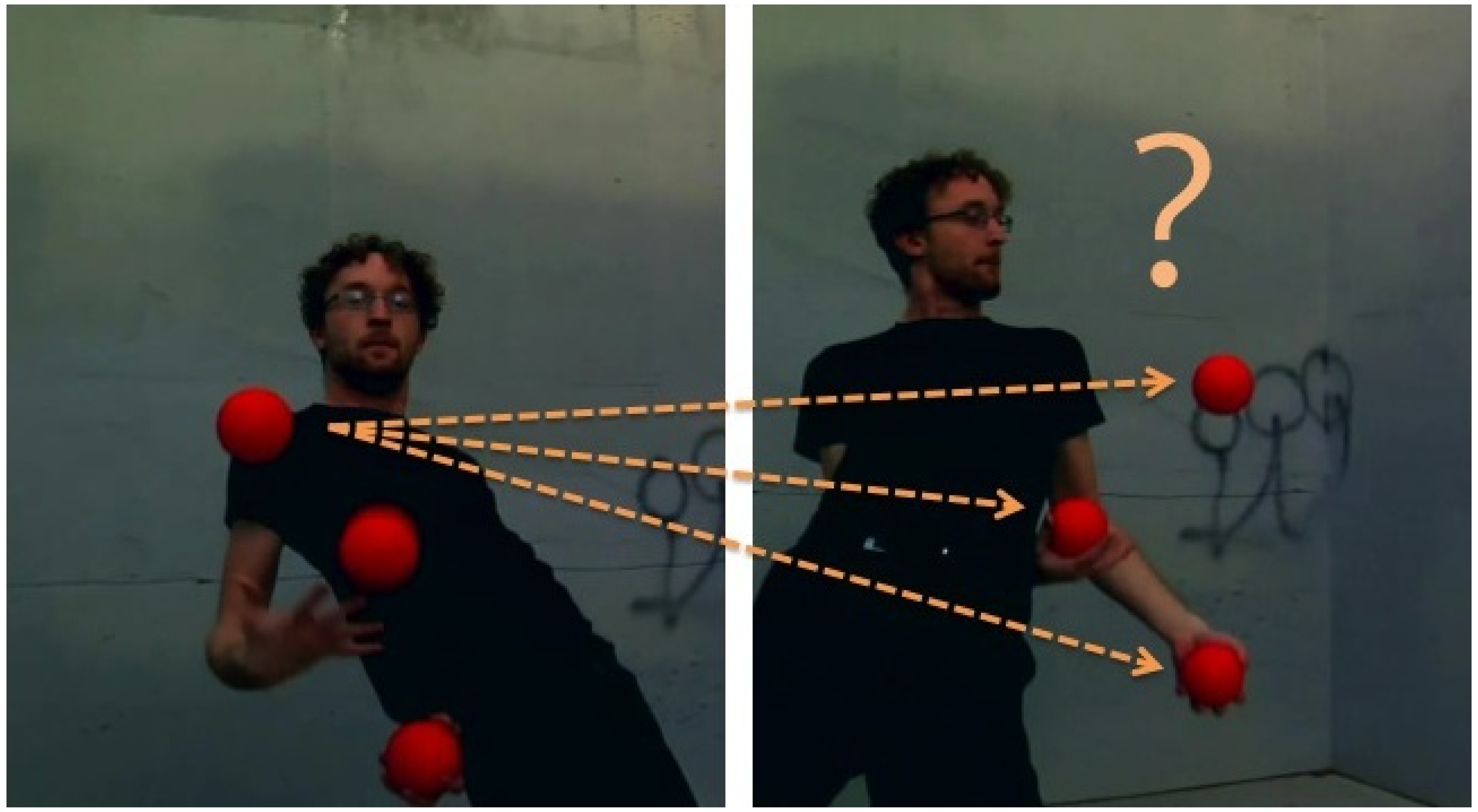
The Way They Move:
Tracking Multiple Targets with Similar Appearance

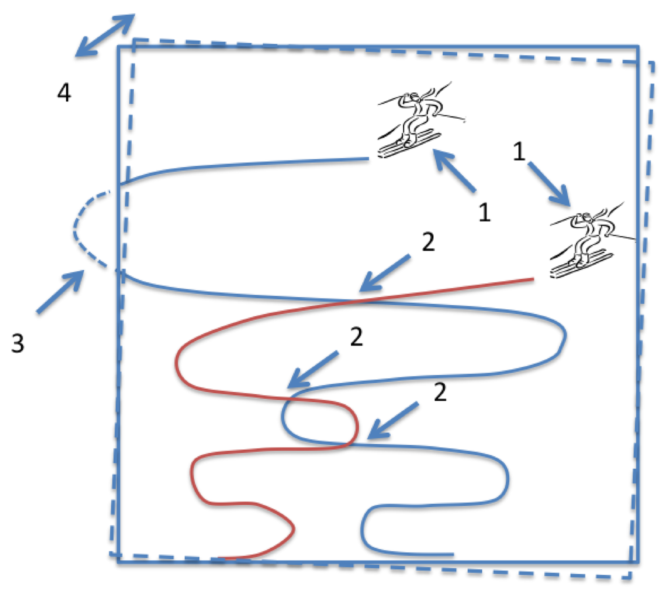
Caglayan Dicle, Octavia Camps, Mario Sznaier
Electrical & Computer Engineering, Northeastern University

Abstract

Motion is a powerful cue to distinguish objects: in many tracking scenarios it is possible to discriminate the targets from each other by only looking at their motion patterns. Yet, most state of the art approaches to multi target tracking rely heavily on appearance to associate detections from frame to frame and often overlook motion cues. In this work, we propose a multi-object tracking framework based on motion dynamics which is capable of tracking alike objects or objects with similar appearance and recover missing data due to long occlusions.



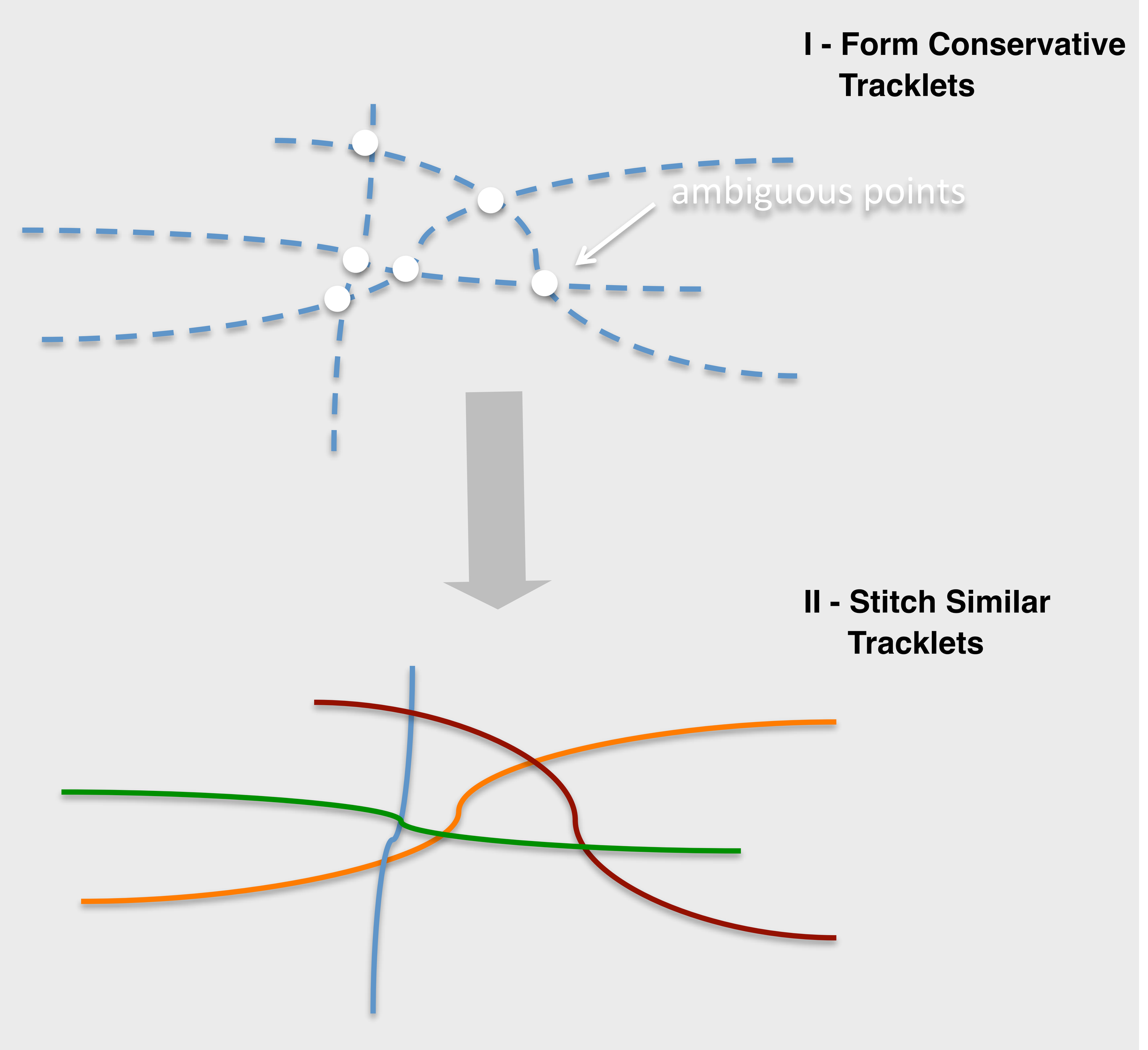
What do we solve?



Problems

1) Similar Objects
2) Crossings
3) Long/Difficult Occlusions
4) Camera Motion

Algorithm



Mechanics

Fact

Tracklets with same motion dynamics can be explained by a single regressor

$$\begin{aligned} y_k^i &= a_1 y_{k-1}^i + a_2 y_{k-2}^i + \ldots + a_5 y_{k-5}^i \\ y_k^j &= a_1 y_{k-1}^j + a_2 y_{k-2}^j + \ldots + a_5 y_{k-5}^j \\ y_k^{ij} &= a_1 y_{k-1}^{ij} + a_2 y_{k-2}^{ij} + \ldots + a_5 y_{k-5}^{ij} \end{aligned}$$

$$\begin{aligned} y_k^i &= a_1 y_{k-1}^i + a_2 y_{k-2}^i + \ldots + a_5 y_{k-5}^i \\ y_k^j &= b_1 y_{k-1}^j + b_2 y_{k-2}^j + \ldots + b_5 y_{k-5}^j \\ y_k^{ij} &= c_1 y_{k-1}^{ij} + c_2 y_{k-2}^{ij} + \ldots + c_7 y_{k-7}^{ij} + c_8 y_{k-8}^{ij} \end{aligned}$$

Remark

Regressor order = rank of Hankel Matrix

$$rank(H_{\alpha_i}) = 5 \quad rank(H_{\alpha_j}) = 5 \quad rank(H_{\alpha_{ij}}) = 5$$

$$rank(H_{\alpha_i}) = 5 \quad rank(H_{\alpha_j}) = 5 \quad rank(H_{\alpha_{ij}}) = 8$$

A Hankel Matrix

$$H_\alpha^{(m)} \doteq \begin{bmatrix} y_s & y_{s+1} & \cdots & y_{s+m-1} \\ y_{s+1} & y_{s+2} & \cdots & y_{s+m} \\ \vdots & \vdots & \vdots & \vdots \\ y_{e-m+1} & y_{e-m} & \cdots & y_e \end{bmatrix}$$

Similarity Measure

Ratio of ranks gives a similarity between tracklets

$$P_{ij} \doteq \begin{cases} -\infty & \text{if } \alpha_i \text{ and } \alpha_j \text{ conflict} \\ \frac{rank(H_{\alpha_i}) + rank(H_{\alpha_j})}{\min_{\beta_i^j} rank(H_{\alpha_{ij}})} - 1 & \text{otherwise} \end{cases}$$

Rank Estimation with Missing Data

Modified Hankel Total Least Squares (HTLS) can handle missing data

$$\begin{aligned} \min \quad & \left\| \Omega \circ [E \mid f] \right\|_F \\ st. \quad & (A+E)x = b+f \\ & [A \mid b], [E \mid f] \in S_H \end{aligned}$$

Sink/Source Free Tracklet Association

Generalized Linear Assignment (GLA) removes sink/source requirement

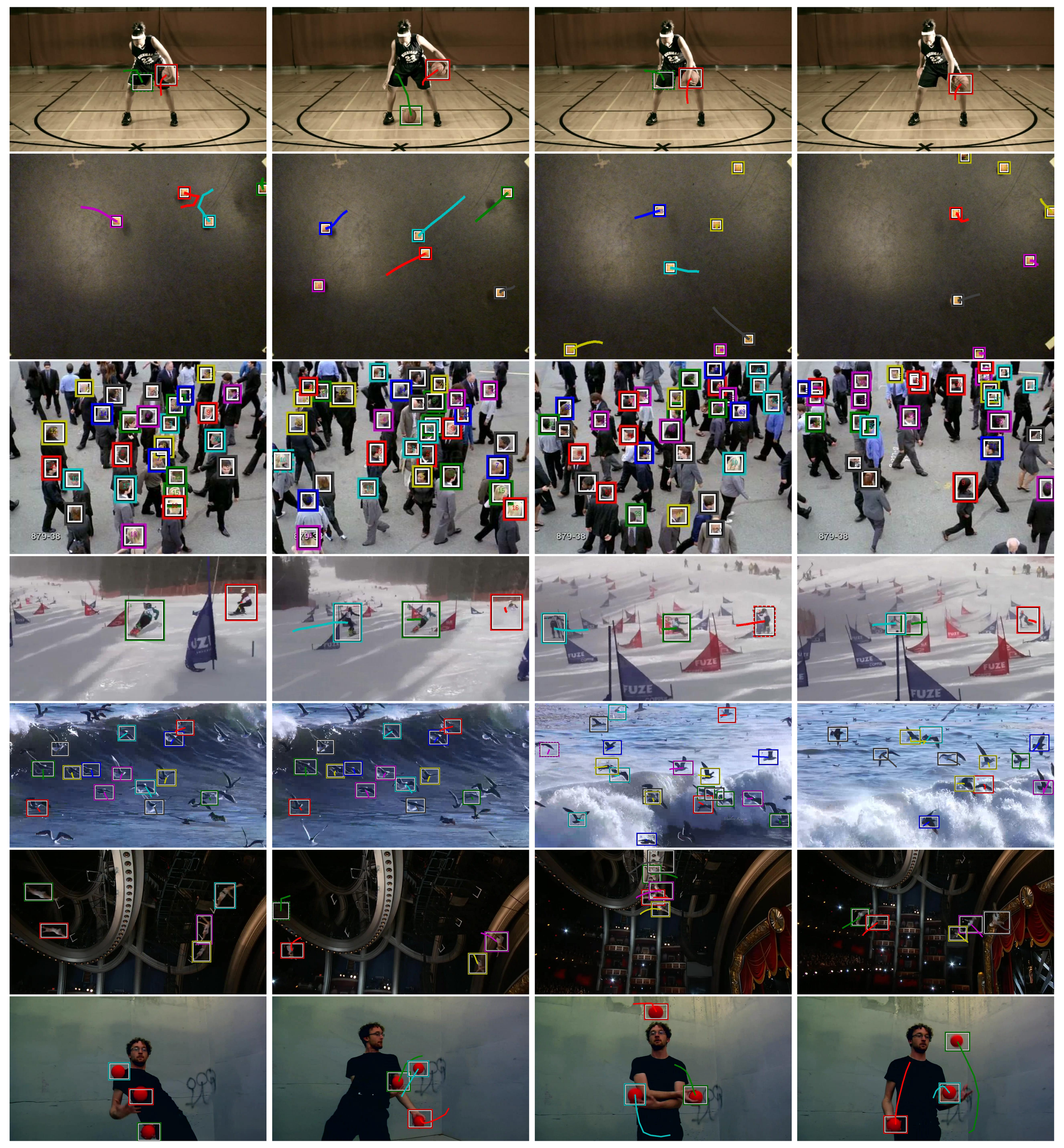
$$\begin{aligned} & \max_X \sum_{i=1}^{N} \sum_{j=1}^{N} P_{ij} X_{ij} \\ & st. \sum_{i=1}^{N} X_{ij} \leq 1 \, ; \sum_{j=1}^{N} X_{ij} \leq 1 \, ; X_{ij} \in \{0,1\} \end{aligned}$$

Results

MOTA Performance for The Proposed Method

Crossings (C), Occlusion (O), Camera Motion(M)

Video Name	SAT (ours)	KSP	MDA
Dribbling (C)	0.992 (0)	0.877 (2)	–
Balls (C)	0.997 (3)	0.572 (6)	–
Crowd (CM)	0.998 (40)	0.870 (1215)	–
Slalom (COM)	0.999 (2)	0.975 (13)	–
Seagulls (COM)	0.993 (23)	0.925 (305)	–
Acrobats (COM)	0.997 (1)	0.957 (12)	–
Juggling (COM)	0.977 (2)	0.422 (15)	–
PSU-sparse (–)	0.9642		1.00
PSU-dense (–)	0.9218		0.87



References

[1]. J. Berclaz, F. Fleuret, E. Turetken, and P. Fua. Multiple object tracking using k-shortest paths optimization. *IEEE Transactions on PAMI*, 33(9):1806–1819, 2011.

[2]. T. Ding, M. Sznaier, and O. Camps. Fast track matching and event detection. In *CVPR*, pages 1–8. IEEE, 2008

[3]. C. Dicle, O. Camps, and M. Sznaier. The way they move: Tracking Multiple Targets with Similar Appearance, to be submitted to ICCV 2013.

This material is based upon work supported by the U.S. Department of Homeland Security under Award Number 2008-ST-061-ED0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied of the U.S. Department of Homeland Security.