Learned Trajectory Embedding for Subspace Clustering

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

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

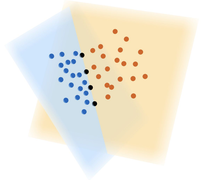
CVPR JUNE 17-21, 2024

WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM WASP

Introduction

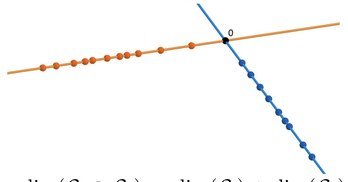
- **Formulation:** cluster 2D point trajectories by rigid motion (no visual inputs).
- **Background:** robust statistical methods + joint optimization, spectral clustering, sparse subspace clustering — most are either sub-optimal or very slow.
- **Observation:** motion models as low dim. subspaces do not intersect in high dim. trajectory space. We build on this **disjoint subspace assumption**.

PLANES IN 3D



vs

MOTIONS IN HIGH DIM. SPACE



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

Learning trajectory representation by contrasting and reconstructing

Method

Feature extraction: f_θ

- Maps a single trajectory to an abstract representation of generating motion
- PointNet style
- 1D convolutions in temporal domain
- No spatial pooling

Subspace estimation: g_φ

- Time-dependent parameterized subspace basis (damped version of cosine) with learnable parameters

$$h^j_\psi(t) = e^{-(\alpha_j(t-\mu_j))^2}\cos(\beta_j t + \gamma_j)$$

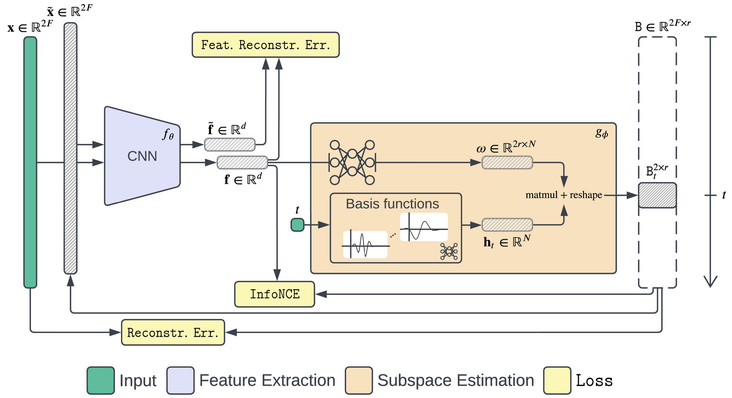
- MLP infers basis coefficients

Training and inference

- **Training:** requires fully observed data
- **Losses:** InfoNCE + reconstruction + reconstruction in feature space
- **Inference:** pass through f_θ + hierarchical clustering

Autoencoder-style architecture

Method



Trajectory completion algorithm

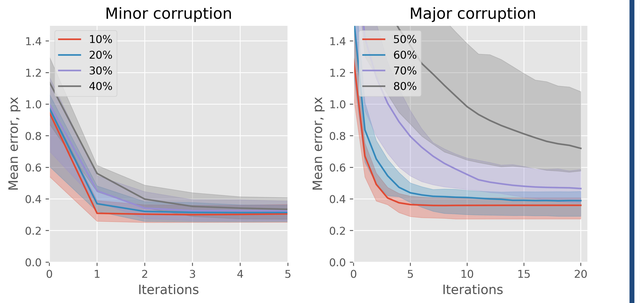
Method

- Objective to fill-in missing values: $\left\|\hat{\mathbf{x}}(\bar{\mathbf{x}}) - \mathrm{B}\mathrm{B}^\dagger\hat{\mathbf{x}}(\bar{\mathbf{x}})\right\|^2 \to \min_{\bar{\mathbf{x}}}$
- Linear solution for fixed subspace: $\bar{\mathbf{x}}^* = \mathrm{A}(\mathrm{B})\mathbf{x}$
- Yields approximate BCD approach,

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

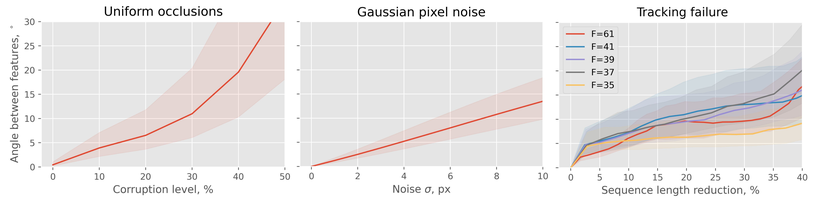
Evaluating completion

Results



Evaluating invariances of f_θ

Results



Clustering error on standard datasets

Results

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
ALC_{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

This work was supported by the Wallenberg Artificial Intelligence, Autonomous Systems and Software Program (WASP), funded by Knut and Alice Wallenberg Foundation