Rigid Multiview Varieties

Joe Kileel

University of California, Berkeley

January 9, 2016
Nonlinear Algebra
JMM, Seattle
arXiv:1509.03257

Michael Joswig

Bernd Sturmfels

André Wagner
Multiview geometry studies 3D scene reconstruction from images. Foundations in projective geometry. Algebraic vision bridges to algebraic geometry (combinatorial, computational, numerical, ...).

Oct 8–9, 2015, Berlin

May 2–6, 2016, San Jose
A **camera** is a full rank $3 \times 4$ real matrix $A$. Determines a projection $\mathbb{P}^3 \rightarrow \mathbb{P}^2$; $X \mapsto AX$

thought of as taking a picture.

A choice of point $C \in \mathbb{P}^3$ (center), plane $\pi \subset \mathbb{P}^3$ (viewing plane), and coordinates on $\pi$ gives a camera.
Given \( n \) cameras \( A = (A_1, \ldots, A_n) \) in generic position, their **multiview variety** \( V_A \) is the closure of the image of the rational map:

\[
\mathbb{P}^3 \rightarrow \mathbb{P}^2 \times \mathbb{P}^2 \times \cdots \times \mathbb{P}^2
\]

\[
X \mapsto (A_1 X, A_2 X, \ldots, A_n X).
\]
Given \( n \) cameras \( A = (A_1, \ldots, A_n) \) in generic position, their **multiview variety** \( V_A \) is the closure of the image of the rational map:

\[
P^3 \rightarrow P^2 \times P^2 \times \cdots \times P^2
\]

\[
X \mapsto (A_1 X, A_2 X, \ldots, A_n X).
\]

- Space of \( n \) consistent views of one world point.
Given \( n \) cameras \( A = (A_1, \ldots, A_n) \) in generic position, their **multiview variety** \( V_A \) is the closure of the image of the rational map:

\[
P^3 \longrightarrow P^2 \times P^2 \times \cdots \times P^2 \\
X \mapsto (A_1X, A_2X, \ldots, A_nX).
\]

- Space of \( n \) consistent views of one world point.
- Irreducible threefold isomorphic to \( P^3 \) blown-up at \( n \) points.
Given \( n \) cameras \( A = (A_1, \ldots, A_n) \) in generic position, their multiview variety \( V_A \) is the closure of the image of the rational map:

\[
\mathbb{P}^3 \rightarrow \mathbb{P}^2 \times \mathbb{P}^2 \times \cdots \times \mathbb{P}^2 \\
X \mapsto (A_1X, A_2X, \ldots, A_nX).
\]

- Space of \( n \) consistent views of one world point.
- Irreducible threefold isomorphic to \( \mathbb{P}^3 \) blown-up at \( n \) points.
- Prime ideal \( I_A \subset \mathbb{R}[u_{i0}, u_{i1}, u_{i2} : i = 1, \ldots, n] \) is \( \mathbb{Z}^n \)-multihomogeneous.
For which $u_j$ and $u_k$, does:

\[
\begin{cases}
A_j X = \lambda_j u_j \\
A_k X = \lambda_k u_k
\end{cases}
\]

have a nonzero solution in $X, \lambda_j, \lambda_k$? Rewrite as:

\[
B^{jk} \begin{bmatrix} X \\ -\lambda_j \\ -\lambda_k \end{bmatrix} = 0 \quad \text{where} \quad B^{jk} := \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix}_{6 \times 6}
\]
For which $u_j$ and $u_k$, does:

$$\begin{cases} 
A_j X = \lambda_j u_j \\
A_k X = \lambda_k u_k 
\end{cases}$$

have a nonzero solution in $X, \lambda_j, \lambda_k$? Rewrite as:

$$B^{jk} \begin{bmatrix}
X \\
-\lambda_j \\
-\lambda_k
\end{bmatrix} = 0 \quad \text{where} \quad B^{jk} := \begin{bmatrix}
A_j & u_j & 0 \\
A_k & 0 & u_k
\end{bmatrix}_{6 \times 6}$$

**Theorem (Heyden-Aström 1997)**

*For $n \geq 4$, the $\binom{n}{2}$ bilinear forms $\det(B^{jk})$ where $1 \leq j < k \leq n$ cut out $V_A$ set-theoretically.*
For which \( u_j \) and \( u_k \), does:

\[
\begin{align*}
A_j X &= \lambda_j u_j \\
A_k X &= \lambda_k u_k
\end{align*}
\]

have a nonzero solution in \( X, \lambda_j, \lambda_k \)? Rewrite as:

\[
B^{jk} \begin{bmatrix} X \\ -\lambda_j \\ -\lambda_k \end{bmatrix} = 0 \quad \text{where} \quad B^{jk} := \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix}_{6 \times 6}
\]

**Theorem (Heyden-Aström 1997)**

*For* \( n \geq 4 \), the \( \binom{n}{2} \) bilinear forms \( \det(B^{jk}) \) where \( 1 \leq j < k \leq n \) *cut out* \( V_A \) set-theoretically.

**Theorem (Aholt-Sturmfels-Thomas 2013)**

*These* \( \binom{n}{2} \) bilinear forms and \( \binom{n}{3} \) trilinear forms *minimally generate* \( I_A \). *Those* and \( \binom{n}{4} \) quadrilinear forms are a *universal Gröbner basis.*
Rigid multiview variety

Given $n$ cameras $A = (A_1, \ldots, A_n)$ in generic position, their **rigid multiview variety** $W_A$ is the closure of the image of the rational map:

$$
V(Q) \hookrightarrow \mathbb{P}^3 \times \mathbb{P}^3 \dashrightarrow \mathbb{P}^2 \times \mathbb{P}^2 \quad (X,Y) \longmapsto ((A_1 X, \ldots, A_n X), (A_1 Y, \ldots, A_n Y)),
$$

$$
Q(X,Y) = (X_0 Y_3 - Y_0 X_3)^2 + (X_1 Y_3 - Y_1 X_3)^2 + (X_2 Y_3 - Y_2 X_3)^2 - X_3^2 Y_3^2.
$$
Rigid multiview variety

Given \( n \) cameras \( A = (A_1, \ldots, A_n) \) in generic position, their \textbf{rigid multiview variety} \( W_A \) is the closure of the image of the rational map:

\[
\begin{align*}
V(Q) & \hookrightarrow \mathbb{P}^3 \times \mathbb{P}^3 \\
(X, Y) & \longmapsto ((A_1 X, \ldots A_n X), (A_1 Y, \ldots A_n Y)),
\end{align*}
\]

\[
Q(X, Y) = (X_0 Y_3 - Y_0 X_3)^2 + (X_1 Y_3 - Y_1 X_3)^2 + (X_2 Y_3 - Y_2 X_3)^2 - X_3^2 Y_3^2.
\]

Irreducible 5-fold inside \( V_A \times V_A \). Prime ideal \( J_A \) in

\[
\mathbb{R}[u_{i0}, u_{i1}, u_{i2}, v_{i0}, v_{i1}, v_{i2} : i = 1, \ldots, n] \text{ is } \mathbb{Z}^{2n}\text{-multihomogeneous}.
\]
Write $Q(X, Y) = T(X, X, Y, Y)$, where $T(\bullet, \bullet, \bullet, \bullet)$ is a quadrilinear form.

**Theorem (Joswig-K.-Sturmfels-Wagner 2015)**

The octics coming from two pairs of cameras:

$$T(\tilde{\Lambda}_5 B_{i_1}^{j_1 k_1}(u), \tilde{\Lambda}_5 B_{i_2}^{j_1 k_1}(u), \tilde{\Lambda}_5 C_{i_3}^{j_2 k_2}(v), \tilde{\Lambda}_5 C_{i_4}^{j_2 k_2}(v))$$

cut out $W_A$ as a subvariety of $V_A \times V_A$ set-theoretically. For this, 16 suffice.
From two views of one world point $X$, recover $X$ by intersecting back-projected lines. Works unless $X$ is collinear with centers.
From two views of one world point $X$, recover $X$ by intersecting back-projected lines. Works unless $X$ is collinear with centers.

For $1 \leq j < k \leq n$ and $1 \leq i \leq 6$, let:

$$B^{jk}(u) = \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix}_{6 \times 6}$$
From two views of one world point $X$, recover $X$ by intersecting back-projected lines. Works unless $X$ is collinear with centers.

For $1 \leq j < k \leq n$ and $1 \leq i \leq 6$, let:

- $B^{jk}(u) = \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix}_{6 \times 6}$

- $B^i_{jk}(u)$ be the $5 \times 6$ matrix that is $B^{jk}(u)$ with its $i^{th}$ row removed
From two views of one world point $X$, recover $X$ by intersecting back-projected lines. Works unless $X$ is collinear with centers.

For $1 \leq j < k \leq n$ and $1 \leq i \leq 6$, let:

- $B^{jk}(u) = \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix}_{6 \times 6}$
- $B^{jk}_i(u)$ be the $5 \times 6$ matrix that is $B^{jk}(u)$ with its $i^{th}$ row removed
- $\land_5 B^{jk}_i(u)$ be the height 6 column of signed maximal minors of $B^{jk}_i(u)$
From two views of one world point $X$, recover $X$ by intersecting back-projected lines. Works unless $X$ is collinear with centers.

For $1 \leq j < k \leq n$ and $1 \leq i \leq 6$, let:

- $B^{jk}(u) = \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix}_{6 \times 6}$
- $B^{jk}_i(u)$ be the $5 \times 6$ matrix that is $B^{jk}(u)$ with its $i^{th}$ row removed
- $\wedge_5 B^{jk}_i(u)$ be the height 6 column of signed maximal minors of $B^{jk}_i(u)$
- $\wedge_5 B^{jk}_i(u)$ be the height 4 column consisting of the top of $\wedge_5 B^{jk}_i(u)$
Triangulation

From two views of one world point $X$, recover $X$ by intersecting back-projected lines. Works unless $X$ is collinear with centers.

For $1 \leq j < k \leq n$ and $1 \leq i \leq 6$, let:

- $B^{jk}(u) = \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix}_{6 \times 6}$

- $B_i^{jk}(u)$ be the $5 \times 6$ matrix that is $B^{jk}(u)$ with its $i^{th}$ row removed

- $\land_5 B_i^{jk}(u)$ be the height 6 column of signed maximal minors of $B_i^{jk}(u)$

- $\tilde{\land}_5 B_i^{jk}(u)$ be the height 4 column consisting of the top of $\land_5 B_i^{jk}(u)$

- $C^{jk}(v)$, $C_i^{jk}(v)$, $\land_5 C_i^{jk}(v)$ and $\tilde{\land}_5 C_i^{jk}(v)$ be the analogs with $v$. 
Write \( Q(X, Y) = T(X, X, Y, Y) \), where \( T(\bullet, \bullet, \bullet, \bullet) \) is a quadrilinear form.

**Theorem (Joswig-K.-Sturmfels-Wagner 2015)**

The octics coming from two pairs of cameras:

\[
T(\tilde{\wedge}_5 B_{i_1}^{j_1 k_1}(u), \tilde{\wedge}_5 B_{i_2}^{j_1 k_1}(u), \tilde{\wedge}_5 C_{i_3}^{j_2 k_2}(v), \tilde{\wedge}_5 C_{i_4}^{j_2 k_2}(v))
\]

*cut out \( W_A \) as a subvariety of \( V_A \times V_A \) set-theoretically. For this, 16 suffice.*
Write $Q(X, Y) = T(X, X, Y, Y)$, where $T(\bullet, \bullet, \bullet, \bullet)$ is a quadrilinear form.

**Theorem (Joswig-K.-Sturmfels-Wagner 2015)**

The octics coming from two pairs of cameras:

$$T(\bar{\Lambda}_5 B_{i_1}^{j_1 k_1}(u), \bar{\Lambda}_5 B_{i_2}^{j_1 k_1}(u), \bar{\Lambda}_5 C_{i_3}^{j_2 k_2}(v), \bar{\Lambda}_5 C_{i_4}^{j_2 k_2}(v))$$

**cut out** $W_A$ **as a subvariety of** $V_A \times V_A$ **set-theoretically**. **For this, 16 suffice.**

**Sketch.**

These octics vanish on $W_A$. Conversely:
Write \( Q(X, Y) = T(X, X, Y, Y) \), where \( T(\bullet, \bullet, \bullet, \bullet) \) is a quadrilinear form.

**Theorem (Joswig-K.-Sturmfels-Wagner 2015)**

The octics coming from two pairs of cameras:

\[
T\left( \tilde{\wedge}_5 B_{i_1}^{j_1 k_1} (u), \tilde{\wedge}_5 B_{i_2}^{j_1 k_1} (u), \tilde{\wedge}_5 C_{i_3}^{j_2 k_2} (v), \tilde{\wedge}_5 C_{i_4}^{j_2 k_2} (v) \right)
\]

cut out \( W_A \) as a subvariety of \( V_A \times V_A \) set-theoretically. For this, 16 suffice.

**Sketch.**

These octics vanish on \( W_A \). Conversely:

- For \( n \geq 3 \), show one of \( B_1^{12}, B_2^{12}, B_1^{12}, B_2^{13} \) has rank 5, similarly with \( C \).
Write $Q(X, Y) = T(X, X, Y, Y)$, where $T(\bullet, \bullet, \bullet, \bullet)$ is a quadrilinear form.

**Theorem (Joswig-K.-Sturmfels-Wagner 2015)**

The octics coming from two pairs of cameras:

\[
T\left( \tilde{\wedge}_5 B_{i_1}^{j_1k_1}(u), \tilde{\wedge}_5 B_{i_2}^{j_1k_1}(u), \tilde{\wedge}_5 C_{i_3}^{j_2k_2}(v), \tilde{\wedge}_5 C_{i_4}^{j_2k_2}(v) \right)
\]

cut out $W_A$ as a subvariety of $V_A \times V_A$ set-theoretically. For this, 16 suffice.

**Sketch.**

These octics vanish on $W_A$. Conversely:

- For $n \geq 3$, show one of $B_{12}^{12}, B_{22}^{12}, B_{12}^{13}, B_{22}^{13}$ has rank 5, similarly with $C$.
- For $n = 2$, need special geometric argument because of world points collinear with centers.
Conjecture (Joswig-K.-Sturmfels-Wagner 2015)

\( J_A \) is minimally generated by \( \frac{4}{9} n^6 - \frac{2}{3} n^5 + \frac{1}{36} n^4 + \frac{1}{2} n^3 + \frac{1}{36} n^2 - \frac{1}{3} n \) polynomials, coming from two triples of cameras, and their number per symmetry class of degrees is:

<table>
<thead>
<tr>
<th>Symmetry Class</th>
<th>Number of Generators</th>
</tr>
</thead>
<tbody>
<tr>
<td>(110..000..)</td>
<td>( 1 \cdot 2 \binom{n}{2} )</td>
</tr>
<tr>
<td>(220..220..)</td>
<td>( 9 \cdot \binom{n}{2}^2 )</td>
</tr>
<tr>
<td>(111..000..)</td>
<td>( 1 \cdot 2 \binom{n}{3} )</td>
</tr>
<tr>
<td>(220..211..)</td>
<td>( 3 \cdot 2 n \binom{n}{2} \binom{n-1}{2} )</td>
</tr>
<tr>
<td>(220..111..)</td>
<td>( 3 \cdot 2 \binom{n}{2} \binom{n}{3} )</td>
</tr>
<tr>
<td>(211..211..)</td>
<td>( 1 \cdot n^2 \binom{n-1}{2}^2 )</td>
</tr>
<tr>
<td>(211..111..)</td>
<td>( 1 \cdot 2 n \binom{n-1}{2} \binom{n}{3} )</td>
</tr>
<tr>
<td>(111..111..)</td>
<td>( 1 \cdot \binom{n}{3}^2 )</td>
</tr>
</tbody>
</table>
Conjecture (Joswig-K.-Sturmfels-Wagner 2015)

\( J_A \) is minimally generated by \( \frac{4}{9}n^6 - \frac{2}{3}n^5 + \frac{1}{36}n^4 + \frac{1}{2}n^3 + \frac{1}{36}n^2 - \frac{1}{3}n \) polynomials, coming from two triples of cameras, and their number per symmetry class of degrees is:

\[
\begin{align*}
(110..000..) & : 1 \cdot 2\left(\frac{n}{2}\right) \\
(220..220..) & : 9 \cdot \left(\frac{n}{2}\right)^2 \\
(111..000..) & : 1 \cdot 2\left(\frac{n}{3}\right) \\
(220..211..) & : 3 \cdot 2n\left(\frac{n}{2}\right)\left(\frac{n-1}{2}\right) \\
(220..111..) & : 3 \cdot 2\left(\frac{n}{2}\right)\left(\frac{n}{3}\right) \\
(211..211..) & : 1 \cdot n^2\left(\frac{n-1}{2}\right)^2 \\
(211..111..) & : 1 \cdot 2n\left(\frac{n-1}{2}\right)\left(\frac{n}{3}\right) \\
(111..211..) & : 1 \cdot n\left(\frac{n-1}{2}\right)^2 \\
(111..111..) & : 1 \cdot \left(\frac{n}{3}\right)^2
\end{align*}
\]

Computational proof.

Up to \( n = 5 \), when there are 4940 minimal generators.
Generalizations

- Images of four coplanar world points.
- Images of rigid world triangles.
- Proposed approach to images of unlabeled world points.


Thank you!