Online multi-target tracking with strong and weak detections

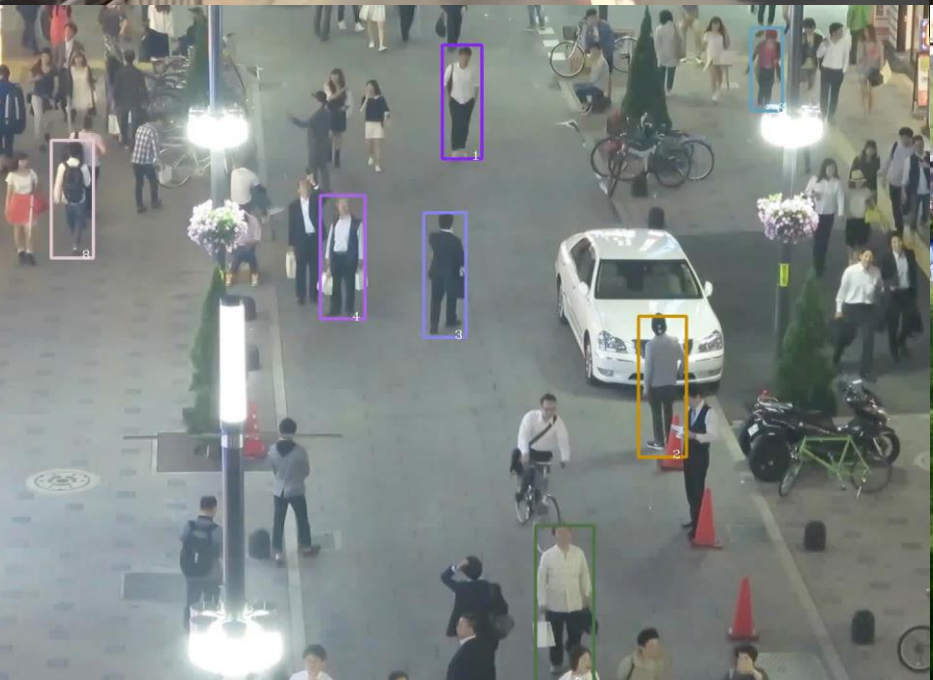
Ricardo Sánchez Matilla, Fabio Poiesi, Andrea Cavallaro

CIS centre for intelligent sensing

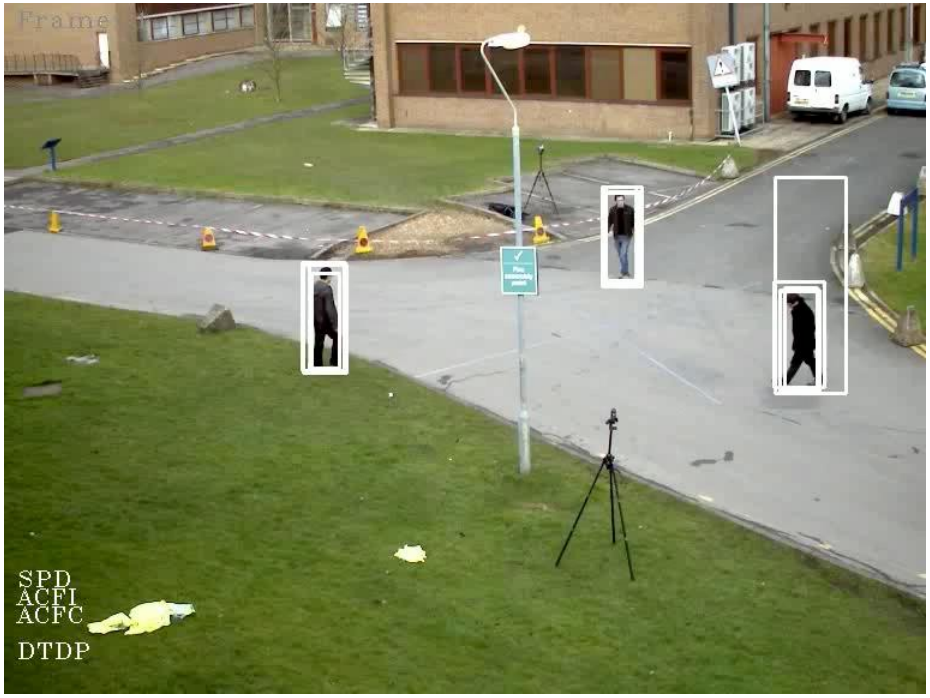
Queen Mary University of London



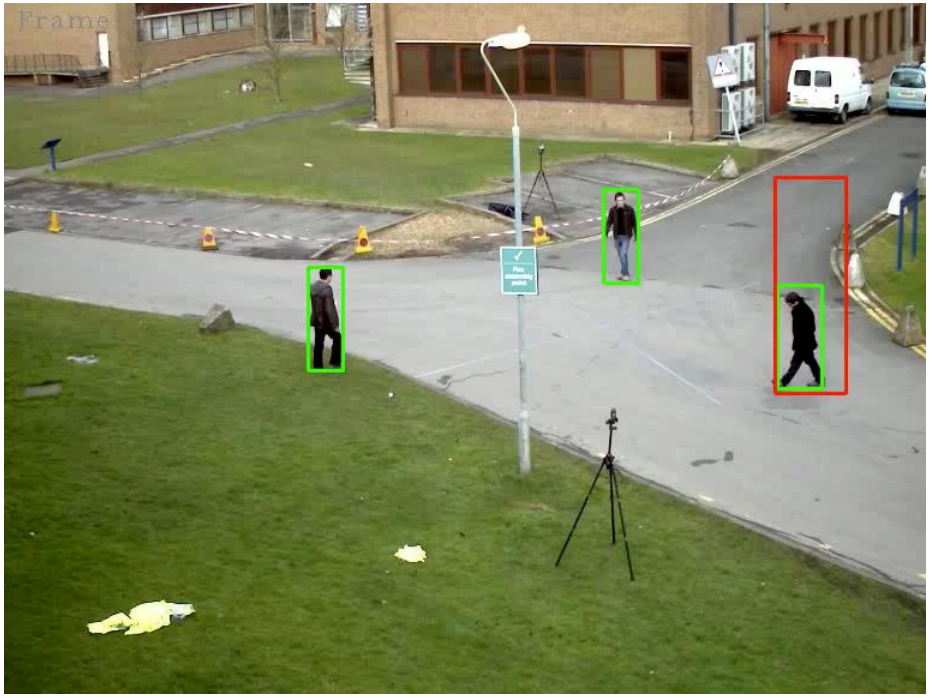





Strong vs weak detections

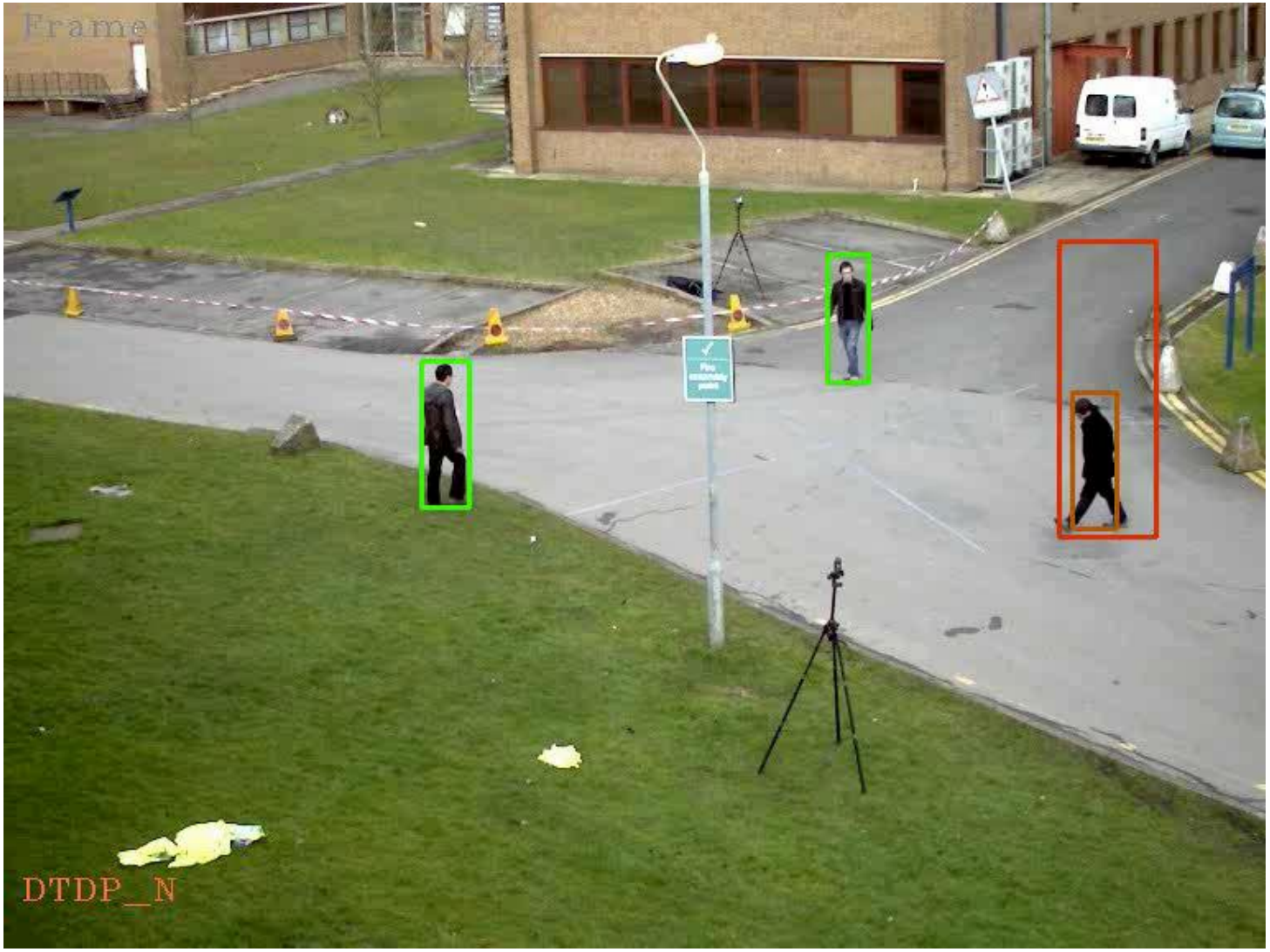


Detections



Z_k^+: strong detections
Z_k^-: weak detections

Confidence score

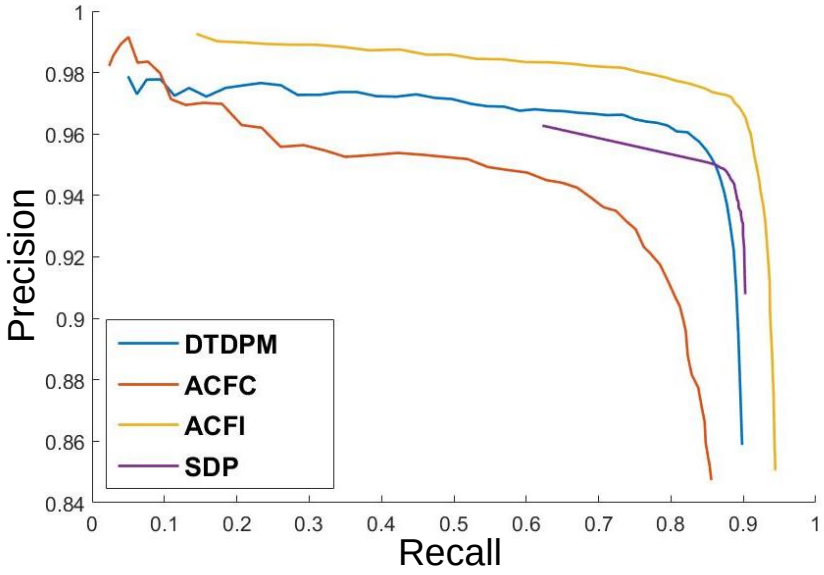


Highest

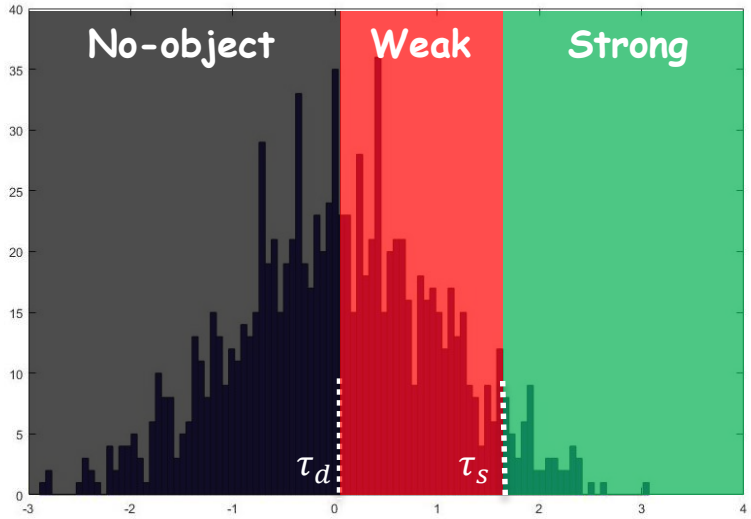
Lowest

Confidence score

Strong vs weak detections: classification

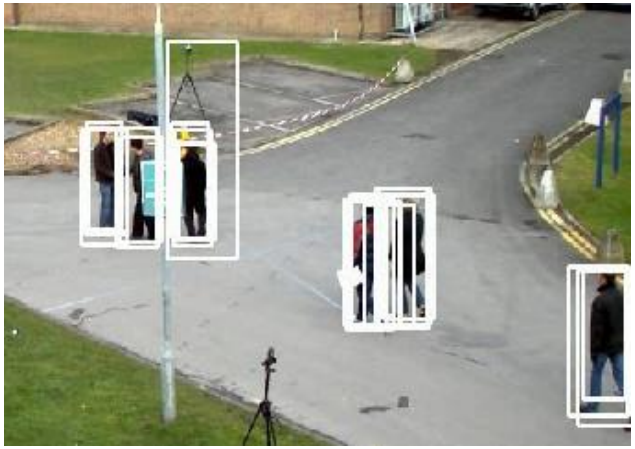


Detection results on PETS09-S2L1



Confidence score histogram

Strong and weak detections



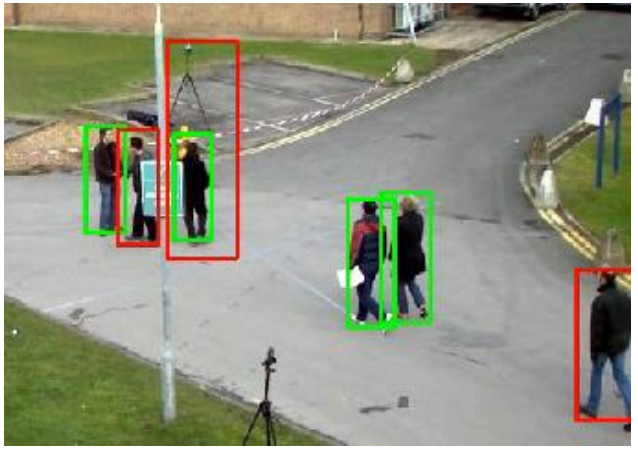
Assumption:
(large) set of
target detections Z_k^*

Combination

$$s_k = \frac{U}{D^2}\sum_{j=1}^{U} s_j \ , s_k \in [0,1]$$

D: # detectors

U: # detectors contributing to an specific combined detection

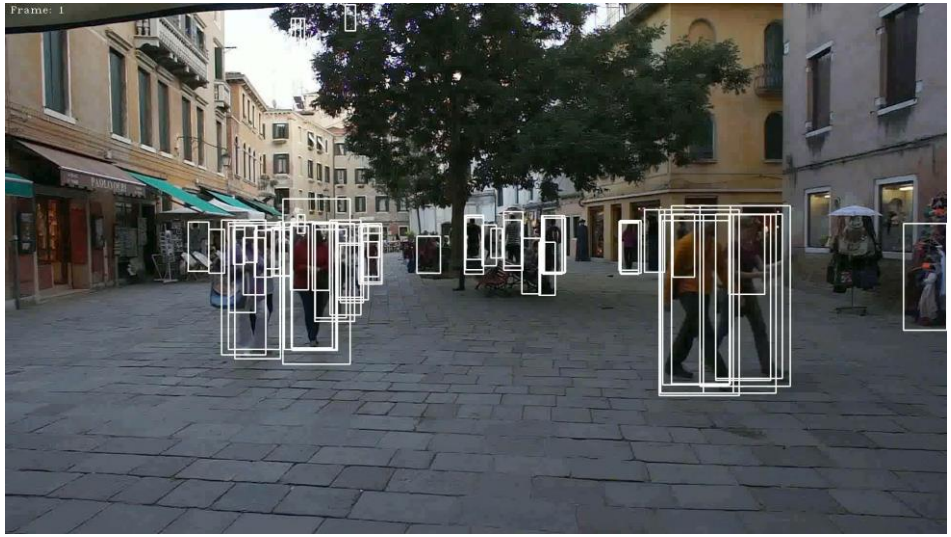


Classification

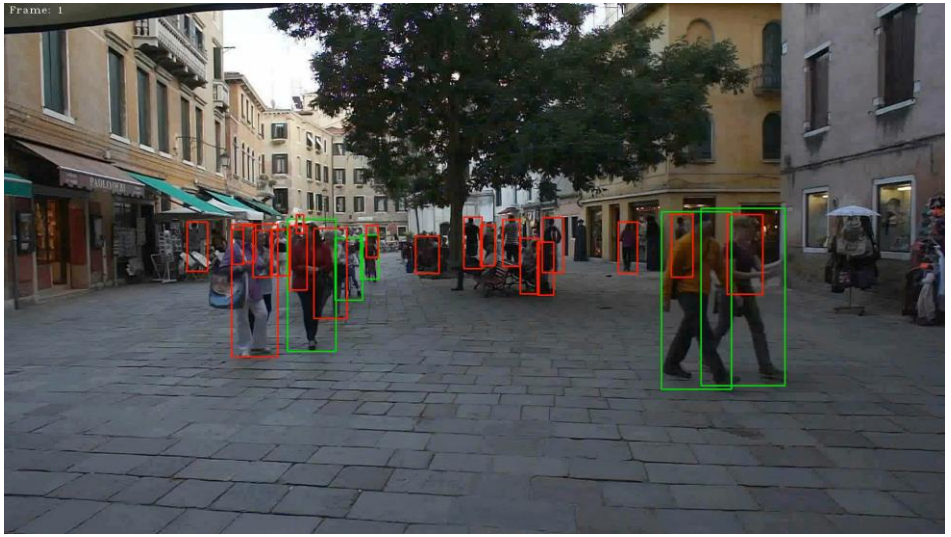
Z_k^+: strong detections
Z_k^-: weak detections

$$Z_k^+ = \{z_k^+ : s_k \geq \tau_s\}$$
$$Z_k^- = \{z_k^- : s_k < \tau_s\}$$

Strong and weak detections

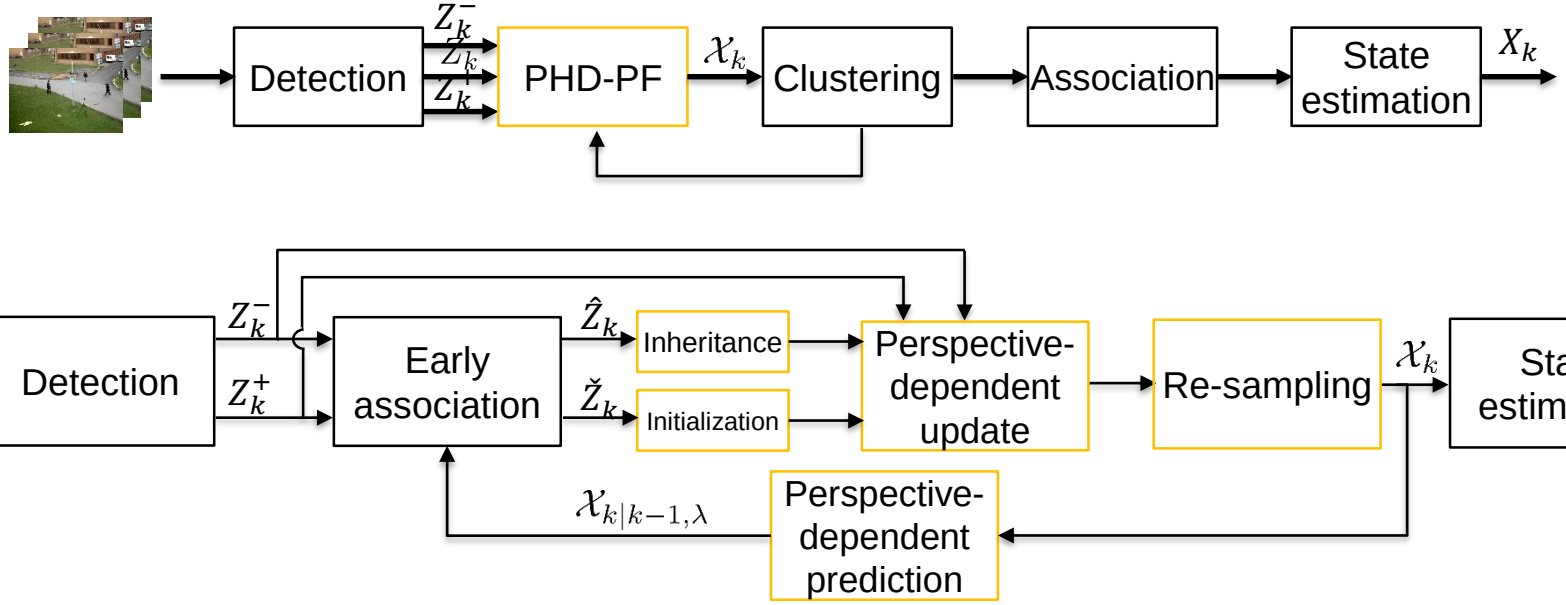


Detections



Z_k^+: strong detections
Z_k^-: weak detections

The framework

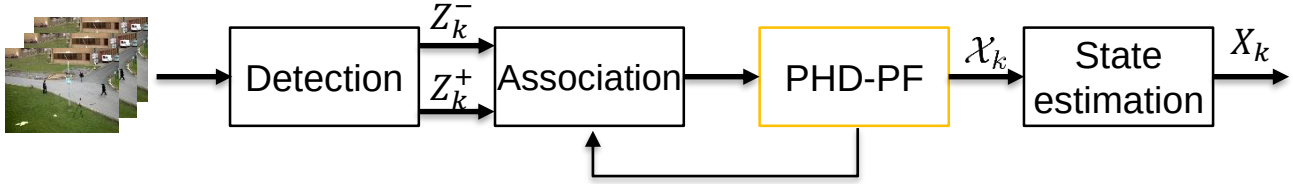


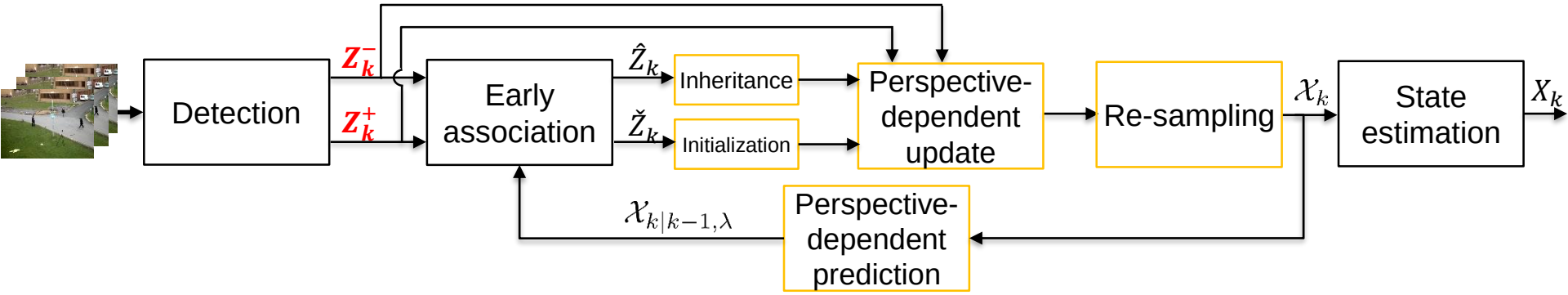
Z_k : detections
Z_k^+: strong detections
Z_k^-: weak detections

$\hat{Z}_k$: associated strong and weak detections
$\check{Z}_k$: un-associated strong detections

$\chi_{k,\lambda}$: particles
X_k: estimated states

The framework



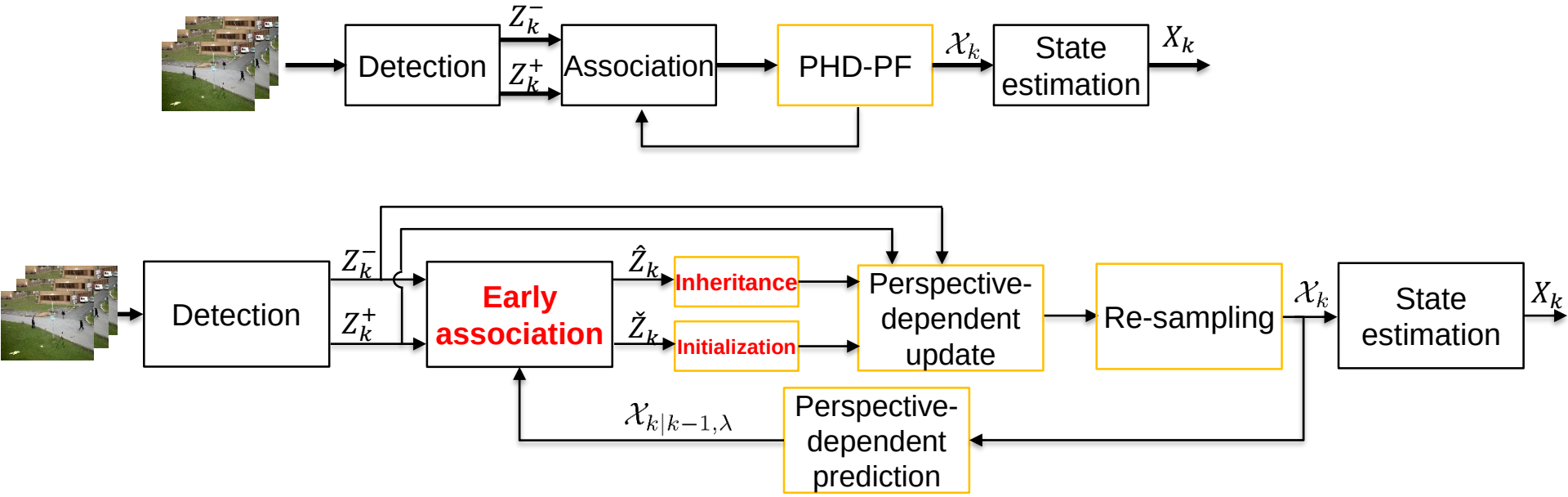


Z_k : detections
Z_k^+: strong detections
Z_k^-: weak detections

$\hat{Z}_k$: associated strong and weak detections
$\check{Z}_k$: un-associated strong detections

$\chi_{k,\lambda}$: particles
X_k: estimated states

The framework

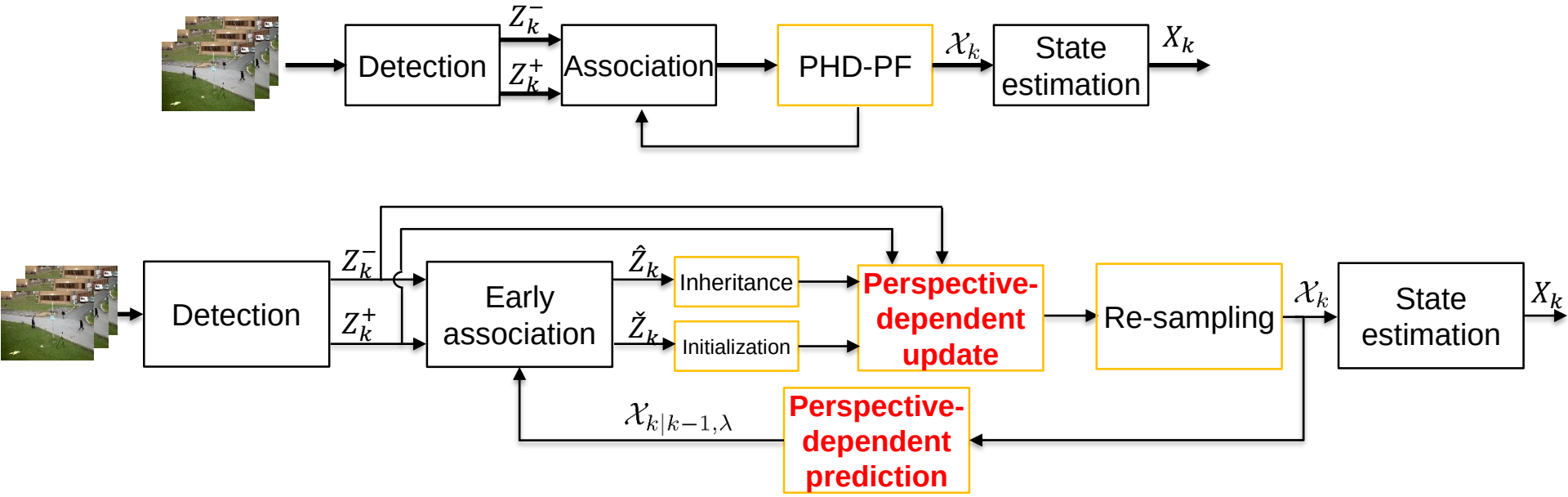


Z_k : detections
Z_k^+: strong detections
Z_k^-: weak detections

$\hat{Z}_k$: associated strong and weak detections
$\check{Z}_k$: un-associated strong detections

$\chi_{k,\lambda}$: particles
X_k: estimated states

The framework



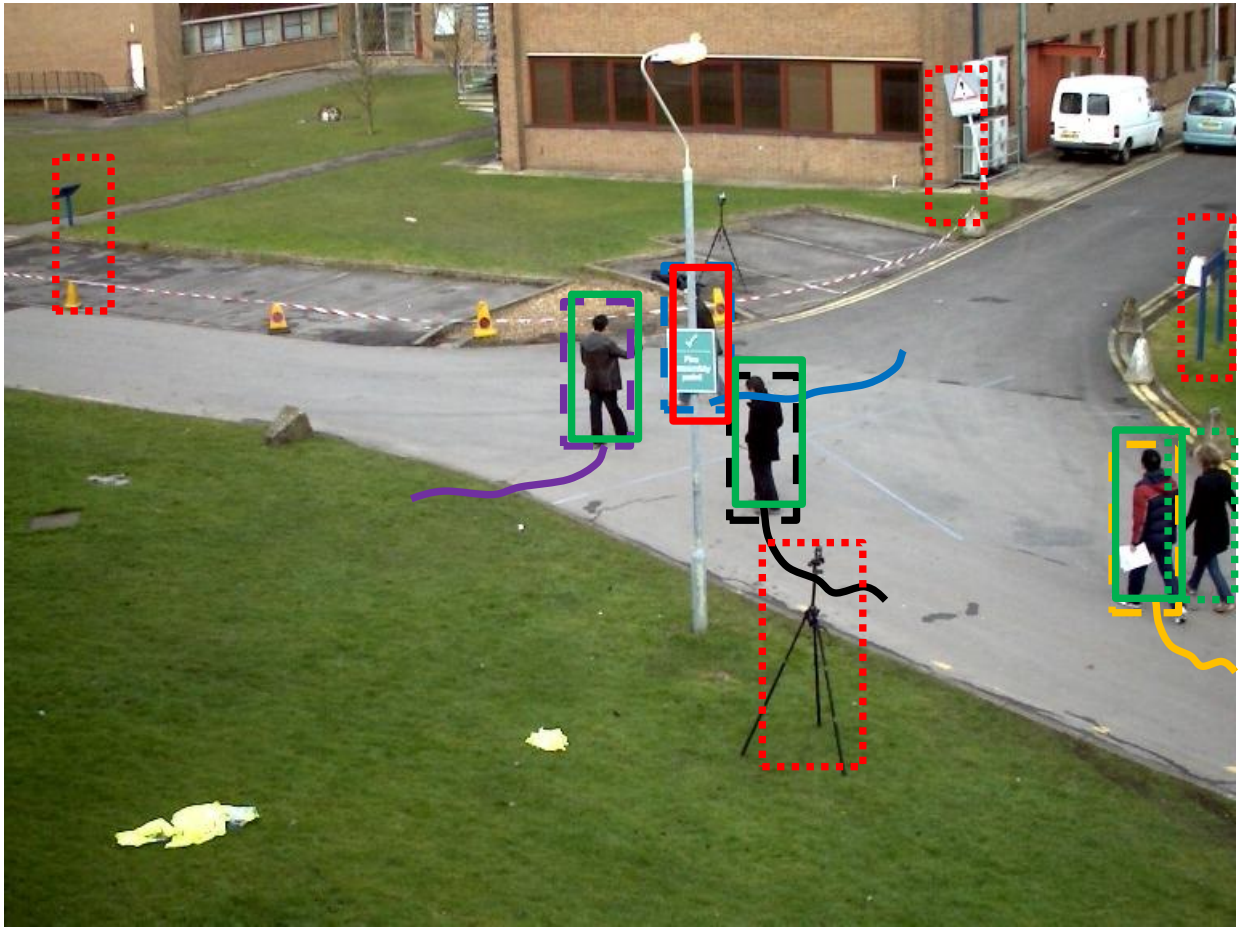
Z_k : detections
Z_k^+: strong detections
Z_k^-: weak detections

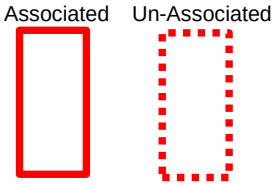
$\hat{Z}_k$: associated strong and weak detections
$\check{Z}_k$: un-associated strong detections

$\chi_{k,\lambda}$: particles
X_k: estimated states

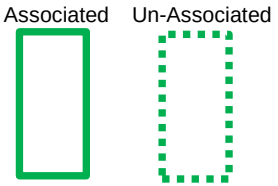
Early Association + Inheritance

- New-Born particles are generated from associated strong & weak detections
  - **Inheriting** existing label

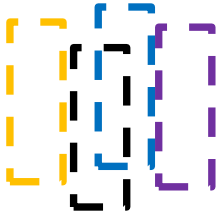




Weak detection



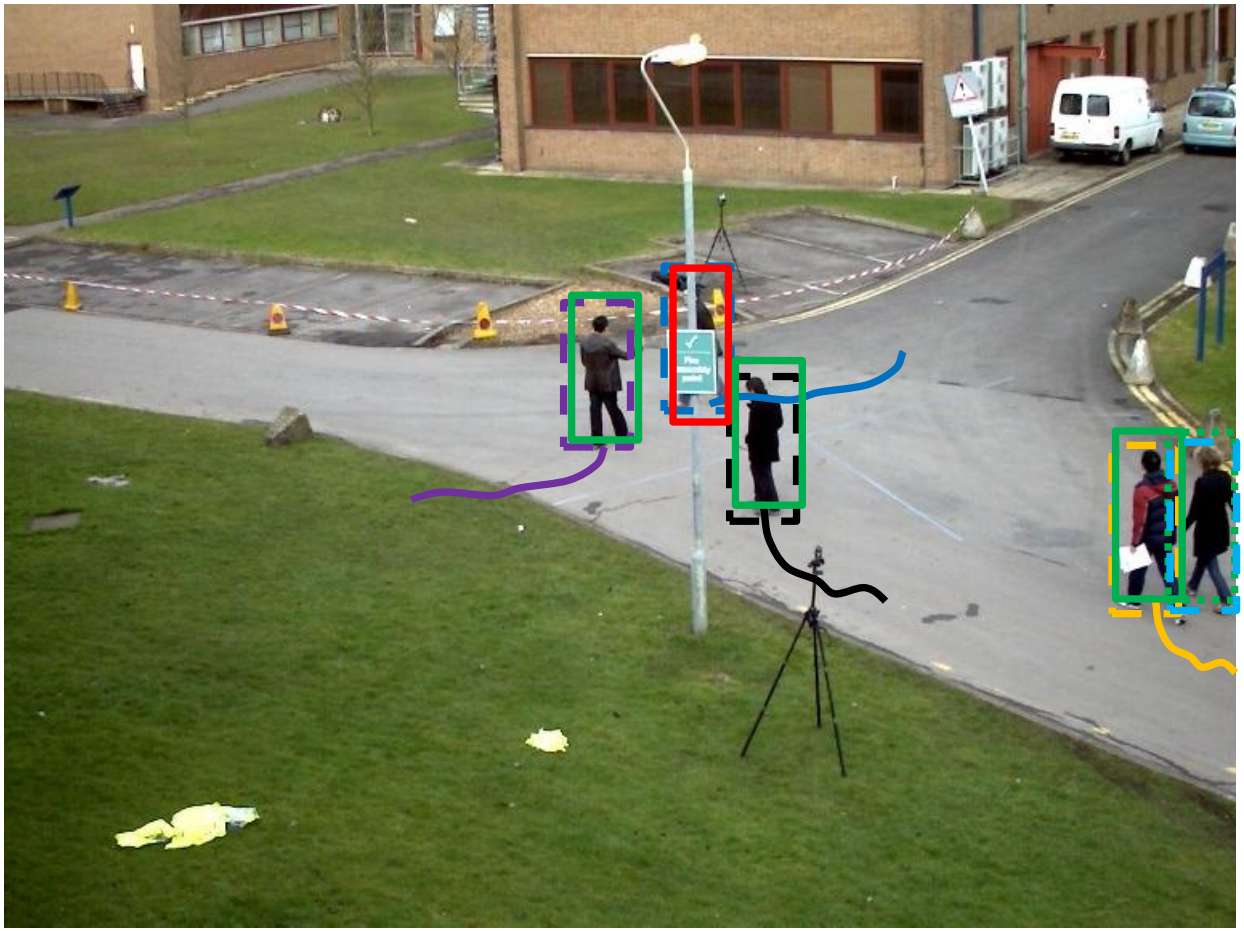
Strong detection

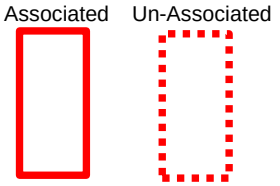


Predicted state

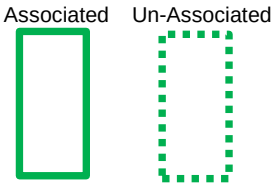
Early Association + Initialization

- New-Born particles are generated from <u>un-associated strong</u> detections
  - **Initializing** new label

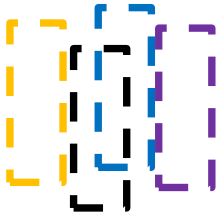




Weak detection



Strong detection

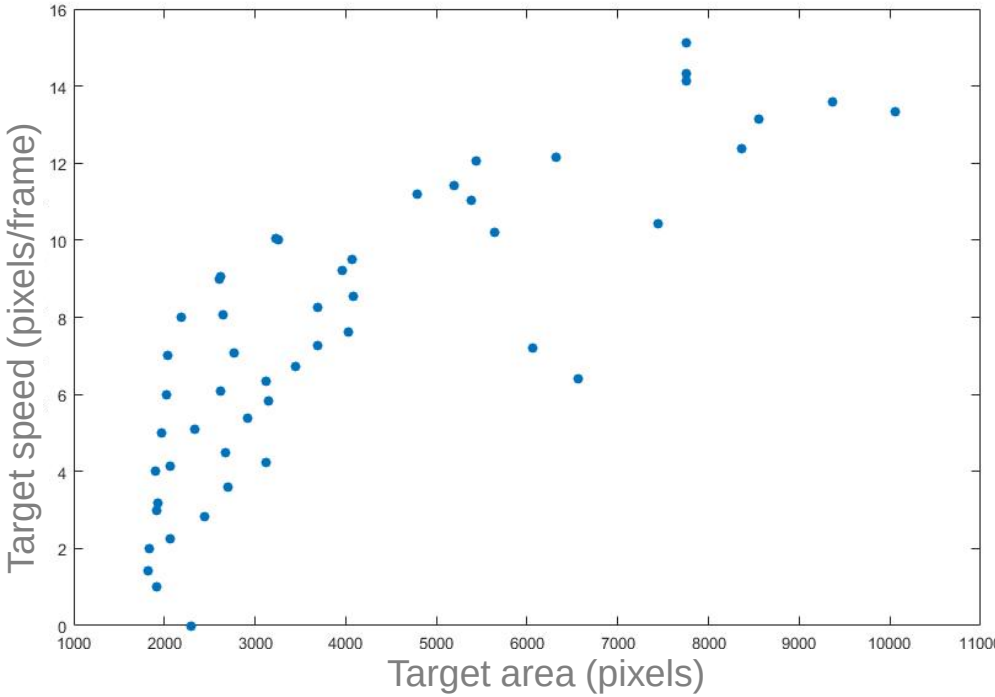


Predicted state

Perspective-dependency

Perspective-dependency

- Learning variations in position, velocity and size from the training dataset



Closer → **Bigger** → **Faster**

Assuming targets with the same size

Evaluation

Dataset		Length (frames)	Length (s)	Density (target/frame)
MOT15	Training	5503	389	7.3
	Test	5783	607	10.6
MOT16	Training	5316	215	20.8
	Test	5919	759	30.8

- The CLEAR MOT Metrics [Bernardin2008]
- Detection set Z_k^* generated with
  - DTDPM [Felzenszwalb2010]
  - ACF trained with Caltech [Dollar2014]
  - ACF trained with INRIA [Dollar2014]
  - SDP [Yang2016]

Online tracking results – private detections

Dataset	Tracker	MOTA	MOTP	FP	FN	IDS	Frag	Hz
MOT15								
MOT16								

Online tracking results – private detections

Dataset	Tracker	MOTA	MOTP	FP	FN	IDS	Frag	Hz
MOT15	EAMTT (proposed)	**53.0 ± 11.1**	**75.3**					11.5
	FOMT	53.0 ± 12.2	74.8					**16.0**
	AMPL	51.9 ± 11.9	75.0					2.8
	LKDAT_CNN	49.3 ± 11.8	74.5					1.2
	MOT_DL	49.1 ± 12.9	73.9					3.9
MOT16	POI	**66.1 ± 13.3**	**79.5**					9.9
	KFILDAwSDP	57.3 ± 15.9	77.5					2.2
	EAMTT (proposed)	52.5 ± 11.4	78.8					**12.2**
	AMPL	50.9 ± 7.1	77.0					1.5

Online tracking results – private detections

Dataset	Tracker	MOTA	MOTP	FP	FN	IDS	Frag	Hz
MOT15	EAMTT (proposed)	**53.0 ± 11.1**	**75.3**	7538	**20590**	776	1269	11.5
	FOMT	53.0 ± 12.2	74.8	6974	20776	1143	2043	**16.0**
	AMPL	51.9 ± 11.9	75.0	6963	22225	**372**	**1130**	2.8
	LKDAT_CNN	49.3 ± 11.8	74.5	**6009**	24550	563	1155	1.2
	MOT_DL	49.1 ± 12.9	73.9	8488	22281	511	1390	3.9
MOT16	POI	**66.1 ± 13.3**	**79.5**	5061	**55914**	805	3093	9.9
	KFILDAwSDP	57.3 ± 15.9	77.5	15682	60252	1873	2664	2.2
	EAMTT (proposed)	52.5 ± 11.4	78.8	4407	81223	910	1321	**12.2**
	AMPL	50.9 ± 7.1	77.0	**3229**	86123	**196**	**639**	1.5

Online tracking results – public detections

Dataset	Tracker	MOTA	MOTP	FP	FN	IDS	Frag	Hz
MOT16								

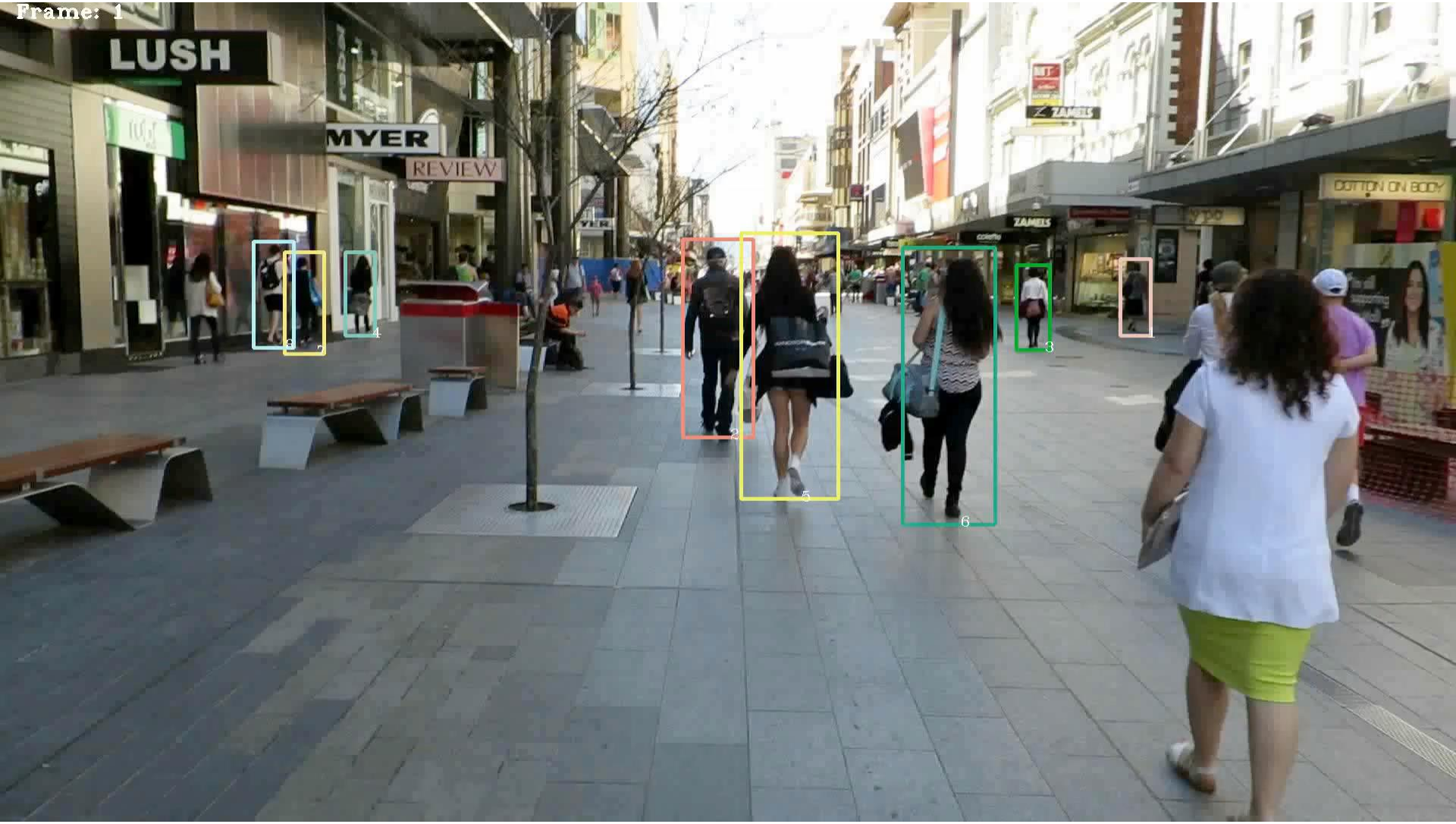
Online tracking results – public detections

Dataset	Tracker	MOTA	MOTP	FP	FN	IDS	Frag	Hz
MOT16	olCF	**43.2 ± 10.2**	74.3					0.4
	EAMTT (proposed)	38.8 ± 8.5	75.1					11.8
	LRIM	36.9 ± 17.7	75.0					10.0
	GMPHD_HDA	29.7 ± 7.3	**75.2**					**13.6**

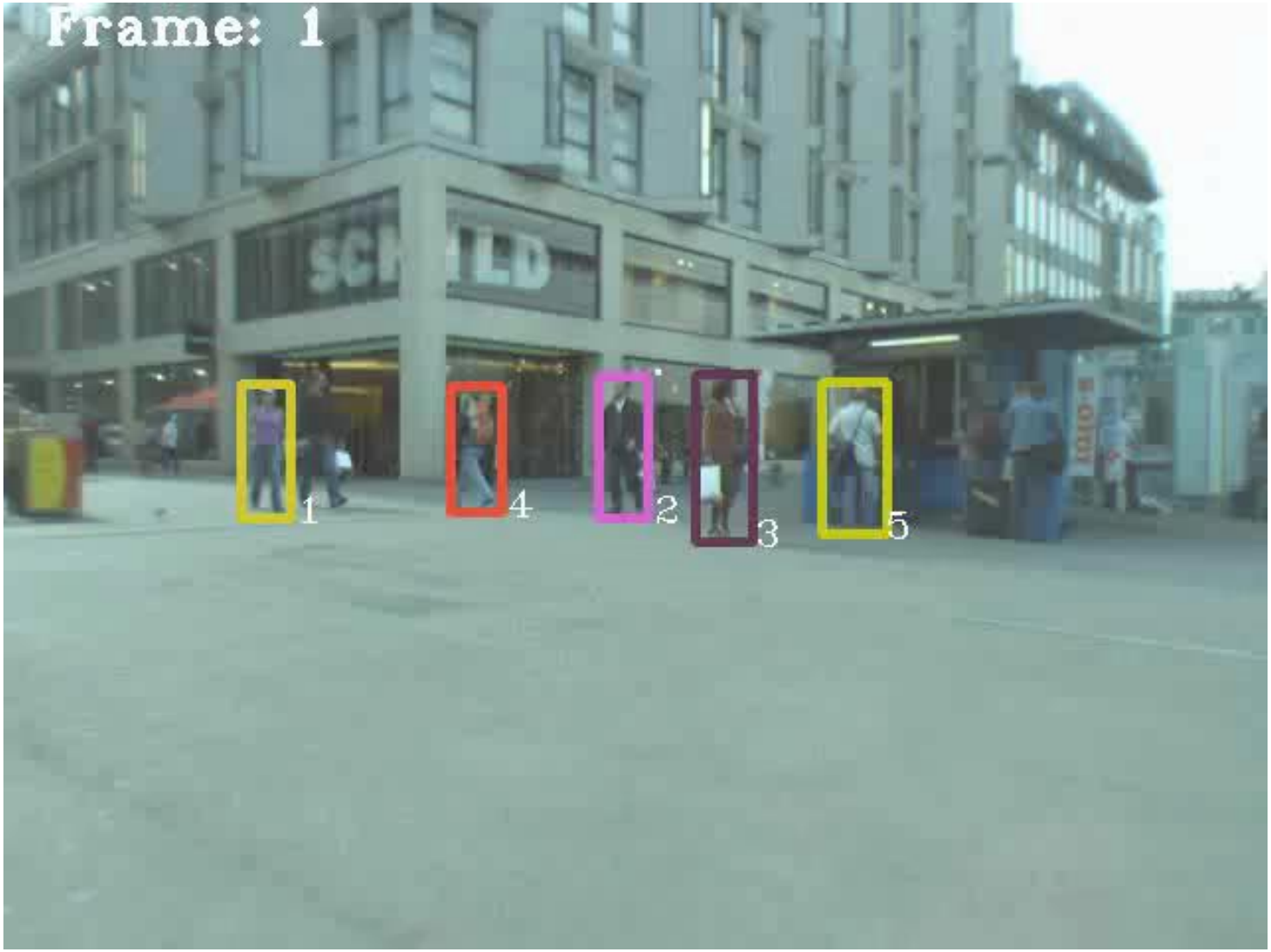
Online tracking results – public detections

Dataset	Tracker	MOTA	MOTP	FP	FN	IDS	Frag	Hz
MOT16	olCF	**43.2 ± 10.2**	74.3	**6651**	**96515**	**381**	**1404**	0.4
	EAMTT (proposed)	38.8 ± 8.5	75.1	8114	102452	965	1657	11.8
	LRIM	36.9 ± 17.7	75.0	14418	97716	2995	4968	10.0
	GMPHD_HDA	29.7 ± 7.3	**75.2**	17426	107552	3180	4483	**13.6**

Results: ADL-Rundle-1

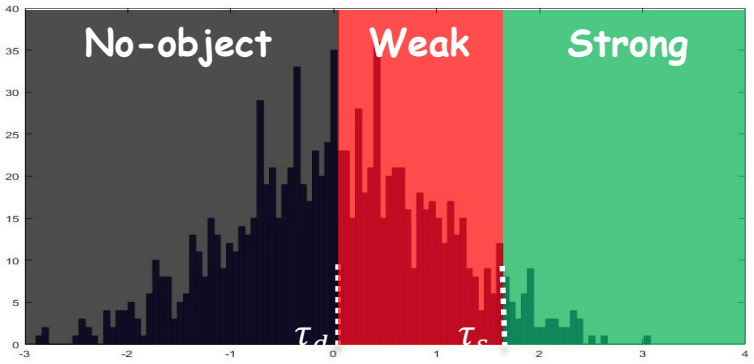


Results: ETH-Linthescher



Weak detections

Strong and weak detections



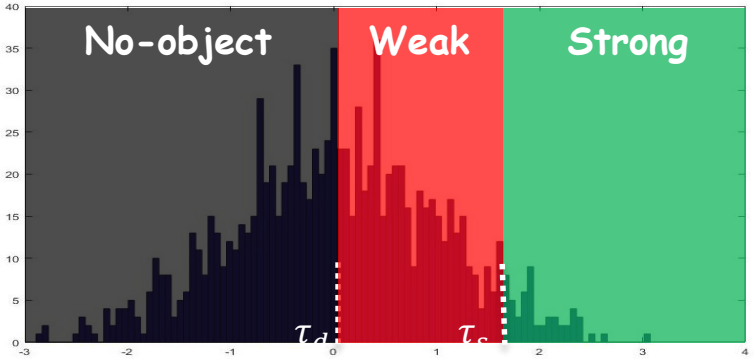
	MOTA	MOTP	FP	FN	IDs	FM
All as strong detections	-102.47	77.13	75217	**4647**	933	940

Results on MOT15 training dataset

Weak detections

Strong and weak detections

Tracking without weak detections



	MOTA	MOTP	FP	FN	IDs	FM
All as strong detections	-102.47	77.13	75217	**4647**	933	940
Without weak detections	44.95	**78.74**	**3185**	18154	627	803

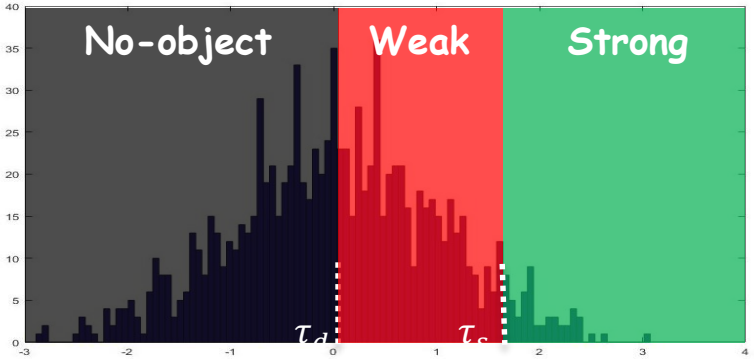
Results on MOT15 training dataset

Weak detections

Strong and weak detections

Tracking without weak detections

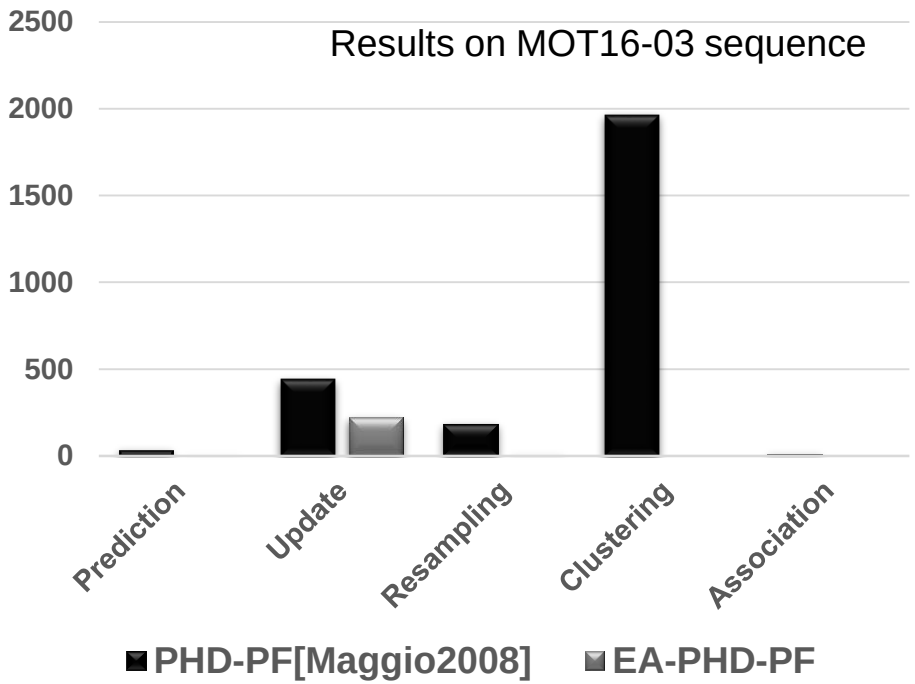
Tracking with weak detections



	MOTA	MOTP	FP	FN	IDs	FM
All as strong detections	-102.47	77.13	75217	**4647**	933	940
Without weak detections	44.95	**78.74**	**3185**	18154	627	803
With weak detections	**56.18**	78.48	5387	**11870**	**229**	**452**

Results on MOT15 training dataset

Execution time



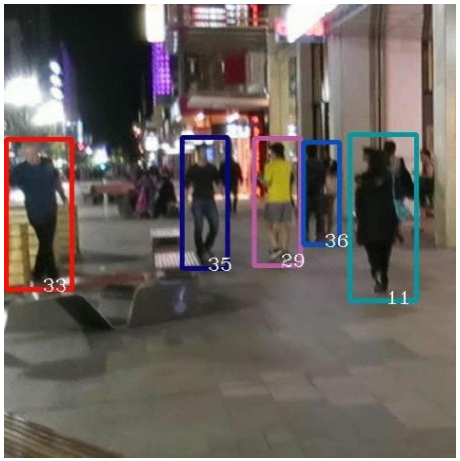
Tracker on NVIDIA Jetson TK1

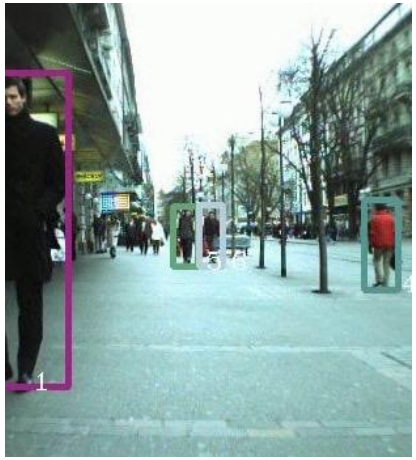
Tracker	Average execution time (fps)	
	MOT16-03 30 fps avg 69.7 targets/frame	PETS09-S2L1 7 fps avg 5.6 targets/frame
PHD-PF [Maggio2008]	0.4	4.9
EA-PHD-PF (proposed)	4.3	45.8

x10 faster

Results on MOT15 training dataset

Perspective dependency

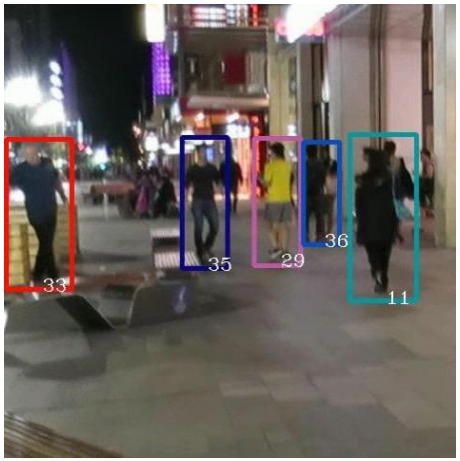


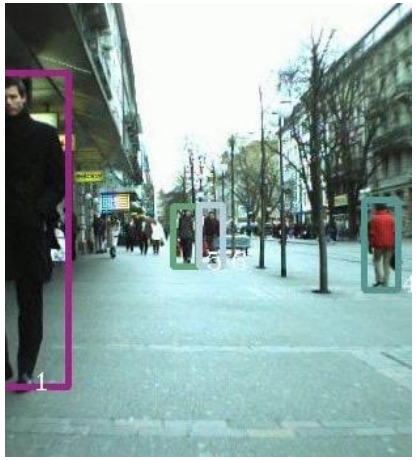


	MOTA	MOTP	FP	FN	IDs	FM
Without perspective-dependency	40.32	73.68	7470	15387	957	1654

Results on MOT15 training dataset

Perspective dependency





	MOTA	MOTP	FP	FN	IDs	FM
Without perspective-dependency	40.32	73.68	7470	15387	957	1654
With perspective-dependency	**56.18**	**78.48**	**5387**	**11870**	**229**	**452**

Results on MOT15 training dataset

Conclusions

- Introduce strong and weak detections
- Improve tracking accuracy
  - exploiting strong and weak detections
  - performing perspective dependent
  - only using spatial/temporal information
- Increase speed execution
  - disregarding clustering in a PHD-PF framework
  - ~12 fps (in a regular machine)
- Early association rely on a great overlap between targets among consecutive frames

- Future work
  - to use appearance features
  - to extend to multiple cameras
  - to perform 3D tracking