Structure from Motion (3)

Algebraic Varieties in Multiple View Geometry*

Anders Heyden and Kalle Åström

Dept of Mathematics, Lund University
Box 118, S-221 00 Lund, Sweden
email: heyden@maths.lth.se, kalle@maths.lth.se

Abstract. In this paper we will investigate the different algebraic varieties and ideals that can be generated from multiple view geometry with uncalibrated cameras. The natural descriptor, $\mathcal{V}_n$, is the image of $\mathcal{P}^3$ in $\mathcal{P}^2 \times \mathcal{P}^2 \times \cdots \times \mathcal{P}^2$ under n different projections. However, we will show that $\mathcal{V}_n$ is not a variety.
Another descriptor, the variety $\mathcal{V}_b$, is generated by all bilinear forms between pairs of views and consists of all points in $\mathcal{P}^2 \times \mathcal{P}^2 \times \cdots \times \mathcal{P}^2$ where all bilinear forms vanish. Yet another descriptor, the variety, $\mathcal{V}_t$, is the variety generated by all trilinear forms between triplets of views. We will show that when $n = 3$, $\mathcal{V}_t$ is a reducible variety with one component corresponding to $\mathcal{V}_b$ and another corresponding to the trifocal plane. In ideal theoretic terms this is called a primary decomposition. This settles the discussion on the connection between the bilinearities and the trilinearities.
Furthermore, we will show that when $n = 3$, $\mathcal{V}_t$ is generated by the three bilinearities and one trilinearity and when $n \geq 4$, $\mathcal{V}_t$ is generated by the $\binom{n}{2}$ bilinearities. This shows that four images is the generic case in the algebraic setting, because $\mathcal{V}_t$ can be generated by just bilinearities.

1 Introduction

When estimating structure and motion from an uncalibrated sequence of images, the bilinear and the trilinear constraints play an important role, see [4], [5], [6], [7], [12], [13] and [14]. One difficulty encountered when using these multilinear constraints is that they are not independent, some of them may be calculated from the others, see [2], [3] and [6]. Thus there is a need to investigate the relations between them and to determine the minimal number of multilinear constraints that are needed to generate all multilinear constraints. These multilinear functions have not before been studied using an ideal theoretic approach.

The simplest multilinear constraint is the bilinear constraint, described by the fundamental matrix between two views. The next step is to consider three images at the same instant. At this stage the so called trilinear functions appear, see [12], [13], [5] [4] and [6]. The coefficients of the trilinearities are elements of the so called trifocal tensor.

* This work has been supported by the ESPRIT project BRA EP 6448, VIVA, and the Swedish Research Council for Engineering Sciences (TFR), project 95-64-222.

The obvious extension of the trilinear constraints is to consider four or more images at the same instant. It turns out that there exist quadrilinear constraints between four different views, see [2] and [14]. However, these constraints follows from the trilinear ones, cf. [6], [2] and [14]. It also became apparent that multilinear constraints between more than four views contain no new information.

One strange thing encountered with the bilinearities and trilinearities is that given the three bilinearities, corresponding to three different views, it is in general possible to calculate the camera matrices and the trilinearities from the components of the fundamental matrices, as described in [6] and [9]. But algebraically the trilinear constraints do not follow from the bilinear ones in the following sense. Consider points on the trifocal plane. The bilinear constraints impose only the condition that the three image points are on the trifocal lines, but the trilinear constraints impose one further condition. The question is now how the fact that it is possible to calculate the trilinear constraints from the bilinear ones, via the camera matrices, correspond to the fact that the trilinear constraints do not follow algebraically from the bilinear ones, that is the trilinearities do not belong to the ideal generated by the bilinearities. This is the key question we will try to answer in this paper. We will try to clarify the meaning of the statement 'the trilinear constraints follows from the bilinear ones, when the camera does not move on a line', where the statement is right or wrong depending of what kind of operations we are allowed to do on the bilinearities. The statement is true if we are allowed to pick out coefficients from the bilinearities and use them to calculate camera matrices and then the trilinearities, but the statement is wrong if we are just allowed to make algebraic manipulations of the bilinearities, where the image coordinates are considered as variables.

In order to understand the relations between the bilinearities and the trilinearities we have to use some algebraic geometry and commutative algebra. A general reference for the former is [10] and for the latter [11].

2 Problem Formulation

Consider the following problem: Given n images taken by uncalibrated cameras of a rigid object, describe the possible locations of corresponding points in the different images. Throughout this paper it is assumed that the views are generic, i.e. the focal points are in general position. Mathematically, this can be formulated as follows. Let $\mathcal{P}^2$ and $\mathcal{P}^3$ denote the projective spaces of dimension 2 and 3 respectively. Denote points in $\mathcal{P}^3$ by $\mathbf{X} = (X, Y, Z, W)$ and points in the i:th $\mathcal{P}^2$ by $\mathbf{x}_i = (x_i, y_i, z_i)$. Let A_i, $i = 1, \dots, n$ be projective transformations, that is linear transformations in projective coordinates, from $\mathcal{P}^3$ to $\mathcal{P}^2$,

$$A_i : \mathcal{P}^3 \ni \mathbf{X} \mapsto A_i\mathbf{X} \in \mathcal{P}^2, \quad i = 1, \dots, n \ . \tag{1}$$

In (1) each A_i is described by a 3×4 matrix or rank 3. The mapping is undefined on the nullspace of this matrix, corresponding to the **focal point**, f_i, of camera i, that is $A_i f_i = 0$. This can be regarded as one transformation,

$\Phi_n = (A_1, A_2, \ldots, A_n)$, from $\dot{\mathcal{P}}^3$ to $\mathcal{P}^2 \times \mathcal{P}^2 \times \cdots \times \mathcal{P}^2 = (\mathcal{P}^2)^n$,

$$\Phi_n : \dot{\mathcal{P}}^3 \ni \mathbf{X} \mapsto (A_1\mathbf{X}, A_2\mathbf{X}, \ldots, A_n\mathbf{X}) \in (\mathcal{P}^2)^n \ , \tag{2}$$

where $\dot{\mathcal{P}}^3 = \mathcal{P}^3 \setminus \{f_1, f_2, \ldots, f_n\}$, that is $\mathcal{P}^3$ with the camera centres omitted. This removal of a finite set of points from $\mathcal{P}^3$ gives a quasiprojective variety, i.e. an open subset of a projective variety. We want to describe the range of Φ_n as a subset of $(\mathcal{P}^2)^n$.

It can be shown, see [10], that $(\mathcal{P}^2)^n$ is indeed a projective variety. It can be embedded in $\mathcal{P}^{3^n-1}$ as a projective subvariety, using the Segre embedding. We will call this projective subvariety $\mathcal{S}_n$, and think of it as n copies of $\mathcal{P}^2$, and do not bother about the actual embedding. However, it is essential to know that $(\mathcal{P}^2)^n$ is indeed a projective variety.

Moreover, this fact has a very important implication on the functions and ideals generating varieties in $(\mathcal{P}^2)^n$. These functions must be homogeneous of the same degree in every triplet of variables corresponding to a factor $(\mathcal{P}^2)$, see [10], pp. 56. For example, there is no meaning in asking the question if the bilinear constraint between two images is contained in some ideal generating a variety in $(\mathcal{P}^2)^3$ for three images. The reason for this is that variables from the third image are not present in the bilinearity between the first two images. Thus this bilinear constraint is not homogeneous of the same degree in every triplet of variables. This difficulty will be overcome in the sequel by considering every multilinear constraint in its homogenised forms, and when we speak of generators of an ideal, describing a variety in $(\mathcal{P}^2)^n$, we implicitly assume that the generators are replaced by their homogenised equivalents.

2.1 Choice of Coordinates

Since we are only interested in algebraic relations between different ideals, we have the freedom to choose coordinates in $\mathcal{P}^3$ and in every $\mathcal{P}^2$ as we like. Consider the n projective transformations A_i in (1) and the n focal points $f_i \in \mathcal{P}^3$. We choose coordinates in $\mathcal{P}^3$ such that the first five points f_i constitute a projective basis with coordinates $f_1 = (0,0,0,-1)$, $f_2 = (1,0,0,-1)$, $f_3 = (0,1,0,-1)$, $f_4 = (0,0,1,-1)$ and $f_5 = (1,1,1,-1)$, where the minus sign in the fourth component will be convenient later. Furthermore we choose coordinates in each $\mathcal{P}^2$ such that the first three columns of each A_i are the columns of the identity matrix. This means that the projection matrices can be written

$$\begin{gathered} A_1 = \begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1&0&0&1 \\ 0&1&0&0 \\ 0&0&1&0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1&0&0&0 \\ 0&1&0&1 \\ 0&0&1&0 \end{bmatrix}, \\ A_4 = \begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&1 \end{bmatrix}, \quad A_5 = \begin{bmatrix} 1&0&0&1 \\ 0&1&0&1 \\ 0&0&1&1 \end{bmatrix}, \quad A_n = \begin{bmatrix} 1&0&0&a_n \\ 0&1&0&b_n \\ 0&0&1&c_n \end{bmatrix}, \quad n \geq 6 \ , \end{gathered} \tag{3}$$

where $A_n f_n = 0$ with $f_n = (-a_n, -b_n, -c_n, 1)$, $n \geq 6$. This coordinate system chosen in (3) will be called a **normalised coordinate system** for the multiple view geometry. This choice of coordinates can be done if the matrices A_i are assumed to be in general position.

The **epipole**, $e_{i,j}$, from camera j in image i is defined by $e_{i,j} = A_i f_j$. For example, with our choice of coordinates, $e_{1,2} = (1,0,0)$, $e_{1,3} = (0,1,0)$, $e_{2,1} = (1,0,0)$, $e_{2,3} = (1,-1,0)$, $e_{3,1} = (0,1,0)$ and $e_{3,2} = (1,-1,0)$. The **trifocal plane**, $TP_{i,j,k}$, for images i, j and k is the plane containing f_i, f_j and f_k. For example, with our choice of coordinates, $TP_{1,2,3}$ is described by $Z = 0$. The **epipolar line**, $EL_{i,j}$, is the line in $\mathcal{P}^3$ containing f_i and f_j. The **trifocal line**, $tl_{i,j,k}$, in image i from the triplet of images i, j and k is the intersection of the trifocal plane, $TP_{i,j,k}$, and image plane i. With our choice of coordinates, $tl_{1,2,3}$ is described by $z_1 = 0$, $tl_{2,1,3}$ by $z_2 = 0$ and $tl_{3,1,2}$ by $z_3 = 0$.

2.2 Multilinear Forms

Consider the equations, obtained from (1),

$$A_i \mathbf{X} = \lambda_i \mathbf{x}_i, \quad i = 1, \ldots n \ , \tag{4}$$

where the λ_i:s are needed because of the homogeneity of the coordinates. These equations can be written

$$Mu = \begin{bmatrix} A_1 & \mathbf{x}_1 & 0 & 0 & \ldots & 0 \\ A_2 & 0 & \mathbf{x}_2 & 0 & \ldots & 0 \\ A_3 & 0 & 0 & \mathbf{x}_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_n & 0 & 0 & 0 & \ldots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ -\lambda_1 \\ -\lambda_2 \\ -\lambda_3 \\ \vdots \\ -\lambda_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \ . \tag{5}$$

Since M has a nontrivial nullspace, it follows that

$$\operatorname{rank}(M) \leq 3 + n \ . \tag{6}$$

The matrix M contains one block with three rows for each image. The **bilinear constraints** for two images are obtained by taking a subdeterminant containing all three rows for the two images and the corresponding nonzero columns. The **trilinear constraints** for three images are obtained from subdeterminants containing three rows from one of the three images and two rows from each of the other two images and the corresponding nonzero columns. The **quadrilinear constraints** are obtained from subdeterminants containing two rows from each of the four images and the corresponding nonzero columns. Observe that all determinants of $(n+4) \times (n+4)$ submatrices are *multihomogeneous* of degree $(1,1,1,\ldots,1)$, that is of the same degree in every triplet of image coordinates. For example the bilinearity, $b_{1,2}$ between image 1 and 2 can be obtained as $x_3 x_4 \ldots x_n b_{1,2}$. This formulation is the same as the one used by Triggs, in [14] and is equivalent to the one used in [6].

2.3 The Varieties

In the sequel we are going to investigate the following subsets of $\mathcal{S}_n$:

Definition 1. The **natural descriptor**, $\mathcal{V}_\mathrm{n}$, is the range of Φ_n in (2), i.e. $\mathcal{V}_\mathrm{n} = \Phi_n(\dot{\mathcal{P}}^3) \subseteq S_n$. □

Definition 2. The **bilinear descriptor** or **bilinear variety** $\mathcal{V}_\mathrm{b}$, is defined as the projective subvariety in $\mathcal{S}_n$, generated by all bilinear constraints. □

Definition 3. The **trilinear descriptor** or **trilinear variety**, $\mathcal{V}_\mathrm{t}$, is defined as the projective subvariety in $\mathcal{S}_n$, generated by all trilinear constraints. □

These definitions raise several questions. It is obvious that $\mathcal{V}_\mathrm{b}$ and $\mathcal{V}_\mathrm{t}$ are projective subvarieties, since they are defined by homogeneous polynomials. $\mathcal{V}_\mathrm{n}$ is a constructible set, see [10], i.e. a rational image of a quasiprojective variety, but is it a variety?. How can these varieties be described as the set of zeros to an ideal of polynomials? Are they irreducible? If not, what are the irreducible components? What are the connections between them? We will answer these questions later.

It is also possible to generate a variety by combining the bilinear and trilinear forms. The projective subvariety, $\mathcal{V}_\mathrm{bt}$, in $\mathcal{S}_n$, generated by all bilinear and trilinear constraints, is called the **bitrilinear variety**. It follows that $\mathcal{V}_\mathrm{bt} = \mathcal{V}_\mathrm{b} \cap \mathcal{V}_\mathrm{t}$. When more than three images are available it is possible to generate a variety from the quadrilinear constraints. The projective subvariety, $\mathcal{V}_\mathrm{q}$, in $\mathcal{S}_n$, generated by all quadrilinear constraints, is called the **quadrilinear variety**. Again, it is obvious that $\mathcal{V}_\mathrm{bt}$ and $\mathcal{V}_\mathrm{q}$ are projective subvarieties and we can of course ask the same questions about connections and of irreducibility.

3 Two Images

Things start to be complicated already in the case of two images. Consider Φ_2 in (2) for $n = 2$. If we make a suitable restriction of Φ_2, we get the following well known theorem, see [1].

Theorem 4 (Fundamental theorem of epipolar geometry). *The mapping*

$$\tilde{\Phi}_2 : \mathcal{P}^3 \setminus \{EL_{1,2}\} \ni \mathbf{X} \mapsto (A_1\mathbf{X}, A_2\mathbf{X}) \in (\mathcal{P}^2 \setminus \{e_{1,2}\}) \times (\mathcal{P}^2 \setminus \{e_{2,1}\}) \qquad (7)$$

is a birational map between the quasiprojective varieties $\mathcal{P}^3 \setminus \{EL_{1,2}\}$ *and* $\big((\mathcal{P}^2 \setminus \{e_{1,2}\}) \times (\mathcal{P}^2 \setminus \{e_{2,1}\})\big) \cap \mathcal{V}_\mathrm{b}$.

Proof. $\big((\mathcal{P}^2 \setminus \{e_{1,2}\}) \times (\mathcal{P}^2 \setminus \{e_{2,1}\})\big) \cap \mathcal{V}_\mathrm{b}$ is a quasiprojective variety since it is the intersection of two quasiprojective varieties. Obviously, $\tilde{\Phi}_2$ is rational and surjective. It remains to prove that its inverse exists and is rational. This can be seen from the fact that given a point $(\mathbf{x}_1, \mathbf{x}_2)$ in the range of $\tilde{\Phi}_2$, it is possible to reconstruct the point in $\mathcal{P}^3 \setminus \{EL_{1,2}\}$ by intersecting the rays $A_1^{-1}(\mathbf{x}_1)$ and $A_2^{-1}(\mathbf{x}_2)$. The reconstructed point can be written as the intersection of these lines. This is clearly a rational map, which concludes the proof. □

The mapping, $\tilde{\Phi}_2$, in (7) is called the **birational restriction** of Φ_2. Note that the inverse image of Φ_2 is 1-dimensional at the point $(e_{1,2}, e_{2,1})$, because every point on the epipolar line projects to $(e_{1,2}, e_{2,1})$. This means that Φ_2 is not bijective between $\mathcal{V}_b$ and $\dot{\mathcal{P}}^3$.

We now turn to the natural descriptor, $\mathcal{V}_n \subseteq \mathcal{P}^2 \times \mathcal{P}^2$. It consists of the following pairs of points:

- one arbitrary point, (x_1, y_1, z_1), in the first $(\mathcal{P}^2 \setminus \{e_{1,2}\})$ and one point, (x_2, y_2, z_2) in the second $(\mathcal{P}^2 \setminus \{e_{2,1}\})$ on the line $b_{1,2} = 0$ (or vice versa),
- the epipole $e_{1,2}$ in the first $\mathcal{P}^2$ and the epipole $e_{2,1}$ in the second $\mathcal{P}^2$,

corresponding to images of points not on the epipolar line, and points on the epipolar line except the focal points.

Next, we turn to the bilinear variety, $\mathcal{V}_b$, generated by the bilinear forms. For two images there is just one bilinear form,

$$b_{1,2}(x_1, y_1, z_1, x_2, y_2, z_2) = y_1 z_2 - z_1 y_2 \ . \tag{8}$$

The projective subvariety $\mathcal{V}_b \subseteq \mathcal{P}^2 \times \mathcal{P}^2$ generated by $b_{1,2}$ consists of the following pairs of points:

- one arbitrary point, (x_1, y_1, z_1), in the first $(\mathcal{P}^2 \setminus \{e_{1,2}\})$ and one point, (x_2, y_2, z_2) in the second $(\mathcal{P}^2 \setminus \{e_{2,1}\})$ on the line $b_{1,2} = 0$ (or vice versa),
- the epipole $e_{1,2}$ in the first $\mathcal{P}^2$ and an arbitrary point in the second $\mathcal{P}^2$,
- the epipole $e_{2,1}$ in the second $\mathcal{P}^2$ and an arbitrary point in the first $\mathcal{P}^2$.

This includes also the point $(e_{1,2}, e_{2,1})$, consisting of the two epipoles. This shows the following theorem

Theorem 5. *For two images we have, with strict inclusion,* $\mathcal{V}_n \subset \mathcal{V}_b$.

We now return to the natural descriptor $\mathcal{V}_n \subseteq \mathcal{P}^2 \times \mathcal{P}^2$. Consider the ideal $\mathcal{I}_n = \mathcal{I}(\mathcal{V}_n) \subseteq \mathbb{R}[x_1, y_1, z_1, x_2, y_2, z_2]$, where $\mathcal{I}(\mathcal{V})$ denotes the ideal generated by all polynomial functions that vanish at all points in $\mathcal{V}$. Thus $\mathcal{I}_n$ is the ideal generated by the polynomial functions in $(x_1, y_1, z_1, x_2, y_2, z_2)$ that vanish at all points in $\mathcal{V}_n$. Since $\mathcal{V}_n$ is a subset of $\mathcal{P}^2 \times \mathcal{P}^2$, these functions must be bihomogeneous of the same degree in (x_1, y_1, z_1) and (x_2, y_2, z_2), see [10], pp. 56. The bilinearity in (8), of degree $(1,1)$, generates $\mathcal{I}_n$ according to the following lemma.

Lemma 6. *The ideal $\mathcal{I}_n$, generated by the natural descriptor, can be described as*

$$\mathcal{I}_n = \mathcal{I}_b = (b_{1,2}) \ .$$

Proof. Use (4) for $n = 2$ and eliminate the scale factors, λ_i, and object coordinates, $\mathbf{X}$. For details, see [8]. □

Remark. This means that the closure of $\mathcal{V}_n$ in the Zariski topology equals $\mathcal{V}_b$ since $\mathcal{V}(\mathcal{I}(\mathcal{V}_n))$ is the closure of $\mathcal{V}_n$. □

Theorem 7. *The natural descriptor $\mathcal{V}_n \in \mathcal{P}^2 \times \mathcal{P}^2$ is not a variety, in the sense that it can not be described as the common zeroes to a system of polynomial equations.*

Proof. Every variety, $\mathcal{V}$, fulfils $\mathcal{V} = \mathcal{V}(\mathcal{I}(\mathcal{V}))$. However, Theorem 5 and Lemma 6 show that $\mathcal{V}_n \subset \mathcal{V}_b = \mathcal{V}(\mathcal{I}(\mathcal{V}_n))$, with strict inclusion. □

Remark. We can also describe $\mathcal{V}_b$ as the range of an extension of the map Φ_2 to a multivalued map $\hat{\Phi}_2$ from the whole $\mathcal{P}^3$ defined as

$$\hat{\Phi}_2(\mathbf{X}) = \begin{cases} (A_1\mathbf{X}, A_2\mathbf{X}) & ; A_1\mathbf{X} \neq 0,\ A_2\mathbf{X} \neq 0 \\ \{(A_1\mathbf{X}, \mathbf{x}_2) \mid \mathbf{x}_2 \in \mathcal{P}^2\} & ; A_2\mathbf{X} = 0 \\ \{(\mathbf{x}_1, A_2\mathbf{X}) \mid \mathbf{x}_1 \in \mathcal{P}^2\} & ; A_1\mathbf{X} = 0\ . \end{cases} \tag{9}$$

Then the range of $\hat{\Phi}_2$ equals exactly the variety, $\mathcal{V}_b$, generated by the bilinear constraint. □

It is obvious that $\mathcal{V}_b$ is irreducible because it is generated by a single irreducible polynomial. Thus we have answered all questions raised above for the two image case.

4 Three Images

4.1 Varieties

Consider Φ_3 in (2) for $n = 3$. Making a suitable restriction of Φ_3, we get the following well known theorem.

Theorem 8 (Fundamental theorem of trifocal geometry). *The mapping*

$$\tilde{\Phi}_3 : \begin{cases} \mathcal{P}^3 \setminus \{TP_{1,2,3}\} \to (\mathcal{P}^2 \setminus \{tl_{1,2,3}\}) \times (\mathcal{P}^2 \setminus \{tl_{2,1,3}\}) \times (\mathcal{P}^2 \setminus \{tl_{3,1,2}\}) \\ \qquad \mathbf{X} \mapsto (A_1\mathbf{X}, A_2\mathbf{X}, A_3\mathbf{X}) \end{cases} \tag{10}$$

is a birational map between the quasiprojective varieties $\mathcal{P}^3 \setminus \{TP_{1,2,3}\}$ and $\big((\mathcal{P}^2 \setminus \{tl_{1,2,3}\}) \times (\mathcal{P}^2 \setminus \{tl_{2,1,3}\}) \times (\mathcal{P}^2 \setminus \{tl_{3,1,2}\})\big) \cap \mathcal{V}_b$.

Proof. In the same way as Theorem 7, see [8]. □

The mapping, $\tilde{\Phi}_3$, in (10) is called the **birational restriction** of Φ_3. It is always possible to reconstruct a point in $\big((\mathcal{P}^2 \setminus \{e_{1,2}, e_{1,3}\}) \times (\mathcal{P}^2 \setminus \{e_{2,1}, e_{2,3}\}) \times (\mathcal{P}^2 \setminus \{e_{3,1}, e_{3,2}\})\big) \cap \mathcal{V}_b$, which is in the range of Φ_3, using a rational map by intersecting two lines. However, we can not in advance choose a function for this. For example, if the point $\mathbf{X}$ is on the epipolar line between image 1 and 2 we can not use just A_1 and A_2 to make a reconstruction. We have to use A_3 also. This indicates why it is impossible to find an inverse rational map which would give a birational equivalence between $\mathcal{P}^3$ and $\mathcal{V}_n$.

The natural descriptor $\mathcal{V}_n$ consists of the following points:

- one arbitrary point in the first $(\mathcal{P}^2 \setminus \{tl_{1,2,3}\})$, one point in the second $(\mathcal{P}^2 \setminus \{tl_{2,1,3}\})$ on the line $b_{1,2} = 0$ and one point in the third $(\mathcal{P}^2 \setminus \{tl_{3,1,2}\})$ on the intersection between the lines $b_{1,3} = 0$ and $b_{2,3} = 0$ (and any permutation of the three images),
- one arbitrary point on $(tl_{1,2,3} \setminus \{e_{1,3}, e_{1,2}\})$ in the first $\mathcal{P}^2$, one arbitrary point on $(tl_{2,1,3} \setminus \{e_{2,3}, e_{2,1}\})$ in the second $\mathcal{P}^2$ and the unique point on $(tl_{3,1,2} \setminus \{e_{3,1}, e_{3,2}\})$ in the third $\mathcal{P}^2$ given by the trilinear constraints or as a projection of the reconstructed point from image 1 and image 2 onto the third image (and any permutation of the three images),
- the epipole $e_{1,3}$ in the first $\mathcal{P}^2$, the epipole $e_{2,3}$ in the second $\mathcal{P}^2$ and an arbitrary point on the trifocal line $tl_{3,1,2}$ in the third $(\mathcal{P}^2 \setminus \{e_{3,1}, e_{3,2}\})$ (and any permutation of the three images),

corresponding to images of points not on the trifocal plane, points in the trifocal plane, not on an epipolar line and points on the epipolar lines.

Consider the variety $\mathcal{V}_{\mathrm{t}} \subseteq \mathcal{P}^2 \times \mathcal{P}^2 \times \mathcal{P}^2$ generated by the trilinear forms. The projective subvariety in $\mathcal{P}^2 \times \mathcal{P}^2 \times \mathcal{P}^2$ corresponding to these forms is given by the same points as $\mathcal{V}_{\mathrm{n}}$ plus the following triplets of points:

- the epipole $e_{1,3}$ in the first $\mathcal{P}^2$, the epipole $e_{2,3}$ in the second $\mathcal{P}^2$ and an arbitrary point, (x_3, y_3, z_3), in the third $\mathcal{P}^2$ (and any permutation of the three images).

Remark. Just as in the case of two images we can also describe this variety, $\mathcal{V}_{\mathrm{t}}$ as the range of an extension of the map Φ_3 to a multivalued map $\hat{\Phi}_3$ from the whole $\mathcal{P}^3$ defined as the obvious extension of (9). Then the range of $\hat{\Phi}_3$ equals exactly the variety, $\mathcal{V}_{\mathrm{t}}$, generated by the trilinear constraints. □

We now turn to the variety, $\mathcal{V}_{\mathrm{b}}$, generated by the bilinear forms. In this case each image pair contributes with a bilinear form,

$$b_{1,2} = y_1 z_2 - z_1 y_2, \; b_{1,3} = x_1 z_3 - z_1 x_3, \; b_{2,3} = (x_2 + y_2) z_3 - z_2 (x_3 + y_3) \; . \quad (11)$$

The projective subvariety $\mathcal{V}_{\mathrm{b}} \subseteq \mathcal{P}^2 \times \mathcal{P}^2 \times \mathcal{P}^2$ generated by these forms consists of the same triplets of points as $\mathcal{V}_{\mathrm{t}}$ plus the following triplets of points:

- one arbitrary point on $tl_{1,2,3}$ in the first $\mathcal{P}^2$, one arbitrary point on $tl_{2,1,3}$ in the second $\mathcal{P}^2$ and one arbitrary point on $tl_{3,1,2}$ in the third $\mathcal{P}^2$.

Theorem 9. *For three images we have, with strict inclusions,*

$$\mathcal{V}_{\mathrm{n}} \subset \mathcal{V}_{\mathrm{t}} \subset \mathcal{V}_{\mathrm{b}} \; . \quad (12)$$

Furthermore, the variety $\mathcal{V}_{\mathrm{b}}$ is reducible and can be written as a union of two irreducible varieties as

$$\mathcal{V}_{\mathrm{b}} = \mathcal{V}_{\mathrm{t}} \cup \mathcal{V}_{\mathrm{tp}} \; , \quad (13)$$

where $\mathcal{V}_{\mathrm{tp}}$ is the variety containing one point on each trifocal line.

One consequence of this theorem is that the bitrilinear variety, $\mathcal{V}_{\mathrm{bt}}$, and the trilinear variety, $\mathcal{V}_t$, coincide, i.e. $\mathcal{V}_{\mathrm{bt}} = \mathcal{V}_t$. In fact, it follows from Theorem 9 that $\mathcal{V}_t \subset \mathcal{V}_b$. Since $\mathcal{V}_{\mathrm{bt}} = \mathcal{V}_t \cap \mathcal{V}_b$, from the definition of $\mathcal{V}_{\mathrm{bt}}$, it follows that $\mathcal{V}_{\mathrm{bt}} = \mathcal{V}_t$.

We now return to the natural descriptor $\mathcal{V}_{\mathrm{n}} \subseteq \mathcal{P}^2 \times \mathcal{P}^2 \times \mathcal{P}^2$. Consider the ideal $\mathcal{I}_{\mathrm{n}} = \mathcal{I}(\mathcal{V}_{\mathrm{n}}) \subseteq \mathbb{R}[x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3]$. Since $\mathcal{V}_{\mathrm{n}}$ is a subset of $\mathcal{P}^2 \times \mathcal{P}^2 \times \mathcal{P}^2$, the functions in $\mathcal{I}_{\mathrm{n}}$ must be trihomogeneous in (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3), see [10], pp. 56. The bilinearities in (11) are functions of degree $(1,1,0)$, $(1,0,1)$ and $(0,1,1)$, which can be extended to functions of degree $(1,1,1)$ as described above. There are also trilinear functions of degree $(1,1,1)$. We have the following lemma, which states that the closure of $\mathcal{V}_{\mathrm{n}}$ is $\mathcal{V}_{\mathrm{t}}$, and theorem (for the proofs, see [8]).

Lemma 10. *The ideal defined by the natural descriptor, can be described as*

$$\mathcal{I}_{\mathrm{n}} = \mathcal{I}(\mathcal{V}_{\mathrm{t}}) = \mathcal{I}_{\mathrm{t}} \ .$$

Theorem 11. *The natural descriptor $\mathcal{V}_{\mathrm{n}} \subseteq \mathcal{P}^2 \times \mathcal{P}^2 \times \mathcal{P}^2$ is not a variety, in the sense that it can not be described as the common zeroes to a system of polynomial equations.*

4.2 Ideals

All multilinear constraints are obtained from (5) with $n = 3$. The trilinearities are obtained as subdeterminants involving at least two rows from each image, for example

$$\begin{aligned} t_{5,7} &= y_1 z_2 z_3 + z_1 x_2 z_3 - x_1 z_2 z_3 - z_1 z_2 y_3 \\ t_{5,9} &= x_1 z_2 x_3 + x_1 z_2 y_3 - y_1 z_2 x_3 - z_1 x_2 x_3 \\ t_{6,7} &= y_1 x_2 z_3 + y_1 y_2 z_3 - x_1 y_2 z_3 - z_1 y_2 y_3 \\ t_{6,9} &= x_1 y_2 x_3 + x_1 y_2 y_3 - y_1 x_2 x_3 - y_1 y_2 x_3 \ , \end{aligned} \tag{14}$$

where $t_{i,j}$ denotes the subdeterminant obtained after removing rows i and j.

We will now describe the relations between the ideals generated by the trilinearities and the bilinearities. The ideal, $\mathcal{I}_{\mathrm{b}}$, in $\mathbb{R}[x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3]$, generated by the bilinearities in (11), is called the **bilinear ideal**. In the same way, the ideal, $\mathcal{I}_{\mathrm{t}}$, generated by the trilinearities, is called the **trilinear ideal**. Finally, the ideal, $\mathcal{I}_{\mathrm{bt}}$, generated by the bilinearities and the trilinearities, is called the **bitrilinear ideal**.

First we are going to study different ways of generating $\mathcal{I}_{\mathrm{b}}$ and $\mathcal{I}_{\mathrm{t}}$. It is obvious that $\mathcal{I}_{\mathrm{b}}$ is generated by the 3 bilinearities, but that no 2 of them are sufficient to generate $\mathcal{I}_{\mathrm{b}}$. Things are more complicated for $\mathcal{I}_{\mathrm{t}}$. Although the trilinear constraint, locally, can be written as the vanishing of 3 trilinear forms among all trilinearities we need 4 forms to generate $\mathcal{I}_{\mathrm{t}}$. Consider first the following simple example, which reveals the difference between a minimal generating set and the codimension of the corresponding variety.

Example 1. The condition that two vectors, $u = (a, b, c)$ and $v = (d, e, f)$, in $\mathbb{R}^3$ are parallel can be written rank $\left[\begin{smallmatrix} a & b & c \\ d & e & f \end{smallmatrix}\right] < 2$, which is equivalent to $p_1 = p_2 = p_3 = 0$, where $p_1 = ae - bd$, $p_2 = bf - ce$ and $p_3 = cd - af$. Introduce the ideal $\mathcal{I}_{\text{ex}} = (p_1, p_2, p_3) \subseteq \mathbb{R}[a, b, c, d, e, f]$. The codimension of the variety $\mathcal{V}(\mathcal{I}_{\text{ex}})$ is 2, since the rank condition above can locally be obtained from 2 polynomial equations. This can be seen from $fp_1 + dp_2 + ep_3 = 0$. However, it is not possible to generate the ideal (p_1, p_2, p_3) by any two of the polynomials, p_1 and p_2, for example, because $u = (1, 0, 1)$ and $v = (1, 0, 2)$ obeys both $p_1 = 0$ and $p_2 = 0$ but $p_3 = -1$. This means that the codimension of the variety $\mathcal{V}(\mathcal{I}_{\text{ex}})$ is 2 and $\{p_1, p_2, p_3\}$ is a minimal generating set for $\mathcal{I}_{\text{ex}}$. □

Theorem 12. *The ideal $\mathcal{I}_{\text{t}}$ can be generated by the bilinearities and one trilinearity, $t_{6,9}$. $\mathcal{I}_{\text{t}}$ can not be generated by three multilinear functions.*

Proof. Using Gröbner basis calculations, see [8]. □

Remark. Using our choice of coordinates, the trilinearity needed apart from the bilinearities can be any of $t_{3,6}$, $t_{3,9}$ or $t_{6,9}$ (they are in fact the same polynomial). Using an arbitrary coordinate system, it is in general possible to choose an arbitrary non-degenerate trilinearity. The condition that must be fulfilled is that the trilinearity does not vanish on the trifocal lines. □

We are now ready to prove our key result describing the relations between $\mathcal{I}_{\text{b}}$ and $\mathcal{I}_{\text{t}}$. First observe that if $z_1 = z_2 = z_3 = 0$, then all bilinearities in (11) vanish but the trilinearity $t_{6,9}$ does not vanish. This means that the trilinear constraint $t_{6,9} = 0$ imposes one further condition on the other image coordinates. The conditions $z_1 = z_2 = z_3 = 0$ describe the intersection of the trifocal plane with the three images, which indicates that it could correspond to an associate prime ideal of $\mathcal{I}_{\text{b}}$. For the proof of the following theorem, see [8].

Theorem 13 (Primary decomposition of the bilinear ideal). *The ideal $\mathcal{I}_{\text{b}}$ is reducible and can be decomposed as*

$$\mathcal{I}_{\text{b}} = \mathcal{I}_{\text{t}} \cap \mathcal{I}_{\text{tp}} \ , \tag{15}$$

where $\mathcal{I}_{\text{tp}} = (z_1, z_2, z_3)$ is the ideal corresponding to the trifocal plane. In (15), $\mathcal{I}_{\text{t}}$ and $\mathcal{I}_{\text{tp}}$ are prime ideals and thus irreducible.

This theorem shows that the trilinear ideal can be obtained from the bilinear ideal in the following way. First make a primary decomposition of the bilinear ideal. This gives two unique primary ideals. Then throw away the ideal that can be generated by linear functions. The remaining one is the trilinear ideal. It follows that the ideal $\mathcal{I}_{\text{bt}}$ generated by the bilinearities and the trilinearities is the same as the ideal $\mathcal{I}_t$ generated by the trilinearities, i.e. $\mathcal{I}_{\text{bt}} = \mathcal{I}_{\text{t}}$.

We conclude this section with the observation that the dimension of the varieties $\mathcal{V}_{\text{t}}$ and $\mathcal{V}_{\text{b}}$ is $3 = 9-3-3$. The number of variables is 9, they are divided into 3 groups of projective vectors and the constraints can locally be written as the vanishing of 3 polynomial equations. This means that the codimension is 3. Thus we would like to have 3 polynomials to generate the variety $\mathcal{V}_{\text{t}}$, unfortunately this is not possible according to Theorem 12.

5 More than Three Images

Because of lack of space we only give the results here, for proofs see [8].

Theorem 14. *For n images, $n \geq 4$, we have with strict inclusion,*

$$\mathcal{V}_{\mathrm{n}} \subset \mathcal{V}_{\mathrm{b}} = \mathcal{V}_{\mathrm{t}} = \mathcal{V}_{\mathrm{q}} \ . \tag{16}$$

Geometrically this can be seen as follows. When we have three images the bilinear constraints fail to distinguish between correct and incorrect point correspondences on the trifocal lines, but when we have another image outside the trifocal plane it is possible to resolve this failure by using the three new bilinear constraints involving the fourth image. Again the closure of $\mathcal{V}_{\mathrm{n}}$ is $\mathcal{V}_{\mathrm{t}}$. Observe that the bilinearities are sufficient to generate the closure of $\mathcal{V}_{\mathrm{n}}$, that is no trilinearities are needed.

Theorem 15. *For n images, $n \geq 4$, we have*

$$\mathcal{I}_{\mathrm{b}} = \mathcal{I}_{\mathrm{t}} = \mathcal{I}_{\mathrm{q}} = \mathcal{I}(\mathcal{V}_{\mathrm{n}}) \ . \tag{17}$$

In the case of 4 images we have 6 bilinearities and

Theorem 16. *The ideal $\mathcal{I}_{\mathrm{b}}$, for 4 images, can be generated by all 6 bilinearities but not by any 5 multilinear functions.*

In the case of 5 images we have 10 bilinearities and since the codimension of $\mathcal{V}_{\mathrm{b}}$ is 3, it would be nice to have $2 * 5 - 3 = 7$ bilinear forms generating $\mathcal{I}_{\mathrm{b}}$. However, this is not sufficient and neither is 8 bilinear forms. In fact, only 1 bilinear form can be removed.

Theorem 17. *The bilinear ideal $\mathcal{I}_{\mathrm{b}}$ for 5 images can be generated by 9 bilinear forms, but not by any 8 multilinear functions.*

Conjecture 18. *It is not possible to generate $\mathcal{I}_{\mathrm{t}}$ for $n \geq 3$ images by $2n - 3$ bilinearities or by $2n - 3$ other multilinear functions.*

6 Conclusions

In this paper we have shown that the image of $\mathcal{P}^3$ in $\mathcal{P}^2 \times \mathcal{P}^2 \times \cdots \times \mathcal{P}^2$ under an n-tuple of projections $\Phi_n = (A_1, A_2, \ldots, A_n)$ is not an algebraic variety, i.e. it can not be described as the set of common zeros to a system of polynomial equations. We have described two different approaches to obtain an algebraic variety. The first one is to extend Φ_n to a multivalued map, defining the image of the focal point f_i of camera i to be the set of points corresponding to the actual epipoles in the other images and an arbitrary point in image i. The second one is to restrict Φ_n by removing an epipolar line or a trifocal plane and the corresponding image points. Moreover, the closure of this image is equal to the trilinear variety.

We have shown that for three images the variety defined by the bilinearities is reducible and can be written as a union of two irreducible varieties; the variety defined by the trilinearities and a variety corresponding to the trifocal lines. For the ideals the situation can be described by saying that the ideal generated by the bilinearities can be written in a primary decomposition as an intersection of two prime ideals; the ideal generated by the trilinearities and an ideal corresponding to the trifocal lines.

Finally, if four or more images are available the ideal generated by the bilinearities is the same as the ideal generated by the trilinearities. This means that it is possible to use only bilinearities to generate the algebraic variety defined by all multilinear forms. We have also shown that the ideal generated by the quadrilinearities is the same as the ideal generated by the trilinearities.

Acknowledgements

The authors would like to thank Gunnar Sparr and Sven Spanne for guidance and support and the reviewers for valuable suggestions.

References

1. Faugeras, O., D., What can be seen in three dimensions with an uncalibrated stereo rig?, *ECCV'92, Lecture notes in Computer Science, Vol 588. Ed. G. Sandini, Springer-Verlag,* 1992, pp. 563-578.
2. Faugeras, O., D., Mourrain, B., On the geometry and algebra on the point and line correspondences between N images, *Proc. ICCV'95, IEEE Computer Society Press,* 1995, pp. 951-956.
3. Faugeras, O., D., Mourrain, B., About the correspondences of points between N images, *Proc. IEEE Workshop on Representation of Visual Scenes,* 1995.
4. Hartley, R., I., Projective Reconstruction and Invariants from Multiple Images, *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 16, no. 10, pp. 1036-1041, 1994.
5. Hartley, A linear method for reconstruction from lines and points, *Proc. ICCV'95, IEEE Computer Society Press,* 1995, pp. 882-887.
6. Heyden, A., Reconstruction from Image Sequences by means of Relative Depths, *Proc. ICCV'95, IEEE Computer Society Press,* 1995, pp. 1058-1063, Also to appear in *IJCV, International Journal of Computer Vision.*
7. Heyden, A., Åström, K., A Canonical Framework for Sequences of Images, *Proc. IEEE Workshop on Representation of Visual Scenes,* 1995.
8. Heyden, A., Åström, K., Algebraic Properties of Multilinear Constraints, *Technical Report, CODEN: LUFTD2/TFMA--96/7001--SE, Lund, Sweden,* 1996.
9. Luong, Q.-T., Vieville, T., Canonic Representations for the Geometries of Multiple Projective Views, *ECCV'94, Lecture notes in Computer Science, Vol 800. Ed. Jan-Olof Eklund, Springer-Verlag,* 1994, pp. 589-599.
10. Schafarevich, I., R., *Basic Algebraic Geometry I - Varieties in Projective Space,* Springer Verlag, 1988.
11. Sharp, R., Y., *Steps in Commutative Algebra,* London Mathematical Society Texts, 1990.
12. Shashua, A., Trilinearity in Visual Recognition by Alignment, *ECCV'94, Lecture notes in Computer Science, Vol 800. Ed. Jan-Olof Eklund, Springer-Verlag,* 1994, pp. 479-484.
13. Shahsua, A., Werman, M., Trilinearity of Three Perspective Views and its Associated Tensor, *Proc. ICCV'95, IEEE Computer Society Press,* 1995, pp. 920-925.
14. Triggs, B., Matching Constraints and the Joint Image, *Proc. ICCV'95, IEEE Computer Society Press,* 1995, pp. 338-343.