrix} A_i & \vec{x}_i & 0 \\ A_j & 0 & \vec{x}_j \end{bmatrix} = 0.$$

Because we are in the projective setting and are viewing the indeterminates as modeling coordinates, there is some choice of each p_k such that $x_k^{p_k}$ will be nonzero. Thus, it is clear that the genuine constraint that arises in this way is the bilinear form

$$b_{ij}(\vec{x}_i,\vec{x}_j) := \det\begin{bmatrix} A_i & \vec{x}_i & 0 \\ A_j & 0 & \vec{x}_j \end{bmatrix} \quad = \quad 0.$$

We denote the ideal in R generated by all such bilinear forms as J_b, and call it the *unprojectivized bilinear ideal*. When we want to consider the unprojectivized bilinear ideal generated by some subset L of the cameras, we write $J_b(L)$.

Remark 1. Despite the radical differences in derivation, the bilinear forms b_{ij} are exactly the bilinear forms that arise by taking the inner product against the fundamental matrix F_{ij} between cameras i,j. That is to say, for any $x_i, x_j \in \mathbb{P}^2$ and fundamental matrix F_{ij}, we have

$$b_{i,j}(\vec{x}_i,\vec{x}_j) = \vec{x}_i^{\top} F_{ij}\vec{x}_j = 0.$$

This can be proved either by direct computation, as is done in [5] or by computing an elimination ideal as is done in [9]. In this way, we see that the algebraicization of the multiview scenario can be exploited to rederive geometric constraints.

The second type of $(m+4)\times(m+4)$ minor arises if one picks one distinguished camera A_i from which all three rows will be taken, two auxilliary cameras A_j and A_k from which two rows will be taken and one row from each of the other $m-3$ cameras. By expanding down each of the columns containing a single indeterminate, one finds the resulting *trilinear constraint* takes the form

$$T_{ij^{pq_j}k^{pq_k}}(\vec{x}_i,\vec{x}_j,\vec{x}_k) := \det \begin{bmatrix} A_i & \vec{x}_i & 0 & 0 \\ A_j^p & 0 & \vec{x}_j^p & 0 \\ A_j^{q_j} & 0 & \vec{x}_j^{q_j} & 0 \\ A_k^p & 0 & 0 & \vec{x}_k^p \\ A_k^{q_k} & 0 & 0 & \vec{x}_k^{q_k} \end{bmatrix} = 0,$$

where $p, q_j, q_k \in \{1,2,3\}$ are the rows that have been chosen from the two auxilliary cameras. We denote the ideal in R generated by all trilinear forms by J_t, and call it the *unprojectivized trilinear ideal*. As with the bilinear ideal, we denote the unprojectivized trilinear ideal generated by some subset of cameras L by $J(L)$.

The third type of minor arises by taking two rows from four cameras. While such a minor does give rise to a genuine constraint, the so-called *quadlinear constraint,* it follows algebraically from the bilinear and trilinear constraints [6].

3. The Ideals and Varieties

Despite their computational simplicity, the ideals J_t and J_b are not suitible to model the multiview scenario. This is because their cooresponding varieties will inevitibly contain the would-be 'origin'. To remedy such complications, we take a colon:

$$\begin{aligned} I_t &= J_t : \left(\prod_{i=1}^m (x_i, y_i, z_i)\right)^\infty \\ I_b &= J_b : \left(\prod_{i=1}^m (x_i, y_i, z_i)\right)^\infty \end{aligned}$$

so as to ensure that the resulting varieties are projectively sensible.

Definition 2. We call I_t the *trilinear ideal*, and I_b the *bilinear ideal*. As with the unprojectivized ideals, if the ideals are defined by a particular subset of the cameras L, then we denote the ideals $I_b(L)$ and $I_t(L)$.

Coorespondingly, we define

Definition 3. We define the *trilinear variety* as $\mathcal{V}_t = \mathcal{V}(I_t) \subset \left(\mathbb{P}^2\right)^m$, and similarly the *bilinear variety* $\mathcal{V}_b = \mathcal{V}(I_b) \subset \left(\mathbb{P}^2\right)^m$.

We are interested in studying the geometry constraining the points in the image $\Phi_A(\mathbb{P}^3 \setminus \{f_1, \ldots, f_m\}) \subset \left(\mathbb{P}^2\right)^m$, which merits its own definition.

Definition 4. The *natural descriptor* $\mathcal{V}_n$ is $\Phi_A(\mathbb{P}^3 \setminus \{f_1, \ldots, f_m\}) \subset \left(\mathbb{P}^2\right)^m$, and the *natural ideal* $I_n = I(\mathcal{V}_n)$ is its corresponding ideal.

Heyden and Astrom demonstrated that the natural descriptor is never a variety, but that its Zariski closure is $\mathcal{V}_b$ in the case of 2 cameras, and $\mathcal{V}_t$ for at least 3 cameras.

4. The Geometry

There are particular collections of points in the camera planes that, with hindsight, deserve definitions. First and foremost are the epipoles.

Definition 5. The j^{th} epipole on the i^{th} camera plane e_{ij} is the projection of the j^{th} focal point f_j onto the i^{th} camera plane. That is $e_{ij} := A_i f_j$.

The 2-view scenario has been solved, and reduces to the geometry induced by the epipoles. Specifically, points in the i^{th} camera plane dualize to lines passing through e_{ji} in the j^{th} camera plane. The dualization is representated by a (singular) matrix, the fundamental matrix. For an elementary treatment, see [5]. As discussed above, the fundamental matrix (as well as the epipolar geometry) is captured within the bilinear forms.

In the $m > 2$ -view scenario one has access to trilinear forms, for which the lines between between epipoles play a central role.

Definition 6. The trifocal line tl_{ijk} in camera plane i between cameras j and k is the line connecting e_{ij} and e_{ik}.

In the way that the epipoles represent degeneracy in the epipolar geometry, the trifocal line represents degeneracy in trifocal geometry. This will become evident in the proof of our central result.

5. Specialization and Simplification

The general (calibrated) camera matrix is typically written in the form

$$A = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot [R| - Rt]$$

for real numbers $\alpha_x, \alpha_y, s, x_0, y_0$, a rotation matrix R, and a 3 vector t. The entries in the matrix on the left account for internal camera parameters (e.g. skew, aspect ratio, film coordiantes), and the matrix on the right describes external parameters (e.g. world coordinates, and camera position). As a consequence of QR decomposition [5], any rank three 3×4 matrix can be written in this form, and can thus be viewed as a camera matrix. That said, the primary objects of study in this paper are ideals, which express little senstivity to the paramaters of the camera matrices in the sense of the following proposition:

Proposition 7. *The master matrix M, defined by m arbitrary rank three 3×4 matrices $\{A_i\}$, induces the same bilinear and trilinear ideals as some master matrix defined by m matrices of the form $\begin{bmatrix} I_{3\times3} | & \vec{a}_i \end{bmatrix}$, for some 3-vector $\vec{a}_i = (a_i, b_i, c_i)^\top$.*

The proof will follow from a series of lemmas. First, we establish that there is flexibility in the manipulation of the columns of each A_i

Lemma 8. *Let k be an infinite field, and let A be an $m \times n$ matrix with entries in k of rank r. Then there exists an invertible $n \times n$ matrix U such that the first r columns in AU are linearly independent. Furthermore, there exists a non-empty Zariski-open subset of k^{n^2} such that for all $(u_{ij}) \in k^{n^2}$, the matrix $U = [u_{ij}]_{i,j}$ has the property above.*

Proof. If columns $c_1, \dots, c_r$ of A are linearly independent, then we can set U to be the matrix obtained from the identity matrix by permuting the rows so that for $j = 1, \dots, r$, the jth column of U is e_{c_j}. Then the first r columns of AU are precisely the columns $c_1, \dots, c_r$ of A, and are thus linearly independent.

In these same columns, let $d_1, \dots, d_r$ be rows that are linearly independent.

Now let U be an $n \times n$ matrix with indeterminate entries X_{ij}. Let D be the determinant of U and let D' be the determinant of the submatrix of AU consisting of the first r columns and of rows $d_1, \dots, d_r$. Then DD' is a non-zero polynomial in $k[X_{ij} : i, j]$. By the first paragraph of the proof, there exists an n^2-tuple (u_{ij}) such that DD' evaluated at it is non-zero. Then the non-empty Zariski-open subset in the statement of the lemma is $k^{n^2} \setminus Z(DD')$. $\square$

Now, the coarseness of the Zariski topology permits enough freedom for to manipulate the columns of each A_i simulataneously, in the sense of

Lemma 9. *Let k be an infinite field. For each $i = 1, \dots, e$ let A_i be an $m_i \times n$ matrix of rank r_i. Then there exists a non-empty Zariski-open subset of k^{n^2} such that for all $(u_{ij}) \in k^{n^2}$, the $n \times n$ matrix $U = [u_{ij}]_{i,j}$ is invertible and for each $i = 1, \dots, e$, the first r_i columns of A_iU are linearly independent.*

Proof. By the previous lemma, for each $i = 1, \dots, e$, there exists a non-empty Zariski-open subset V_i of n^2 such that for all $(u_{ij}) \in k^{n^2}$, the $n \times n$ matrix $U = [u_{ij}]_{i,j}$ is invertible and the first r_i columns of A_iU are linearly independent. Let $V = V_1 \cap \cdots \cap V_e$. Then V is Zariski-open and non-empty as it is a finite intersection of non-empty Zariski-open sets over an infinite field. $\square$

We now have the tools to prove the above proposition.

Proof. (of proposition) By Lemma 8 there exists an invertible 4×4 matrix U such that the first three columns in each of $A_1U, \dots, A_mU$ are linearly independent. Let U' be the $(4+m) \times (4+m)$ matrix whose top left 4×4 block is U, the bottom right $m \times m$ block is the identity matrix, and all other entries are 0. Then U' is an invertible matrix. Note that M and MU' are identical in columns 5 through $4+m$, so the upshot is that MU' is still in the same master-matrix form as M, with camera matrices being A_iU instead of A_i. By construction the first three columns of MU' have the property that for $i = 0, \dots m-1$, rows $3i+1, 3i+2, 3i+3$ are linearly independent. Since elementary column operations on the first 4 columns do not change the ideal of all 6×6 minors, we have that that the ideals generated by bilinear and trilinear forms will be the same for M and MU'. Thus to prove the theorem it suffices to assume that in each A_i the first three columns are linearly independent.

With this assumption there exists a 3×3 invertible matrix V_i such that the submatrix of the first three columns in V_iA_i is the identity matrix. Let V be the block-diagonal $(3m) \times (3m)$ matrix whose ith diagonal block is V_i. Then V is an invertible matrix, and

$$VM = \begin{bmatrix} [I_{3\times 3}|\vec{a}_1] & V_1\vec{x}_1 & \cdots & 0_{3\times 1} \\ \vdots & \vdots & \ddots & \vdots \\ [I_{3\times 3}|\vec{a}_m] & 0_{3\times 1} & \cdots & V_m\vec{x}_m \end{bmatrix}$$

where $\vec{a}_1, \dots, \vec{a}_m$ are some linear combination of the original 4 rows of each A_i. Note that the entries in $V_i\vec{x}_i$ are variables, independent of all other $\vec{x}_j$ for $j \neq i$. So we name these entries $\vec{x'}_i$. Certainly the ideals (x_i, y_i, z_i) and (x_i', y_i', z_i') are identical, and the polynomial rings $\mathbb{R}[x_1, y_1, z_1, \dots, x_m, y_m, z_m]$ and $\mathbb{R}[x_1', y_1', z_1', \dots, x_m', y_m', z_m']$ are identical via a linear change of variables lifting each linear change of variables $\mathbb{R}[x_i, y_i, z_i] = \mathbb{R}[x_i', y_i', z_i']$ by V_i. By row operations restricted to single cameras, the bilinear and trilinear ideals will be the same for M and VM. Thus, we may assume that every camera matrix is of the form $A_i = [I_{3\times 3}|\vec{a}_i]$. □

Remark 10. With the simplified master matrix, it is easy to derive simple explicit formulas for the bilinear and trilinear forms. Indeed we compute by subtracting the first row from the fourth, the second from the fifth, and the third from the sixth, and expanding by cofactors along the upper right 3×3 diagonal:

$$\begin{aligned} b_{ij}(\vec{x}_i,\vec{x}_j) &= \det\begin{bmatrix} [I_{3\times 3}|\vec{a}_i] & \vec{x}_i & 0_{3\times 1} \\ [I_{3\times 3}|\vec{a}_j] & 0_{3\times 1} & \vec{x}_j \end{bmatrix} = 0 \\ &= \det\begin{bmatrix} \vec{a}_i - \vec{a}_j & -\vec{x}_i & \vec{x}_j \end{bmatrix} = 0. \end{aligned}$$

This form elucidates the geometry of the 2 view scenario, in the following sense. Recall that the the j^{th} epipole in the i^{th} camera frame is $e_{ij} = A_i f_j$. Here, we are assuming each A_i is of the form $[I_{3\times 3}|\vec{a}_i]$, and the generator of the kernel of A_j is $f_j = \left(\vec{a}_j^\top, -1\right)^\top$. We now compute $e_{ij} = A_i f_j = \vec{a}_i - \vec{a}_j$. Recall that three points in the projective plane x_1, x_2, x_3, represented by vectors $\vec{x}_1, \vec{x}_2, \vec{x}_3$ are colinear if and only if $\det\begin{bmatrix}\vec{x}_1 & \vec{x}_2 & \vec{x}_3\end{bmatrix} = 0$. Thus, the bilinear form can be seen as a *colinearity requirement between the image points and the epipole.*

To compute a formula for the trilinear form, pick $p, q_j, q_k \in \{1,2,3\}$ and as a above, subtract the p^{th} row from the $(3+\text{p})^{th}$ row, the q_j^{th} row from the $(3+q_j)^{th}$ row, the p^{th} row from the $(5+p)^{th}$ row, and the q_k^{th} row from the $(5+q_k)^{th}$ row, then repeatedly expand expand by cofactors to get

$$\begin{aligned} T_{ij^{pq_j}k^{pq_k}}(\vec{x}_i,\vec{x}_j,\vec{x}_k) &= \det\begin{bmatrix} [I_{3\times 3}|\vec{a}_i] & \vec{x}_i & 0 & 0 \\ [I_{3\times 3}|\vec{a}_j]^p & 0 & \vec{x}_j^p & 0 \\ [I_{3\times 3}|\vec{a}_j]^{q_j} & 0 & \vec{x}_j^{q_j} & 0 \\ [I_{3\times 3}|\vec{a}_k]^p & 0 & 0 & \vec{x}_k^p \\ [I_{3\times 3}|\vec{a}_k]^{q_k} & 0 & 0 & \vec{x}_k^{q_k} \end{bmatrix} = 0 \\ &= \det\begin{bmatrix} \vec{a}_{ji}^p & -\vec{x}_i^p & \vec{x}_j^p & 0 \\ \vec{a}_{ji}^{q_j} & -\vec{x}_i^{q_j} & \vec{x}_j^{q_j} & 0 \\ \vec{a}_{ki}^p & -\vec{x}_i^p & 0 & \vec{x}_k^p \\ \vec{a}_{ki}^{q_k} & -\vec{x}_i^{q_k} & 0 & \vec{x}_k^{q_k} \end{bmatrix} = 0 \\ &= \vec{a}_{ji}^p \det\begin{bmatrix} -\vec{x}_i^{q_j} & \vec{x}_j^{q_j} & 0 \\ -\vec{x}_i^p & 0 & \vec{x}_k^p \\ -\vec{x}_i^{q_k} & 0 & \vec{x}_k^{q_k} \end{bmatrix} - \vec{a}_{ji}^{q_j} \det\begin{bmatrix} -\vec{x}_i^p & \vec{x}_j^p & 0 \\ -\vec{x}_i^p & 0 & \vec{x}_k^p \\ -\vec{x}_i^{q_k} & 0 & \vec{x}_k^{q_k} \end{bmatrix} \\ &\quad + \vec{a}_{ki}^p \det\begin{bmatrix} -\vec{x}_i^p & \vec{x}_j^p & 0 \\ -\vec{x}_i^{q_j} & \vec{x}_j^{q_j} & 0 \\ -\vec{x}_i^{q_k} & 0 & \vec{x}_k^{q_k} \end{bmatrix} - \vec{a}_{ki}^{q_k} \det\begin{bmatrix} -\vec{x}_i^p & \vec{x}_j^p & 0 \\ -\vec{x}_i^{q_j} & \vec{x}_j^{q_j} & 0 \\ -\vec{x}_i^p & 0 & \vec{x}_k^p \end{bmatrix} = 0 \\ &= \left(\vec{a}_{ji}^p \vec{x}_j^{q_j} - \vec{a}_{ji}^{q_j}\vec{x}_j^p\right) \det\begin{bmatrix} -\vec{x}_i^p & \vec{x}_k^p \\ -\vec{x}_i^{q_k} & \vec{x}_k^{q_k} \end{bmatrix} \\ &\quad + \left(\vec{a}_{ki}^p \vec{x}_k^{q_k} - \vec{a}_{ki}^{q_k}\vec{x}_k^p\right) \det\begin{bmatrix} -\vec{x}_i^p & \vec{x}_j^p \\ -\vec{x}_i^{q_j} & \vec{x}_j^{q_j} \end{bmatrix} = 0, \end{aligned}$$

where we have introduced the shorthand $\vec{a}_{ij} := \vec{a}_i - \vec{a}_j$. The explicit computation of this formula will allow us to recognize a polynomial as a trilinear form in a later computation. One should note by looking at the form in the second line in the computation, if $q_j = q_k = q$, then the form is invariant (up to sign) under permutation of ijk . This is to say that for $p, q \in \{1, 2, 3\}$, $T_{ij^{pq}k^{pq}} = T_{ji^{pq}k^{pq}} = T_{ki^{pq}j^{pq}}$. This fact will prove useful in a later computation.

6. The Primary Decomposition

To motivate the central theorem, we begin with a dimension count. Let $m \geq 4$ and consider the m-view scenario. Because the bilinear $\mathcal{V}_b$ is the Zariski closure of the the image of a nondegenerate rational map from $\mathbb{P}^3$ into the m-fold product of $\mathbb{P}^2$, we know that its (projective) dimension is 3. One would hope that, because the codimension of a (projective) variety is the cardinality of the maximum number of algebraically independent functions in its coordinate ring, I_b can be generated by $2m-3 = \dim\left(\mathbb{P}^2\right)^m - \dim I_b$. Unfortunately this cannot always be the case, as Heyden and Astrom demonstrated in the case $m = 4, 5$, the variety generated by 5 and 7 bilinear forms (respectively) is always strictly larger than that generated by I_b.

This discrepancy leads to the study of a particular subset of the bilinearities that uniformly uses all cameras. We define

$$\tilde{J}_b = \sum_{i=1}^{m-1}(b_{i,i+1}) + \sum_{i=1}^{m-2}(b_{i,i+2})$$

and set

$$\tilde{I}_b = \tilde{J}_b : \left(\prod_{i=1}^{m}(x_i, y_i, z_i)\right)^{\infty}$$

for $i \in \{1, ..., m-1\}$. When we want to be explicit about which cameras we are using to generate $\tilde{I}_b$, we write $\tilde{I}_b(L)$ for a list of cameras L. Despite $\tilde{I}_b$ proveably not being the whole of I_b, Heyden and Astrom conjectured that it captures much of the complete m view scenario, in the sense of the following theorem.

Theorem 11. *I_t is a primary component of $\tilde{I}_b$ for all $m \geq 4$.*

Corollary 12. *$\mathcal{V}_t = \overline{\mathcal{V}_n}$ is an irreducible subvariety of $\mathcal{V}(\tilde{I}_b)$ for all $m \geq 4$.*

The proof of Theorem 11 will rely on the following lemmas.

Lemma 13. *For $w := \det\begin{bmatrix}\vec{x}_1 & \vec{x}_2 & \vec{x}_3\end{bmatrix}$ we have $w \cdot I_t(1,2,3) \in \tilde{I}_b(1,2,3)$.*

Proof. Consider the following element of $\tilde{I}_b$, where $p, q_2, q_3 \in \{1, 2, 3\}$

$$r = w_1^{q_3} w_2^{q_2} b_{12} + w_2^{q_2} w_3^{q_3} b_{23} - w_1^{q_2} w_3^{q_3} b_{13},$$

where w_j^i is the cofactor of w with the i^{th} row and j^{th} column removed. Rewriting each b_{ij} in the explicit form derived in Remark 10, we have

$$\begin{aligned} r \quad = \quad & w_1^{q_3} w_2^{q_2}\left(\vec{a}_{21}^{q_2} w_3^{q_2} - \vec{a}_{21}^{q_3} w_3^{q_3} + \vec{a}_{21}^{p} w_3^{p}\right) \\ & + w_2^{q_2} w_3^{q_3}\left(\vec{a}_{32}^{q_2} w_1^{q_2} - \vec{a}_{32}^{q_3} w_1^{q_3} + \vec{a}_{32}^{p} w_1^{p}\right) \\ & - w_1^{q_2} w_3^{q_3}\left(\vec{a}_{31}^{q_2} w_2^{q_2} - \vec{a}_{31}^{q_3} w_2^{q_3} + \vec{a}_{31}^{p} w_2^{p}\right) \end{aligned}$$

Then replacing $\vec{a}_{31}$ by $\vec{a}_{32}+\vec{a}_{21}$ and collecting by coefficients, we have

$$\begin{aligned} r \;=\; & \vec{a}_{21}^{q_2} w_2^{q_2}\left(w_1^{q_3} w_3^{q_2}-w_1^{q_2} w_3^{q_3}\right) \\ & -\vec{a}_{21}^{q_3} w_3^{q_3}\left(w_1^{q_3} w_2^{q_2}-w_1^{q_2} w_2^{q_3}\right) \\ & +\vec{a}_{21}^{p}\left(w_3^{p} w_1^{q_3} w_2^{q_2}-w_1^{q_2} w_3^{q_3} w_2^{p}\right) \\ & +\vec{a}_{32}^{q_2} w_1^{q_2}\cdot 0 \\ & -\vec{a}_{32}^{q_3} w_3^{q_3}\left(w_1^{q_3} w_2^{q_2}-w_1^{q_2} w_2^{q_3}\right) \\ & +\vec{a}_{32}^{p} w_3^{q_3}\left(w_1^{p} w_2^{q_2}-w_1^{q_2} w_2^{p}\right) \end{aligned}$$

To induce further simplifications, we invoke a lemma from [7] .

Lemma 14. *Let M be an $n\times n$ matrix with $n\geq 3$ and denote the determinant of the submatrix obtained by deleting the i^{th} row and j^{th} column by M_j^i. Fix $1\leq i,j\leq n$, with $i\neq j$. Let N be the $n-2\times n-2$ matrix obtained from M by deleting the i^{th} and j^{th} row and column. Then $\det M \det N = M_i^i M_j^j - M_i^j M_j^i$.*

Specializing the lemma to w, we further reduce

$$\begin{aligned} r \;=\; & -\vec{a}_{21}^{q_2} x_2^{p} w_2^{q_2} \det w \\ & +\vec{a}_{31}^{q_3} x_3^{p} w_3^{q_3} \det w \\ & -\vec{a}_{21}^{p}\left(w_1^{q_2} w_3^{q_3} w_2^{p}-w_2^{q_2} w_1^{q_3} w_3^{p}\right) \\ & +\vec{a}_{32}^{p} x_3^{q_3} w_3^{q_3} \det w. \end{aligned}$$

Then adding a suggestive $0=-\vec{a}_{21}^{p} w_2^{q_2} w_1^{p} w_3^{q_3}+\vec{a}_{21}^{p} w_2^{q_2} w_1^{p} w_3^{q_3}$, and using the lemma again

$$\begin{aligned} r \;=\; & -\vec{a}_{21}^{q_2} x_2^{p} w_2^{q_2} \det w \\ & +\vec{a}_{31}^{q_3} x_3^{p} w_3^{q_3} \det w \\ & -\vec{a}_{21}^{p} x_3^{q_3} w_3^{q_3} \det w \\ & -\vec{a}_{21}^{p} x_2^{q_2} w_2^{q_2} \det w \\ & +\vec{a}_{32}^{p} x_3^{q_3} w_3^{q_3} \det w. \end{aligned}$$

Rearranging, one sees

$$\begin{aligned} r &= \det w\cdot\left(\left(\vec{a}_{21}^{q_2} x_2^{p}-\vec{a}_{21}^{p} x_2^{q_2}\right) w_2^{q_2}+\left(\vec{a}_{31}^{q_3} x_3^{p}-\vec{a}_{31}^{p} x_3^{q_3}\right) w_3^{q_3}\right) \\ &= \det w\cdot T_{12^{p q_2} 3^{p q_3}} \in \tilde{I}_b. \end{aligned}$$

By symmetry, $\det w\cdot I_t\subset \tilde{I}_b$. □

Remark 15. By the lemma we have $I_t\subset \tilde{I}_b : w$ in the case of 3 cameras. Furthermore, in [6] it was demonstrated that I_t is always prime, so we immediately have the other containment $\tilde{I}_b : w\subset I_t : w = I_t$. Consequently $\tilde{I}_b : w = I_t$ for 3 cameras. With this in mind, Lemma 13 tells a rich geometric story. When the cameras in a 3-view scenario are not colinear, the trilinear constraints are valuable because they are nonzero when the bilinear constraints have common *false* zeroes. Indeed, consider the following points which make up $\mathcal{V}_n\subset\left(\mathbb{P}^2\right)^3$.

- An arbitrary point $x_1\in\mathbb{P}^2 - tl_{123}$ in the first image constrains $x_2\in\mathbb{P}^2 - tl_{213}$ to the line of points satisfying $b_{12}(x_1,x_2)=0$ in the second image, so then the back-projections of x_1 and x_2 intersect at a unique point, which projects to a unique point in the third image. Similarly for any permutation of $1,2,3$.

- If $x_1 = e_{12}$ in the first image, and $x_2 = e_{21}$ in the second image, then x_3 can lie anywhere on tl_{123} in the third. This cooresponds to the projection of an image on the epipolar line between e_{12} and e_{21}. Similarly for any permutation of $1, 2, 3$.
- If $\vec{x}_1 = e_{12}$ and $\vec{x}_2 = e_{21}$, then for any $\vec{x}_3 \in tl_{321}$, we will have $(\vec{x}_1, \vec{x}_2, \vec{x}_3) \in \mathcal{V}_n$. This is because the back-projections from $\vec{x}_1$ and $\vec{x}_2$ degenerate to a whole line, so that the trilinearities do not constrain $\vec{x}_3 \in tl_{312}$. Similarly for any permutation of $1, 2, 3$.

Looking at the explicit form of the generators of I_b,

$$\begin{aligned} b_{12}(\vec{x}_1, \vec{x}_2) &= \det \begin{bmatrix} e_{12} & \vec{x}_1 & \vec{x}_2 \end{bmatrix} = 0 \\ b_{23}(\vec{x}_2, \vec{x}_3) &= \det \begin{bmatrix} e_{23} & \vec{x}_2 & \vec{x}_3 \end{bmatrix} = 0 \\ b_{13}(\vec{x}_1, \vec{x}_3) &= \det \begin{bmatrix} e_{13} & \vec{x}_1 & \vec{x}_3 \end{bmatrix} = 0 \end{aligned}$$

We see that the preceeding points will be in $\mathcal{V}_b$, but also the *false* zeroes:

- Due to the relation $e_{13} = e_{12} + e_{23}$, if for any $t, p, q \in \mathbb{R}$
$$\begin{aligned} \vec{x}_1 &= te_{12} + (1-t)e_{13} \\ \vec{x}_2 &= pe_{21} + (1-p)e_{23} \\ \vec{x}_3 &= qe_{31} + (1-q)e_{32} \end{aligned} \tag{6.1}$$
then all of the bilinearities will be zero. This is to say that $\mathcal{V}_b$ contains all $(\vec{x}_1, \vec{x}_2, \vec{x}_3) \in tl_{123} \times tl_{213} \times tl_{312} \subset \left(\mathbb{P}^2\right)^3$, even if $\vec{x}_1, \vec{x}_2, \vec{x}_3$ are not mutually projections of one point in $\mathbb{P}^3$.
- If $\vec{x}_1 = e_{13}$ and $\vec{x}_2 = e_{23}$, then $e_{13} = e_{12} + e_{23}$ forces all three bilinearities to be zero, regardless of the value of $\vec{x}_3$. Thus $\mathcal{V}_b$ contains all $(e_{12}, e_{23}, \vec{x}_3) \in \{e_{12}\} \times \{e_{23}\} \times \mathbb{P}^2 \subset \left(\mathbb{P}^2\right)^3$. Geometrically, this is because the intersection of the back-projection of e_{13} and e_{23} is exactly f_3, camera center of camera 3. The projection of such a point does not exist in $\mathbb{P}^2$, so $\vec{x}_3$ is unconstrained. Similarly for any permutation of $1, 2, 3$.

The Zariski closure of $\mathcal{V}_n$ is $\mathcal{V}_t$. Because $\mathcal{V}_n \neq \mathcal{V}_t$, we know $\mathcal{V}_t$ also contains *false* zeroes. In addition to the three types of points coming from $\mathcal{V}_n \subset \mathcal{V}_t$, the second situation of $\mathcal{V}_b$ can arise in $\mathcal{V}_t$.

Thus
$$\mathcal{V}_n \subset \mathcal{V}_t \subset \mathcal{V}_b,$$
and coorespondingly
$$I_n \supseteq I_t \supset \tilde{I}_b.$$
Notice that $\tilde{I}_b : w_{123}^2 = (\tilde{I}_b : w_{123}) : w_{123} = I_t : w_{123} = I_t = \tilde{I}_b : w_{123}$, meaning that $\tilde{I}_b : w_{123} = \tilde{I}_b : w_{123}^\infty$. In general, taking (saturated) colons of ideals translates to closures of differences of varieties

$$\begin{aligned} \mathcal{V}(\tilde{I}_b : w^\infty) \overset{I_t \text{ is prime}}{=} \mathcal{V}(\tilde{I}_b : w) &= \overline{\mathcal{V}(\tilde{I}_b) - \mathcal{V}(w)} \\ &\overset{\text{lemma}}{=} \mathcal{V}(I_t). \end{aligned}$$

One notices that the first situation from $\mathcal{V}_b$ does not arise in $\mathcal{V}_t$, but that the relation $e_{13} = e_{12} + e_{23}$ demands that $w = \det \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \vec{x}_3 \end{bmatrix} = 0$ whenever equations 6.1 hold. Thus, for any $T \in I_t$, we must have $w \cdot T(\vec{x}_1, \vec{x}_2, \vec{x}_3) = 0$ for arbitrary points on respective trifocal lines. This is to say that w_{123} is zero exactly when one would hope to use the trilinear constraints.

Somewhat suprisingly, Lemma 13 extends to the scenario with four cameras.

Lemma 16. *For* $w_{123} := \det\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$ *and* $w_{234} := \det\begin{bmatrix} x_2 & x_3 & x_4 \end{bmatrix}$, *we have* $w_{123} \cdot w_{234} \cdot I_t(1,2,3,4) \subset \tilde{I}_b(1,2,3,4)$.

Proof. The result is true by Lemma 13 for trilinearities using consecutive triples of cameras, and by symmetry it suffices to all trilinearities using cameras $1,2,4$: $T_{12^{pq_2}4^{pq_4}}$, $T_{21^{pq_1}4^{pq_4}}$, $T_{41^{pq_1}2^{pq_2}}$ for all choices of $p, q_1, q_2, q_4 \in \{1,2,3\}$. Further, for all of the trilinearities of a fixed distinguished camera (say, camera 1), there is no algebraic distinction between each of the cases where $q_2 \neq q_4$, and likewise no distinction between the cases with $q_2 = q_4$. Finally, by the computation in Remark 10, if $q_2 = q_4 = q$ then the trilinearity $T_{12^{pq}4^{pq}}$ is the same as $T_{21^{pq}4^{pq}} = T_{41^{pq}2^{pq}}$. Thus it suffices to check one cases for each distinguished camera, in addition to the degenerate case where $q_2 = q_4$. The computations are similar to those from Lemma 13, so we will forego the explict computation, and instead offer each of the four polynomials written as an R linear combination of the bilinearities. We see

$$\begin{aligned}
w_{123} \cdot w_{234} T_{12^{12}4^{12}} \quad = \quad & \begin{vmatrix} x_2 & x_2 \\ y_2 & y_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} \cdot w_{234} \cdot b_{12} + \begin{vmatrix} x_2 & x_4 \\ y_2 & y_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \cdot w_{134} \cdot b_{23} \\
& + \begin{vmatrix} x_2 & x_4 \\ y_2 & y_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \cdot w_{123} b_{34} - \begin{vmatrix} x_2 & x_4 \\ y_2 & y_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \cdot w_{234} \cdot b_{13} \\
& - \begin{vmatrix} x_3 & x_4 \\ y_3 & y_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \cdot w_{123} \cdot b_{14} \\
w_{123} \cdot w_{234} T_{12^{12}4^{23}} \quad = \quad & \left(x_1 y_2 \begin{vmatrix} y_3 & y_4 \\ z_3 & z_4 \end{vmatrix} - x_3 y_2 \begin{vmatrix} y_1 & y_4 \\ z_1 & z_4 \end{vmatrix} + x_2 y_4 \begin{vmatrix} y_1 & y_3 \\ z_1 & z_3 \end{vmatrix} \right) w_{234} b_{12} \\
& + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \cdot \begin{vmatrix} y_2 & y_4 \\ z_2 & z_4 \end{vmatrix} w_{134} b_{23} + \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} \cdot \begin{vmatrix} y_2 & y_4 \\ z_2 & z_4 \end{vmatrix} w_{123} b_{34} \\
& - \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \cdot \begin{vmatrix} y_2 & y_4 \\ z_2 & z_4 \end{vmatrix} w_{234} b_{13} - \begin{vmatrix} y_3 & y_4 \\ z_3 & z_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} w_{123} b_{24} \\
w_{123} \cdot w_{234} T_{21^{23}4^{13}} \quad = \quad & \begin{vmatrix} y_1 & y_3 \\ z_1 & z_3 \end{vmatrix} \cdot \begin{vmatrix} x_2 & x_4 \\ y_2 & y_4 \end{vmatrix} w_{234} b_{12} + \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \cdot \begin{vmatrix} x_2 & x_4 \\ z_2 & z_4 \end{vmatrix} w_{134} b_{23} \\
& + \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \cdot \begin{vmatrix} x_2 & x_4 \\ z_2 & z_4 \end{vmatrix} w_{123} b_{34} - \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \cdot \begin{vmatrix} x_2 & x_4 \\ z_2 & z_4 \end{vmatrix} w_{234} b_{13} \\
& - \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \cdot \begin{vmatrix} x_3 & x_4 \\ z_3 & z_4 \end{vmatrix} w_{123} b_{24} \\
w_{123} \cdot w_{234} T_{41^{12}2^{23}} \quad = \quad & \begin{vmatrix} y_2 & y_4 \\ z_2 & z_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_3 \\ z_1 & z_3 \end{vmatrix} \cdot w_{234} \cdot b_{12} + \begin{vmatrix} y_2 & y_4 \\ z_2 & z_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} \cdot w_{134} b_{23} \\
& + \begin{vmatrix} y_2 & y_4 \\ z_2 & z_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} \cdot w_{123} b_{34} + \begin{vmatrix} y_2 & y_4 \\ z_2 & z_4 \end{vmatrix} \cdot \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} \cdot w_{234} \cdot b_{13} \\
& + \left(-x_4 z_1 \begin{vmatrix} y_2 & y_3 \\ z_2 & z_3 \end{vmatrix} + x_3 z_1 \begin{vmatrix} y_2 & y_4 \\ z_2 & z_4 \end{vmatrix} - x_1 z_2 \begin{vmatrix} y_3 & y_4 \\ z_3 & z_4 \end{vmatrix} \right) w_{123} b_{14}.
\end{aligned}$$

One can see from the form of the coefficients, that each of these cases represent an earnestly different algebraic phenomena. By symmetry, these phenomenad exhaust the possibilities. $\square$

With Lemmas 13 and 16 in hand, we are equipped to prove Theorem 11.

Proof. (Of Theorem 11) It is a general fact that for an ideal J and ring element p such that $J : p$ is primary, then $J : p$ is a primary component of J. The first step in the proof will thus be to demonstrate some ring element $w \notin I_t$ such that $w \cdot I_t \subset \tilde{I}_b$. Then, noting that $\tilde{I}_b \subset I_b = I_t$ and that I_t is a prime ideal, we will have $\tilde{I}_b : w \subset I_t : w = I_t$ forcing $\tilde{I}_b : w = I_t$.

The proof will follow by induction, with the base cases $m = 3, 4$ dispatched in Lemmas 13 and 16.

Let $m > 4$ and suppose there exists some w such that $w \cdot I_t \subset \tilde{I}_{bt}$ for any configuration of $m - 1$ cameras. By the induction hypothesis, we know that there exists some w_1,$w_2 \in R$ such that

$$w_1 \cdot I_t(1, ..., m-1) \subset \tilde{I}_b(1, .., m-1) \tag{6.2}$$

$$w_2 \cdot I_t(2, ..., m) \subset \tilde{I}_b(2, ..., m) \tag{6.3}$$

Thus, it will suffice to demonstrate that there is some w' such that

$$w' \cdot I_t(1, s, m) \subset \tilde{I}_b(1, ..., m)$$

for each $1 < s < m$. First, let $s < m - 1$. Then the case when $m = 4$, and by the trivial containment $I_t(1, s, m) \subset I_t(1, s, s+1, m)$, there must exist some w'_s such that

$$w'_s \cdot I_t(1, s, m) \quad \subset \quad \tilde{I}_b(1, s, s+1, m).$$

Granting that $\tilde{I}_b(1, s, s+1, m) \subseteq \tilde{I}(1, s, s+1) + \tilde{I}_b(s, s+1, m)$, and the trivial containment $\tilde{I}_b(L) \subset I_t(L')$ for $L \subset L'$, we find

$$w'_s \cdot I_t(1, s, m) \subset I_t(1, .., m-1) + I_t(2, .., m).$$

Using 6.2 and 6.3

$$\begin{aligned} w_1 \cdot w_2 \cdot w'_s \cdot I_t(1, s, m) \quad &\subset \quad \tilde{I}_b(1, .., m-1) + \tilde{I}_b(2, .., m) \\ &= \quad \tilde{I}_b. \end{aligned}$$

Doing this for all s, gives

$$w_1 \cdot w_2 \cdot \prod_{s=2}^{m-2} w'_s \cdot I_t \subset \tilde{I}_b,$$

completing the induction. □

References

[1] Chris Aholt, Bernd Sturmfels, and Rekha Thomas. A hilbert scheme in computer vision, 2012.

[2] Michael Atiyah and Ian Grant Macdonald. Introduction to commutative algebra, 1969.

[3] David A. Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

[4] Olivier Faugeras and Bernard Mourrain. On the geometry and algebra of the point and line correspondences between n images, 1995.

[5] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[6] Anders Heyden and Kalle Åström. Algebraic properties of multilinear constraints, 1996.

[7] Greg Martin and Erick B. Wong. Almost all integer matrices have no integer eigenvalues. *The American Mathematical Monthly*, 116(7):588–597, 2009.

[8] E.G. Rees. *Notes on geometry*. Universitext (1979). Springer-Verlag, 1983.

[9] Michael Werman and Amnon Shashua. The study of 3d-from-2d using elimination. In *In Proc. of Europe-China Workshop on Geometrical Modeling and Invariants for Computer Vision*, pages 94–101, 1995.