Multiobject Tracking as Maximum Weight Independent Set

William Brendel

Mohamed Amer

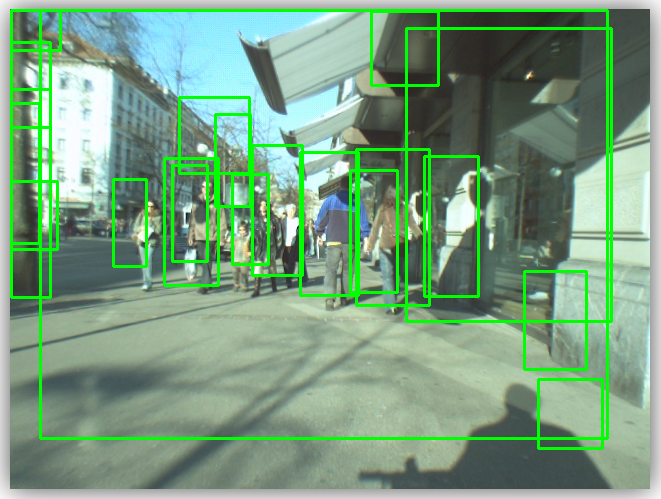
Sinisa Todorovic

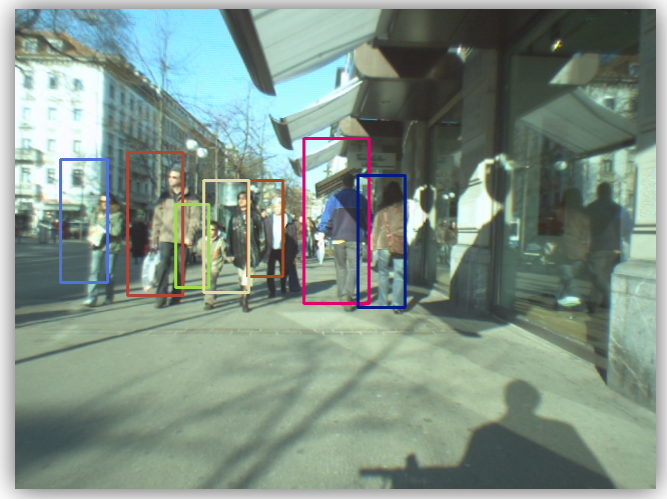
Oregon State University

Our Goal

- Multiple targets from distinct classes
- Occlusion
- Dynamic cluttered background
- Camera motion

Setting





Known	Unknown
Appearance of object classes	• Number of targets • Interactions • Motions • Spatial configuration

Problem Statement

Given noisy object detections:

- Identify the total number of unique tracks,
- Associate right detections with the tracks,
- Maintain unique IDs of the tracks

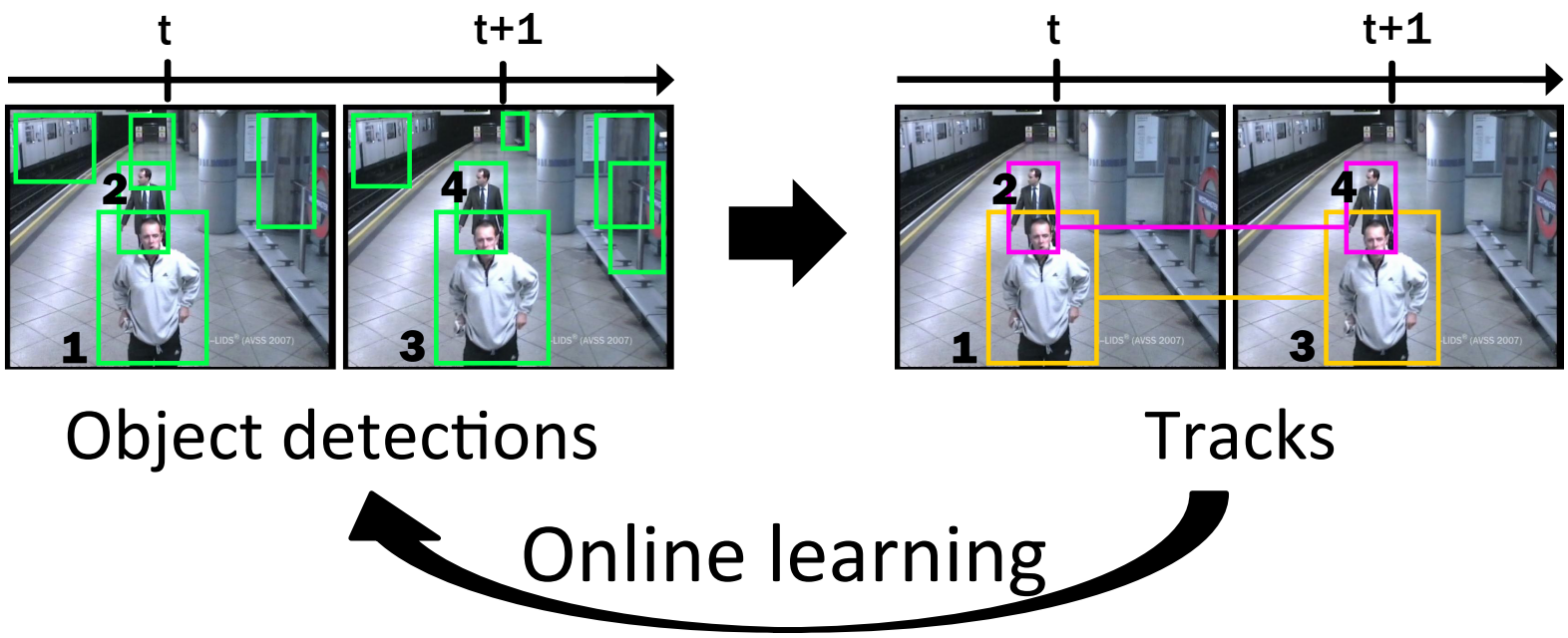
Target Properties

Soft constraints	Hard constraints
Appearance, motion	2 objects cannot be at the same location
Spatiotemporal interactions	An object cannot be at 2 distinct locations

Overview

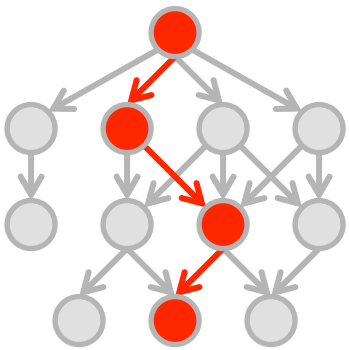
Given object detections in frames, we conduct:

1. Data association
2. Online learning of data affinities

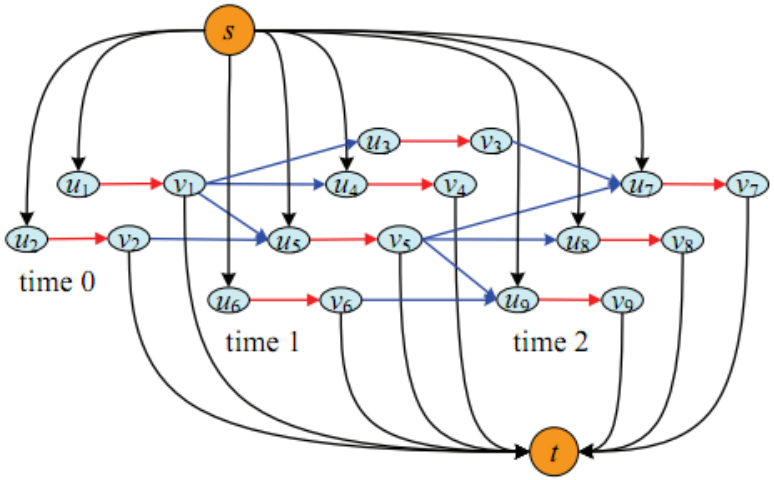


Prior Work – Data Association

- Nevatia et al. 2010, 2009,…
- Soatto et al. 2010
- Xing et al. 2009
- van Gool et al. 2009
- Andriluka et al. 2008
- Shafique et al. 2008
- Pavan & Pelillo 2007
- Ren & Malik 2007
- Perera et al. 2006
- Reid 79
- Bar-Shalom & Tse 75
- …



Bayesian network – ITST 09



Network flow – CVPR 08

Data Association as Integer Quadratic Optimization

- Accounting for target interactions is necessary
- However, IQP is intractable
- Existing relaxation methods:
  - Linear constraints
  - Convex quadratic

Poor guarantees of tightness to the original

Contributions – Data Association

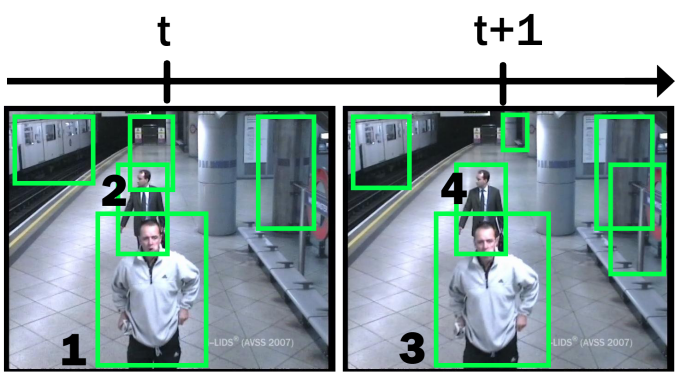
Our relaxation of the NP-hard problem:

- Provides strong guarantees of tightness
- Preserves all hard constraints

Maximum Weight Independent Set

Our Approach

Object detection at each frame

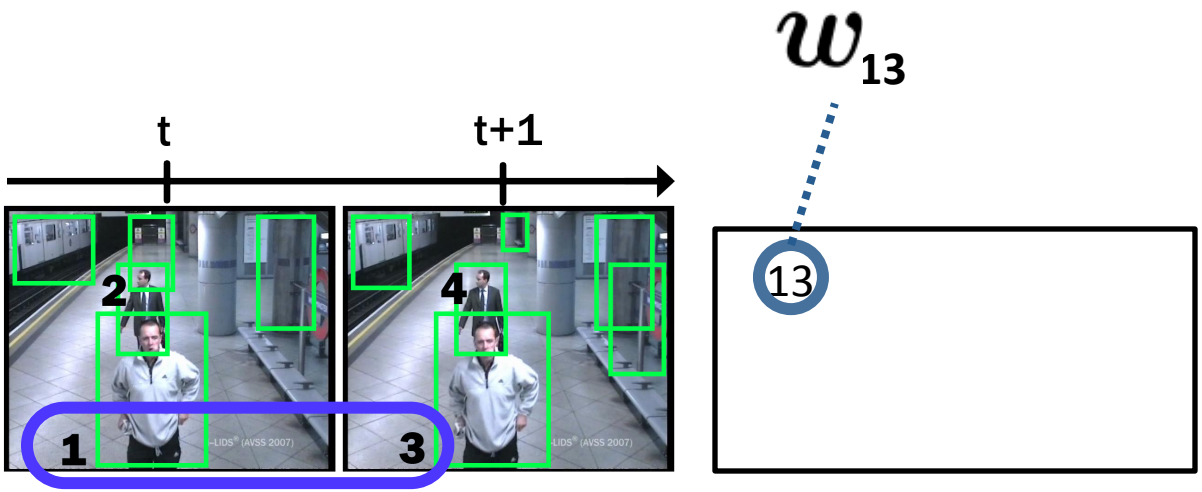


Detection

Our Approach

Nodes = Pairs of detections

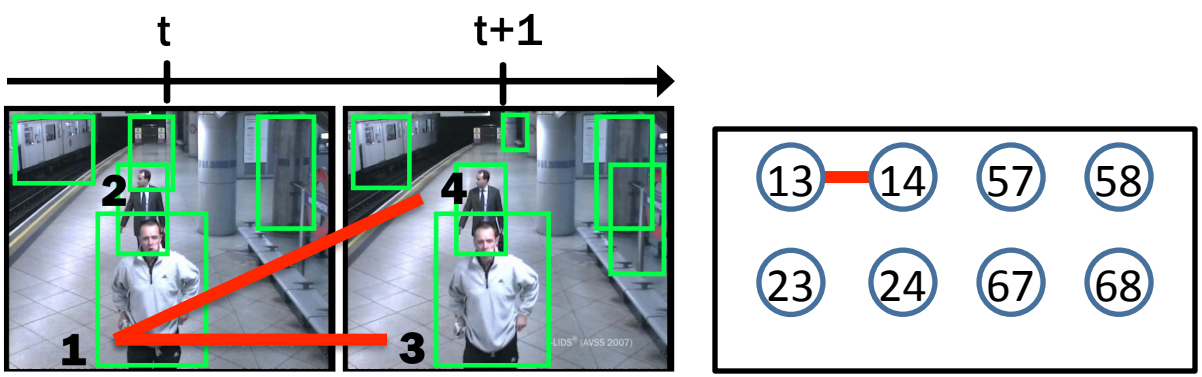
Weights of nodes = Similarity



Detection

Our Approach

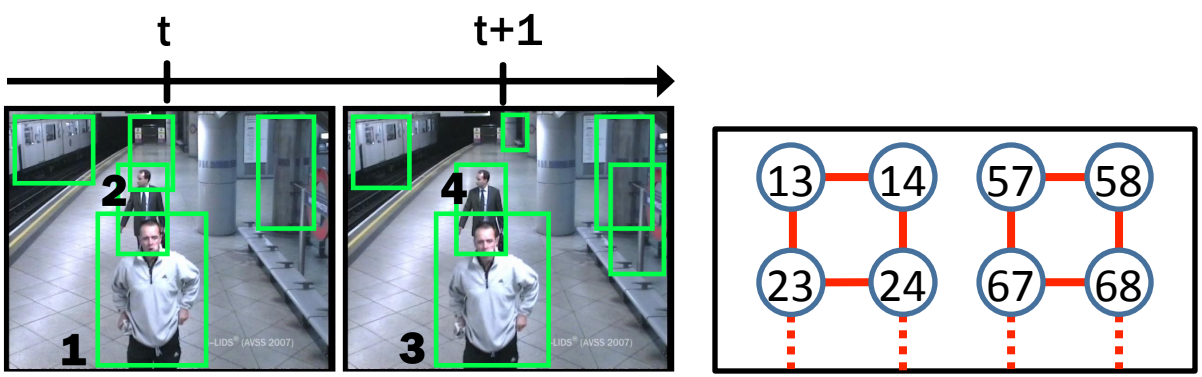
Edges = Hard constraints between detection pairs



Detection

Our Approach

Edges = Hard constraints between detection pairs



Detection

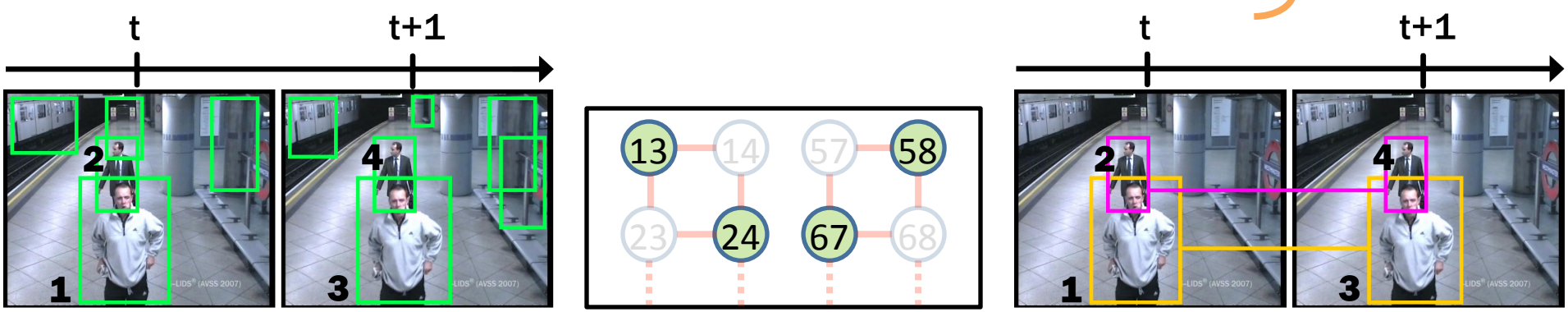
Our Approach

<u>Proposition:</u>

Tracking = Select pairs of detections that

- Do not violate the constraints
- Maximize their total similarity

MWIS



Detection → MWIS → Tracks

Maximum-Weight Independent Set

MWIS = **Heaviest** subset of **non-adjacent** nodes

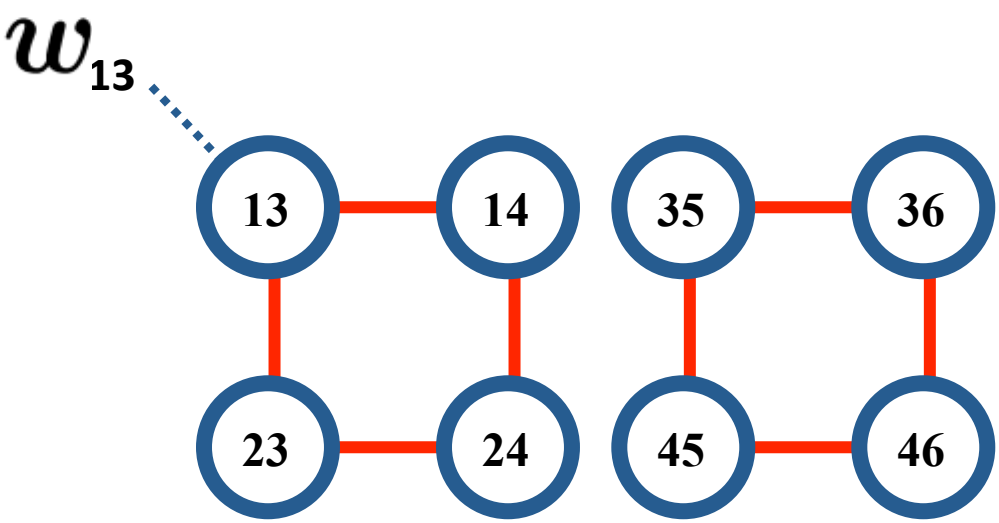
NP-hard, and hard to approximate

Sanghavi et al. 2007
Palubeckis 2006
Warrier et al. 2005
Trevisan 2004
Busygin et al. 2002
Bomze & Pelillo 2000
⋮

Linearization of the MWIS Problem

$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ = Total sum of weights in the MWIS

Similarity

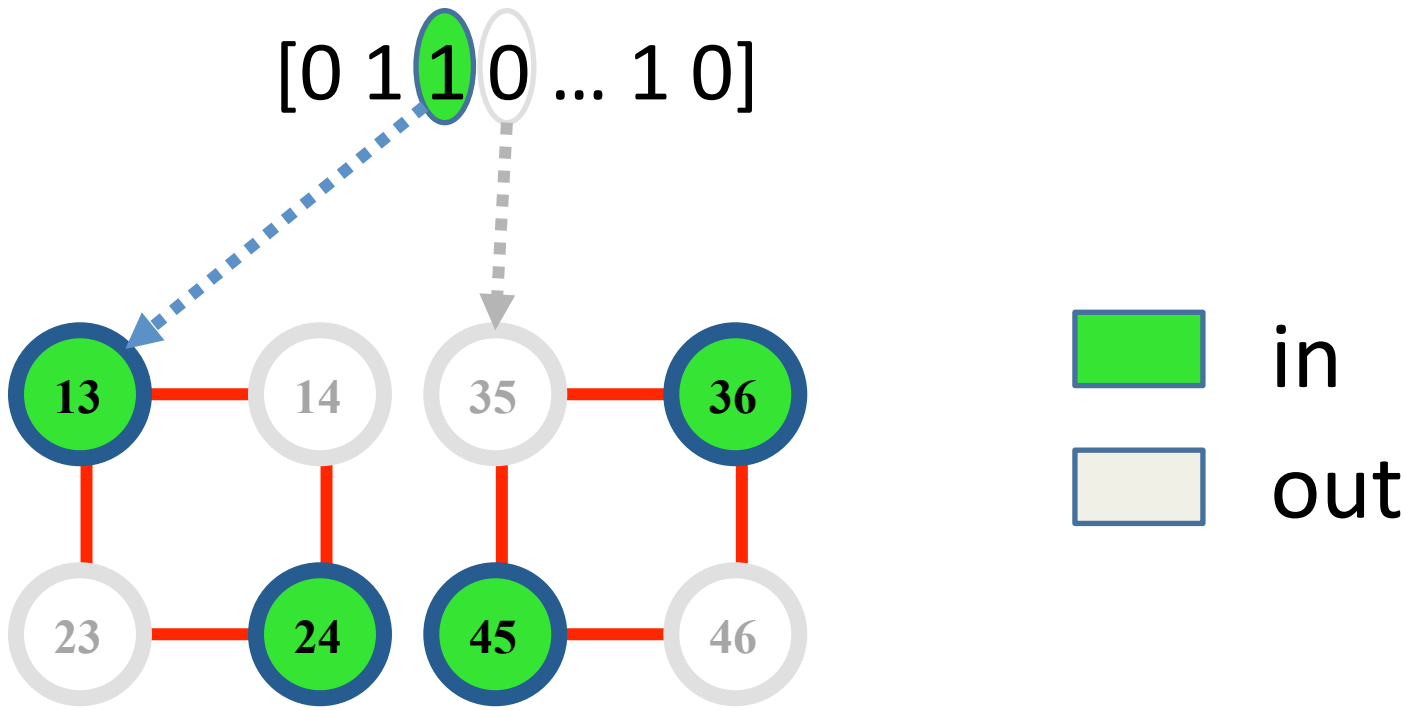


Linearization of the MWIS Problem

$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ = Total sum of weights in the MWIS

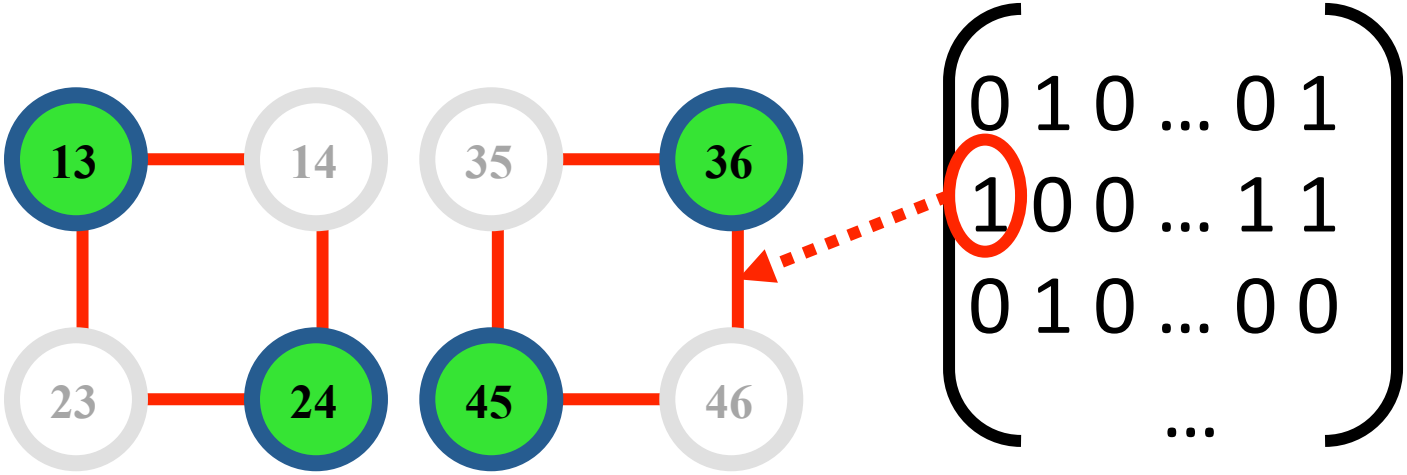
$\boldsymbol{x}$ = binary indicator vector

[0 1 1 0 … 1 0]



Encoding Hard Constraints

$\mathbf{B}$ = adjacency matrix



MWIS Formulation

$$\text{maximize} \quad \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x},$$

Find the heaviest subset

MWIS Formulation

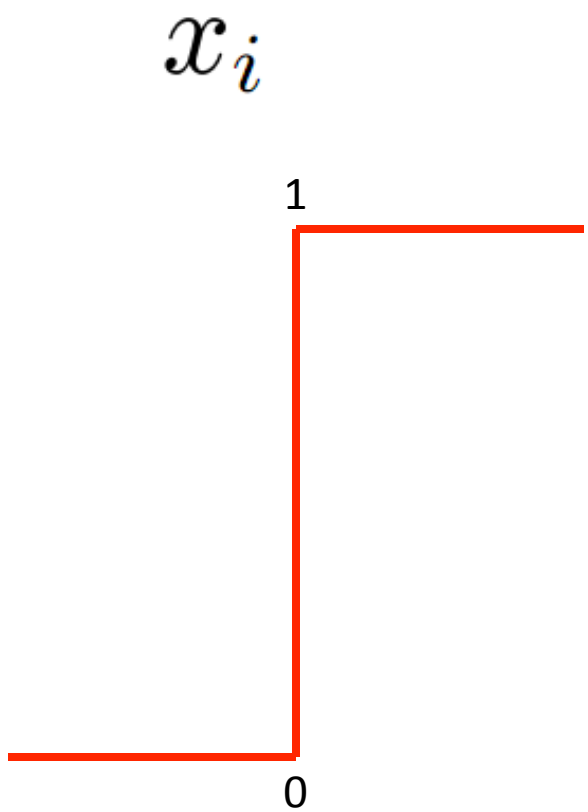
$$\text{maximize} \quad \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x},$$

$$\text{s.t.} \quad \boxed{\forall i, \quad x_i \in \{0, 1\}},$$

$$\boxed{\forall (i,\ j), \quad \mathbf{B}_{ij} \cdot x_i \cdot x_j = 0}$$

Constraints

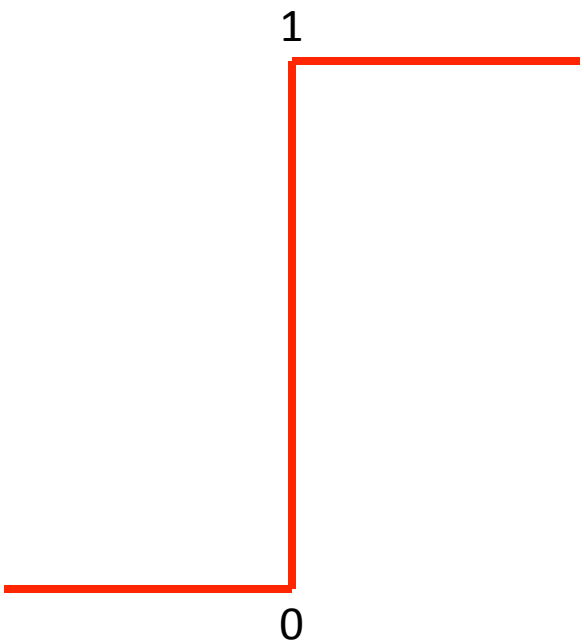
Continuous Relaxation

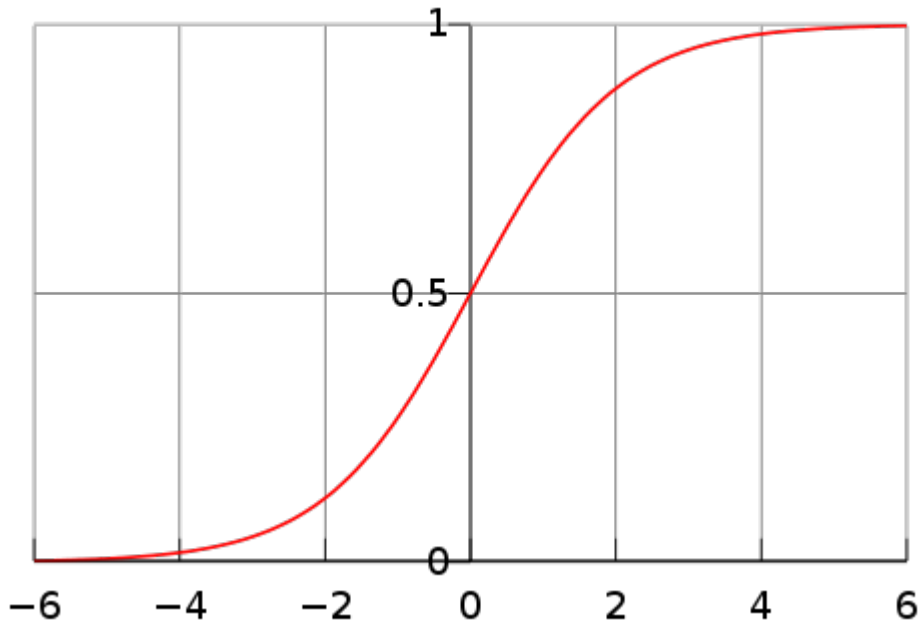


Step function

Continuous Relaxation

$$x_i = \sigma(y_i) = (1 + e^{-\beta y_i})^{-1}$$





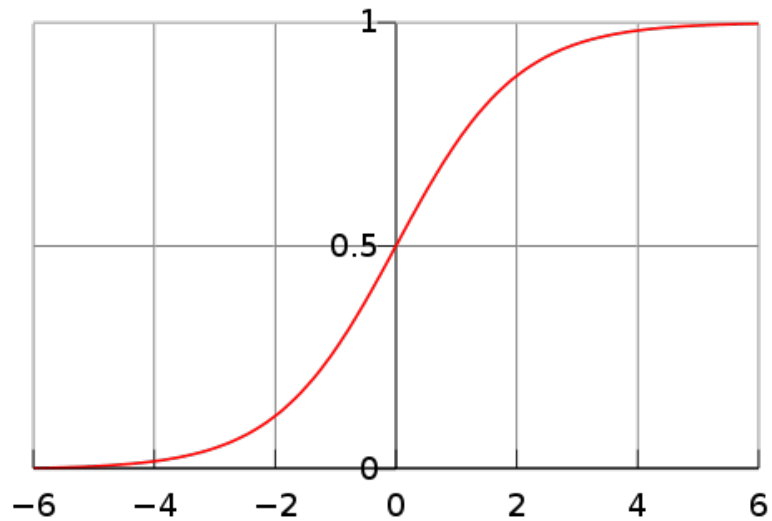
Step function

Sigmoid function

Continuous Relaxation

$$\boldsymbol{x} \in \{0,1\}^n \rightarrow \boldsymbol{y} \in \mathbb{R}^n$$

$$x_i \approx \sigma(y_i) = (1 + e^{-\beta y_i})^{-1}$$



Sigmoid function

Hard Constraints

$$\forall(i,\ j),\quad \mathbf{B}_{ij} \cdot x_i \cdot x_j = 0$$

Hard Constraints

$$\forall(i,\ j),\quad \mathbf{B}_{ij} \cdot x_i \cdot x_j = 0$$

Prior work: $x_i + x_j \leq 1$

E.g.

x_i = 0.5 x_j = 0.5 → **Violation of the original constraints**

Hard Constraints → Polynomial Form

$$\forall (i,\ j),\quad \mathbf{B}_{ij} \cdot x_i \cdot x_j = 0$$

$$\Downarrow$$

$$\forall i,\quad \boldsymbol{x}_i = 1 \ \Rightarrow\ \boldsymbol{x}_i \prod_j (1 - \boldsymbol{x}_j)^{\mathbf{B}_{ij}} = 1$$

Polynomial MWIS Formulation

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}, \\ \text{s.t.} \quad & \forall i, \quad x_i \in \{0,1\}, \\ & \forall (i,\ j), \quad \mathbf{B}_{ij} \cdot x_i \cdot x_j = 0 \end{aligned}$$

$$\boxed{\text{maximize} \sum_i w_i \sigma(y_i) \prod_j (1-\sigma(y_j))^{\mathbf{B}_{ij}}}$$

Our Algorithm

$$\boldsymbol{y}^{(\tau)} = \boldsymbol{y}^{(\tau-1)} + \Delta\tau\, \dot{\boldsymbol{y}}^{(\tau-1)}$$

$\Delta\tau$ → Iteration step

$\dot{\boldsymbol{y}}$ → First order dynamic

Theorem:
Converges to an optimum

Key Contribution

Theorem:

$$\max_{\substack{\boldsymbol{y} \in \mathbb{R}^n \\ \beta \to \infty}} \sum_i w_i \sigma(y_i) \prod_j (1 - \sigma(y_j))^{\mathbf{B}_{ij}}$$

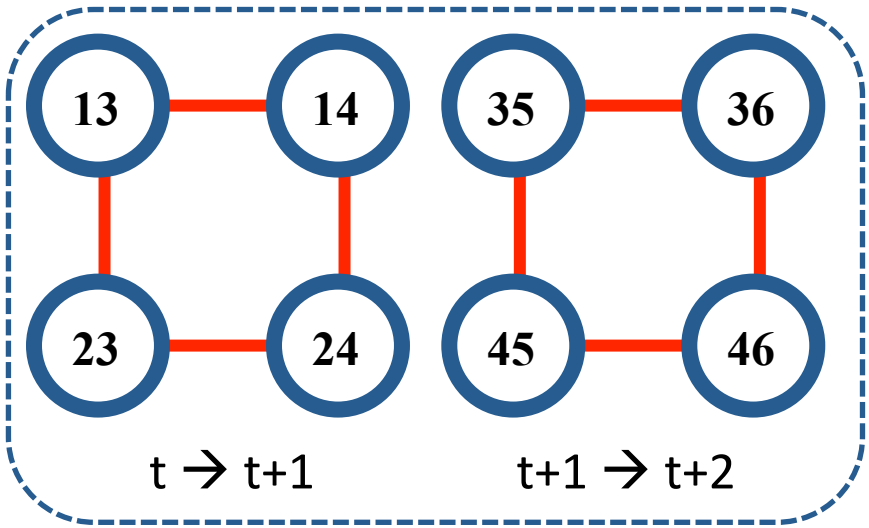
$$\parallel$$

Stability number of the graph

→ Guarantee of tightness

MWIS Property

MWIS of the entire graph

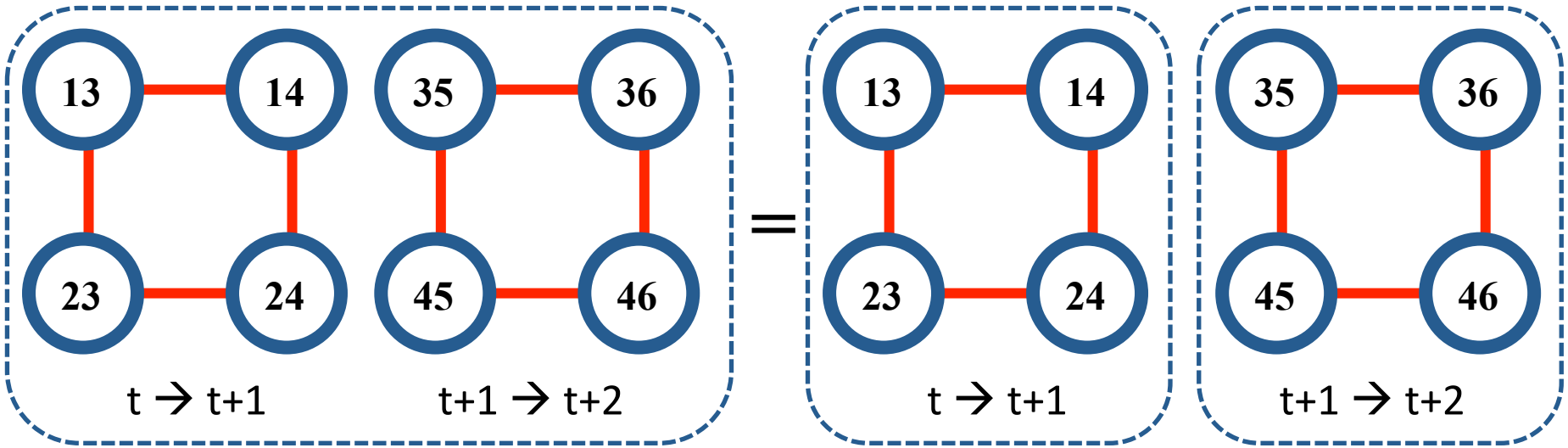


MWIS of t → t+2

MWIS Property

MWIS of the entire graph =

Set of the MWIS's of independent subgraphs

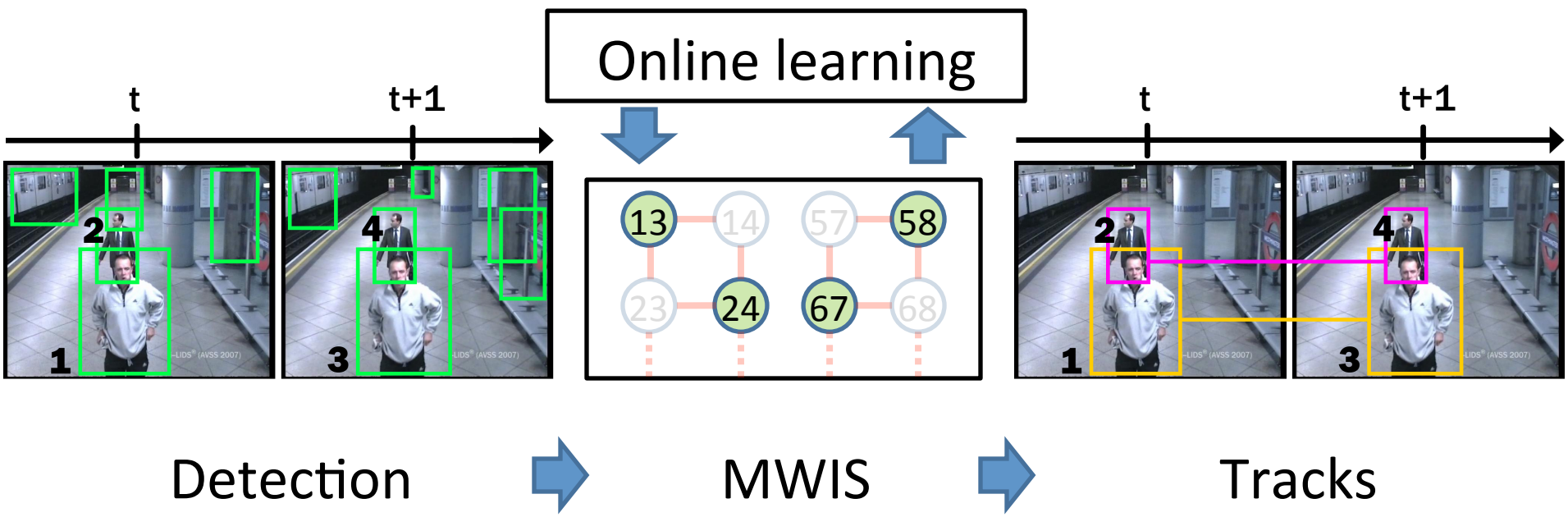


MWIS of t → t+2 MWIS of t → t+1 MWIS of t+1 → t+2

Our Approach

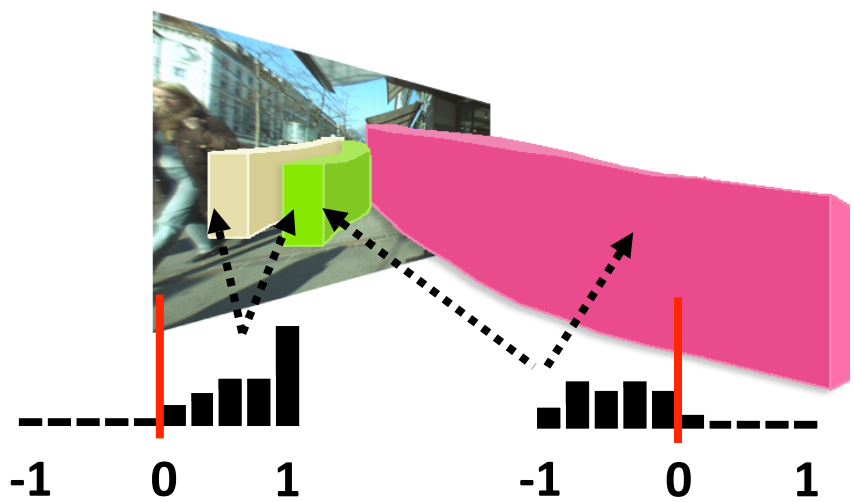
Online learning of:

- Appearance, motion
- Spatiotemporal interactions



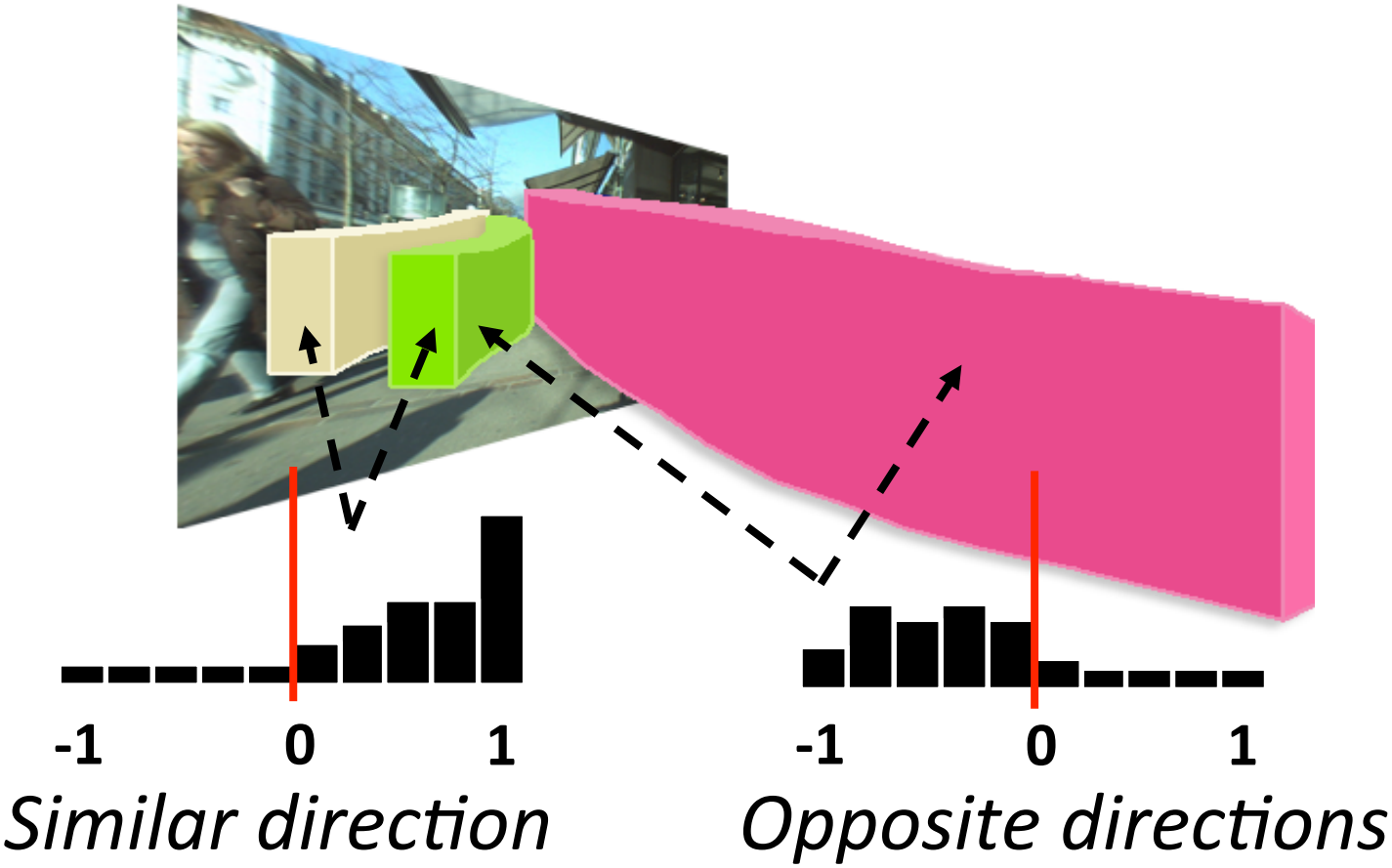
Target Interactions

- Account for ***all*** target interactions
- Estimate the distribution of motion correlations



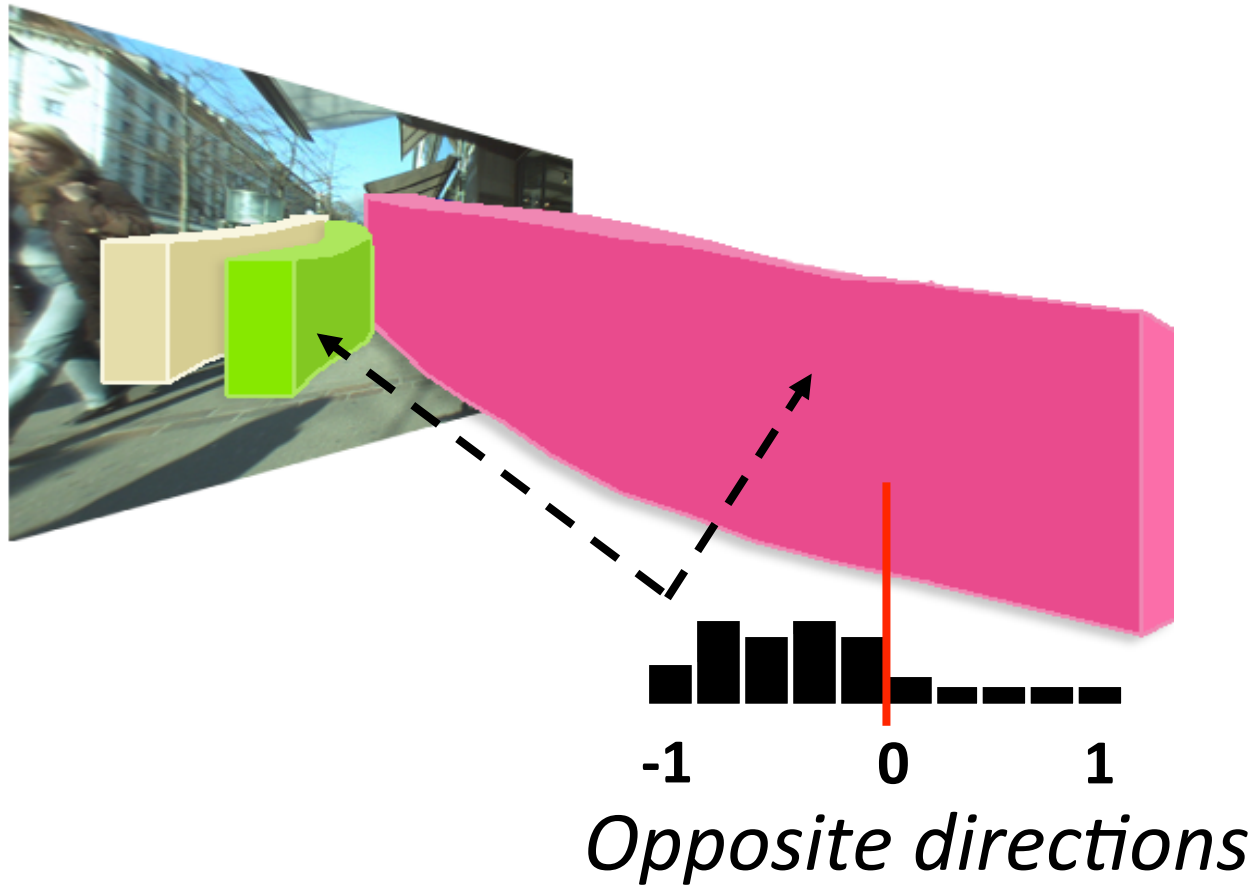
Similar direction *Opposite directions*

Learning Spatiotemporal Interactions



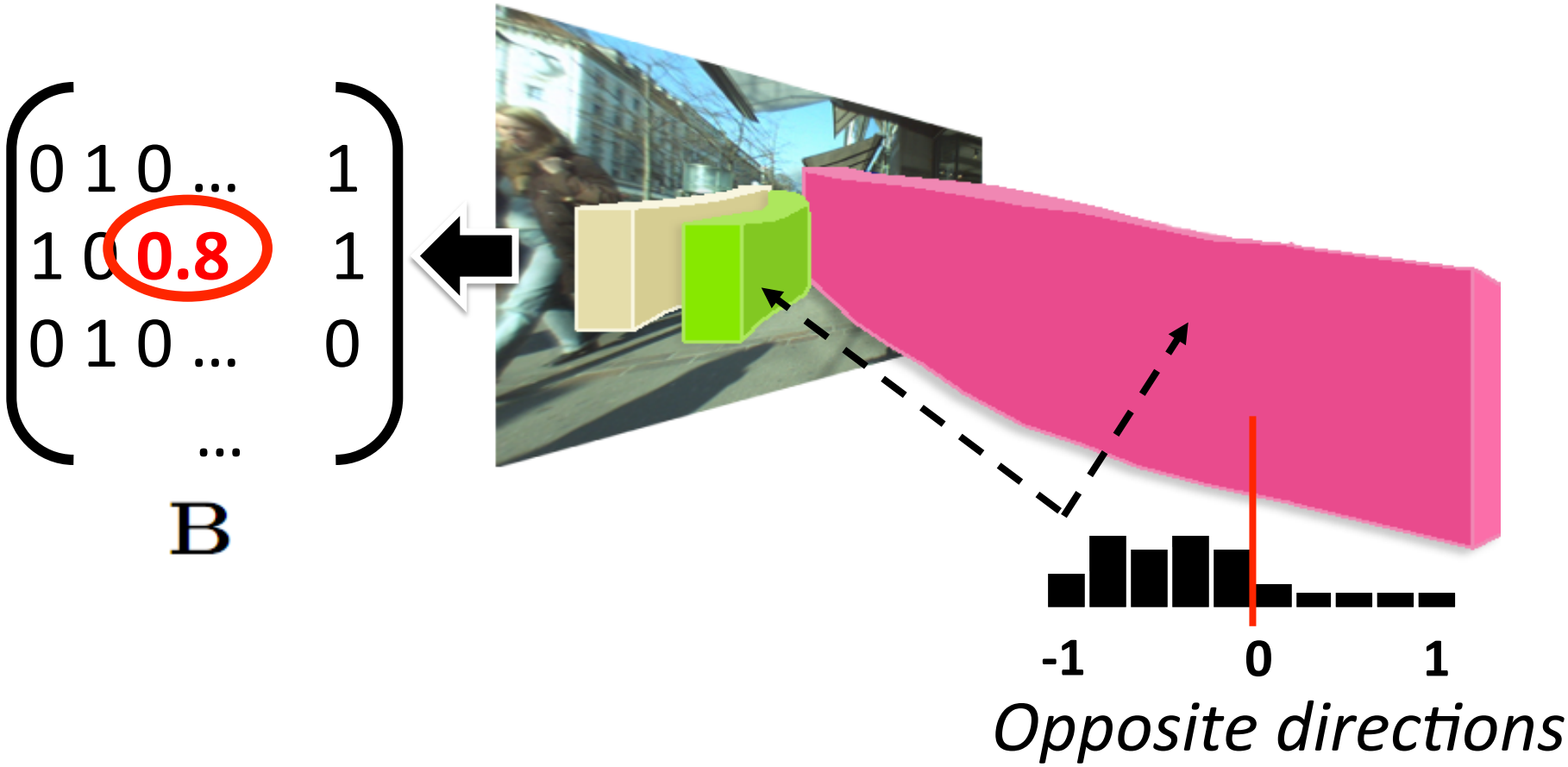
Distribution of the motion correlation between two targets

Learning Spatiotemporal Interactions



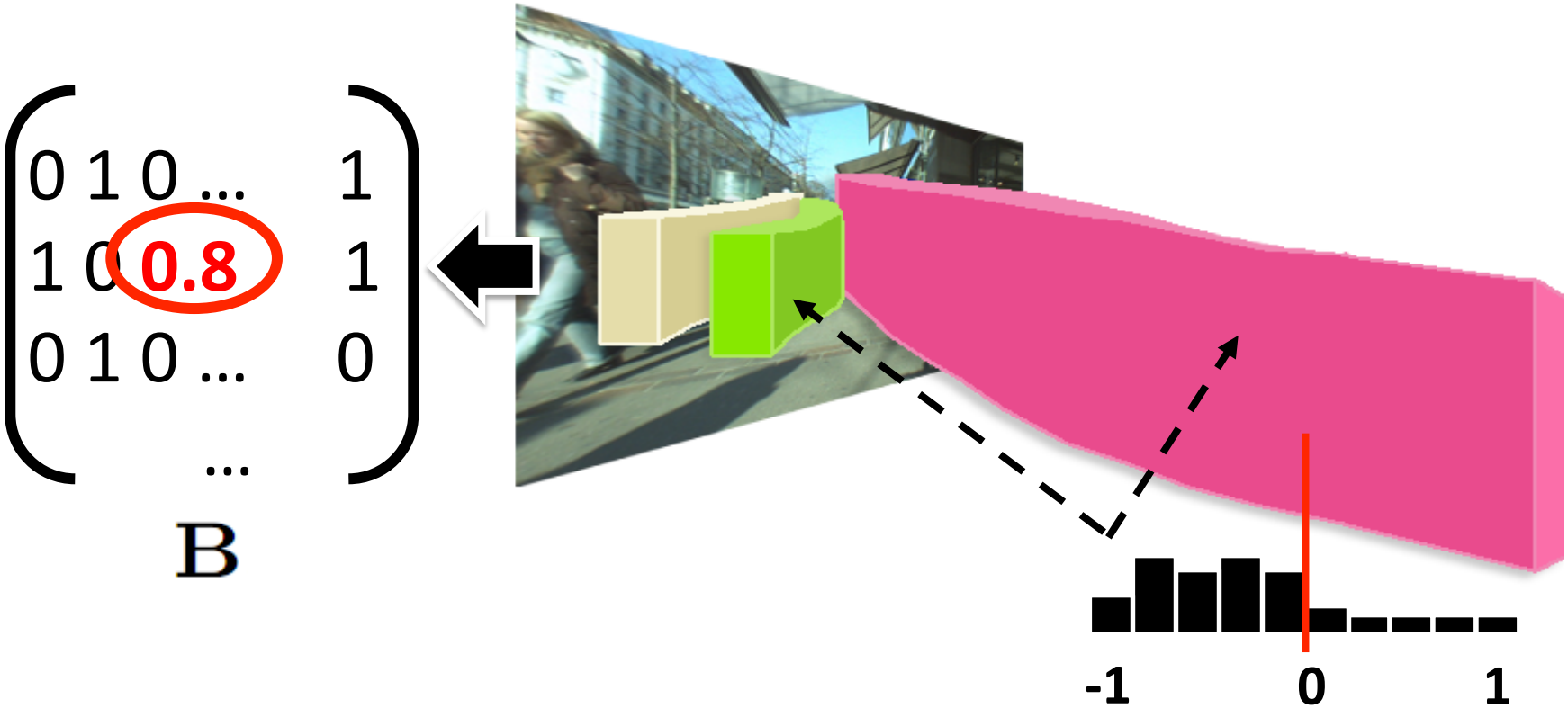
Distribution of the motion correlation between two targets

Learning Spatiotemporal Interactions



Distribution of the motion correlation between two targets

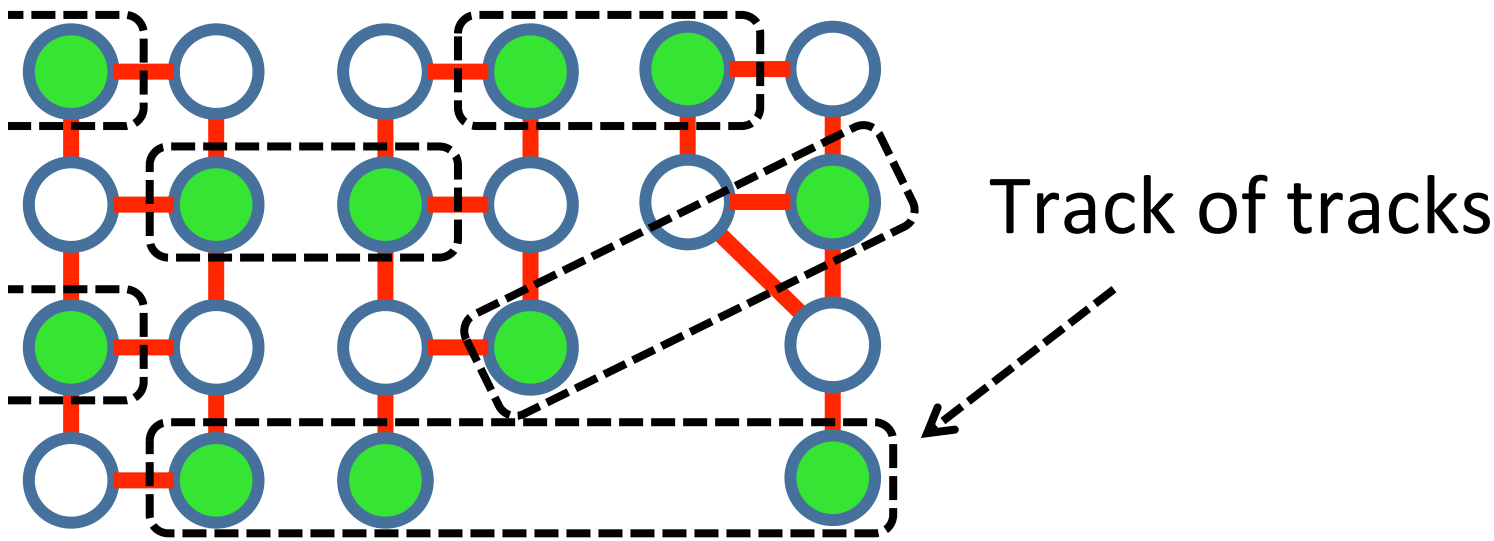
Learning Spatiotemporal Interactions



Distribution of motion correlations is incorporated in the adjacency matrix

Handling Gaps in Detection

Iteratively apply the MWIS



Pairs of tracks from previous MWIS

Correcting Errors in Tracking

Iteratively apply the MWIS algorithm to tracks obtained in the previous iteration

Handles:

- Occlusion
- Appearance/disappearance
- Missing detections

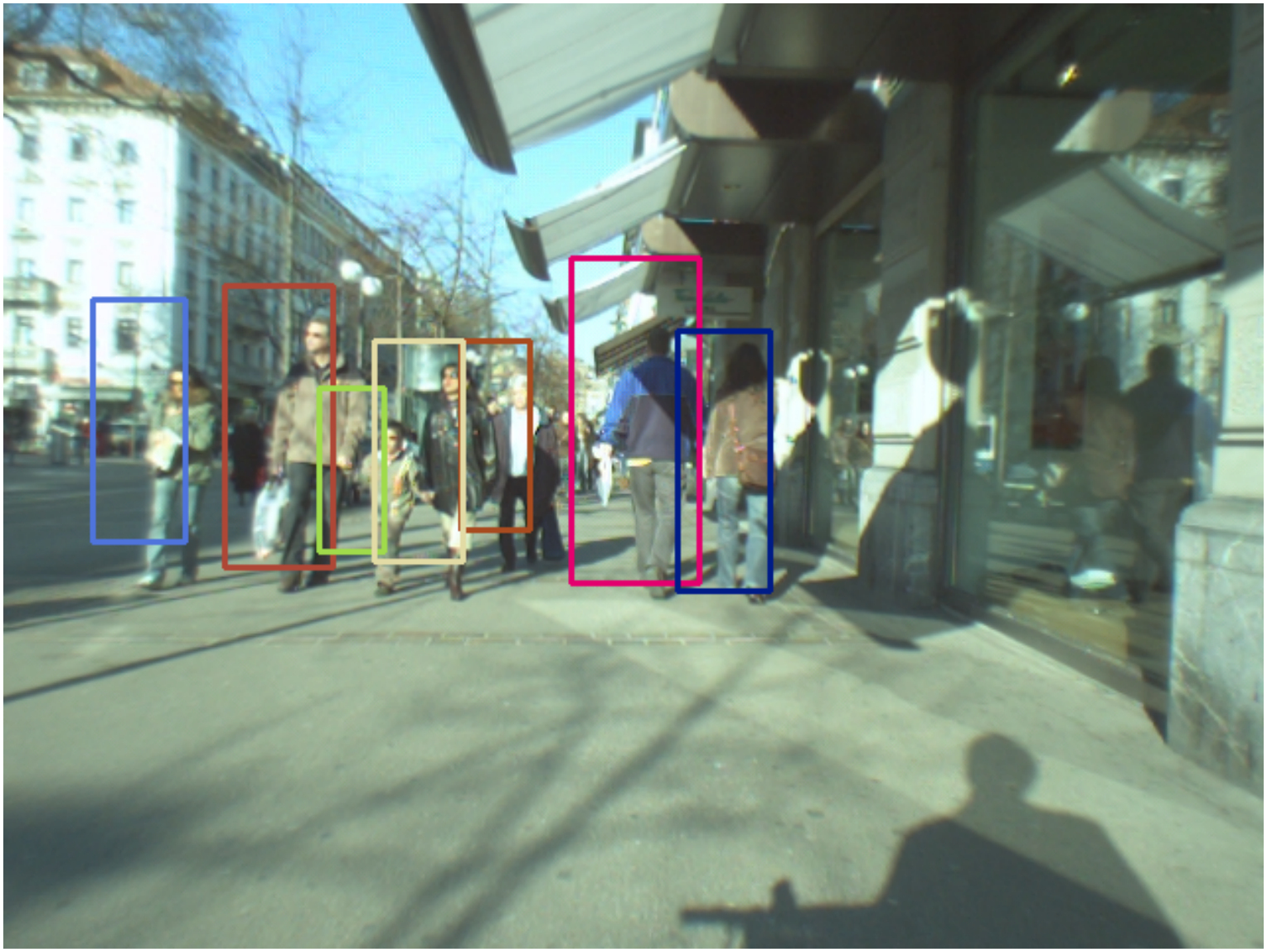
Datasets

- ETHZ datasets
- TUD Crossing
- i-Lids AB
- UBC Hockey
- OSU Campus

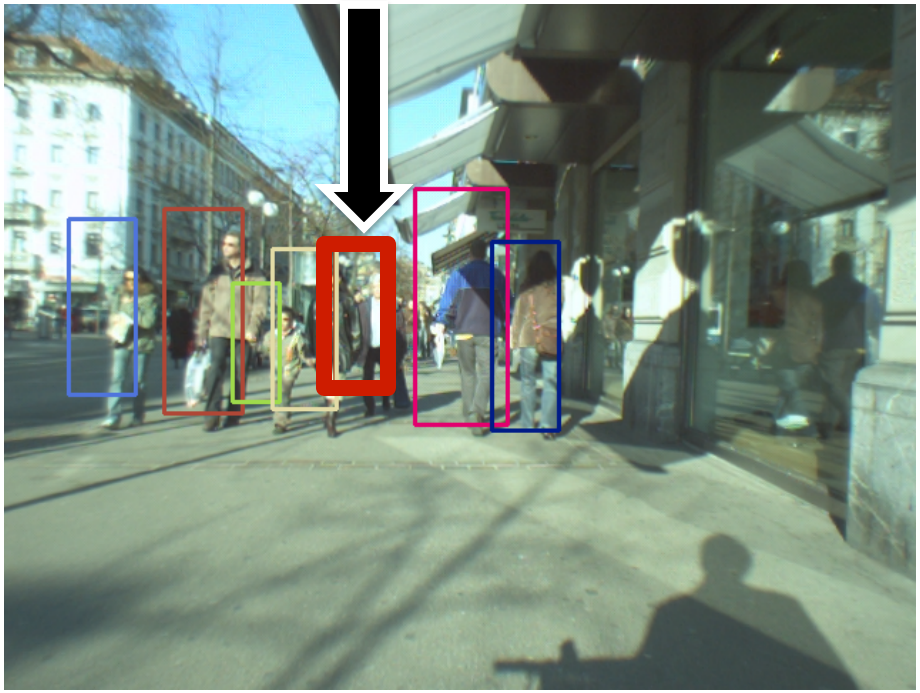
Evaluation Metrics: CLEAR MOT

- Precision
- Accuracy
- False Negatives
- False Positives
- ID Switches

ETHZ Moving Vehicle

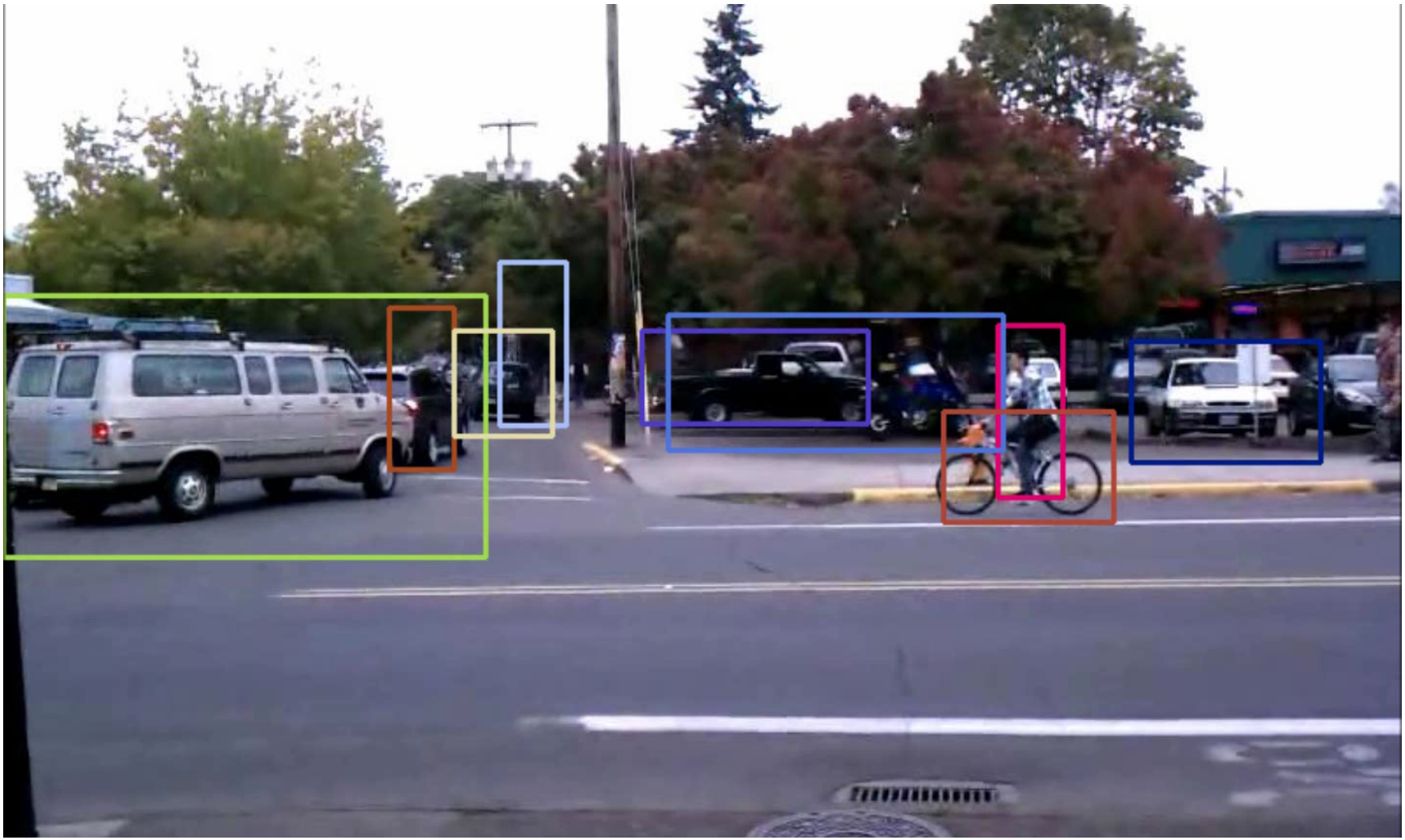


Qualitative – ETHZ Moving Vehicle

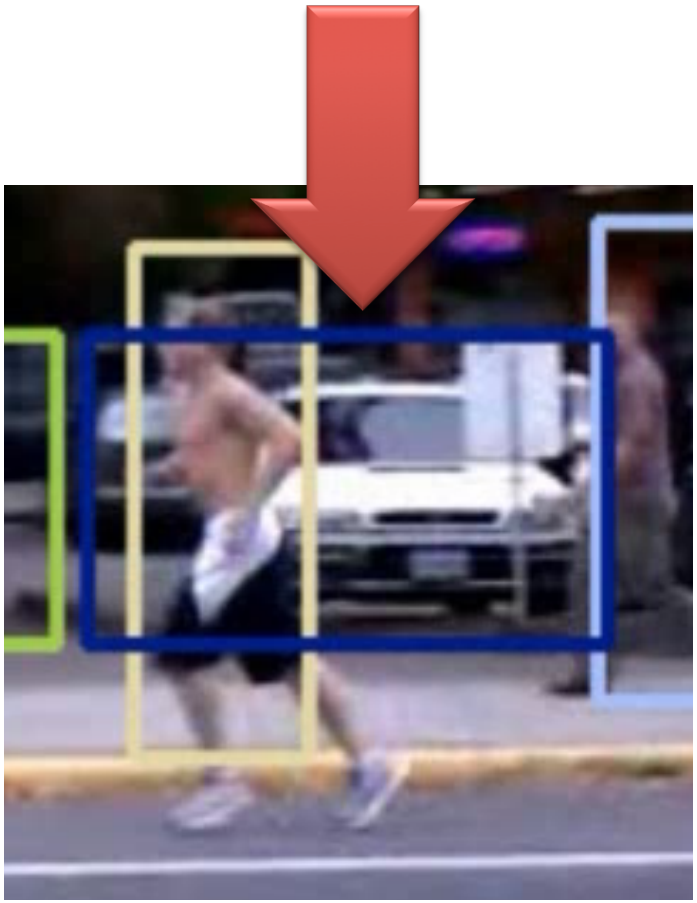


Handle change of scale
and camera motion

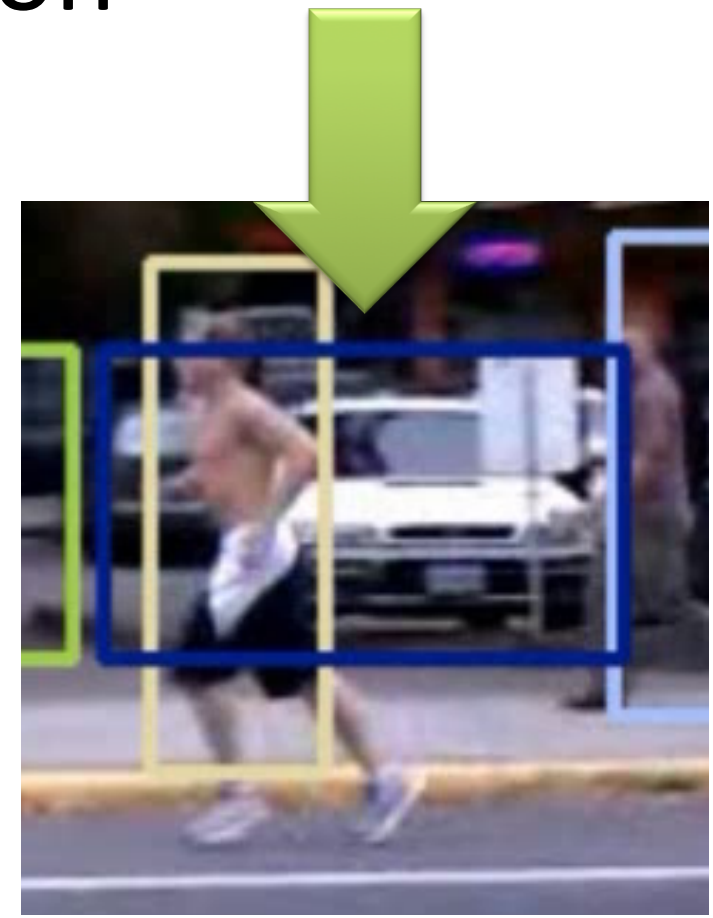
OSU Campus Dataset



Occlusion



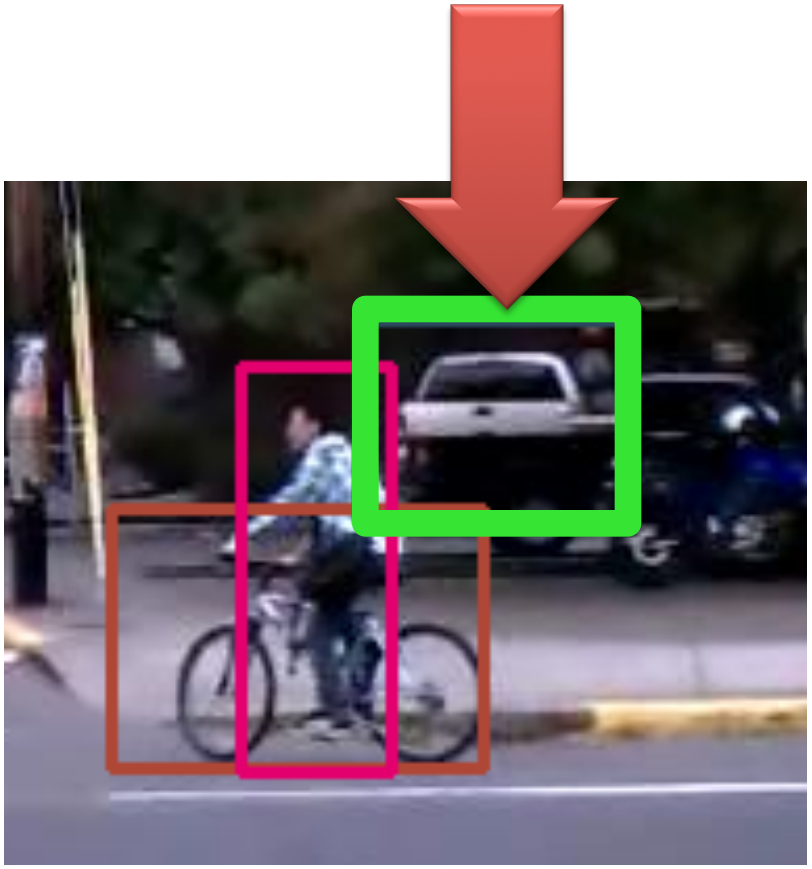
Without



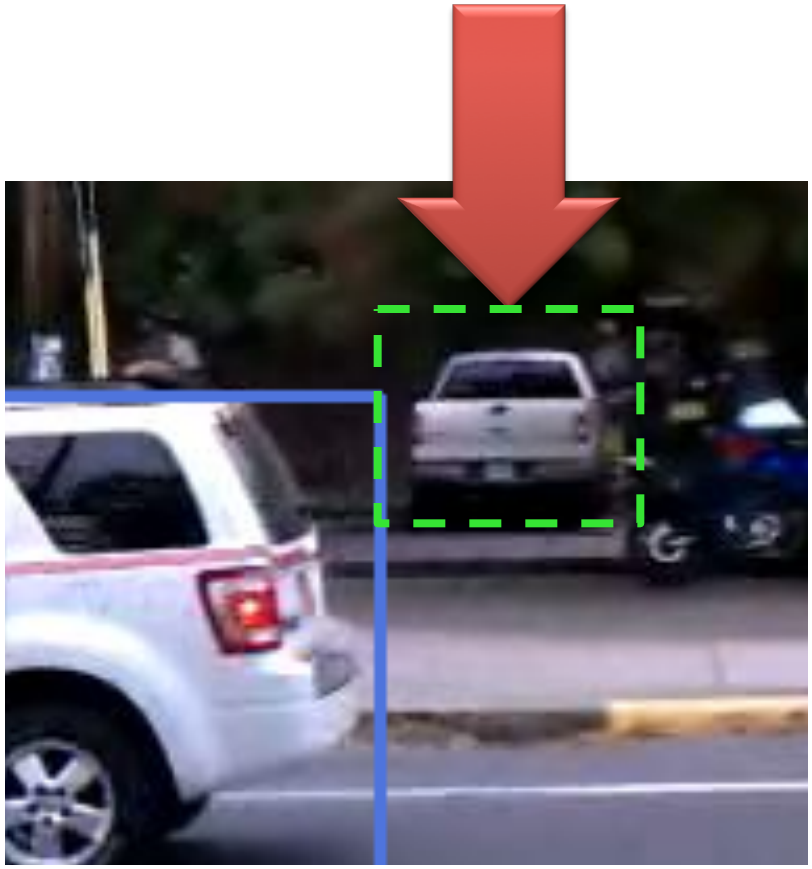
With

Iterative application of MWIS

Failure: No Detections



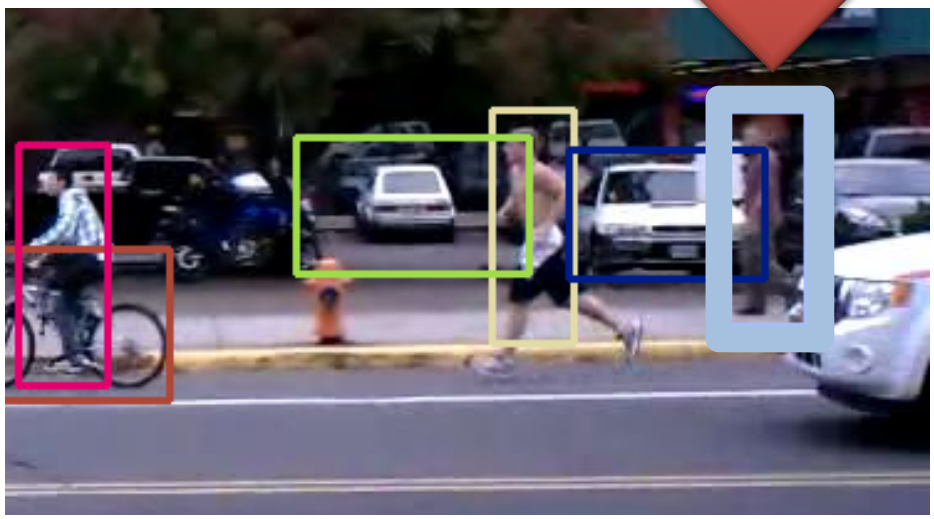
Frame 210

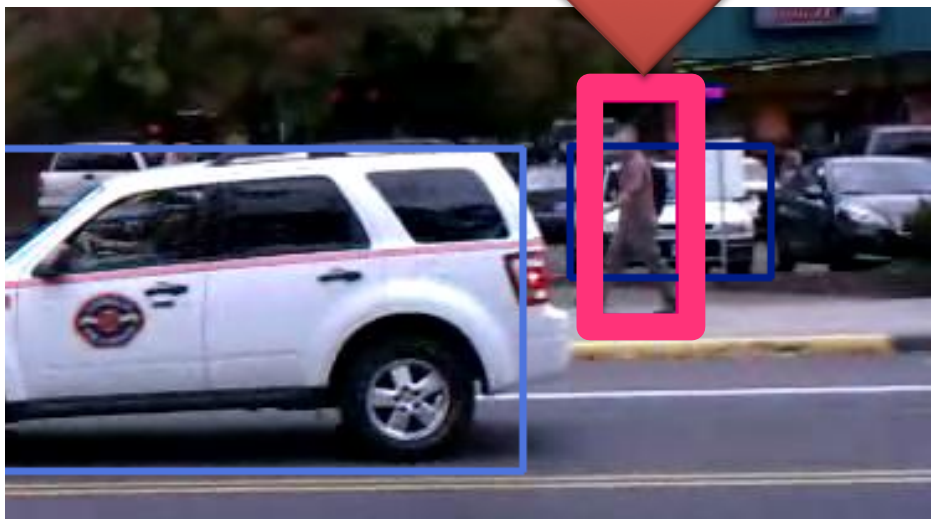


Frame 248

Failure: Occlusion

Switch IDs

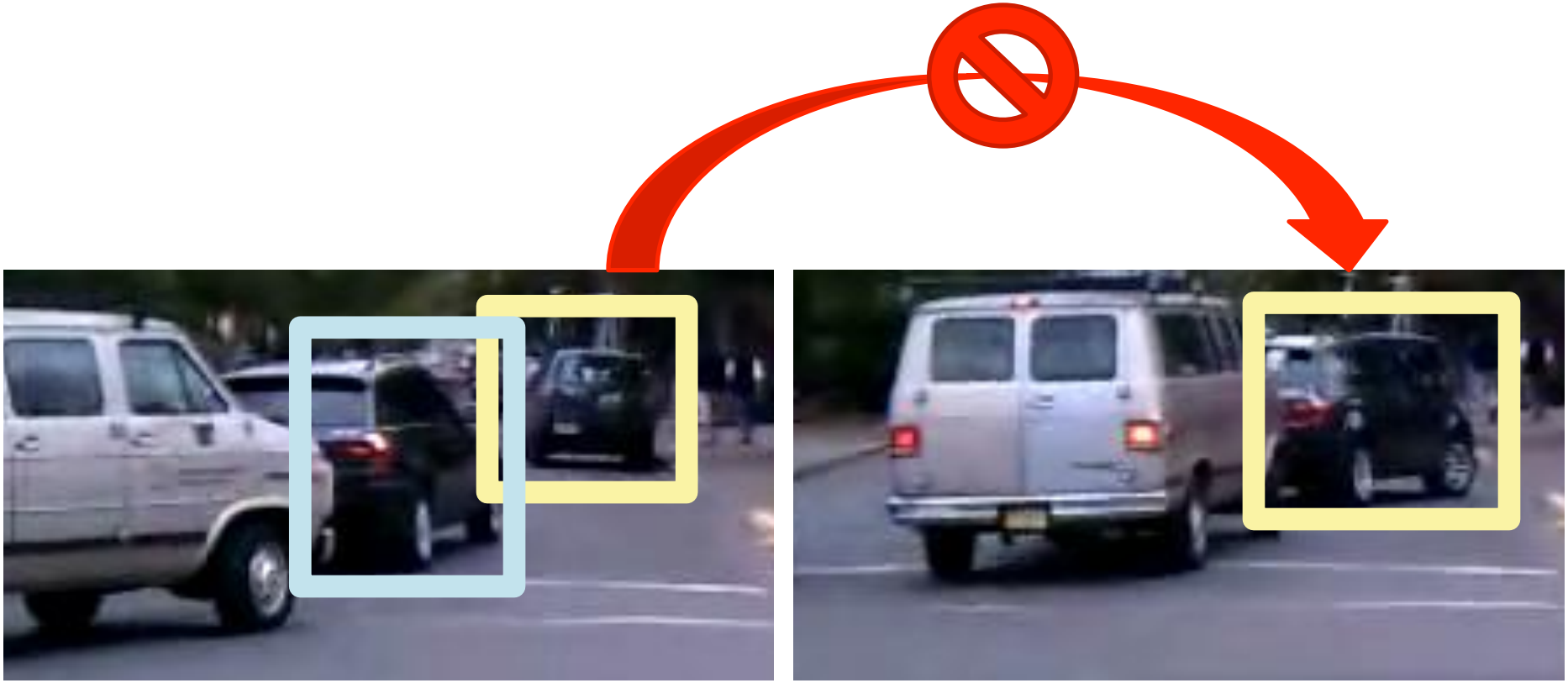




Frame 201

Frame 226

Failure: Drifting



Frame 164 Frame 228

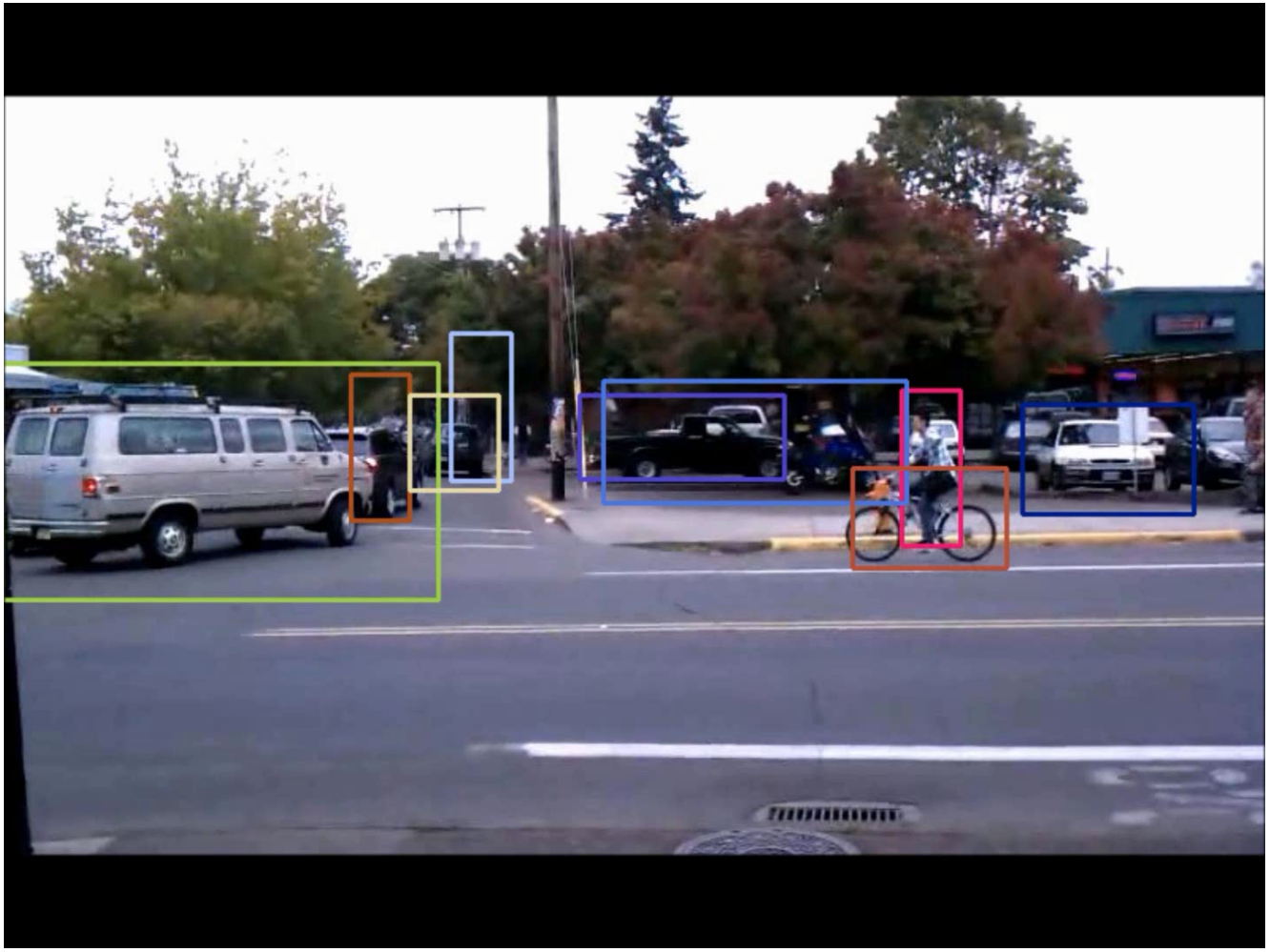
Two target merged into one

CLEAR MOT Metrics

We outperform state of the art

Dataset	Prec.	Accur.	False Neg.	False Pos.	ID Switch
Central	**72.0%**	**74.2%**	**21.7%**	**0.7%**	**0**
Central[3]	70.0%	72.9%	26.8%	0.3%	0
Central[13]	66.0%	33.8%	51.3%	14.7%	5
Hockey	**60.0%**	**79.7%**	**19.5%**	**1.1%**	**0**
Hockey[3]	57.0%	76.5%	22.3%	1.2%	0
Hockey[16]	51.0%	67.8%	31.3%	0.0%	11
i-Lids	**70.0%**	**78.6%**	**19.4%**	**1.5%**	**1**
i-Lids[3]	66.0%	76.0%	22.0%	2.0%	2
i-Lids[10]	-	68.4%	29.0%	13.7%	-
i-Lids[22]	-	55.3%	37.0%	22.8%	-
Crossing	**73.0%**	**85.9%**	**10.8%**	**1.2%**	**2**
Crossing[3]	71.0%	84.3%	14.1%	1.4%	2
Soccer	**70.0%**	**87.2%**	**6.1%**	**4.9%**	**3**
Soccer[3]	67.0%	85.7%	7.9%	6.2%	4

Conclusion

- Data association formulated as MWIS
- New formulation of MWIS that:
  - Preserves hard constraints
  - Guarantees tightness to the original problem
- New MWIS algorithm and theoretical analysis
- Favorable results on challenging datasets

Thank You