CIS centre for intelligent sensing

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Queen Mary University of London

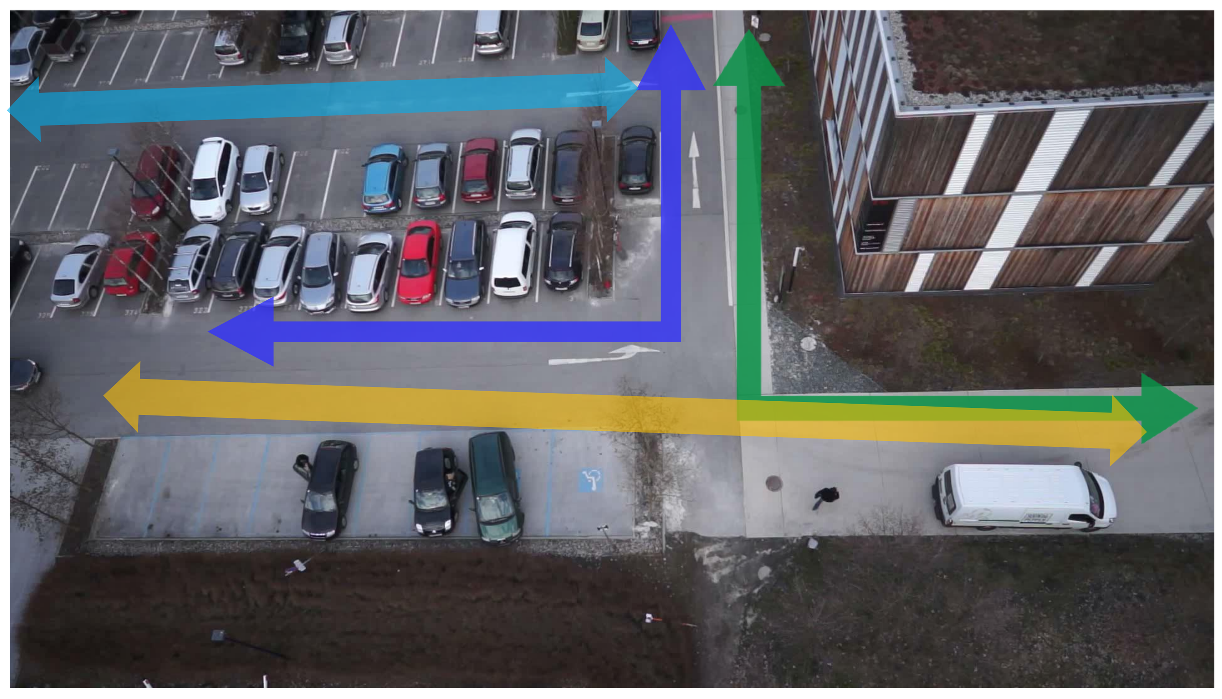
TRAJECTORY CLUSTERING FOR MOTION PATTERN EXTRACTION IN AERIAL VIDEOS

Tahir Nawaz, Andrea Cavallaro, Bernhard Rinner

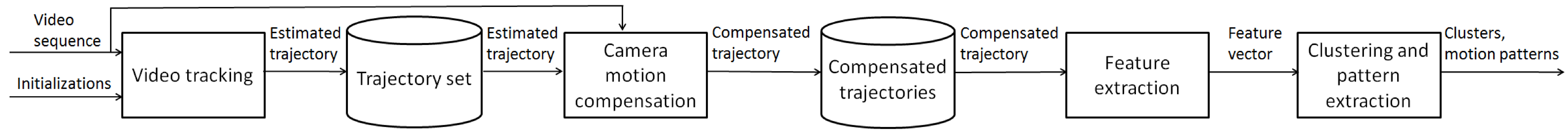
{tahir.nawaz,andrea.cavallaro}@eecs.qmul.ac.uk, bernhard.rinner@aau.at

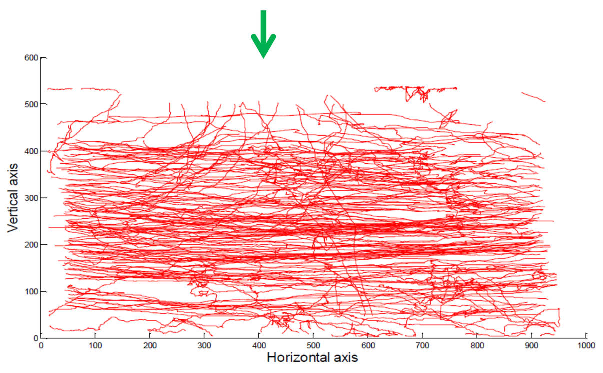
1. Introduction

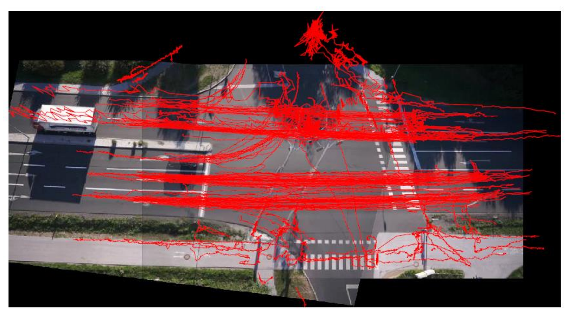
- Motion pattern: representative spatio-temporal trend of object motion in a scene
- Methods using motion information between frame pairs [1, 2]
  - Suitable for short-range patterns
  - Stationary camera setting
- Methods using motion information across multiple frames (object trajectories) [3, 4]
  - Suitable also for long-range patterns
  - Stationary camera setting
- Aerial videos: need to compensate camera motion in trajectories

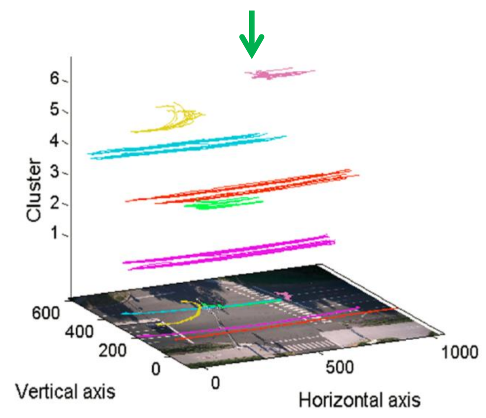


2. Proposed pipeline









Camera motion compensation

- Trajectory i: $\mathcal{X}_i = [X_i^k]_{k=k_i^S}^{k_i^E}$
- Estimated state at frame k: $X_i^k = [(x_i^k, y_i^k)]$
- Motion compensation using homography ($H_{k,r}$) between kth frame (I_k) and ref. frame (I_r):

$$[w\,\hat{x}_i^k, w\,\hat{y}_i^k, w]^T = H_{k,r}\,[x_i^k, y_i^k, 1]^T$$

Feature extraction

- Feature vector for trajectory i: $\mathbf{f}_i = (\mathbf{f}^{\hat{x}_i}, \mathbf{f}^{\hat{y}_i})$

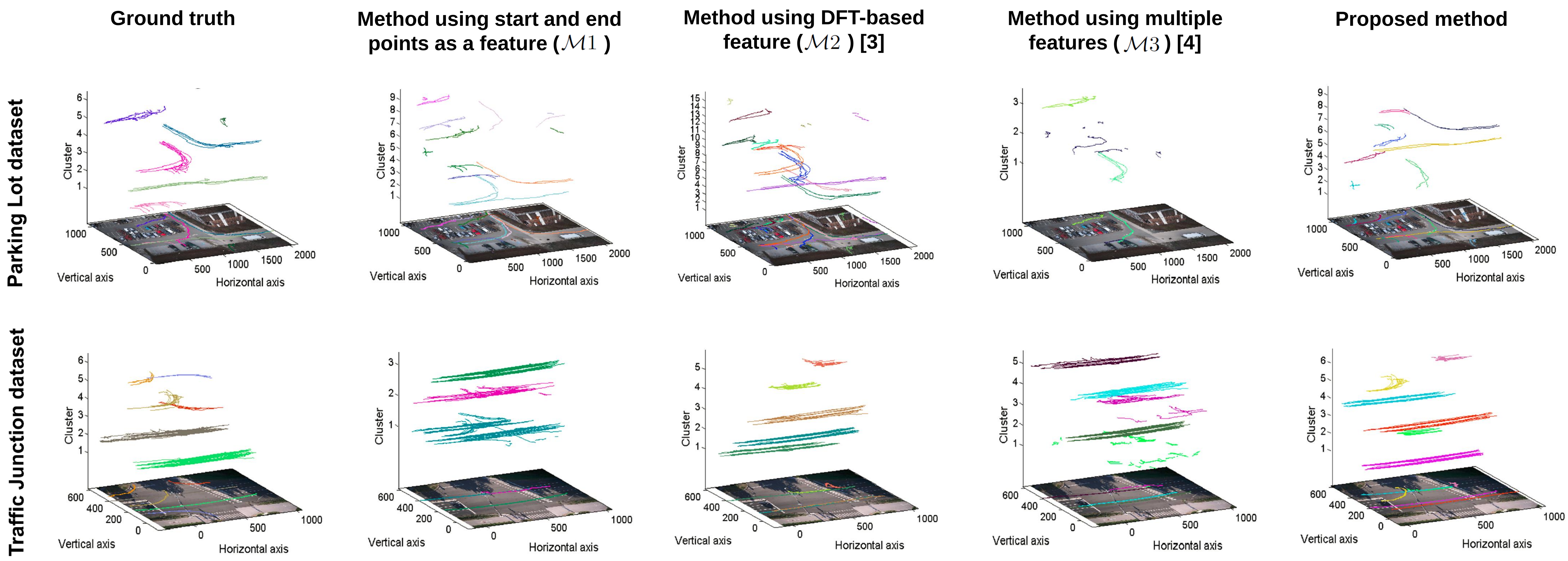
$$\mathbf{f}^{\hat{x}_i} = (\min(\mathcal{C}^{\hat{x}_i}), Q_{25}^{\hat{x}_i}, Q_{50}^{\hat{x}_i}, Q_{75}^{\hat{x}_i}, \max(\mathcal{C}^{\hat{x}_i}))$$
$$\mathbf{f}^{\hat{y}_i} = (\min(\mathcal{C}^{\hat{y}_i}), Q_{25}^{\hat{y}_i}, Q_{50}^{\hat{y}_i}, Q_{75}^{\hat{y}_i}, \max(\mathcal{C}^{\hat{y}_i}))$$

- $\mathcal{C}^{\hat{x}_i}$: set of DWT coefficients for x-coordinates of trajectory i;
- $Q_{25}^{\hat{x}_i}$: 25th percentile for $\mathcal{C}^{\hat{x}_i}$

Clustering and motion pattern extraction

- Mean-shift clustering with adaptive bandwidth selection [4]
- Trajectory that minimizes distance from the cluster centroid represents the motion pattern

3. Experimental results



Evaluation in terms of clustering accuracy (A) and precision (P) and recall (R) of the extracted motion patterns.

Method	Parking Lot			Traffic Junction			Students003			Train Station		
	A	P	R	A	P	R	A	P	R	A	P	R
$\mathcal{M}1$	.64	.48	.53	.67	.67	.27	.41	.90	.40	.32	.60	.18
$\mathcal{M}2$	.72	.41	1	.82	.46	.40	.78	.74	.43	.80	.46	.36
$\mathcal{M}3$	.56	.63	.33	.70	.40	.33	.51	.60	.28	.33	.35	.18
Proposed	.89	.65	1	.88	.52	.50	.90	.58	.51	.82	.45	.50

Effect of inducing tracking failures to p% randomly selected trajectories on clustering accuracy (A); p =0, 10, …, 50. μ: mean A; σ: standard deviation of A.

Method	Parking Lot $\mu\,(\sigma)$	Traffic Junction $\mu\,(\sigma)$	Students003 $\mu\,(\sigma)$	Train Station $\mu\,(\sigma)$
$\mathcal{M}1$	.45 (.18)	.53 (.13)	.29 (.10)	.23 (.07)
$\mathcal{M}2$	.57 (.13)	.66 (.19)	.66 (.21)	.60 (.19)
$\mathcal{M}3$	.41 (.20)	.54 (.15)	.37 (.13)	.20 (.09)
Proposed	.64 (.27)	.67 (.19)	.69 (.22)	.52 (.22)

References

[1] S. Wu, B. Moore, and M. Shah, *Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes*, in Proc. of IEEE CVPR, San Francisco, 2010.
[2] P.-M. Jodoin, Y. Benezeth, and Y. Wang, *Meta-tracking for video scene understanding*, in Proc. of IEEE AVSS, Krakow, 2013.
[3] W. Hu, X. Li, G. Tian, S. Maybank, and Z. Zhang, *An incremental DPMM-based method for trajectory clustering, modeling, and retrieval*, IEEE Trans. on PAMI, 35(5), 2013.
[4] N. Anjum and A. Cavallaro, *Multi-feature object trajectory clustering for video analysis*, IEEE Trans. on CSVT, 18(11), 2008.

Acknowledgement

This work was supported in part by the EACEA Agency of the European Commission under EMJD ICE FPA no 2010-0012 and by Lakeside Labs with funding from ERDF and KWF under grant KWF-20214/24272/36084.

Datasets available: http://uav.lakeside-labs.com/publications/test-data/data2/