

CVPR 2020 Tutorial

Synchronization & Cycle Consistency

in Computer Vision

Tolga Birdal

in collaboration with
Federica Arrigoni, Qixing Huang, Leonidas Guibas





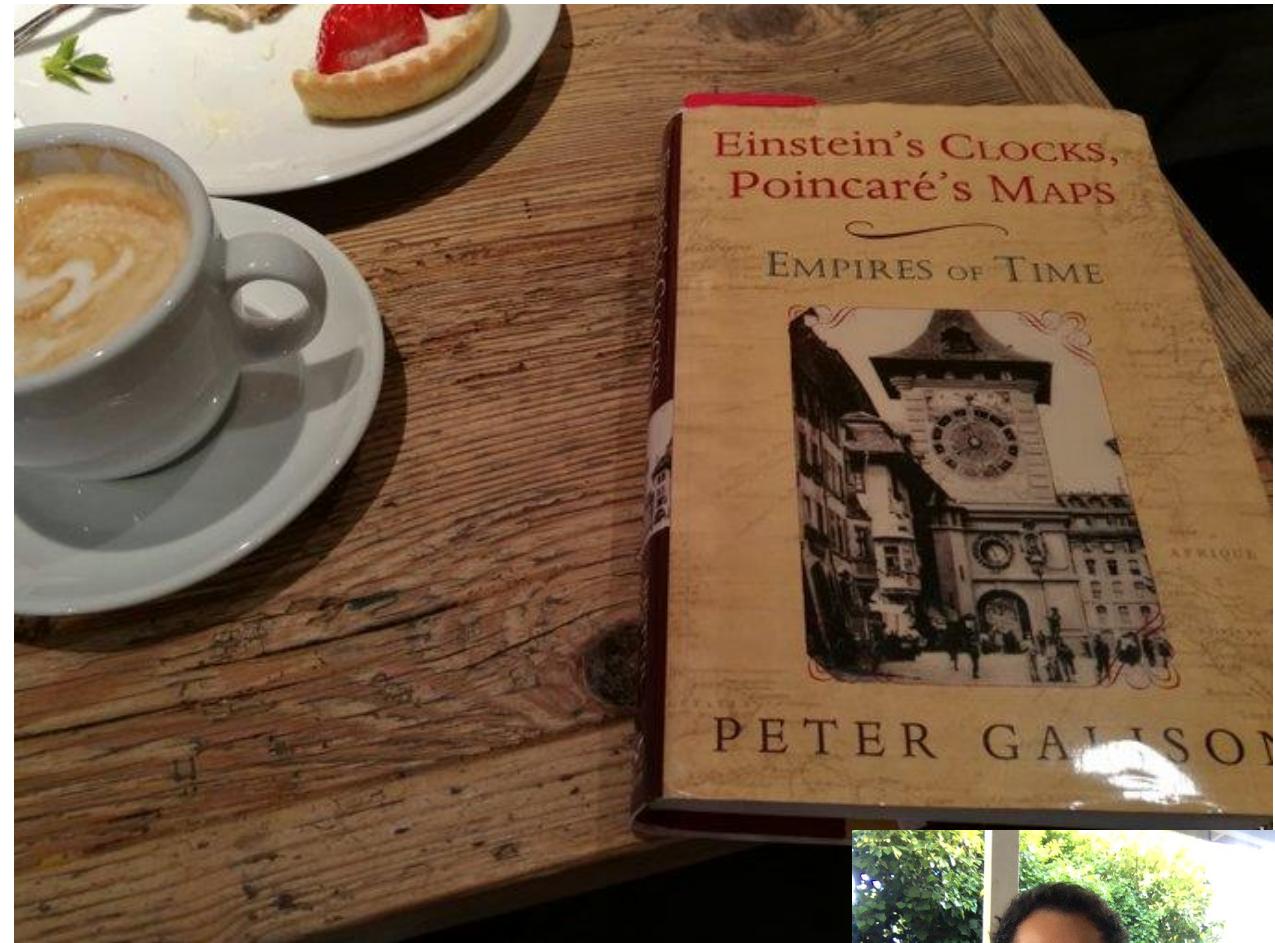
Synchronization

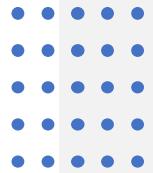
- The coordination of events to operate a system in unison
- Finding simultaneity



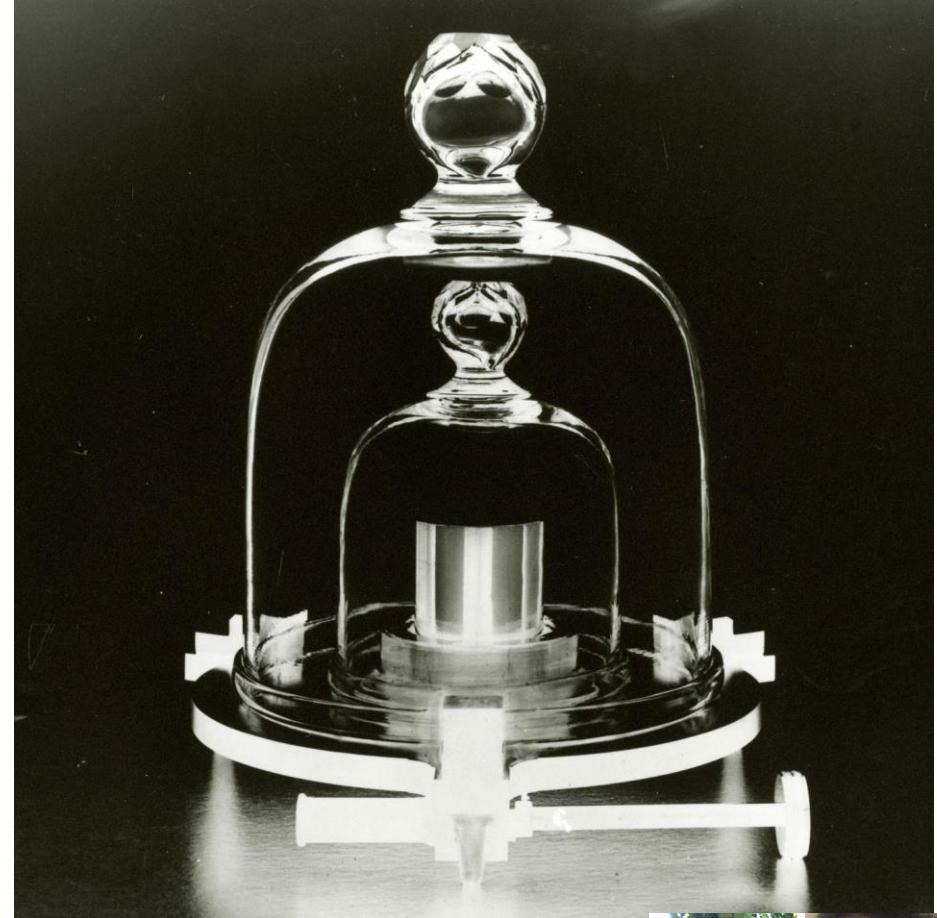
It all began with...

Clocks.





The Metric System

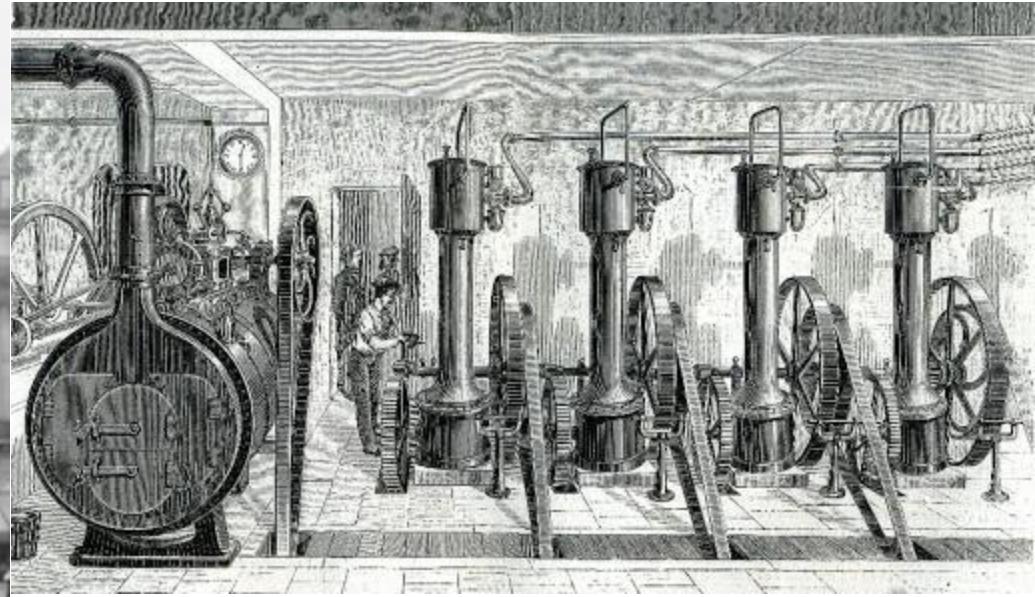


How do we meter-ize thi



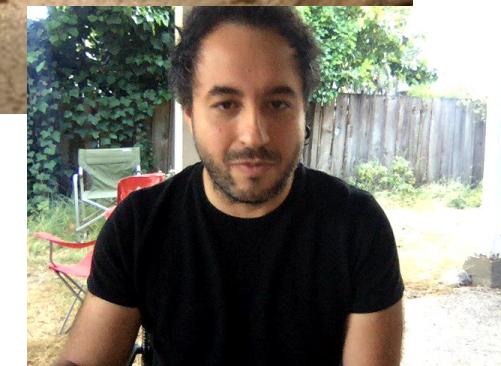
<https://qz.com/1458672/the-history-of-the-international-prototype-kilogram/>

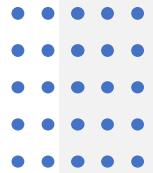
Parisian Pneumatic Clocks



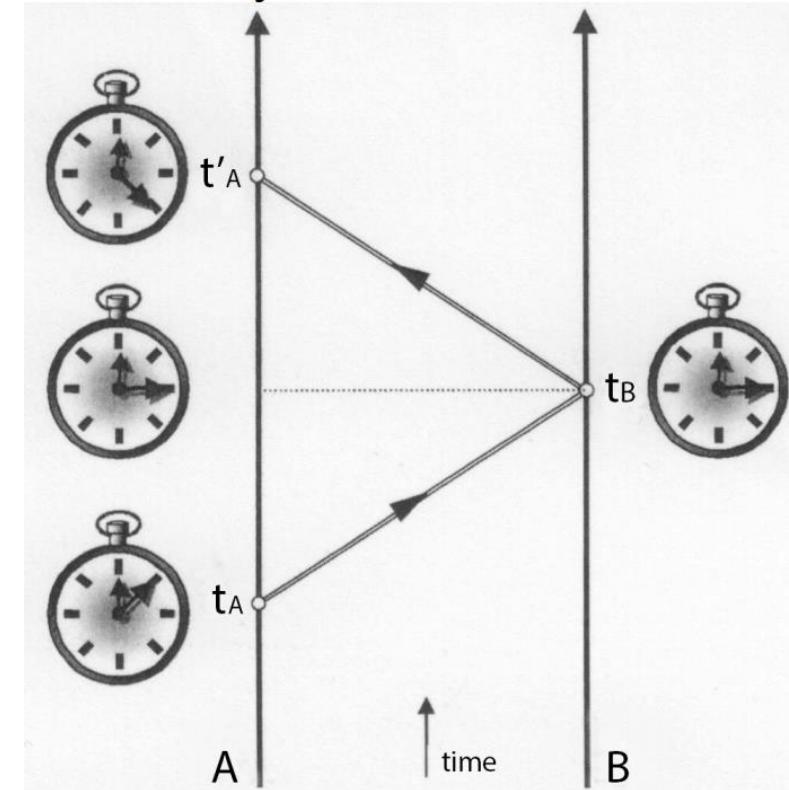
<http://www.douglas-self.com/MUSEU>

The town of Einstein,
Bern was no short of
clocks.

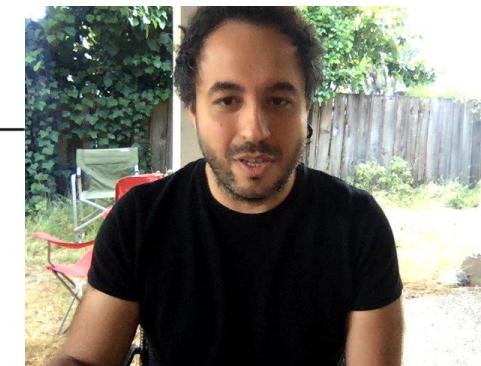




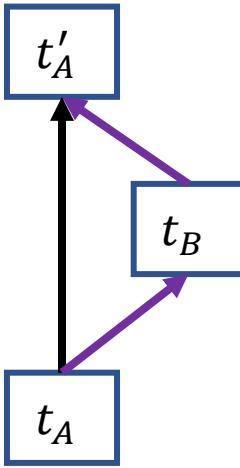
Einstein - Poincaré Synchronization



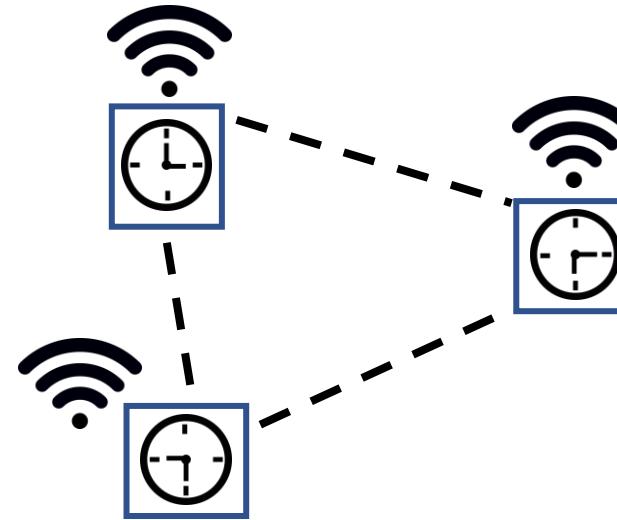
$$t_B = t_A + 1/2(t'_A -$$



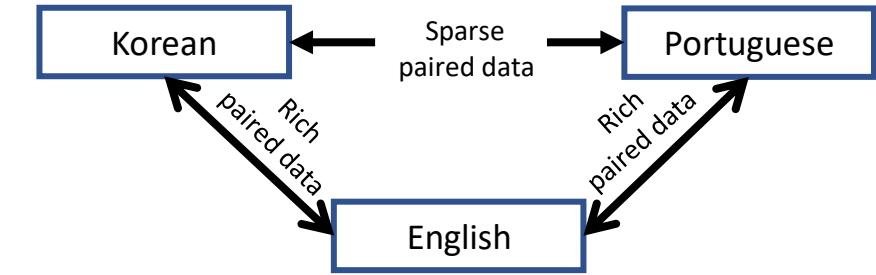
Synchronization & Cycle Consistency pre-Computer Vision



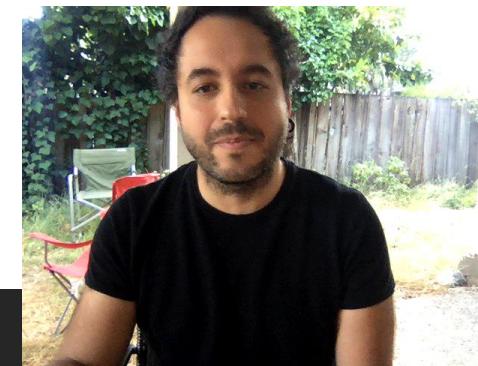
Clock Synchronization



Wireless Communication



Language Translation





Applications of Cycle Consistency in Computer Vision

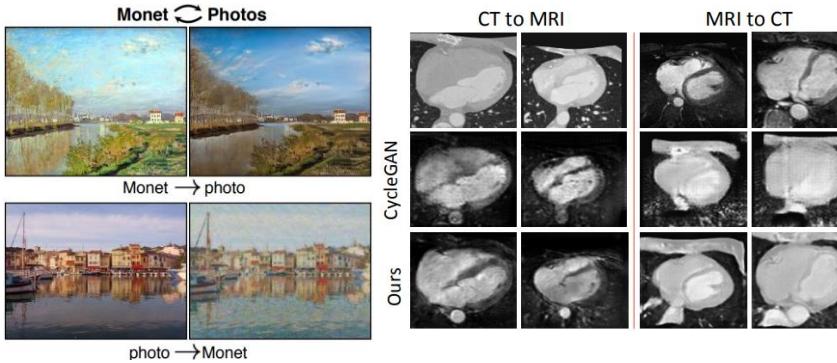
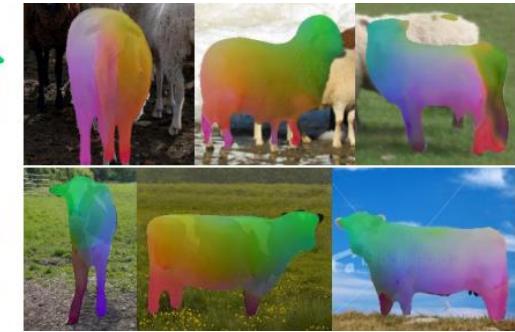


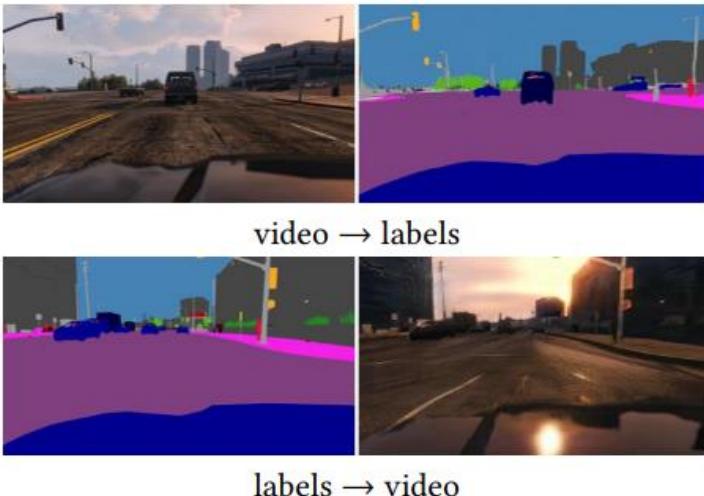
Image-to-Image Translation
Zu et al. ICCV'17



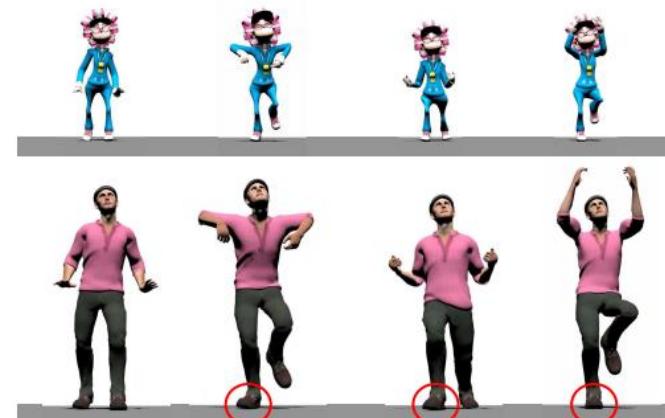
Video Understanding
Lai & Xie BMVC'19



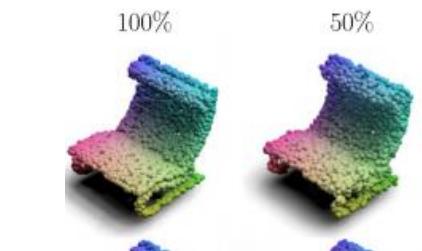
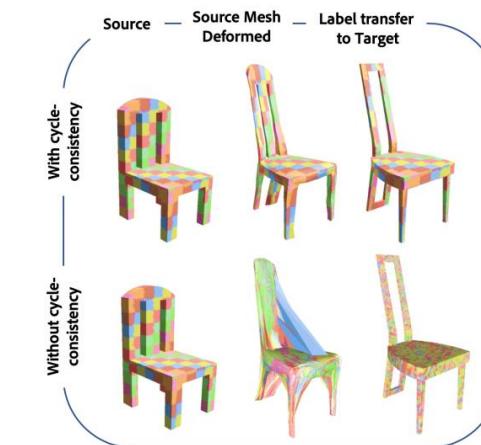
3D Geometric Registration
Kulkarni ICCV'19



Video-to-Video Translation
Chen et al. ACM MM'19

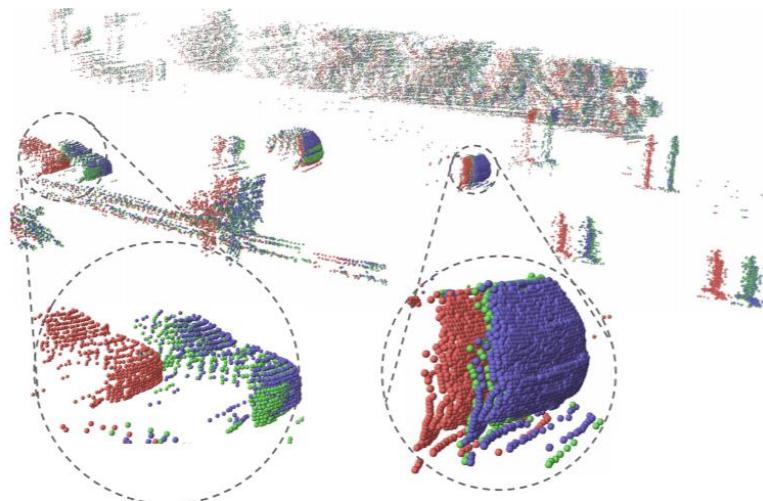


Motion Retargeting
Villegas et al. '19

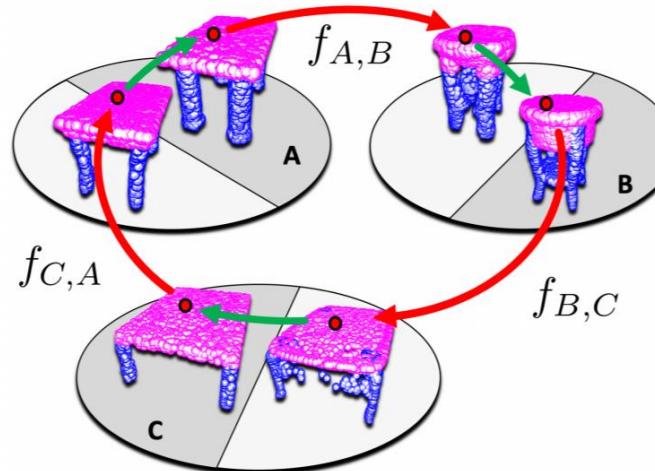




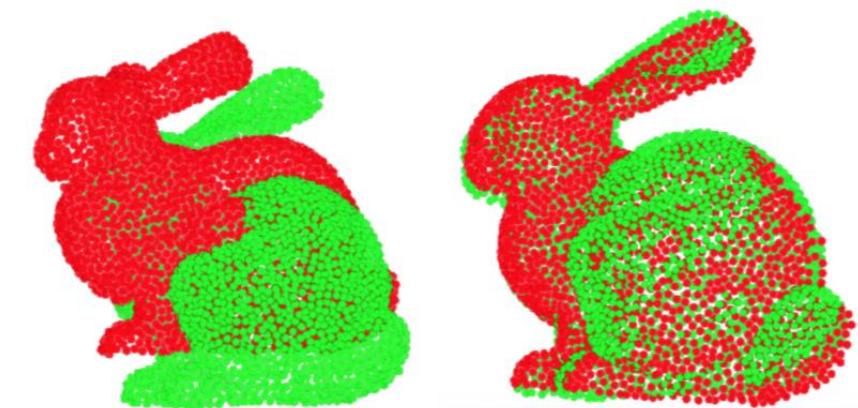
Applications of Cycle Consistency in Computer Vision



FlowNet3D
Liu et al. CVPR'19



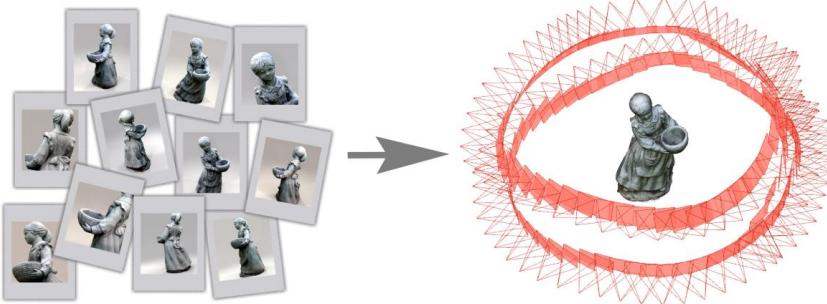
Grouiex ESGP'19



PRNet
Wang & Solomon NeurIPS'19



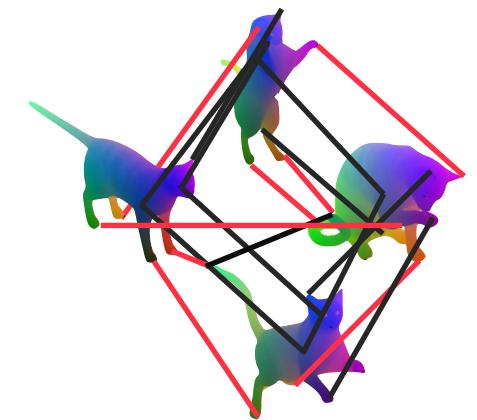
Applications of Synchronization in Computer Vision



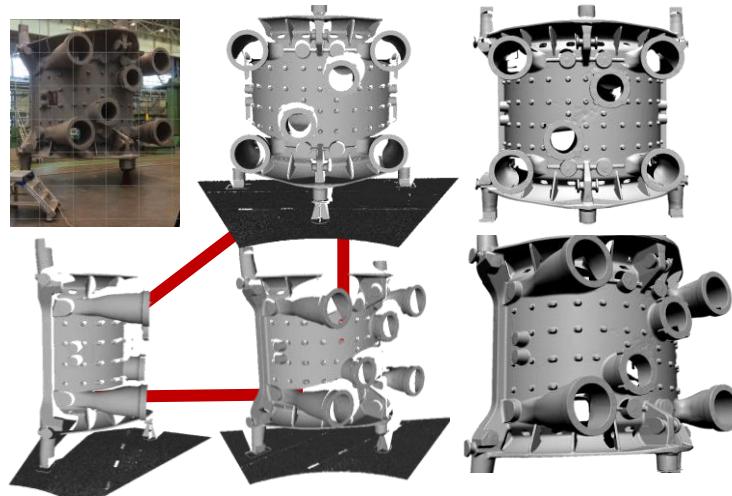
Structure from Motion
Bianco '18



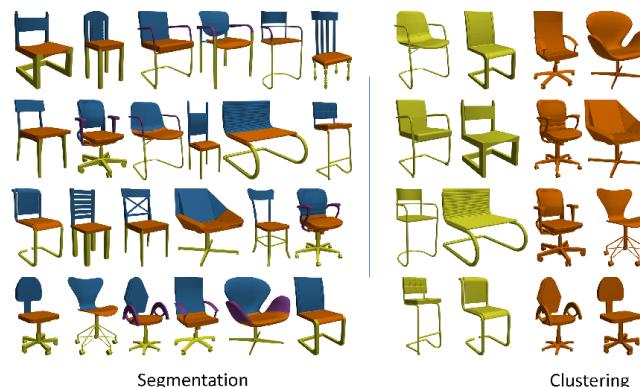
Feature Matching
Birdal et al. CVPR'19



3D Shape Matching
Birdal et al. CVPR'19



3D Reconstruction from Scans
Birdal et al. ICCV'17



Co-Segmentation / Clustering
Huang ACM'19



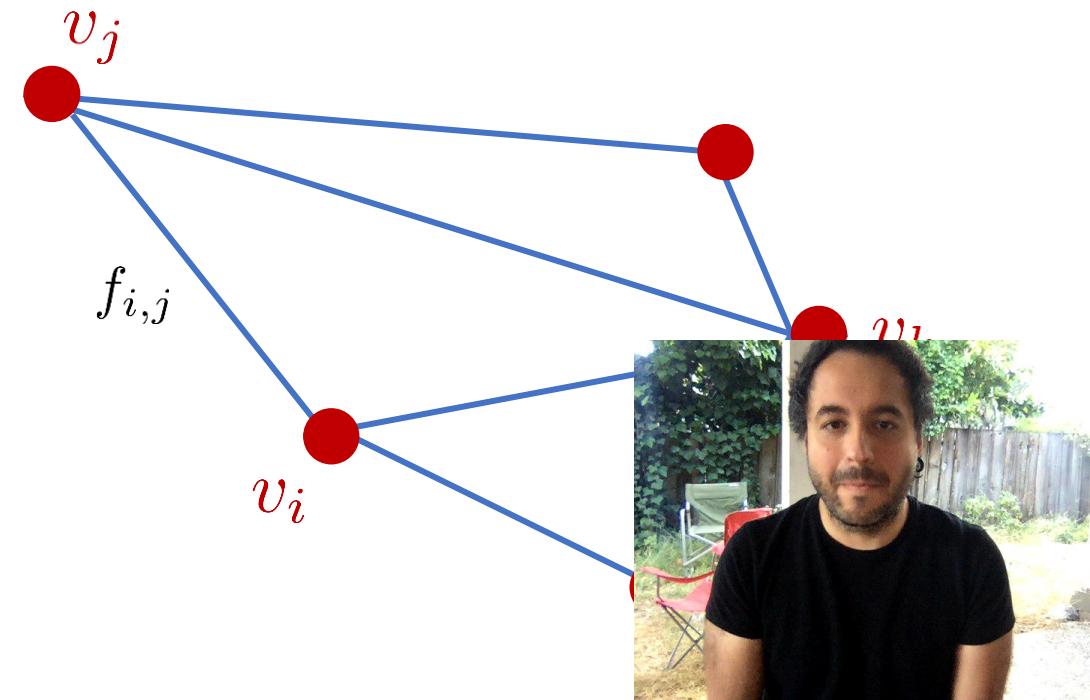
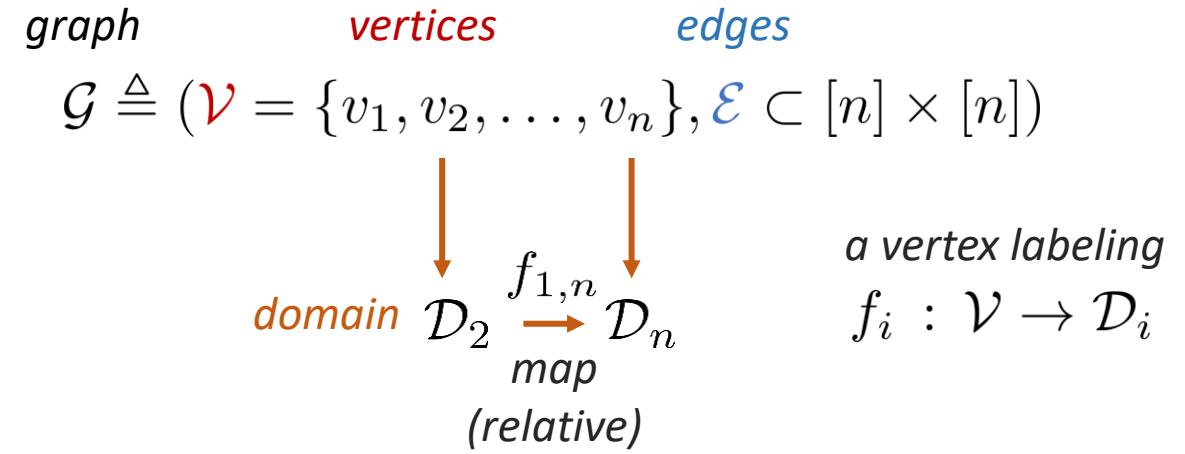
Simultaneous
Localization and
Mapping
Arras et al. CVPR'19



Notation

A connected graph

Domain, map, labeling



Path, Cycle, Null Cycle

Path:

$$p = \{(i_1, i_2), (i_2, i_3), \dots, (i_{n-1}, i_n)\} \in \mathcal{P}$$

Cycle:

$$c = \{(i_1, i_2), (i_2, i_3), \dots, (i_{n-1}, i_n), (i_n, i_1)\} \in \mathcal{C}$$

Null-cycle:

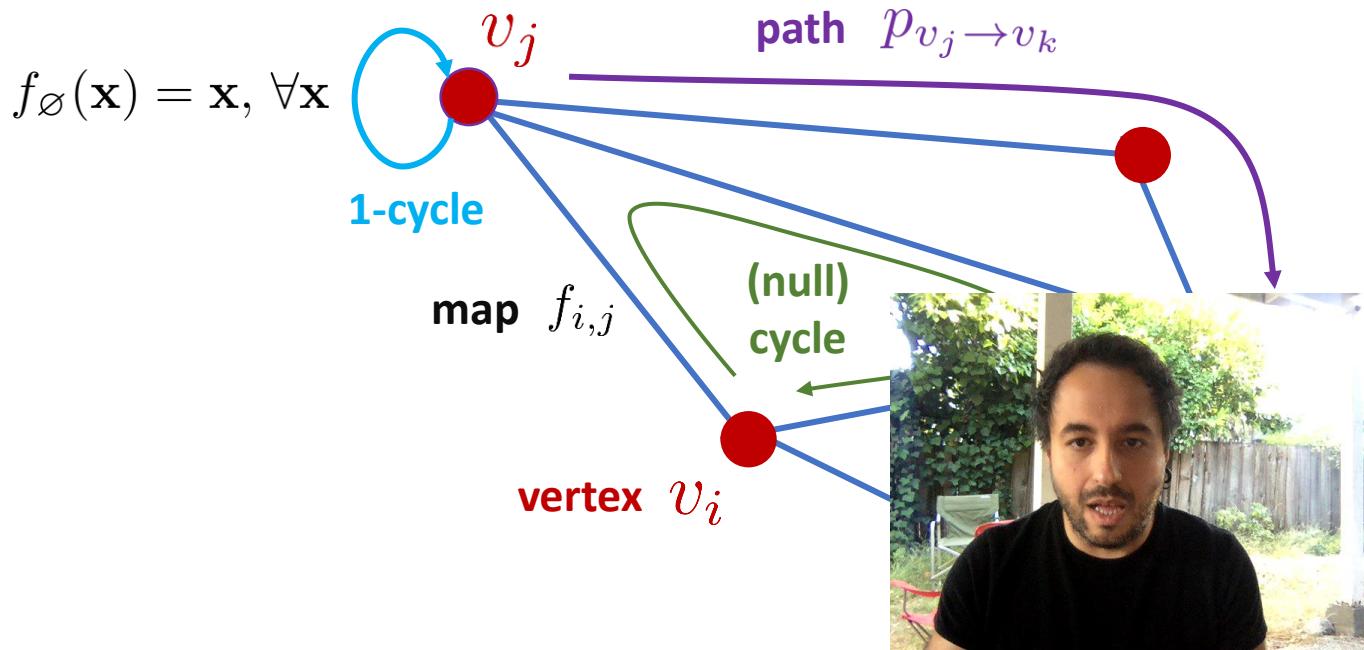
A cycle where the composition of functions lead to identity:

$$f_c = f_{1,2} \circ f_{2,3} \cdots \circ f_{(n-1),n} \circ f_{n,1} = f_\emptyset$$

k-cycle:

$$\text{1-cycle: } f_{i,i} = f_\emptyset$$

$$\text{2-cycle: } f_c = f_{i,j} \circ f_{j,i}$$



Path Invariance (PI) & Cycle Consistency (CC)

All possible paths between u and v :

All possible path pairs:

Graph is Path Invariant if:

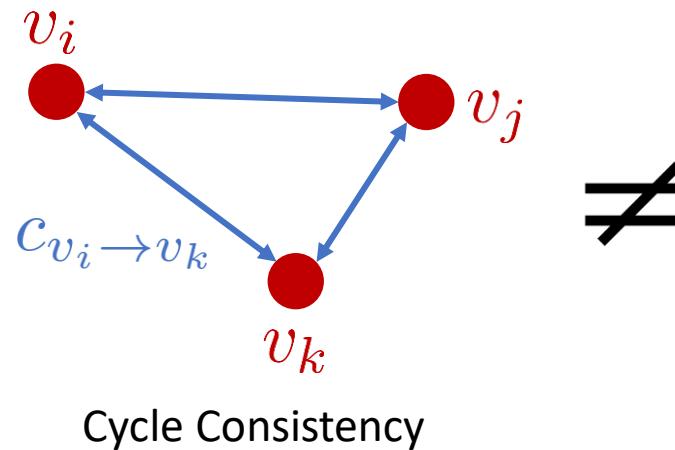
Graph is Cycle Consistent if:

$$\mathcal{G}_{\text{path}}(u, v)$$

$$\mathcal{G}_{\text{pairs}} = \bigcup_{(u,v) \in \mathcal{V}^2} \{(p, q) : p, q \in \mathcal{G}_{\text{path}}(u, v)\}$$

$$f_p = f_q \quad \forall (p, q) \in \mathcal{G}_{\text{pairs}}$$

$$f_c = f_\emptyset \quad \forall c \in \mathcal{C}$$



Common Assumptions on f

Cycle Consistency by Construction

1. $f_{i,j}$ is assumed to be an *isomorphism* i.e. $f_{j,i} = f_{i,j}^{-1}$.
2. $f_{i,j}$ are linear.
3. $f_{i,j}$ is *group valued* i.e. relative maps are *closed, invertible* and *associative* under the group operation \circ with the existence of an identity element. Note that this induces (1).



Consistent vertex labeling: $f_{i,j} = f_i \circ f_j^{-1} \quad \forall (i,j) \in \mathcal{E}$

Simple proof:

$$\begin{aligned}f_c &= f_{1,2} \circ f_{2,3} \circ \cdots \circ f_{(n-1),n} \circ f_{n,1} \\&= (f_1 \circ f_2^{-1}) \circ (f_2 \circ f_3^{-1}) \circ \cdots \circ \\&= f_1 \circ (f_2^{-1} \circ f_2) \circ (f_3^{-1} \circ f_3) \circ \cdots \\&= f_1 \circ f_1^{-1} = f_{\text{null}} \quad \forall c \in \mathcal{C}.\end{aligned}$$





Gauge Freedom

“Conventionality of Simultaneity”

Without loss of generality the following holds for group-valued f :

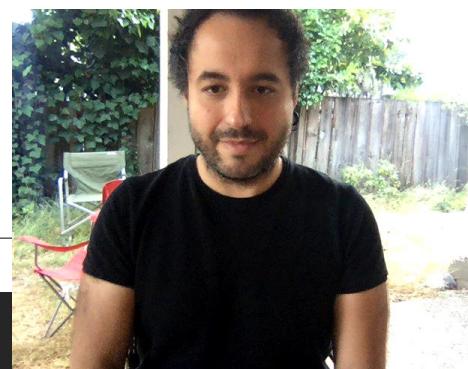
$$(f_i \circ f_g) \circ (f_j \circ f_g)^{-1} = (f_i \circ f_g) \circ (f_g^{-1} \circ f_j) = f_i \circ f_j^{-1}.$$

“That light requires the same time to traverse the path A → M as for the path B → M is in reality neither a supposition nor a hypothesis about the physical nature of light, but a stipulation which I can make of my own free will in order to arrive at a definition of simultaneity ” – Albert Einstein

(“Ueber die spezielle und die allgemeine Relativitaetstheorie”, Vieweg, Braunschweig (1988))

Can fix one of the functions arbitrarily to reduce the DoF.

Anderson, R., I. Vetharaniam, and G. Stedman, 1998. “Conventionality of Synchronisation, Gauge Dependence and Test Theories of Relativity,” *Physics Reports*, 295: 93–



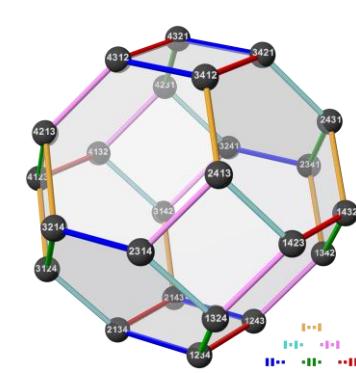


What are those maps f ?



Rotation Matrices

Zhao et al.'20



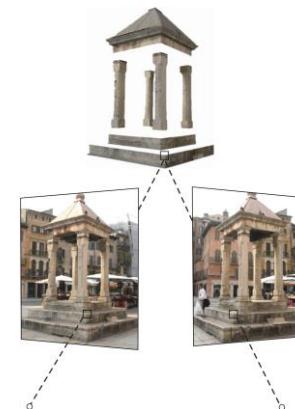
Permutations

wikipedia.org/wiki/Permutohedron



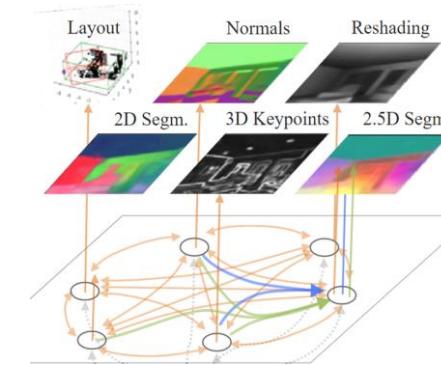
Functional Maps

Guibas '20



Essential Matrices

3dflow.net



Neural Networks

Zamir et al.'18



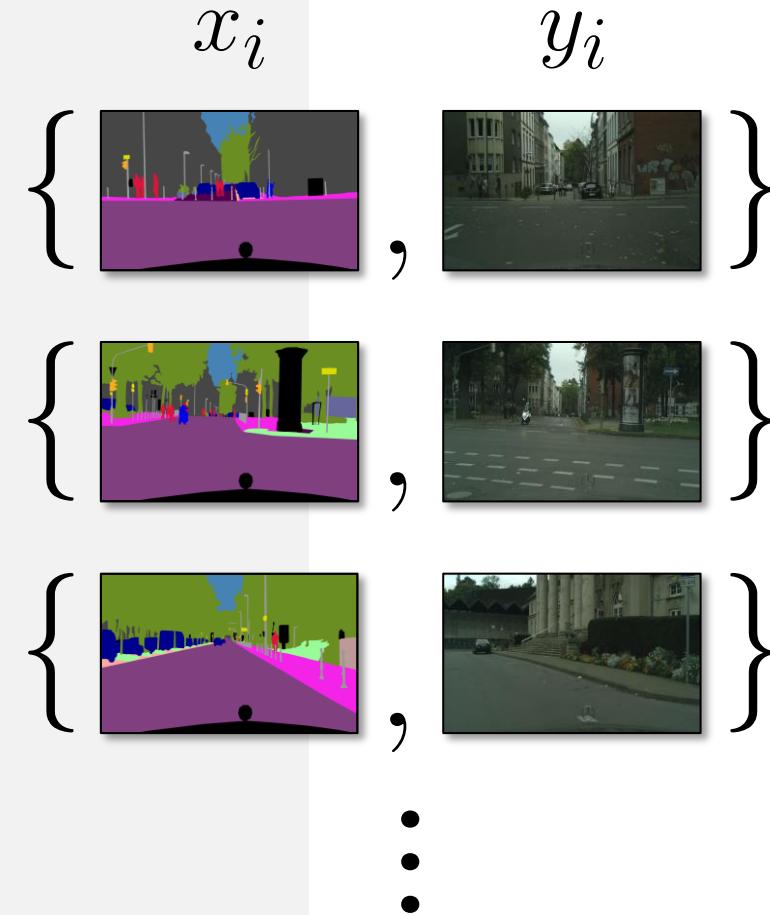
Tensor Maps

webstockreview.net

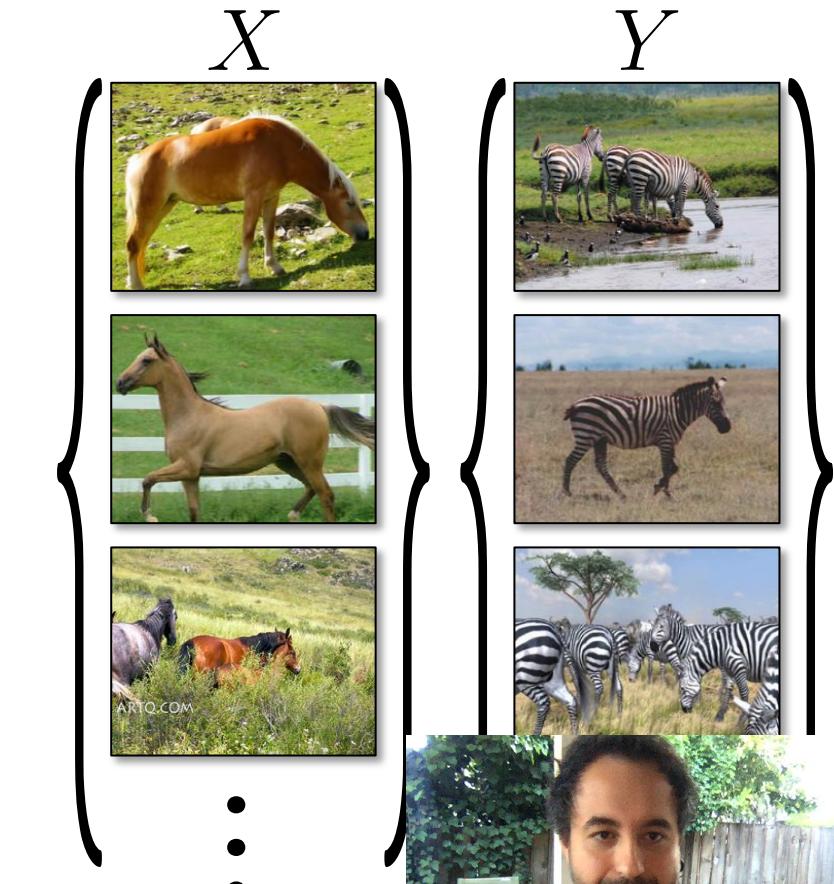


When to Use
Path Invariance
(PI) and Cycle
Consistency
(CC)

Supervised

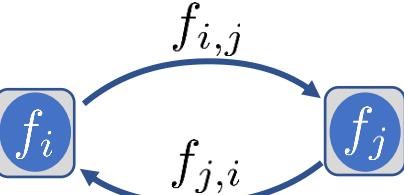
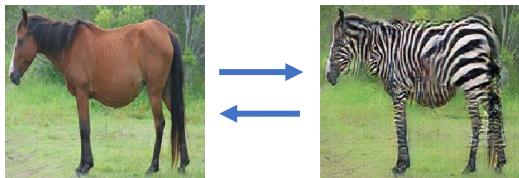


Unsupervised

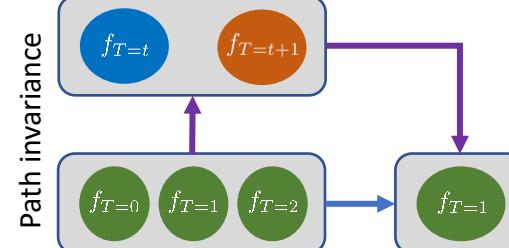
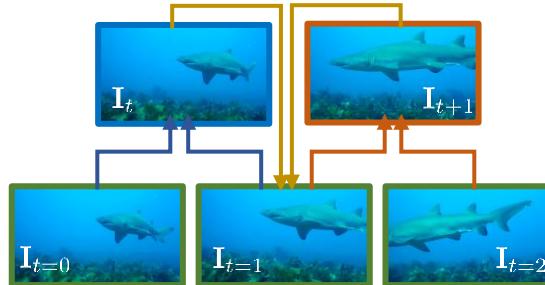


PI & CC as Unsupervised Losses

Cycle-GAN, ICCV'17

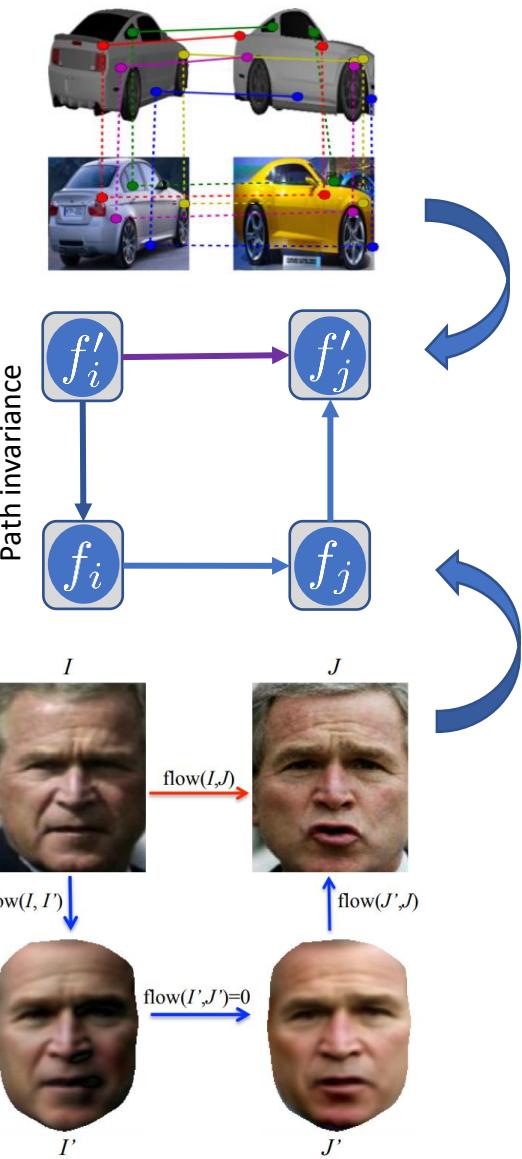


Vertex label is an image, relative map is a neural network.



Unsupervised Video Interpolation Using Cycle Consistency, Reda et al. ICCV 2019

Learning Dense Correspondence via 3D-guided Cycle Consistency, Zhou et al.



Collection Flow, Shlizerman et al.



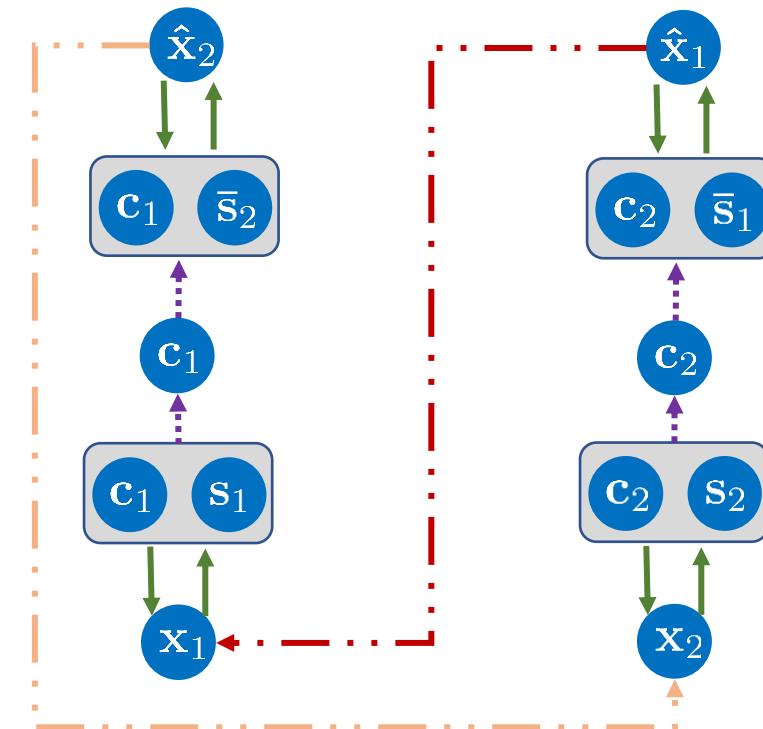
PI & CC as Unsupervised Losses



— · · · GAN (distributional)
— — cycle (learnt maps)

— · · · GAN (distributional)
· · · · path (fixed maps)

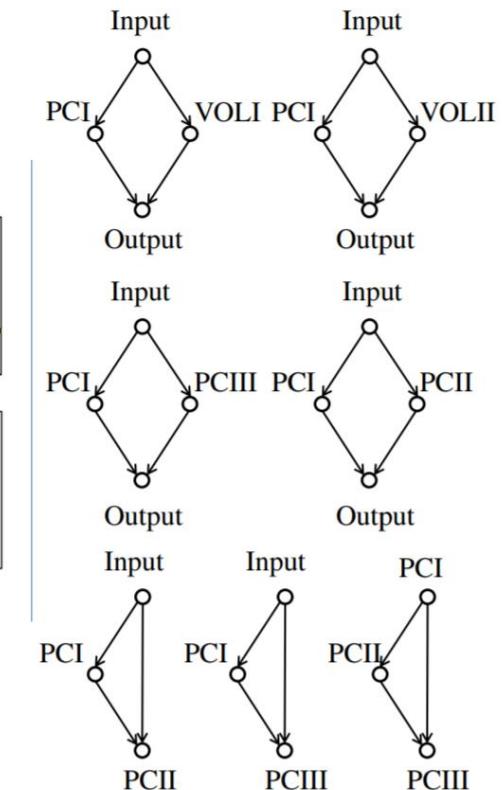
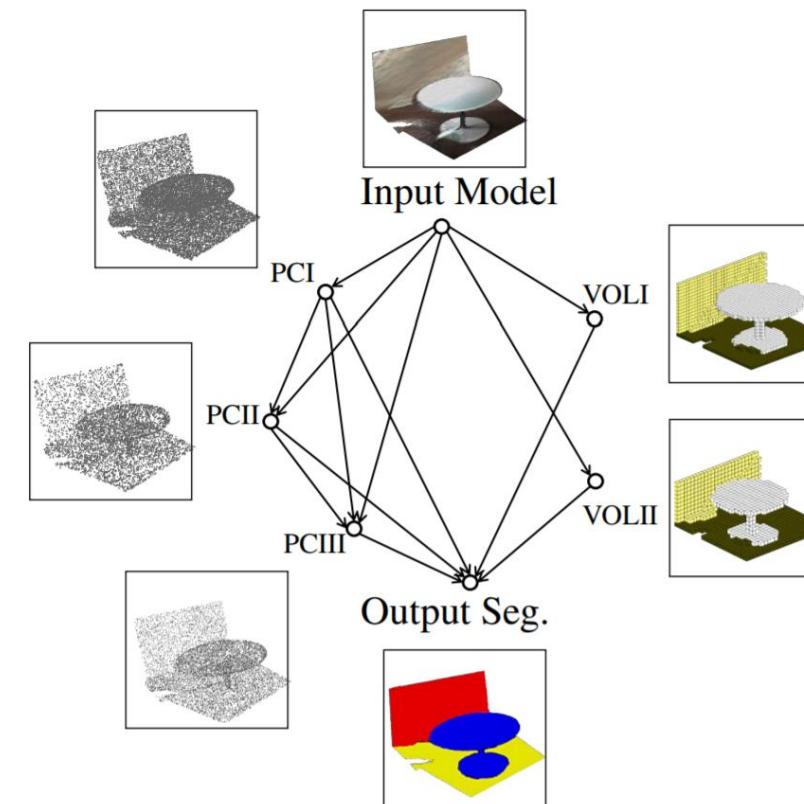
x: Image c: Content s: Style \bar{s} : Random Style



Huang, Xun, et al. "Multimodal unsupervised image-to-image translation." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.



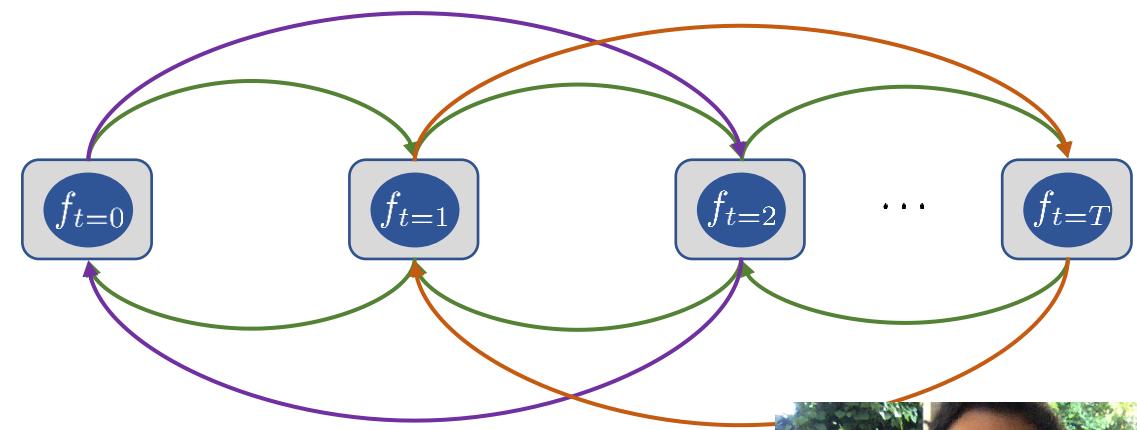
PI & CC as Unsupervised Losses



Zhang, Zaiwei, et al. "Path-invariant map networks."
Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.



PI & CC as Unsupervised Losses



Wang, Xiaolong et al. "Learning correspondence from the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition." 2018.





Synchronization

the art of handshaking

Given noisy relative maps $\{f_{i,j}\}$ as observations, recover the absolute maps $\{f_i\}$.

Given relative maps, find a consistent vertex labeling.





Synchronization

the art of handshaking

Naïve, over all the paths:

$$\arg \min_{\{f_k\}_k} \sum_{c \in \bar{\mathcal{C}}} \epsilon(f_c, f_\emptyset) + \sum_{(i,j) \in \mathcal{E}} d_f(f_i \circ f_{i,j}, f_j)$$

All cycles are null.

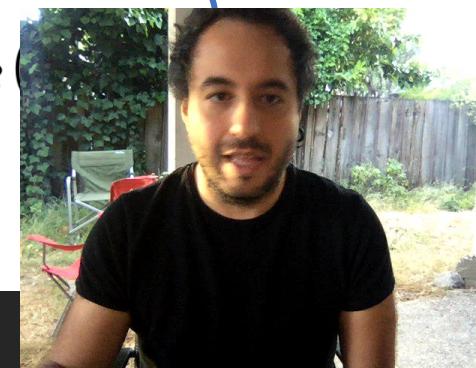
Group valued version:

$$\arg \min_{\{f_k\}_k} \sum_{(i,j) \in \mathcal{E}} \rho(d(f_i \circ f_j^{-1}, f_{i,j}))$$

Path Invariance

Sync via path invariance:

$$\arg \min_{\{f_k\}_k} \sum_{(i,j) \in \mathcal{E}} d_f(f_i \circ f_{i,j}, f_j) + \sum_{(p,q) \in \mathcal{B}} E\epsilon_B($$





Synchronization: Algorithms

The case of invertible linear maps.

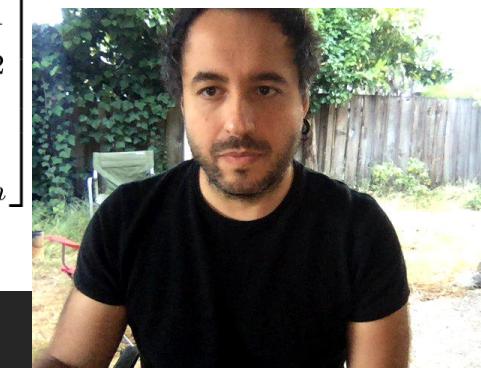
$$\arg \min_{\{f_k\}_k} \sum_{(i,j) \in \mathcal{E}} \rho(d(f_i \circ f_j^{-1}, f_{i,j})) \quad \xrightarrow{\hspace{1cm}} \quad \begin{array}{l} \text{Similar to a Dirichlet energy when } \rho(\mathbf{x}) = \mathbf{x} \\ \text{i.e. we seek to find a \textbf{harmonic}.} \end{array}$$

Matrix point of view for linear maps:

- $\mathbf{F}_{i,j} \in G$ represents $f_{i,j}$
- \mathbf{I} is the identity

$$\xrightarrow{\hspace{1cm}} \mathbf{G} = \begin{bmatrix} \mathbf{0} & \mathbf{F}_{1,2} & \mathbf{F}_{1,3} & \cdots & \mathbf{F}_{1,n} \\ \mathbf{F}_{2,1} & \mathbf{0} & \mathbf{F}_{2,3} & \cdots & \mathbf{F}_{2,n} \\ \cdots & \cdots & \vdots & \cdots & \cdots \\ \mathbf{F}_{n,1} & \mathbf{F}_{n,2} & \mathbf{F}_{n,3} & \cdots & \mathbf{F}_{n,n} \end{bmatrix} \succcurlyeq 0 \quad \xrightarrow{\hspace{1cm}} \quad \mathbf{Z} = \mathbf{I} + \mathbf{G}$$

$$\mathbf{Z} = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \\ \vdots \\ \mathbf{F}_n \end{bmatrix} \quad \begin{array}{c} \text{A man's face} \\ \text{represented by a vector} \end{array}$$

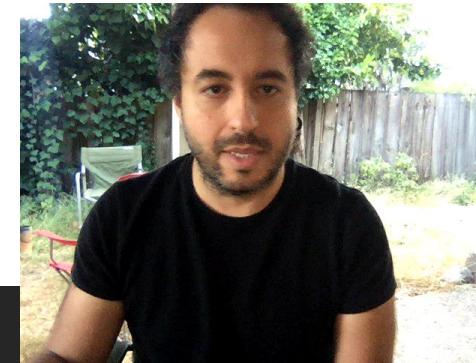




Synchronization: Algorithms

The case of invertible linear maps.

$$\arg \min_{\{f_k\}_k} \sum_{(i,j) \in \mathcal{E}} \rho(d(f_i \circ f_j^{-1}, f_{i,j})) \quad \equiv \quad \arg \min_{\{\mathbf{F}_k \in \mathcal{M}\}_{k=1}^K} \sum_{(i,j) \in \mathcal{E}} w_{ij} \|\mathbf{F}_i \mathbf{F}_j^{-1} - \mathbf{F}_{ij}\|_F^2$$





Semidefinite Relaxation (SDP)

*The case of **orthogonal** linear maps.*

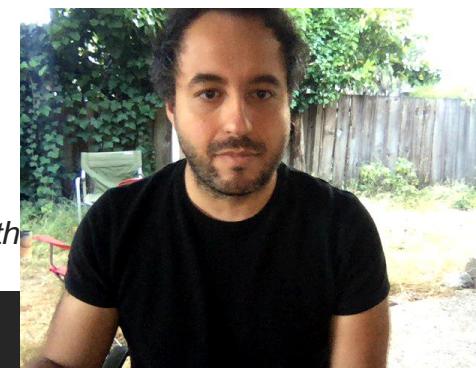
$$\arg \min_{\{\mathbf{F}_k \in \mathcal{M}\}_{k=1}^K} \sum_{(i,j) \in \mathcal{E}} w_{ij} \|\mathbf{F}_i \mathbf{F}_j^{-1} - \mathbf{F}_{ij}\|_F^2 \quad \equiv \quad \arg \max_{\{\mathbf{F}_k \in \mathcal{M}\}_{k=1}^K} \sum_{(i,j) \in \mathcal{E}} w_{ij} \text{Tr}(\mathbf{F}_i^\top \mathbf{F}_{i,j} \mathbf{F}_j)$$

which can be re-arranged into:

$$\max_{\mathbf{F}} \text{Tr}(\mathbf{F}^\top \mathbf{G} \mathbf{F})$$

semi-definite relaxation:

$$\max_{\hat{\mathbf{G}} \succcurlyeq 0, \mathbf{G}_{ii} = \mathbf{I}} \text{Tr}(\mathbf{G} \hat{\mathbf{G}}) \qquad \text{where} \qquad \hat{\mathbf{G}} = \mathbf{F}^\top \mathbf{F}$$





Spectral Solutions

$$\max_{\mathbf{F}} \text{Tr}(\mathbf{F}^\top \mathbf{G} \mathbf{F}) \text{ s.t. } \mathbf{F}^\top \mathbf{F} = \mathbf{I}$$

Rayleigh Problem

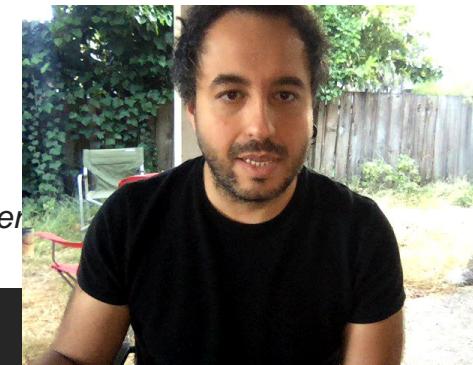
Solution is given by the eigenvectors of \mathbf{G} .

Both SDP and Spectral methods must project back on the **feasible set**.



1. Use constrained optimization
2. Use Riemannian optimization

Arrigoni, Federica, Beatrice Rossi, and Andrea Fusiello. "Spectral synchronization of multiple views in SE (3)." *SIAM Journal on Imaging Sciences*.





Optimization on Riemannian Manifolds

Riemannian Metric :

$\mathbf{G} = \langle \cdot, \cdot \rangle_{\mathcal{T}_X M} : \mathcal{T}_X M \times \mathcal{T}_X M \rightarrow \mathbb{R}$. e. g. In \mathbb{R}^2 , $G = dx^2 + dy^2$.

Exponential Map :

$\exp_X : \mathcal{T}_X M \rightarrow X$ maps a tangent vector to the endpoint of a geodesic path.

Logarithmic Map :

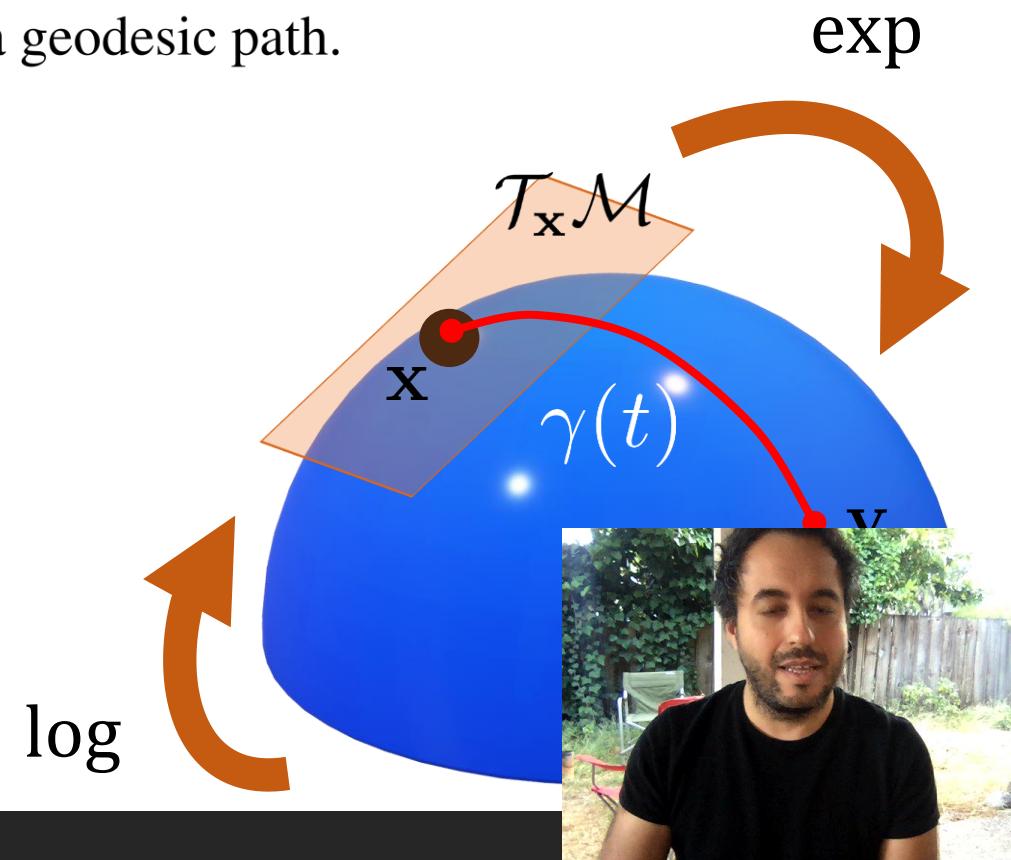
Inverse of the exponential map $\log_X : X \rightarrow \mathcal{T}_X M$

Isometry is a metric-preserving deformation.

A geodesic path γ is length minimizing and constant speed

$$d(x, y) = \min l_x^y(\gamma)$$

$$\text{On } \mathcal{M} : l_x^y = \int_x^y \sqrt{G(\dot{\gamma}(t), \dot{\gamma}(t))} dt$$





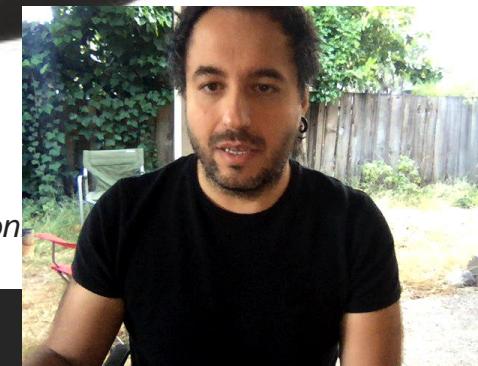
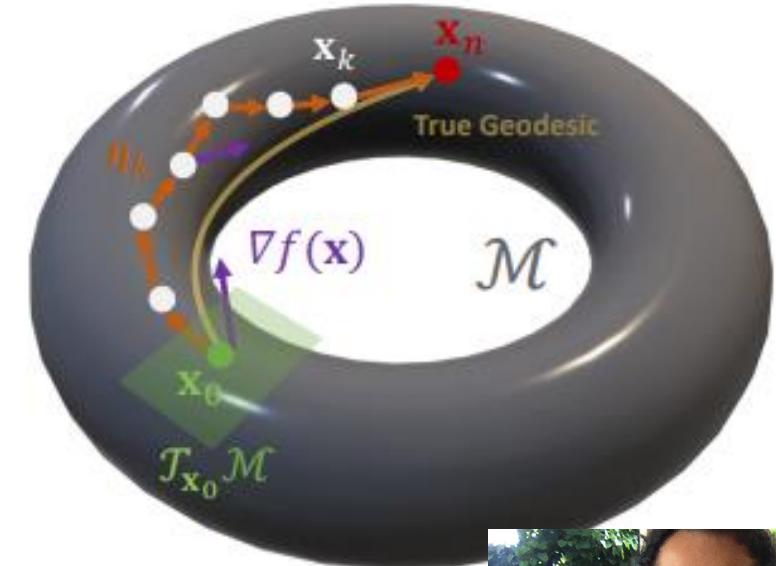
Riemannian Descent

Riemannian Manifolds + Group Structure = Lie Group

$$\arg \min_{\{\mathbf{F}_k \in \mathcal{M}\}_{k=1}^K} \sum_{(i,j) \in \mathcal{E}} w_{ij} \|\mathbf{F}_i \mathbf{F}_j^{-1} - \mathbf{F}_{ij}\|_F^2$$

Algorithm 1: General Riemannian Line Search Minimizer

- 1 **input:** A Riemannian manifold \mathcal{M} , a retraction operator R and initial iterate $\mathbf{x}_k \in \mathcal{M}$ where $k = 0$.
 - 2 **while** \mathbf{x}_k does not sufficiently minimize f **do**
 - 3 Pick a gradient related descent direction $\eta_k \in \mathcal{T}_{\mathbf{x}_k} \mathcal{M}$.
 - 4 Choose a retraction $R_{\mathbf{x}_k} : \mathcal{T}_{\mathbf{x}_k} \mathcal{M} \rightarrow \mathcal{M}$.
 - 5 Choose a step length $\tau_k \in \mathbb{R}$.
 - 6 Set $\mathbf{x}_{k+1} \leftarrow R_{\mathbf{x}_k}(\tau_k \eta_k)$.
 - 7 $k \leftarrow k + 1$.
-



Absil, P-A., Robert Mahony, and Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

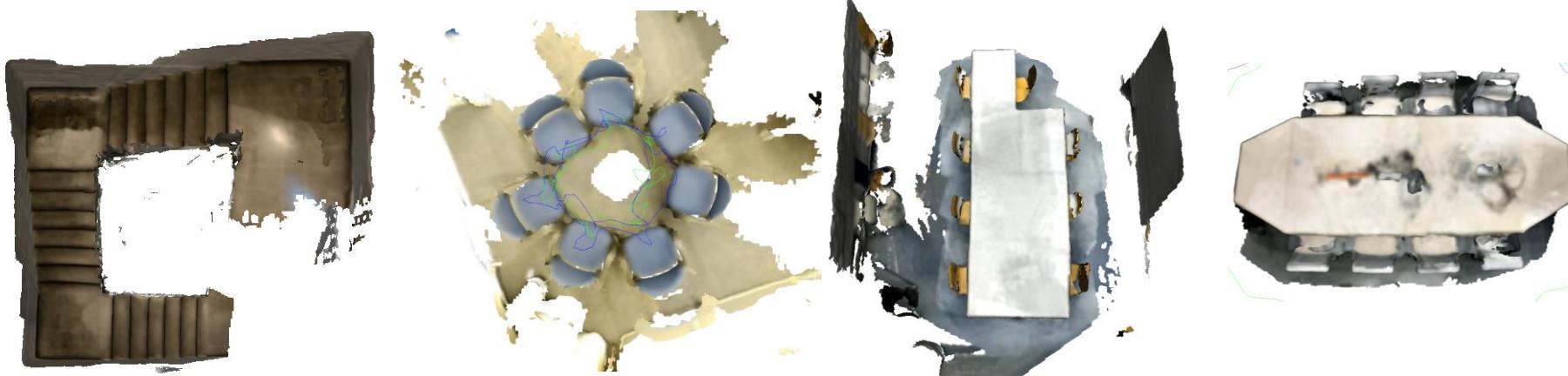
Birdal, Tolga, et al. "Probabilistic Permutation Synchronization using the Riemannian Structure of the Birkhoff Polytope" *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.



Real life is challenging.



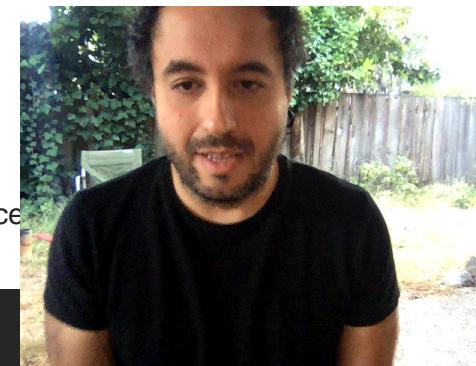
Ambiguous Views

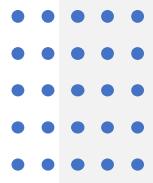


Uncertainty

Ambiguities

Bui, Mai, Tolga Birdal, Haowen Deng, Shadi Albarqouni, Leonidas Guibas, Slobodan Ilic, and Nassir Navab. "6D Camera Relocalization in Ambiguous Scenes via Multimodal Inference." *arXiv preprint arXiv:2004.04807* (2020).





Multiple Rotation or Motion Averaging





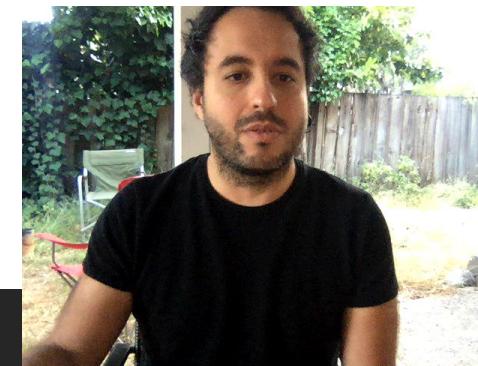
$$M = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \in SE(3)$$

Camera
poses form
non-Euclidean
parameter
spaces

$$R \in SO(3) \quad t \in \mathbb{R}^3$$

$$\mathbf{x}' = \mathbf{Rx} + \mathbf{t}$$

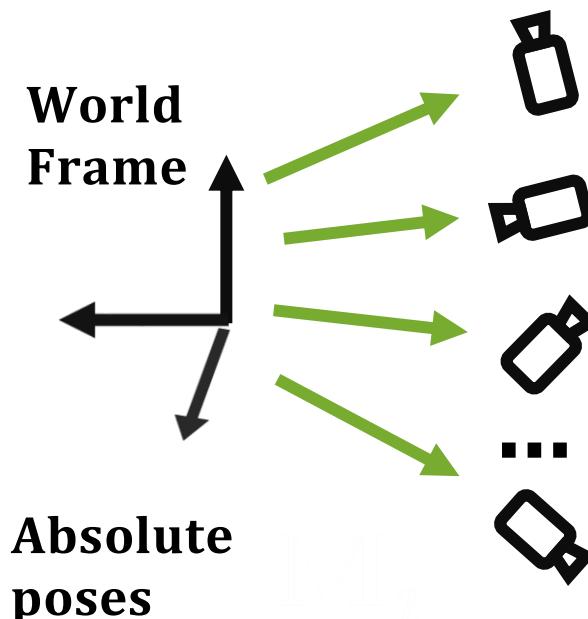
The manifold of rotations can be
parameterized in many ways.
We will use **quaternions**.



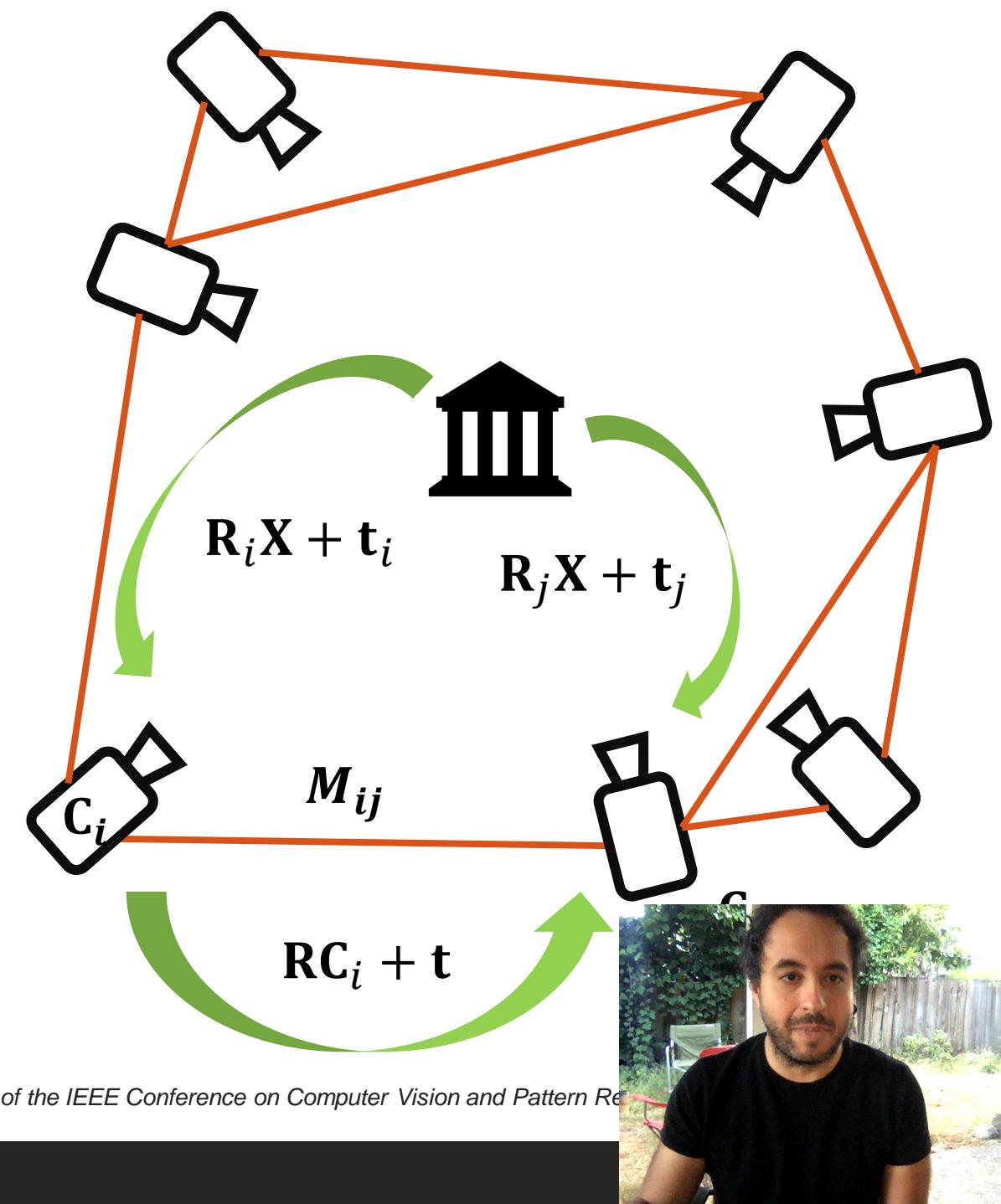
Multiple Motion Averaging

$$\mathbf{M}_{ij} \approx \mathbf{M}_j \mathbf{M}_i^{-1}, \forall i \neq j$$

(cycle consistency constraint)



+ estimate
uncertainties





Quaternions

- Introduced by Hamilton
- Extends complex numbers:
- Quaternions lie on a 4D unit sphere:
- Quaternions are antipodally symmetric
- Natural and good way to parameterize
- Rotate a point by quaternion:
- Quaternion distance:
- Easily relate to axis-angle:

$$\mathbf{q} = \cos\left(\frac{\theta}{2}\right) + (u_x \mathbf{i} + u_y \mathbf{j} + u_z \mathbf{k}) \sin\left(\frac{\theta}{2}\right)$$

$$\mathbf{q} = q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k}$$

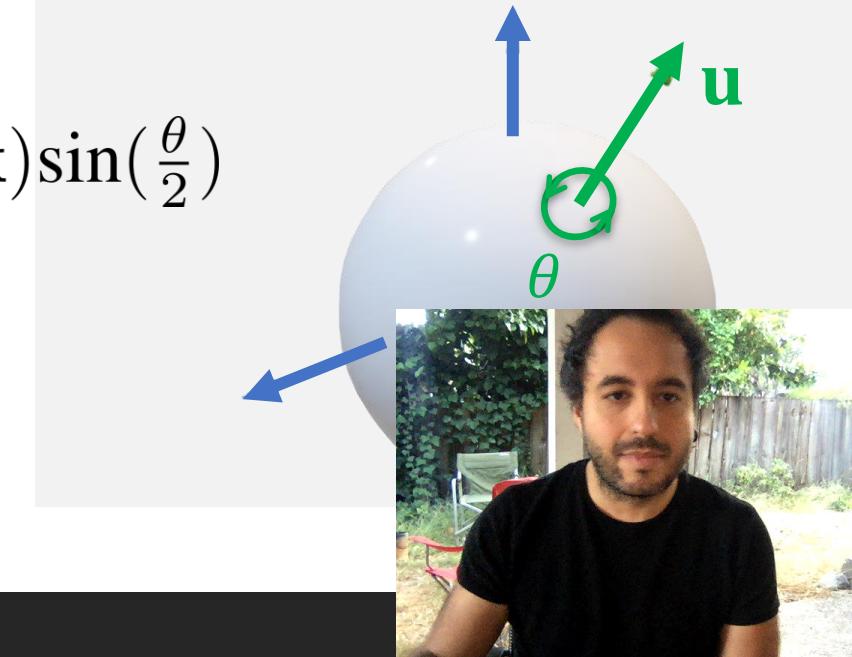
$$\|\mathbf{q}\| = 1$$

$$\mathbf{q} \equiv -\mathbf{q}$$

$$\mathbf{R} \leftrightarrow \mathbf{q}$$

$$\mathbf{p}' = \mathbf{q} \mathbf{p} \mathbf{q}^{-1}$$

$$d(\mathbf{q}_1, \mathbf{q}_2) = 2 \arccos(|\mathbf{q}_1 \bar{\mathbf{q}}_2|)$$



A Classical Algorithm

Govindu, Venu Madhav. "Combining two-view constraints for motion estimation." *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001.* Vol. 2. IEEE, 2001.

$$\mathbf{R}_{i,j} \approx \mathbf{R}_j \mathbf{R}_i^{-1}, \forall i \neq j$$

↓ Convert to quaternions.

$$\mathbf{q}_{i,j} \approx \mathbf{q}_j \otimes \mathbf{q}_i^{-1}, \forall i \neq j$$

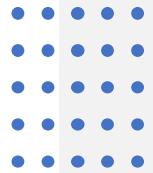
↓ Linearize using the adjoint.

$$\mathbf{Q}_{i,j} \mathbf{q}_i \approx \mathbf{q}_j, \forall i \neq j$$

↓ Write as a linear system.

$$[\cdots \mathbf{Q}_{i,j} \cdots \mathbf{I} \cdots] \Phi \mathbf{q}_j = \mathbf{q}_j$$





A Classical Algorithm

Govindu, Venu Madhav. "Combining two-view constraints for motion estimation." *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001.* Vol. 2. IEEE, 2001.

$$[\dots \mathbf{Q}_{i,j} \dots - \mathbf{I} \dots] \mathbf{q} = 0$$

$$\Phi \mathbf{q} = 0$$

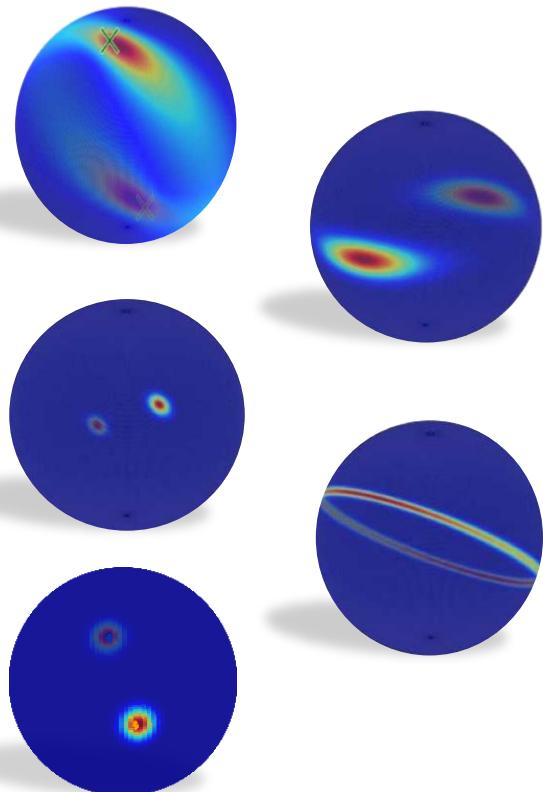
Solve!

Cannot enforce group constraints.





Bingham Distribution



$$\begin{aligned}\mathcal{B}(\mathbf{x}; \boldsymbol{\Lambda}, \mathbf{V}) &= (1/F) \exp(\mathbf{x}^T \mathbf{V} \boldsymbol{\Lambda} \mathbf{V}^T \mathbf{x}) \\ &= (1/F) \exp\left(\sum_{i=1}^d \lambda_i (\mathbf{v}_i^T \mathbf{x})^2\right)\end{aligned}$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0, & & & \\ & \lambda_1 & & \\ & & \lambda_2 & \\ & & & \ddots \\ & & & & \lambda_{d-1} \end{bmatrix}_{d \times d}$$

$(0 \geq \lambda_1 \geq \dots \geq \lambda_{d-1})$

$$F \triangleq |S_{d-1}| \cdot {}_1F_1\left(\frac{1}{2}, \frac{d}{2}, \boldsymbol{\Lambda}\right)$$

Bingham, Christopher. "An antipodally symmetric distribution on the sphere." *The Annals of Statistics* (1974): 1201-1225.

- $\mathbf{V} \in \mathbb{R}^{d \times d}$: an orthogonal matrix
- $\boldsymbol{\Lambda} \in \mathbb{R}^{d \times d}$: *concentration matrix*, a diagonal matrix with a set of non-positive values
- $F(\boldsymbol{\Lambda})$: the normalization constant
- A variant of zero-mean Gaussian distribution condition



- Tempered Geodesic MCMC
- Framework
- for Sampling and Optimization

Birdal, Tolga, et al. "Bayesian Pose Graph Optimization via Bingham Distributions and Tempered Geodesic MCMC" *Proceedings of the Neural Information Processing Systems*. 2018.





A Bayesian Formulation for Synchronization on $SE(3)$

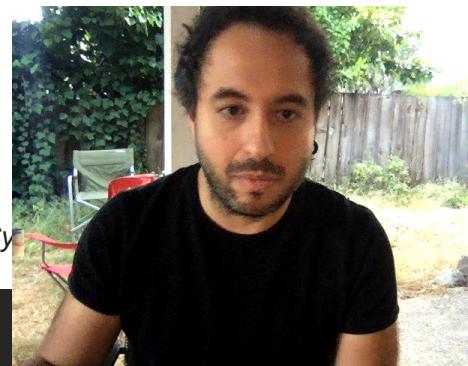
1. The maximum a-posteriori (MAP) estimate:

$$(\mathbf{Q}^*, \mathbf{T}^*) = \arg \max_{\mathbf{Q}, \mathbf{T}} p(\mathbf{Q}, \mathbf{T} | \mathcal{D}) =$$

$$\arg \max_{\mathbf{Q}, \mathbf{T}} \left(\sum_{(i,j) \in E} \{ \log p(\mathbf{q}_{ij} | \mathbf{Q}, \mathbf{T}) + \log p(\mathbf{t}_{ij} | \mathbf{Q}, \mathbf{T}) \} + \sum_i \log p(\mathbf{q}_i) + \sum_i \log p(\mathbf{t}_i) \right),$$


2. The full posterior distribution:

$$p(\mathbf{Q}, \mathbf{T} | \mathcal{D}) \propto p(\mathcal{D} | \mathbf{Q}, \mathbf{T}) \times p(\mathbf{Q}) \times p(\mathbf{T})$$





SG-MCMC

- Sampling from Bayesian posterior distributions defined on Riemann manifolds.
- No need of a global coordinate system.
- Define the Riemannian-Hamiltonian (negative log of the joint density):

$$H_\lambda(\varphi) \triangleq \underbrace{U_\lambda(\tilde{\mathbf{x}})}_{\text{potential energy}} + \log |\mathbf{G}(\tilde{\mathbf{x}})| + \underbrace{\mathbf{p}^\top \mathbf{G}(\tilde{\mathbf{x}})^{-1} \mathbf{p}}_{\text{kinetic energy}}$$

- Liu, Chang, Jun Zhu, and Yang Song. "Stochastic gradient geodesic MCMC methods." *Advances in Neural Information Processing Systems*. 2016.
- Girolami and Calderhead, *Riemannian manifold Hamiltonian Monte Carlo*, 2011)



A Modified Riemannian Hamiltonian Monte Carlo

$$\mathcal{E}_\lambda(\varphi) \triangleq \beta U_\lambda(\tilde{\mathbf{x}}) + \frac{\beta}{2} \log |\mathbf{G}(\tilde{\mathbf{x}})| + \frac{\beta}{2} \mathbf{p}^\top \mathbf{G}(\tilde{\mathbf{x}})^{-1} \mathbf{p}$$



Inverse temperature

Hamiltonian Paths converge to a measure with density:

$$\exp(-\beta U(\mathbf{x}))$$

β is the Optimization vs Sampling knob.

$\beta = 1$ we perform SG – MCMC

$\beta \rightarrow \infty$ we perform gradient descent with momentum.





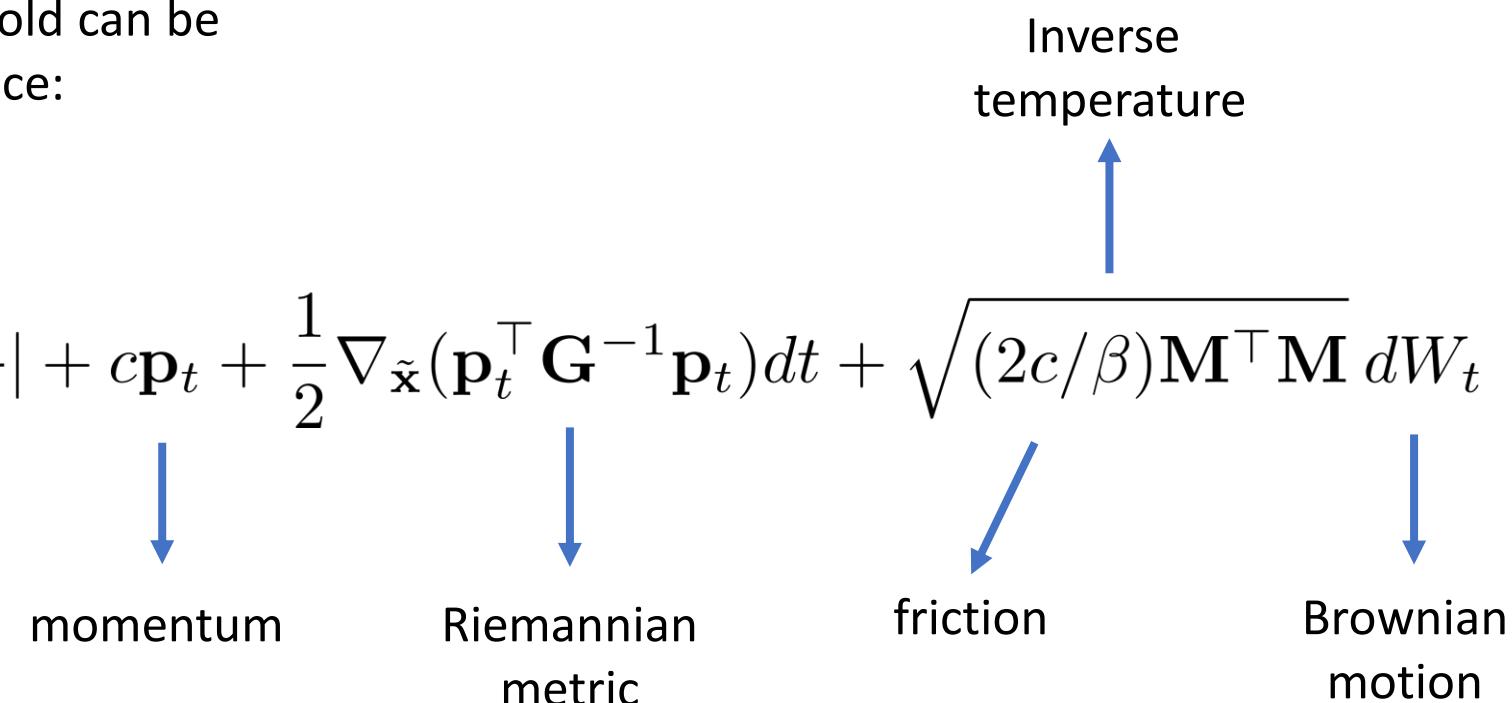
Inference: Tempered Geodesic MCMC

Markov process on the Riemannian manifold can be simulated in the embedded Euclidean space:

$$d\tilde{\mathbf{x}}_t = \mathbf{G}(\tilde{\mathbf{x}}_t)^{-1} \mathbf{p}_t dt$$

$$d\mathbf{p}_t = -\left(\nabla_{\tilde{\mathbf{x}}} U_\lambda(\tilde{\mathbf{x}}_t) + \frac{1}{2} \nabla_{\tilde{\mathbf{x}}} \log |\mathbf{G}| + c \mathbf{p}_t + \frac{1}{2} \nabla_{\tilde{\mathbf{x}}} (\mathbf{p}_t^\top \mathbf{G}^{-1} \mathbf{p}_t) dt + \sqrt{(2c/\beta) \mathbf{M}^\top \mathbf{M}} dW_t \right)$$

Hamiltonian Dynamics

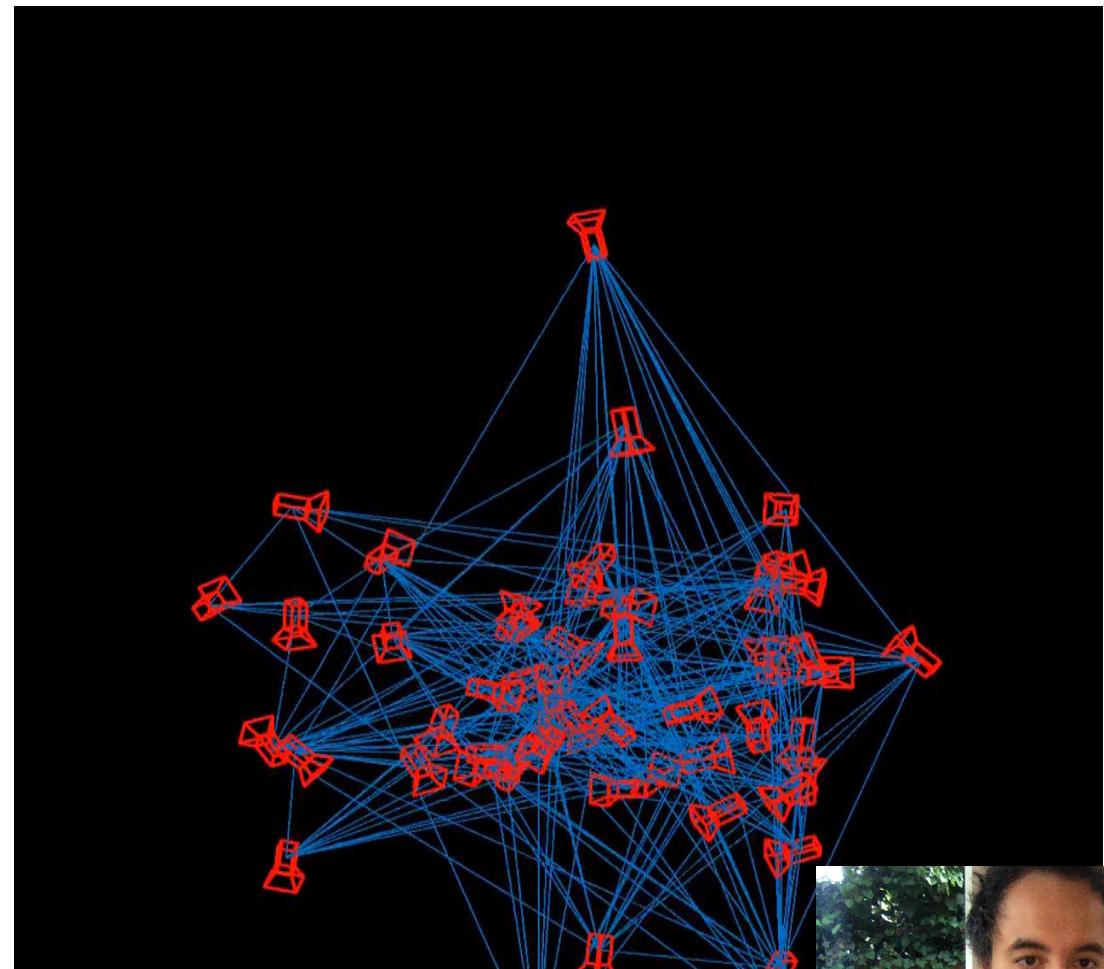
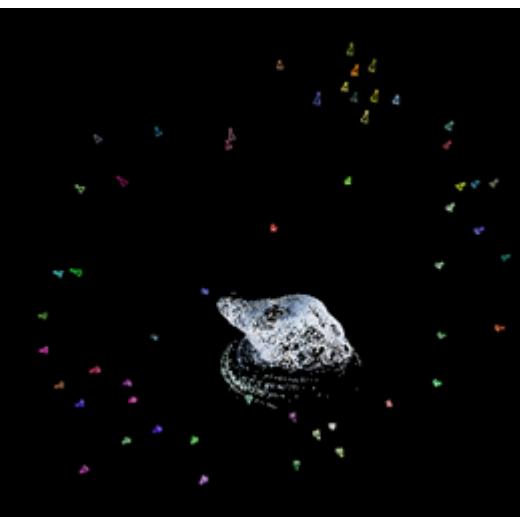


1. Eventually, we will provide samples close to the global minimum (even when non-convex).
2. Suffers from *meta-stability phenomenon* (exponential time required).



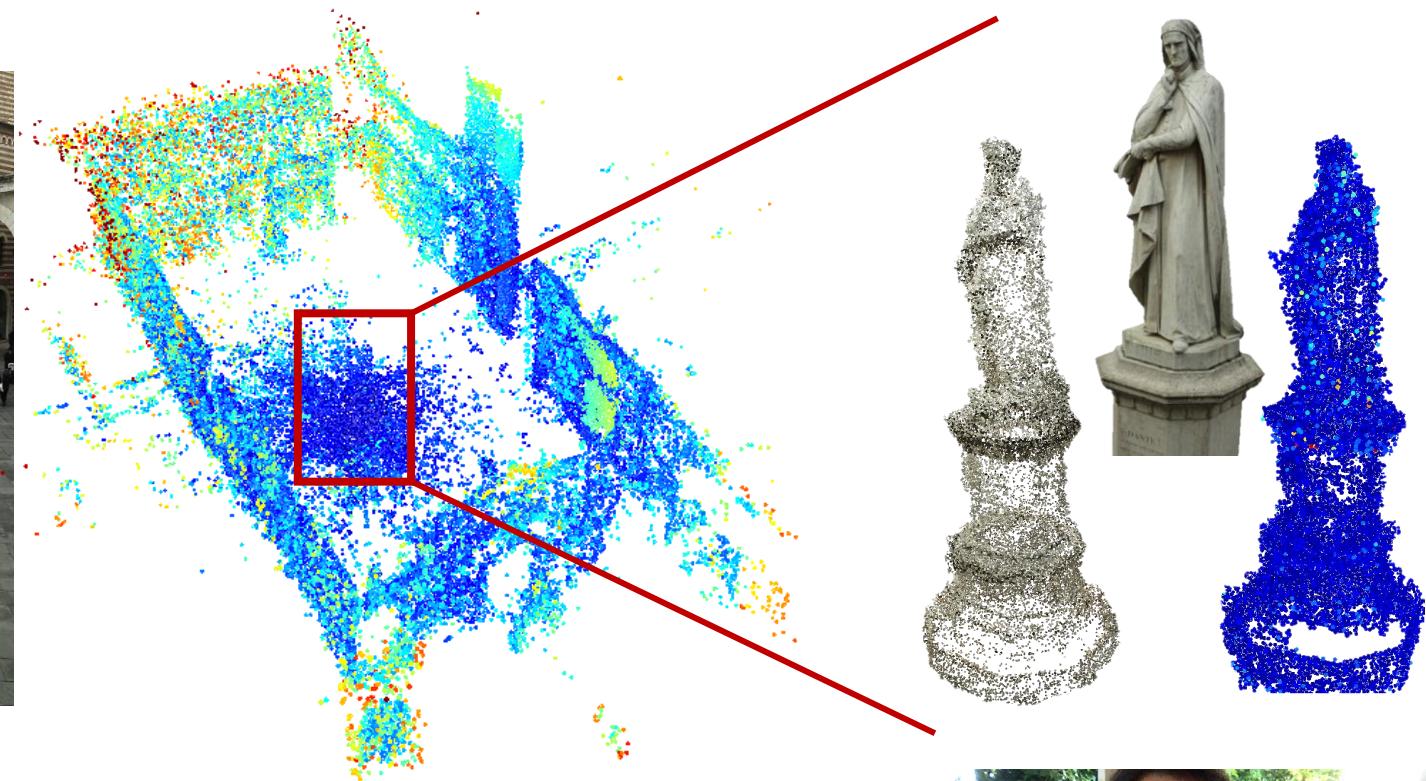
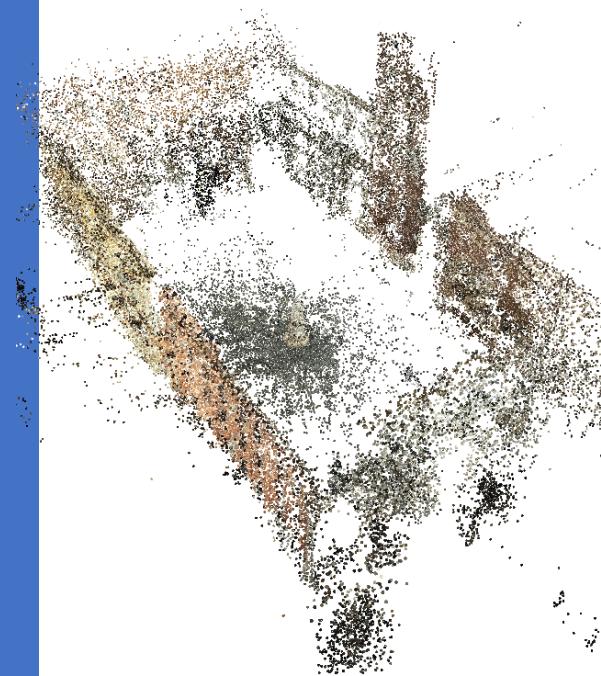


Solving the Classical Problem

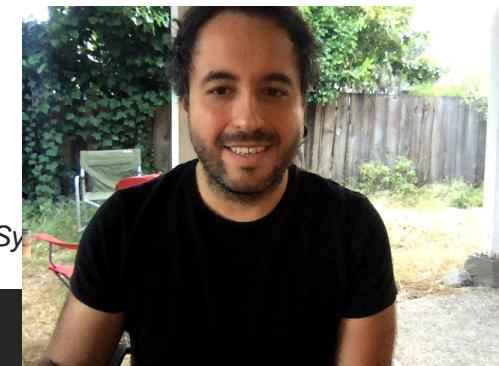


Birdal, Tolga, et al. "Bayesian Pose Graph Optimization via Bingham Distributions and Tempered Geodesic MCMC" *Proceedings of the Neural Information Processing Systems*

Estimating Uncertainty



Birdal, Tolga, et al. "Bayesian Pose Graph Optimization via Bingham Distributions and Tempered Geodesic MCMC" *Proceedings of the Neural Information Processing Systems*

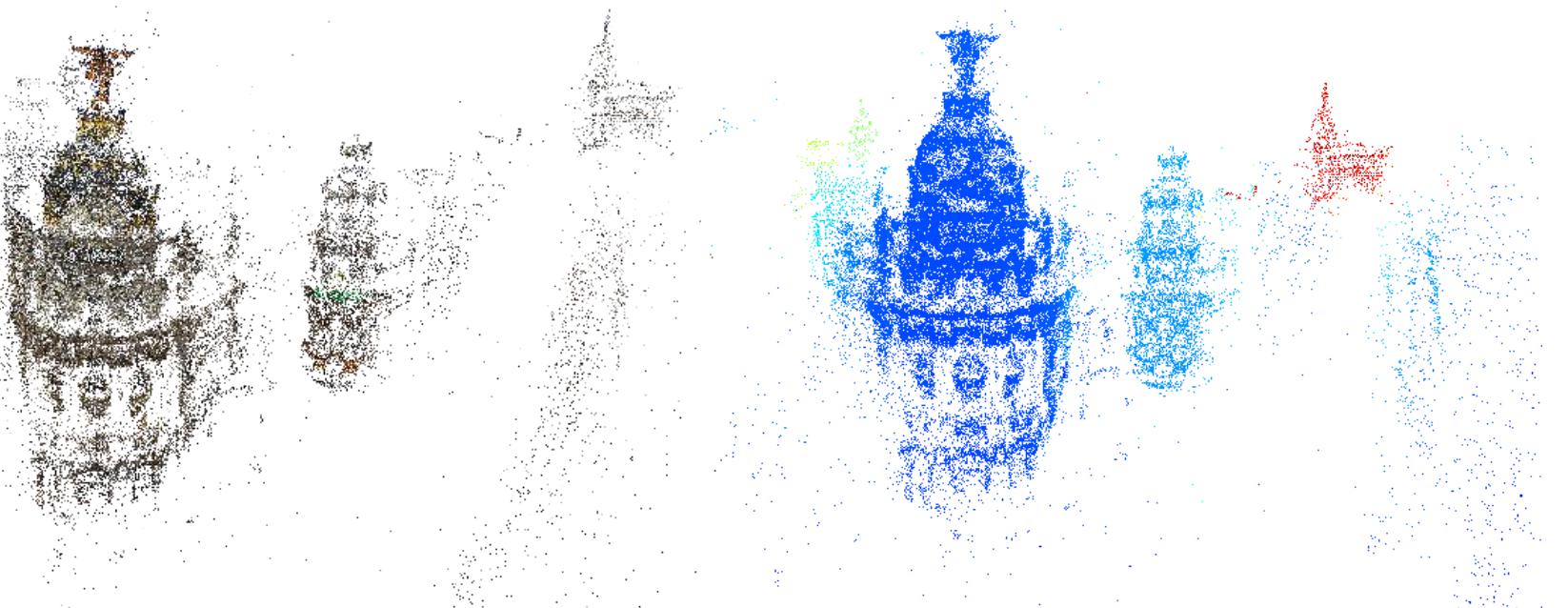




Estimating Uncertainty



(a) Madrid Metropolis



(b) 3D Reconstruction

(c) Uncertainty Map





A New Probabilistic Look to Synchronization

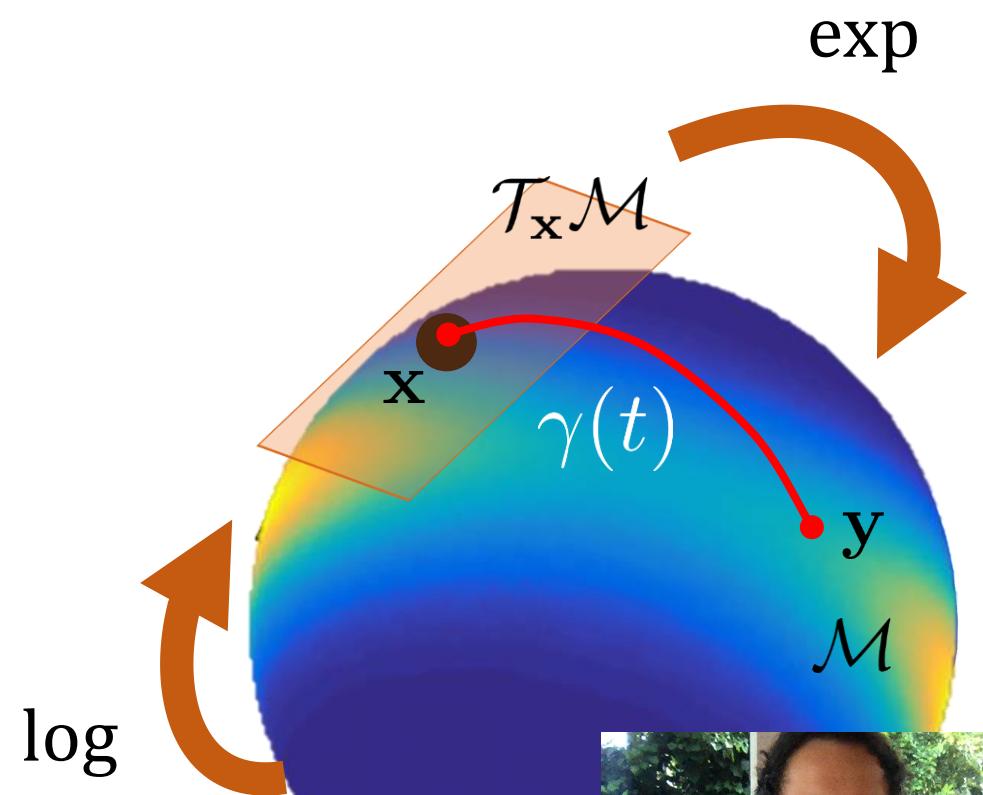
Given a Riemannian Manifold \mathcal{M}

Analytical Geodesic Flow

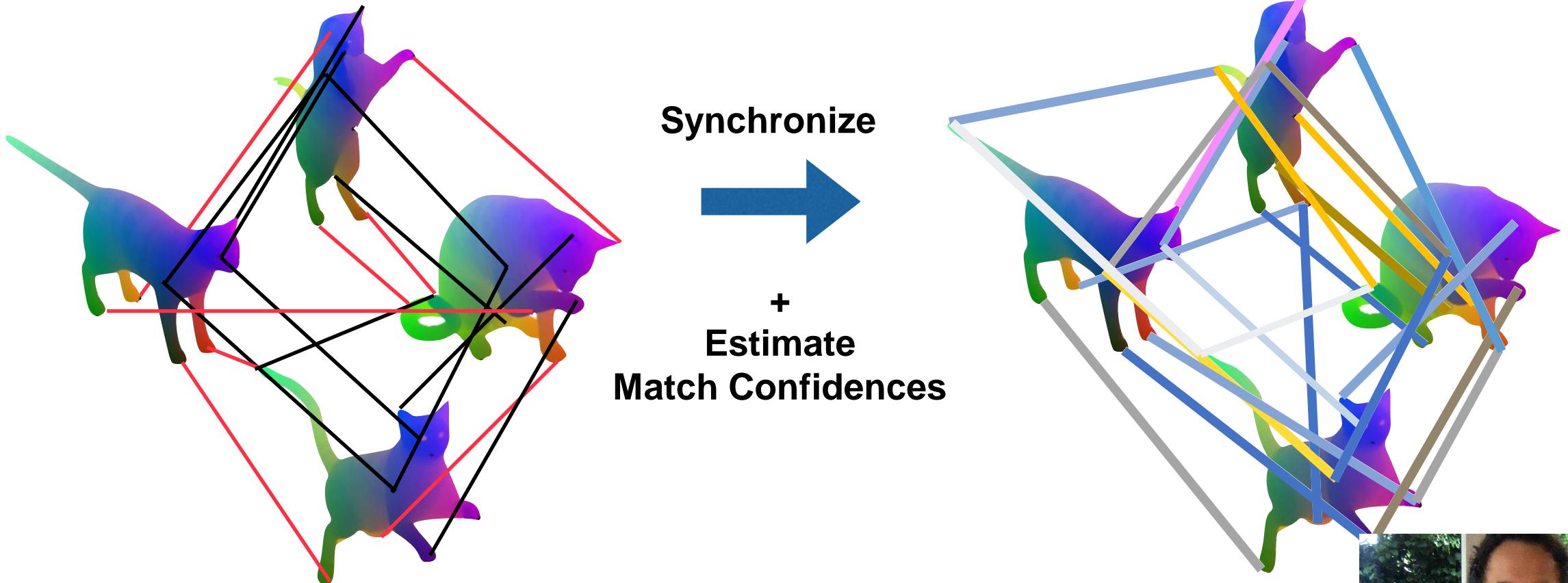
A Suitable Probability Distribution



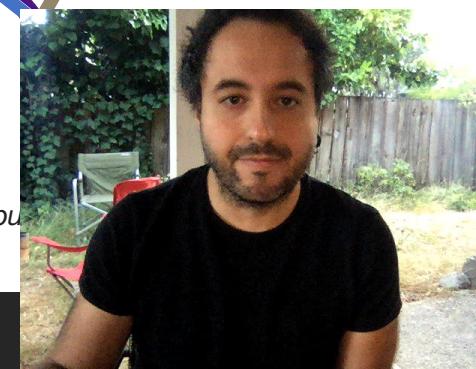
Configurable
Sampling & Optimization



Extension to Correspondence Synchronization (Multi-Graph Matching)



Birdal, Tolga, et al. "Probabilistic Permutation Synchronization using the Riemannian Structure of the Birkhoff Polytope" *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.



Extension to Correspondence Synchronization (Multi-Graph Matching)

$$\mathbf{X}^* = \arg \min_{\mathbf{X} \in \mathcal{P}_n^N} \left\{ U(\mathbf{X}) := \sum_{(i,j) \in E} \|\mathbf{P}_{ij} - \mathbf{X}_i \mathbf{X}_j^\top\|_F^2 \right\}$$

Direct optimization on permutations is difficult.

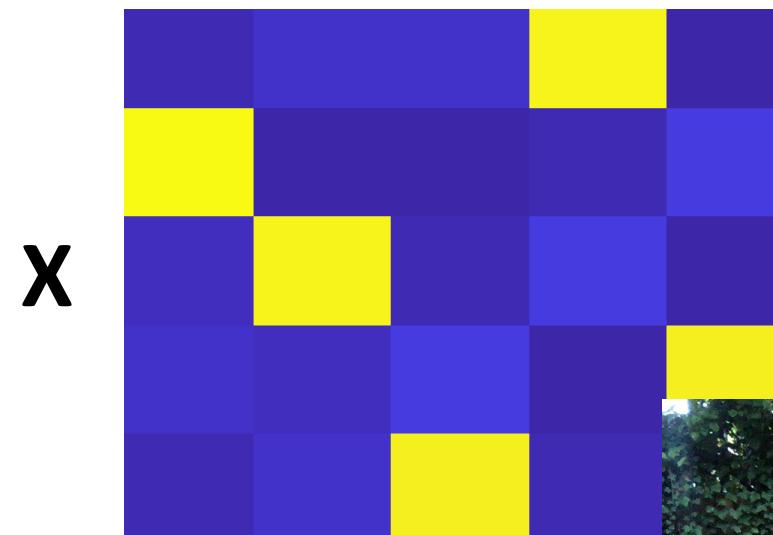
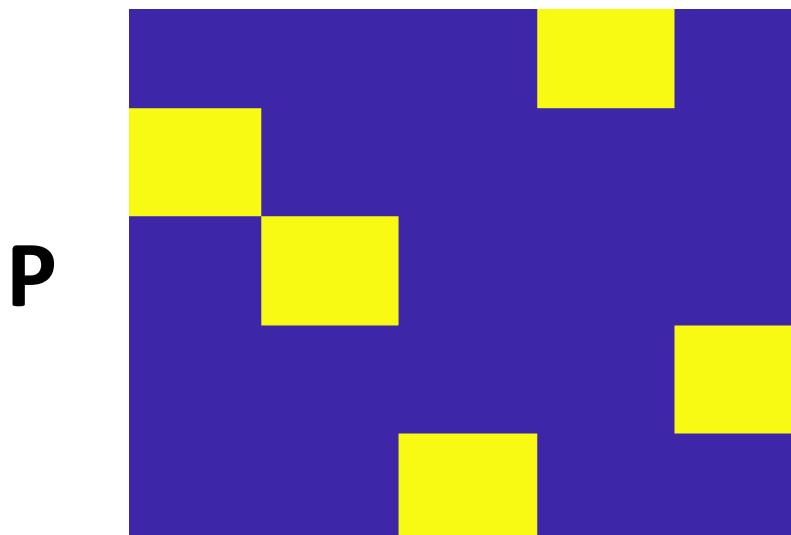


• • • • •

Extension to Correspondence Synchronization (Multi-Graph Matching)

Relaxation onto Doubly Stochastic Matrices

$$\mathcal{DP}_n = \{ \mathbf{X} \in \mathbb{R}_+^{n \times n} : \sum_{i=1}^n x_{ij} = 1 \wedge \sum_{j=1}^n x_{ij} = 1 \}.$$



•••••

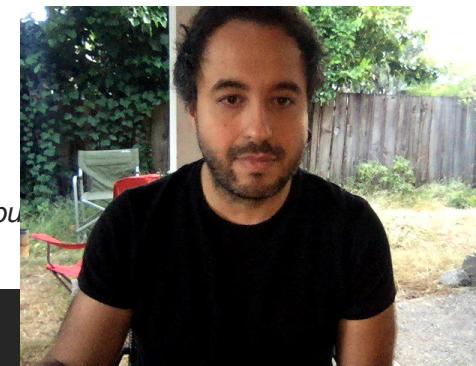
Extension to Correspondence Synchronization (Multi-Graph Matching)

$$\mathbf{X}^* = \arg \min_{\mathbf{X} \in \mathcal{DP}_n^N} \left\{ U(\mathbf{X}) := \sum_{(i,j) \in E} \|\mathbf{P}_{ij} - \mathbf{X}_i \mathbf{X}_j^\top\|_F^2 \right\}$$

Optimize using Riemannian LBFGS

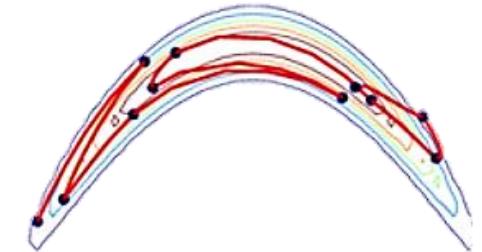
Sample using Birkhoff RLMC

Birdal, Tolga, et al. "Probabilistic Permutation Synchronization using the Riemannian Structure of the Birkhoff Polytope" *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.





Riemannian Langevin Monte Carlo



Riemannian Metric

$$d\tilde{\mathbf{X}}_t = (-\mathbf{G}^{-1} \nabla_{\tilde{\mathbf{X}}} U_\lambda(\tilde{\mathbf{X}}_t) + \boldsymbol{\Gamma}_t) dt + \sqrt{2/\beta \mathbf{G}^{-1}} dB_t$$



Euclidean Embedding

Riemannian Correction

Brownian Motion

Embedded Posterior Density

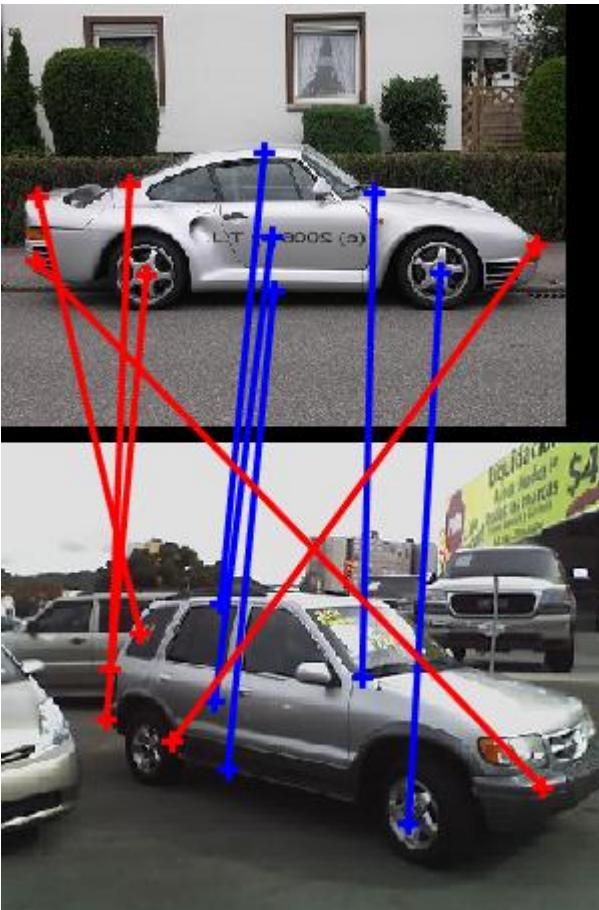
$$\pi_{\mathcal{H}}(\mathbf{x}) = \pi_\lambda(\tilde{\mathbf{X}}) / \sqrt{|\mathbf{G}(\tilde{\mathbf{X}})|}$$



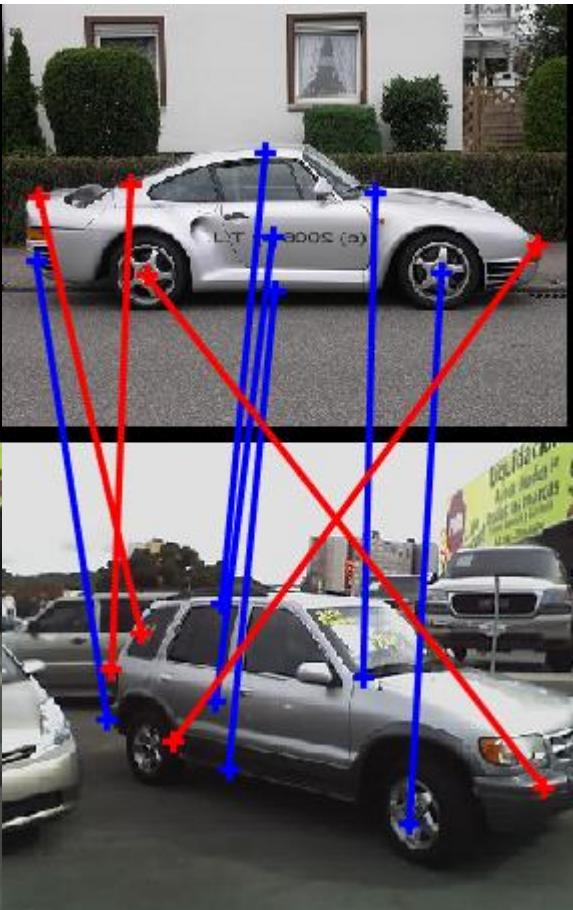


Minimizing Objective & Estimating Uncertainty

(a) Initialization



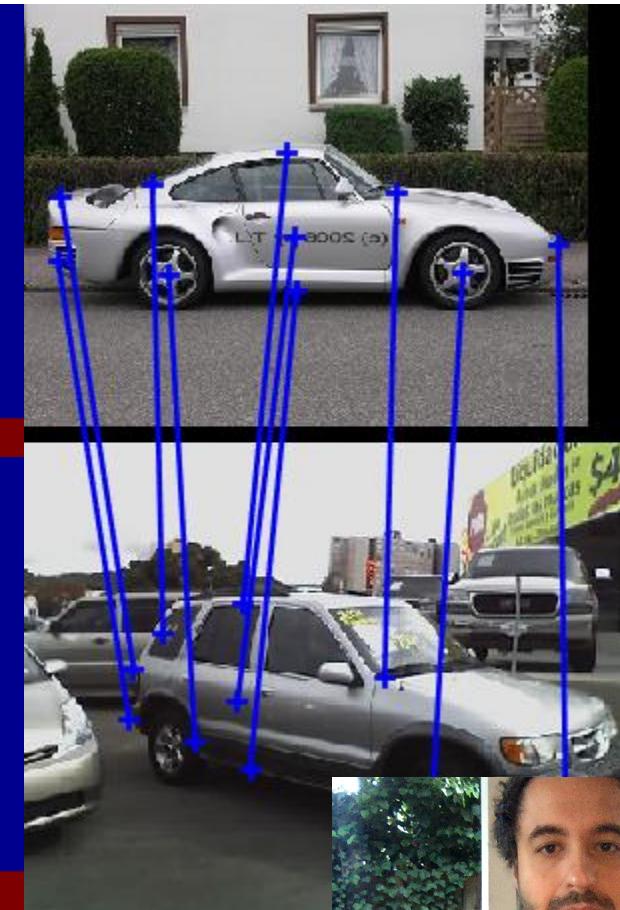
(b) Solution



(c) Confidence

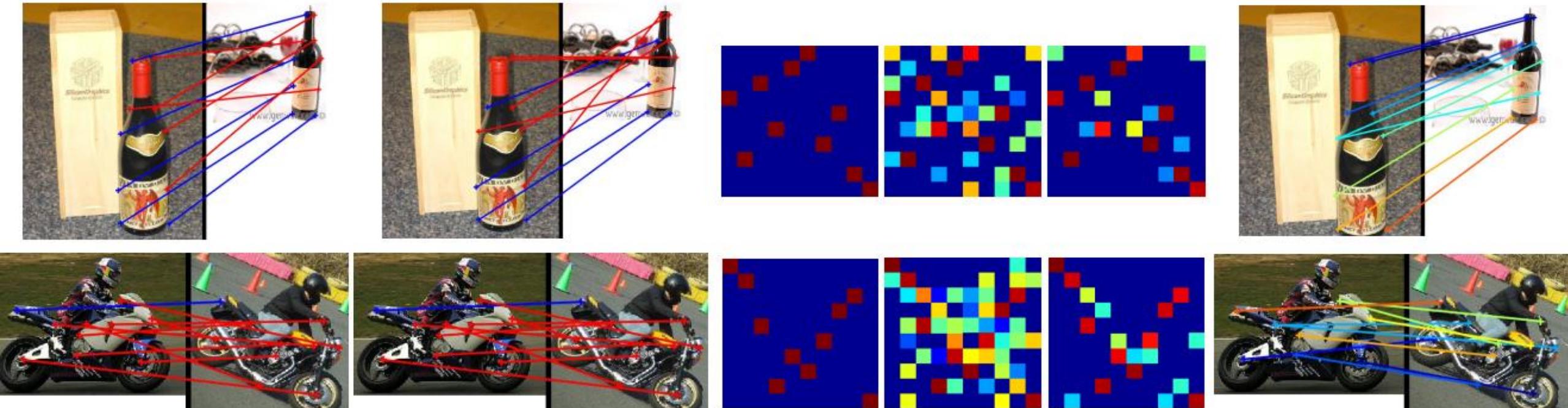


(d) Multiple Hypotheses Solution



Birdal, Tolga, et al. "Probabilistic Permutation Synchronization using the Riemannian Structure of the Birkhoff Polytope" *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

Minimizing Objective & Estimating Uncertainty



Birdal, Tolga, et al. "Probabilistic Permutation Synchronization using the Riemannian Structure of the Birkhoff Polytope" *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

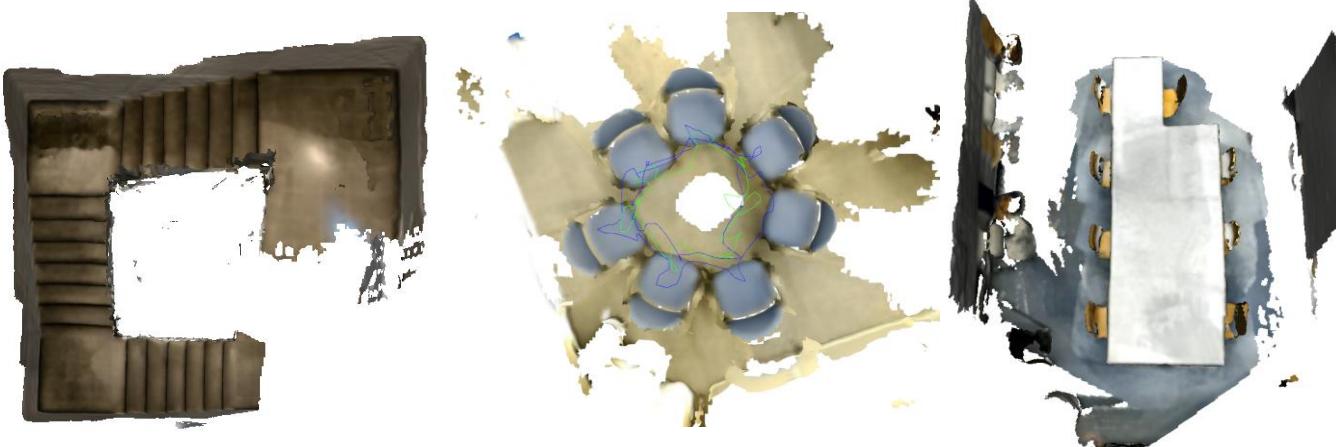




Real life is challenging.



Ambiguous Views



Uncertainty



Ambiguities

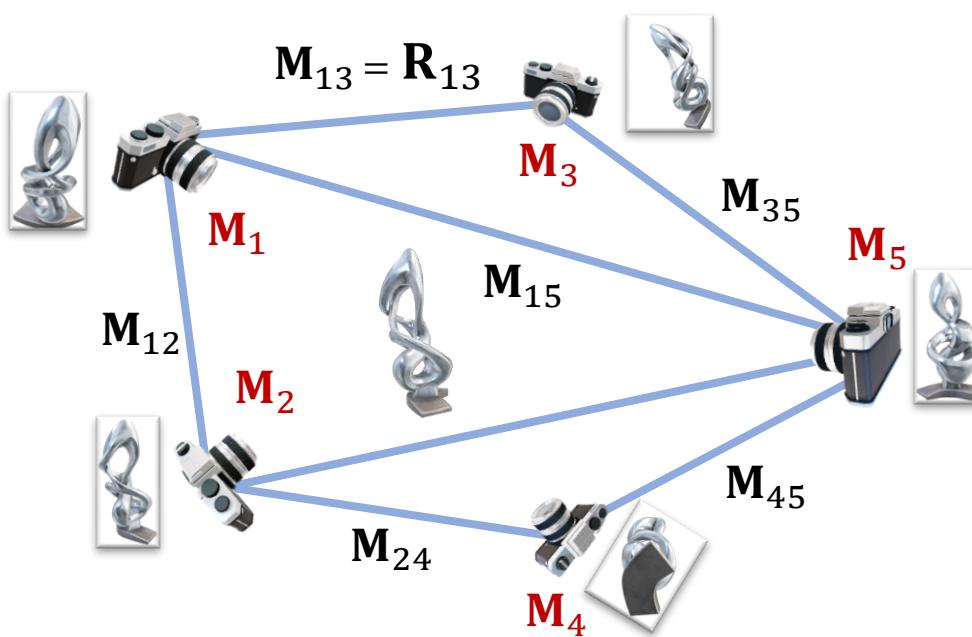
Bui, Mai, Tolga Birdal, Haowen Deng, Shadi Albarqouni, Leonidas Guibas, Slobodan Ilic, and Nassir Navab. "6D Camera Relocalization in Ambiguous Scenes via Multimodal Inference." *arXiv preprint arXiv:2004.04807* (2020).





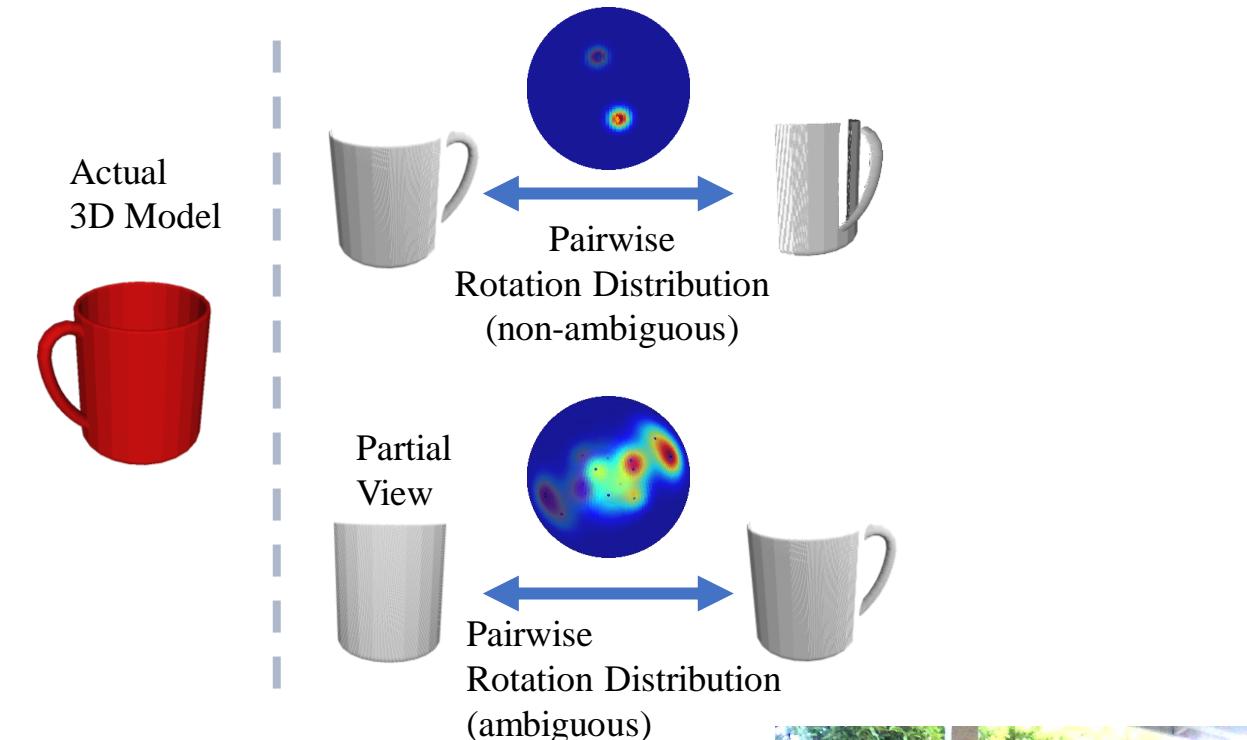
Synchronizing Distributions

Classical Synchronization



M_i : Absolute rotation M_{ij} : Relative rotation

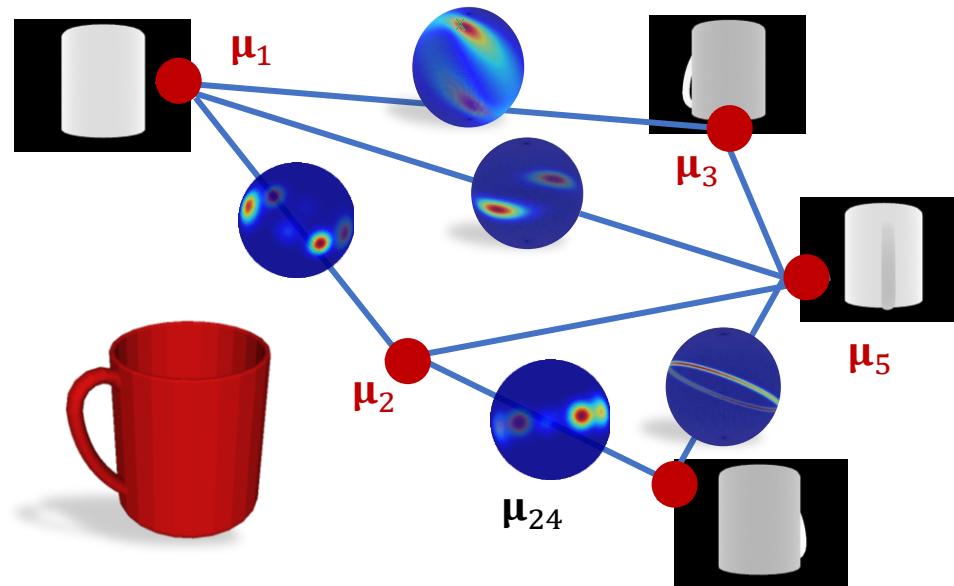
The Ambiguity Problem





Synchronizing Distributions = Measure Synchronization

Classical Rotation Synchronization



**relative empirical
probability measure**

$$\mu_{ij} \triangleq \sum_{k=1}^{K_{ij}} w_{ij}^{(k)} \delta_{\mathbf{q}_{ij}^{(k)}}$$

**absolute (sought)
probability measure**

$$\mu_i \triangleq \sum_{k=1}^{K_i} w_i^{(k)} \delta_{\mathbf{q}_i^{(k)}}$$

$$\mathbf{R}_{ij} \approx \mathbf{R}_j \mathbf{R}_i^{-1}, \forall i \neq j$$



Measure Synchronization

$$\mathcal{D}(g_{ij}(\mu), \mu_{ij}) \rightarrow 0$$

Divergence
(e.g.
Wasserstein
distance)

composition
functions
(relative
rotation)

coupling
(of absol
measure)



How to Solve?

Classical Rotation Synchronization

$$\arg \min_{\{\mathbf{R}_i\}_i} \sum_{(i,j) \in \mathcal{E}} d_{\text{geo}}(\mathbf{R}_{ij}, \mathbf{R}_j \mathbf{R}_i^T)$$

s. t. $\mathbf{R}_i \in SO(3), \forall i \in \{1, \dots, n\}$

Riemannian Measure Synchronization

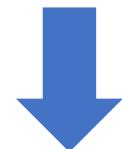
$$\min_{\boldsymbol{\mu}} \sum_{(i,j) \in \mathcal{E}} \mathcal{D}(g_{ij}(\boldsymbol{\mu}), \boldsymbol{\mu}_{ij}) + \mathcal{R}(\boldsymbol{\mu})$$

s. t. $[w_i^1, \dots, w_i^{K_i}]^\top \in \mathcal{C}_i, \quad \forall i \in \{1, \dots, n\}$
 $\boldsymbol{\mu}_i \in \mathbb{H}_1, \quad \forall i \in \{1, \dots, n\}$

Multiple Optimal Transport Problems
on the Lie Group of Quaternions

Solved by many existing methods:

*Spectral, Low-rank, Riemannian descent,
Semi-definite programming, Lie algebraic
and etc.*



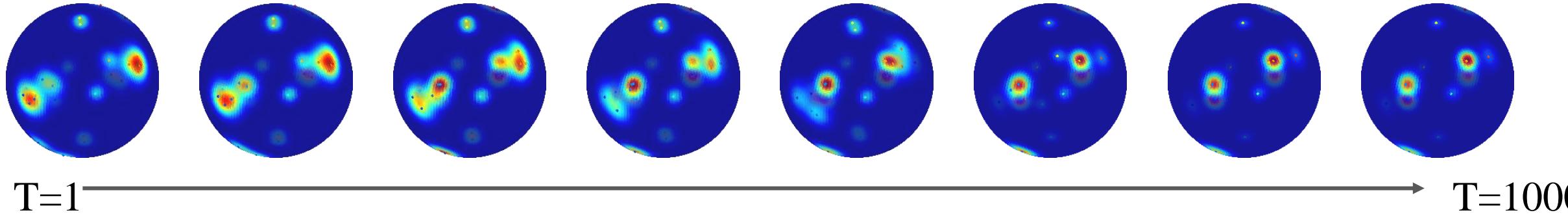
Proposed

Riemannian Particle Gradient
via **Geodesic Sinkhorn Divergence**
with convergence guarantee





Theoretical Guarantees

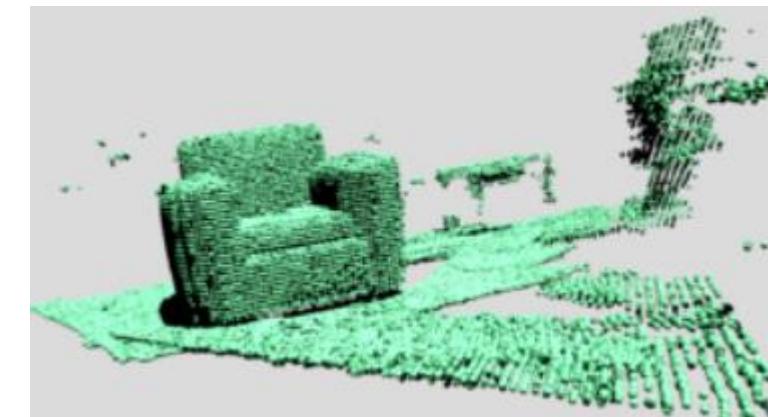
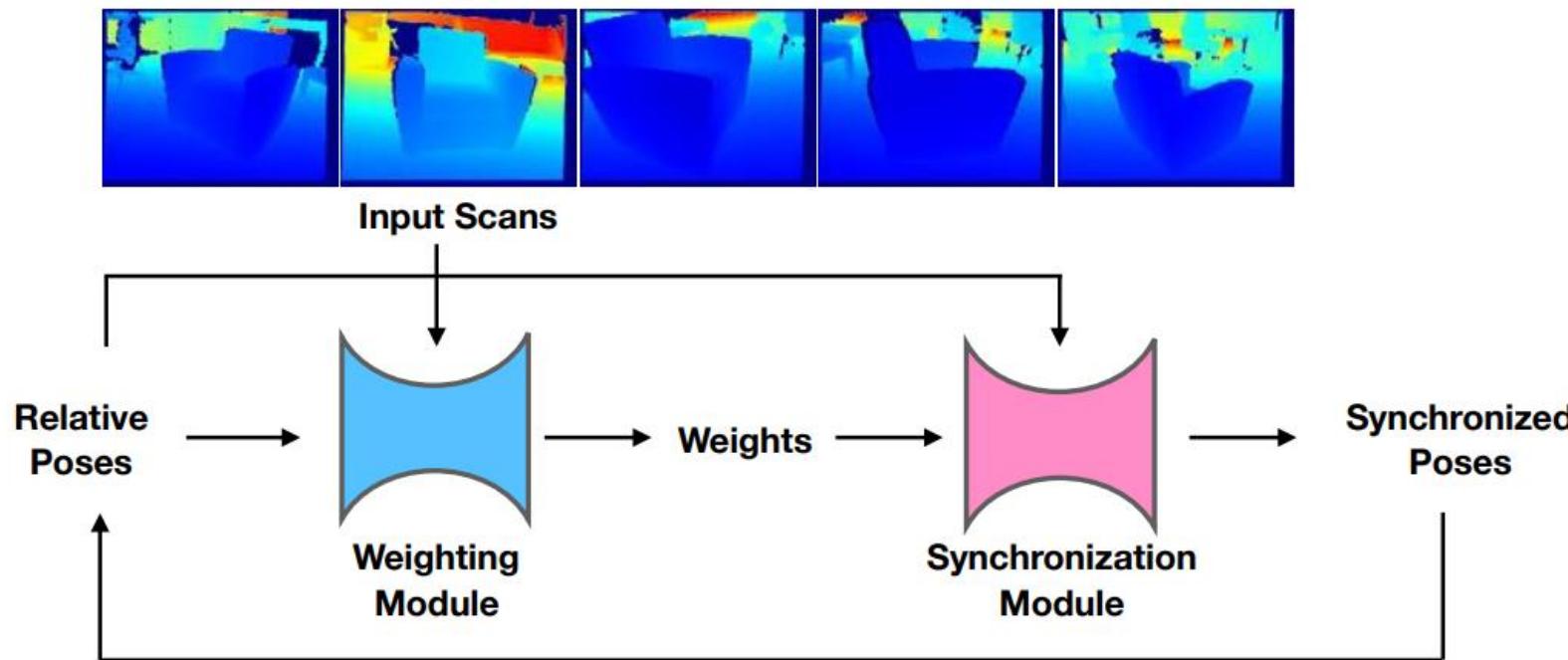


Our algorithm is shown to converge to the global solution under mild assumptions.





Use of Differentiable Synchronization in Neural Networks

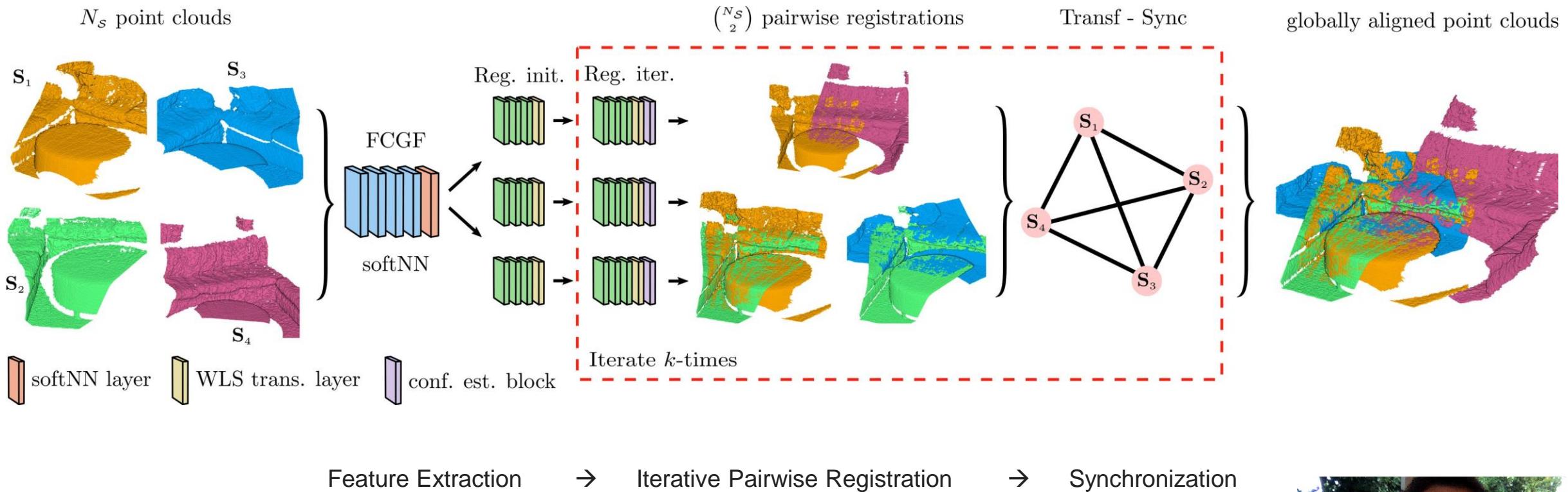


Huang, Xiangru, et al. "Learning transformation synchronization." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.





Use of Differentiable Synchronization in Neural Networks



Feature Extraction

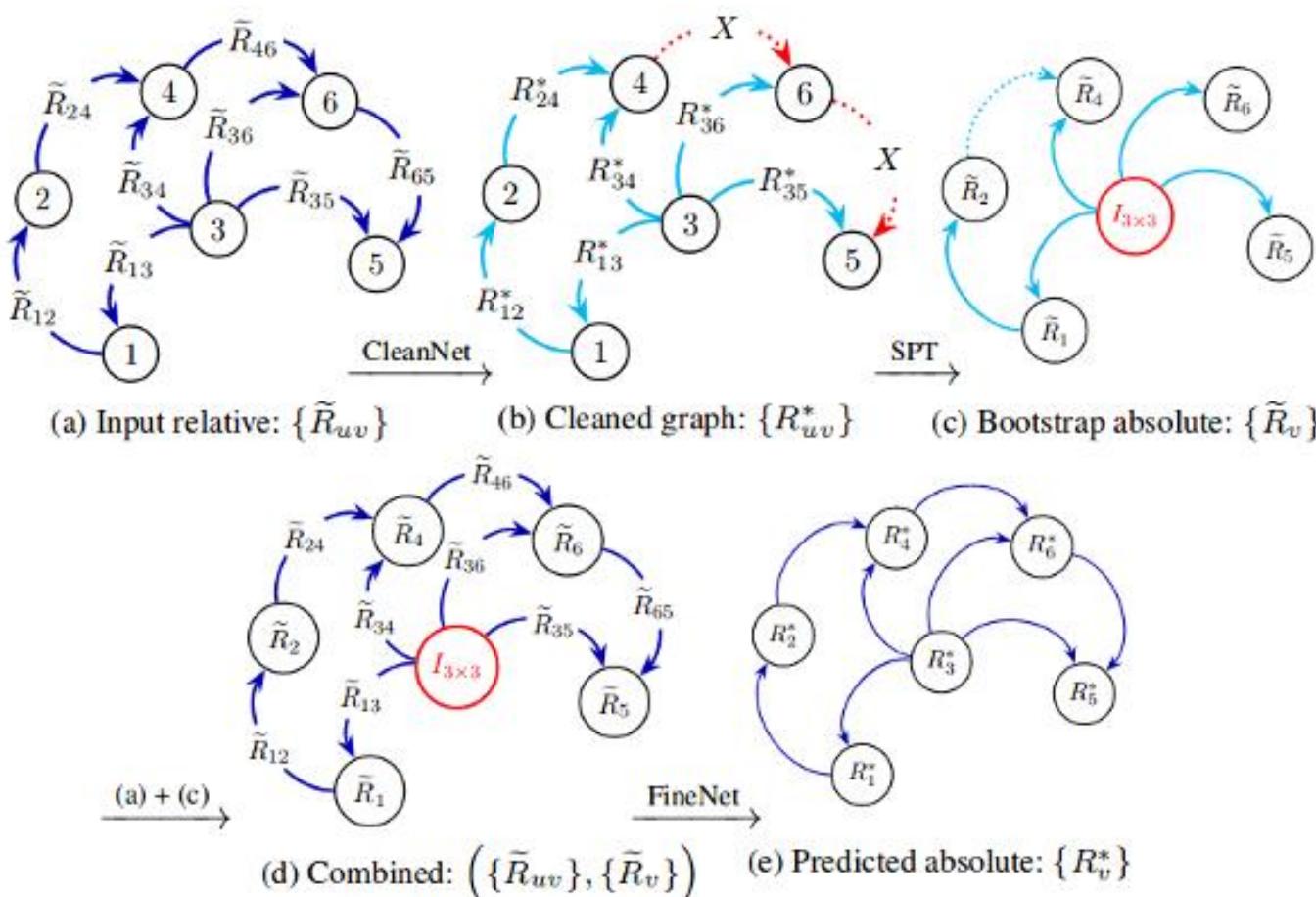
→ Iterative Pairwise Registration

→ Synchronization

Gojcic, Zan, Caifa Zhou, Jan D. Wegner, Leonidas J. Guibas, and Tolga Birdal. "Learning multiview 3D point cloud registration." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2020.



Learning to Synchronize Robustly



Purkait, Pulak, Tat-Jun Chin, and Ian Reid. "NeuRoRA: Neural Robust Rotation Averaging." *arXiv preprint arXiv:1912.04485* (2019).

Table 2. Results of Rotation averaging on a test synthetic dataset. The average angular error on all the view-graphs in our dataset is displayed. The proposed method NeuRoRA is remarkably faster than the baselines while producing better results. NeuRoRA-v2 is a variation of NeuRoRA where the initialization from CleanNet is fine-tuned in a two-step of FineNet. There is no GPU implementations of [5] and [16] available, thus these methods are excluded from the runtime comparisons on cuda.

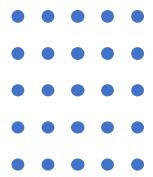
	Angular Error			Runtime (seconds)	
Baseline Methods	mn	md	rms	cpu	cuda
Chatterjee [5]	2.17°	1.25°	4.55	5.38s	(1×)
Weiszfeld [16]	3.35°	1.02°	9.74	50.92s	(0.11×)
Proposed Methods	mn	md	rms	cpu	cuda
CleanNet-SPT + [5]	2.11°	1.26°	4.04	5.41s	(0.99×)
CleanNet-SPT + [16]	1.74°	1.01°	3.53	50.36s	(0.11×)
NeuRoRA	1.45°	0.74°	3.53	0.21s	(24×)
NeuRoRA-v2	1.30°	0.68°	3.28	0.30s	(18×)
Other Methods	mn	md	rms	cpu	cuda
CleanNet-SPT	2.93°	1.47°	5.34	0.11s	(47×)
SPT-FineNet	3.00°	1.57°	6.12	0.11s	(47×)
SPT-FineNet + [5]	2.12°	1.26°	4.11	5.41s	(0.99×)
SPT-FineNet + [16]	1.78°	1.01°	3.95	50.36s	(0.11×)
NeuRoRA + [5]	2.11°	1.26°	4.04	5.51s	(0.97×)
NeuRoRA + [16]	1.73°	1.01°	3.51	50.46s	(0.10×)

mn: mean of the angular error, md: median of the angular error,

rms: root mean square angular error,

MethodA + MethodB: MethodB is initialized by MethodA





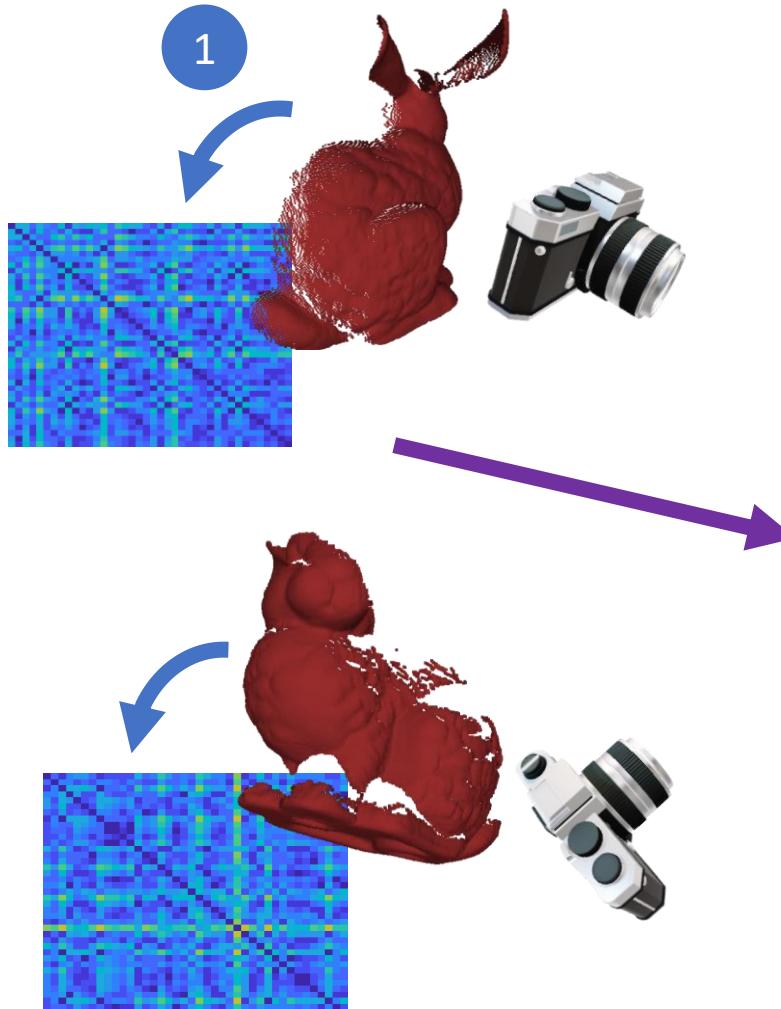
Where to go next?

- Training Neural Networks via Synchronization
- A Complete Treatment of Uncertainty

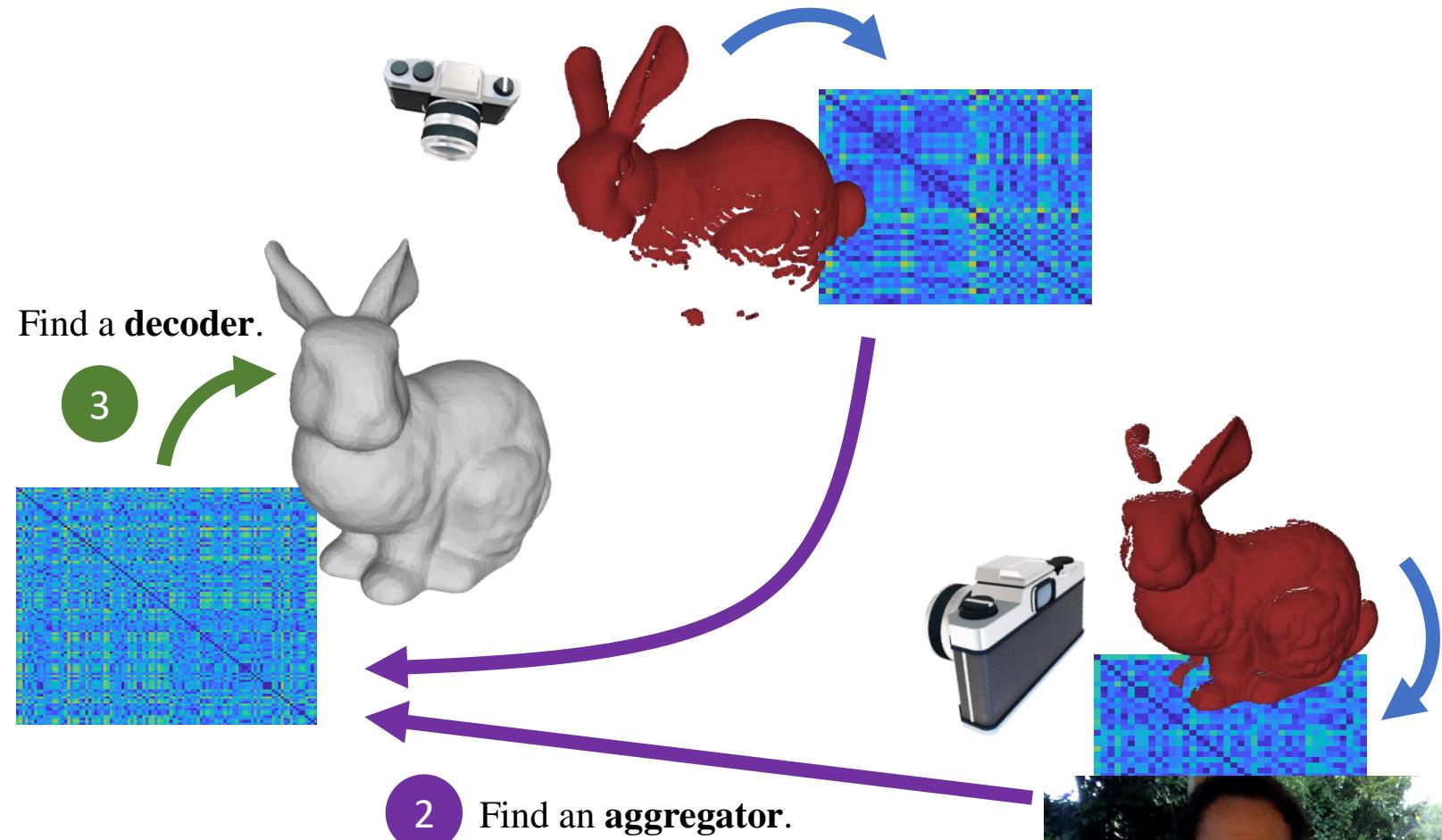


Synchronization without Synchronization

Find an invariant **embedding**.



Find a **decoder**.

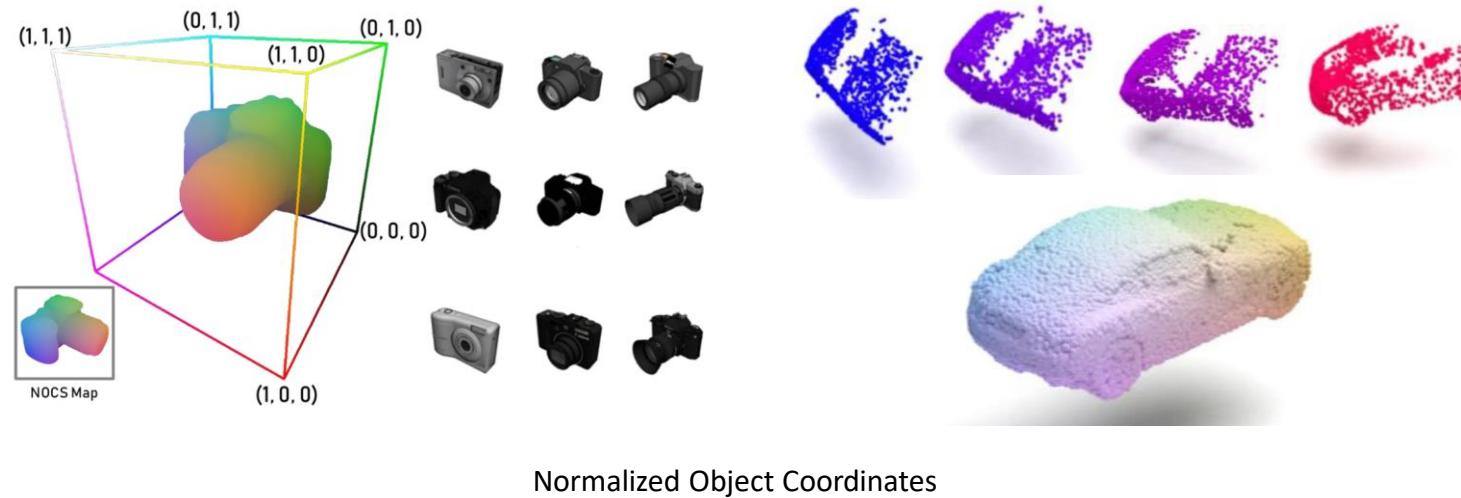


2 Find an **aggregator**.

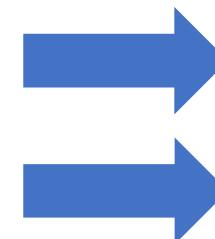




Synchronization without Synchronization: Canonicalizers



Canonicalization
Synchronization

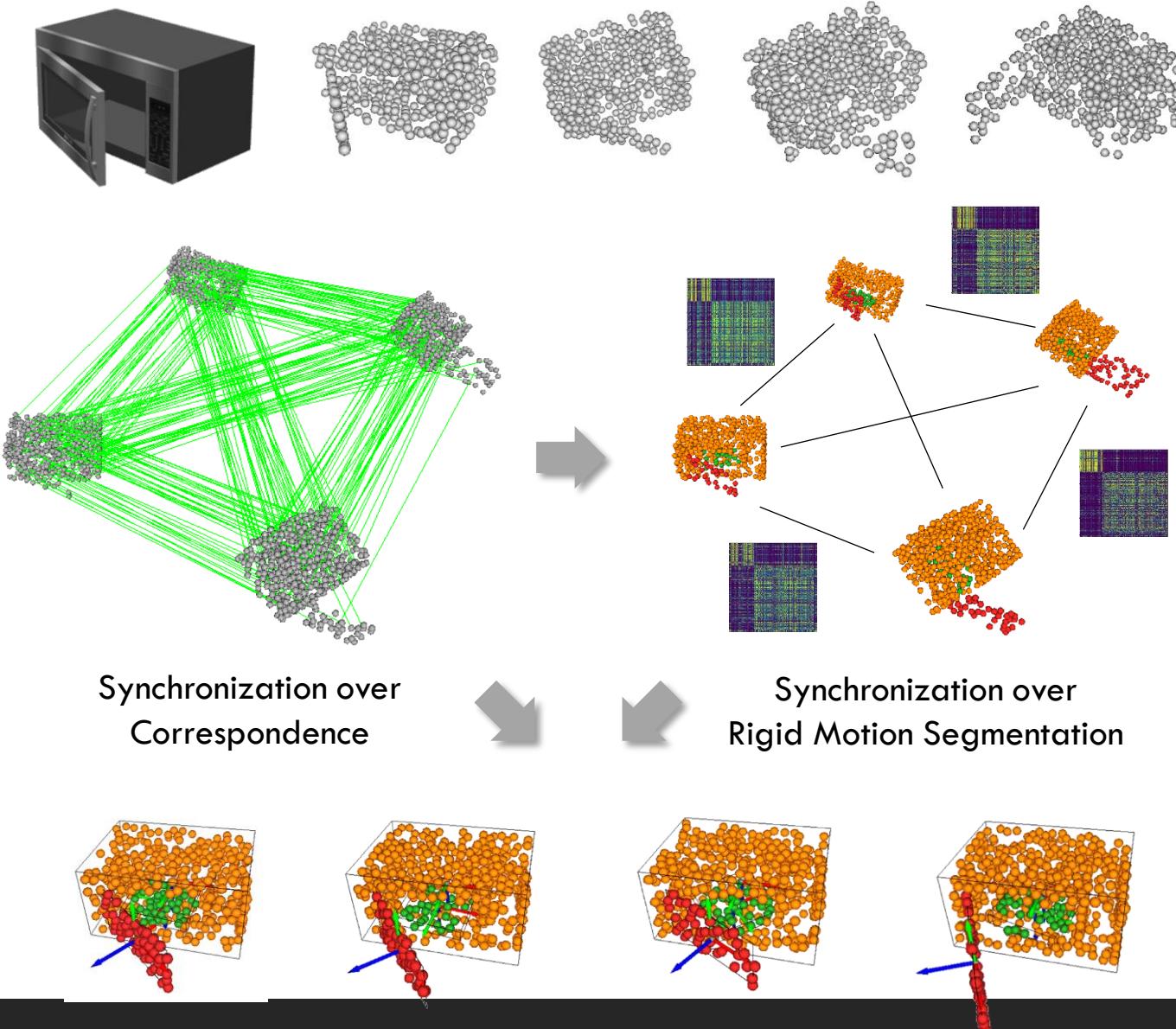


Invariance
Equivariance?

Hybrid methods?

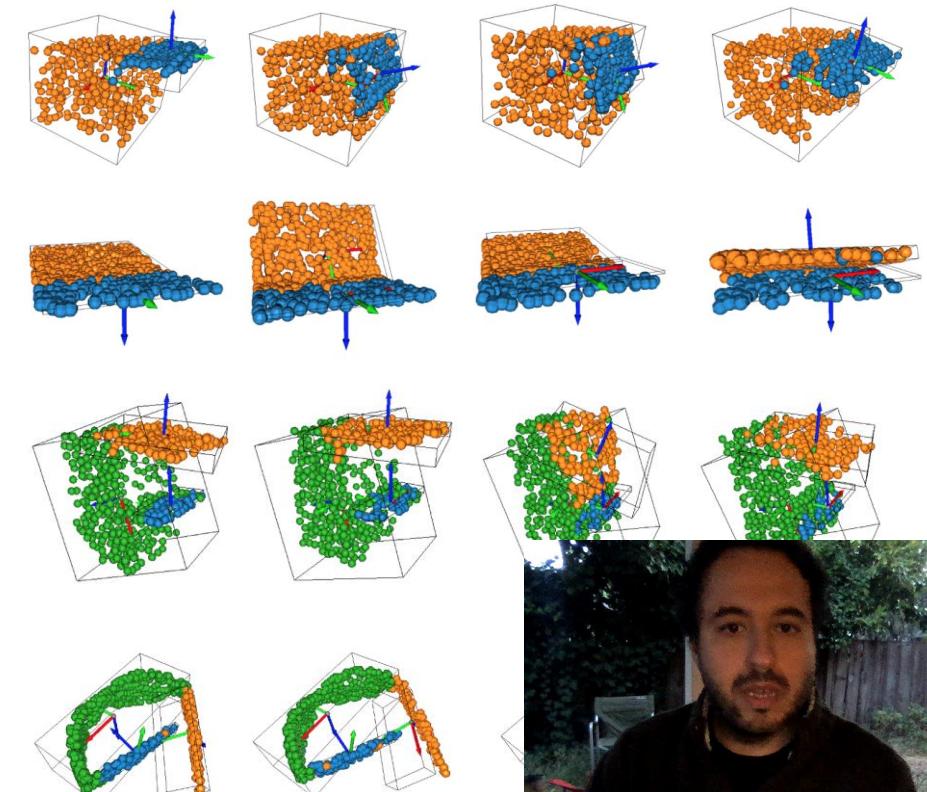


Dealing with Articulation: Simultaneous Synchronization of Motion and Segmentation



Motion based analysis can naturally generalize across shapes from different categories!

Results



That's all folks.

