# 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

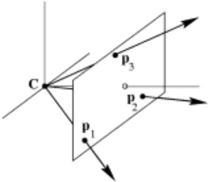
*Multiview geometry* studies 3D scene reconstruction from images. Foundations in projective geometry. *Algebraic vision* bridges to algebraic geometry (combinatorial, computational, numerical, ...).


Oct 8–9, 2015, Berlin


May 2–6, 2016, San Jose

# What is a camera?

A **camera** is a full rank $3 \times 4$ real matrix $A$.
Determines a projection $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$; $X \mapsto AX$




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.

# Multiview variety

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:

$$\begin{array}{ccc} \mathbb{P}^3 & \dashrightarrow & \mathbb{P}^2 \times \mathbb{P}^2 \times \cdots \times \mathbb{P}^2 \\ X & \mapsto & (A_1 X, A_2 X, \ldots, A_n X). \end{array}$$

# Multiview variety

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:

$$\begin{array}{ccc} \mathbb{P}^3 & \dashrightarrow & \mathbb{P}^2 \times \mathbb{P}^2 \times \cdots \times \mathbb{P}^2 \\ X & \mapsto & (A_1X, A_2X, \ldots, A_nX). \end{array}$$

- Space of $n$ consistent views of one world point.

# Multiview variety

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:

$$\begin{array}{ccc} \mathbb{P}^3 & \dashrightarrow & \mathbb{P}^2 \times \mathbb{P}^2 \times \cdots \times \mathbb{P}^2 \\ X & \mapsto & (A_1X, A_2X, \ldots, A_nX). \end{array}$$

- Space of $n$ consistent views of one world point.
- Irreducible threefold isomorphic to $\mathbb{P}^3$ blown-up at $n$ points.

# Multiview variety

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:

$$\begin{array}{ccc} \mathbb{P}^3 & \dashrightarrow & \mathbb{P}^2 \times \mathbb{P}^2 \times \cdots \times \mathbb{P}^2 \\ X & \mapsto & (A_1X, A_2X, \ldots, A_nX). \end{array}$$

- 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.

# Equations

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:

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

# Equations

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:

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

**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**.

# Equations

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:

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

**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)**

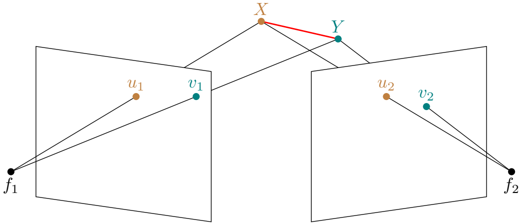
*These $\binom{n}{2}$ bilinear forms and $\binom{n}{3}$ trilinear forms* **minimally generate** *$I_A$. Those and $\binom{n}{4}$ quadrilinear forms are a* **universal Gröbner basis**.

# 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:

$$\begin{array}{ccccc} V(Q) & \hookrightarrow & \mathbb{P}^3 \times \mathbb{P}^3 & \dashrightarrow & (\mathbb{P}^2)^n \times (\mathbb{P}^2)^n \\ (X, Y) & & \longmapsto & & \big((A_1 X, \ldots A_n X), (A_1 Y, \ldots A_n Y)\big), \end{array}$$

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

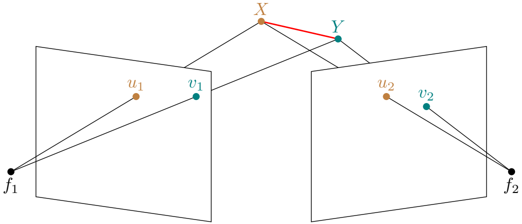


# Rigid multiview variety

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

$$\begin{array}{ccccc} V(Q) & \hookrightarrow & \mathbb{P}^3 \times \mathbb{P}^3 & \dashrightarrow & (\mathbb{P}^2)^n \times (\mathbb{P}^2)^n \\ (X,Y) & & \longmapsto & & \big((A_1X, \ldots A_nX), (A_1Y, \ldots A_nY)\big), \end{array}$$

$$Q(X,Y) = (X_0Y_3 - Y_0X_3)^2 + (X_1Y_3 - Y_1X_3)^2 + (X_2Y_3 - Y_2X_3)^2 - {X_3}^2{Y_3}^2.$$




Irreducible 5-fold inside $V_A \times V_A$. Prime ideal $J_A$ in $\mathbb{R}[u_{i0}, u_{i1}, u_{i2}, v_{i0}, v_{i1}, v_{i2} : i = 1, \ldots, n]$ is $\mathbb{Z}^{2n}$-multihomogeneous.

# Equations

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:*

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

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

# Triangulation

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

# 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:

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

# 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 $i^{\text{th}}$ row removed

# 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 $i^{\text{th}}$ row removed
- $\wedge_5 B_i^{jk}(u)$ be the height 6 column of signed maximal minors of $B_i^{jk}(u)$

# 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 $i^{\text{th}}$ row removed
- $\wedge_5 B_i^{jk}(u)$ be the height 6 column of signed maximal minors of $B_i^{jk}(u)$
- $\widetilde{\wedge}_5 B_i^{jk}(u)$ be the height 4 column consisting of the top of $\wedge_5 B_i^{jk}(u)$

# 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 $i^{\text{th}}$ row removed
- $\wedge_5 B_i^{jk}(u)$ be the height 6 column of signed maximal minors of $B_i^{jk}(u)$
- $\widetilde{\wedge}_5 B_i^{jk}(u)$ be the height 4 column consisting of the top of $\wedge_5 B_i^{jk}(u)$
- $C^{jk}(v)$, $C_i^{jk}(v)$, $\wedge_5 C_i^{jk}(v)$ and $\widetilde{\wedge}_5 C_i^{jk}(v)$ be the analogs with $v$.

# Equations

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:*

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

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

# Equations

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:*

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

*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:

# Equations

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:*

$$T\left(\widetilde{\bigwedge}_5 B_{i_1}^{j_1 k_1}(u)\,,\,\widetilde{\bigwedge}_5 B_{i_2}^{j_1 k_1}(u)\,,\,\widetilde{\bigwedge}_5 C_{i_3}^{j_2 k_2}(v)\,,\,\widetilde{\bigwedge}_5 C_{i_4}^{j_2 k_2}(v)\right)$$

*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 $C$.

# Equations

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:*

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

*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 $C$.
- For $n = 2$, need special geometric argument because of world points collinear with centers. $\square$

**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:*

$$
\begin{array}{cc}
(110..000..) : 1 \cdot 2\binom{n}{2} & (220..111..) : 3 \cdot 2\binom{n}{2}\binom{n}{3} \\
(220..220..) : 9 \cdot \binom{n}{2}^2 & (211..211..) : 1 \cdot n^2\binom{n-1}{2}^2 \\
(111..000..) : 1 \cdot 2\binom{n}{3} & (211..111..) : 1 \cdot 2n\binom{n-1}{2}\binom{n}{3} \\
(220..211..) : 3 \cdot 2n\binom{n}{2}\binom{n-1}{2} & (111..111..) : 1 \cdot \binom{n}{3}^2
\end{array}
$$

# Computations

## 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:*

$$(110..000..) : 1 \cdot 2\binom{n}{2} \qquad (220..111..) : 3 \cdot 2\binom{n}{2}\binom{n}{3}$$
$$(220..220..) : 9 \cdot \binom{n}{2}^2 \qquad (211..211..) : 1 \cdot n^2\binom{n-1}{2}^2$$
$$(111..000..) : 1 \cdot 2\binom{n}{3} \qquad (211..111..) : 1 \cdot 2n\binom{n-1}{2}\binom{n}{3}$$
$$(220..211..) : 3 \cdot 2n\binom{n}{2}\binom{n-1}{2} \qquad (111..111..) : 1 \cdot \binom{n}{3}^2$$

## Computational proof.

Up to $n = 5$, when there are 4940 minimal generators. $\square$

# Generalizations

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

# References


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

[2] D. Grayson and M. Stillman: *Macaulay2, a software system for research in algebraic geometry*, available at `www.math.uiuc.edu/Macaulay2/`.

[3] R. Hartley and A. Zisserman: *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2003.

[4] A. Heyden and K. Åström: *Algebraic properties of multilinear constraints*, Mathematical Methods in the Applied Sciences **20** (1997) 1135–1162.

[5] M. Joswig, J. Kileel, B. Sturmfels and A. Wagner: *Rigid Multiview Varieties*, `arXiv:1509.032571`.

[6] B. Li: *Images of rational maps of projective spaces*, `arXiv:1310.8453`.

[7] E. Miller and B. Sturmfels: *Combinatorial Commutative Algebra*, Graduate Texts in Mathematics, Springer Verlag, New York, 2004.

Thank you!