Distortion Varieties

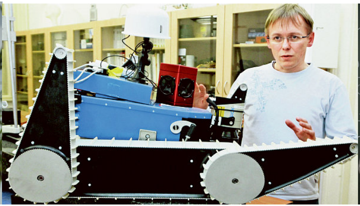
Joe Kileel
UC Berkeley

AMS Sectional @ NCSU
November 12, 2016

Preprint

Distortion Varieties, J. Kileel, Z. Kukelova, T. Pajdla, B. Sturmfels

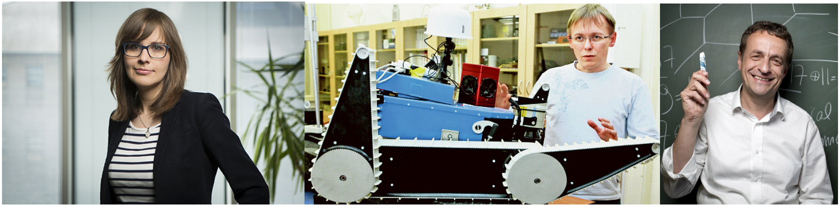
`arXiv:1610.01860`


Preprint

Distortion Varieties, J. Kileel, Z. Kukelova, T. Pajdla, B. Sturmfels

`arXiv:1610.01860`



Today, I want to tell you about our work on lifting varieties in projective spaces to rational normal scrolls. This models configurations of cameras subject to radial distortion.

First the background: 3*D* reconstruction

Example from *Building Rome in a Day* (2009) by S. Agarwal et al.

Input: 2106 `Flickr` images tagged “Colosseum”

First the background: 3*D* reconstruction

Example from *Building Rome in a Day* (2009) by S. Agarwal et al.

Input: 2106 `Flickr` images tagged "Colosseum"

Output: configuration of cameras and 819,242 3*D* points

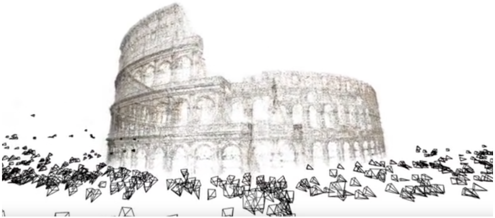


Figure: 3*D* model of the Colosseum in Rome from 2106 `Flickr` images

How does Google do it?

- Identify pairs or triplets of images that overlap.
- Do robust reconstruction with pairs or triplets of images.
- Piece together.

Google solves polynomial systems

- Identify pairs or triplets of images that overlap.
- Do robust reconstruction with pairs or triplets of images.
- Piece together.

For mathematicians: 'Tiny' reconstructions are subroutines in large-scale reconstructions. The 'tiny' reconstructions rely on super-fast, specialized polynomial equation solvers.

[Fischler-Bolles: *Random Sample Consensus: a Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography*, 1981]

[Kúkelová-Bujnak-Pajdla: *Automatic Generator of Minimal Problem Solvers*, 2008]

What is a camera?

A **camera** is a full rank 3×4 real matrix A (up to scale).

Determines a projection $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$; $X \mapsto AX$.

Math	**Interpretation**
$\mathbb{P}^3$	world
$\mathbb{P}^2$	image plane
$\ker(A)$	camera center

Definition

A **configuration** of n cameras is an orbit of the action of the group $\mathrm{SL}(4)$ on the set:

$$\{(A_1, \ldots, A_n) : A_i \text{ is a camera}\}$$

via simultaneous right multiplication.

What is a camera?

A **camera** is a full rank 3×4 real matrix A (up to scale).

Determines a projection $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$; $X \mapsto AX$.

Math	**Interpretation**
$\mathbb{P}^3$	world
$\mathbb{P}^2$	image plane
$\ker(A)$	camera center

Definition

A **configuration** of n cameras is an orbit of the action of the group $\mathrm{SL}(4)$ on the set:

$$\{(A_1, \ldots, A_n) : A_i \text{ is a camera}\}$$

via simultaneous right multiplication.

3D reconstruction: 1) recover configuration, 2) triangulate world scene.

Example: Two Views

$$A_1 = \begin{bmatrix} \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \end{bmatrix} \qquad A_2 = \begin{bmatrix} \diamond & \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond & \diamond \end{bmatrix}$$

The image of $(A_1, A_2) : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2 \times \mathbb{P}^2$ is the hypersurface defined by:

$$f(x,y,z) \quad = \quad \det \begin{bmatrix} \star & \star & \star & \star & x_1 & 0 \\ \star & \star & \star & \star & y_1 & 0 \\ \star & \star & \star & \star & z_1 & 0 \\ \diamond & \diamond & \diamond & \diamond & 0 & x_2 \\ \diamond & \diamond & \diamond & \diamond & 0 & y_2 \\ \diamond & \diamond & \diamond & \diamond & 0 & z_2 \end{bmatrix} \qquad \text{(multiview variety)}$$

This polynomial is bilinear:

$$f(x,y,z) \quad = \quad \begin{bmatrix} x_1 & y_1 & z_1 \end{bmatrix} \cdot \begin{bmatrix} \square & \square & \square \\ \square & \square & \square \\ \square & \square & \square \end{bmatrix} \cdot \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$$

This 3×3-matrix is the *fundamental matrix* of the two cameras.

Fact: it is in bijection with (A_1, A_2) up to $\mathrm{SL}(4)$.

What is its rank?

How to write $\square$ in terms of $\star$ and $\diamond$?

Image Distortion

In real life, images are subject to distortion.

This can happen e.g. if the cameras are in motion while taking the picture, or if the cameras use fisheye lenses for a wide view.

Image Distortion

In real life, images are subject to distortion.

This can happen e.g. if the cameras are in motion while taking the picture, or if the cameras use fisheye lenses for a wide view.

In vision, this is dealt with by introducing distortion parameters to the model for higher order terms. We built a math framework.

Scrolls and Distortions

Fix $u = (u_0, u_1, \ldots, u_n) \in \mathbb{N}^{n+1}$ and $N = |u| + n$. The *rational normal scroll* $\mathcal{S}_u$ is a smooth projective toric variety of dimension $n+1$ and (minimal) degree $|u|$ in $\mathbb{P}^N$. It has the parametrization

$\big(x_0 : x_0\lambda : \cdots : x_0\lambda^{u_0} : x_1 : x_1\lambda : \cdots : x_1\lambda^{u_1} : \ \cdots : x_n : x_n\lambda : \cdots : x_n\lambda^{u_n}\big).$

Ideal given by 2×2-minors of *multi-Hankel matrix* of size $2 \times |u|$.

Example

$n = 2,\ u = (1, 2, 3)$. The scroll $\mathcal{S}_u$ is the threefold in $\mathbb{P}^8$ defined by

$$\begin{pmatrix} a_0 & b_0 & b_1 & c_0 & c_1 & c_2 \\ a_1 & b_1 & b_2 & c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} x_0 & x_1 & x_1\lambda & x_2 & x_2\lambda & x_2\lambda^2 \\ x_0\lambda & x_1\lambda & x_1\lambda^2 & x_2\lambda & x_2\lambda^2 & x_2\lambda^3 \end{pmatrix}$$

Scrolls and Distortions

Fix $u = (u_0, u_1, \ldots, u_n) \in \mathbb{N}^{n+1}$ and $N = |u| + n$. The *rational normal scroll* $\mathcal{S}_u$ is a smooth projective toric variety of dimension $n+1$ and (minimal) degree $|u|$ in $\mathbb{P}^N$. It has the parametrization

$$(x_0 : x_0\lambda : \cdots : x_0\lambda^{u_0} : x_1 : x_1\lambda : \cdots : x_1\lambda^{u_1} : \cdots : x_n : x_n\lambda : \cdots : x_n\lambda^{u_n}).$$

Ideal given by 2×2-minors of *multi-Hankel matrix* of size $2 \times |u|$.

Example

$n = 2$, $u = (1, 2, 3)$. The scroll $\mathcal{S}_u$ is the threefold in $\mathbb{P}^8$ defined by

$$\begin{pmatrix} a_0 & b_0 & b_1 & c_0 & c_1 & c_2 \\ a_1 & b_1 & b_2 & c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} x_0 & x_1 & x_1\lambda & x_2 & x_2\lambda & x_2\lambda^2 \\ x_0\lambda & x_1\lambda & x_1\lambda^2 & x_2\lambda & x_2\lambda^2 & x_2\lambda^3 \end{pmatrix}$$

Consider an arbitrary subvariety X of dimension d in $\mathbb{P}^n$.

Its *distortion variety* $X_{[u]} \subset \mathbb{P}^N$ is the closure of the preimage of X under the map $\mathcal{S}_u \dashrightarrow \mathbb{P}^n$. We have $\dim(X_{[u]}) = d + 1$.

Points on $X_{[u]}$ come from the model X but have been distorted by an unknown parameter λ. Want to recover $x \in X$ and λ from data.

In vision, image data defines linear sections of $X_{[u]}$

Distorting varieties of 3×3-matrices

Basic models X in computer vision have $n = 8$ and $d = 5, 6, 7$.
For two cameras with radial distortion,

$$x = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \qquad \text{and} \qquad u = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

The distortion varieties $X_{[u]}$ live in $\mathbb{P}^{14}$ since $N = n + |u| = 14$.

Example ($d = 7$)

The variety $X = \mathrm{V}(\det)$ of fundamental matrices (configurations of 2 uncalibrated cameras) has degree 3. Its distortion variety $X_{[u]}$ has degree **16** $< 3 \cdot 6$.
Defined by 15 quadrics (from $\mathcal{S}_u$) and 3 cubics (from $\det$).

Degree Formula

Fix $X \subset \mathbb{P}^n$ of codimension c, and $u \in \mathbb{N}^{n+1}$ with $u_0 \leq u_1 \leq \cdots \leq u_n$.

Theorem (KKPS)

The degree of the distortion variety $X_{[u]}$ is the maximum value of the linear functional $w \mapsto u \cdot w$ on the Chow polytope of X. This can be computed by the formula

$$\mathrm{degree}(X_{[u]}) \quad = \quad \sum_{j=0}^{n} u_j \cdot \mathrm{degree}\big(\, \mathrm{in}_{-u}(X) : \langle x_j \rangle^\infty \,\big),$$

where $\mathrm{in}_{-u}(X)$ is an initial monomial ideal of X with respect to $-u$.

Corollary

The distortion variety satisfies

$$\mathrm{degree}(X_{[u]}) \;\leq\; \mathrm{degree}(X) \cdot (u_c + u_{c+1} + \cdots + u_n).$$

Equality holds when X is in general position in $\mathbb{P}^n$.

Equations

Fix $n \in \mathbb{N}$, a vector $u \in \mathbb{N}^{n+1}$, and $N := |u| + n$.

Lemma

The 2×2-minors that define the rational normal scroll $\mathcal{S}_u$ form a Gröbner basis with respect to the diagonal monomial order. The initial monomial ideal is squarefree.

A monomial m is *standard* if it does not lie in this initial ideal. The *weight* of a monomial m is the degree in λ of the monomial in $N+1$ variables that arises from m when substituting in the parametrization of $\mathcal{S}_u$.

Lemma

Consider any monomial $x^\nu = x_0^{\nu_0} x_1^{\nu_1} \cdots x_n^{\nu_n}$ of degree $|\nu|$ in the coordinates of $\mathbb{P}^n$. For any $i \leq \nu \cdot u$ there exists a unique monomial m in the coordinates on $\mathbb{P}^N$ such that m is standard and maps to $x^\nu \lambda^i$ under the parametrization of the scroll $\mathcal{S}_u$.

We call the standard monomial m above the i^{th} *distortion* of x^ν.

Example ($u_0 = 1, u_1 = 2, u_2 = 3$)

Let $x^\nu = x_0^3 x_1^2 x_2^2$. Its various distortions, for $0 \leq i \leq 13$, are the monomials:

$$a_0^3 b_0^2 c_0^2,\ a_0^3 b_0^2 c_0 c_1,\ a_0^3 b_0^2 c_0 c_2,\ a_0^3 b_0^2 c_0 c_3,\ a_0^3 b_0^2 c_1 c_3,\ a_0^3 b_0^2 c_2 c_3,\ a_0^3 b_0^2 c_3^2,$$
$$a_0^3 b_0 b_1 c_3^2,\ a_0^3 b_0 b_2 c_3^2,\ a_0^3 b_1 b_2 c_3^2,\ a_0^3 b_2^2 c_3^2,\ a_0^2 a_1 b_2^2 c_3^2,\ a_0 a_1^2 b_2^2 c_3^2,\ a_1^3 b_2^2 c_3^2.$$

Given any homogeneous polynomial p in $x_0, x_1, \ldots, x_n$ we write $p_{[i]}$ for the polynomial on $\mathbb{P}^N$ that is obtained by replacing each monomial in p by its ith distortion.

Equations

Fix $X \subset \mathbb{P}^n$ and a distortion vector $u \in \mathbb{N}^{n+1}$.

Theorem (KKPS)

The ideal $I(X_{[u]})$ is generated by the $\binom{N-n}{2}$ quadrics that define $\mathcal{S}_u$ together with the distortions $p_{[i]}$ of the elements p in the reduced Gröbner basis of X for a term order that refines the weight $-u$.

Corollary

The ideal $I(X_{[u]})$ is generated by polynomials whose degree is at most the maximal degree of monomial generators of $\mathrm{in}_{-u}(X)$.

Future project: Multi-parameter distortions; lift to different toric variety than $\mathcal{S}_u$. *Tropical geometry* useful.

Thank you!