The joint image handbook

Supplementary material

Matthew Trager[*] — Inria
Martial Hebert — Carnegie Mellon University
Jean Ponce[*] — École Normale Supérieure / PSL Research University

Contents

This supplementary material collects some technical proofs that were only sketched in the main body of the paper, and clarifies a few aspects that were not treated in full detail due to space limitations. For convenience of the reader, we have also provided in Section A a brief introduction to elementary notions of algebraic geometry, that could be useful for a better understanding of the theory underlying our results. Propositions that appear in the main body of the paper will be stated here with the same numbering, while auxiliary results will be numbered independently.

A. Algebraic varieties and ideals

We give a brief and informal introduction to some notions of algebraic geometry. Proofs and more details can be found in any basic reference on the subject. For example, see [3] for an accessible presentation with a focus on computational aspects.

Affine and projective algebraic varieties. Let $\mathbb{K}$ be a field; for our purposes, we will only be interested in $\mathbb{K} = \mathbb{C}$ (complex numbers) or $\mathbb{K} = \mathbb{R}$ (real numbers). Let $\mathbb{K}[x_1, \ldots, x_n]$ denote the ring of polynomials in n variables with coefficients in $\mathbb{K}$. An *affine algebraic variety* is defined as the zero-locus of a family of polynomial equations:

$$V(S) = \{(s_1, \ldots, s_n) \mid f(s_1, \ldots, s_n) = 0, \ \forall f \in S\} \subseteq \mathbb{K}^n \tag{1}$$

where $S \subseteq \mathbb{K}[x_1, \ldots, x_n]$.[1] We will say that such a variety $V(S)$ is *defined* by the set S. It is actually customary to consider only special families of polynomials called *ideals*: these are sets that form additive groups, and that are closed under multiplication with any other polynomial (so that if f belongs to an ideal $\mathfrak{i}$, then the product gf is in $\mathfrak{i}$ for *any* polynomial g). For example, the set

$$\langle f, g \rangle = \{h_1 f + h_2 g \mid h_1, h_2 \in \mathbb{K}[x_1, \ldots, x_n]\} \tag{2}$$

is the ideal *generated* by f and g, *i.e.*, the smallest ideal that contains both f and g. It is quite easy to realize that the affine variety defined by a set of polynomials S coincides with the variety defined by the ideal generated by S, so we can always assume that a variety is defined by an ideal. The association between algebraic varieties and polynomial ideals is the first brick in the foundation of algebraic geometry. We collect some fundamental properties of this correspondence:

- Every polynomial ideal can be generated by a *finite* number of polynomials, so varieties are always defined by finite sets of equations.
- Different ideals can define the same variety V. To clarify this behavior, we need to mention that any ideal $\mathfrak{i}$ is contained in its so-called *radical* $\sqrt{\mathfrak{i}}$ (defined as the set of polynomials f for which there exists a $m \geq 1$ such that $f^m \in \mathfrak{i}$). Now, if two ideals $\mathfrak{i}_1$ and $\mathfrak{i}_2$ have the same radical $\sqrt{\mathfrak{i}_1} = \sqrt{\mathfrak{i}_2}$, then they define the same variety $V(\mathfrak{i}_1) = V(\mathfrak{i}_2)$. If $\mathbb{K}$ is an algebraically closed field, then the converse also holds. This important result is known as *Hilbert's Nullstellensatz.*
- The largest ideal defining an algebraic variety V, containing *all* polynomials which vanish on V, is denoted with $\mathfrak{i}(V)$. This ideal is a always a *radical ideal*, *i.e.*, it coincides with its radical $\mathfrak{i}(V) = \sqrt{\mathfrak{i}(V)}$. In summary, the association between varieties and ideals can be descrbed as follows:

$$\begin{array}{ccc} \text{affine varieties} & & \text{radical ideals} \\ V & \xrightarrow{\mathfrak{i}} & \mathfrak{i}(V) \\ V(\mathfrak{i}) & \xleftarrow{V} & \mathfrak{i} \end{array}$$

If $\mathbb{K}$ is an algebraically closed field, then these maps are bijections and inverses of each other (however the map "$\mathfrak{i}$" is actually injective over any field).

[*] Willow project team. DI/ENS, ENS/CNRS/Inria UMR 8548.

[1] For some authors this is actually definition of an affine *algebraic set*, since they require algebraic varieties to be also irreducible (see the definition in the following paragraph).

In the paper, we usually deal with subsets of projective space, so we now introduce the notion of *projective algebraic varieties*.[2] A projective variety is defined as a subset of $\mathbb{P}^n(\mathbb{K})$ (n-dimensional projective space) that is the zero-set of family of *homogeneous* polynomials:

$$W(S) = \{[s_0; s_1; \ldots; s_n] \mid f(s_0, \ldots, s_n) = 0 \;\; \forall f \in S\} \subseteq \mathbb{P}^n(\mathbb{K}) \tag{3}$$

where $S \subseteq \mathbb{K}[x_0, x_1, \ldots, x_n]$ is a set of homogeneous polynomials. We recall that a homogenous polynomial has terms of the same degree d, and has the property that $f(\lambda s_0, \ldots, \lambda s_n) = \lambda^d f(s_0, \ldots, s_n)$ for all $\lambda \in \mathbb{K}$: note that this condition is necessary in order for the evaluation on homogeneous coordinates to be well-defined. There exists a correspondence between projective varieties and homogeneous radical ideals (radical ideals generated by sets of homogenous polynomials) that has some technical complications but is conceptually completely analogous to the affine case. Finally, we point out that the intersection of projective variety $W \subseteq \mathbb{P}^n(\mathbb{K})$ with an affine chart (say $U_0 = \{x_0 = 1\}$) is an affine variety, obtained simply by substituting $x_0 = 1$ for all polynomials defining W; conversely, an affine variety $V \subseteq \mathbb{K}^n$ can always be viewed as the affine portion of its *projective closure* $\overline{V} \subseteq \mathbb{P}^n(\mathbb{K})$, that is obtained by *homogenizing* all polynomials defining V.

Example 1. Consider the affine variety $V \subseteq \mathbb{R}^2$ defined by the following polynomial conditions:

$$\begin{cases} f_1 : x^2 + x = 0 \\ f_2 : xy + y = 0 \end{cases} \tag{4}$$

One can prove that the projective closure of V is the projective variety $\overline{V} \subseteq \mathbb{P}^2(\mathbb{R})$ defined by the homogenized generators:

$$\begin{cases} f_1^h : x^2 + xz = 0 \\ f_2^h : xy + yz = 0 \end{cases} \tag{5}$$

If we restrict to the affine chart $\{z = 1\}$ we see that $\overline{V}$ coincides with V, but $\overline{V}$ also contains one point "at infinity": $\overline{V} \cap \{z = 0\} = \{[0 : 1 : 0]\}$.

Irreducible decomposition, dimension. An algebraic variety (affine or projective) is said to be *irreducible* if it is not the union of two proper algebraic varieties. A fundamental result states that every algebraic variety can be expressed as the union of a finite number of irreducible components, and moreover this decomposition is unique. Given a decomposition of a variety V into irreducible components

$$V = W_1 \cup \ldots \cup W_k \tag{6}$$

we obtain an associated decomposition of ideals (notice the union now becomes an intersection):

$$\mathfrak{i}(V) = \mathfrak{i}(W_1) \cap \ldots \cap \mathfrak{i}(W_k). \tag{7}$$

Ideals of the form $\mathfrak{i}(W)$ where W is irreducible can be characterized in purely algebraic terms, and are called *prime ideals*. The reason for this terminology is that ideals were first introduced as a generalization of the integers ("ideal numbers"), and the decomposition (7) generalizes the factorization of integers. However, a general non-radical ideal cannot be expressed as an intersection of primes, and requires a more elaborate decomposition based on so-called *primary ideals* ("primary decomposition"); this decomposition is also not completely unique. Since for our purposes we will only be interested in radical ideals associated to varieties, we can actually always find a decomposition into primes, and the decomposition is unique (basically because it corresponds to an irreducible decomposition of varieties, as illustrated above).

If V is an irreducible variety, we can define its *dimension* to be the maximal length d of chains $V_0 \subset \ldots \subset V_d = V$ of irreducible varieties contained in V. There are also a number of equivalent algebraic definitions (e.g., the longest chain of prime ideals ideals containing $\mathfrak{i}(V)$).

Example 2. If we consider again the variety $V \subseteq \mathbb{R}^2$ defined by (4), one can show that $V = W_1 \cup W_2$ where

$$\begin{aligned} W_1 &= \{x = 0, y = 0\} \\ W_2 &= \{x + 1 = 0\} \end{aligned} \tag{8}$$

See Figure 1. Note that W_1 is a point (a 0-dimensional irreducible variety) and W_2 is a line (a 1-dimensional irreducible variety).

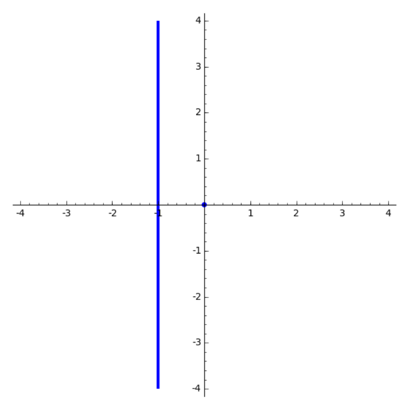

Figure 1. The affine variety V from Examples 1 and 2.

[2]The joint image is actually a subset of a *product* of projective spaces $\mathbb{P}^2 \times \ldots \times \mathbb{P}^2$. However, we can still talk about projective varieties since one can show that a product of projective spaces can always be embedded in a higher dimensional projective space $\mathbb{P}^N$ (using the so-called *Segre embbedding*). One can also work directly in $\mathbb{P}^2 \times \ldots \times \mathbb{P}^2$, with a few technical adjustments.

Gröbner bases. The properties of ideals discussed above do not provide us with a practical ways for dealing with algebraic varieties computationally. For example, how can we verify if a polynomial g is contained in a given ideal $\mathfrak{i} = \langle f_1, \ldots, f_s \rangle$ ("ideal membership problem")? More generally, how can we establish whether two ideals $\mathfrak{i}_1$ and $\mathfrak{i}_2$ coincide ("ideal equality problem")? The basic tool for dealing with these kinds of practical problems is given by *Gröbner bases*. In order to define a Gröbner basis we have to first fix a total *monomial ordering* (*e.g.*, lexicographic ordering or graded lexicographic ordering), so that for every polynomial we can define a leading term. For any ideal $\mathfrak{i} \subseteq \mathbb{K}[x_1, \ldots, x_n]$, we can consider the associated *leading term ideal* $LT(\mathfrak{i})$, generated by all the leading terms of elements of $\mathfrak{i}$. Now we can define a Gröbner basis for $\mathfrak{i}$ as a special set $f_1, \ldots, f_n$ of generators of $\mathfrak{i}$ whose leading terms $LT(f_1), \ldots, LT(f_n)$ are generators for the leading term ideal $LT(\mathfrak{i})$. Some important properties of Gröbner bases are:

- A Gröbner basis can be computed from any set of generators of an ideal (*Buchberger's algorithm*).
- Gröbner bases provide a practical *division algorithm*, that computes the division of a polynomial f by a set of polynomials $f_1, \ldots, f_s$. This results in an expression of the form $f = q_1 f_1 + \ldots q_s f_s + r$ for some remainder polynomial r, and provides a simple solution to the ideal membership problem (an element f belongs to an ideal $\mathfrak{i} = \langle f_1, \ldots, f_s \rangle$ if and only if its remainder under division is zero).
- For a fixed monomial ordering, every ideal has a unique *reduced* Gröbner basis, where "reduced" essentially means that each generator does not contain any monomial that is the leading term of some other generator. This allows to easily verify when two ideals are the same.

Gröbner bases are also very important for solving systems of polynomial equations (*elimination method*), and for computing the dimension of ideals. Moreover, Gröbner bases have been used to design algorithms for computing the radical of an ideal, and for recovering primary decompositions [5]. These methods are implemented in computer algebra systems such as Macaulay2 [6], Singular [4] (used for our computations), or Sage [16] (which actually contains Singular).

Example 3. Considering the lexicographic ordering $x \succ y$, one can show that any element of the form $a(x,y)f_1 + b(x,y)f_2$, lying in the ideal generated by f_1 and f_2, will have a leading term that is either multiple of $LT(f_1) = x^2$ or of $LT(f_2) = xy$ (basically one needs to prove that no cancellation can occur): this means f_1, f_2 form a Gröbner basis. However, if we had considered for example the polynomials f_1 together with

$$\tilde{f}_2 = xy - y^2 \tag{9}$$

then a simple computation shows that

$$\tilde{f}_3 = y f_1 - (x - y + 1) f_2 = y^3 - y^2 \tag{10}$$

whose leading term y^3 cannot be generated using $LT(f_1) = x^3, LT(\tilde{f}_2) = x^2 y$. A basic evaluation using Singular shows that in fact $f_1, \tilde{f}_2$ together with $\tilde{f}_3$ actually form a Gröbner basis.

B. Proofs and technical material

B1. The closure of the joint image

We prove Proposition 1 from the main part of the paper, which describes the relationship between the joint image $\mathcal{I}_n$ and its closure $\overline{\mathcal{I}}_n$. We also argue that the set $\mathcal{C}_n = \overline{\mathcal{I}}_n \setminus \mathcal{I}_n$ is a distinguishable set.

Proposition 1. *Given $n \geq 3$ cameras with non-collinear (distinct) pinholes, one has $\mathcal{I}_n = \overline{\mathcal{I}}_n \setminus \mathcal{C}_n$, where*

$$\mathcal{C}_n = \bigcup_{i=1}^{n} (\boldsymbol{e}_{i1} \times \ldots \times \mathbb{P}^2_{(i)} \times \ldots \times \boldsymbol{e}_{in}). \tag{11}$$

Here $\mathbb{P}^2_{(i)}$ indicates $\mathbb{P}^2$ at position i in the product, and $\boldsymbol{e}_{ij}$ denotes the epipole in image j relative to image i. If $n = 2$, or more generally if the cameras have collinear pinholes, then one must remove from $\mathcal{C}_n$ the n-tuple of epipoles $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ (in this case there is only one epipole in each image).

Proof. In the main part of the paper we have introduced the matrix

$$\mathcal{U}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = \begin{pmatrix} \mathcal{M}_1 & \boldsymbol{u}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathcal{M}_2 & \mathbf{0} & \boldsymbol{u}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{M}_n & \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{u}_n \end{pmatrix} \tag{12}$$

and pointed out that joint image variety can be set-theoretically characterized by:

$$\overline{\mathcal{I}}_n = \{(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \in (\mathbb{P}^2)^n \mid \mathcal{U}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \text{ is not full rank}\}.$$

See for example [1, 11]. If the cameras are not all collinear, the nullspace of $\mathcal{U}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ has dimension at most one, and $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ belongs to $\mathcal{C}_n = \overline{\mathcal{I}}_n \setminus \mathcal{I}_n$ if and only if annihilating vectors $[\boldsymbol{p}; \lambda_1; \ldots; \lambda_n] \in \mathbb{R}^{4+n}$ are such that $\lambda_i = 0$ for some i: in this case $\boldsymbol{p}$ must be one of the camera centers, and we see that $\mathcal{C}_n$ is given by (11). If the cameras are collinear, then the same reasoning applies except when $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ (because the nullspace of $\mathcal{U}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ has dimension two): in this case, it is easy to realize that $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) \in \mathcal{I}_n$. $\square$

Proposition b1. *The set $\mathcal{C}_n$ defined in the previous proposition is a distinguished set in $\overline{\mathcal{I}}_n$,* i.e., *it can be computed from any algebraic characterization of $\overline{\mathcal{I}}_n$.*

Proof. For $n = 2$, the epipoles are distinguished points since $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ is the singular locus of $\overline{\mathcal{I}}_2$ (see Proposition 3); otherwise, more explicitly, they can be computed from the epipolar constraint as the null-spaces of the associated fundamental matrix. For any number of views, it is sufficient to note that the "projection" of $\overline{\mathcal{I}}_n(\mathcal{M}_1, \dots, \mathcal{M}_n)$ onto a chosen copy of $\mathbb{P}^2 \times \mathbb{P}^2$ is exactly the joint image of the corresponding views (and, algebraically, the epipolar constraint can be obtained via variable elimination). □

B2. The joint image and multilinear forms

We now give a proof of Proposition 3. The result is implied by Theorem 3.6 in [1], but the case of multilinear constraints (multidegree $(1, \dots, 1)$) is not pointed out. Our proof is also adapted from the one given in [1], but we use more direct argument to show Proposition b3.

Proposition 3. *Given n cameras $\mathcal{M}_1, \dots, \mathcal{M}_n$, the multilinear constraints which vanish on $\mathcal{V}_n$ form a vector space of dimension $d_n = 3^n - \binom{n+3}{3} + n$.*

Remark. In the rest of this section, we will assume that the camera matrices $\mathcal{M}_1, \dots, \mathcal{M}_n$ satisfy a generality assumption, namely that all 4×4 matrices formed by any four distinct rows from $\mathcal{M}_1, \dots, \mathcal{M}_n$ have full rank. If the camera centers are distinct, this condition can be guaranteed by changing bases appropriately in the different image planes. Since this kind of operation preserves the dimension of spaces generated by multilinearities (because the vector spaces are mapped isomorphically onto themselves), this assumption does not affect the general validity of the result.

Proof. Let W_n be the vector space generated by multlinear relations arising as maximal minors of $\mathcal{U}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ defined in (12); for simplicity we will refer to these constraints as the "fundamental" multilinearities. We will first compute the dimension of W_n; then we will show that W_n coincides with the space of *all* multilinear constraints which vanish on $\overline{\mathcal{I}}_n$. Our proof heavily relies on the fact that the fundamental multilinearities form a *Gröbner basis* for lexicographic ordering (and in fact, for all admissible term orderings): this will be shown in Proposition b3.

To compute the dimension of W_n we will use the following well known property of Gröbner bases (see for example [3]):

(P) If a set of homogeneous polynomials $f_1, \dots, f_N$ forms a Gröbner basis (for a fixed monomial ordering), then $\dim_{\mathbb{R}}(f_1, \dots, f_N) = \dim_{\mathbb{R}}(In(f_1), \dots, In(f_N))$, where $In(f_i)$ denotes the initial term of f_i.

To use this result, we first introduce a lexicographic ordering for the monomials in $\mathbb{R}[x_1, y_1, z_1, \dots, x_n, y_n, z_n]$, so that $x_1 \succ \dots \succ x_n \succ y_1 \succ \dots \succ y_n \succ z_1 \dots \succ z_n$. In Lemma b3, we will show that for this ordering the set of leading terms of the fundamental mutlilinearities is given by $T = M_n \setminus S$, where

$$M_n = \{x^\alpha y^\beta z^\gamma \mid \alpha, \beta, \gamma \in \mathbb{N}^n, \;\; \alpha + \beta + \gamma = [1; \dots; 1]\}, \tag{13}$$

is the set of all multilinear monomials, and

$$\begin{aligned} S = &\{x^\alpha y^\beta z^\gamma \in M_n \mid \alpha = 0, |\beta| \leq 3\} \\ &\cup \{x^\alpha y^\beta z^\gamma \in M_n \mid |\alpha| = 1, |\beta| \leq 1\}, \end{aligned} \tag{14}$$

where we used multi-index notation, so that for example $x^\alpha = \prod_k x_k^{\alpha(k)}$. According to property (P), we know that $\dim(W_n) = \dim(Span(T))$. Since all monomials are linearly independent, we can now compute the dimension of W_n using a simple counting argument:

$$\begin{aligned} \dim(W_n) &= |T| = |M_n| - |S| = \\ &= 3^n - \left(1 + n + \binom{n}{2} + \binom{n}{3}\right) - n\Big(1 + (n-1)\Big) \\ &= 3^n - \left(1 + n + \binom{n}{2} + \binom{n}{3} + n(n+1)\right) + n \\ &= 3^n - \binom{n+3}{3} + n. \end{aligned} \tag{15}$$

Finally, to show that W_n coincides with the space of all multilinear constraints, we recall that the fundamental multilinearities set-theoretically define the joint image variety $\overline{\mathcal{I}}_n$ (they give necessary and sufficient conditions for correspondence). Using the *Nullstellensatz* theorem, it is enough to prove that they generate a radical ideal[3]. This will also be shown in Proposition b3.

□

Lemma b2. *Using the ordering $x_1 \succ \dots \succ x_n \succ y_1 \succ \dots \succ y_n \succ z_1 \dots \succ z_n$, and assuming the condition in the Remark, the leading terms of the fundamental multilinearities are given by $T = M_n \setminus S$, where M_n and S are defined in* (13) *and* (14).

[3] Assuming the ideal $\mathfrak{I}$ generated by the fundamental multilinearities is radical, we can write $\mathfrak{I} = \mathfrak{P} \cap \bigcap_i (x_i, y_i, z_i)$ where $\mathfrak{P}$ is the prime ideal describing $\overline{\mathcal{I}}_n$ (defined by bilinear, trilinear, and quadrilinear relations [1]). Any multilinearity that vanishes on $\overline{\mathcal{I}}_n$ lies in $\mathfrak{P}$ and in (x_i, y_i, z_i) for all i, thus it must belong to $\mathcal{I}$.

Proof. From (13), we see that a monomial $x^\alpha y^\beta z^\gamma$ in M_n can be also identified by the vector $\nu = \alpha + 2\beta + 3\gamma \in [1, 2, 3]^n$; in this case we will write $\mu_\nu = x^\alpha y^\beta z^\gamma$. One can verify that μ_{ν_1} can be a leading term of a fundamental multilinearity if and only if there exist ν_2, ν_3, ν_4 such that $\mu_{\nu_1} \succ \mu_{\nu_i}$ and $|\nu_1 - \nu_i| \leq 4$ for $i = 2, 3, 4$. Indeed, a fundamental mutilinearity P is defined by choosing $n + 4$ rows of $\mathcal{U}$ from (12), and the monomials μ_ν appearing in P correspond to subsets of n rows (which are expressed by the vectors $\nu \in [1, 2, 3]^n$); note also that the generality assumption guarantees that all possible coefficients are non-zero. It is straightforward to verify that μ_{ν_1} does not satisfy the previous property if and only if it belongs to the set S defined in (14).

□

Proposition b3. *The fundamental multilinearities form a universal Gröbner basis (*i.e.*, a Gröbner basis relative to any monomial ordering). Moreover, they define a radical ideal.*

Proof. Our proof is motivated by a classical result, which states that in the ring $R = \mathbb{R}[x_{ij}]$, where $i = 1, \ldots, m$, $j = 1, \ldots, n$, and $m \leq n$, the maximal $m \times m$ minors of $X = (x_{i,j})$ form a universal Gröbner basis (generating the *classical determinantal ideal*) [14].

In order to obtain a similar statement for the maximal minors of the matrix $\mathcal{U}$, we use the following two results:

1. If Z is a subset of variables in $\mathbb{R}[x_{ij}]$, and X' is the matrix obtained from $X = (x_{i,j})$ by setting the variables in Z to zero, then the nonzero maximal minors of X' form a universal Gröbner basis [2].

2. Let $f_1, \ldots, f_N$ be a set of polynomials forming a universal Gröbner basis in $\mathbb{R}[s_1, \ldots, s_k, t_1, \ldots, t_j]$, and consider the evaluation map
$$\begin{aligned} \phi : \mathbb{R}[s_1, \ldots, s_k, t_1, \ldots, t_j] &\to \mathbb{R}[s_1, \ldots, s_k] \\ g(s_1, \ldots, s_k, t_1, \ldots, t_j) &\mapsto g(s_1, \ldots, s_k, c_1, \ldots, c_j) \end{aligned} \tag{16}$$
for a fixed set $c_1, \ldots, c_j \in \mathbb{R}$. Thinking $f_1, \ldots, f_N$ as polynomials with ring coefficients, *i.e.*, as elements of $R[s_1, \ldots, s_k]$ with $R = \mathbb{R}[t_1, \ldots, t_j]$, we denote with $Lc(f_i) \in R$ the leading coefficient of f_i. If $\phi(Lc(f_1)), \ldots, \phi(Lc(f_N))$ are non-zero, then $\phi(f_1), \ldots, \phi(f_N)$ form a universal Gröbner basis in $\mathbb{R}[s_1, \ldots, s_k]$ [12].

We now consider the matrix $\mathcal{U}'$ as in (12), but with indeterminate camera matrices: according to the first result given above, the maximal minors form a universal Gröbner basis. The fundamental multilinearities are obtained from these minors by substituting the actual entries of the camera matrices. If the leading coefficients of the minors are not mapped to zero under this specialization, then we can use the second result and obtain the claim. However, it is easy to realize that the leading coefficients are given by the determinants of the matrices obtained by selecting four rows from the camera matrices; assuming the generality condition from the Remark, these do not vanish for the actual camera entries. This implies that the fundamental multilinearities form a universal Gröbner basis.

Finally, we now see that the monomials in $T = M_n \setminus S$ (defined in (13) and (14)) generate the *initial ideal* for the ideal generated by fundamental multilinearities. The fact that this monomial ideal is squarefree, allows us to conclude that the fundamental multilinearities generate a radical ideal; see [10, Corollary 2.2].

□

B3. Geometric properties of the joint image

In this section we show that the joint image variety is closely related to the *blow-up*, a fundamental construction in algebraic geometry [8]. For our purposes we can give the following definition for the blow-up:

Definition 1. *Consider a smooth algebraic variety X of dimension n, and a subvariety $Z \subseteq X$ of dimension d. Locally, we may write X as $Z \times W$ where W has dimension $n - d$ and is transversal to Z. Let $\lambda : X \setminus Z \to \mathbb{P}^{n-d-1}$ be the map associating with $x = (z, w) \in Z \times W$ the line ℓ_x passing through w and z. The* blow-up *of X in Z is given by closure of the graph of λ:*
$$\tilde{X}_Z = \overline{\{(x, \ell_x), x \in X \setminus Z\}} \subseteq X \times \mathbb{P}^{n-d-1}. \tag{17}$$
This is a smooth variety, with a natural map $\pi : \tilde{X}_Z \to X$, known as the blow-up map. *The* exceptional locus *is the inverse image $\tilde{Z} = \pi^{-1}(Z) = Z \times \mathbb{P}^{n-d-1}$ of the center Z. It is is the set of points where π fails to be a local isomorphism: in fact, π always contracts the second factor $\mathbb{P}^{n-d-1}$ of $\tilde{Z}$ to a point. See Figure 2 in the main part of paper.*

Basically, the blow-up replaces a subvariety $Z \subset X$ with all the directions in X pointing out of Z. Blow-ups are an extremely important tool for the *resolution of singularities*, that is, the operation of constructing suitable smooth models for varieties with singularities. We refer to [9] for a nice and accessible presentation.

Proposition b4. *If $n \geq 3$ cameras do not have collinear pinholes, the joint image variety $\overline{\mathcal{I}}_n$ is isomorphic to the* blow-up *of $X = \mathbb{P}^3$ at $Z = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n\}$, where $\boldsymbol{c}_i$, $i = 1, \ldots, n$ are the camera pinholes.*

Proof. Iterating the construction given in Definition 1, we see that the blow-up of $\mathbb{P}^3$ at $Z = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n\}$ is given by
$$\tilde{X}_Z = \overline{\{(\boldsymbol{p}, \ell^1_{\boldsymbol{p}}, \ldots, \ell^n_{\boldsymbol{p}}),\ \boldsymbol{p} \in \mathbb{P}^3 \setminus Z\}} \subseteq \mathbb{P}^3 \times (\mathbb{P}^2)^n, \tag{18}$$

where $\ell^i_{\boldsymbol{p}}$ denotes the line through $\boldsymbol{p}$ and $\boldsymbol{c}_i$. In particular, the projection of $\tilde{X}_Z$ onto $(\mathbb{P}^2)^n$ is exactly the joint image variety $\overline{\mathcal{I}}_n$, so we have a natural map

$$f : \tilde{X}_Z \to \overline{\mathcal{I}}_n. \tag{19}$$

In order to prove that f is an isomorphism we have to show that: 1) f is injective 2) the differential Tf of f is injective.

The fact that f is injective for $(\boldsymbol{p}, \ell^1_{\boldsymbol{p}}, \ldots, \ell^n_{\boldsymbol{p}})$ with $\boldsymbol{p} \in \mathbb{P}^3 \setminus Z$ follows from the observation that the "joint projection map" $\mathbb{P}^3 \to \mathbb{P}^2 \times \ldots \times \mathbb{P}^2$ is injective if one assumes non-collinear pinholes (see Section 2.3 of the main part of the paper). Moreover, if $\boldsymbol{p} = \boldsymbol{c}_i$ for some i, then the exceptional set $\pi^{-1}(\boldsymbol{c}_i)$ is mapped isomorphically onto $\{(\boldsymbol{e}_{1i}, \ldots, \ell, \ldots \boldsymbol{e}_{ni}), \ell \in \mathbb{P}^2\} \subseteq \overline{\mathcal{I}}_n$, where $\boldsymbol{e}_{ji}$ is the epipole in the j-th image for the center $\boldsymbol{c}_i$ (note that the image of the exceptional locus $f(\pi^{-1}(Z))$ is exactly the set $\mathcal{C}_n$ given by Proposition A1).

To prove that the differential Tf is injective at points $(\boldsymbol{p}, \ell^1_{\boldsymbol{p}}, \ldots, \ell^n_{\boldsymbol{p}})$ with $\boldsymbol{p} \in \mathbb{P}^3 \setminus Z$, it is enough to observe that all lines in $\mathbb{P}^3$ are mapped injectively by the joint projection map (since f is locally an affine map, if the differential were not injective, some line would have to be contracted). Similarly, if $\boldsymbol{p} = \boldsymbol{c}_i$, then all lines contained in $\pi^{-1}(\boldsymbol{c}_i)$ are mapped isomorphically on some image $\mathbb{P}^2$, and the same holds for all lines passing through $\boldsymbol{c}_i$ in $\mathbb{P}^3$. ☐

Proposition 4. (Singularities of the joint image variety) *When the camera pinholes are not collinear, $\overline{\mathcal{I}}_n$ is smooth. When they are collinear (in particular, for $n = 2$ views), then $\overline{\mathcal{I}}_n$ has a unique singular point given by the n-tuple of epipoles $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$.*

Proof. If the camera pinholes are not collinear, then Proposition A3 immediately implies that the joint image variety is smooth. If the pinholes are collinear, then the projection map $f : \tilde{X}_Z \to \overline{\mathcal{I}}_n$ considered in the previous proof is still well defined, and is a local isomorphism except at points that lie on the baseline $\ell \subseteq \mathbb{P}^3$ containing the centers (more precisely, on its *strict transform* in $\tilde{X}_Z$, given by $\overline{\pi^{-1}(\ell \setminus Z)}$). In fact, f contracts this set onto $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$. In general, however, birational morphisms between smooth varieties can only contract sets of codimension 1 [13, Theorem 2.2], while the baseline has codimension 2: we conclude that $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ must be the only singular point of $\overline{\mathcal{I}}_n$. ☐

B4. Dependencies among multi-view constraints

In this section we prove Proposition 5 of the main part of the paper.

Proposition 5. *Assume n cameras are given.*

1. Bilinearities and trilinearities always *strongly characterize $\overline{\mathcal{I}}_n$, independently of the camera configurations.*

2. Bilinear constraints alone *strongly characterize $\overline{\mathcal{I}}_n$, if and only if the pinholes are* not all coplanar.

3. Bilinear constraints alone *weakly characterize $\overline{\mathcal{I}}_n$ if and only if the pinholes are* not all collinear.

Proof. Let us first assume that $n = 4$.

If $\mathcal{M}_1, \ldots, \mathcal{M}_4$ are cameras with pinholes $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ in *general position* (*i.e.*, non coplanar), then up to homographies of $\mathbb{P}^3$ (that do not affect the joint image) and homographies of image planes (that simply result in linear changes of variables that map k-linearities isomorphically onto themselves) we may assume that:

$$\begin{aligned} \mathcal{M}_1 &= \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \ \mathcal{M}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\ \mathcal{M}_3 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \ \mathcal{M}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \end{aligned} \tag{20}$$

In this case, one can verify using Gröbner bases that the six epipolar relations are sufficient to generate, up to irrelevant components, the ideal associated to the joint image variety $\overline{\mathcal{I}}_4$ (see also [11]).

If $\mathcal{M}_1, \ldots, \mathcal{M}_4$ are cameras with pinholes $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ in *general coplanar position* (*i.e.*, coplanar with no subset of three that are collinear), then we may similarly assume that:

$$\begin{aligned} \mathcal{M}_1 &= \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \ \mathcal{M}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\ \mathcal{M}_3 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \ \mathcal{M}_4 = \begin{pmatrix} 1 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & 1/3 \\ 0 & 0 & 1 & 1/3 \end{pmatrix}. \end{aligned} \tag{21}$$

This time, using Gröbner basis we see that the epipolar conditions describe an algebraic set $\mathcal{W}$ that is strictly larger than $\overline{\mathcal{I}}_4$. In fact, $\mathcal{W}$ decomposes as

$$\mathcal{W} = \overline{\mathcal{I}}_4 \cup \mathcal{V}_t \tag{22}$$

where $\mathcal{V}_t = \{x_i + y_i + z_i = 0 \,|\, i = 1, \ldots, 4\}$ is the product of the "trifocal lines" (*i.e.*, the projection of the plane containing the pinholes). We see that the epipolar constraints are only a *weak* characterization of the joint image: camera geometry can be recovered (for example by considering subsets of three cameras, see [7]) although correspondences are not directly characterized. On the other hand, one can verify that including trilinear relations is sufficient to exclude the spurious solutions and yields a strong characterization of $\overline{\mathcal{I}}_4$

If $\mathcal{M}_1, \ldots, \mathcal{M}_4$ are cameras with pinholes $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ that are *coplanar, and with a subset of three that are collinear*, then we can consider

$$\begin{aligned}\mathcal{M}_1 &= \begin{pmatrix}1&0&0&1\\0&1&0&0\\0&0&1&0\end{pmatrix}, \ \mathcal{M}_2 = \begin{pmatrix}1&0&0&0\\0&1&0&1\\0&0&1&0\end{pmatrix},\\ \mathcal{M}_3 &= \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&1\end{pmatrix}, \ \mathcal{M}_4 = \begin{pmatrix}1&0&0&0\\0&1&0&1/2\\0&0&1&1/2\end{pmatrix}.\end{aligned} \tag{23}$$

Once again, we see that that the bilinear and trilinear constraints completely characterize $\overline{\mathcal{I}}_4$, while only bilinear conditions describe a set $\mathcal{W}$ that also decomposes as

$$\mathcal{W} = \overline{\mathcal{I}}_4 \cup \mathcal{V}_t \tag{24}$$

where $\mathcal{V}_t = \{x_i + y_i + z_i = 0 \,|\, i = 1, \ldots, 4\}$ is the product of the “trifocal lines”.

Finally, if all $\mathcal{M}_1, \ldots, \mathcal{M}_4$ have collinear pinholes, we cannot simply change reference frame in $\mathbb{P}^3$ and assume the cameras to be fixed, since four collinear points are not always projectively equivalent (the projective invariant is given by the cross-ratio). However, we can use a single parameter to describe all such camera configurations:

$$\begin{aligned}\mathcal{M}_1 &= \begin{pmatrix}1&0&0&1\\0&1&0&0\\0&0&1&0\end{pmatrix}, \ \mathcal{M}_2 = \begin{pmatrix}1&0&0&0\\0&1&0&1\\0&0&1&0\end{pmatrix},\\ \mathcal{M}_3 &= \begin{pmatrix}1&0&0&1/2\\0&1&0&1/2\\0&0&1&0\end{pmatrix}, \ \mathcal{M}_4 = \begin{pmatrix}1&0&0&t\\0&1&0&1-t\\0&0&1&0\end{pmatrix}.\end{aligned} \tag{25}$$

In this case, computing the primary decomposition we see that bilinear and trilinear constraints have spurious components only for $t = 0, 1/2, 1$ which correspond to non-valid values for t (we assume the camera pinholes are distinct). Bilinear constraints, on the other hand, characterize (for feasible values of of t) a *single irreducible component* $\mathcal{W}$, that strictly contains contains the joint image. With our choice of camera matrices, this component is described by

$$\begin{aligned}\mathcal{W} = \{ &- z_3x_4 - z_3y_4 + x_3z_4 + y_3z_4 = 0,\\ &- z_2x_4 - z_2y_4 + x_2z_4 + y_2z_4 = 0,\\ &- z_2x_3 - z_2y_3 + x_2z_3 + y_2z_3 = 0,\\ &- z_1x_4 - z_1y_4 + x_1z_4 + y_1z_4 = 0,\\ &- z_1x_3 - z_1y_3 + x_1z_3 + y_1z_3 = 0,\\ &- z_1x_2 - z_1y_2 + x_1z_2 + y_1z_2 = 0\}.\end{aligned} \tag{26}$$

Note that these expressions *do not depend on the parameter* t, which shows that camera matrices cannot be determined by the epipolar conditions. We conclude that in this case the bilinearities do not give a weak characterization of the joint image.

The previous analysis proves Proposition 4 for $n = 4$, since we have exhausted all possible configurations of the four pinholes in $\mathbb{P}^3$. The case $n > 4$ now follows easily Indeed, bilinear and trilinear are always strongly sufficient because we have shown this to be true for all sets of four cameras (and this is enough thanks to Proposition 1 in the main part of the paper). If the pinholes are non-coplanar, we may assume that cameras $(1, 2, 3)$ have non-collinear pinholes and span a plane that does not contain any other pinhole, so applying the previous argument for all quadruplets of cameras $(1, 2, 3, i)$ we can conclude that bilinearities give a strong characterization of the joint image. If the pinholes are only non-collinear, we can assume that pinholes $(1, 2)$ span a line that doesn't contain other pinholes, and similarly use this to conclude that bilinearities give a weak characterization of the joint image (see the proof of Proposition 5 in the main part of the paper). Finally, whenever the pinholes are all aligned, camera geometry cannot be uniquely determined from bilinearities: for example, the expressions in (26) show that, using appropriate image coordinates, *all* pairs of cameras with pinholes on a given line yield the same epipolar constraint. More generally, it is easy to realize that if $\mathcal{M}_1, \ldots, \mathcal{M}_n$ have collinear pinholes, then any set of camera matrices $\mathcal{M}'_1, \ldots, \mathcal{M}'_n$, where $\mathcal{M}'_i$ is obtained from $\mathcal{M}_i$ by adding to its columns multiples of the coordinate vector for the epipole, will yield the same set of epipolar constraints. □

B5. Three views

We now focus on the case of $n = 3$ views. We point out that any three cameras with non-collinear pinholes can be transformed by homographies of $\mathbb{P}^3$ and of the image planes into the triplet

$$\begin{aligned}\mathcal{M}_1 = \begin{pmatrix}1&0&0&1\\0&1&0&0\\0&0&1&0\end{pmatrix}, \ M_2 = \begin{pmatrix}1&0&0&0\\0&1&0&1\\0&0&1&0\end{pmatrix},\\ \mathcal{M}_3 = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&1\end{pmatrix}.\end{aligned} \tag{27}$$

Similarly, three cameras with collinear pinholes can be transformed into

$$\begin{aligned}\mathcal{M}_1 = \begin{pmatrix}1&0&0&1\\0&1&0&0\\0&0&1&0\end{pmatrix}, \ \mathcal{M}_2 = \begin{pmatrix}1&0&0&0\\0&1&0&1\\0&0&1&0\end{pmatrix},\\ \mathcal{M}_3 = \begin{pmatrix}1&0&0&1/2\\0&1&0&1/2\\0&0&1&0\end{pmatrix}.\end{aligned} \tag{28}$$

These camera matrices will be used for our explicit computations.

We now consider the trilinear constraints encoded in the trifocal tensor. We recall that these conditions can be deduced from the maximal minors of the matrix

$$\mathcal{U}(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3)=\begin{pmatrix}\mathcal{M}_1 & \boldsymbol{u}_1 & \mathbf{0} & \mathbf{0}\\ \mathcal{M}_2 & \mathbf{0} & \boldsymbol{u}_2 & \mathbf{0}\\ \mathcal{M}_3 & \mathbf{0} & \mathbf{0} & \boldsymbol{u}_3\end{pmatrix}. \tag{29}$$

More precisely, the nine trilinearities in the tensor that distinguishes the first view are given by

$$T_{ij},\ \ i\in\{4,5,6\}, j\in\{7,8,9\} \tag{30}$$

where T_{ij} is obtained by considering all rows of $\mathcal{U}$ excluding i and j.

Proposition 8. *The nine trilinearities* (30) *that are encoded in the trifocal tensor span a vector space of dimension* 8.

Proof. For the triplets of camera matrices (27) or (28) the claim can be verified by direct computation (it is enough to consider a reduced Gröbner basis of the generated ideal and count the number of trilinear polynomials). Thus, we only need to show that changing reference frames does not affect the dimension of the vector space spanned by (30). This is clear for homographies of $\mathbb{P}^3$, since these do not affect the expressions of the trilinear constraints (see Section 2.3 in the main part of the paper). On the other hand, a change of coordinates in the different image planes corresponds to a linear change of variables in the multilinearities and it is easy to realize that the space generated by the trilinearities defined by the trifocal tensor is mapped isomorphically onto itself. Thus, the dimension of the space generated by this set of constraints is preserved by changes of reference frames in the different images. □

Proposition 9. *If $\mathcal{M}_1,\mathcal{M}_2,\mathcal{M}_3$ do not have collinear pinholes, then the constraints* (30) *describe a set $\mathcal{W}=\overline{\mathcal{I}}_3\cup\mathcal{S}_{12}\cup\mathcal{S}_{13}$, where*

$$\begin{aligned}\mathcal{S}_{12}&=\{\boldsymbol{e}_{12}\times\boldsymbol{e}_{21}\times\boldsymbol{u}_3\in(\mathbb{P}^2)^3\,|\,\boldsymbol{u}_3\in\mathbb{P}^2\},\\ \mathcal{S}_{13}&=\{\boldsymbol{e}_{13}\times\boldsymbol{u}_2\times\boldsymbol{e}_{21}\in(\mathbb{P}^2)^3\,|\,\boldsymbol{u}_2\in\mathbb{P}^2\},\end{aligned} \tag{31}$$

and $\boldsymbol{e}_{ij}$ is the epipole in image i relative to the camera j.

Proof. Considering the camera matrices (27), one can verify using primary decomposition that the nine relations describe a space $\mathcal{W}$ that decomposes as $\mathcal{W}=\overline{\mathcal{I}}_3\cup\mathcal{S}_{12}\cup\mathcal{S}_{13}$, where

$$\begin{aligned}\mathcal{S}_{12}=\{x_1+y_1&=0;\\ x_2+y_2&=0;\\ z_1&=0;\\ z_2&=0\}\\ \mathcal{S}_2=\{x_1+y_1&=0;\\ x_3+y_3&=0;\\ z_1&=0;\\ z_3&=0\}.\end{aligned} \tag{32}$$

For our choice of cameras, these sets are exactly the ones described by (31), which in turn are a characterization that is independent of the choice of reference frames. □

By using the collinear cameras of (28) one can also show the following

Proposition b5. *If $\mathcal{M}_1,\mathcal{M}_2,\mathcal{M}_3$ have collinear pinholes, the constraints* (30) *strongly characterize the joint image $\overline{\mathcal{I}}_3$.*

Let us now clarify a subtle point relative to this last result: from Proposition 1 in the paper, we know that 9 trilinearities can never generate all of the trilinear relations which vanish on $\mathcal{V}_3$ (since this is always a vector space of dimension 10). However, this does not exclude the possibility for these to *set-theoretically* define $\mathcal{V}_3$: in other words, the zero-set of some relations can still describe $\mathcal{V}_3$, even though they do not algebraically generate the whole space of the trilinearities. This is the case for the nine trilinearities (30) for cameras with collinear pinholes. For example, one can verify that for the matrices (28) the trilinearity

$$T_{26}=-z_1x_2x_3-2z_1y_2x_3+z_1x_2y_3+x_1x_2z_3+x_1y_2z_3 \tag{33}$$

does not belong to the vector space generated by (30), although its zero-set clearly contains that of these nine trilinearities. To motivate this fact, one can show that T_{26}^2 *does* belong to the ideal generated by (30) (*i.e.*, T_{26}^2 can be written as an algebraic combination of the nine trilinearities) even if T_{26} doesn't: this clearly implies that adding the constraint T_{26} to (30) does not impose any additional conditions on the zero-set (since $T_{26}=0$ if and only if $T_{26}^2=0$).

Proposition 11. *Consider three cameras $\mathcal{M}_1,\mathcal{M}_2,\mathcal{M}_3$.*

- *If the pinholes are non-collinear:*

  1. *For* any *trilinearity T that does not vanish on the product of the trifocal lines, $\{B_{12},B_{13},B_{23},T\}$ gives a strong characterization of the joint image.*
  2. *The epipolar constraints $\{B_{12},B_{13},B_{23}\}$ uniquely determine camera geometry,* i.e.*, they give a weak characterization of the joint image.*

- *If the cameras have collinear pinholes:*

  1. *A strong characterization of the joint image is given by* $\{B_{12}, B_{13}, B_{23}, T_1, T_2\}$ *where* T_1 *and* T_2 *are (sufficiently general) trilinear constraints.*
  2. *Two epipolar constraints together with one (sufficiently general) trilinearity* $\{B_{12}, B_{13}, T\}$ *uniquely determine camera geometry,* i.e.*, they give a weak characterization of the joint image.*

Proof. Let us first assume that $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ have non-collinear pinholes.

1. The epipolar constraints describe the same projective locus as nine trilinearities (namely, $T_{12}, T_{13}, T_{23}, T_{45}, T_{46}, etc.$), that are obtained by multiplying bilinearities with a variable associated to the excluded image. In the case of non-collinear pinholes, one can verify that these constraints span a vector space of dimension 9. Any trilinearity that does not vanish on the product of the trifocal lines cannot belong to this space, so the three bilinearities together with such a trilinearity must describe the same zero-set as the whole space of trilinear constraints, *i.e.*, exactly $\overline{\mathcal{I}}_3$. See also [15].
2. This is well known: see [7, 11], or the discussion in Example 3 in the main part of the paper.

We now assume that $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ have collinear pinholes.

1. Consider constraints $\{B_{12}, B_{13}, B_{23}, T_{ij}\}$, where T_{ij} is chosen from (30). In *generic conditions* these describe a set $\mathcal{W}$ that decomposes as $\mathcal{W} = \overline{\mathcal{I}}_3 \cup \mathcal{V}_1 \cup \mathcal{V}_2$, where $\mathcal{V}_1$ and $\mathcal{V}_2$ are each products of three corresponding (epipolar) lines in each image, respectively passing through the $(i-2)$-th coordinate points in the second image and through the $(j-5)$-th coordinate point in the third (*e.g.*, if we consider T_{48} then $\mathcal{V}_1$ is the product of the three epipolar lines containing $[1; 0; 0] \in \mathbb{P}^2$ in the second image and $\mathcal{V}_2$ is the product of the three epipolar lines containing $[0; 1; 0] \in \mathbb{P}^2$ in the third). Moreover, in general, if another trilinearity T_{kl} among (30) is such that $k \neq i$ and $l \neq j$, then $\{B_{12}, B_{13}, B_{23}, T_{ij}, T_{kl}\}$ are strongly sufficient, since the spurious sets associated with each trilinearity are excluded by the other constraint. The "generic" assumption that was used in this argument is that the triplets of epipolar lines associated with the fundamental points are all distinct: in other words, we assumed that 1) no fundamental point in each image corresponds to any other fundamental point in another 2) no pair of fundamental points in a given image lies on the same epipolar line[4]. However, even if these conditions are not satisfied, there will *always* be at least two trilinearities among (30) that can be used together with the epipolar constraints to characterize $\overline{\mathcal{I}}_3$: this follows from the fact that in each image one can always choose two fundamental points that do not lie on the same epipolar line. See Example 2 in Section 6.
2. Let T be any trilinearity whose zero-set does not contain that of $\{B_{12}, B_{13}\}$: then we may be sure that $\mathcal{W} = \{B_{12} = B_{13} = T = 0\}$ is a set of dimension 3, which must contain $\overline{\mathcal{I}}_3$ as a component of maximal dimension. By excluding the spurious components we can recover a strong characterization of the joint image, and thus camera geometry can be determined.

□

[4] We point out that the matrices (28) *do not* satisfy these assumptions.

B6. Transfer using multilinear constraints

We now present a brief discussion that relates the degeneracies of some classical approaches to point transfer with the spurious components arising from weak characterizations of the joint image.

Let us first consider some bilinear and/or trilinear constraints $P_1, \ldots, P_s$ among three views that provide a *strong* characterization of the joint image $\overline{\mathcal{I}}_3$. Given a pair of corresponding points $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ in the first two images (we may assume for simplicity that the epipolar constraint between two views is known), it is possible to find a corresponding point in the third image by solving a *linear system* in the affine coordinates of the third point:

$$P_i(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{y}_3) = 0, \quad i = 1, \ldots, s. \tag{34}$$

The coordinates of a point $\boldsymbol{u}_3$ will satisfy (34) *if and only if* $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ is a correspondence.

A similar approach can actually be applied even if $P_1, \ldots, P_s$ do not strongly characterize the joint image, but still describe a set of dimension 3. Indeed, this assumption guarantees that for *generic* corresponding pairs $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ (formally, pairs in an open dense set), the space of solutions to (34) will be zero-dimensional, *i.e.*, a point (since it must be a linear space). The extent to which this approach will fail for non-generic pairs depends on the extraneous correspondences characterized by $P_1, \ldots, P_s$: indeed, in general, solutions to (34) will not necessarily coincide with correspondences. In particular, we can consider the following variations of transfer based on (34):

- Using $\{B_{12}, B_{13}, B_{23}\}$ yields *epipolar transfer*, which is known to fail for pairs that lie on the trifocal lines, or when the pinholes are collinear.
- Using the nine trilinearities (30) is equivalent to *transfer based on the trifocal tensor*: this will fail when $\boldsymbol{u}_1, \boldsymbol{u}_2$ are epipoles, giving no conditions on $\boldsymbol{u}_3$ [7].

- Using any strong characterization of $\overline{\mathcal{I}}_3$ (for example the ones given in Proposition 11) is equivalent to *triangulation* based on the first two views followed by *re-projection* onto the third: in this case we may be sure that solutions to (34) give actual correspondences. For example, if $\boldsymbol{u}_1, \boldsymbol{u}_2$ are corresponding epipoles, then even though $\boldsymbol{u}_3$ is not uniquely determined, the constraints (correctly) impose that it must lie on the trifocal line (contrary to trifocal transfer).

B7. Examples

We conclude by discussing two simple examples that illustrate in practice the results discussed in the previous section.

1) Consider the three cameras with non-collinear pinholes given in (27). The sets $\mathcal{W}_1 = \{B_{12} = B_{13} = B_{23} = 0\}$ and $\mathcal{W}_2 = \{B_{12} = B_{13} = T_{47} = 0\}$ decompose as

$$\begin{aligned} \mathcal{W}_1 &= \overline{\mathcal{I}}_3 \cup \mathcal{V}_t \\ \mathcal{W}_2 &= \overline{\mathcal{I}}_3 \cup \mathcal{V}_a \cup \mathcal{V}_b \end{aligned} \tag{35}$$

where $\mathcal{V}_t = \{x_i + y_i + z_i = 0 \,|\, i = 1,2,3\}$ is the product of the trifocal lines, and

$$\begin{aligned} \mathcal{V}_a &= \{B_{12} = y_1 = y_3 = 0\} \\ \mathcal{V}_b &= \{B_{13} = z_1 = z_2 = 0\}. \end{aligned} \tag{36}$$

In particular, we see that both sets of constraints $\{B_{12}, B_{13}, B_{23}\}$ and $\{B_{12}, B_{13}, T_{47}\}$ are weak characterizations of the joint image $\overline{\mathcal{I}}_3$ (and they give *minimal* and *local* descriptions of $\overline{\mathcal{I}}_3$, see Example 2 in the main part of the paper). On the other hand

$$\{B_{12} = B_{13} = B_{23} = T_{47} = 0\} = \mathcal{W}_1 \cap \mathcal{W}_2 = \overline{\mathcal{I}}_3 \tag{37}$$

so $\{B_{12}, B_{13}, B_{23}, T_{47}\}$ give a strong characterization of the joint image, confirming Proposition 11.

2) Consider now the three cameras with collinear pinholes given in (27). We introduce the following sets:

$$\begin{aligned} \mathcal{W}_3 &= \{B_{12} = B_{13} = B_{23} = 0\} \\ \mathcal{W}_4 &= \{B_{12} = B_{13} = T_{47} = 0\} \\ \mathcal{W}_5 &= \{B_{12} = B_{13} = B_{23} = T_{47} = 0\} \\ \mathcal{W}_6 &= \{B_{12} = B_{13} = B_{23} = T_{47} = T_{58} = 0\} \\ \mathcal{W}_7 &= \{B_{12} = B_{13} = B_{23} = T_{47} = T_{69} = 0\}. \end{aligned} \tag{38}$$

One can verify that:

- $\mathcal{W}_3$ is an *irreducible* set, and in fact the epipolar constraints do not yield a weak characterization of $\overline{\mathcal{I}}_3$ (Proposition 4).

-$\mathcal{W}_4$ decomposes as $\mathcal{W}_4 = \overline{\mathcal{I}}_3 \cup \mathcal{V}_c \cup \mathcal{V}_d \cup \mathcal{V}_e$ where

$$\begin{aligned} \mathcal{V}_c &= \{z_1 = z_2 = z_3 = 0\} \\ \mathcal{V}_d &= \{x_1 + y_1 = z_1 = z_3 = 0\} \\ \mathcal{V}_e &= \{x_1 + y_1 = z_1 = z_2 = 0\}. \end{aligned} \tag{39}$$

In particular, $\{B_{12}, B_{13}, T_{47}\}$ give a weak characterization of $\overline{\mathcal{I}}_3$, as stated in Proposition 11.

-$\mathcal{W}_5$ decomposes as $\mathcal{W}_5 = \overline{\mathcal{I}}_3 \cup \mathcal{V}_c$, so, contrary to the non-collinear case, the conditions $\{B_{12} = B_{13} = B_{23} = T_{47} = 0\}$ are *not* a strong characterization of the joint image. Note also that there is only one spurious component, because the first fundamental points (*i.e.*, $[1;0;0] \in \mathbb{P}^2$) in the second and third image correspond (see the proof of Proposition 11).

-$\mathcal{W}_6$ also decomposes as $\mathcal{W}_6 = \overline{\mathcal{I}}_3 \cup \mathcal{V}_c$, because the generic assumption in the proof of Proposition 11 is not satisfied (the spurious sets associated with T_{47} and T_{58} coincide).

-$\mathcal{W}_7 = \overline{\mathcal{I}}_3$, so that $\{B_{12} = B_{13} = B_{23} = T_{47} = T_{69} = 0\}$ is a strong characterization of the joint image.

References

[1] C. Aholt, B. Sturmfels, and R. Thomas. A Hilbert scheme in computer vision. *Canadian Journal of Mathematics*, 65:961–988, 2013. 3, 4

[2] A. Boocher. Free resolutions and sparse determinantal ideals. *arXiv preprint arXiv:1111.0279*, 2011. 5

[3] D. A. Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Springer, 2007. 1, 4

[4] W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 4-0-2 — A computer algebra system for polynomial computations. http://www.singular.uni-kl.de. 3

[5] P. Gianni, B. Trager, and G. Zacharias. Gröbner bases and primary decomposition of polynomial ideals. *Journal of Symbolic Computation*, 6(2):149–167, 1988. 3

[6] D. R. Grayson and M. E. Stillman. Macaulay2, a software system for research in algebraic geometry. http://www.math.uiuc.edu/Macaulay2/. 3

[7] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2000. 6, 9

[8] R. Hartshorne. *Algebraic Geometry*. Encyclopaedia of mathematical sciences. Springer, 1977. 5

[9] H. Hauser. The hironaka theorem on resolution of singularities. *Bull.(New Series) Amer. Math. Soc*, 40:323–403, 2003. 5

[10] J. Herzog, T. Hibi, F. Hreinsdóttir, T. Kahle, and J. Rauh. Binomial edge ideals and conditional independence statements. *Advances in Applied Mathematics*, 45(3):317–333, 2010. 5

[11] A. Heyden and K. Astrom. Algebraic properties of multilinear constraints. *Mathematical Methods in the Applied Sciences*, 20(13):1135–1162, 1997. 3, 6, 9

[12] M. Kalkbrener. On the stability of gröbner bases under specializations. *Journal of Symbolic Computation*, 24(1):51–58, 1997. 5

[13] Q. Liu and R. Erne. *Algebraic geometry and arithmetic curves*. Oxford university press, 2002. 6

[14] E. Miller and B. Sturmfels. *Combinatorial commutative algebra*, volume 227. Springer, 2005. 5

[15] J. Ponce and M. Hebert. Trinocular geometry revisited. In *CVPR*, 2014. 9

[16] W. Stein et al. Sage Mathematics Software (Version 6.8). http://www.sagemath.org. 3