Multiview varieties and reconstruction problems

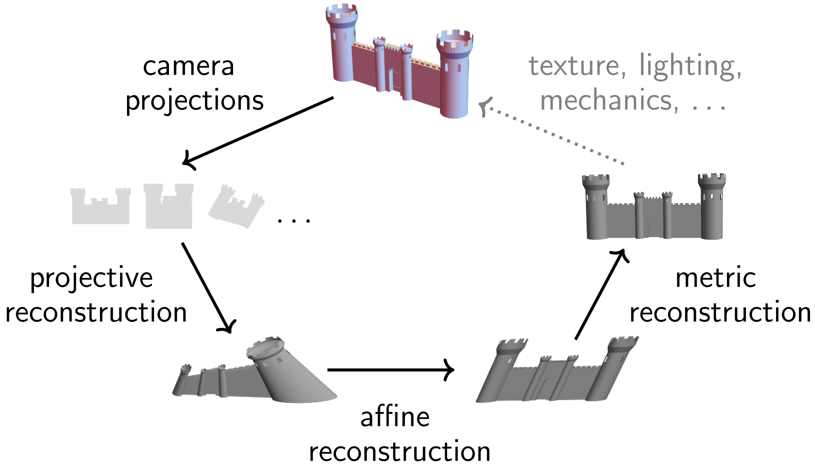
Makoto Miura

joint work with Atsushi Ito and Kazushi Ueda

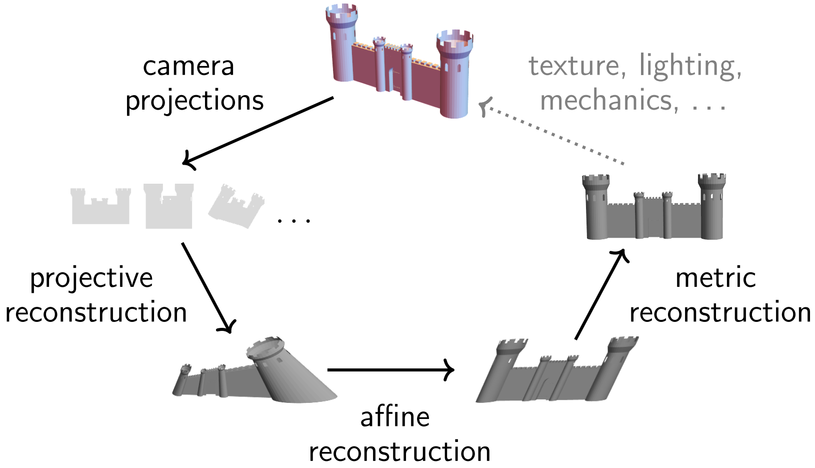
Korea Institute for Advanced Study

October 6, 2018

Reconstruction problems in computer vision

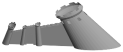
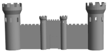


Reconstruction problems in computer vision



How to describe the problems in arbitrary dimensions?

Model: world spaces

projective	affine	metric
		
$(\mathbb{P}^n, \operatorname{Aut}\mathbb{P}^n)$	$(\mathbb{R}^n, \operatorname{Aff}\mathbb{R}^n)$	$(\mathbb{R}^n, \operatorname{Sim}\mathbb{R}^n)$

- $\mathbb{P}^n$: projective space over $\mathbb{R}$
- $\mathbb{R}^n = \mathbb{P}^n \setminus H$: Euclidean space as a set
- $\operatorname{Aut}\mathbb{P}^n \simeq PGL(n+1)$
- $\operatorname{Aff}\mathbb{R}^n = \{g \in \operatorname{Aut}\mathbb{P}^n \mid g \cdot H \subset H\} \simeq GL(n) \ltimes \mathbb{R}^n$
- $\operatorname{Sim}\mathbb{R}^n = \{g \in \operatorname{Aff}\mathbb{R}^n \mid g \cdot Q \subset Q\} \simeq \mathbb{R}^\times O(n) \ltimes \mathbb{R}^n$

$H \in |\mathcal{O}_{\mathbb{P}^n}(1)|$ — *hyperplane at infinity*

smooth definite $Q \in |\mathcal{O}_H(2)|$ — *absolute quadric*

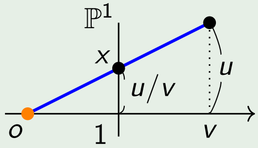
Model: cameras

$s\colon \mathbb{R}^{n+1} \twoheadrightarrow \mathbb{R}^{m+1}$: a surjective linear map $(n > m)$

$l_x \subset \mathbb{R}^{m+1}$: line corresponding to $x \in \mathbb{P}^m$

$\bar{s}\colon \mathbb{P}^n \dashrightarrow \mathbb{P}^m$	— *camera projection*
$\mathbb{P}(\ker s) \subset \mathbb{P}^n$	— *focal locus* (dim : $n - m - 1$)
$\mathbb{P}(s^{-1}(l_x)) \subset \mathbb{P}^n$	— *back-projected plane* (dim : $n - m$)

Example (Pinhole camera model)



$$H = \{w = 0\} \supset Q = \{u^2 + v^2 = 0\}$$

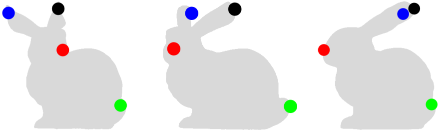
$$\bar{s}\colon \mathbb{P}^2 \ni \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} u \\ v \end{bmatrix} \equiv \begin{bmatrix} u/v \\ 1 \end{bmatrix} \in \mathbb{P}^1$$

$\bar{s}$: *realistic* $\Leftrightarrow$ $n = 3$ and $m = 2$

Additional information for reconstruction

$\bar{s}_i \colon \mathbb{P}^n \dashrightarrow \mathbb{P}^{m_i}$ $(i = 1, \ldots, r)$: camera projections

- point correspondences



$\varphi := (\bar{s}_1, \ldots, \bar{s}_r) \colon \mathbb{P}^n \dashrightarrow \prod_i \mathbb{P}^{m_i}$
$\varphi(u) = (\bar{s}_1(u), \ldots, \bar{s}_r(u))$ stands for a correspondence.

- camera motions
- prior knowledge of the scene

Multiview varieties

$X_\varphi = \varphi(\mathbb{P}^n)$ — *multiview variety*
(moduli space of point correspondences)

$\varphi(\mathbb{P}^n) := \overline{\varphi(\mathbb{P}^n \setminus Z)} \subset \prod_i \mathbb{P}^{m_i}$: the image of φ

Z: the union of focal loci $Z_1, \ldots, Z_r$

Multiview varieties

$X_\varphi = \varphi(\mathbb{P}^n)$ — *multiview variety*
(moduli space of point correspondences)

$\varphi(\mathbb{P}^n) := \overline{\varphi(\mathbb{P}^n \setminus Z)} \subset \prod_i \mathbb{P}^{m_i}$: the image of φ

Z: the union of focal loci $Z_1, \ldots, Z_r$

Questions

- How to describe X_φ?
- Can we recover φ (up to $\operatorname{Aut} \mathbb{P}^n$) from $X_\varphi \subset \prod_i \mathbb{P}^{m_i}$?

Description 1: via back-projected planes

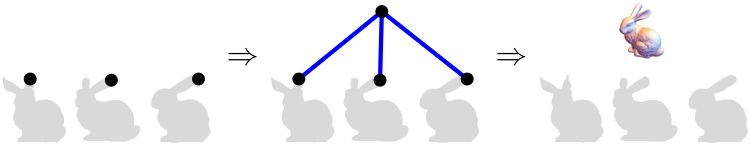
Each $x = (x_1, \ldots, x_r) \in \prod_i \mathbb{P}^{m_i}$ corresponds to an r-tuple of back-projected planes $(P_1, \ldots, P_r)$, where $P_i = \overline{\bar{s}_i^{-1}(x_i)}$.

Description 1

Assume φ is generic.

$$X_\varphi = \{x \in \textstyle\prod_i \mathbb{P}^{m_i} \mid \bigcap_i P_i \neq \varnothing\}.$$

Plotting $\varphi^{-1} \colon X_\varphi \dashrightarrow \mathbb{P}^n$ is refered to as *triangulation*:



Description 2: via Grassmann tensors

Assume φ is generic, and $|\mathbf{m}| := \sum_i m_i > n$.

Write $\mathbf{s} := (s_1, \ldots, s_r)\colon V \hookrightarrow \bigoplus_i W_i$.

Fix $\alpha \in \mathbb{Z}^r$ such that $1 \leq \alpha_i \leq m_i$ and $\sum_i \alpha_i = n+1$.

$\mathrm{pr}_\alpha\colon \bigwedge^{n+1} \bigoplus_i W_i \simeq \bigoplus_\beta \left(\bigotimes_i \bigwedge^{\beta_i} W_i\right) \twoheadrightarrow \bigotimes_i \bigwedge^{\alpha_i} W_i$

$[A^{\sigma_1,\ldots,\sigma_r}] := \left[\mathrm{pr}_\alpha \bigwedge^{n+1} \mathbf{s}(V)\right]$ — *Grassmann tensor* of profile α

$(U_1, \ldots, U_r) \in \prod_i Gr(m_i - \alpha_i, \mathbb{P}(W_i)) \xrightarrow{\prod_i p^i} \prod_i \mathbb{P}(\bigwedge^{\alpha_i} W_i^*)$

Description 2

Assume φ is generic. For any profile α,

$$X_\varphi \cap \prod_i U_i \neq \varnothing \Leftrightarrow \sum_{\sigma_1,\ldots,\sigma_r} A^{\sigma_1,\ldots,\sigma_r} p^1_{\sigma_1} \ldots p^r_{\sigma_r} = 0.$$

Moduli spaces for $|\mathbf{m}| > n$

$$\begin{array}{ccc}
\prod_i \mathbb{P}(V^* \otimes W_i) & & \\
\Big\downarrow{\scriptstyle /\!\!/ PGL(V)} & & \\
\mathrm{Cam} := Gr(n+1, \bigoplus_i W_i) /\!\!/ \mathbb{G}_m^r & \overset{\gamma}{\dashrightarrow} & \mathrm{Hilb}\left(\prod_i \mathbb{P}(W_i)\right) \\
{\scriptstyle \text{Plücker}}\Big\downarrow & \overset{\pi_\alpha}{\searrow} & \\
\mathbb{P}\left(\bigwedge^{n+1} \bigoplus_i W_i\right) /\!\!/ \mathbb{G}_m^r & \underset{\overline{\mathrm{pr}}_\alpha}{\dashrightarrow} & \mathbb{P}\left(\bigotimes_i \bigwedge^{\alpha_i} W_i\right)
\end{array}$$

Cam	moduli space of camera projections
$\mathcal{X} := \gamma(\mathrm{Cam})$	—"— of multiview varieties
$\mathcal{A}_\alpha := \pi_\alpha(\mathrm{Cam})$	—"— of Grassmann tensors of profile α

Examples for realistic cameras:

$\mathcal{A}_{2,2}, \mathcal{A}_{2,1,1}, \mathcal{A}_{1,1,1,1}$: the *epipolar*, *trifocal*, *quadrifocal* varieties

Projective reconstruction theorem

$\mathbf{m} = (m_1, \ldots, m_r)$, $\alpha = (\alpha_1, \ldots, \alpha_r)$

Theorem

Assume φ is generic, and $|\mathbf{m}| > n$.

- *(Hertley–Schaffalitzky '09)*

  If $\mathbf{m} \neq (1^{n+1})$, $\pi_\alpha\colon \mathrm{Cam} \dashrightarrow \mathcal{A}_\alpha$ is generically injective.

  If $\mathbf{m} = (1^{n+1})$, π_α is generically $2:1$.

- *(Aholt–Sturmfels–Thomas '13, Ito–M–Ueda '17+)*

  If $\mathbf{m} \neq (1^{n+1})$, $\gamma\colon \mathrm{Cam} \dashrightarrow \mathcal{X}$ is generically injective.

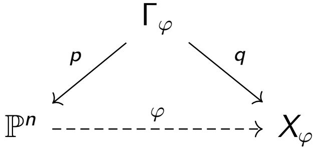
  If $\mathbf{m} = (1^{n+1})$, γ is generically $2:1$.

  If $|\mathbf{m}| \geq 2n-1$, $\mathcal{X}$ is an irreducible component of $\mathrm{Hilb} \prod_i \mathbb{P}^{m_i}$.

Projective reconstruction for $\mathbf{m} \neq (1^{n+1})$

Assume φ is generic, and $|\mathbf{m}| > n$.

$\Gamma_\varphi := (\mathrm{id} \times \varphi)(\mathbb{P}^n) \subset \mathbb{P}^n \times \prod_i \mathbb{P}^{m_i}$: the graph of φ



$$\mathcal{L}_1 = \iota^* \mathrm{pr}_1^* \mathcal{O}_{\mathbb{P}^{m_1}}(1)$$
$$E_1 = p^{-1}(Z_1)$$
$$D_1 = q(E_1) = \mathbb{P}^{m_1} \times \text{something}$$

- $\mathrm{Sing}\, X_\varphi = \{\dim \bigcap_i P_i \geq 1\}$: singular locus
- X_φ is normal and q is small.
- If $\mathbf{m} \neq (1^{n+1})$, D_1 is uniquely determined only from X_φ.

Theorem (Ito–M–Ueda '17+)

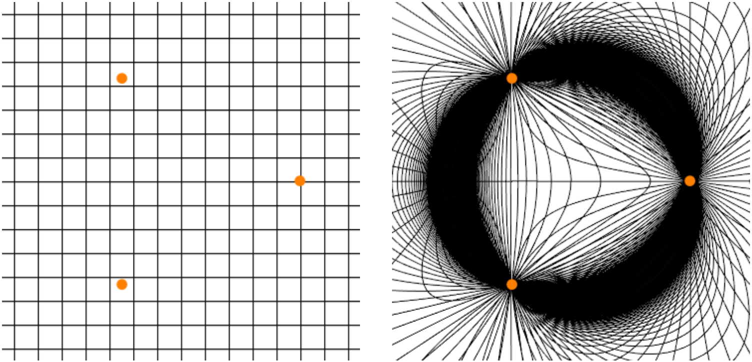
$\varphi^{-1} \colon X_\varphi \dashrightarrow \mathbb{P}^n$ *is given by* $|\mathcal{O}(D_1) \otimes \mathcal{L}_1|$.

Projective reconstruction for $\mathbf{m} = (1^{n+1})$

$\mathbf{s}' \colon \left(\bigoplus_i W_i / \mathbf{s}(V) \right)^* \hookrightarrow \bigoplus_i W_i^*$ gives the dual reconstruction.

Theorem (Ito–M–Ueda '17+)

$\varphi^{-1} \circ \varphi' \in \operatorname{Bir} \mathbb{P}^n$ *is given by* $|\mathcal{O}(n) \otimes I_{Z_1 \cup \cdots \cup Z_{n+1}}|$.

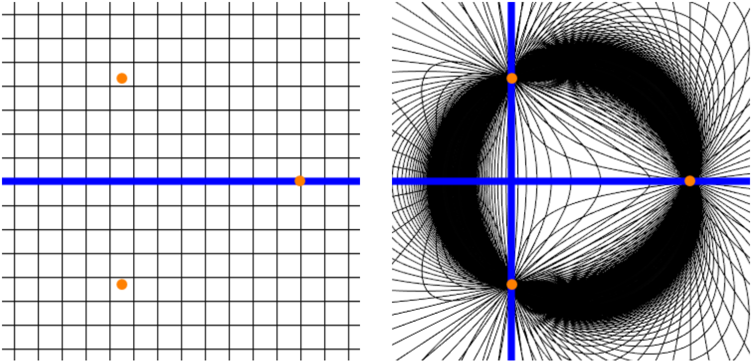


Projective reconstruction for $\mathbf{m} = (1^{n+1})$

$\mathbf{s}' \colon \left(\bigoplus_i W_i/\mathbf{s}(V)\right)^* \hookrightarrow \bigoplus_i W_i^*$ gives the dual reconstruction.

Theorem (Ito–M–Ueda '17+)

$\varphi^{-1} \circ \varphi' \in \mathrm{Bir}\,\mathbb{P}^n$ is given by $|\mathcal{O}(n) \otimes I_{Z_1 \cup \cdots \cup Z_{n+1}}|$.

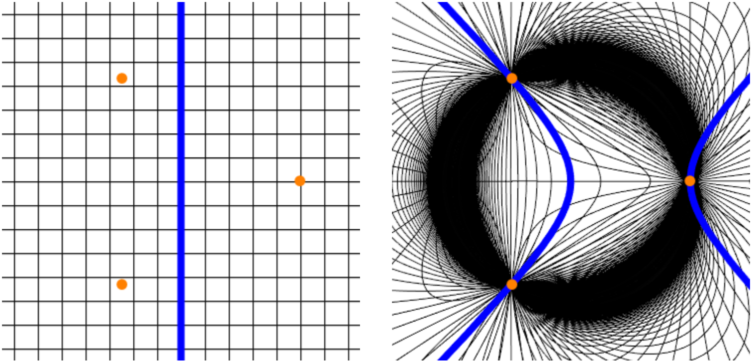


Projective reconstruction for $\mathbf{m} = (1^{n+1})$

$\mathbf{s}' \colon \left(\bigoplus_i W_i/\mathbf{s}(V)\right)^* \hookrightarrow \bigoplus_i W_i^*$ gives the dual reconstruction.

Theorem (Ito–M–Ueda '17+)

$\varphi^{-1} \circ \varphi' \in \mathrm{Bir}\,\mathbb{P}^n$ is given by $|\mathcal{O}(n) \otimes I_{Z_1 \cup \cdots \cup Z_{n+1}}|$.

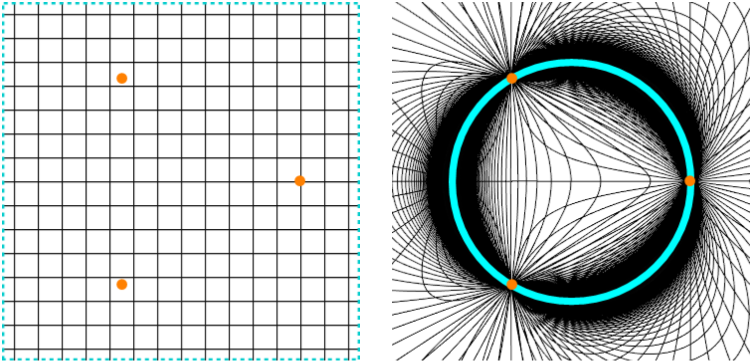


Projective reconstruction for $\mathbf{m} = (1^{n+1})$

$\mathbf{s}' \colon \left(\bigoplus_i W_i/\mathbf{s}(V)\right)^* \hookrightarrow \bigoplus_i W_i^*$ gives the dual reconstruction.

Theorem (Ito–M–Ueda '17+)

$\varphi^{-1} \circ \varphi' \in \mathrm{Bir}\,\mathbb{P}^n$ *is given by* $|\mathcal{O}(n) \otimes I_{Z_1 \cup \cdots \cup Z_{n+1}}|$.



Affine reconstruction theorem

$H \in \mathcal{O}_{\mathbb{P}^n}(1)$: hyperplane at infinity

$T(H) := \{g \in \mathrm{Aut}(\mathbb{P}^n) \mid g = \mathrm{id} \text{ or } (\mathbb{P}^n)^g = H\} \simeq \mathbb{R}^n$
— the group of *pure translations* on $(\mathbb{R}^n = \mathbb{P}^n \setminus H, \mathrm{Aff}\,\mathbb{R}^n)$

$\bar{s}_1, \bar{s}_2$: related by a G-motion $\Leftrightarrow \mathbb{P}^{m_1} = \mathbb{P}^{m_2}$ and $\bar{s}_2 \in \bar{s}_1 G$

$\Delta \subset \mathbb{P}^m \times \mathbb{P}^m$: the diagonal set

Theorem (Ito–M–Ueda, in preparation)

$\varphi = (\bar{s}_1, \bar{s}_2)\colon \mathbb{P}^n \dashrightarrow \mathbb{P}^m \times \mathbb{P}^m$: *related by a* $T(H)$*-motion*

Assume $X_\varphi \not\subset \Delta$, *then* H *is uniquely reconstructed from* φ.

The affine structure is compatible with the factoring,

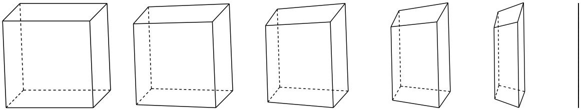
$$\mathbb{P}(V)^n \dashrightarrow \mathbb{P}\left(V/(\ker s_1 \cap \ker s_2)\right)^{m+1} \dashrightarrow \mathbb{P}(W)^m \times \mathbb{P}(W)^m.$$

Summary

- Two formulations for projective reconstruction are obtained for a generic φ and $|\mathbf{m}| > n$.
- Affine reconstruction is described by some degenerate φ and $|\mathbf{m}| \leq n$ in general.

Future topics:

- Degenerate configurations
- Metric reconstruction problems (the difference between $\mathbb{R}$ and $\mathbb{C}$ may give difficulty.)



rotating cube in $\mathbb{R}^4$