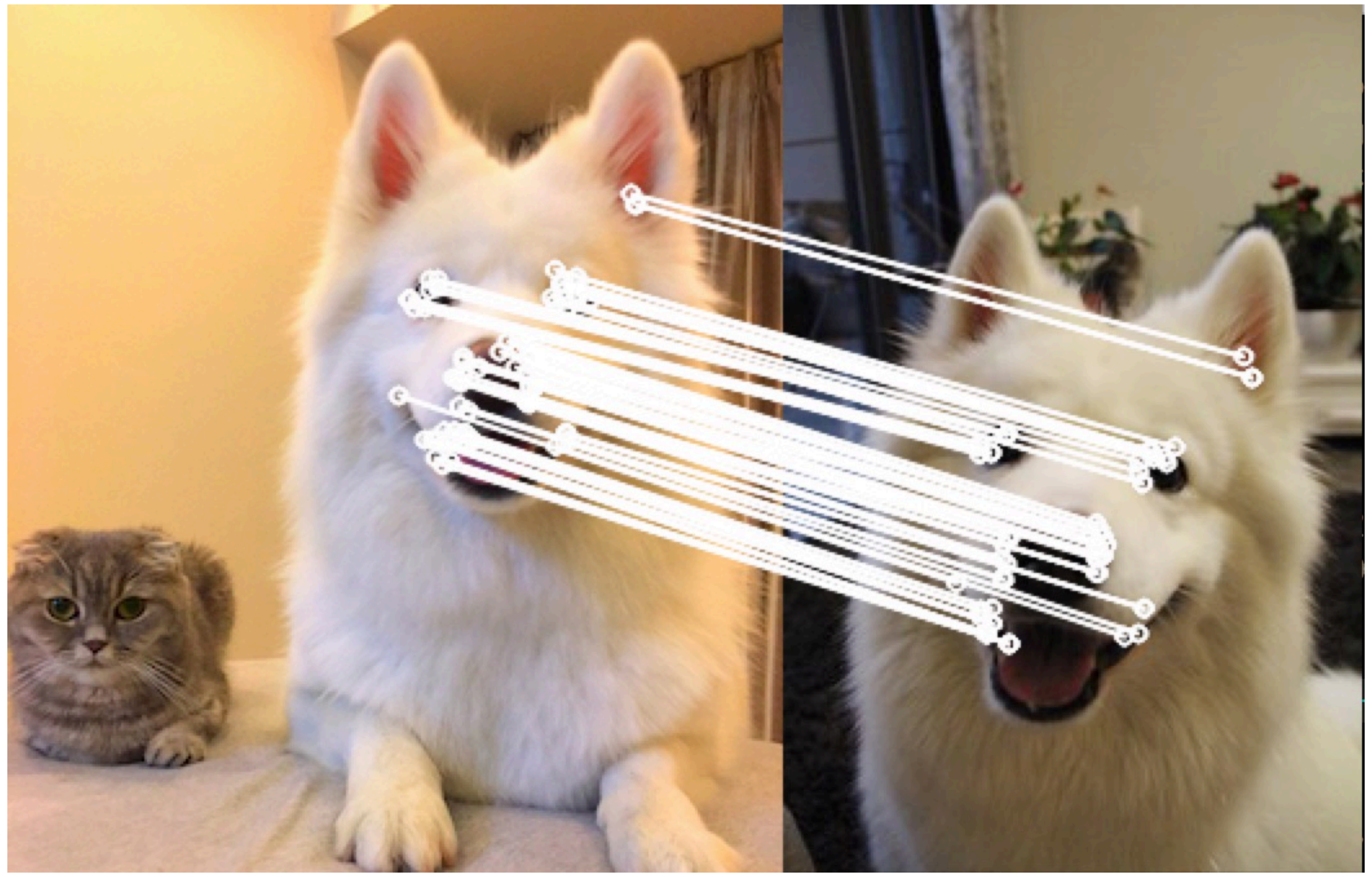
CVPR LONG BEACH CALIFORNIA June 16-20, 2019

浙江大学 Zhejiang University

Map Synchronization for Poses and Correspondences

Xiaowei Zhou
Zhejiang University

Geometric transformation between images



Local correspondences

Global relative pose

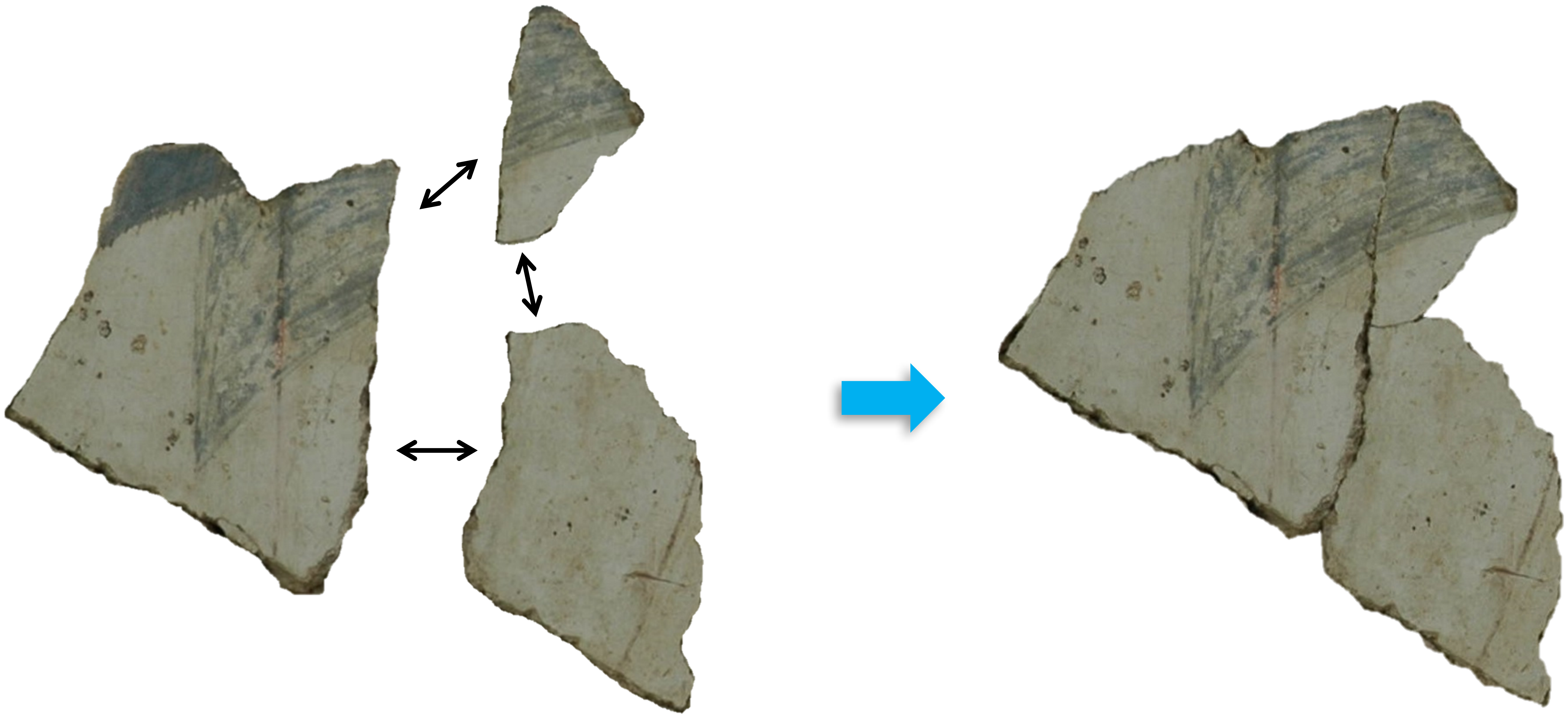
Why joint analysis?

Horse Mule Donkey

Why joint analysis?

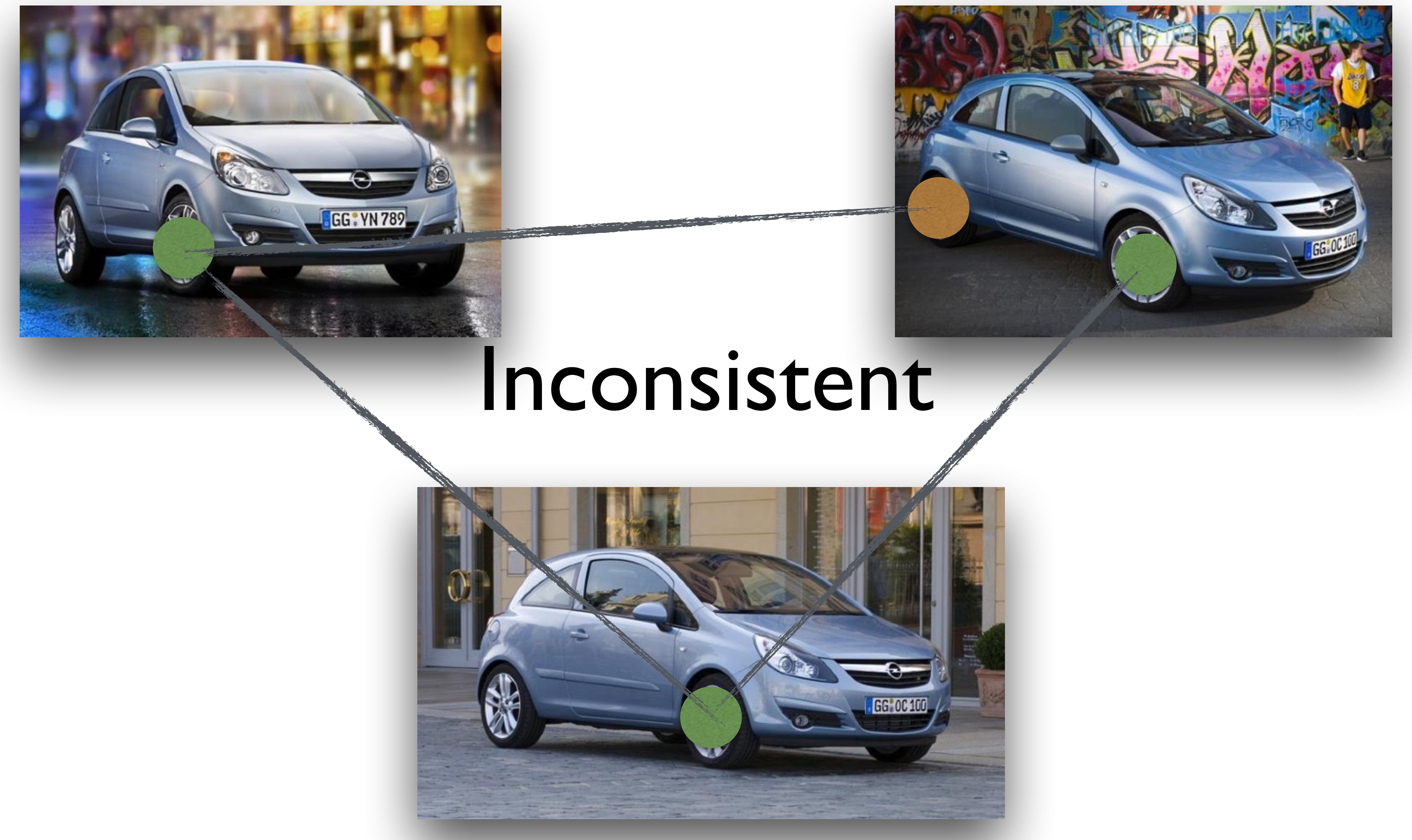
Ambiguities exist when matching two pieces

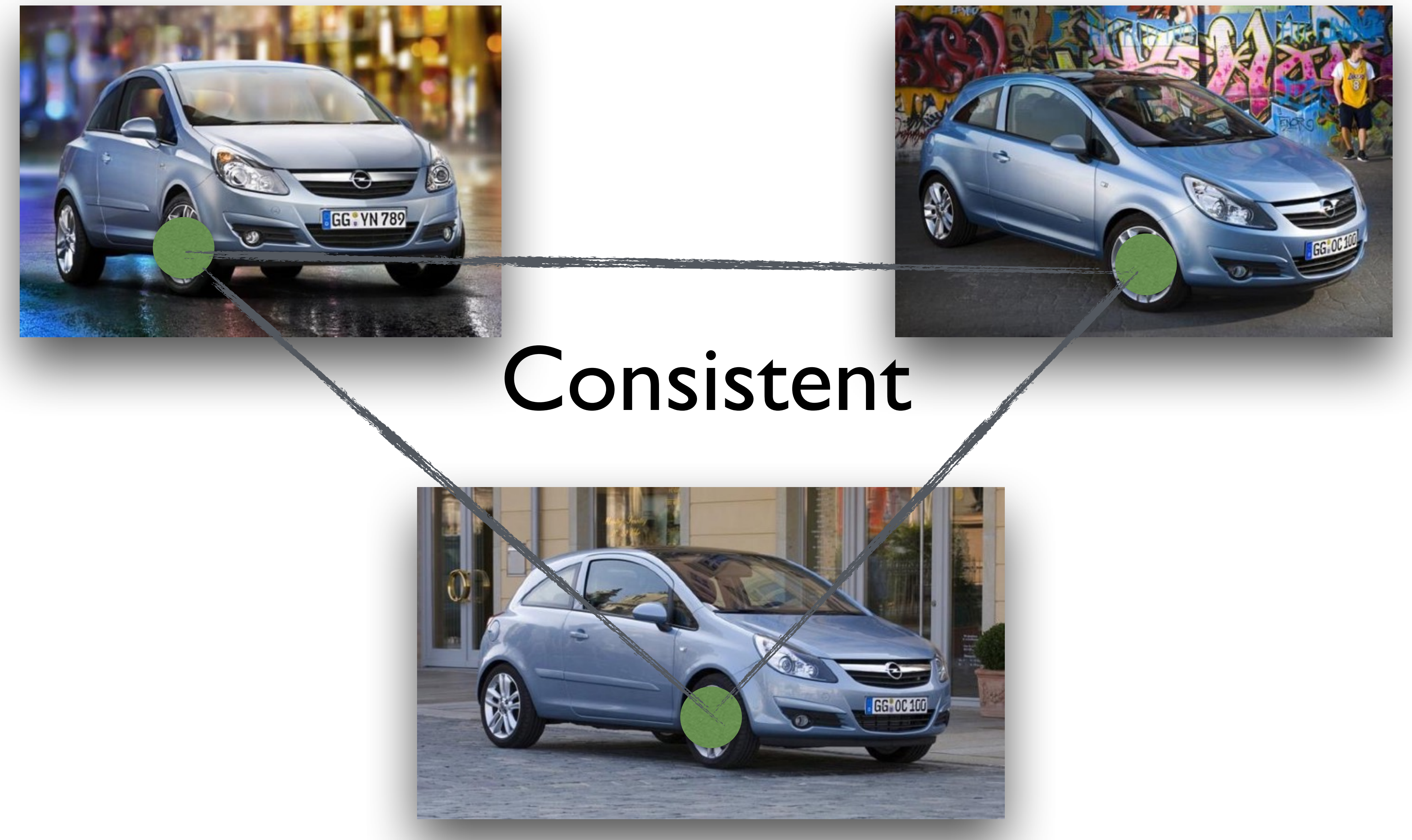
Why joint analysis?



Ambiguities resolved when looking at additional piece

Cycle consistency





The composition of maps along a cycle should be identity

$$m_{12} \circ m_{23} \circ \cdots \circ m_{n1} = I$$

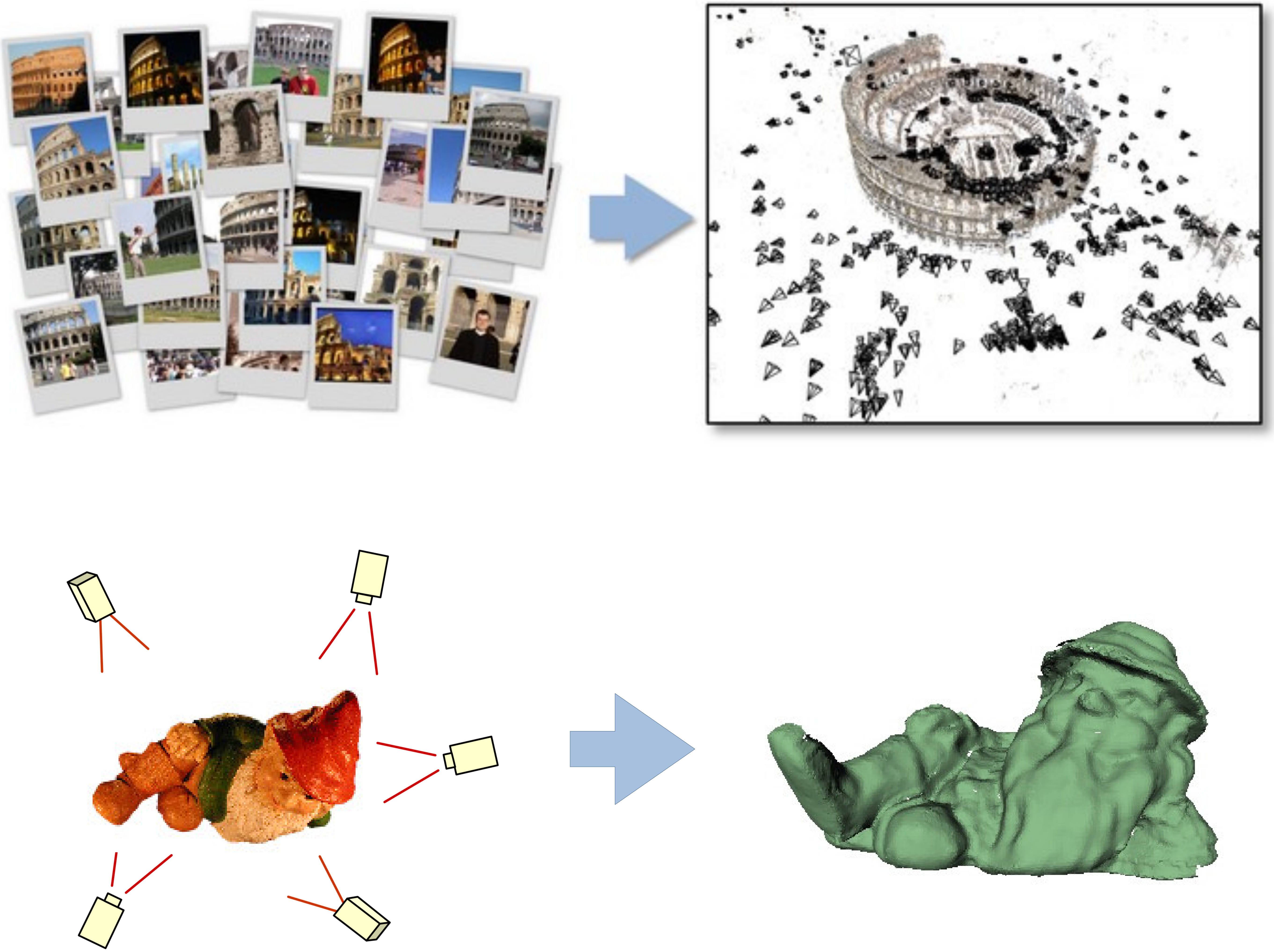
Part I
Map Synchronization for Pose Estimation

Pipeline of 3D reconstruction

1. Collect data from different viewpoints

2. Recover relative poses between views

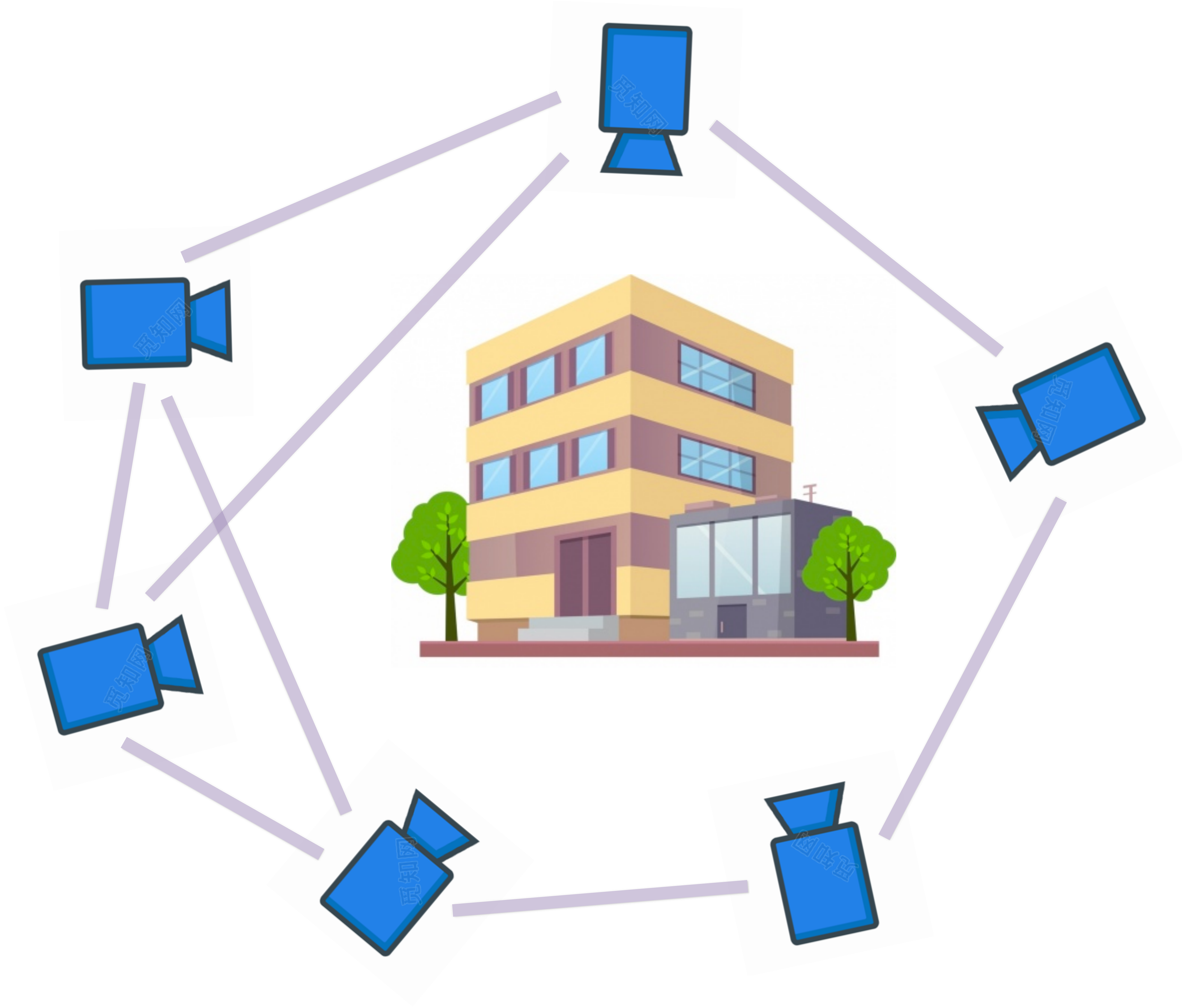
3. Reconstruct 3D model



Pipeline of 3D reconstruction

2. Recover relative poses between views

- Compute for each pair separately
- Pairwise estimation might be inaccurate or failed
- Joint optimization required



Pose optimization (synchronization)

Goal

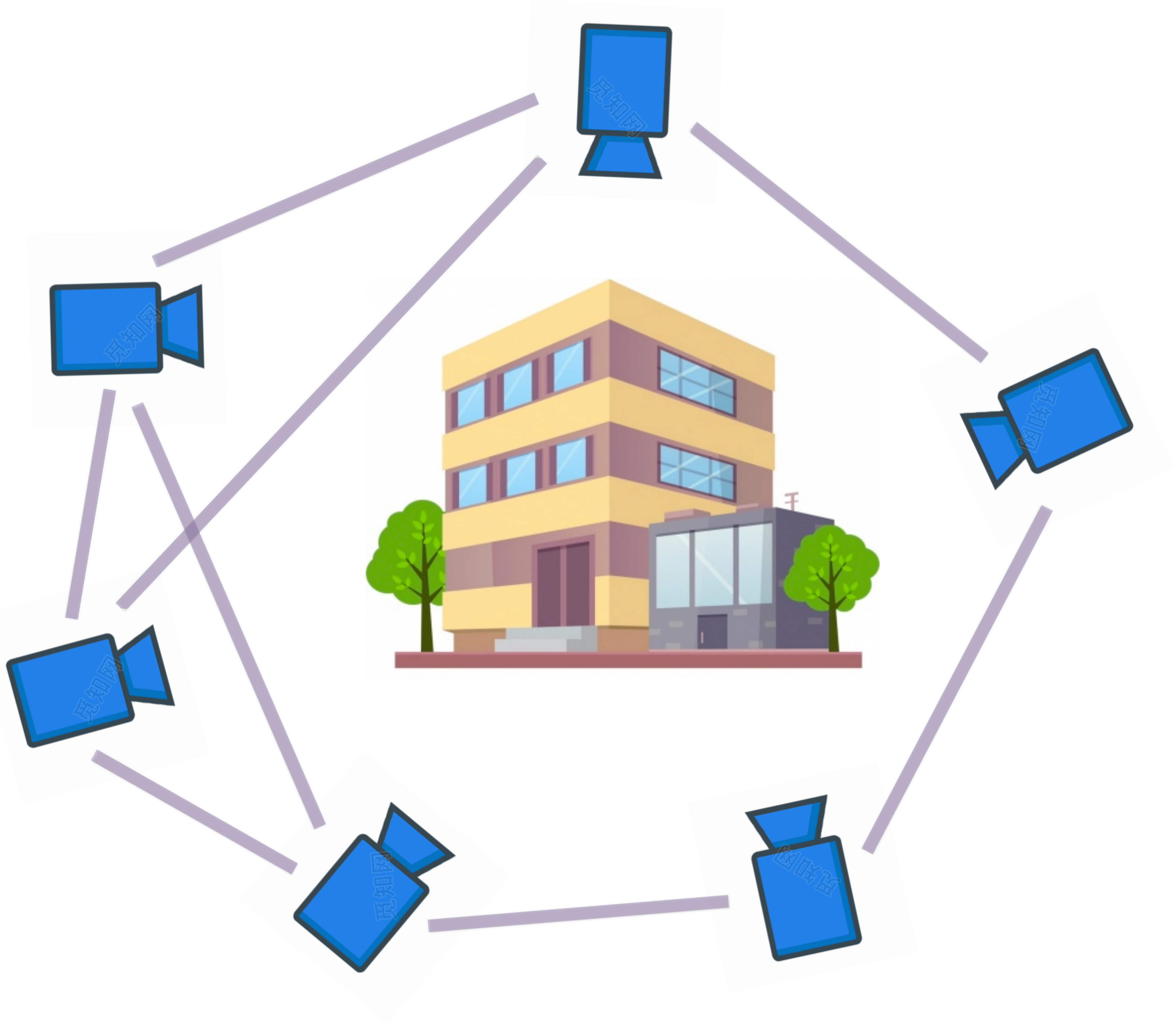
Given noisy pairwise pose measurements, jointly optimizing all of them in order to improve accuracy and reject outliers

How

Cycle consistency on pose graph!

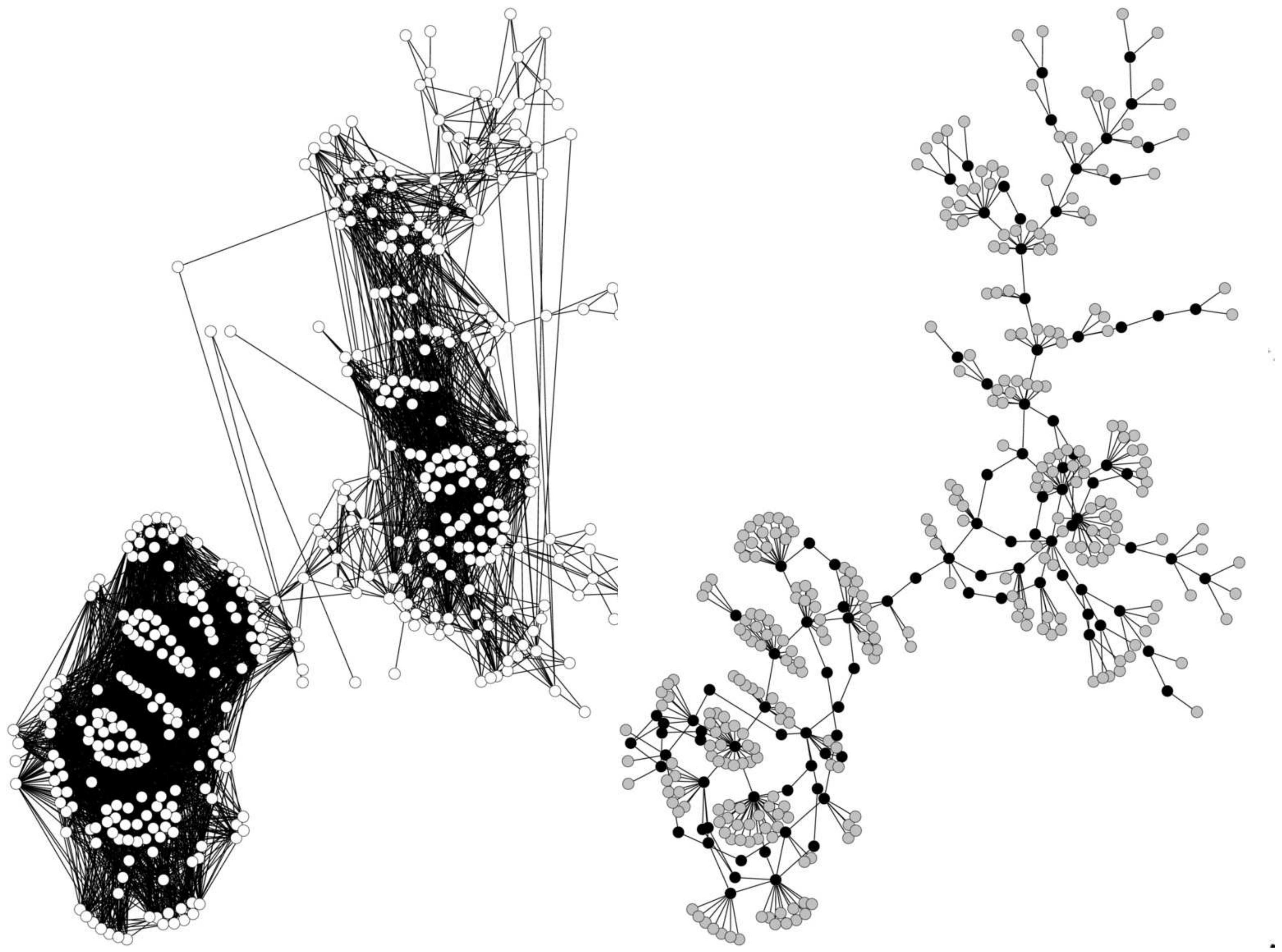
Three types of approaches

- Inlier/outlier inference
- Local, iterative optimization
- Global, factorization-based optimization

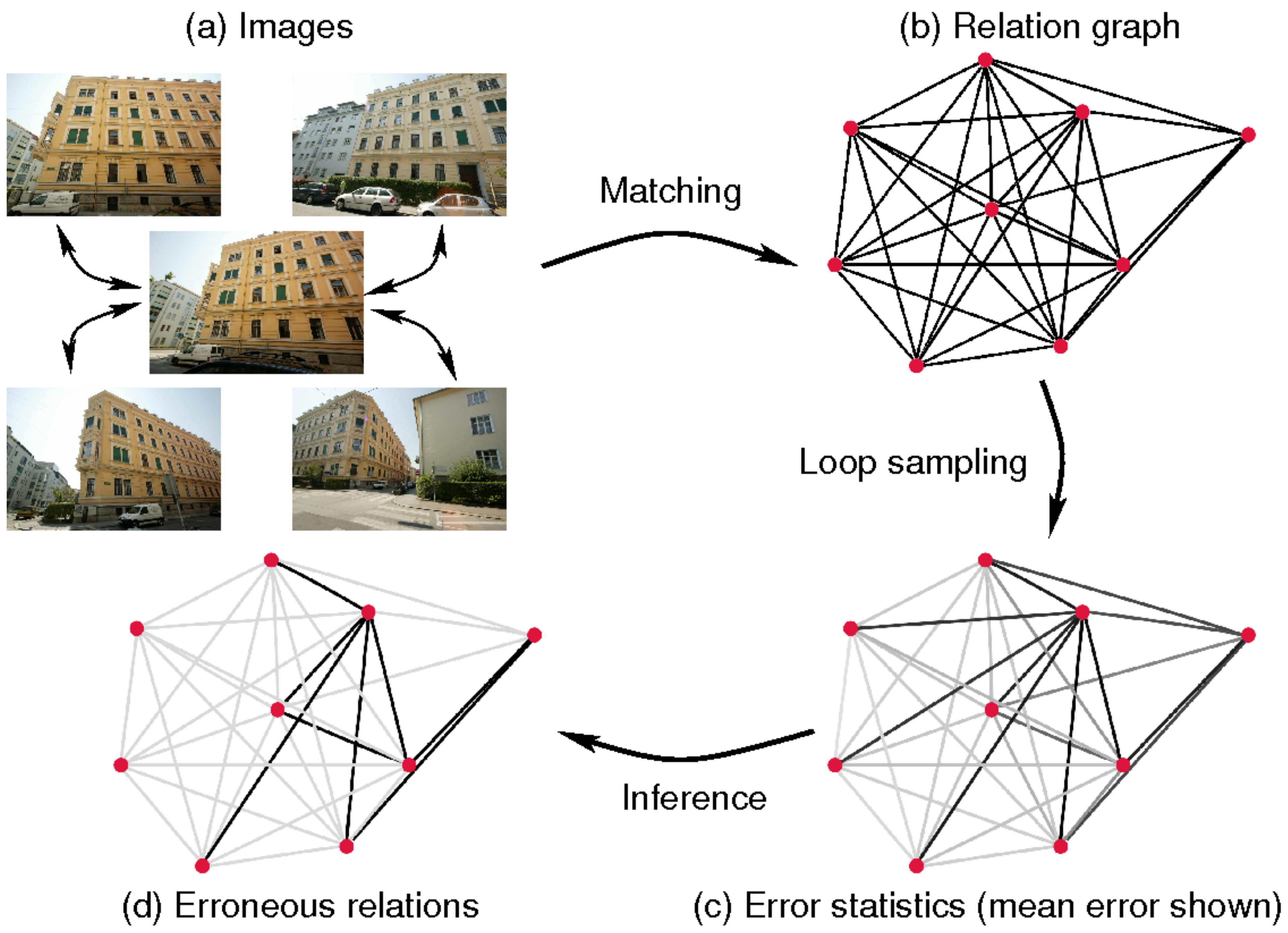


Inlier/outlier inference

Detect "good or bad" edges in the graph



Snavely et al. 2008



Zach et al. 2010

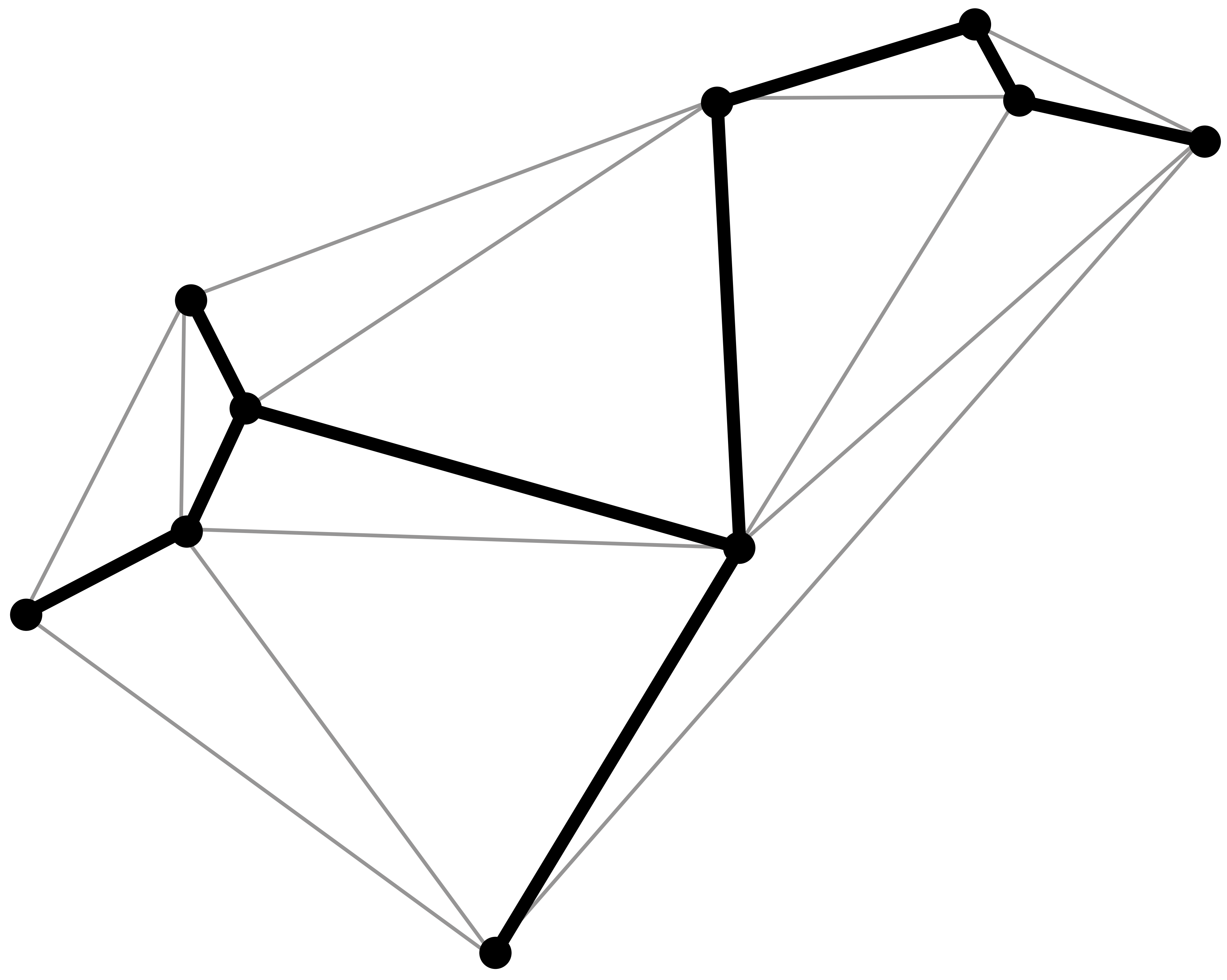
Inlier/outlier inference by spanning tree

<u>Find a tree of confident maps and composite others</u>

Node — image

Edge — map between images

Edge weight — confidence of map



<u>Limitation</u>

A single incorrect map can destroy everything

Optimization-based approach

Optimize pose variables subject to the cycle consistency constraint

Noisy measurements

$$\min_{X_{ij}} \sum_{(i,j)\in e} d\left(X_{ij} - M_{ij}\right)$$

$$st. \quad X_{ij} X_{jk} \cdots X_{zi} = I \quad \forall (i, j, k, \cdots, z) \in c$$

Limitation

Number of constraints (cycles) grows quickly with the number of nodes

Problem reformation

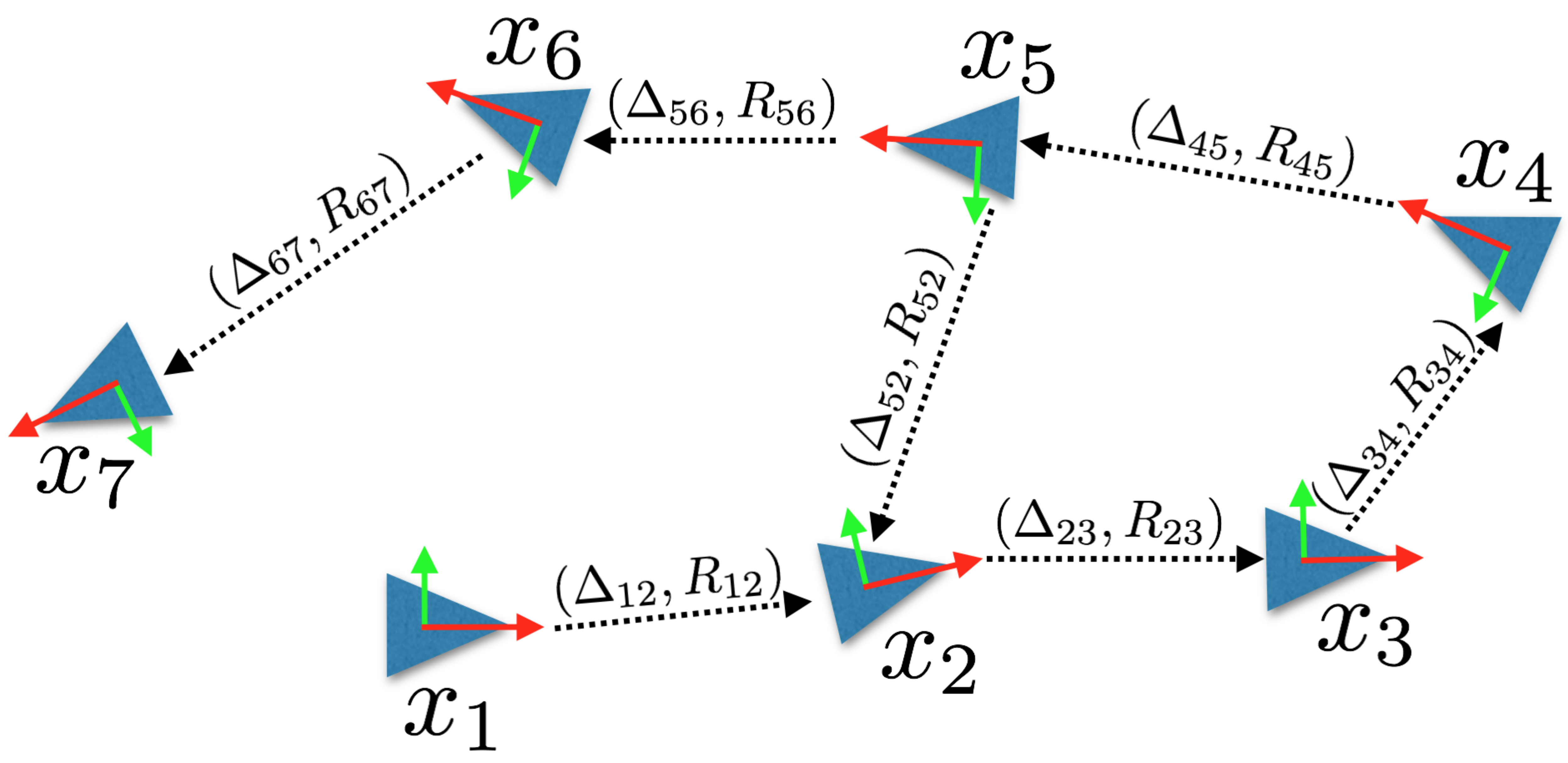
Estimate absolute pose for each node

$$\boldsymbol{x}_i \doteq (\boldsymbol{p}_i, \boldsymbol{R}_i)$$

which respect relative measurements

$$\boldsymbol{R}_{ij} = \boldsymbol{R}_i^\top \boldsymbol{R}_j$$
$$\boldsymbol{\Delta}_{ij} = \boldsymbol{R}_i^\top (\boldsymbol{p}_j - \boldsymbol{p}_i)$$

where $(\boldsymbol{\Delta}_{ij}, \boldsymbol{R}_{ij})$ are pairwise measurements



Carlone et al. (2015)

Cycle consistency is satisfied by construction!

e.g. $\boldsymbol{R}_{ij}\boldsymbol{R}_{jk}\boldsymbol{R}_{ki} = (\boldsymbol{R}_i^T\boldsymbol{R}_j)(\boldsymbol{R}_j^T\boldsymbol{R}_k)(\boldsymbol{R}_k^T\boldsymbol{R}_i) = \boldsymbol{I}$

Pose graph optimization

Noisy relative pose measurements

$$\min_{\substack{\{\boldsymbol{R}_i\}\in \mathrm{SO}(3)\\ \{\boldsymbol{t}_i\}\in\mathbb{R}^3}} \sum_{(i,j)\in\mathcal{E}} \mathrm{d}_{\mathbb{R}^3}\left(\boldsymbol{t}_{ij}, \boldsymbol{R}_i^\top(\boldsymbol{t}_j - \boldsymbol{t}_i)\right)^2 + \mathrm{d}_{\mathrm{SO}(3)}\left(\boldsymbol{R}_{ij}, \boldsymbol{R}_i^\top \boldsymbol{R}_j\right)^2$$

Relative poses constructed from absolute poses

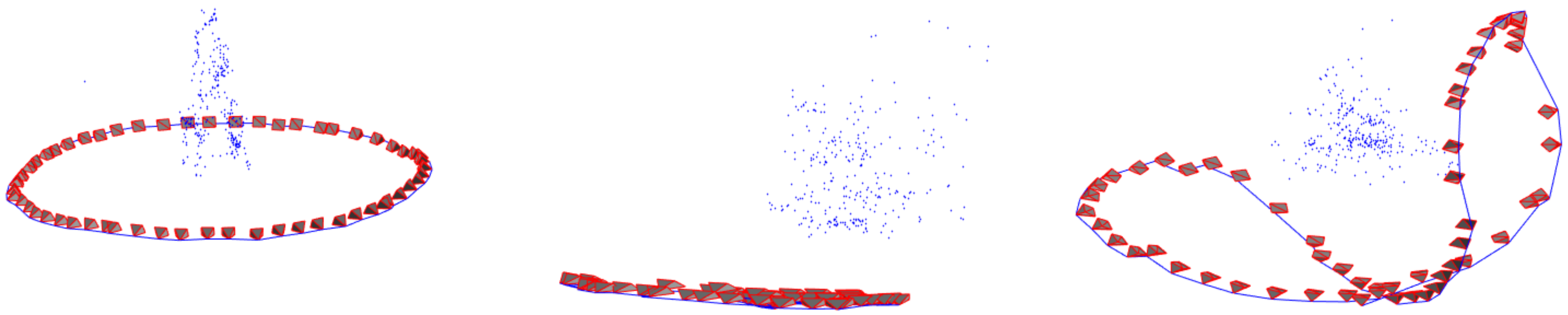
- Many choices of distance metrics
- Usually first solve rotation and then solve translation

Rotation optimization (averaging)

$$\min_{\{R_i\}_{i\in V}\in SO(3)^N} \sum_{(i,j)\in E} \ell(R_i R_i^{\mathrm{T}}, \tilde{R}_{ij})$$

- Many choices of loss functions, parameterizations, and optimization methods
  Examples: Crandall (2011), Chatterjee (2013), Tron (2014)
  Surveys: Hartley(2013), Carlone(2015), Tron (2016)

- Nonconvex, different initialization leads to different local minima



Eriksson (2018)

Spectral relaxation [Arie-Nachimson 12, Bernard 15, Arrigoni 16]

Use least-squares loss and write loss in matrix form

$$\left[\tilde{G}\right]_{ij;3\times 3} = \begin{cases} \tilde{R}_{ij} & \text{if } (i,j) \in \mathcal{E} \\ \mathbf{0} & \text{otherwise} \end{cases}$$

$$\min \sum_{(i,j)\in\mathcal{E}} \left\| \widetilde{\boldsymbol{R}}_{ij} - \boldsymbol{R}_i^{\mathrm{T}} \boldsymbol{R}_j \right\|_{\mathrm{F}}^2 \longrightarrow \max\ \mathrm{tr}(\mathbf{R}^{\mathrm{T}} \tilde{G} \mathbf{R})$$

$$\mathbf{R} = \begin{bmatrix} \boldsymbol{R}_1 \\ \boldsymbol{R}_2 \\ \vdots \\ \boldsymbol{R}_n \end{bmatrix}$$

Relax constraints on rotations

$$\{\boldsymbol{R}_i\} \in \mathrm{SO}(3) \longrightarrow \mathbf{R}^{\mathrm{T}}\mathbf{R} = I$$

Then the problem becomes

$$\max\ \mathrm{tr}(\mathbf{R}^{\mathrm{T}} \tilde{G} \mathbf{R}) \quad \text{Subject to } \mathbf{R}^{\mathrm{T}}\mathbf{R} = I$$

Analytically solved by Eigenvalue decomposition!

SDP relaxation

[Arie-Nachimson 12, Fredriksson 12, Wang 13, Rosen 16, Carlone 15, Eriksson 18]

Rewrite the loss

$$\mathrm{tr}(\mathbf{R}^{\mathrm{T}}\tilde{G}\mathbf{R}) \longrightarrow \mathrm{tr}(\tilde{G}G) \qquad \text{where } G = \mathbf{R}\mathbf{R}^{\mathrm{T}}$$

Ignore SO(3) constraints, the problem becomes semidefinite program (SDP)

$$\max_{G\succeq 0, [G]_{ii;3\times 3}=I} \mathrm{tr}(\tilde{G}G)$$

Convex and provable exact recovery [Wang 13, Rosen 16, Ericsson 18]

Robust factorization [Wang and Singer, 2013]

Solution 1: Robust loss function in SDP formation

$$\min_{G} \sum_{(i,j)\in\mathcal{E}} \|G_{ij} - R_{ij}\| \text{ s.t. } G_{ii} = I_d, \text{ and } G \succcurlyeq 0.$$

Solution 2: Reweighted spectral decomposition (= reweighted least squares)

$$\max_{\mathbf{R}^{\mathrm{T}}\mathbf{R} = I} \operatorname{tr}(\mathbf{R}^{\mathrm{T}}\tilde{G}\mathbf{R})$$

$$\left[\tilde{G}\right]_{ij;3\times 3} = \begin{cases} w_{ij}\tilde{R}_{ij} & \text{if } (i,j)\in\mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

Weights indicate the confidence of pairwise measurement and computed from the residual in the previous iteration

Summary

Three types of methods

- Inlier/outlier inference
- Local, iterative optimization
- Global, factorization-based optimization

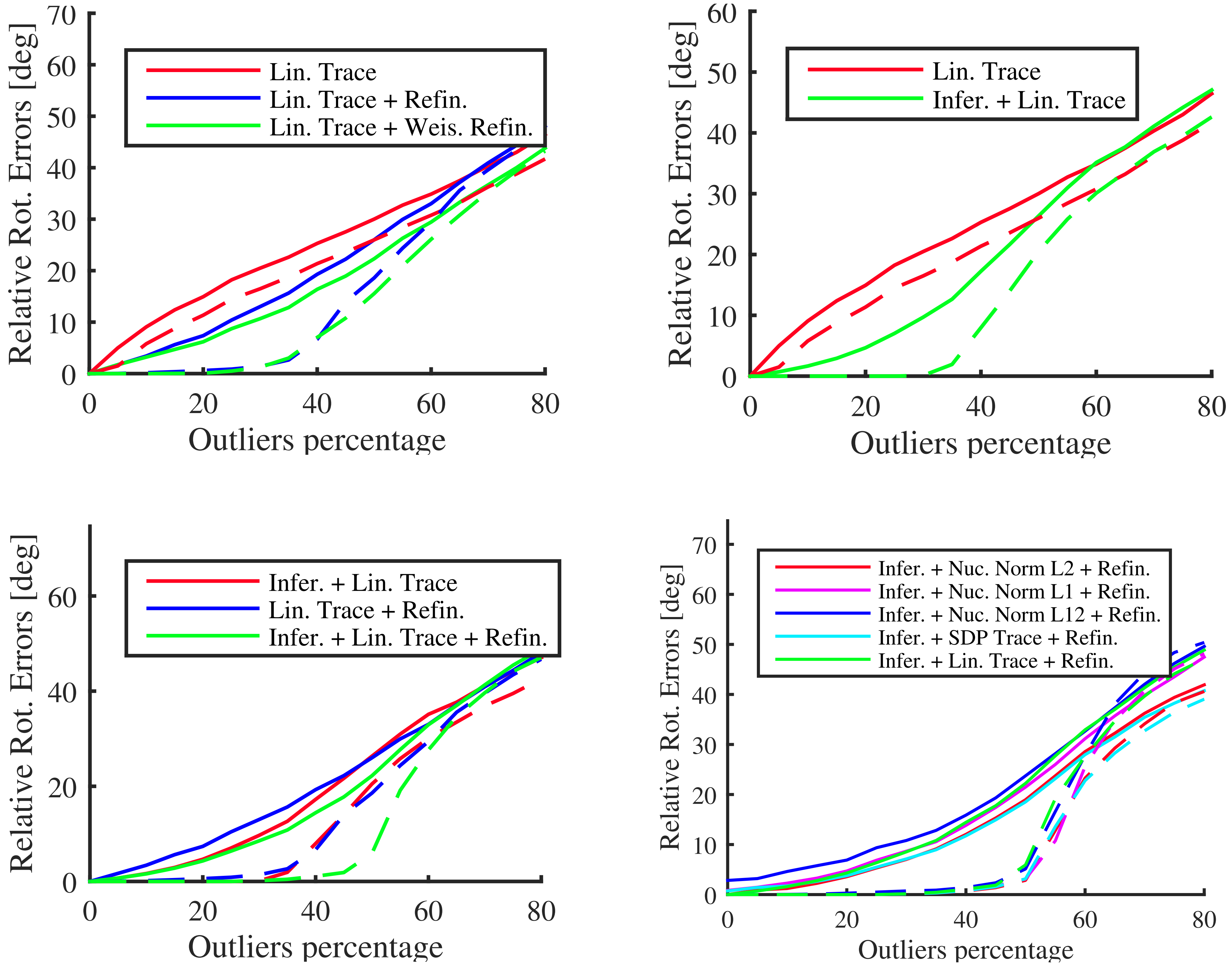
The best practice?

The combination of three!

1. Prune outliers by inference
2. Initialize by factorization
3. Refine by local optimization

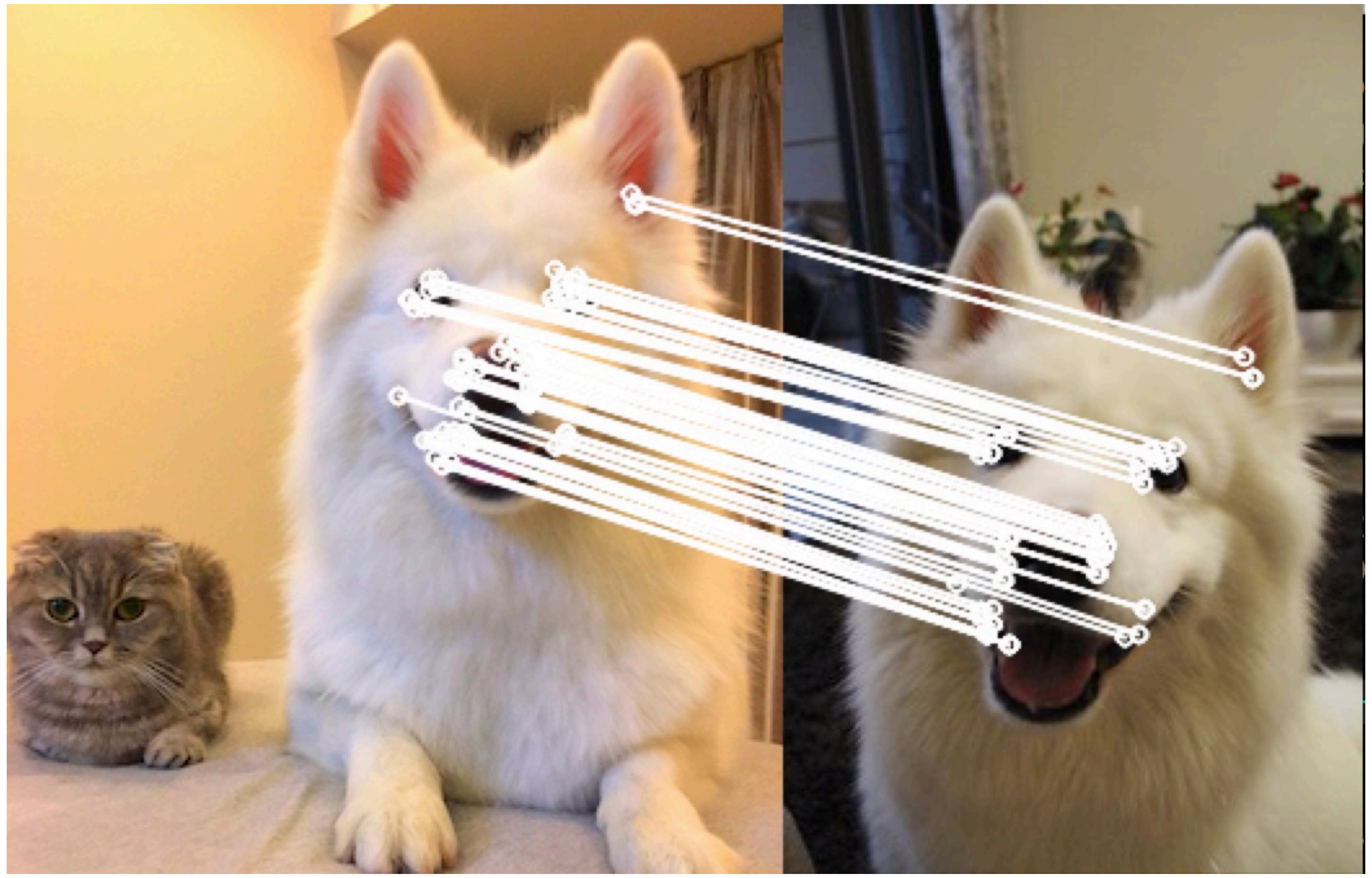
Translation synchronization

e.g. [Brand 04, Jiang 13, Wilson 16, Huang 17]



Tron et al., 2016. A Survey on Rotation Optimization in Structure from Motion

Part II
Map Synchronization for Correspondence Estimation



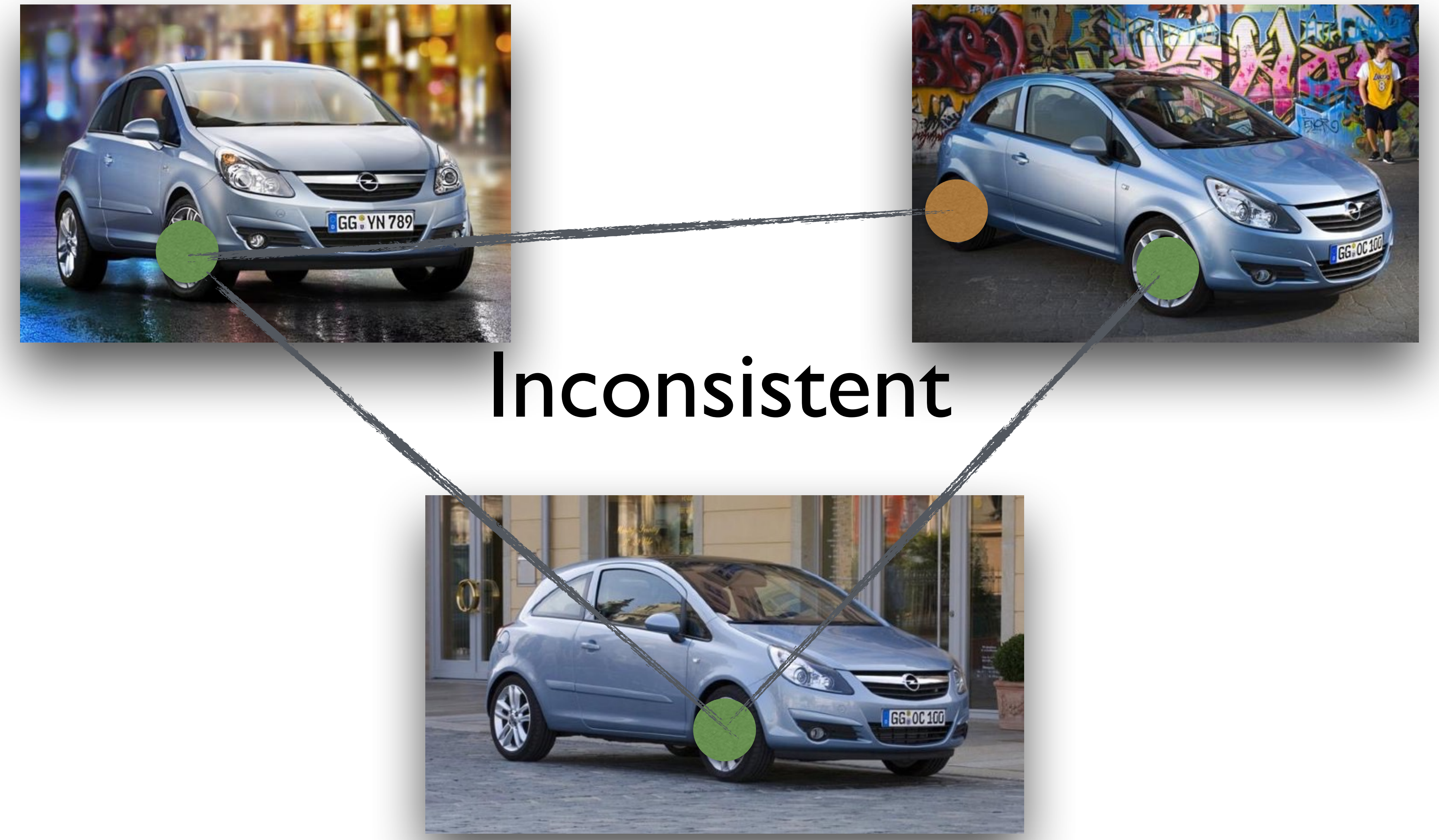
Correspondence problem

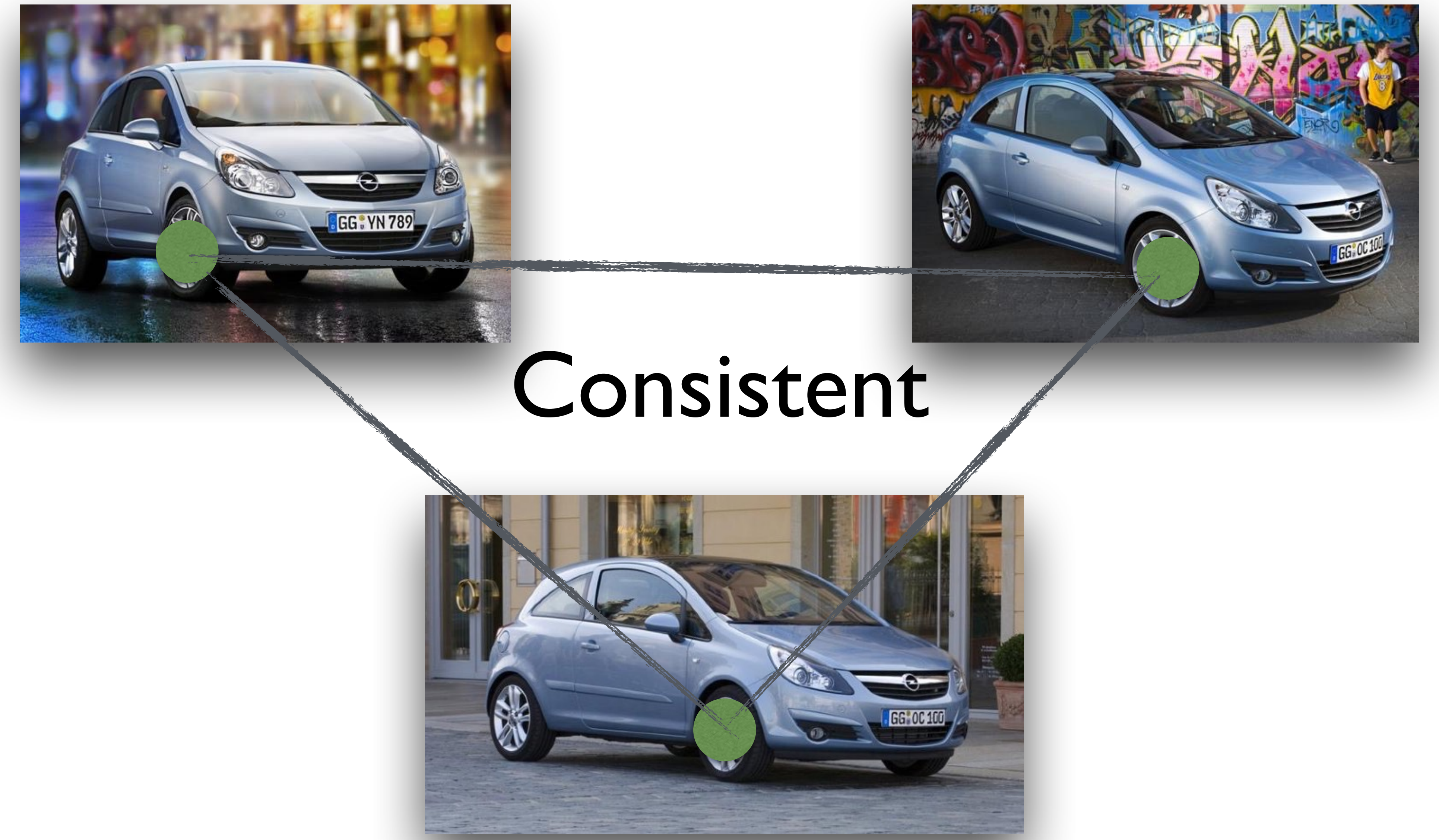
Image $\boldsymbol{i}$ Image $\boldsymbol{j}$

Represented by
permutation matrix

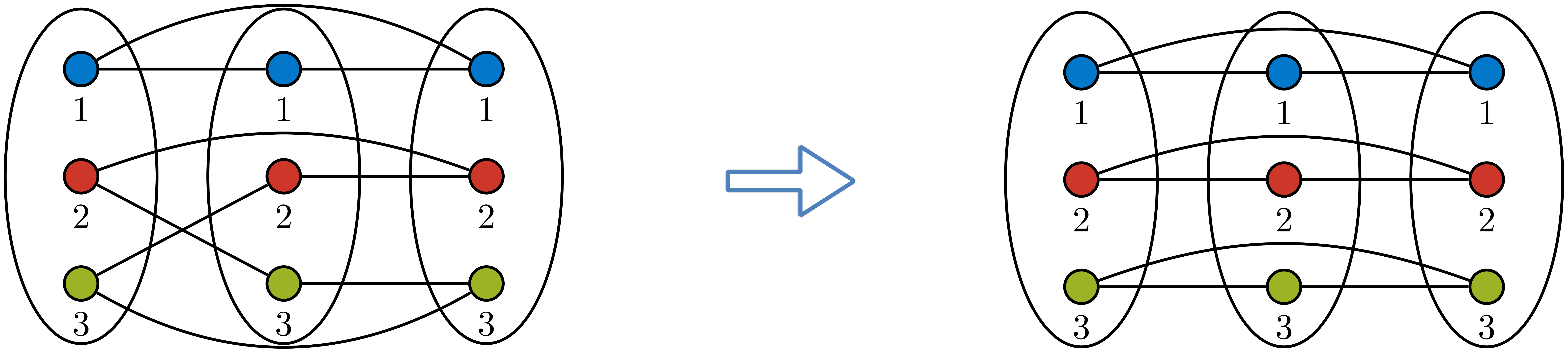
$$\boldsymbol{X}_{ij} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Cycle consistency is also desired





Permutation synchronization



Input: pairwise correspondences from existing algorithms which may be noisy

Output: cycle-consistent correspondences which respect the input

NP-Complete [Huber 2002]

Approaches

Inliner/outlier inference

[Huber 01, Huang 06, Cho 08, Zach10, Nguyen11, Crandel 11, Huang 12, Zhou15]

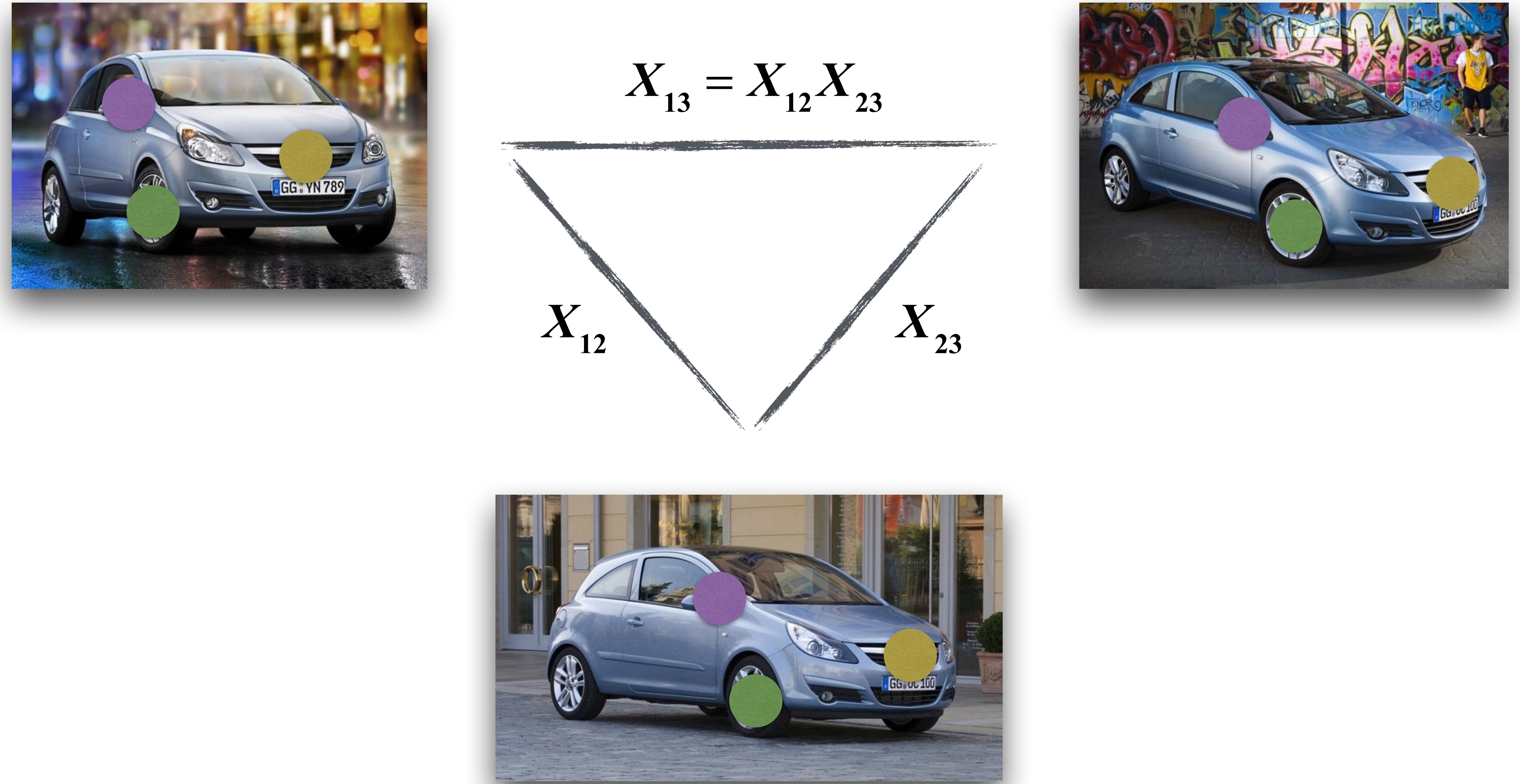
Local, iterative optimization

[Yan 13, 14, 15]

Global, factorization-based optimization

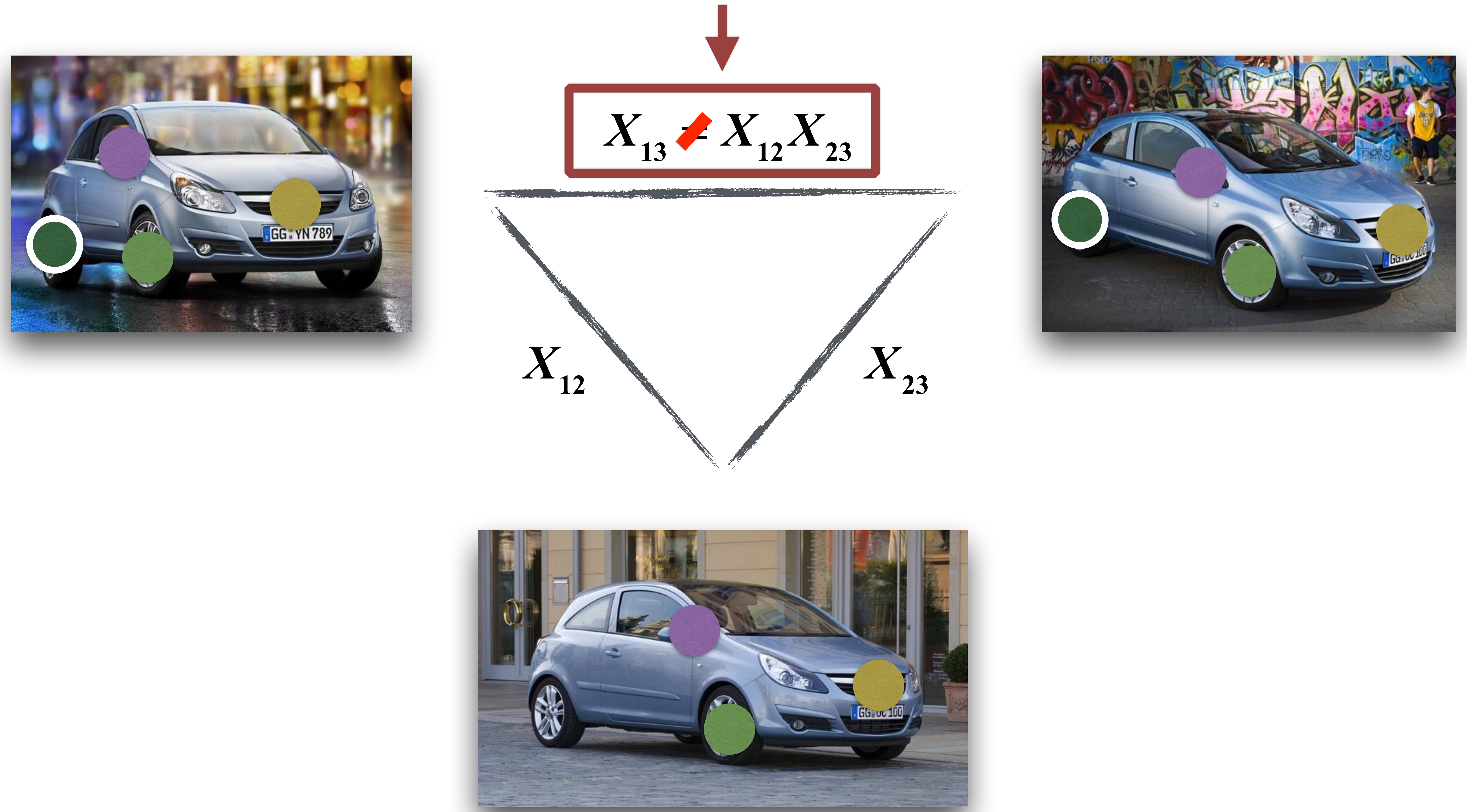
[Huang 13, Pachauri 13, Chen 14, Zhou15]

Cycle consistency for permutation matrices



Cycle consistency for permutation matrices

Not true if the feature sets are not the same



Partial similarity

Feature sets are not always the same

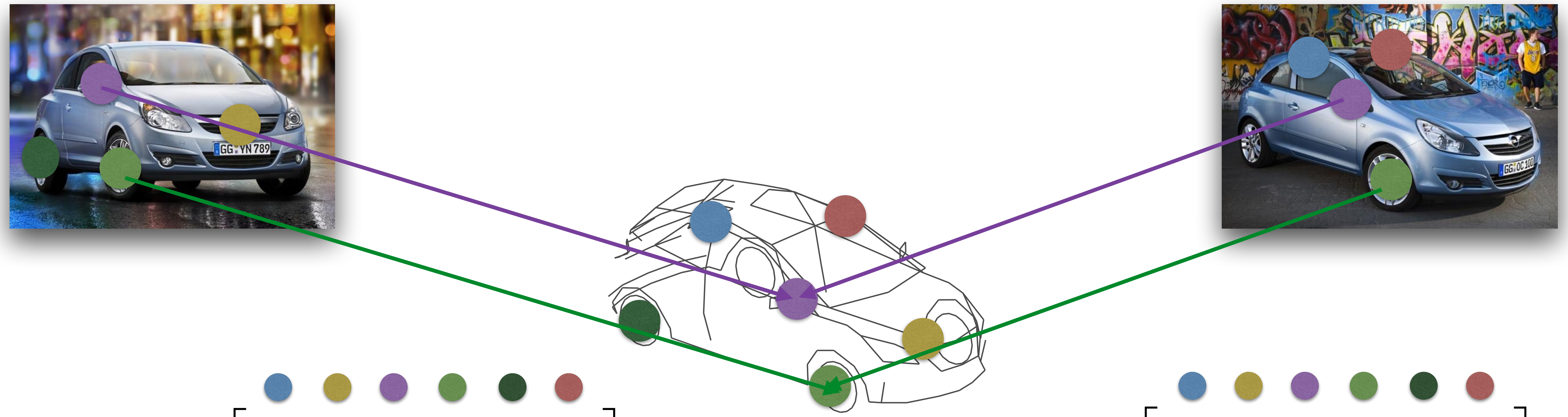
— correspondences represented by partial permutation matrix

$$X_{13} \neq X_{12}X_{23}$$

How can we represent the cycle consistency for partial permutation matrices?

Cycle consistency under partial similarity

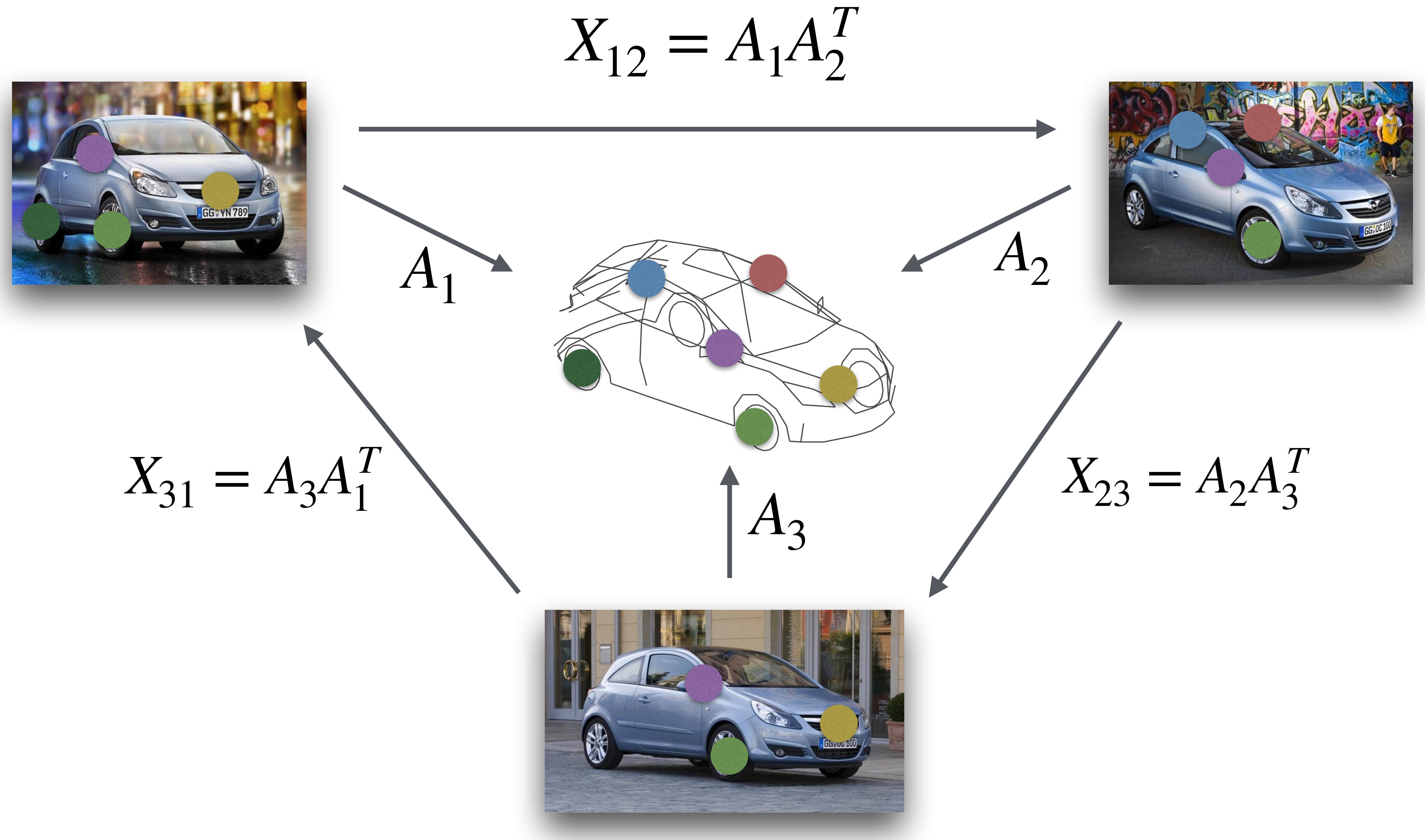
Map to a latent feature space (**universe**) [Huang 13, Chen 14]



$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Cycle consistency under partial similarity



A and ***X*** are conceptually similar to absolute pose and relative pose

Cycle consistency = positive semidefinite + low rank

Map between images

Map from image to universe

$$\underbrace{\begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \begin{bmatrix} A_1^T & A_2^T & \cdots & A_n^T \end{bmatrix}}_{}$$

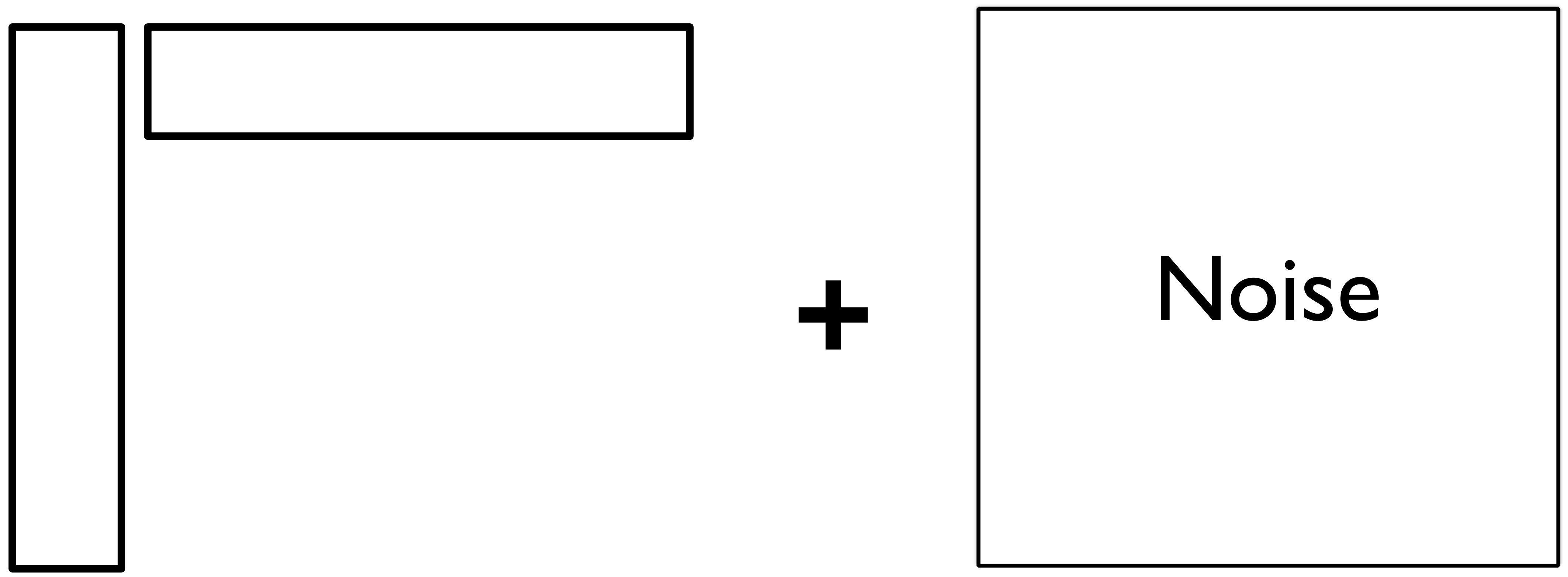
$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{A}^T$$

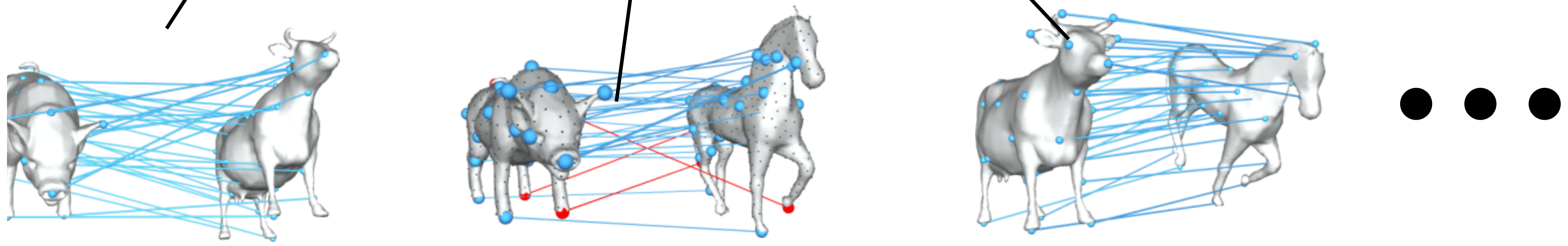
The maps are cycle consistent [Huang et al. 2013]

= X can be factorized as above = X is positive semidefinite and low-rank

Permutation synchronization by matrix decomposition

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





Noisy measurements of matrix blocks

[Huang et al. 2013]

Permutation synchronization by spectral method

Input: the objective matrix $\mathcal{T}$ $= X^{input}$

Compute the n leading eigenvectors $(v_1, v_2, \ldots, v_n)$ of $\mathcal{T}$ and set $U = \sqrt{m}\,[v_1, v_2, \ldots, v_n]$

for $i = 1$ to m **do**

$P_{i1} = U_{(i-1)n+1:in,\,1:n}\, U^{\top}_{1:n,\,1:n}$

$\sigma_i = \arg\max_{\sigma \in \mathbb{S}_n} \langle P_{i1}, \sigma \rangle$ [Kuhn-Munkres]

end for

for each (i, j) **do**

$\tau_{ji} = \sigma_j \sigma_i^{-1}$

end for

Output: the matrix $(\tau_{ji})_{i,j=1}^{m}$ of globally consistent matchings

Eigenvalue decomposition

Discretization

Pachauri et al. (2013). Solving the multi-way matching problem by permutation synchronization.

Permutation synchronization by convex optimization

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

Constraints: $X \succeq 0$ ← Positive semidefinite (cycle consistency)

$$X_{ii} = I_m, \quad 1 \le i \le n$$

$$X_{ij}\mathbf{1} = \mathbf{1}, X_{ij}^T\mathbf{1} = \mathbf{1}, \quad 1 \le i < j \le n$$

$$0 \le X \le 1$$

↖ Relaxed constraints on permutation matrices

Huang and Guibas (2013). Consistent shape maps via semidefinite programming.

Permutation synchronization by convex optimization

$$\begin{aligned} \text{minimize} \quad & \sum_{(i,j)\in\mathcal{E}} \|\mathbf{X}_{ij}^{input} - \mathbf{X}_{ij}\|_1 \\ \text{subject to} \quad & \mathbf{X}_{ii} = I_{m_i}, \quad 1 \le i \le n \\ & \begin{bmatrix} m & \mathbf{1}^T \\ \mathbf{1} & \mathbf{X} \end{bmatrix} \succeq 0 \\ & \mathbf{X} \ge 0 \end{aligned}$$

Chen et al. (2014). Near-optimal joint object matching via convex relaxation.

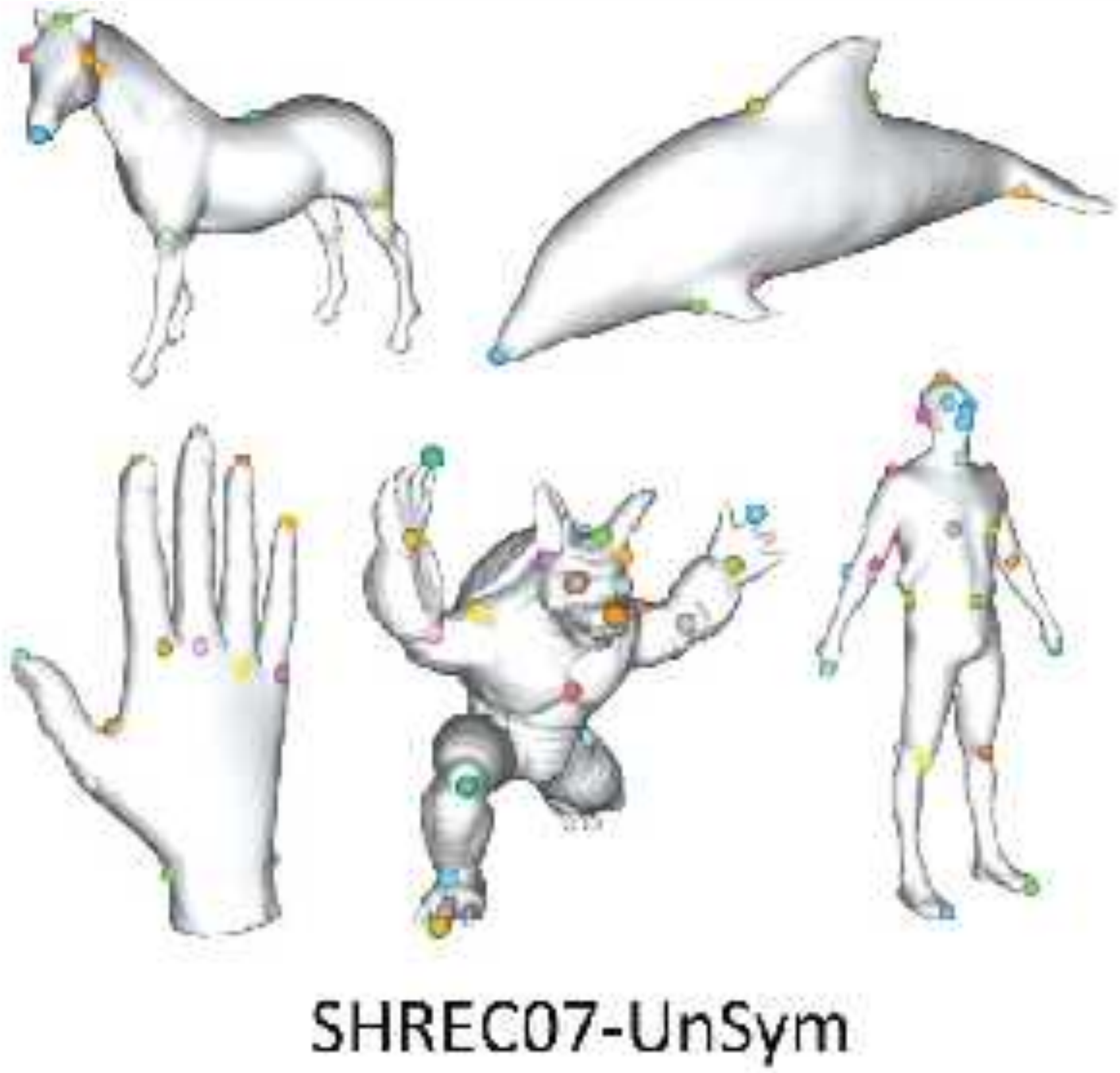
Provable exact recovery of MatchLift [Chen 2014]

- Theorem [CGH'14]: *The underlying permutations can be recovered w.h.p if*

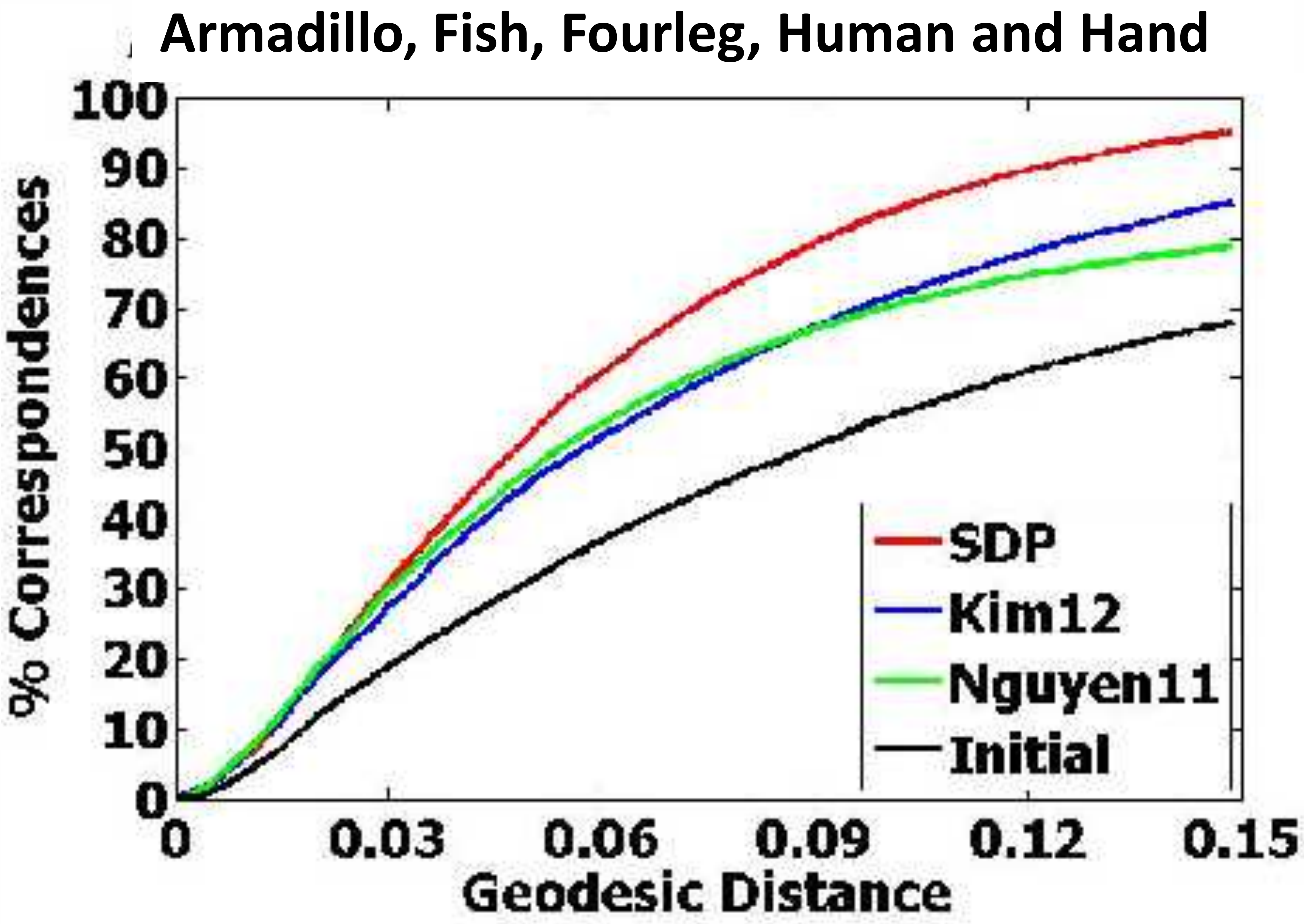
$$p_{\text{true}} \geq c\frac{\log^2(mn)}{\sqrt{np_{\text{obs}}}}$$

  - p_{obs}: the probability that two objects connect
  - p_{true}: the probability that a pair-wise map is correct
  - Incorrect maps are random permutations

Comparison to previous methods



20 objects, 128 points per object



Solving optimization

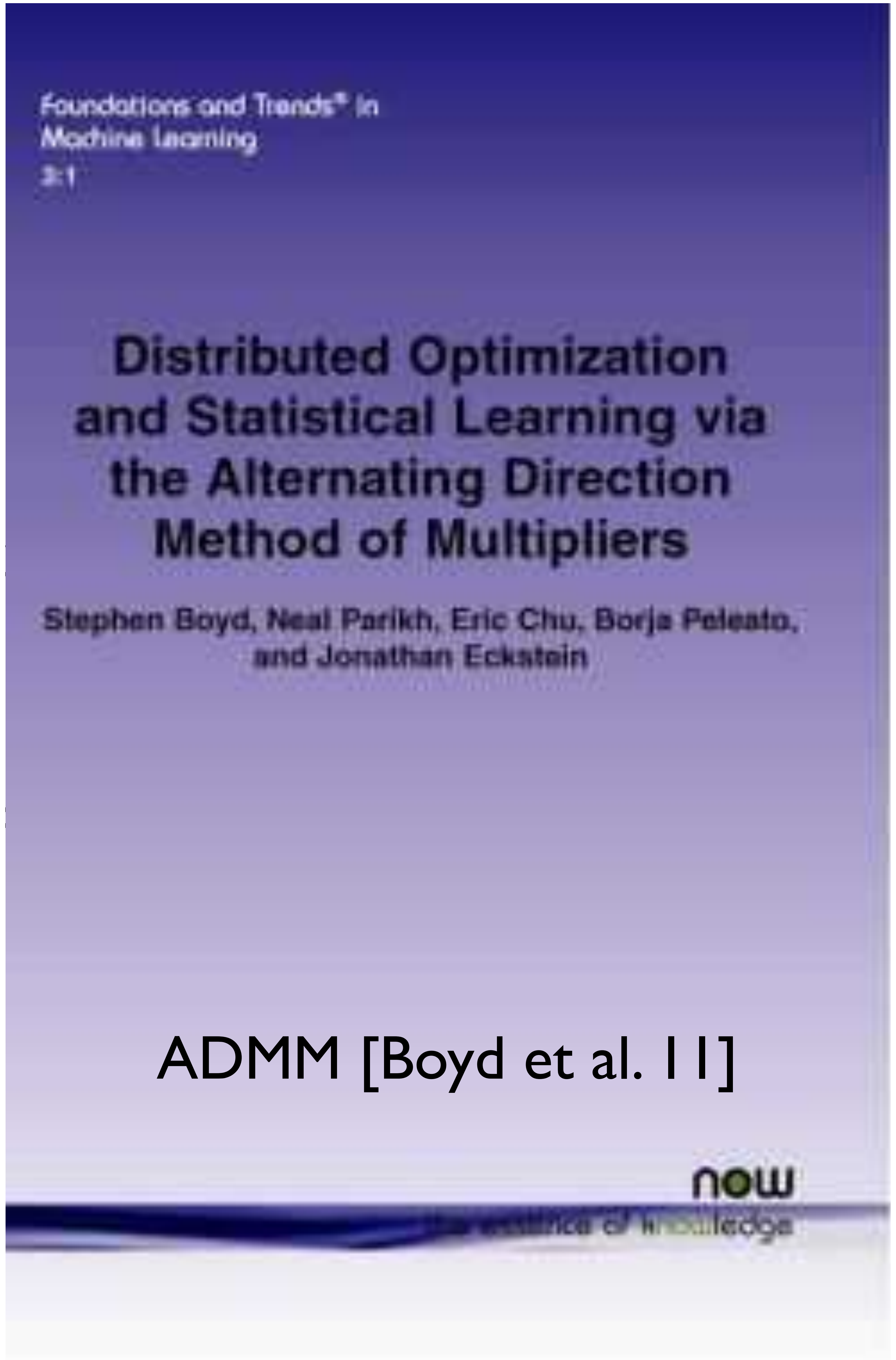
$$\text{minimize} \quad \sum_{(i,j)\in\mathcal{E}} \|\mathbf{X}_{ij}^{input} - \mathbf{X}_{ij}\|_1$$

$$\text{subject to} \quad \mathbf{X}_{ii} = I_m, \quad 1 \le i \le n$$

$$\mathbf{X}_{ij}\mathbf{1} = \mathbf{1}, \mathbf{X}_{ij}^T\mathbf{1} = \mathbf{1}, \quad 1 \le i < j \le n$$

$$\mathbf{X} \succeq 0$$

$$\mathbf{X} \ge 0$$



ADMM [Boyd et al. 11]

Make the optimization more efficient?

$$X = \underbrace{\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}}_{\text{universe size}} \begin{bmatrix} A_1^T & A_2^T & \cdots & A_n^T \end{bmatrix}$$

X is both positive semidefinite and low-rank
Rank is bounded by the size of universe

Low-rank formulation (MatchALS)

Pairwise matching cost

$$\sum_{i,j} \text{trace}\left(\boldsymbol{W}_{ij}^T \boldsymbol{X}_{ij}\right)$$

Nuclear norm
(sum of singular values)

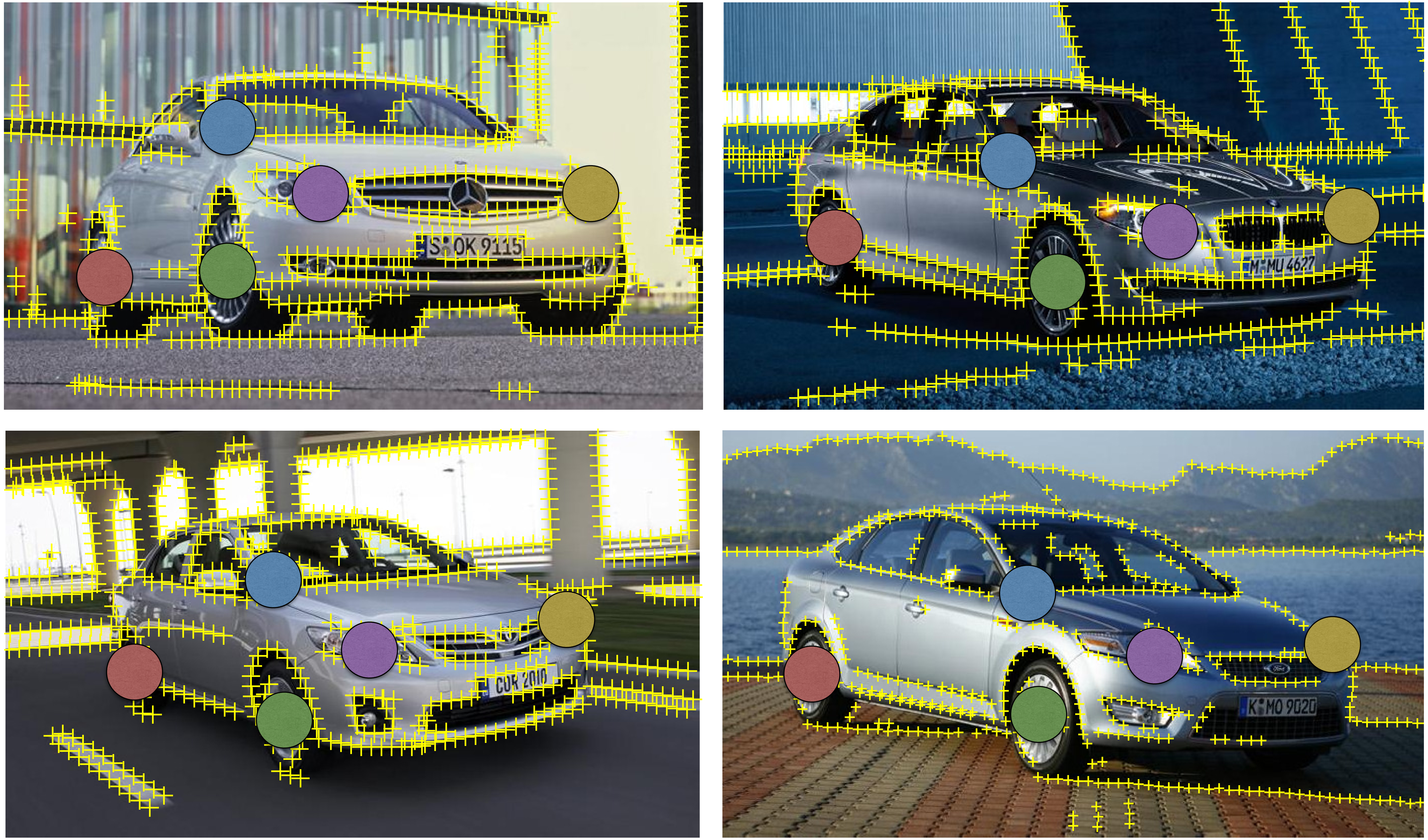
$$\min_{\boldsymbol{X}} \ \langle \boldsymbol{W}, \boldsymbol{X} \rangle + \lambda \|\boldsymbol{X}\|_*$$

$$\text{s.t. } \boldsymbol{X} \in \mathcal{C}$$

$$\boldsymbol{X}_{ij} = \boldsymbol{X}_{ji}^T \quad \boldsymbol{X}_{ii} = \boldsymbol{I}_{p_i} \quad \boldsymbol{0} \leq \boldsymbol{X} \leq \boldsymbol{1}$$

Zhou, Zhu and Daniilidis (2015). Multi-image matching via fast alternating minimization.

Mining consistent features



Feature point selection by optimization

Input pairwise maps

$$\min_{\boldsymbol{X}} \quad \frac{1}{4}\|\boldsymbol{W} - \boldsymbol{X}\boldsymbol{X}^T\|_F^2$$

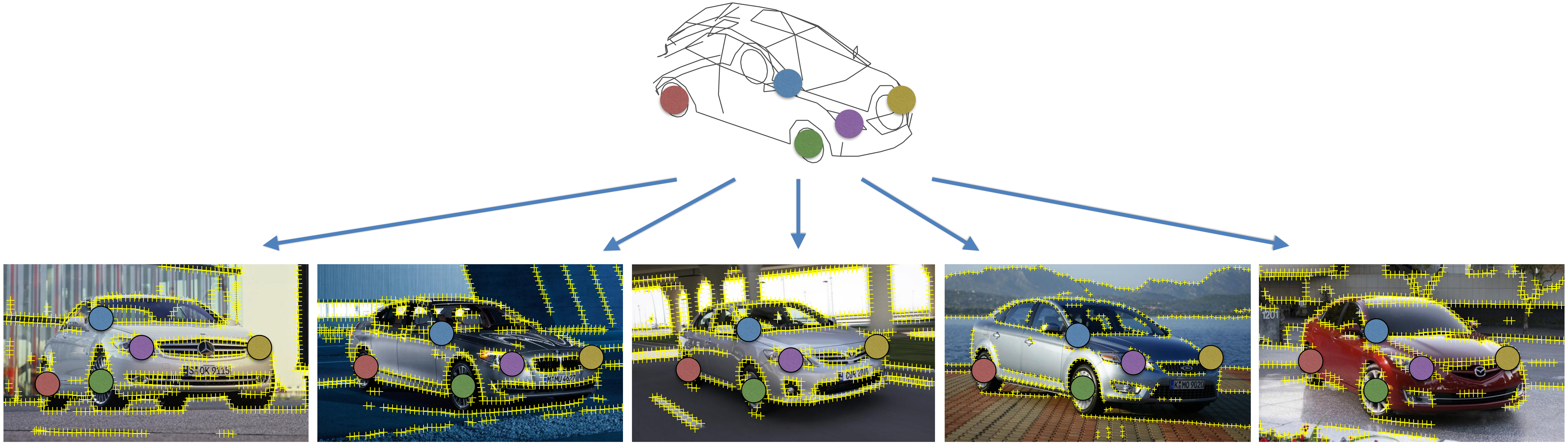
$$\text{s.t.} \quad \boldsymbol{X}_i \in \mathbb{P}^{p_i \times k}, 1 \leq i \leq n$$

$$\boldsymbol{X}_i = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{bmatrix}$$

Variable size is small as k<<p

Wang, Zhou, & Daniilidis. (2017). Multi-Image Semantic Matching by Mining Consistent Features.

Geometric constraint



The consistent feature points are 2D projections of the same or similar 3D structures

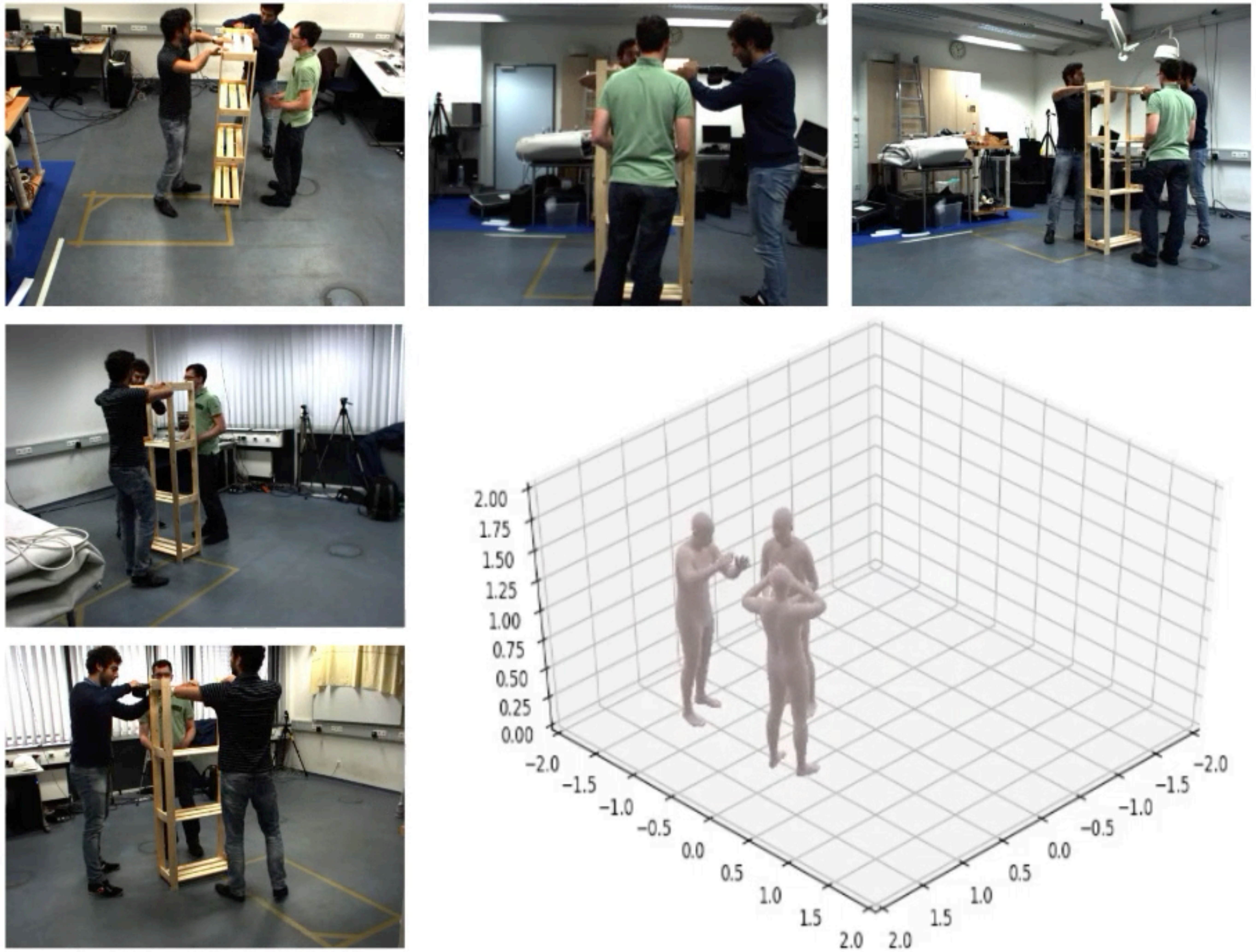
Wang, Zhou, & Daniilidis. (2017). Multi-Image Semantic Matching by Mining Consistent Features.

Application — automatic landmark annotation

Application — automatic landmark annotation

1000 cat head images

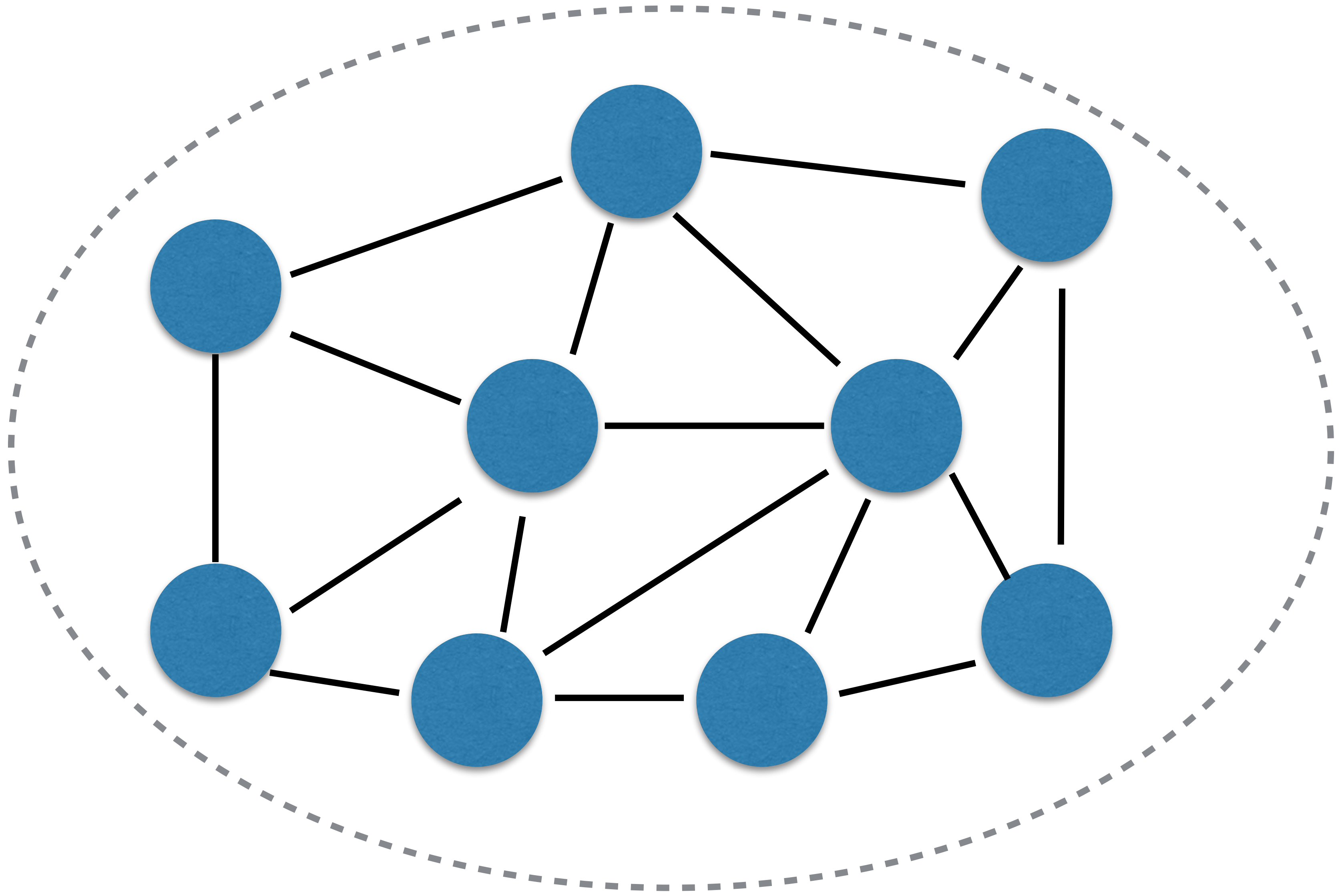
Application — cross-view matching for marker-less motion capture



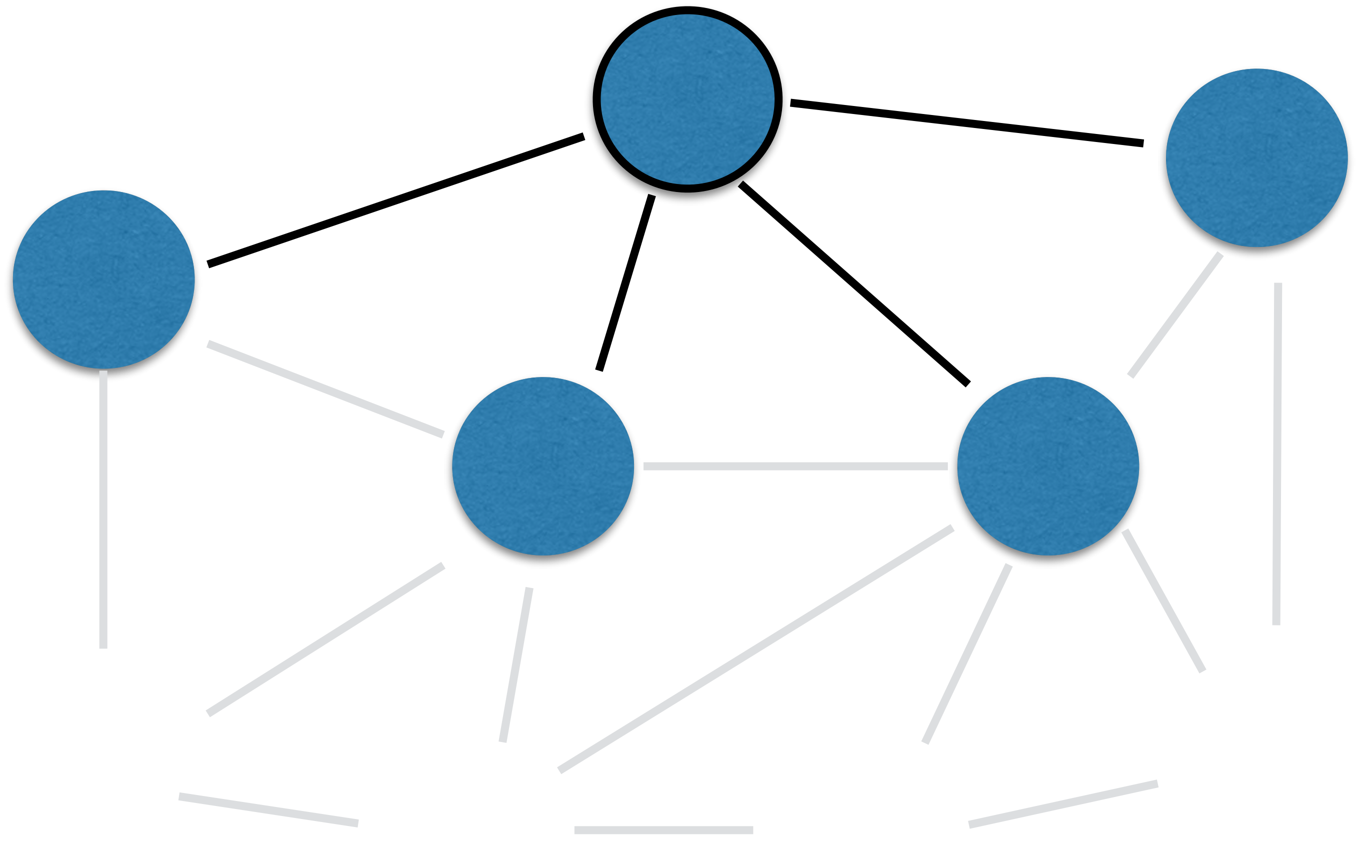
Dong et al., CVPR 2019. Fast and Robust Multi-Person 3D Pose Estimation from Multiple Views.

Part III
More research directions

Distributed algorithms



Centralized: all data is available and processed at once



Distributed: data in a subgraph is available and processed locally

Distributed multi-way matching

Algorithm

Previous estimate

Estimates from neighbors

$$\Pi_i(t+1) = \frac{1}{|\mathcal{N}_i| + 1}\left(\Pi_i(t) + \sum_{j \in \mathcal{N}_i} \widetilde{\Pi}_{ij}\Pi_j(t)\right)$$

Results

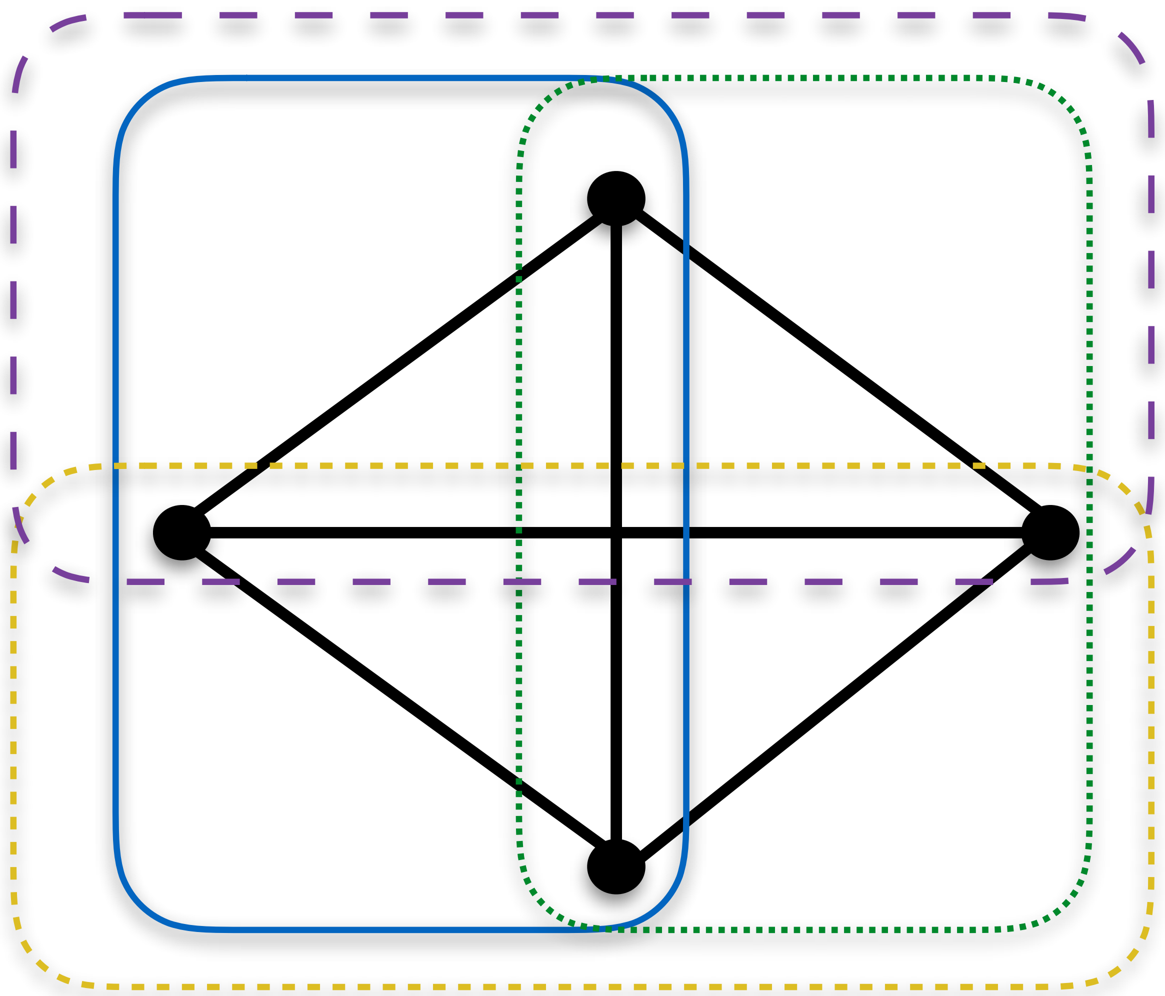
- Guaranteed to converge to true labels in noiseless case
- Solutions are always doubly stochastic matrices

Leonardos, Zhou, Daniilidis (2016). Distributed consistent data association.

Distributed multi-way matching

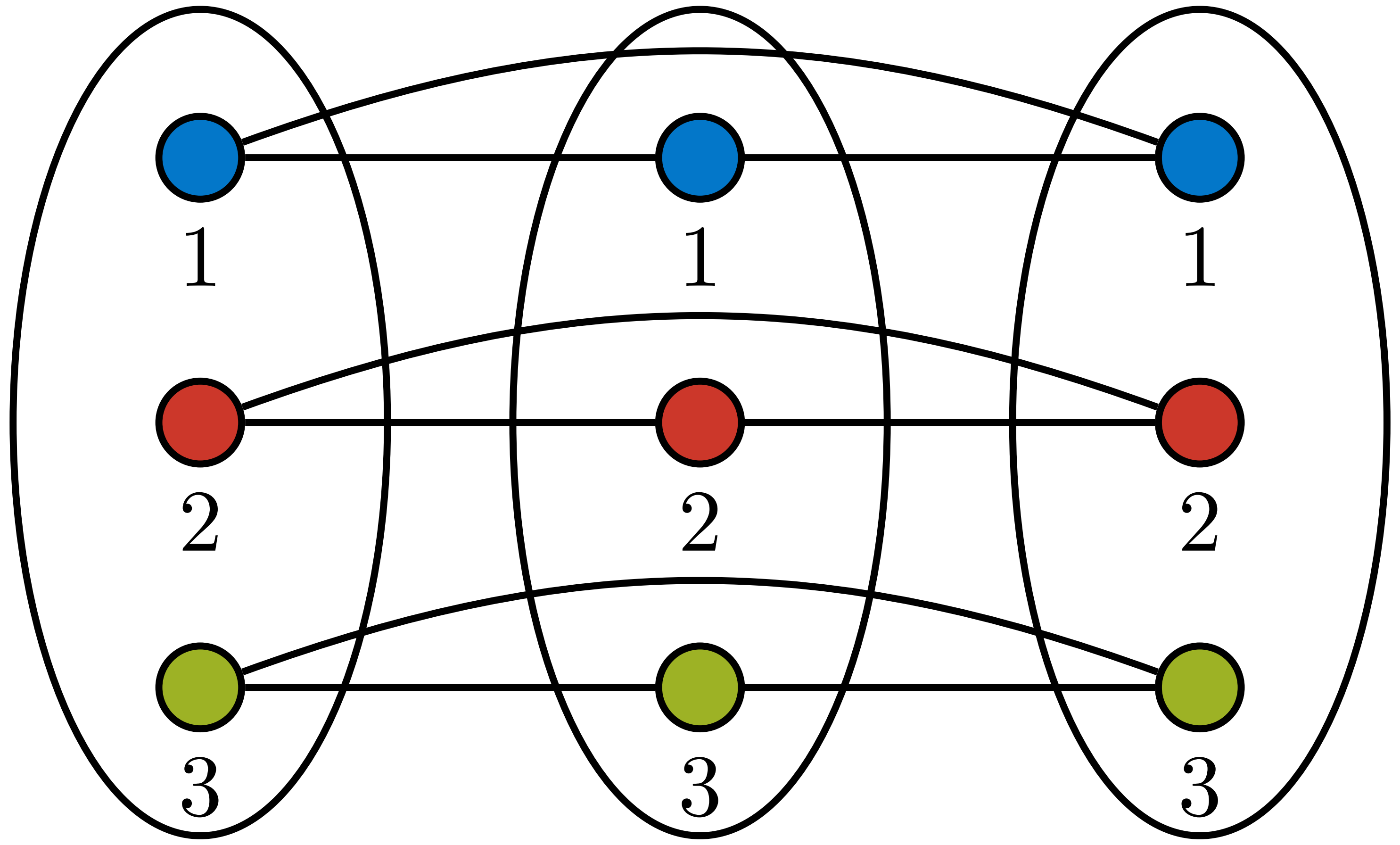
Divide the entire graph into overlapped subgraphs
Optimizing the permutation matrices for each subgraph with consensus constraints

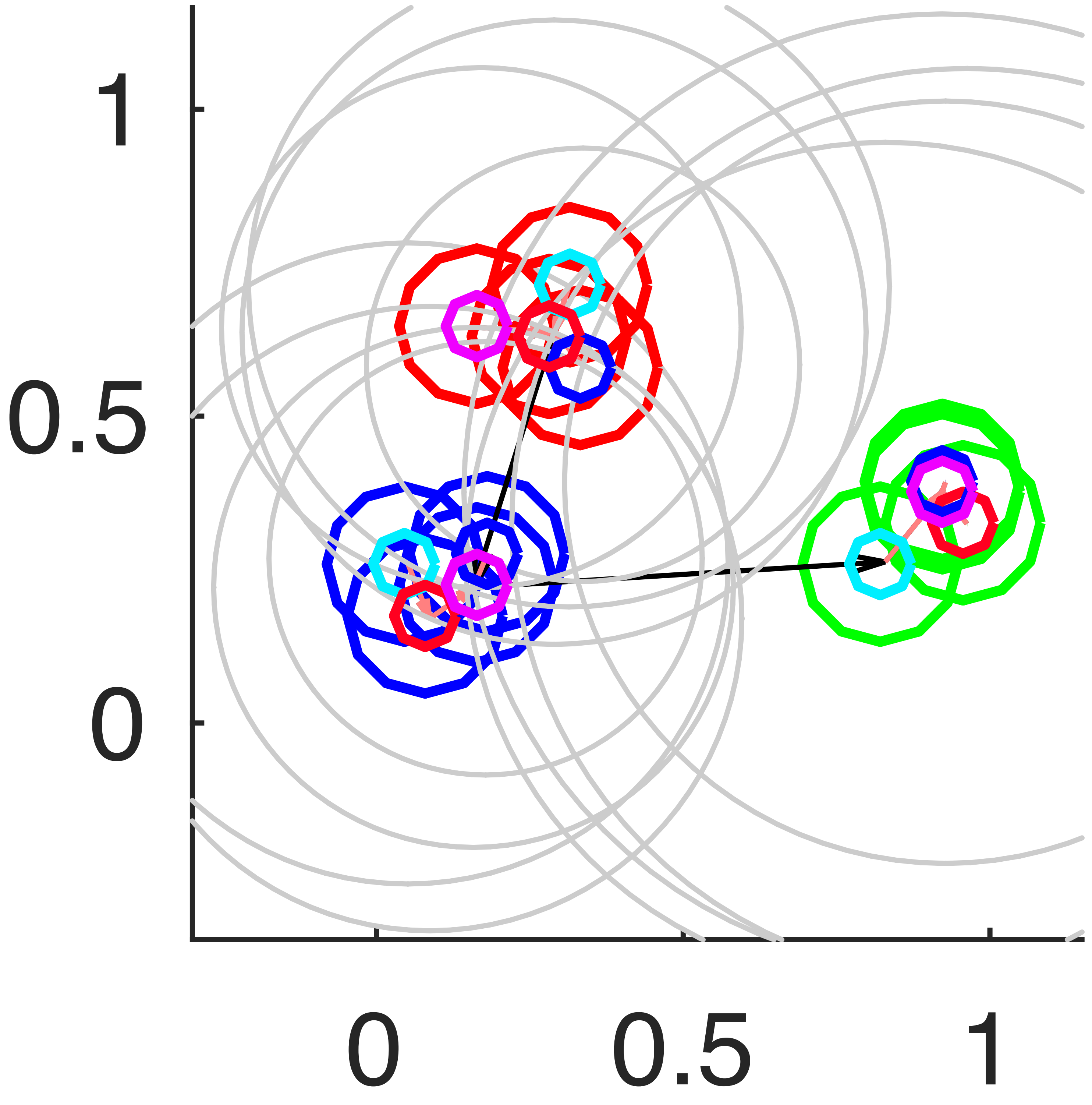
$$\begin{aligned} \min \quad & \textstyle\sum_i \left(\langle \mathbf{W}_{\mathcal{V}_i}, \mathbf{X}_{\mathcal{V}_i} \rangle + \lambda \|\mathbf{X}_{\mathcal{V}_i}\|_*\right) \\ \text{s.t.} \quad & \mathbf{X}_{\mathcal{V}_i} \in \mathcal{C}_i \\ & \mathbf{X}_{\mathcal{V}^i_{i\cap j}} = \mathbf{X}_{\mathcal{V}^j_{i\cap j}}, \forall (i,j) \in \mathcal{E} \end{aligned}$$



Hu, Huang, Thibert, Alpes, & Guibas (2018). Distributable Consistent Multi-Object Matching.

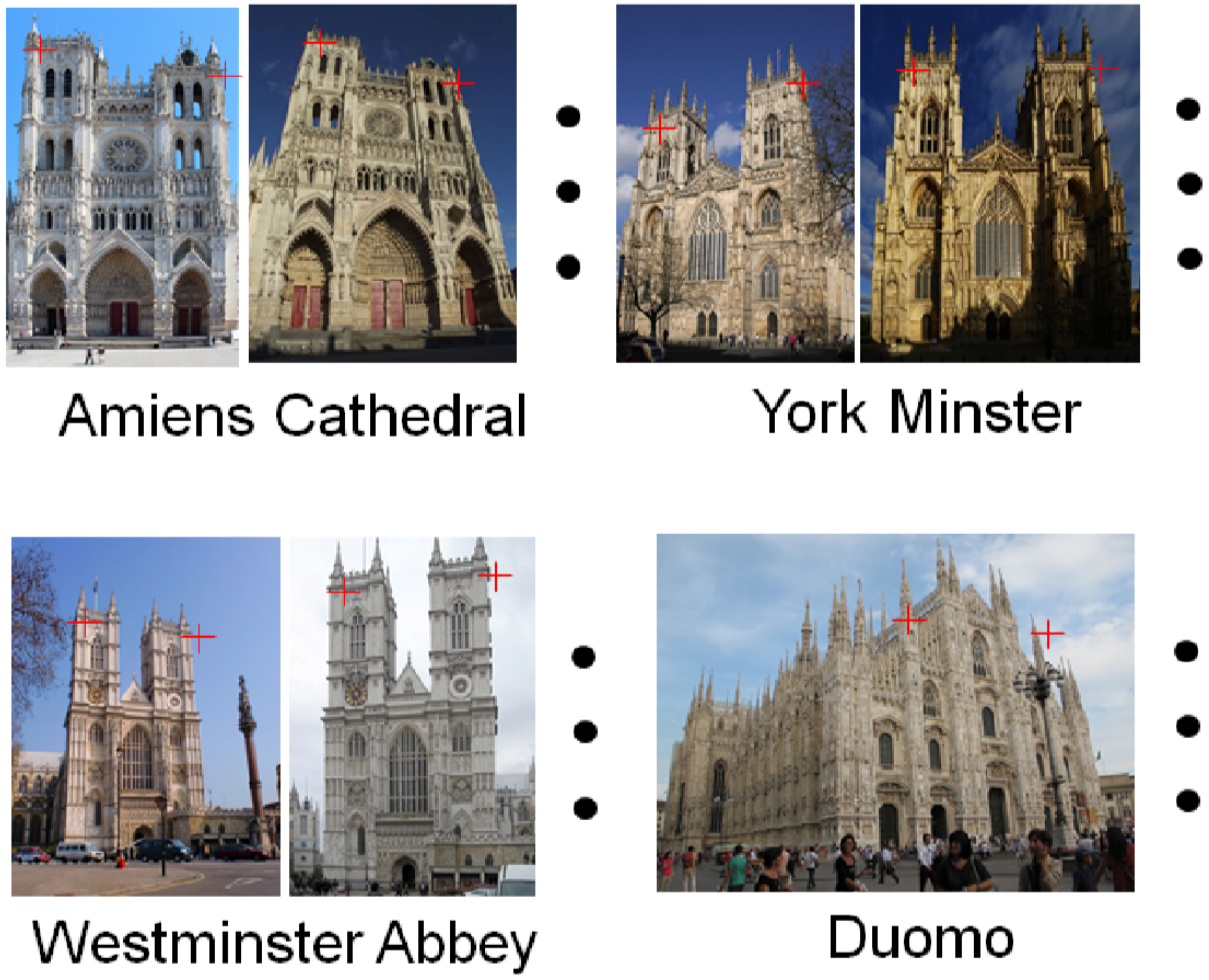
Multi-image matching as a clustering problem

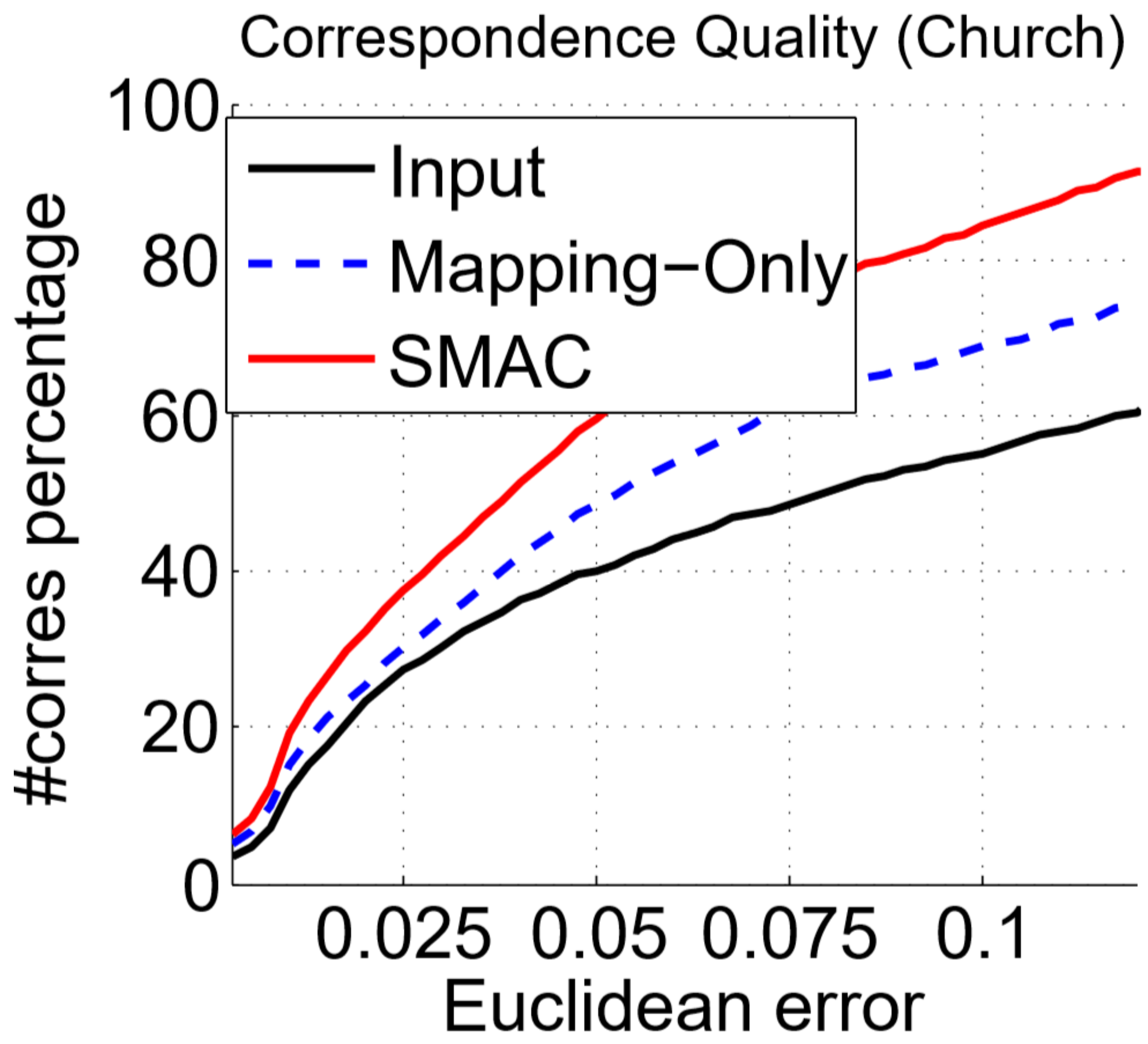




Tron, Zhou, Esteves, & Daniilidis (2017). Fast multi-image matching via density-based clustering.

Simultaneous mapping and clustering





Bajaj et al (2018). Simultaneous Mapping and Clustering via Spectral Decompositions.

Matching symmetric objects

Multiple plausible self-maps and pairwise maps

Sun et al (2018). Joint Map and Symmetry Synchronization.

Learning map synchronization

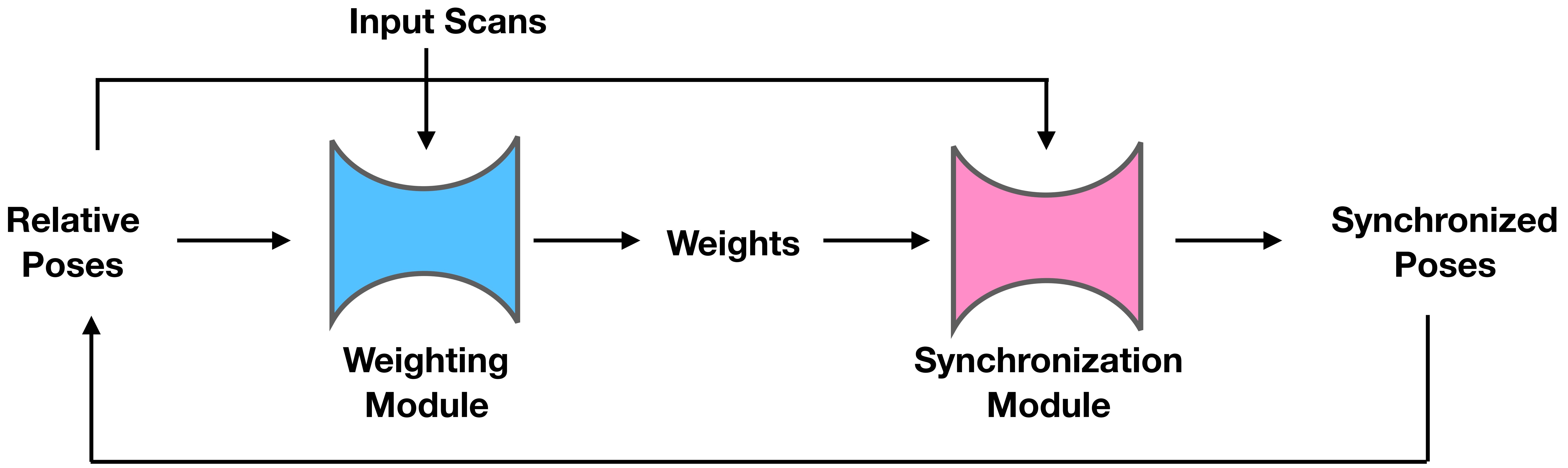
Recap: reweighed least squares for robust pose synchronization

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

Weights determined by previous guess and “hand-crafted” loss function

Can we make the weighting scheme learnable?

Learning map synchronization



<u>Iterative reweighted least squares with learned weights</u>

- Weighting module is a neural network
- Synchronization module is solving weighted spectral decomposition

Huang et al., CVPR 2019. Learning Transformation Synchronization.

Summary

Cycle consistency

Composition of maps along a cycle equals to identity

Synchronization for correspondences and relative poses

- inlier/outlier inference
- Local, iterative optimization
- Global, factorization-based optimization

Methods applied to more types of transformations

Open problems

- Scalability to large dataset
- Learning synchronization

Acknowledgements

Some slides borrowed from Qixing Huang (UT Austin)

Joint work with Spyridon Leonardos (UPenn), Roberto Tron (Boston Univ), Qianqian Wang (Cornell), Menglong Zhu (Google), Kostas Daniilidis (UPenn)

Slides and paper list will be available on the tutorial website

https://www.cs.utexas.edu/~huangqx/cvpr19_tutorial_map_sync.html