Learned trajectory embedding for subspace clustering

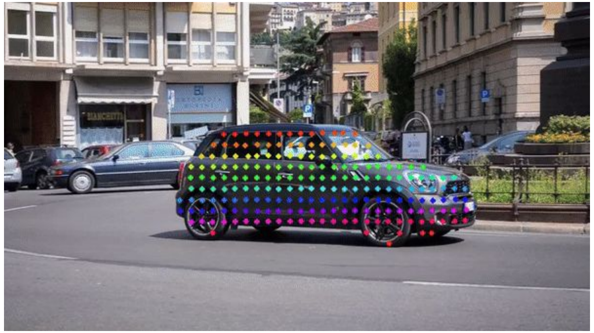
Yaroslava Lochman [1]   Carl Olsson [1,2]   Christopher Zach [1]

[1] Chalmers University of Technology   [2] Lund University

SEATTLE, WA
CVPR
JUNE 17-21, 2024

Learned trajectory embedding for subspace clustering

Point trajectories



Courtesy of **Tumanyan, Singer et al.** *DINO-tracker: taming DINO for self-supervised point tracking in a single video*

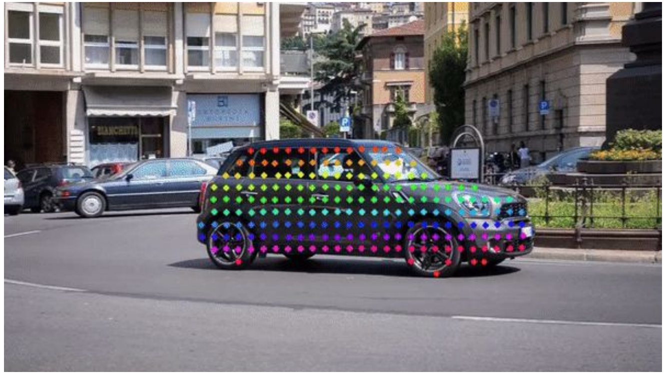
Courtesy of **Wang et al.** *Tracking everything everywhere all at once*

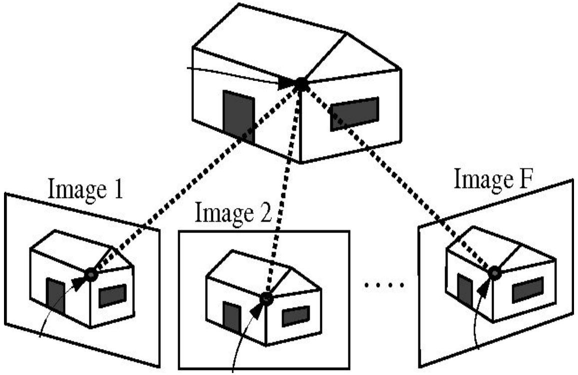
Learned trajectory embedding for subspace clustering

Rigid motion estimation

Matrix factorization for shape and motion reconstruction

$$\underbrace{\mathtt{M}}_{2F\times P} \approx \overbrace{\underbrace{\mathtt{B}}_{2F\times r}}^{\text{motions}} \overbrace{\underbrace{\mathtt{C}}_{r\times P}}^{\text{shapes}}, \qquad r \ll 2F$$



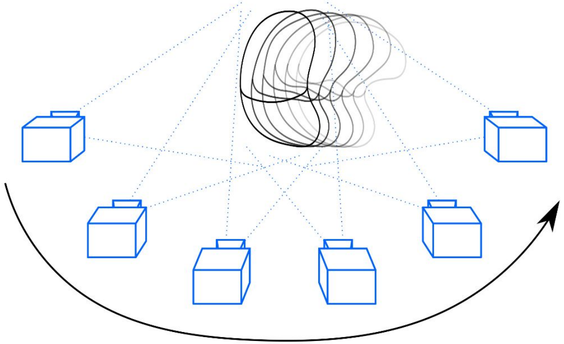


Learned trajectory embedding for subspace clustering

Nonrigid motion estimation

Matrix factorization for shape and motion reconstruction

$$\underbrace{\mathtt{M}}_{2F \times P} \approx \overbrace{\underbrace{\mathtt{B}}_{2F \times r}}^{\text{“motions”}} \overbrace{\underbrace{\mathtt{C}}_{r \times P}}^{\text{“shapes”}}, \qquad r \ll 2F$$



Courtesy of **Badias et al.** *MORPH-DSLAM: model order reduction for physics-based deformable SLAM*

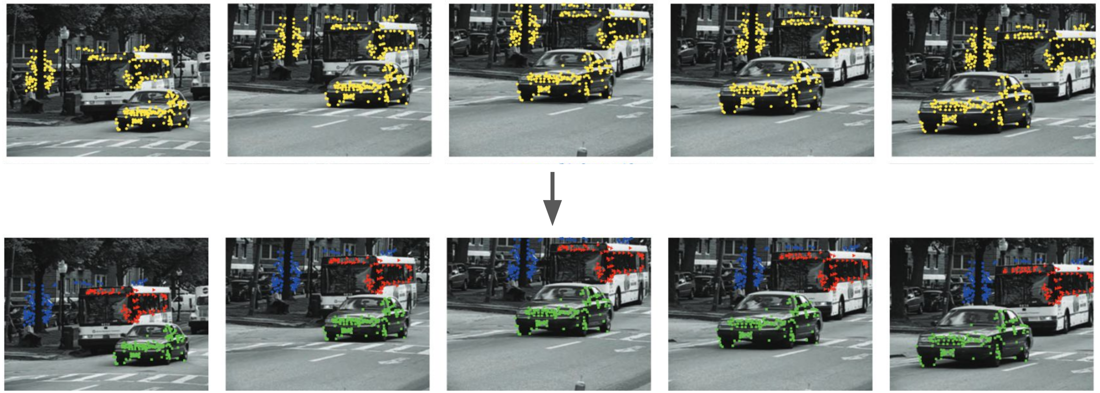
Learned trajectory embedding for subspace clustering

Motion segmentation

Multiple independent motions. Important for dynamic scene understanding

Chicken-and-egg problem (even for **rigid motion**) + contaminated by outliers and missing points

$$\mathrm{M}\,\underbrace{\mathrm{P}_{\pi}}_{P\times P} \approx \Big[\underbrace{\mathrm{B}_1\mathrm{C}_1}_{\text{group 1}} \quad \dots \quad \underbrace{\mathrm{B}_c\mathrm{C}_c}_{\text{group c}}\Big]$$



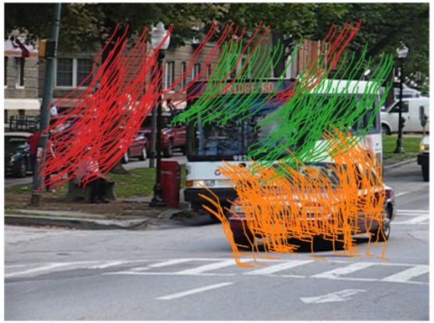
Learned trajectory embedding for subspace clustering

Motion segmentation methods

- Hypothesis generation-based *Robust statistical methods, joint optimization*
- Spectral clustering *Affinity matrix design, pairwise or multi-view relations*
  - Sparse subspace clustering *Use self-expressiveness, sparse optimization*

*“Unfortunately, traditional cluster-based trajectory segmentation methods rely on **heavy optimization** and **hand-crafted features**, and are hard to scale with dense trajectories”* — **Zhao et al.** ParticleSfM

Heavy test-time optimization



Learned trajectory embedding for subspace clustering

Proposed method

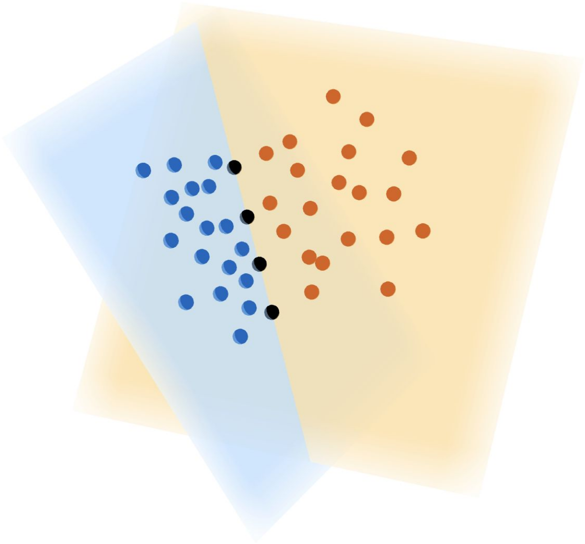
Goal: learn trajectory feature representation useful for clustering
so that no simultaneous grouping and motion estimation at test-time is needed

Learned trajectory embedding for subspace clustering

Disjoint subspace assumption

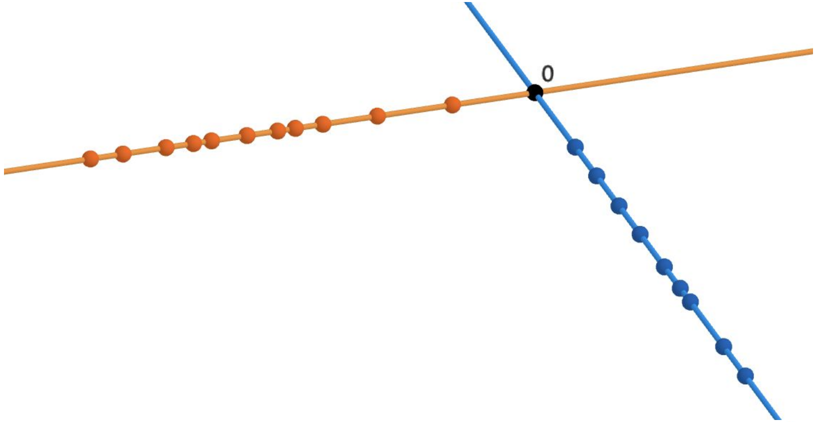
- Motion models do not intersect in high dimensional trajectory space
- In this work, we build on this **disjoint subspace assumption**

Planes in 3D



vs.

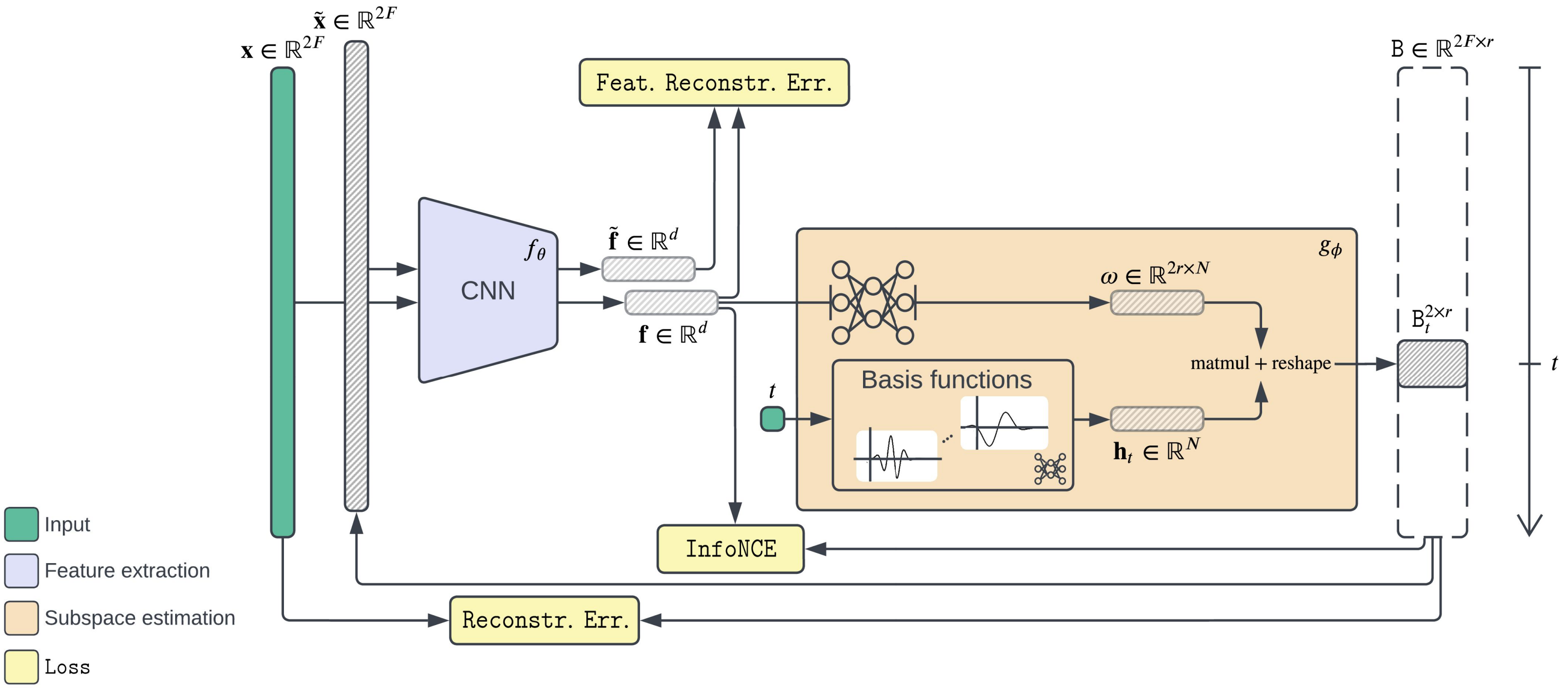
Motions as low dim subspaces in high dim space



$$\dim(\mathcal{S}_i \oplus \mathcal{S}_j) = \dim(\mathcal{S}_i) + \dim(\mathcal{S}_j)$$

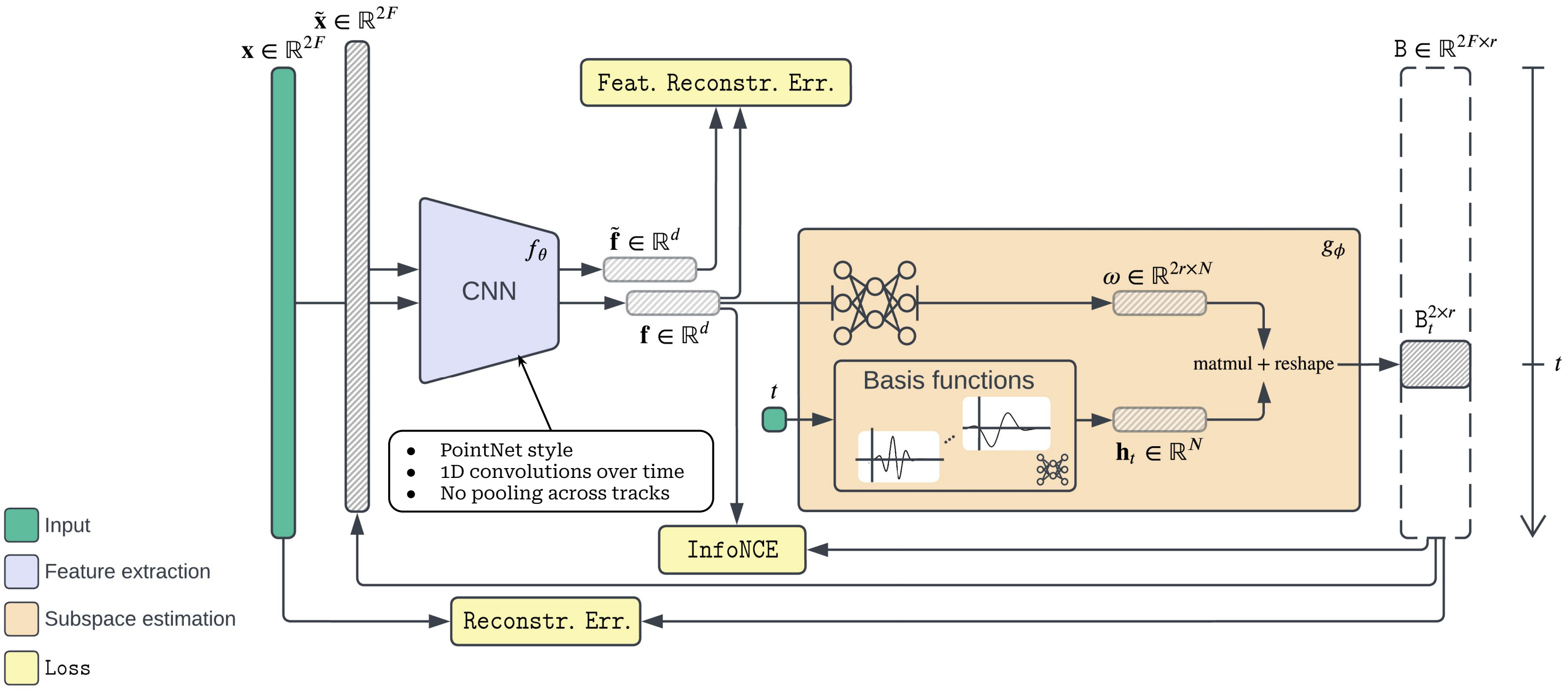
Learned trajectory embedding for subspace clustering

Proposed approach



Learned trajectory embedding for subspace clustering

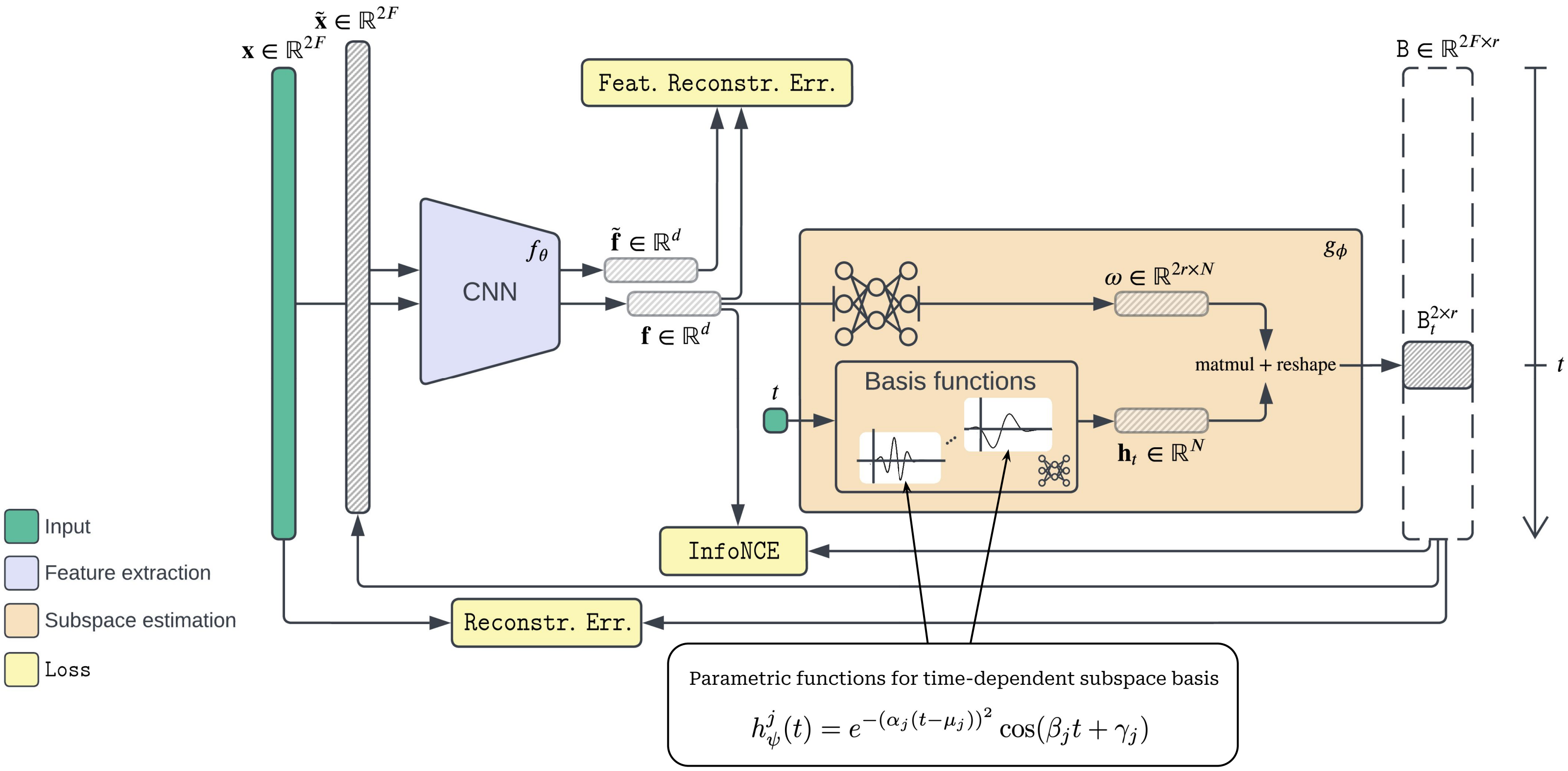
Proposed approach



CHALMERS UNIVERSITY OF TECHNOLOGY
LUND UNIVERSITY
WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

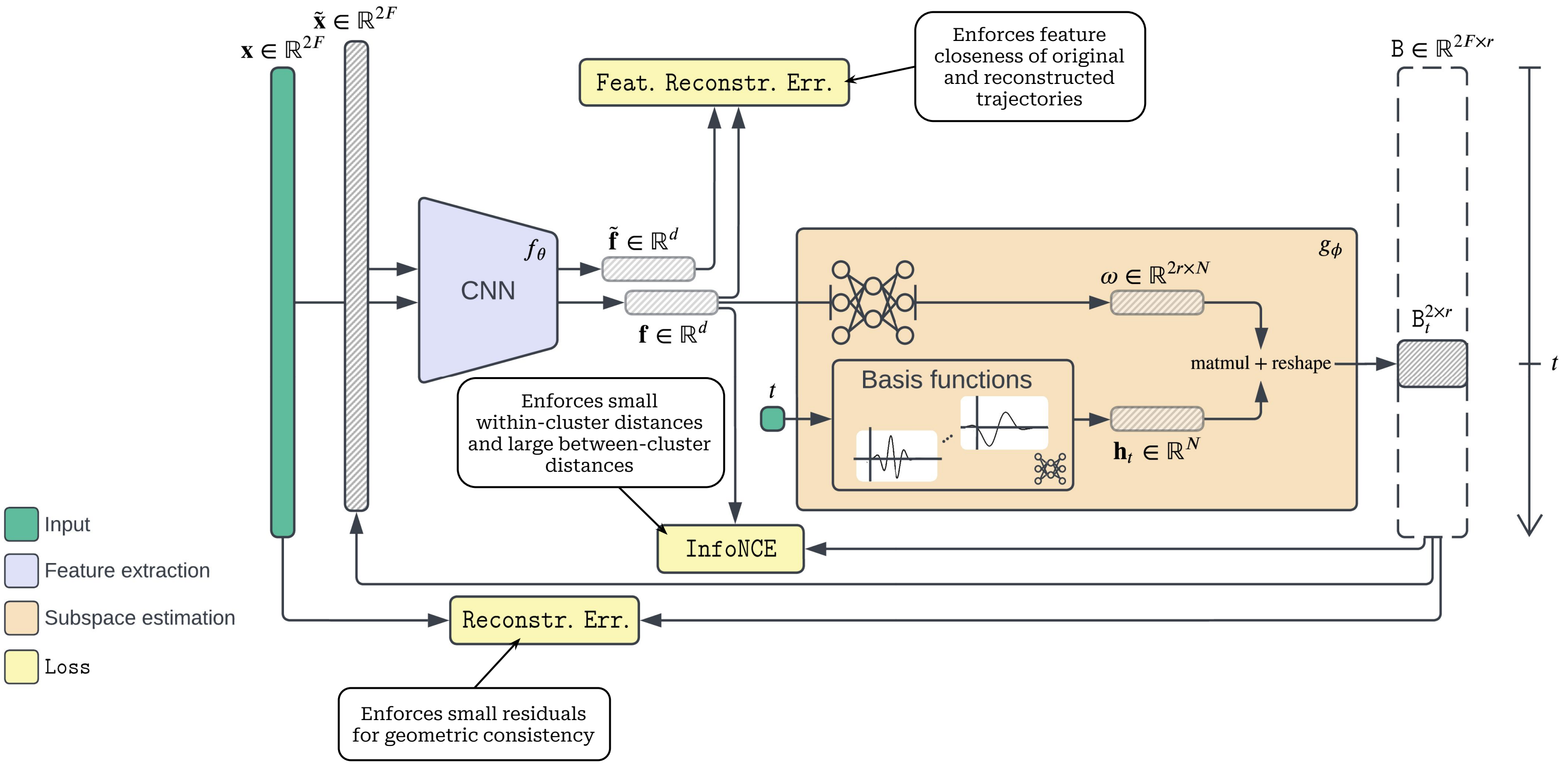
Learned trajectory embedding for subspace clustering

Proposed approach



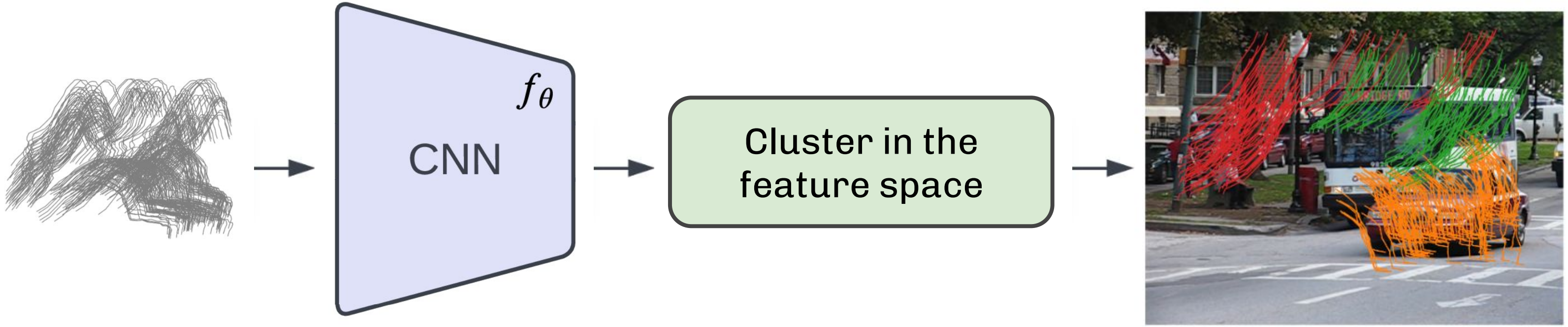
Learned trajectory embedding for subspace clustering

Proposed approach



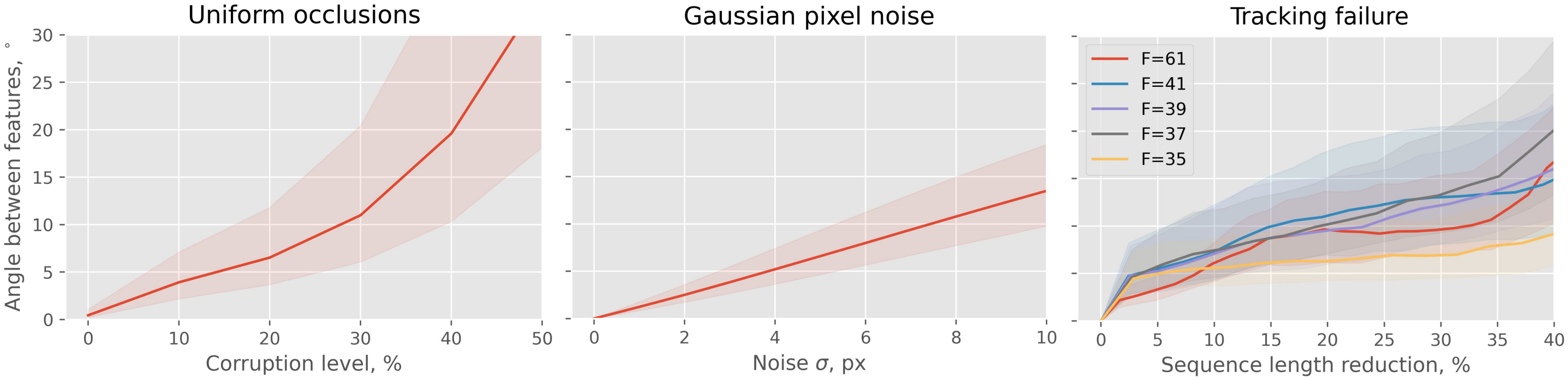
Learned trajectory embedding for subspace clustering

Inference: fully observed trajectory



Learned trajectory embedding for subspace clustering

Approximate invariances of f_θ



Learned trajectory embedding for subspace clustering

Inference: trajectory completion

- Formulate an objective to fill-in missing values

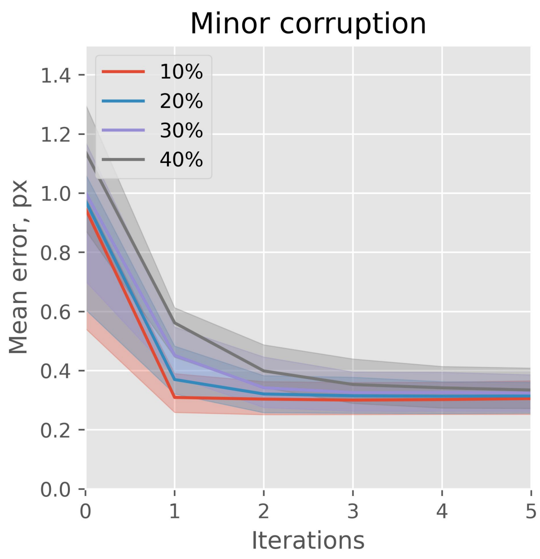
$$\left\|\hat{\mathbf{x}}(\bar{\mathbf{x}}) - \mathtt{B}\mathtt{B}^{\dagger}\hat{\mathbf{x}}(\bar{\mathbf{x}})\right\|^2 \rightarrow \min_{\bar{\mathbf{x}}}$$

- Obtain linear solution for a fixed subspace

$$\bar{\mathbf{x}}^* = \mathtt{A}(\mathtt{B})\mathbf{x}$$

- Yields an iterative procedure

$$\begin{aligned}\mathtt{B}_0 &\leftarrow B_{\theta,\phi}(\mathbf{x}_{\mathsf{vis}}, \mathbf{t})\\ \bar{\mathbf{x}}_i &\leftarrow \mathtt{A}(\mathtt{B}_{i-1})\mathbf{x}\\ \mathtt{B}_i &\leftarrow B_{\theta,\phi}(\mathbf{w}\odot\mathbf{x}+\bar{\mathbf{w}}\odot\bar{\mathbf{x}}_i, \mathbf{t})\end{aligned}$$



Learned trajectory embedding for subspace clustering

Inference: trajectory completion

- Formulate an objective to fill-in missing values

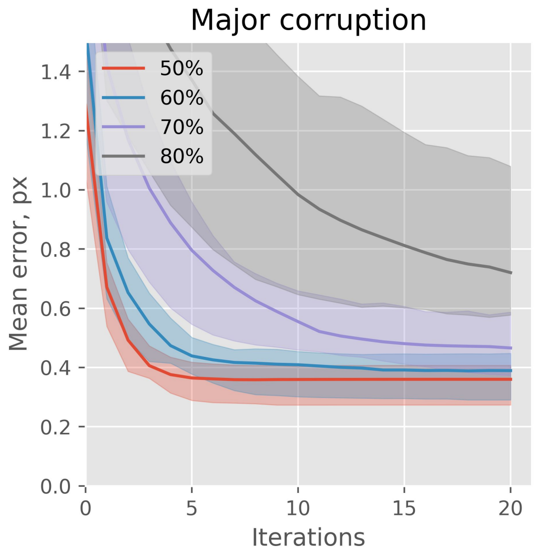
$$\left\|\hat{\mathbf{x}}(\bar{\mathbf{x}}) - \mathtt{B}\mathtt{B}^{\dagger}\hat{\mathbf{x}}(\bar{\mathbf{x}})\right\|^2 \to \min_{\bar{\mathbf{x}}}$$

- Obtain linear solution for a fixed subspace

$$\bar{\mathbf{x}}^* = \mathtt{A}(\mathtt{B})\mathbf{x}$$

- Yields an iterative procedure

$$\begin{aligned}\mathtt{B}_0 &\leftarrow B_{\theta,\phi}(\mathbf{x}_{\text{vis}}, \mathbf{t})\\ \bar{\mathbf{x}}_i &\leftarrow \mathtt{A}(\mathtt{B}_{i-1})\mathbf{x}\\ \mathtt{B}_i &\leftarrow B_{\theta,\phi}(\mathbf{w}\odot\mathbf{x}+\bar{\mathbf{w}}\odot\bar{\mathbf{x}}_i, \mathbf{t})\end{aligned}$$



Learned trajectory embedding for subspace clustering

Clustering error on standard datasets

Method	Hopkins155 Mean	Hopkins155 Median	Hopkins155 Time	Hopkins12 Mean	Hopkins12 Median	KT3DMoSeg Mean	KT3DMoSeg Median
RANSAC	9.76	3.21	194ms	-	-	-	-
GPCA	10.34	2.54	417ms	-	-	34.60	33.95
MSL	5.03	0.00	19h 11m	-	-	-	-
LSA	4.94	0.90	9.47s	-	-	38.30	38.58
ALC_5	3.76	0.26	5m 15s	3.81	0.17	24.31	19.04
$\mathrm{ALC}_{\mathrm{sp}}$	3.37	0.49	6m 11s	1.28	1.07	-	-
LRR	5.41	0.53	1.1s	-	-	33.67	36.01
SSC	2.45	0.20	920ms	-	-	33.88	33.54
RSIM	1.01	0.00	176ms	0.68	0.70	-	-
MultiCons	4.40	-	40ms	-	-	-	-
Ours	0.62	0.0	9ms	5.12	2.04	5.85	0.80