MEETING ON ALGEBRAIC VISION 2015

SAMEER AGRAWAL, MICHAEL JOSWIG, AND REKHA R. THOMAS

1. Organizers

- Sameer Agrawal (Google Inc.)
- Michael Joswig (TU Berlin)
- Rekha R. Thomas (Washington U)

2. Background

The meeting was be held at the Technical University in Berlin from Oct 8-9, 2015.

Over the last decade, algebraic geometry and polynomial optimization techniques have been used to formulate and solve a number of problems in computer vision. This collaboration has been mostly one-way with computer vision researchers borrowing algebraic geometric and optimization techniques to solve their problems. Much of the vision literature is not known to algebraic geometers and optimization researchers. We believe that all three fields can benefit substantially by talking more to each other and that the time is right for such a collaboration.

To this end, we brought together a small group of experts in computer vision, optimization and algebraic geometry to the first official meeting of an area that we'd like to call *Algebraic Vision*.

The main goal of this meeting was to discuss and establish a core set of problems in computer vision that can benefit from the use of optimization and algebraic geometry and establish a research program for them.

3. Talks

Peter Bürgisser (TU Berlin, Germany)
Numerical condition in polynomial equation solving
Numerical computations are affected by errors (e.g., due to round-off), hence it is important to understand the sensitivity of the results with regard to perturbations. This can be quantified by condition numbers. This concept is well-known in numerical linear algebra, but less so in optimization and polynomial equation solving, even though it plays an important role not only for sensitivity analysis, but also for designing and understanding the complexity of iterative methods. The goal of the talk is to provide a high-level overview of this, focussing on polynomial equation solving.

Didier Henrion (University of Toulouse, France)
SPECTRA: solving exactly linear matrix inequalities
The set of real points such that a symmetric pencil is positive semidefinite is a convex semi-algebraic set called spectrahedron, described by a linear matrix inequality (LMI). After recalling the specific geometric features of spectrahedra and their versatility in convex algebraic geometry, we describe our Maple package SPECTRA that computes an exact algebraic representation of at least one point in a given spectrahedron, or decides that it is empty. In particular, our algorithm does not assume the existence of an interior point, and the computed point minimizes the rank of the pencil on the spectrahedron. Our experiments show significant improvements with respect to state-of-the-art computer algebra algorithms. This is joint work with Simone Naldi and Mohab Safey El Din.

Fredrik Kahl (Chalmers University of Technology, Gothenburg, Sweden)
Critical Configurations for Projective Reconstruction from Multiple Views
In this talk, a classical problem in computer vision is investigated: Given corresponding points in multiple images, when is there a unique projective reconstruction of the 3D geometry of the scene points and the camera positions? A set of points and cameras is said to be critical when there is more than one way of realizing the resulting image points. For two views, it has been known for almost a century that the critical configurations consist of points and camera lying on a ruled quadric surface. We give a classification of all possible critical configurations for any number of points in three images, and show that in most cases, the ambiguity extends to any number of cameras. The theoretical results are accompanied by many examples and illustrations.

Tomas Pajdla (Czech Technical University, Prague, Czech Republic)
Minimal Problems - What works and what does not
We will review some our recent developments in solving minimal problems in computer vision (radial distortion with homography, rolling shutter absolute camera pose, L2 three view triangulation) and will point to issues that seem to be difficult for us and could motivate further development of algebraic methods for solving computer vision problems.

Pablo Parrilo (MIT, Cambridge, USA)
Chordal structure and polynomial systems

Marc Pollefeys (ETH Zurich, Switzerland)
(radial) multi-focal tensors concept, internal constraints and geometric insight

Mohab Safey El Din (Uni. Pierre et Marie Curie, Paris, France)
Solving polynomial systems over the reals exactly: why and how?
Polynomial system solving appear in many problems of engineering sciences, including computer vision. Most of the time, the end-user expects to get some information on the real solution set of such systems. We will explain why algebraic computation (i.e. computer algebra) becomes an essential ingredient for solving some families of such systems, the main geometric ideas on which these algorithms rely and how to use them efficiently.

Josef Schicho (RICAM, Linz, Austria)
Photogrammetry Profiles and Classical Invariant Theory
For a fixed object, its profile is defined as the set of all possible image (from all possible camera positions). Typically, images are defined up to some equivalence (e.g. projective transformations). Therefore the profiles can naturally be embedded in quotient varieties. Tasks in photogrammetry, such as identifying an object from a fixed number of images, reduce to interpolating specific subvarities in these quotient varieties. In this talk we explain this connection with several examples (Segre cubic, Igusa quartic).

Bernd Sturmfels (University of California, Berkeley, USA)
Varieties, Parameters, and Moduli
Multiview varieties, camera matrices, and various tensors play a central role in algebraic vision. This lecture introduces these objects from a perspective that is entirely natural to those studying algebraic geometry.

René Vidal (Johns Hopkins University, Baltimore, USA)
Subspace Arrangements in Vision and Learning
The problem of clustering data drawn from multiple linear subspaces can be posed as one of estimating and decomposing a subspace arrangement into its irreducible components. Over the past decade, this problem has found widespread applications in computer vision, including image and video segmentation, face and digit clustering, and hybrid system identification. This lecture will review the classical GPCA algorithm for solving this problem, which is based on estimating and factorizing multivariate homogeneous polynomials. The lecture will also present extensions to the classical GPCA algorithm that can handle data corrupted by noise and outliers by constructing a descending subspace filtration. Applications to motion segmentation will also be presented.

4. Posters

- Tomáš Bajbar (Karlsruher Institut für Technologie)
- Diego Cifuentes (Massachusetts Institute of Technology)
- Joe Kileel (University of California, Berkeley)
- André Wagner (TU Berlin)

Rigid Multiview Varieties

Michael Joswig, Joe Kileel*, Bernd Sturmfels, André Wagner*

Multiview Variety

What is the space of pictures of one world point?

Given n generic 3×4 camera matrices $A_1, \ldots, A_n$.

Multiview map ▶

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

Multiview variety ▶ $V_A := \overline{\mathrm{im}(\phi_A)} \subseteq (\mathbb{P}^2)^n$.
Irreducible three-fold.

Multiview ideal ▶ $I_A := I(V_A) \subseteq \mathbb{R}[u_{i0}, u_{i1}, u_{i2} : i = 1, \ldots, n]$.
$\mathbb{Z}^n$-multihomogeneous prime ideal in a polynomial ring with $3n$ variables.
Here $(u_{i0} : u_{i1} : u_{i2})$ are homogeneous coordinates on the i^{th} $\mathbb{P}^2$.

Linear system ▶ For which u_j and u_k, does:

$$\begin{cases} A_jX = \lambda_j u_j \\ A_kX = \lambda_k u_k \end{cases}$$

have a nonzero solution in X, λ_j, λ_k ? Rewrite as:

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

Bilinear equations ▶ For all $1 \leq j < k \leq n$, $\det(B^{jk}) \in I_A$.
It equals $u_j^T F_{jk} u_k$, where F_{jk} is the fundamental matrix.

Theorem 1 (Heyden-Åström 1997)

For $n \geq 4$, the $\binom{n}{2}$ bilinear forms cut out V_A **set-theoretically**:

$$V_A = V\left(u_j^T F_{jk} u_k : \forall j, k\right).$$

Trilinear equations ▶ Maximal minors of $B^{jk\ell} := \begin{bmatrix} A_1 & u_1 & 0 & 0 \\ A_2 & 0 & u_2 & 0 \\ A_3 & 0 & 0 & u_3 \end{bmatrix}_{9\times 7}$

Quadrilinear equations ▶ Maximal minors of $B^{jk\ell m} := \begin{bmatrix} A_1 & u_1 & 0 & 0 & 0 \\ A_2 & 0 & u_2 & 0 & 0 \\ A_3 & 0 & 0 & u_3 & 0 \\ A_4 & 0 & 0 & 0 & u_4 \end{bmatrix}_{12\times 8}$

Theorem 2 (Aholt-S.-Thomas 2013)

- $\binom{n}{2}$ binilinear and $\binom{n}{3}$ trilinear forms above **minimally generate** I_A.
- These bilinear, trilinear, quadrilinear forms are **universal Gröbner basis**.

Rigid Multiview Variety

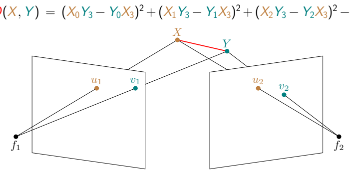
What is the space of pictures of two distance-constrained world points?

Rigid multiview map ▶

$$\begin{aligned} \psi_A :\ V(Q) \hookrightarrow \mathbb{P}^3 \times \mathbb{P}^3 &\dashrightarrow (\mathbb{P}^2)^n \times (\mathbb{P}^2)^n, \\ (X, Y) &\mapsto ((A_1X, \ldots A_nX), (A_1Y, \ldots A_nY)). \end{aligned}$$

where

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



Rigid multiview variety ▶ $W_A := \overline{(\mathrm{im}(\psi_A))} \subseteq \mathbb{P}^{2n}$.
Irreducible 5-fold inside $V_A \times V_A$.

Rigid multiview ideal ▶ $J_A := I(V_A) \subseteq \mathbb{R}[u_{i0}, u_{i1}, u_{i2}, v_{i0}, v_{i1}, v_{i2} : i = 1, \ldots, n]$.
$\mathbb{Z}^{2n}$-multihomogeneous prime ideal in a polynomial ring with $6n$ variables.

Triangulate with Cramer's Rule ▶ For $1 \leq j < k \leq n$ and $1 \leq i \leq 6$, let:

- $B_i^{jk}(u)$ be the 5×6 matrix that is $B^{jk}(u)$ with its i^{th} row removed
- $\widetilde{\wedge}_5 B_i^{jk}(u)$ be the height 6 column of signed maximal minors of $B_i^{jk}(u)$
- $C_i^{jk}(v)$ and $\widetilde{\wedge}_5 C_i^{jk}(v)$ 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 3 (J.-K.-S.-W. 2015)

The octics coming from two pairs of cameras:

$$T\left(\widetilde{\wedge}_5 B_{i_1}^{j_1k_1}, \widetilde{\wedge}_5 B_{i_2}^{j_1k_1}, \widetilde{\wedge}_5 C_{i_3}^{j_2k_2}, \widetilde{\wedge}_5 C_{i_4}^{j_2k_2}\right)$$

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

Ideals ▶ Above octics together with $I_A(u) + I_A(v)$ do not generate J_A.

Conjecture 4 (J.-K.-S.-W. 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 class of degrees is:

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

Computational proof ▶ Up to $n = 5$, when there are 4940 minimal generators.

Other Constraints, More Points, and No Labels

More points, one polynomial constraint ▶ Take n pictures of m world points constrained by a single irreducible multihomogeneous polynomial equation $Q(X^{(1)}, \ldots, X^{(m)}) = 0$. Then **Theorem 3 holds verbatim**: to cut out the image set-theoretically, equations from pairs of cameras suffice. For example if $m = 4$ and $Q = \det(X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)})$, the constraint is four points in $\mathbb{P}^3$ are coplanar, and $16\binom{n}{2}^2$ polynomials cut out set-theoretically.

More rigid points ▶ Impose distances between all pairs of m world points:

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

When $m = 3$, the image of $V(Q_{ij} : \forall i, j)$ in $(\mathbb{P}^2)^{mn}$ is six-dimensional unless:

$$(d_{12} + d_{13} + d_{23})(d_{12} + d_{13} - d_{23})(d_{12} - d_{13} + d_{23})(-d_{12} + d_{13} + d_{23}) = 0.$$

It is cut out by $27\binom{n}{2}^2$ biquadratics set-theoretically, coming from pairs of points and pairs of cameras.

No labels on world points ▶ Suppose images of m world points are **unlabeled**.
Future work: Study the *unlabeled multiview variety*, i.e. the image of:

$$(\mathbb{P}^3)^m \dashrightarrow ((\mathbb{P}^2)^m)^n \to (\mathrm{Sym}_m(\mathbb{P}^2))^n.$$

Here $\mathrm{Sym}_m(\mathbb{P}^2)$ is the *Chow variety* of ternary forms that are products of m linear forms. Some known equations for it inside the space $\mathbb{P}^{\binom{m+2}{2}-1}$ of all ternary forms of degree m are Brill's equations.

References

- C. Aholt, B. Sturmfels and R. Thomas: *A Hilbert scheme in computer vision*, Canadian Journal of Mathematics **65** (2013) 961–988.
- A. Heyden and K. Åström, *Algebraic properties of multilinear constraints*, Mathematical Methods in the Applied Sciences **20** (1997) 1135–1162.
- M. Joswig, J. Kileel, B. Sturmfels and A. Wagner, *Rigid Multiview Varieties*, `arXiv:1509.032571`.

5. Discussion Session

Sameer Agarwal. Two view reconstruction Problem. That is to reconstruct cameras A_1, A_2 and world points X_i, from given data

$$x_i \in \mathbb{R}^2,$$
$$y_i \in \mathbb{R}^2$$

such that

$$\sum \left(||x_i - \hat{x}_i||^2 + ||y_i - \hat{y}_i||^2 \right)$$

is minimized as a function of x_i, y_i and F, with respect to

$$\hat{y}_i^\top F \hat{x}_i = 0, \text{ rk}(F) = 2.$$

If only 7 image points per picture are given, then $x_i = \hat{x}_i$, $y_i = \hat{y}_i$ and the reprojection error is zero.

A version with a robust error norm might be more interesting for applications.

Max Lieblich. What is the functor of points of multiview geometry? What is an interesting parameter space?

Josef Schicho. How to reconstruct an algebraic curve from it's picture $\text{Disc}_z(F(x, y, z)$ and the degree of the curve. For example:

$$\begin{aligned} F(x, y, z, w) &= x^2 + y^2 + z^2 - w^2 &= 0 \\ \text{Disc}_z(x, y, w) &= x^2 + y^2 - w^2 &= 0 \end{aligned}$$

Sameer Agarwal. Given seven image point correspondences $(x_i, y_i) \in (\mathbb{R}^2 \times \mathbb{R}^2)$ when are there three different fundamental matrices (rank 2 Matrices) F_1, F_2, F_3, such that:

$$\begin{aligned} \begin{bmatrix} y_i \\ 1 \end{bmatrix}^\top F_1 \begin{bmatrix} x_i \\ 1 \end{bmatrix} &= 0 \\ \begin{bmatrix} y_i \\ 1 \end{bmatrix}^\top F_2 \begin{bmatrix} x_i \\ 1 \end{bmatrix} &= 0 \\ \begin{bmatrix} y_i \\ 1 \end{bmatrix}^\top F_3 \begin{bmatrix} x_i \\ 1 \end{bmatrix} &= 0 \end{aligned}$$

is satisfied for all seven correspondences.

Michael Joswig. Consider the classical 7-point algorithm plus RANSAC. How to select the 7-tuple efficiently?

SUPPORTED BY:

ECMath
Einstein Center
for Mathematics Berlin

Algorithmic and Experimental Methods
in Algebra, Geometry, and Number Theory
DFG Priority Project SPP 1489