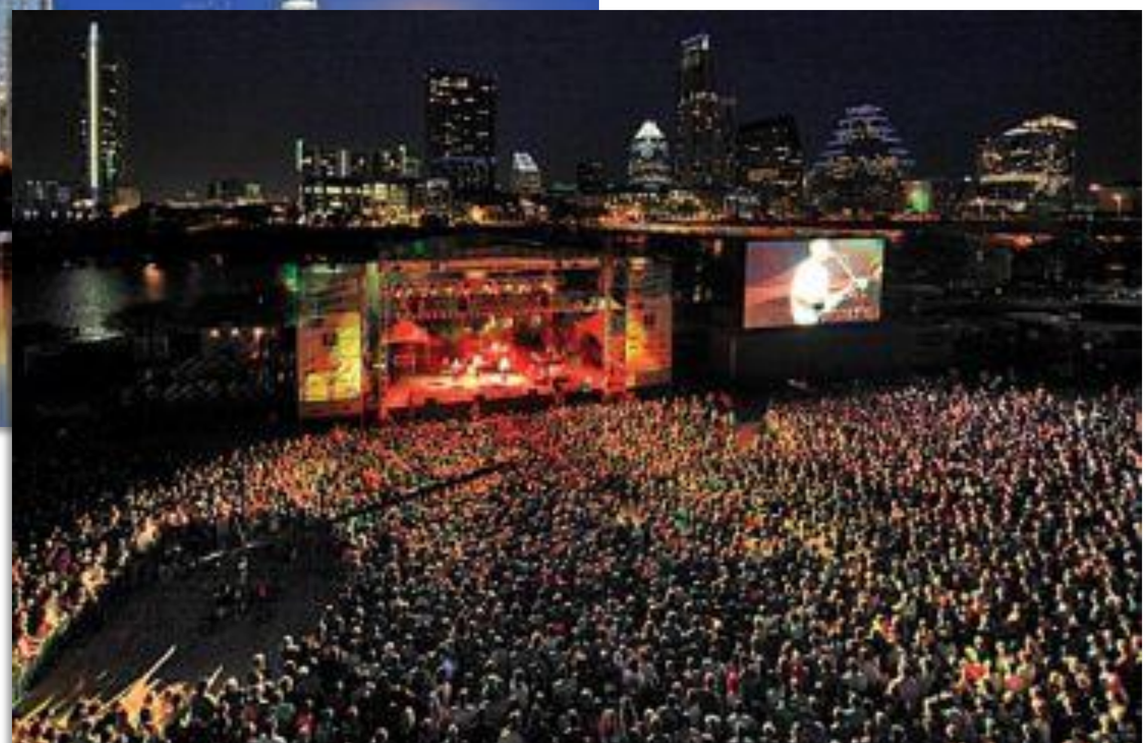
Map Synchronization: from Object Correspondences to Neural Networks

Qixing Huang

March 7th 2019

Qixing Huang

Live in Austin



The University of Texas at Austin

Computer Science Department

Ranking

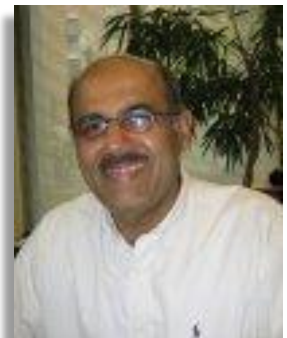
#10 in **Theory**

The only top-10 program within 1000 miles

Computer Science Department

- 45 tenured or tenure-track faculties
- Plan to grow to 60 faculties in 5 years

Computer Graphics at UT



Chandrajit Bajaj

Don Fussell

Etienne Vouga

Qixing Huang

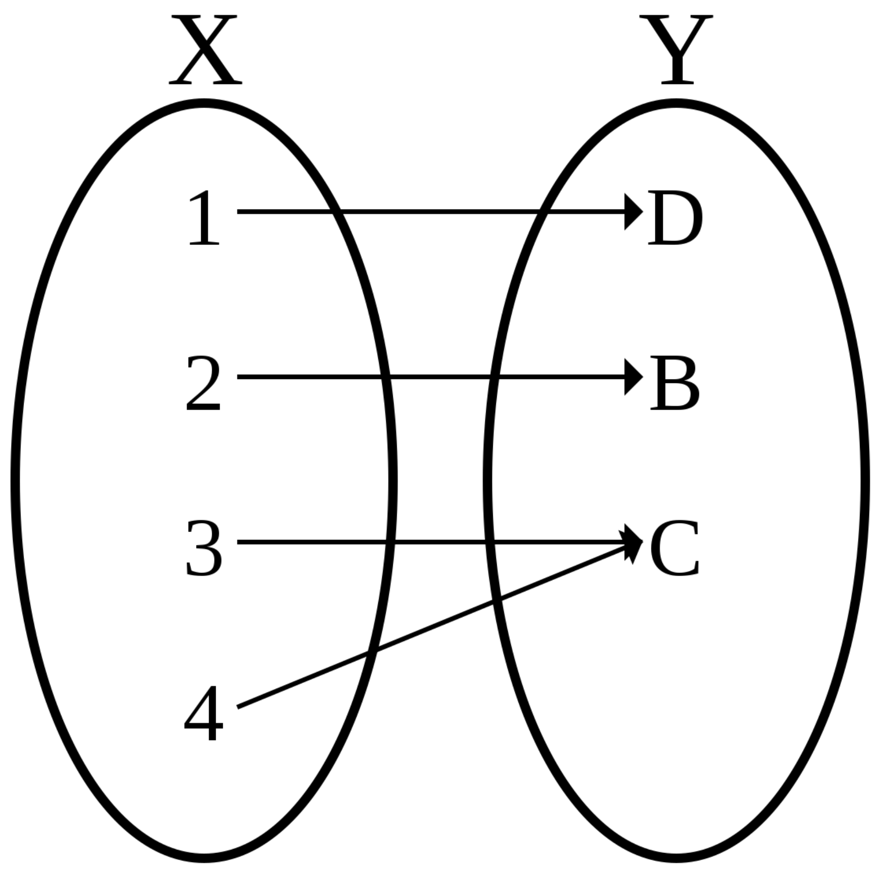
Vision group is growing --- two more members in the next two years

Kristen Grauman

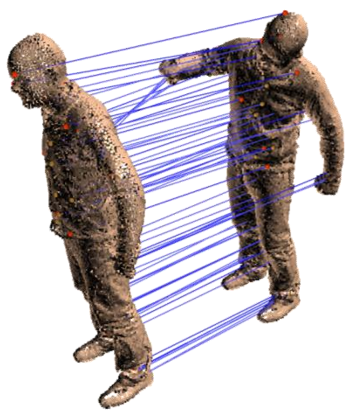
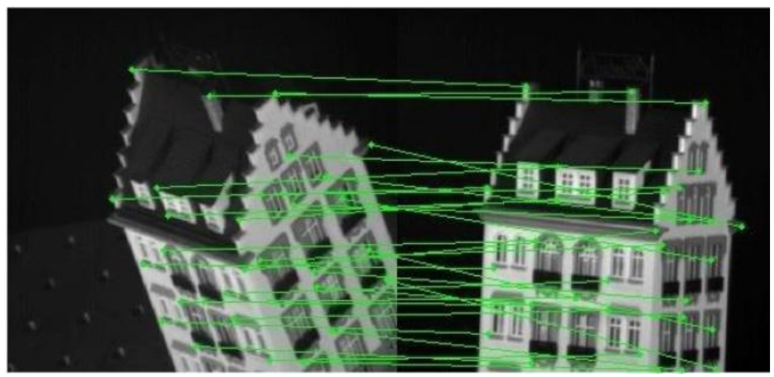
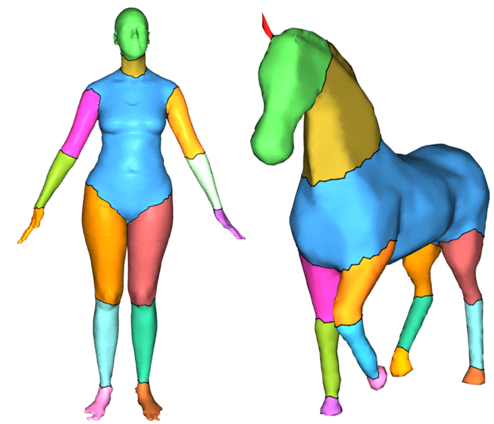
Philipp Krähenbühl

Qixing Huang

Maps between Sets



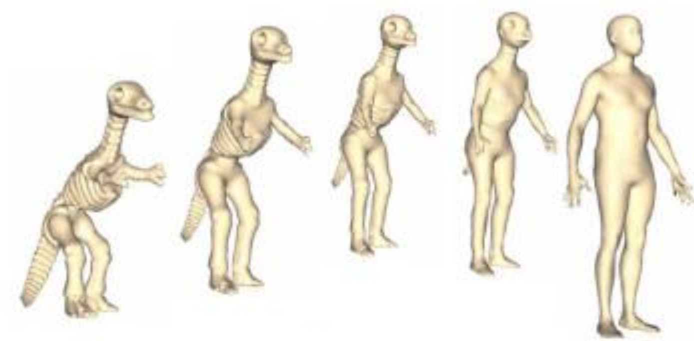
Maps between objects

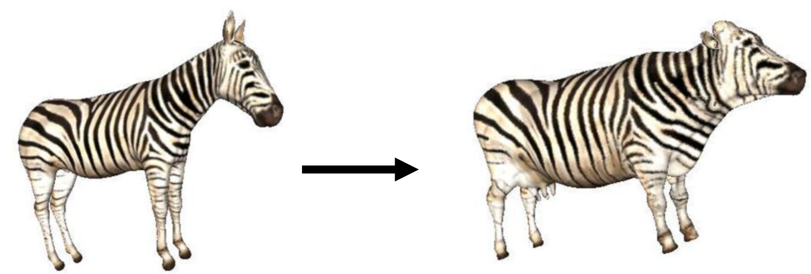
Dense correspondences between images pixels/mesh vertices

Sparse correspondences between features points/parts/segments

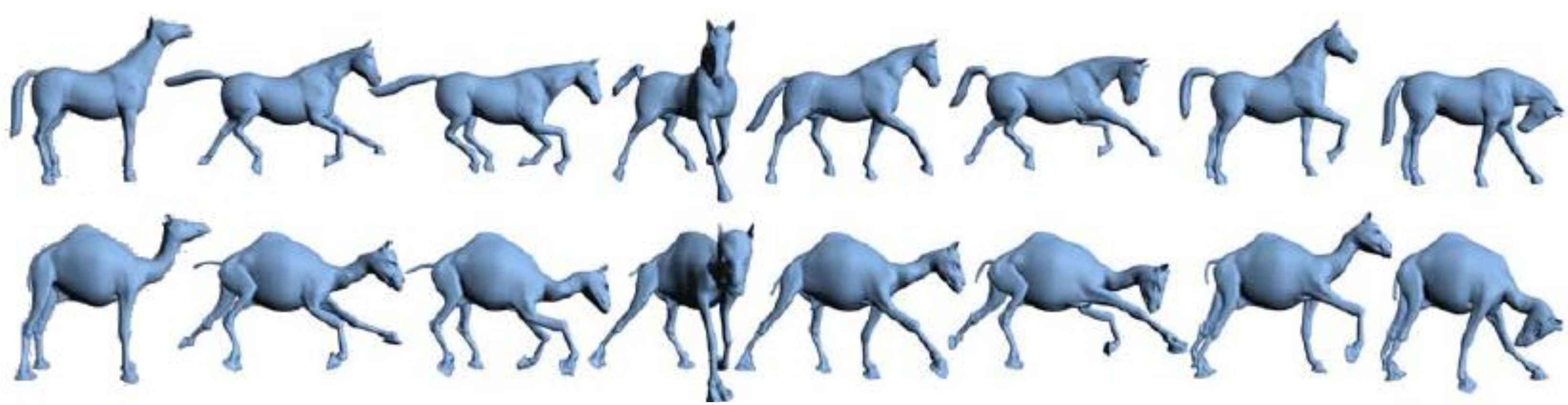
Application in information propagation



Morphing [Zhang et al. 08]

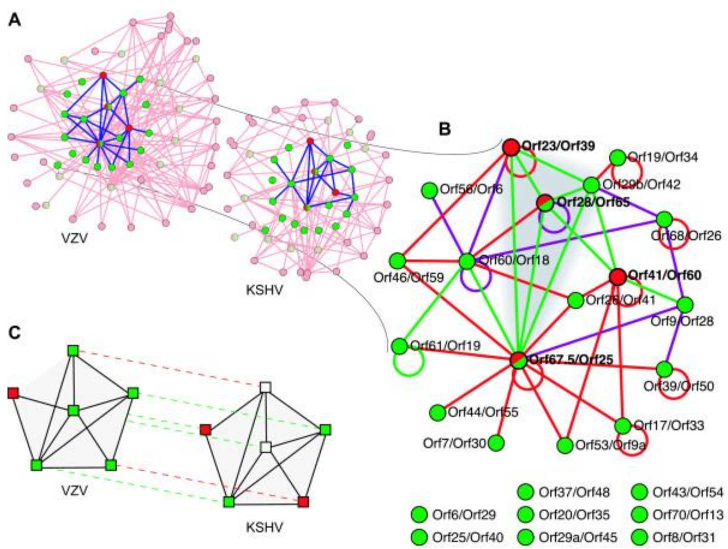


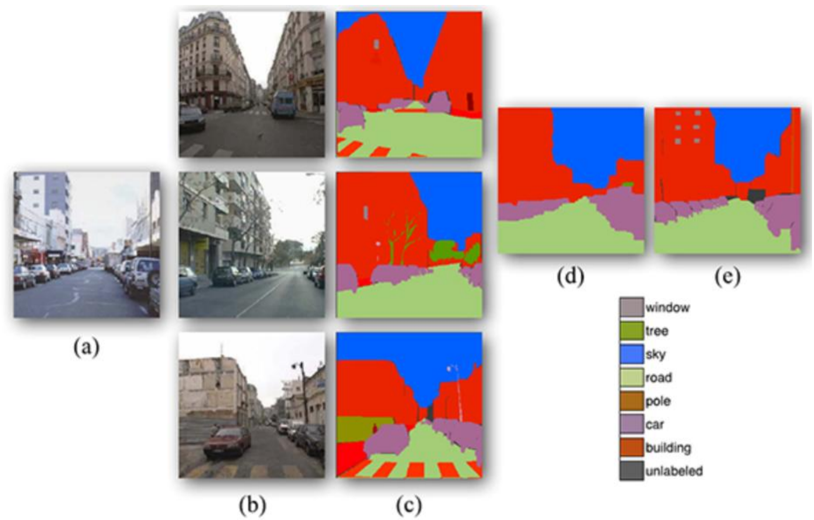
Texture transfer [Chen et al. 12]



Deformation transfer [Sumner and Popovic 04]

Application in information propagation

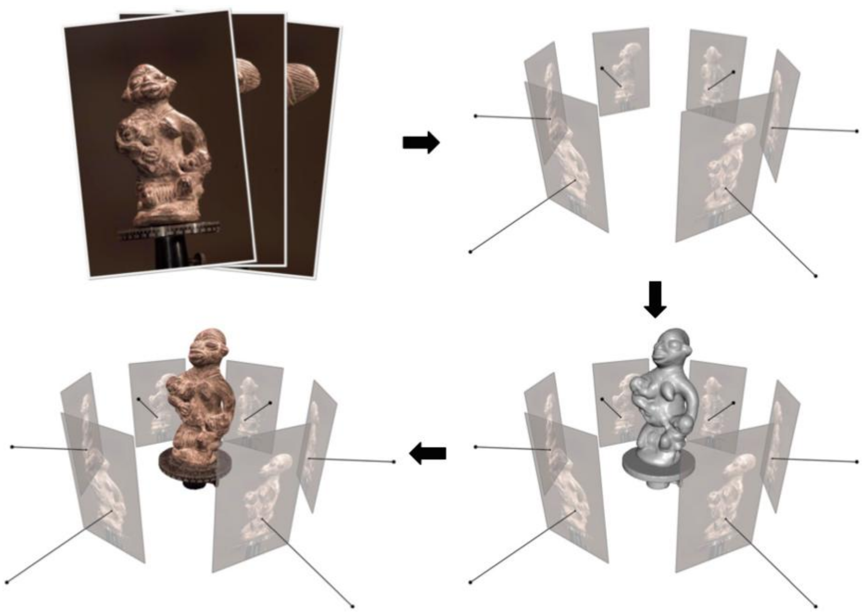


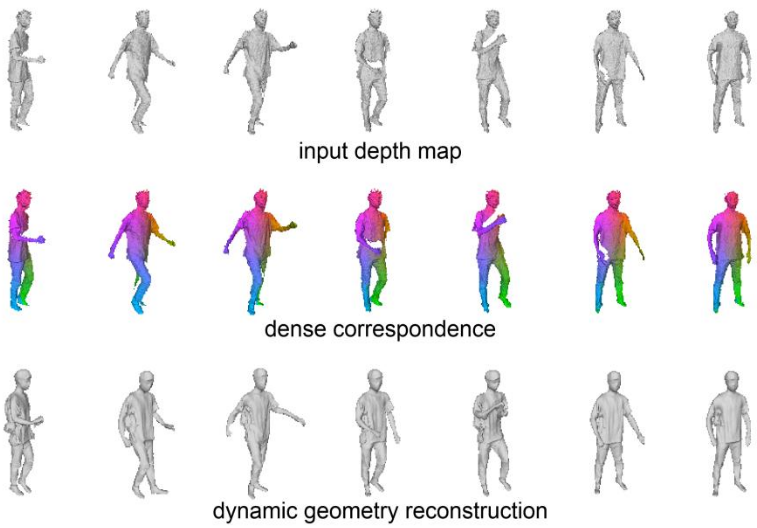


Protein-protein interaction network alignment
[Kolar et al. 08]

Nonparametric Scene Parsing
[Liu et al. 11]

Application in 3D Reconstruction



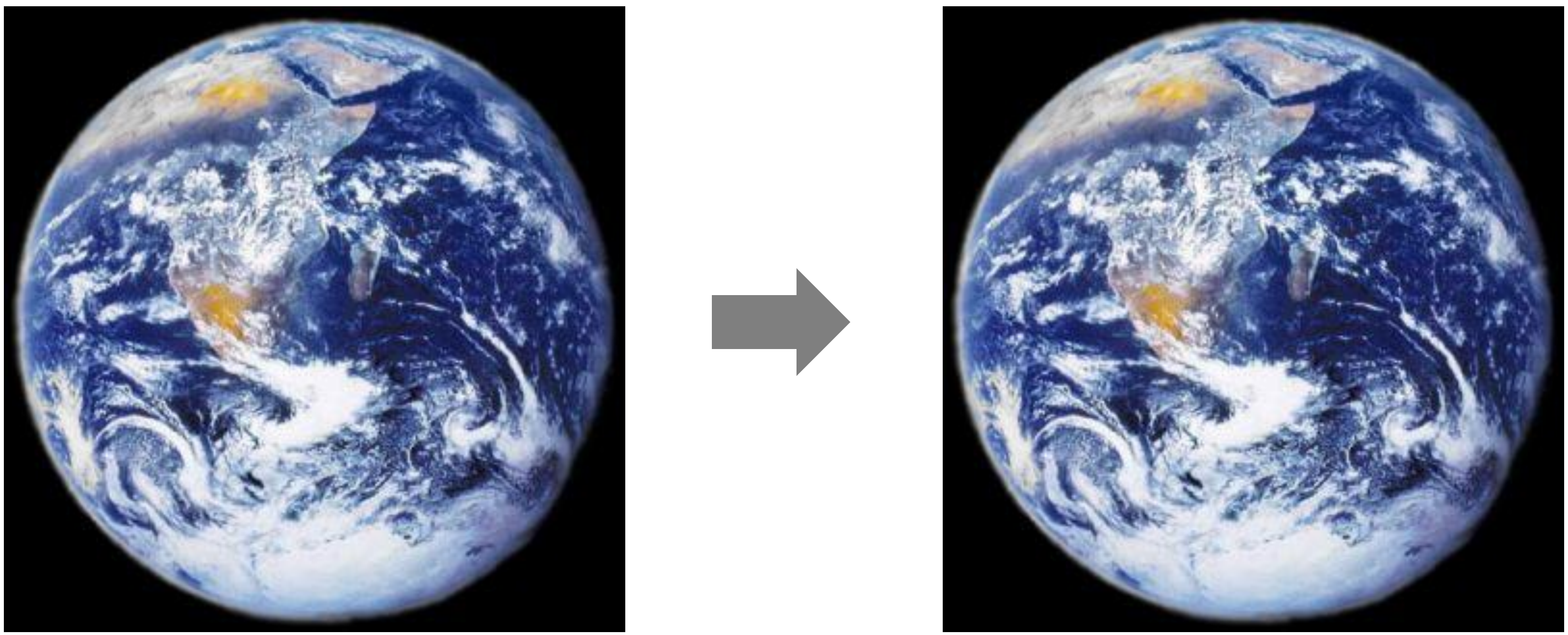


Multiview Stereo
[Furukawa and Hernandez 15]

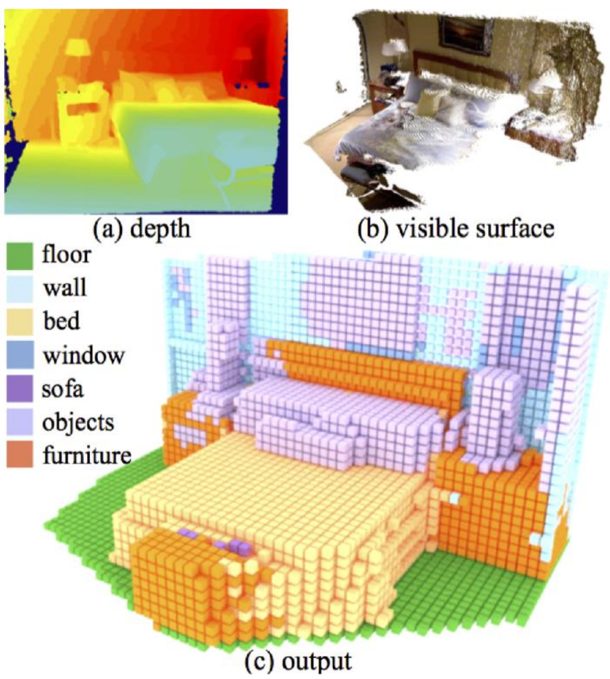
Dynamic geometry reconstruction
[Li et al. 15]

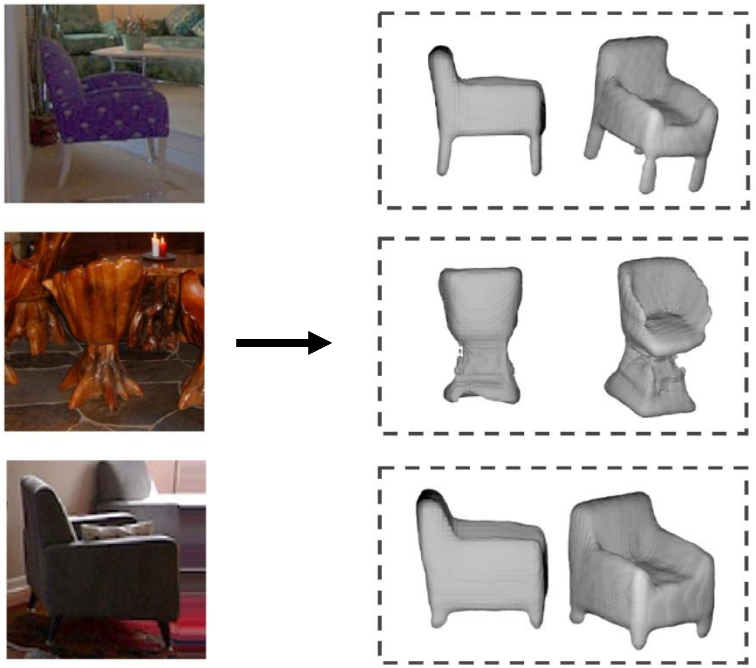
Neural networks are maps

- Approximate any function given sufficient data



Monocular reconstruction





Semantic scene completion [Song et al. 17]

MarrNet [Wu et al. 17]

Space of images → Space of 3D models

Image Captioning

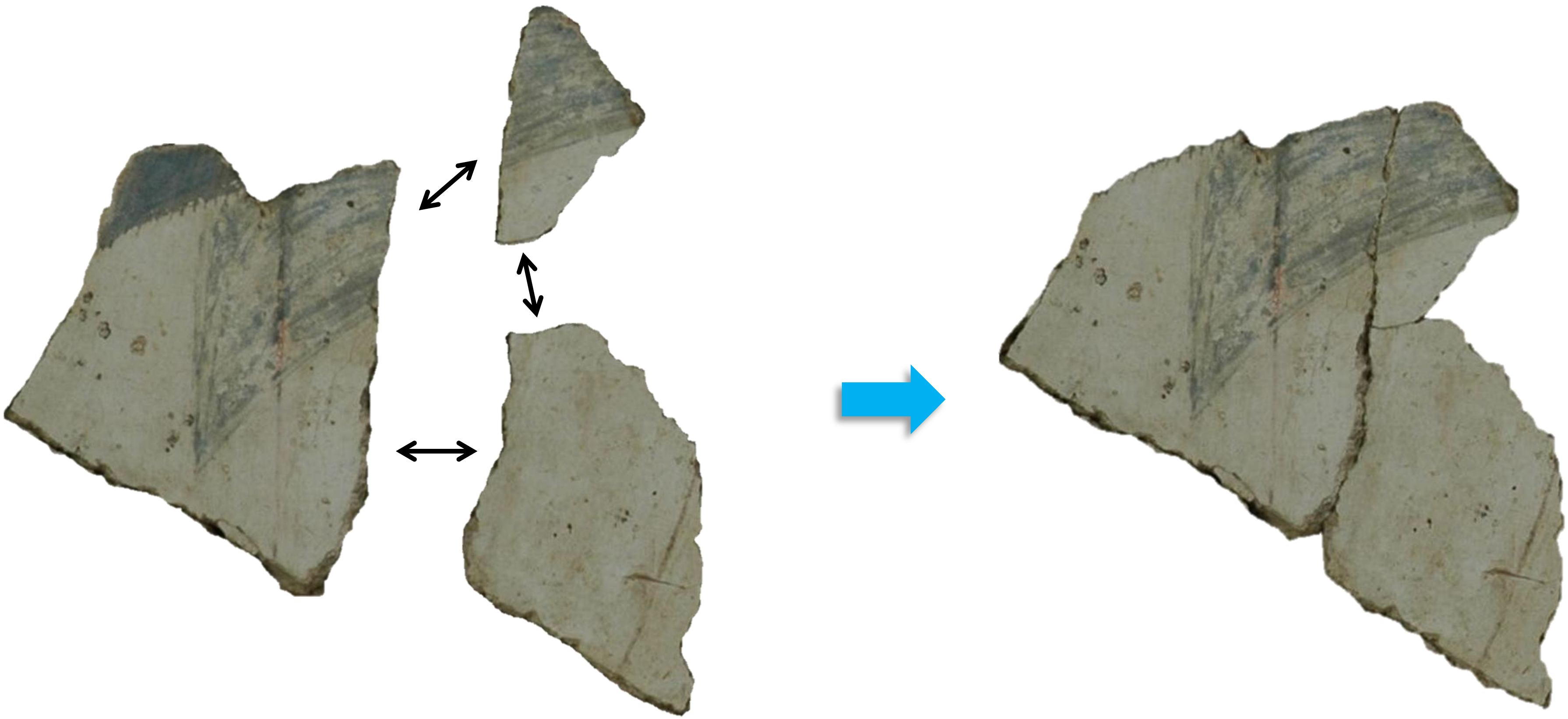
Space of images → Space of natural language descriptions

Map Computation is Difficult

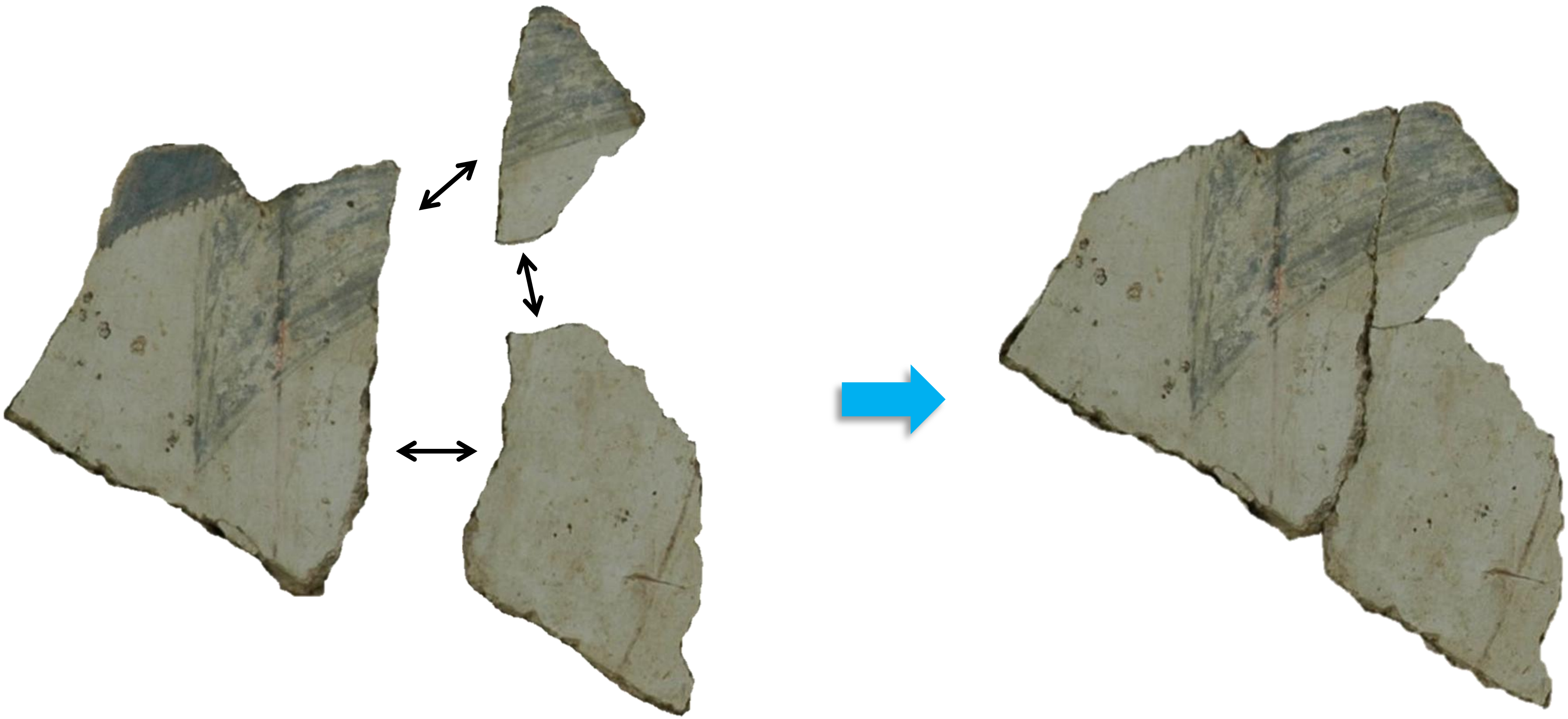
Joint Map Computation
(Map Synchronization)

Ambiguities in assembling pieces

Resolving ambiguities by looking at additional pieces

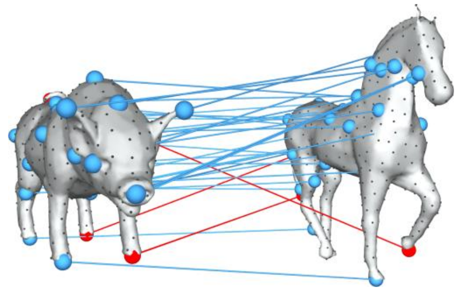


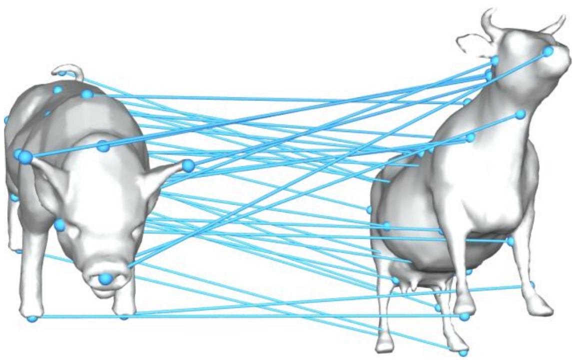
Resolving ambiguities by looking at additional pieces



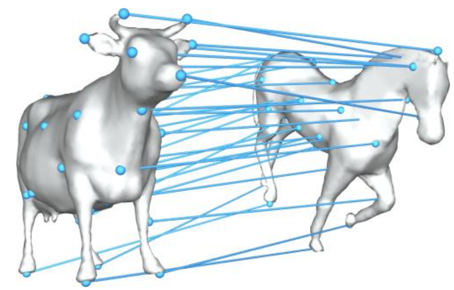
Matching through intermediate objects --- map propagation

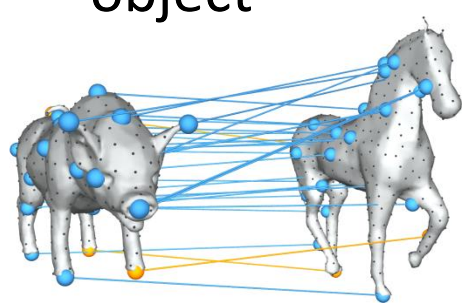
Blended intrinsic maps
[Kim et al. 11]





← Intermediate object →

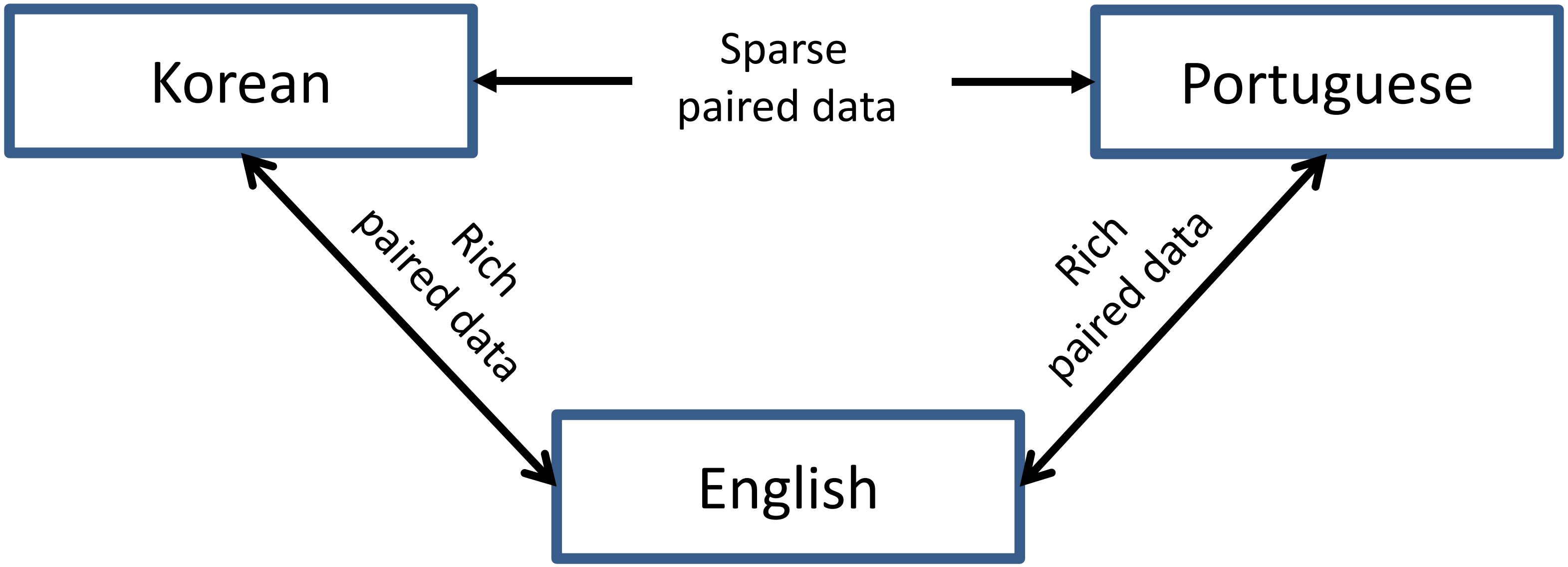




Composite

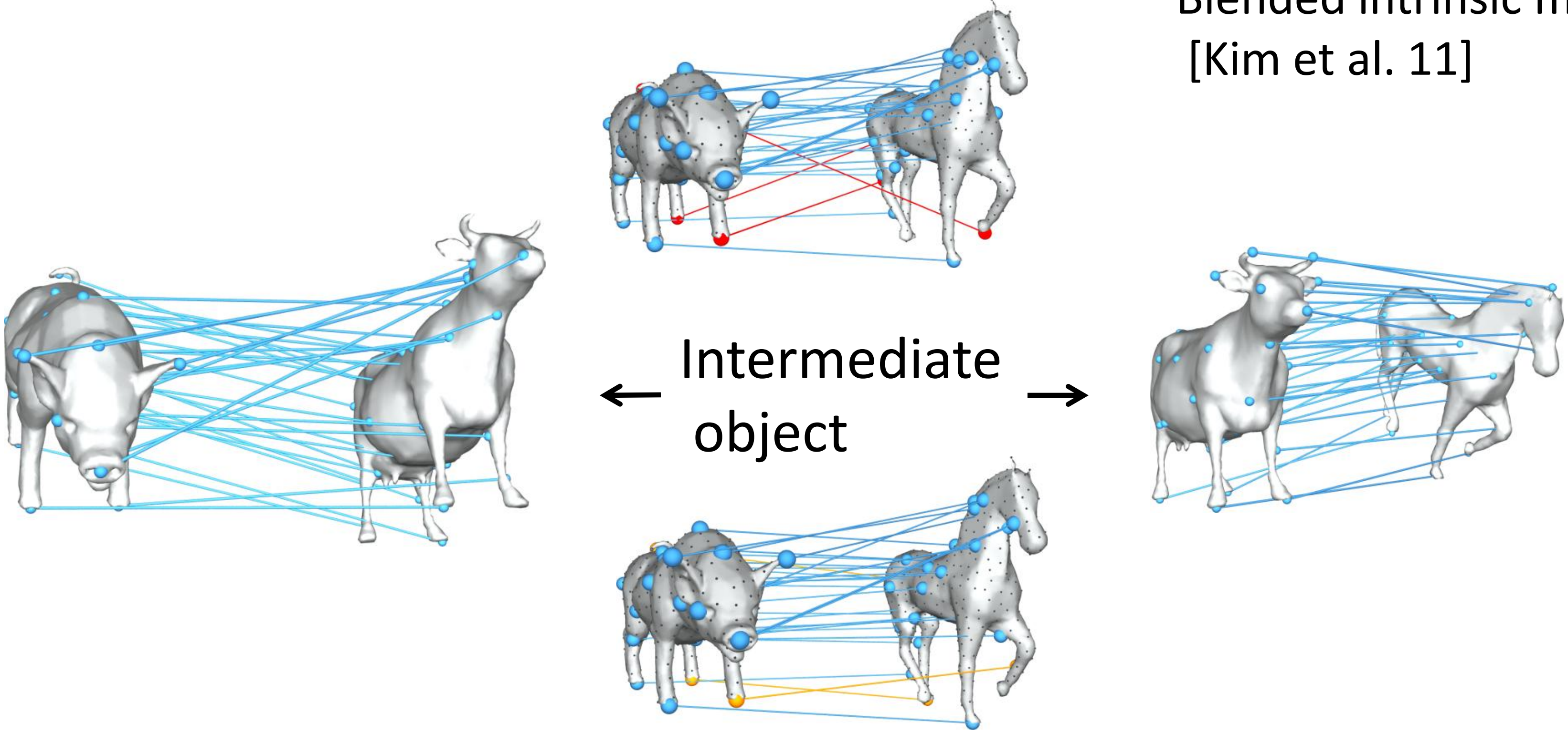
Multi-lingual translation

[Johnson et al. 16]



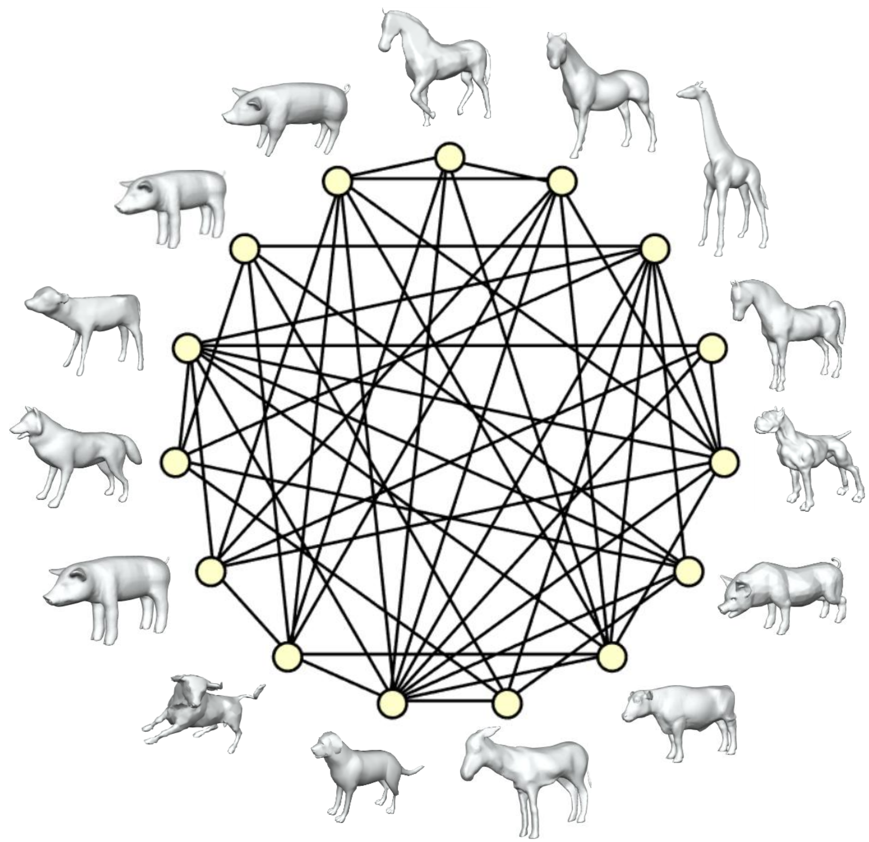
Matching through intermediate objects --- map propagation

Blended intrinsic maps
[Kim et al. 11]



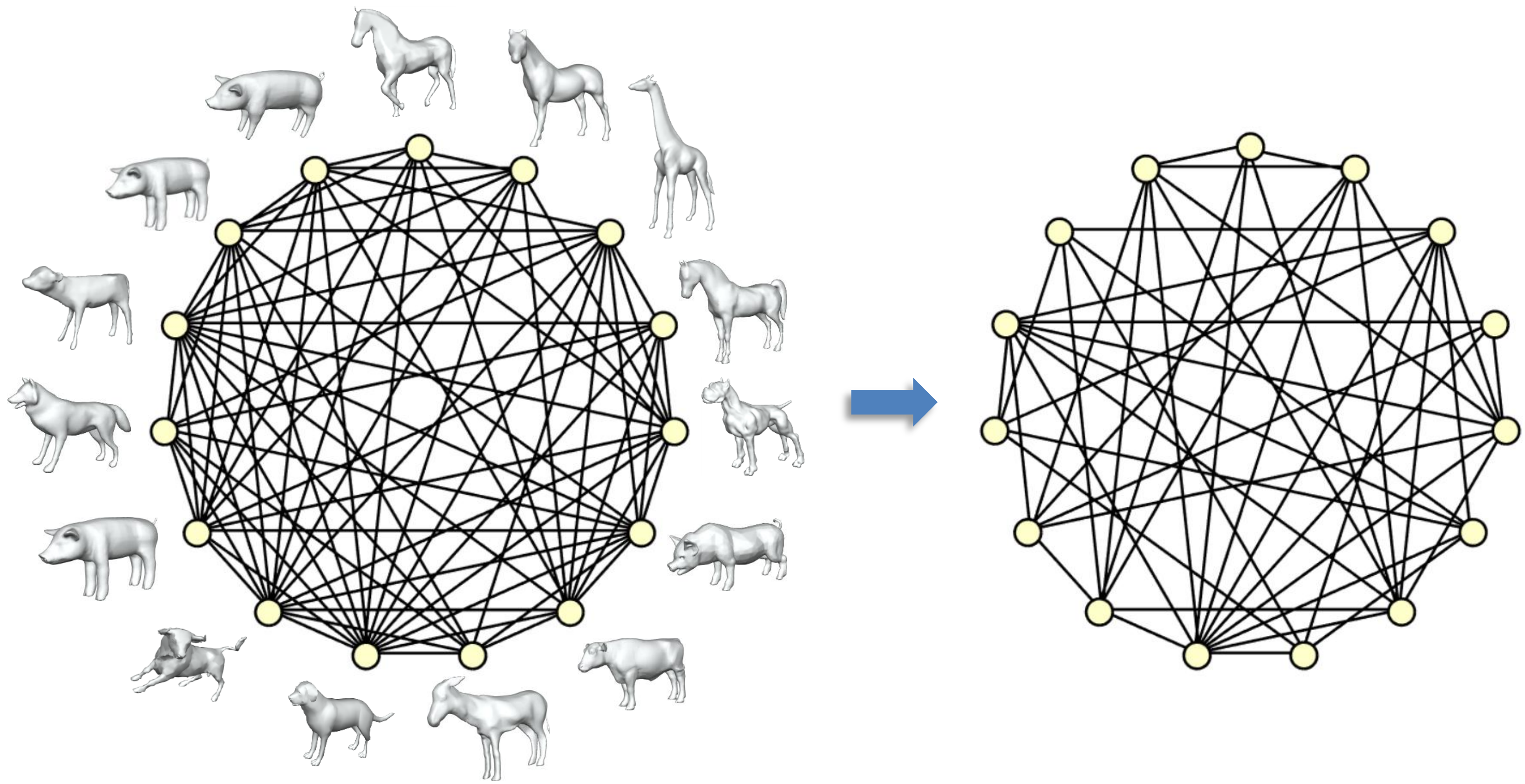
Composite

Pair-wise maps usually contain sufficient information



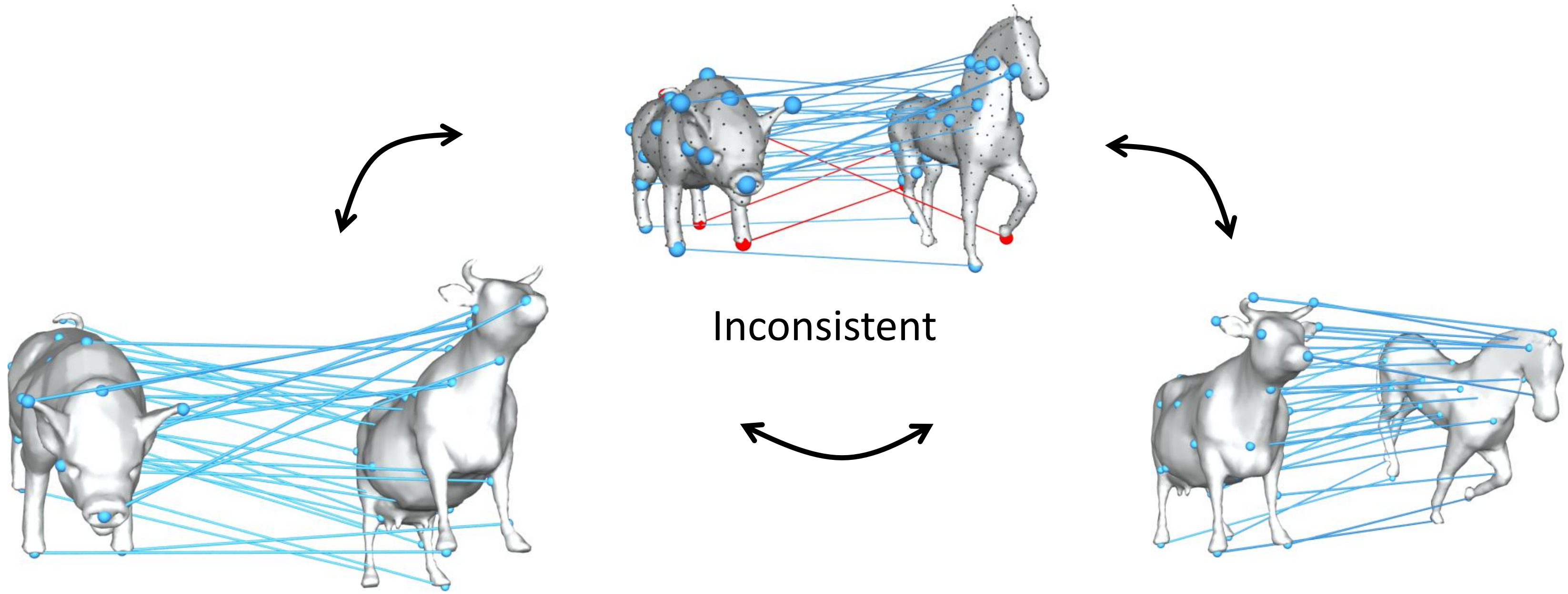
Network of approximately correct blended intrinsic maps

Map synchronization problem



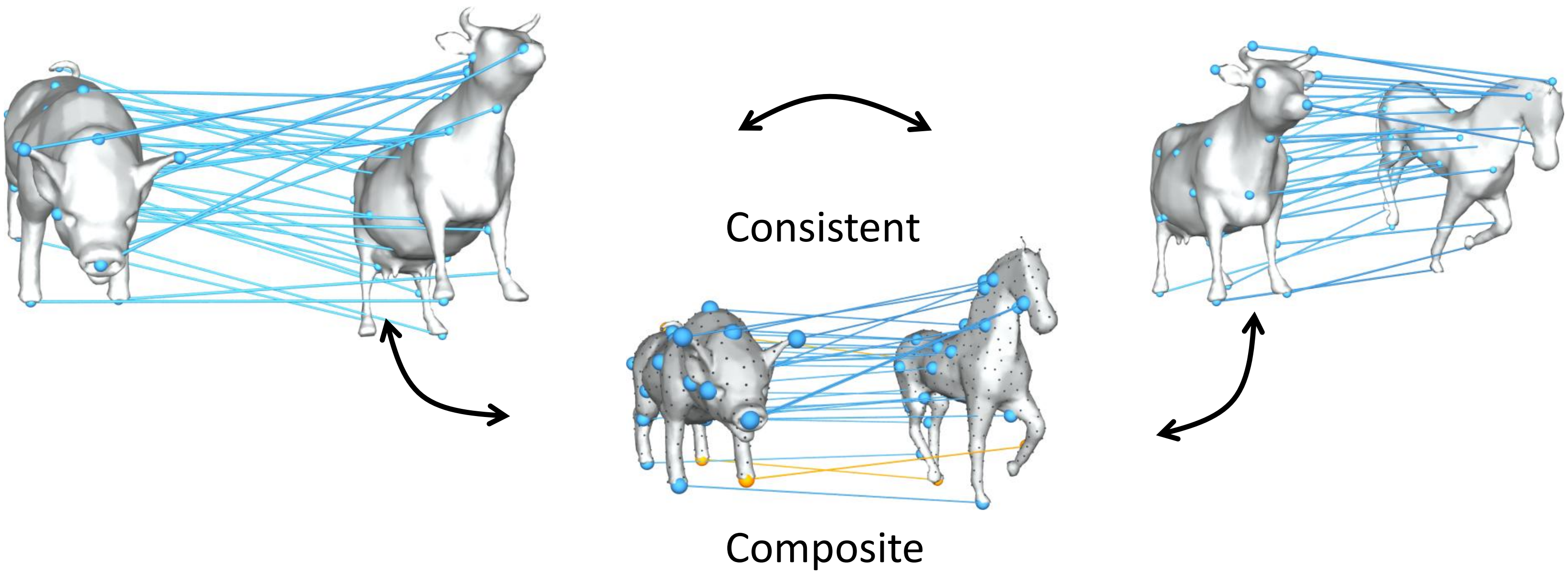
Identify correct maps among a (sparse) network of maps

A natural constraint on maps is that they should be consistent along cycles



Q. Huang, G. Zhang, L. Gao, S. Hu, A. Bustcher, and L. Guibas. *An Optimization Approach for Extracting and Encoding Consistent Maps in a Shape Collection*, SIGGRAPHAsia' 12

A natural constraint on maps is that they should be consistent along cycles



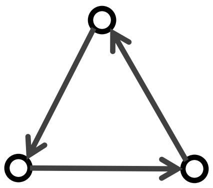
Q. Huang, G. Zhang, L. Gao, S. Hu, A. Bustcher, and L. Guibas. *An Optimization Approach for Extracting and Encoding Consistent Maps in a Shape Collection*, SIGGRAPHAsia' 12

Literature on utilizing the cycle-consistency constraint

- Spanning tree optimization [Huber et al. 01, Huang et al. 06, Cho et al. 08, Crandel et al. 11, Huang et al. 12]

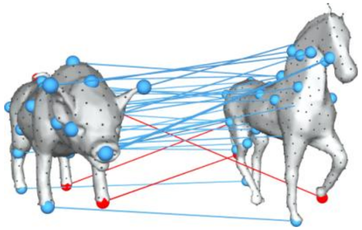
- Sampling inconsistent cycles [Zach et al. 10, Nyugen et al. 11, Zhou et al. 15]

Compressive sensing view of map synchronization



Cycle-consistency

Compressible

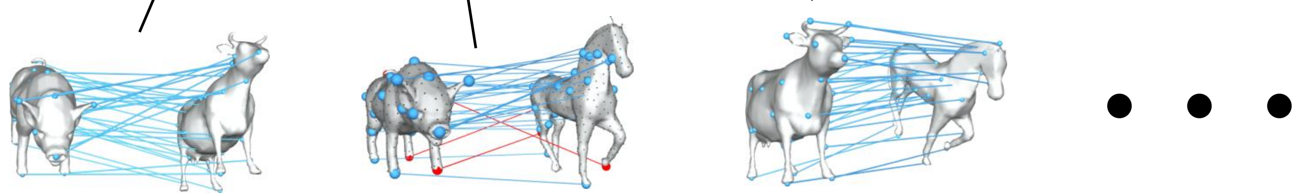


Input maps

Noisy observations

Map synchronization as constrained matrix optimization

$$X = \begin{bmatrix} I_m & X_{12} & \cdots & X_{1n} \\ X_{21} & I_m & \cdots & \vdots \\ \vdots & \vdots & \ddots & X_{n-1,n} \\ X_{n1} & \cdots & X_{n,n-1} & I_m \end{bmatrix} = \boxed{\begin{matrix} I_m \\ X_{21} \\ \vdots \\ X_{n1} \end{matrix}} \boxed{I_m \, X_{12} \cdots X_{1n}}$$



Noisy measurements of matrix blocks

Example: permutation synchronization

Objective function:

[Huang et al. 13]

$$\text{minimize} \sum_{(i,j)\in\mathcal{G}} \|X_{ij}^{\text{input}} - X_{ij}\|_1$$

Observation graph

Constraints:

$$X \succeq 0$$

cycle-consistency

$$\boxed{\begin{array}{l} X_{ii} = I_m, \quad 1 \leq i \leq n \\ X_{ij}\mathbf{1} = \mathbf{1}, X_{ij}^T\mathbf{1} = \mathbf{1}, \quad 1 \leq i < j \leq n \\ 0 \leq X \leq 1 \end{array}}$$

mapping constraint

Robust recovery for maps

$$X^{\text{observation}} = X^{\text{ground-truth}} + X^{\text{noise}}$$

Recovery if **In some reduced space**

$$\text{spectral-gap}(X^{\text{ground-truth}}) \geq c\|X^{\text{noise}}\|$$

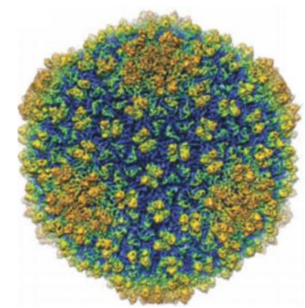
[Huang and Guibas 13, Wang and Singer 13, Bandeira et al. 14, Chen et al. 14, Zhou et al. 15, Chen and Candes 16, Shen et al. 16, Bajaj et al. 18,….]

Joint Map and Symmetry Synchronization

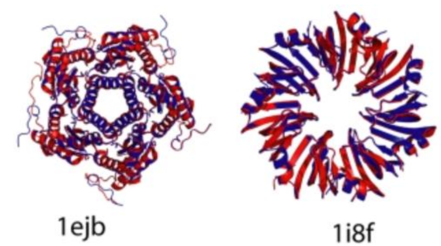
[Y. Sun, Z. Liang, X. Huang, Q. Huang. ECCV 2018]

Symmetric objects are ubiquitous

Daily objects



[Ranson and Stockley 10]



[André et al. 07]

Biological/chemical objects

Multiple plausible self-maps and pair-wise maps

No separation in the standard formulation

$$X^{\text{observation}} = X^{\text{ground-truth}} + X^{\text{noise}}$$

$$\begin{pmatrix} I_2 & I_2 & I_2 & \cdots & I_2 \\ I_2 & I_2 & I_2 & \cdots & \vdots \\ I_2 & I_2 & I_2 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_2 & \cdots & \cdots & \cdots & I_2 \end{pmatrix} \quad \text{v.s.} \quad \begin{pmatrix} I_2 & I_2 & I_2 & \cdots & -I_2 \\ I_2 & I_2 & -I_2 & \cdots & \vdots \\ I_2 & -I_2 & I_2 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -I_2 & \cdots & \cdots & \cdots & I_2 \end{pmatrix}$$

$$O(\sqrt{n})$$

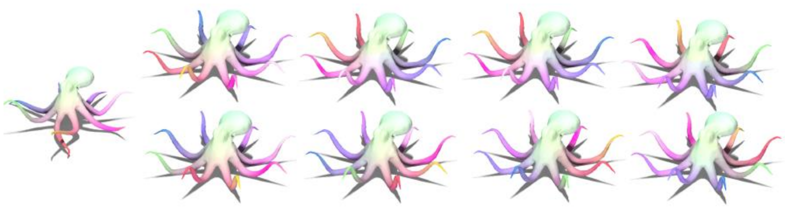
Symmetry detection first?

- Symmetry detection is difficult, particularly in the presence of partial observations

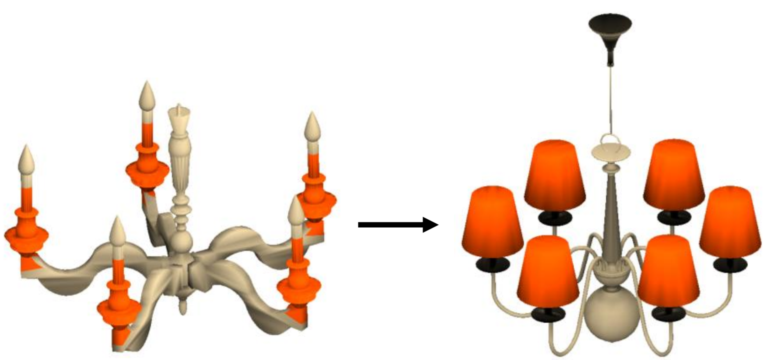
Dome of the Rock

Two correlated problems

Symmetry detection improves matching

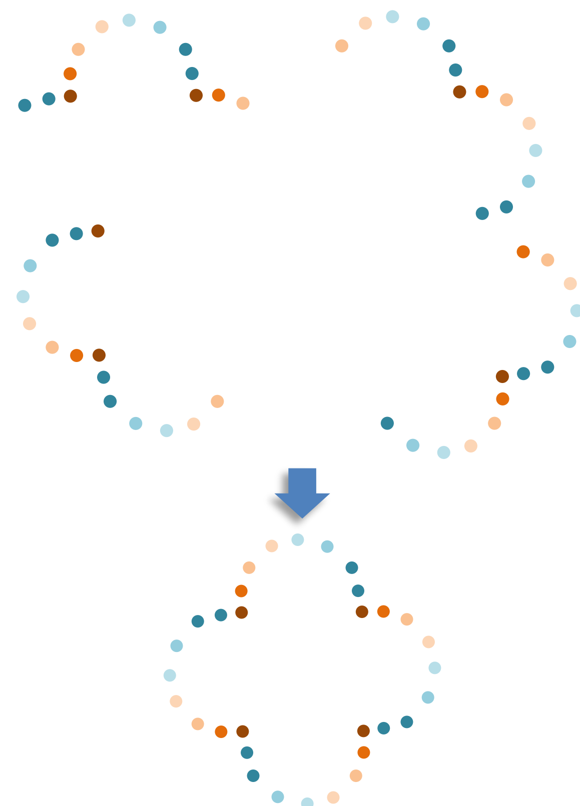


[Ovsjanikov et al. 13]



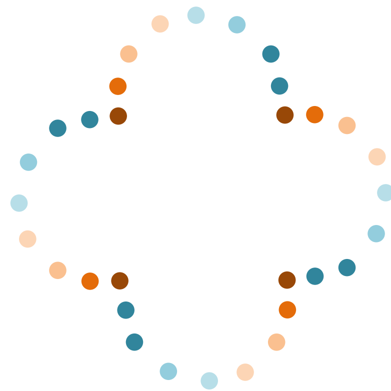
[Tevs and Huang et al. 14]

Better symmetry detection through information aggregation



Using the product operator - lifting

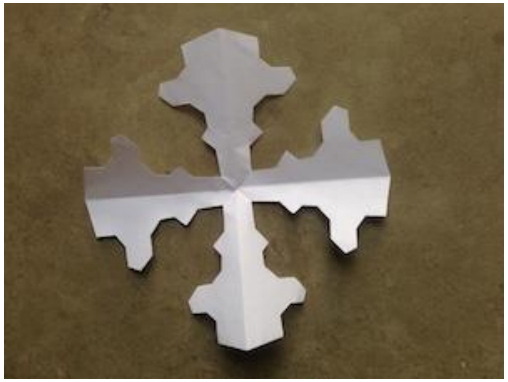
$$Q = \sum_{P \in \mathcal{G}} (P \otimes P)$$



Linear programming or semidefinite programming relaxations for MAP inference
[Wainwright and Jordan 08, Kumar et al. 09, Huang et el. 14,….]

Properties of lifting

- Proposition: *If the orbit size is equal to the group size, then we can recover G from Q*



Low-rank representation

- Define

$$\mathcal{F} : \mathbb{R}^{m_1^2 \times m_2^2} \to \mathbb{R}^{m_1 m_2 \times m_1 m_2}$$

$$\mathcal{F}(A)_{i_1 m_2 + i_2, j_1 m_2 + j_2} = A_{i_1 m_1 + j_1, i_2 m_2 + j_2}, \quad \begin{matrix} 0 \leq i_1, j_1 \leq m_1 - 1, \\ 0 \leq i_2, j_2 \leq m_2 - 1. \end{matrix}$$

- Then

$$\mathcal{F}(Q) = \sum_{P \in \mathcal{G}} \operatorname{vec}(P) \cdot \operatorname{vec}(P)^T$$

Low-rank

Observation induces a linear constraint

$$Q_{12} = \sum_{P_{12} \in \mathcal{M}_{12}} P_{12} \otimes P_{12}$$

$$\Downarrow$$

$$\mathcal{F}(Q_{12}) = \sum_{P_{12} \in \mathcal{M}_{12}} \mathrm{vec}(P_{12})\mathrm{vec}(P_{12})^T$$

+

$$\angle(\mathrm{vec}(P_{12}), \mathrm{vec}(P'_{12})) \approx \frac{\pi}{2}, \qquad \forall P_{12} \neq P'_{12} \in \mathcal{M}_{12}$$

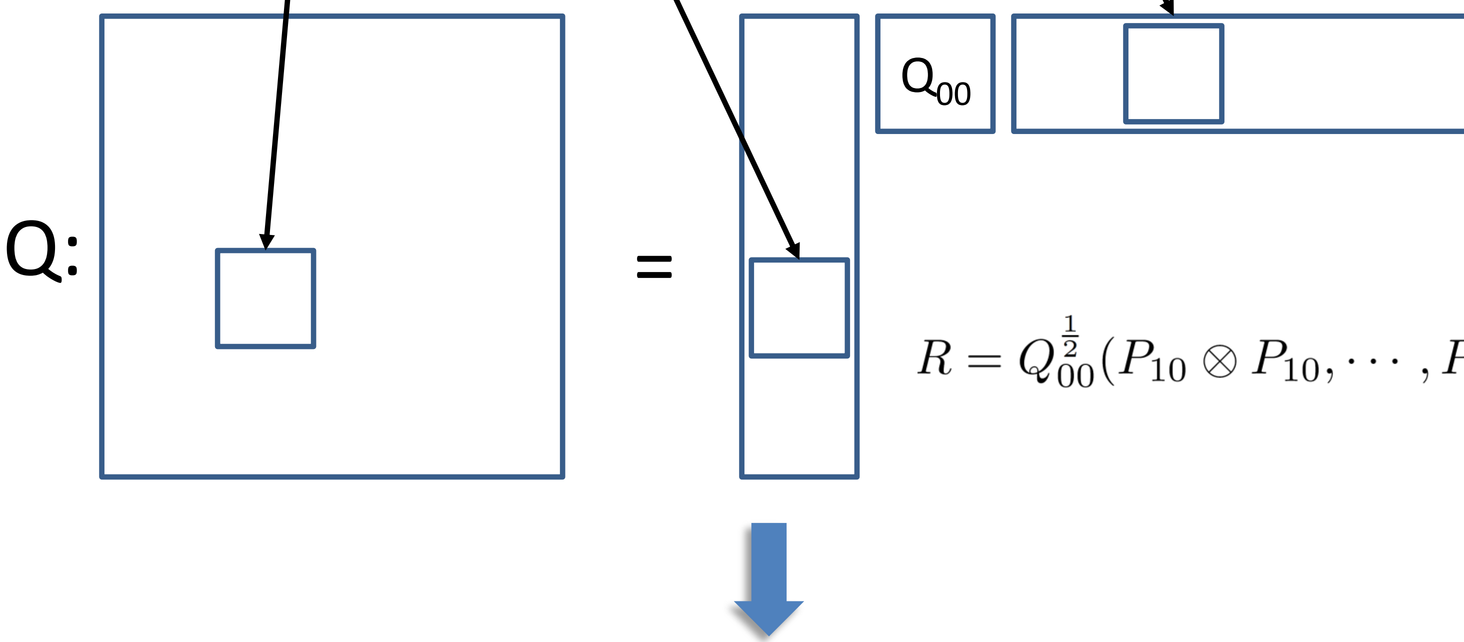
$$\Downarrow$$

$$\mathcal{F}(Q_{12})\mathrm{vec}(P_{12}) \approx \|\mathrm{vec}(P_{12})\|^2 \mathrm{vec}(P_{12})$$

Low-rank factorization

- Low-rank factorization

$$Q_{ij} = (P_{j0} \otimes P_{j0})^T Q_{00} (P_{i0} \otimes P_{i0})$$



$$R = Q_{00}^{\frac{1}{2}} (P_{10} \otimes P_{10}, \cdots, P_{n0} \otimes P_{n0})$$

$$Q = R^T R$$

Low-rank matrix recovery

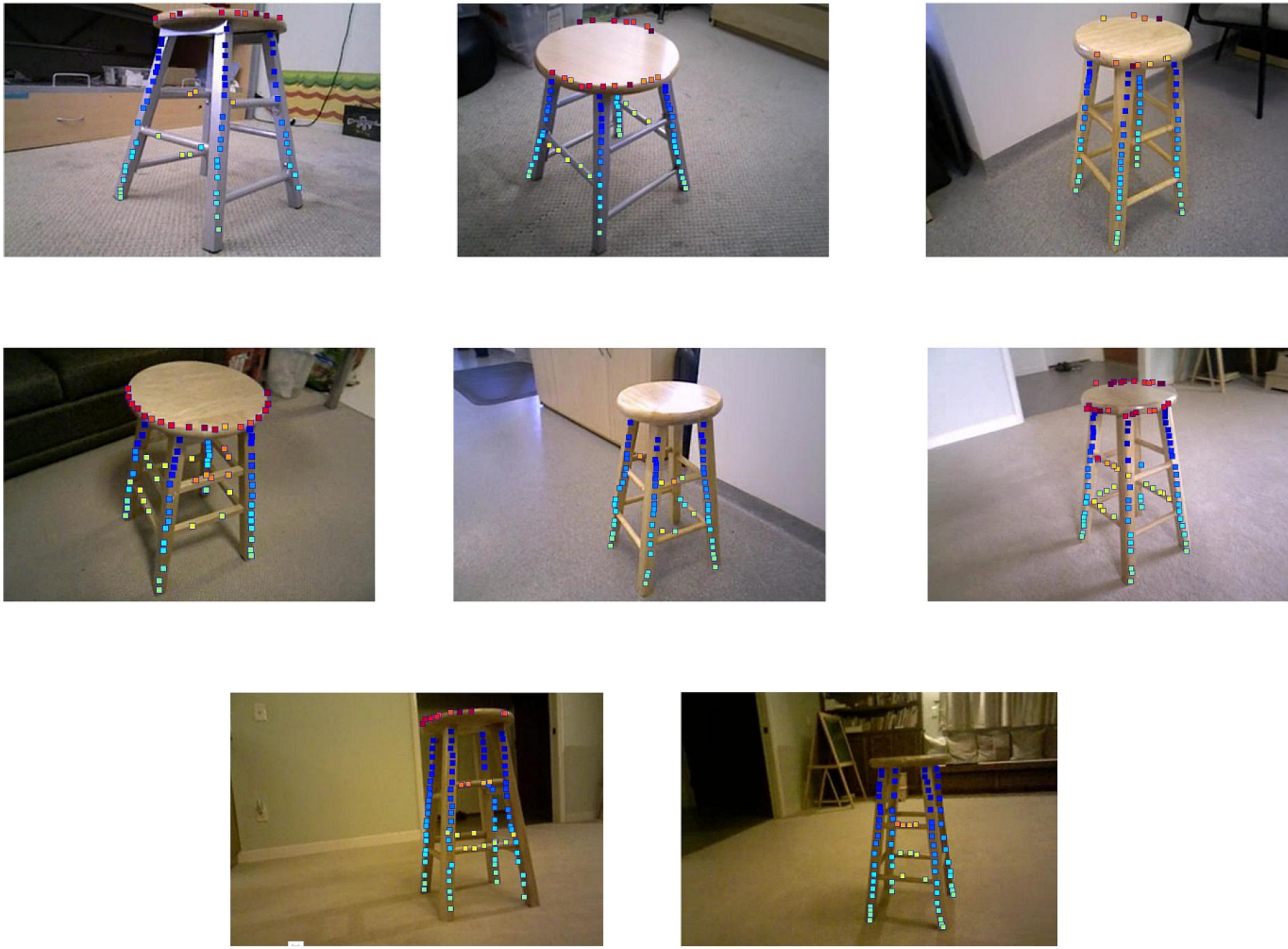
$$\min_{Q,R} \sum_{(i,j)\in\mathcal{E}} \|\mathcal{F}(Q_{ij})\mathrm{vec}(P_{ij}^{in}) - \|\mathrm{vec}(P_{ij}^{in})\|^2 \mathrm{vec}(P_{ij}^{in})\| + \lambda \|Q - R^T R\|_{\mathcal{F}}^2$$

Block-wise L1-norm for robust recovery

Low-rank constraint

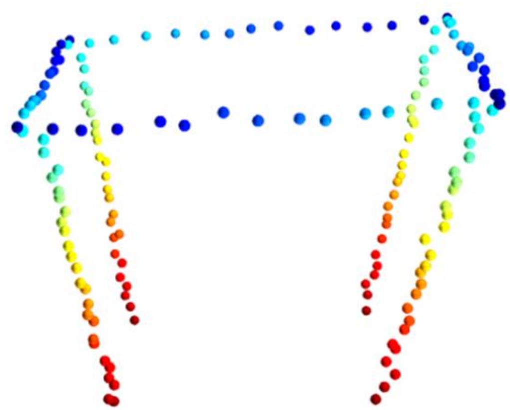
- Spectral initialization
- Alternating minimization
- Greedy rounding

Stool dataset



Trash Container Dataset

Similar but non-identical objects

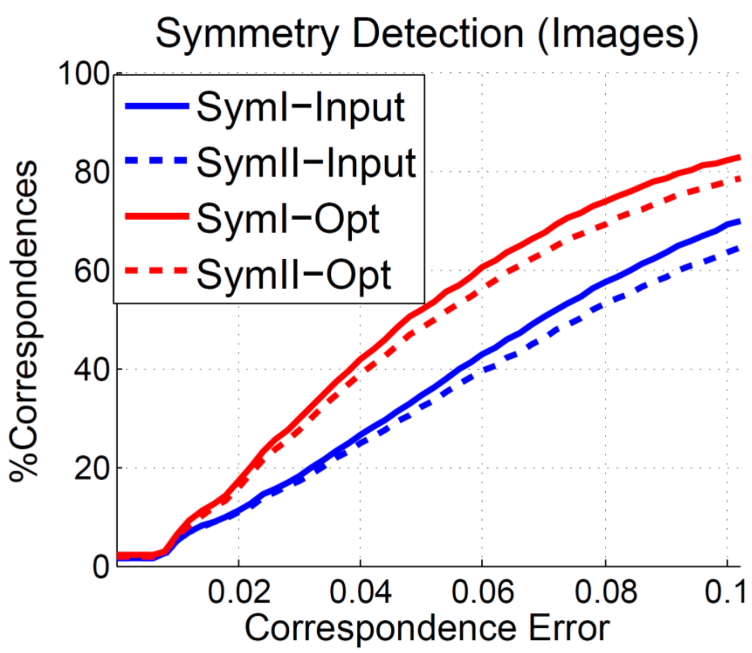


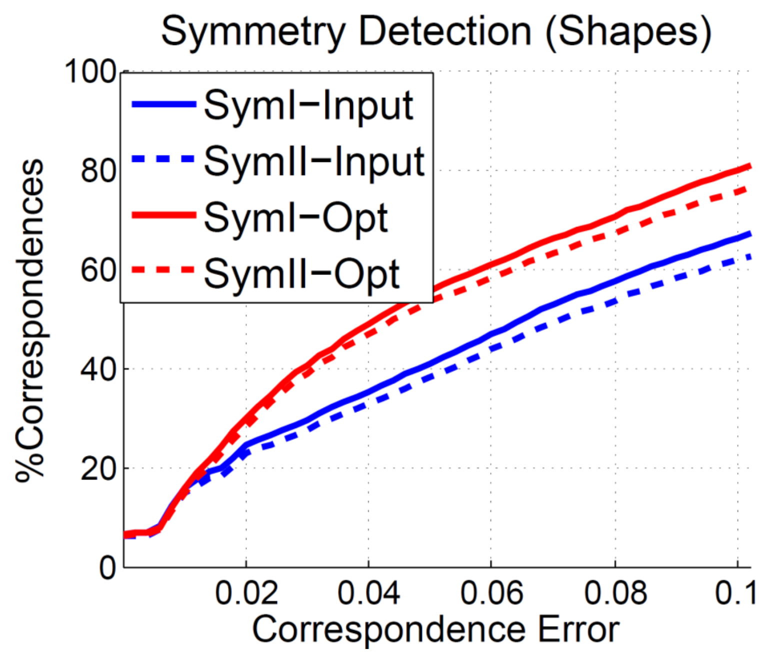

VK013290
WE352010


RECYCLE

Quantitative Evaluations

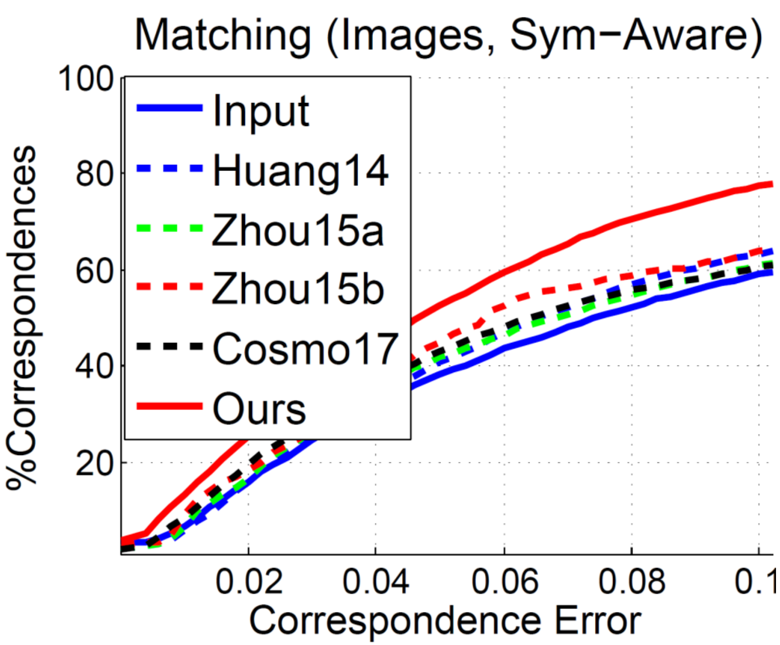
- Joint map and symmetry synchronization improves symmetry detection

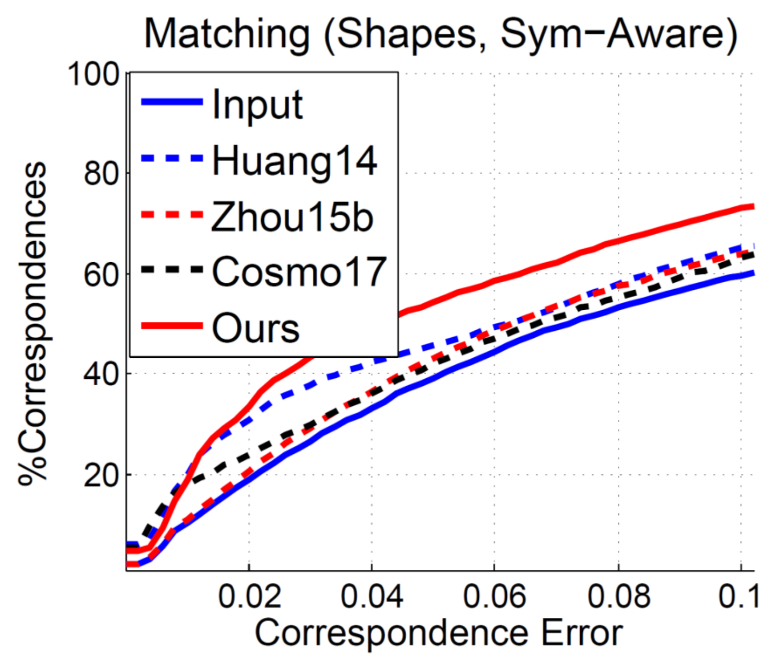




Quantitative Evaluations

- Joint map and symmetry synchronization improves mapping
  - With respect to the closest map (not correspondence)



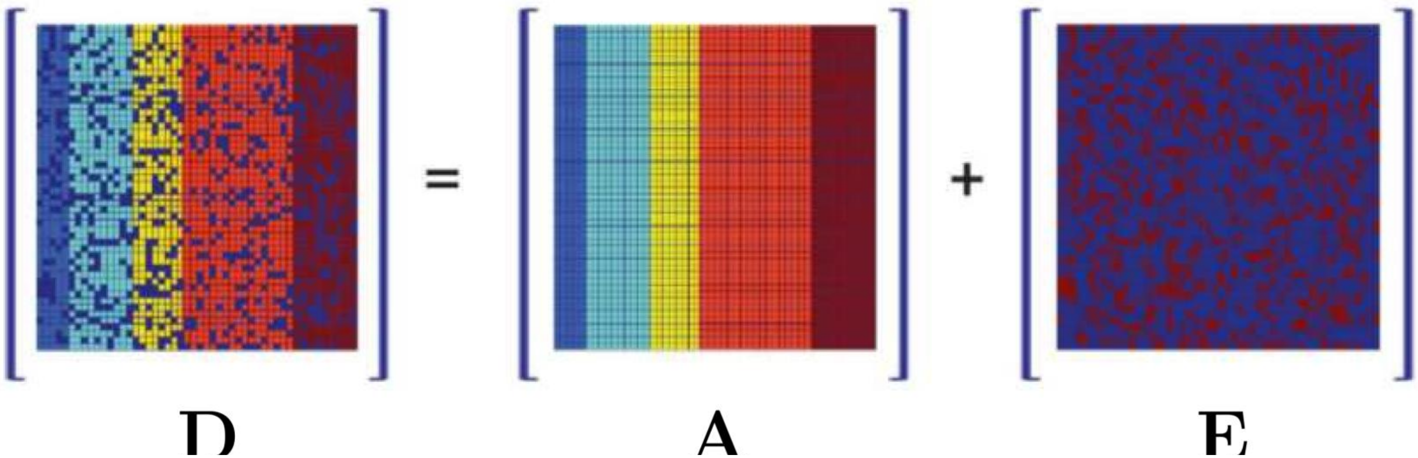


Learning Transformation Synchronization

[X. Huang, Z. Liang, X. Zhou, X. Yao, L. Guibas, Q.H., CVPR' 2019]

Hand-crafted objective function

[Candes et al. 11]



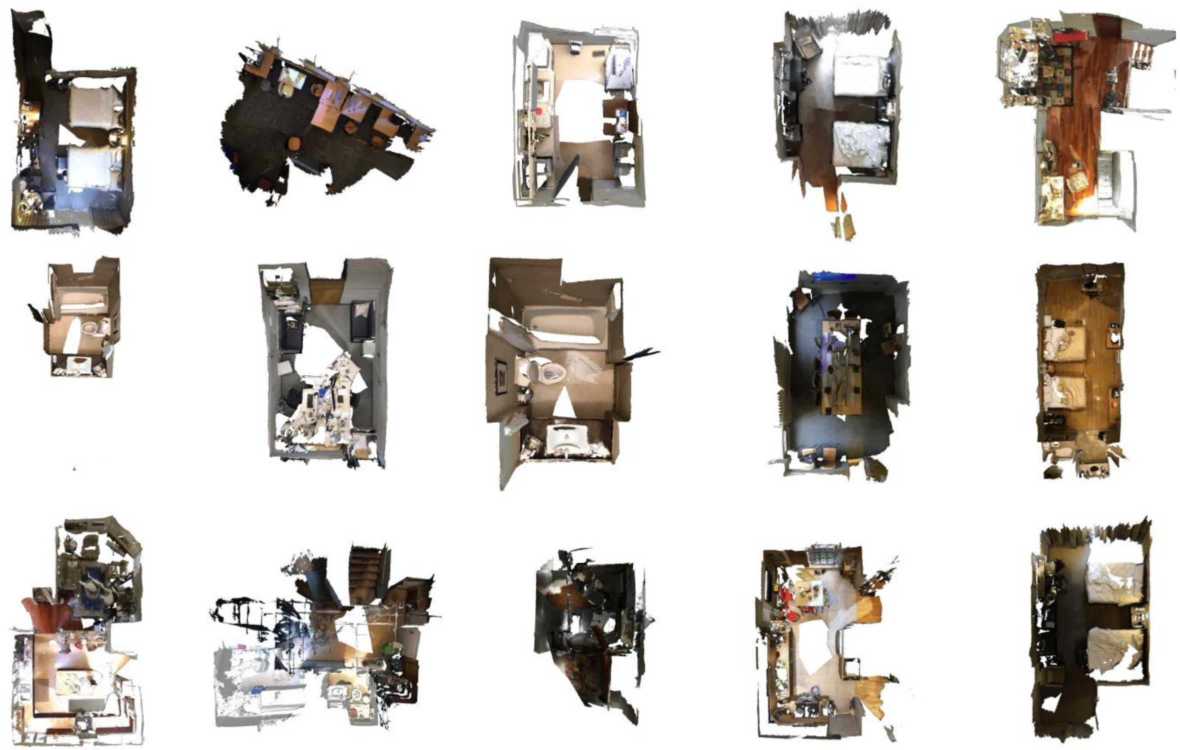
$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1, \text{ subject to } \mathbf{A} + \mathbf{E} = \mathbf{D}.$$

$\|\mathbf{A}\|_*$: nuclear norm, sum of singular values of $\mathbf{A}$; surrogate for rank($\mathbf{A}$).

$\|\mathbf{E}\|_1$: l_1-norm, sum of absolute values of elements of $\mathbf{E}$; surrogate for $\|\mathbf{E}\|_0$.

3D scene reconstruction from depth scans

[Dai et al. 17]



- Similar noise sources
  - Scanning noise, frame rate, and symmetry structures

Reweighted least square synchronization

Rotation:

$$\underset{R_i \in SO(3), 1 \le i \le n}{\text{minimize}} \sum_{(i,j) \in \mathcal{E}} w_{ij} \|R_{ij} R_i - R_j\|_{\mathcal{F}}^2$$

Solved by the first 3 eigenvectors of a Connection Laplacian

$$L_{ij} := \begin{cases} \sum_{j \in \mathcal{N}(i)} w_{ij} I_3 & i = j \\ -w_{ij} R_{ij}^T & (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

Translation:

$$\underset{\boldsymbol{t}_i, 1 \le i \le n}{\text{minimize}} \sum_{(i,j) \in \mathcal{E}} w_{ij} \|R_{ij} \boldsymbol{t}_i + \boldsymbol{t}_{ij} - \boldsymbol{t}_j\|^2$$

Linear system:

$$\boldsymbol{t}^\star = L^+ \boldsymbol{b}$$

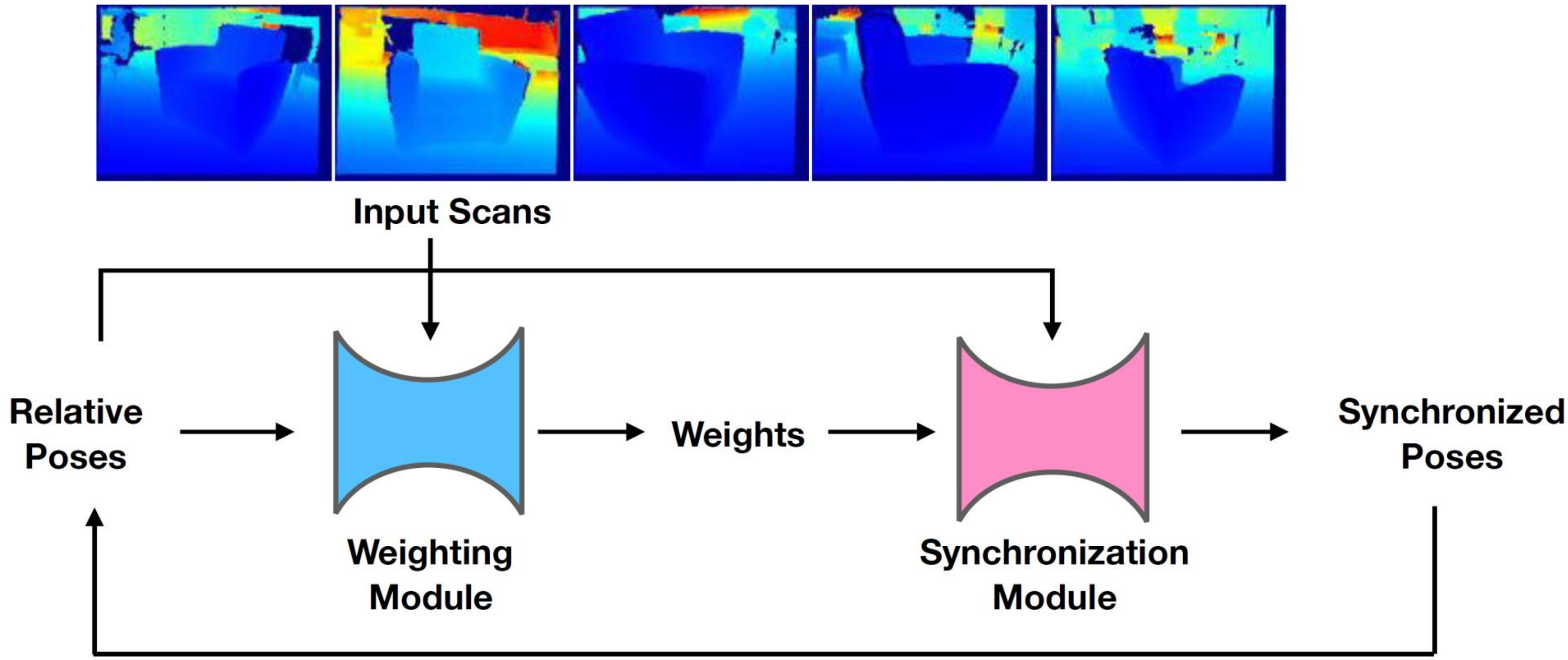
Where

$$\boldsymbol{b}_i := -\sum_{j \in \mathcal{N}(i)} w_{ij} R_{ij}^T \boldsymbol{t}_{ij}$$

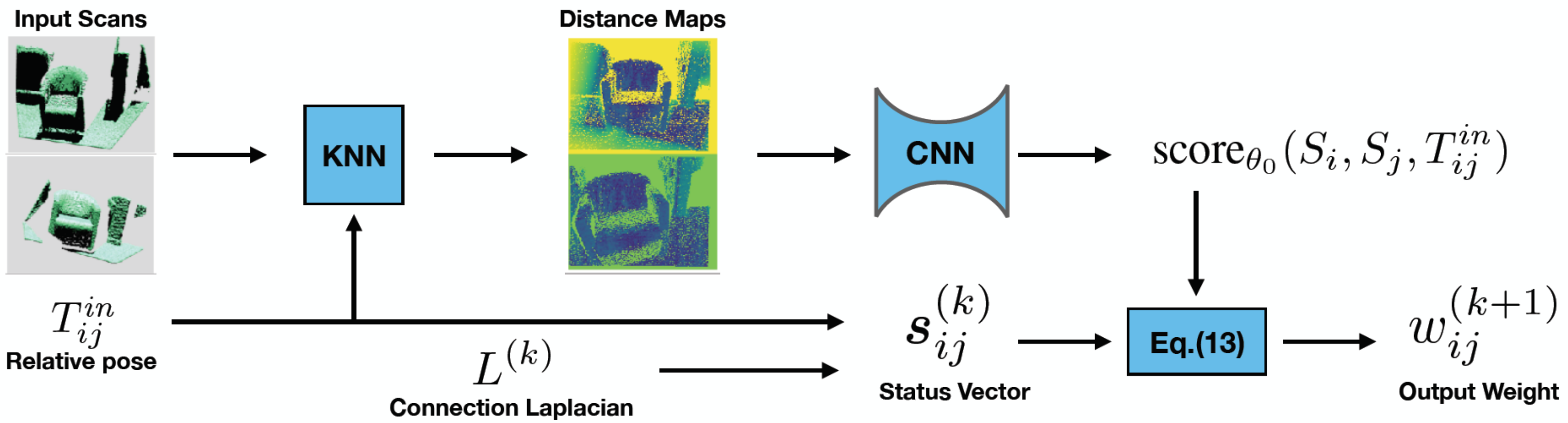
Robust recovery under a constant fraction of adversarial noise if

$$w_{ij} = \rho(\|R_{ij} R_i^{(k)} - R_j^{(k)}\|) \quad \text{where} \quad \rho(x) = \frac{\epsilon^2}{\epsilon^2 + x^2}$$

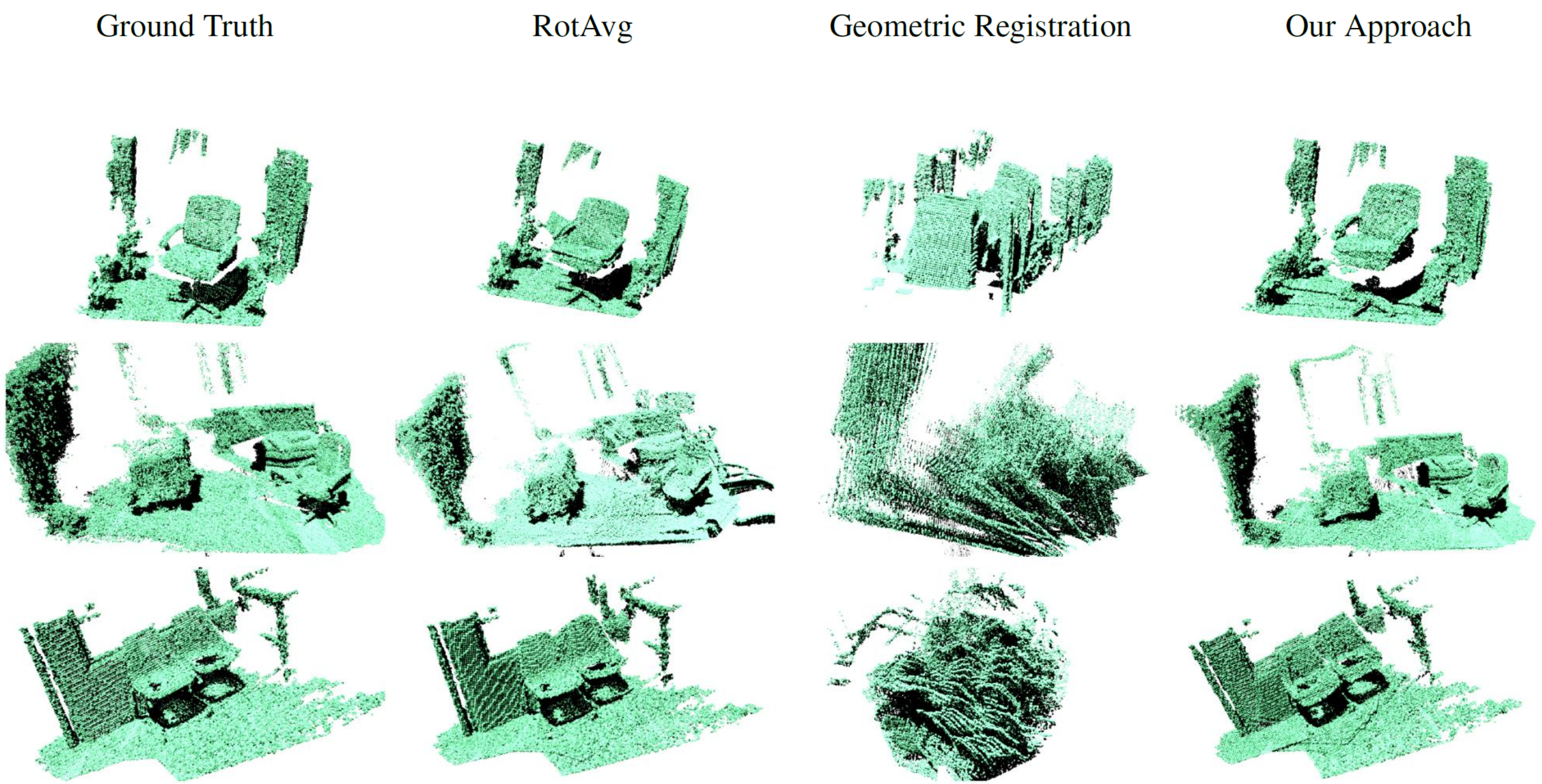
Network design



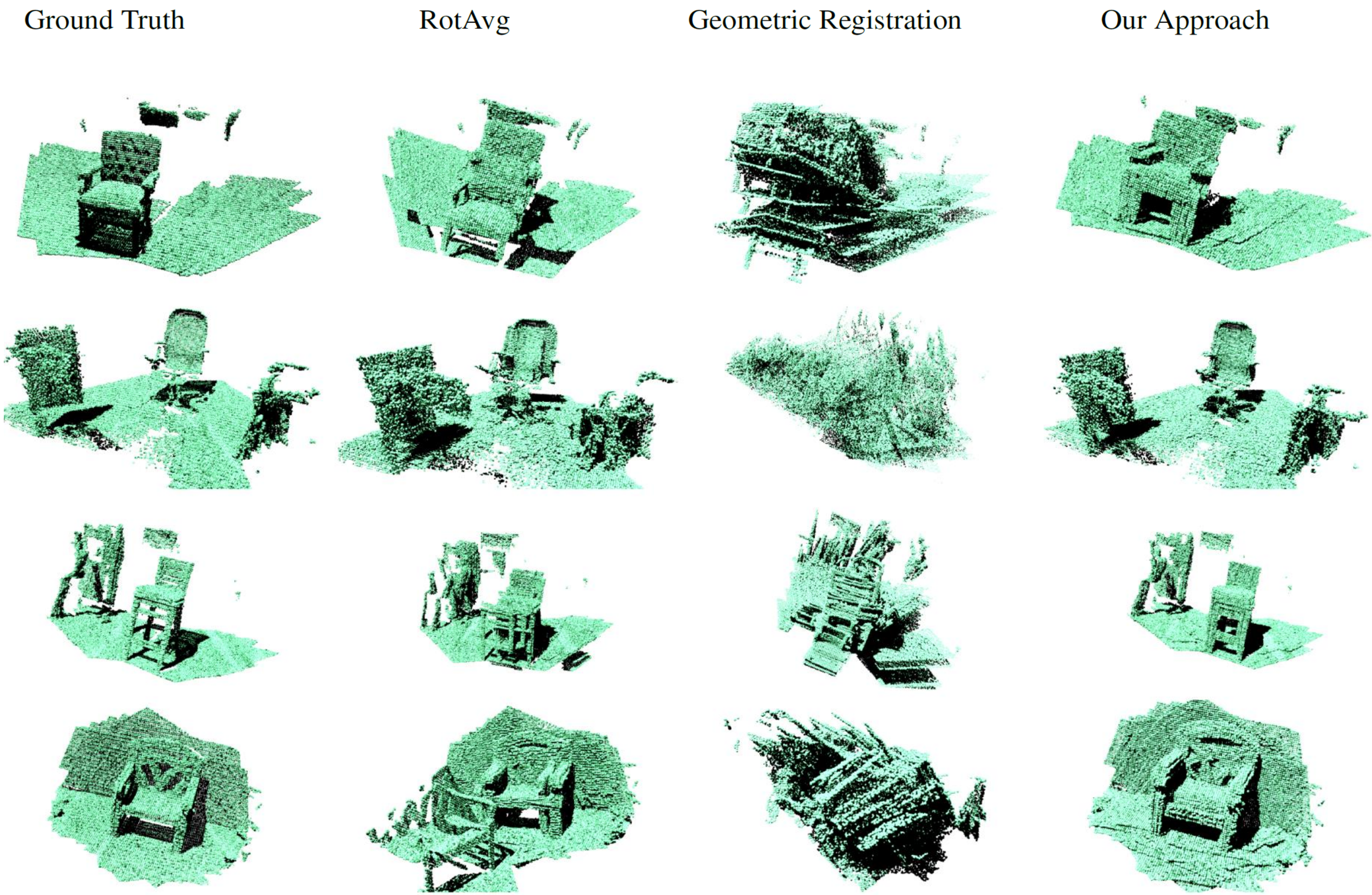
Weighting module



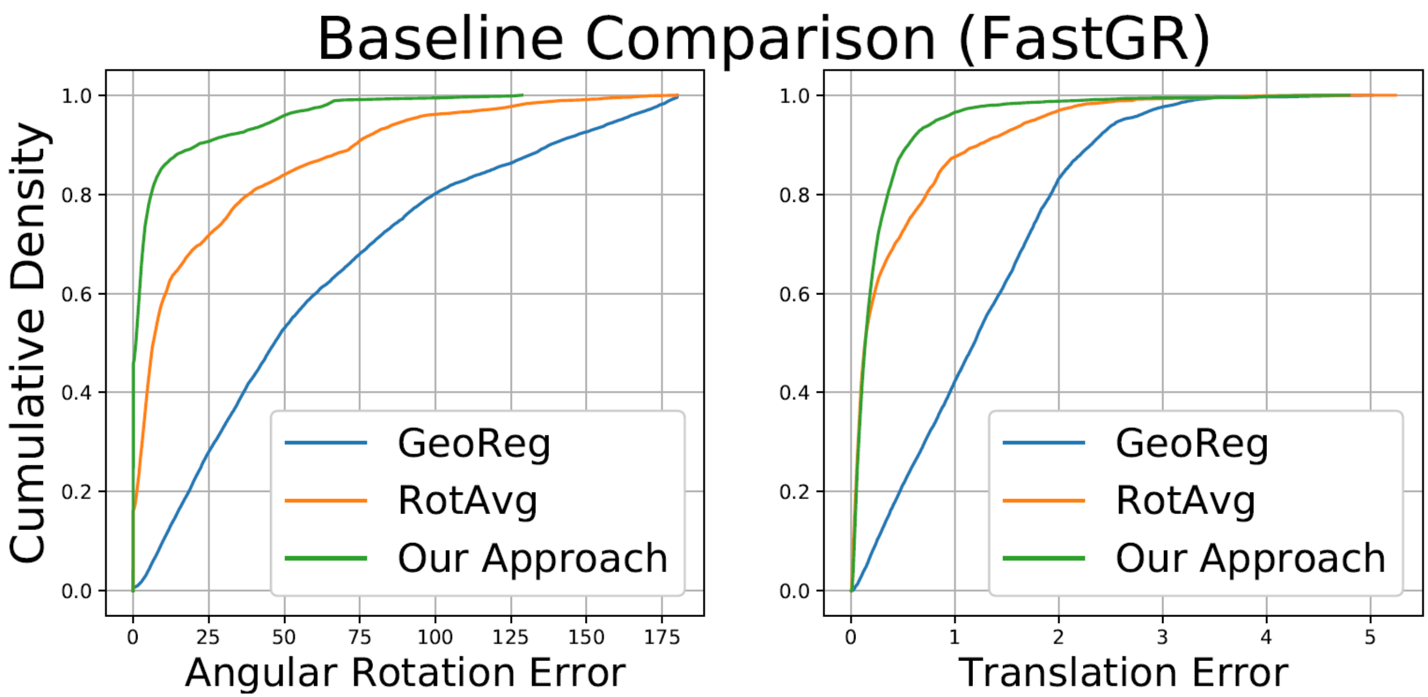
Qualitative results



Qualitative results

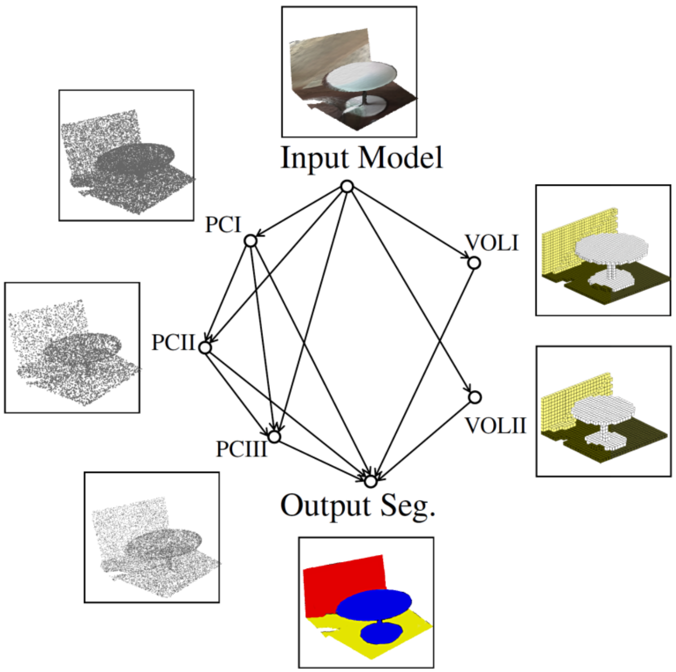


Quantitative results



Redwood dataset

Limitations of low-rank approaches



Neural networks

Directed maps

$$X = \begin{bmatrix} I_m & X_{12} & \cdots & X_{1n} \\ X_{21} & I_m & \cdots & \vdots \\ \vdots & \vdots & \ddots & X_{n-1,n} \\ X_{n1} & \cdots & X_{n,n-1} & I_m \end{bmatrix}$$

Matrix representations

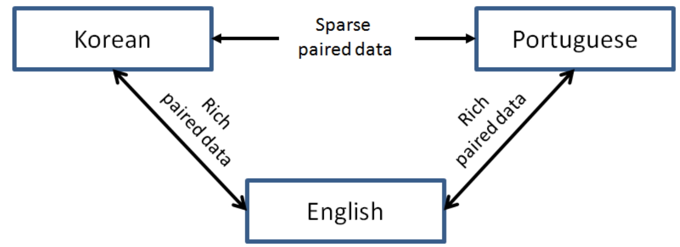
Undirected maps

Path-Invariant Map Networks

[Z. Zhang, Z. Liang, L. Wu, X. Zhou, Q. H, CVPR 2019' Oral]

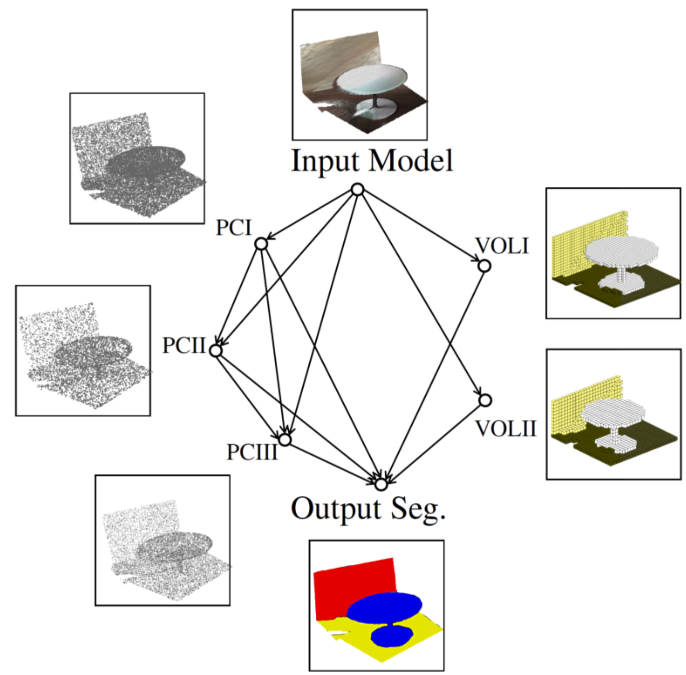
Benefits of Joint of Learning of Neural Networks

- Leverage additional training data



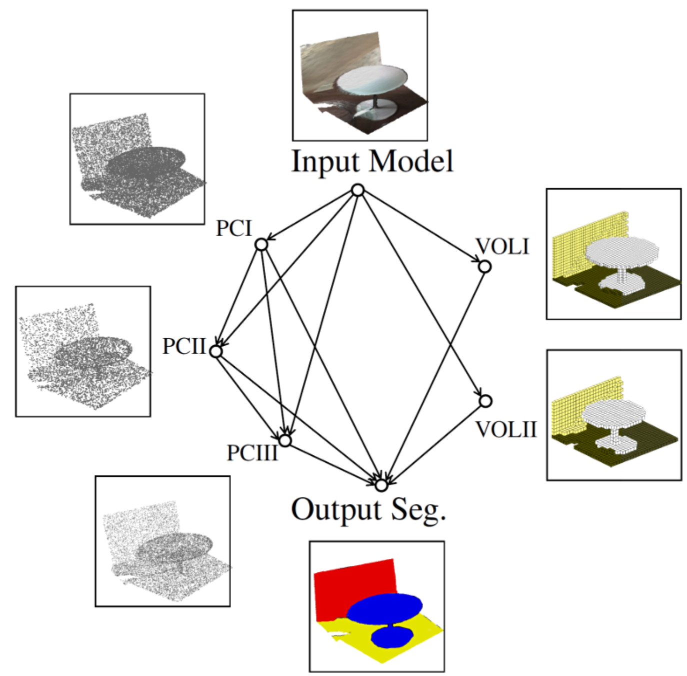
- Fuse patterns learned under individual representations

- Utilize unlabeled data

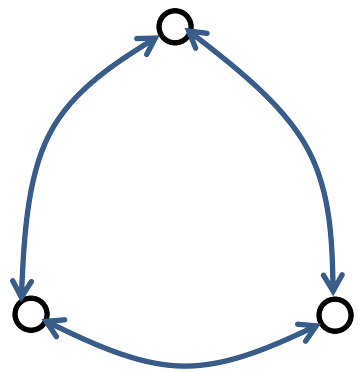


Challenges

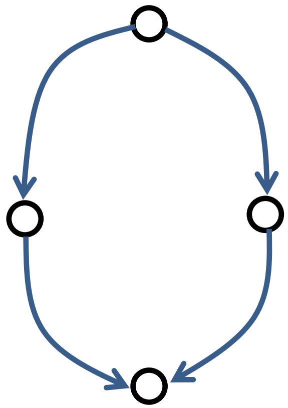
- Cannot represent neural networks as matrices
- Need to regularization constraint for directed maps
- Need a concise representation



From cycle-consistency to path-invariance



Cycle-consistency



Path-invariance

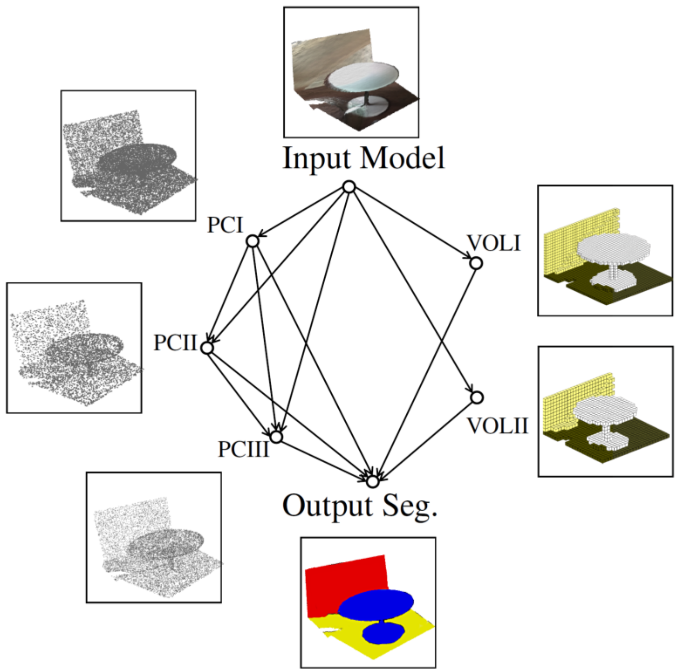
Definition *Let $\mathcal{G}_{\text{path}}(u,v)$ collect all paths in $\mathcal{G}$ that connect u to v. We define the set of all possible path pairs of $\mathcal{G}$ as*

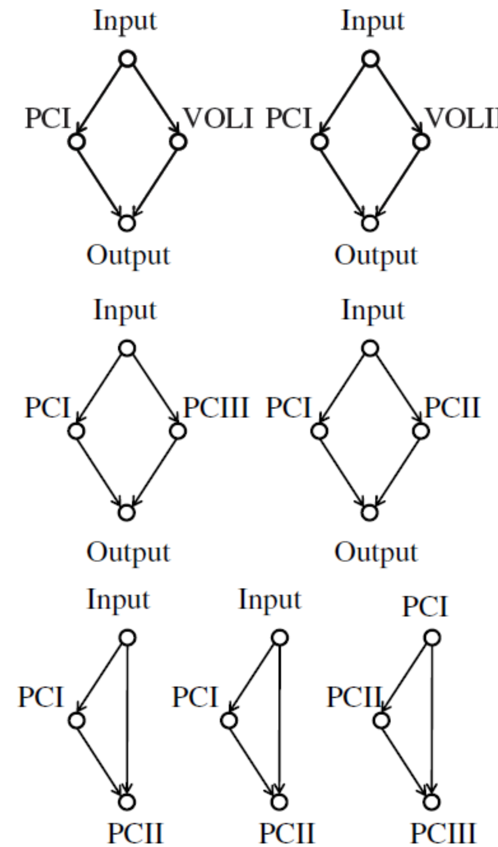
$$\mathcal{G}_{\text{pair}} = \bigcup_{u,v\in\mathcal{V}} \{(p,q)|p,q \in \mathcal{G}_{\text{path}}(u,v)\}.$$

We say $\mathcal{F}$ is path-invariant if

$$f_p = f_q, \qquad \forall (p,q) \in \mathcal{G}_{\text{pair}}.$$

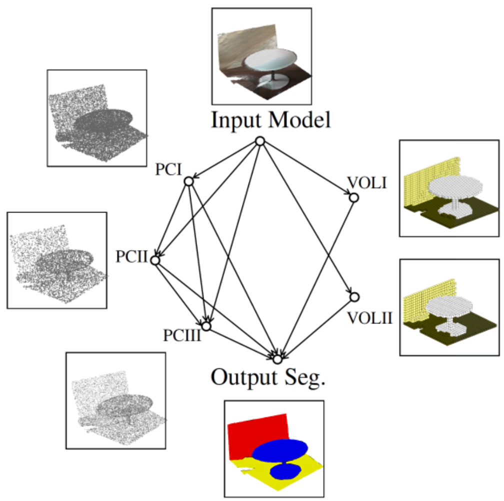
Path-invariance basis

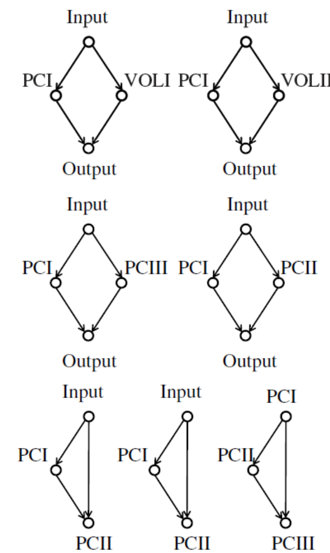




Can induce the path-invariance property of the entire graph

Path-invariance provides a regularization for training neural networks





$$\min_{\Theta} \sum_{(i,j)\in\overline{\mathcal{E}}} l_{ij}(f^{\Theta}_{\widehat{v_i v_j}}) + \lambda \sum_{(p,q)\in\mathcal{B}} \underset{v\sim P_{p_t}}{E} d_{\mathcal{D}_{p_t}}(f_p^{\Theta}(v), f_q^{\Theta}(v))$$

Supervised loss

Unsupervised loss

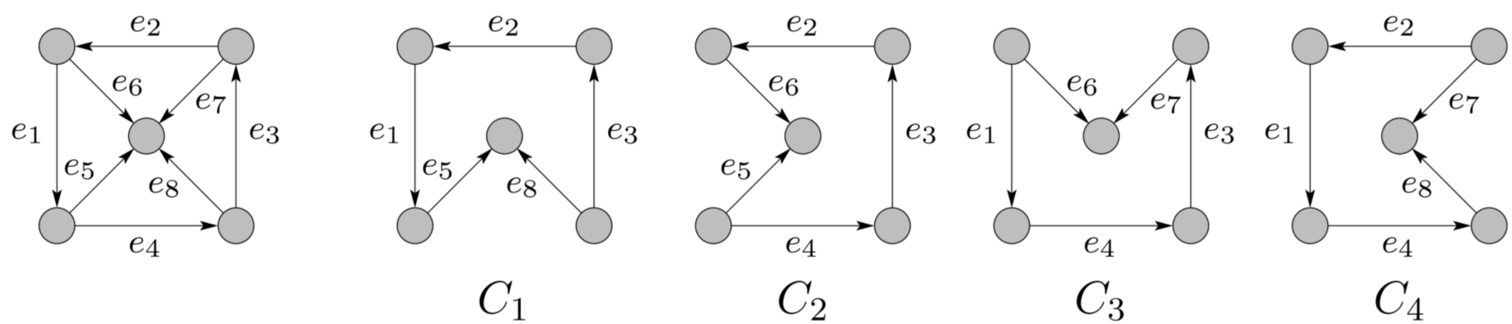
Main result

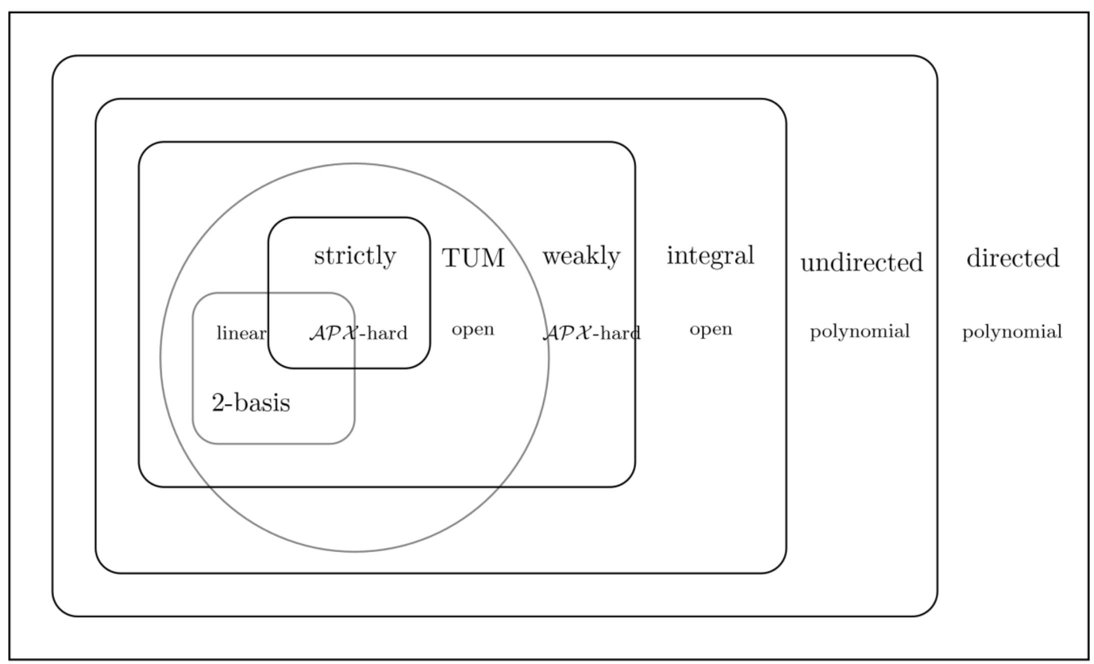
- Theorem: *Given a directed graph with n vertices and m edges, there exists a path-invariance basis with size at most O(nm)*

- Main idea for the proof
  - A directed graph is a directed acyclic graph (DAG) of strongly connected components
  - Use a vertex order to construct a path-invariance basis for DAG

Cycle-basis

[Kavitha et al. 09]

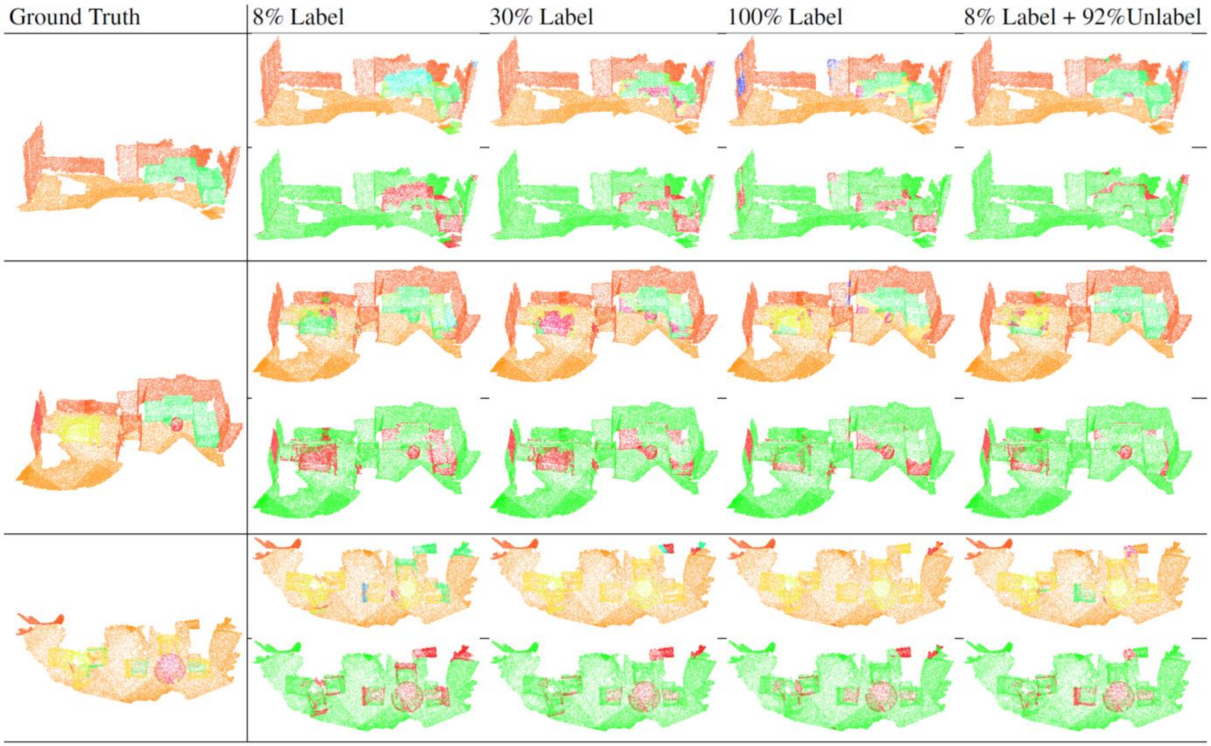




Cycle basis/Cycle-consistency basis/Path-invariance basis

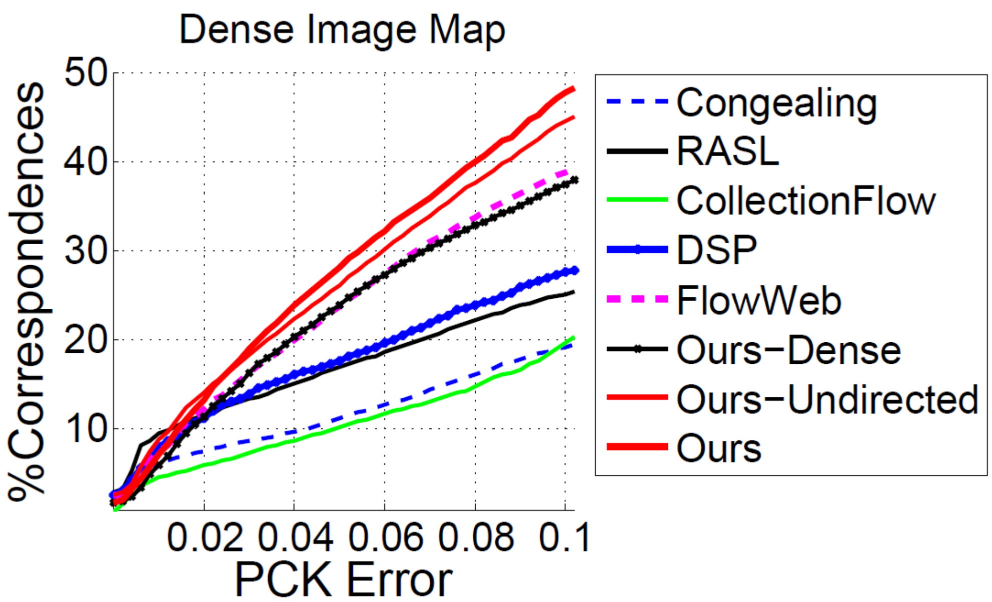
- Undirected cycle basis generalizes to cycle-consistency basis
  - Size: #edges - #vertices + #components
- It is an open question whether other bases generalize
- The minimum size of a path-invariance basis is an open problem

Semantic segmentation on ScanNet

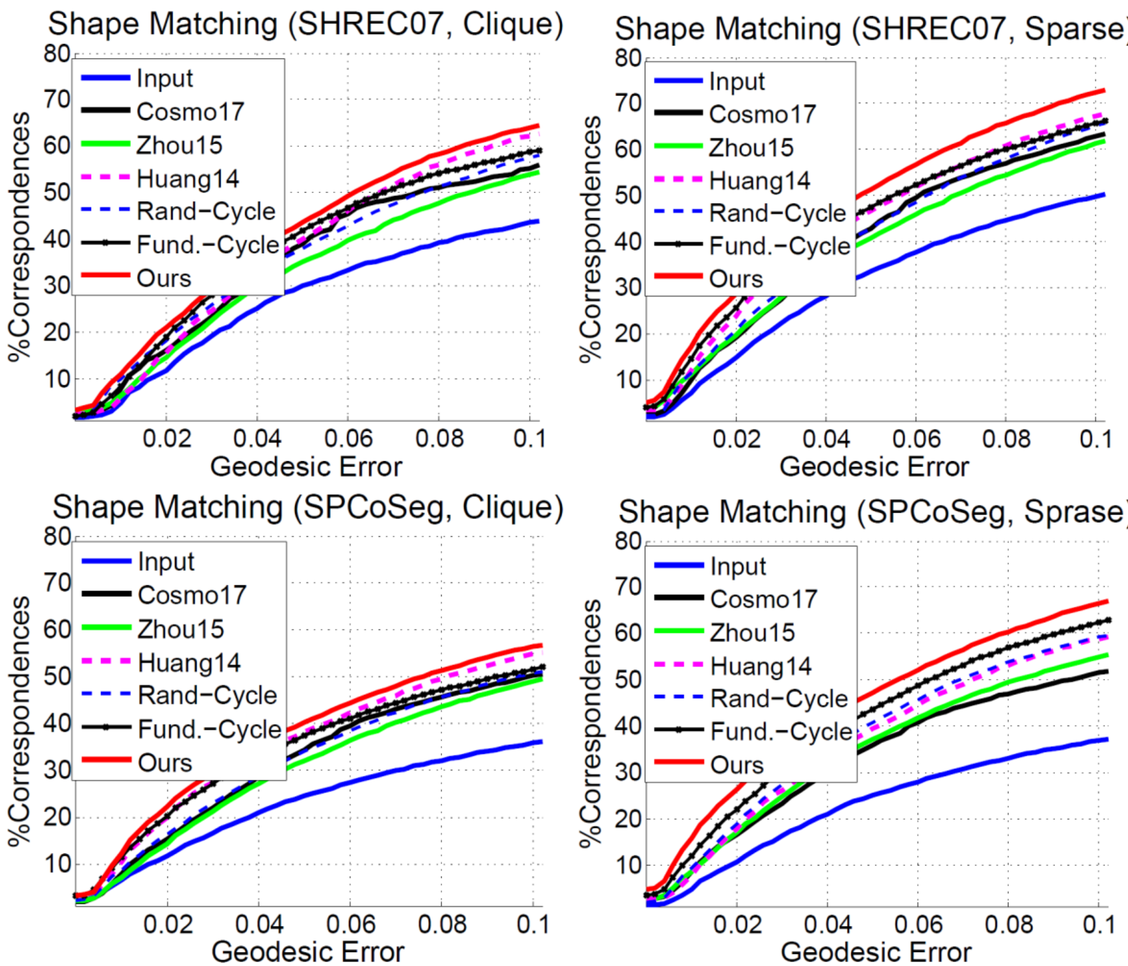


	PCI	PCII	PCIII	VOLI	VOLII
100% Label (Isolated)	84.2	83.3	83.4	81.9	81.5
8% Label (Isolated)	79.2	78.3	78.4	78.7	77.4
8% Label + 92%Unlabel (Joint)	81.7	81.7	81.4	81.1	78.7
30% Label (Isolated)	80.8	81.9	81.2	80.3	79.5

Application in learning image flows



Comparisons on computing object correspondences



Concluding remarks

- Map synchronization is a powerful tool for computing high-quality maps across a data collection
- The interplay between cycle-consistency/path-invariance and low-rank representations is important
- Many open questions:
  - Path-invariance basis
  - Theoretical guarantees of optimizing path-invariant map networks
  - Uncertainty quantification

Acknowledgments

Questions?