Rewind to Track: Parallelized Apprenticeship Learning with Backward Tracklets

Jiang Liu[1,2], Jia Chen[2], De Cheng[2], Chenqiang Gao[1], Alexander G. Hauptmann[2]

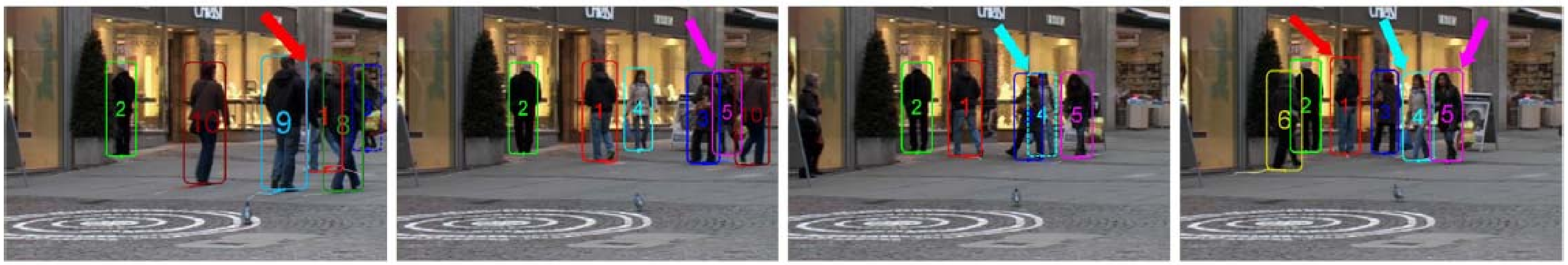
[1]Chongqing University of Posts and Telecommunications

[2]Carnegie Mellon University

Carnegie Mellon University
Language Technologies Institute

Background

Task of multiple object tracking: given a video sequence and corresponding object detections in key frames, the algorithm needs to associate detections among different frames into trajectories.



Frame 7051 Frame 7126 Frame 7164 Frame 7198

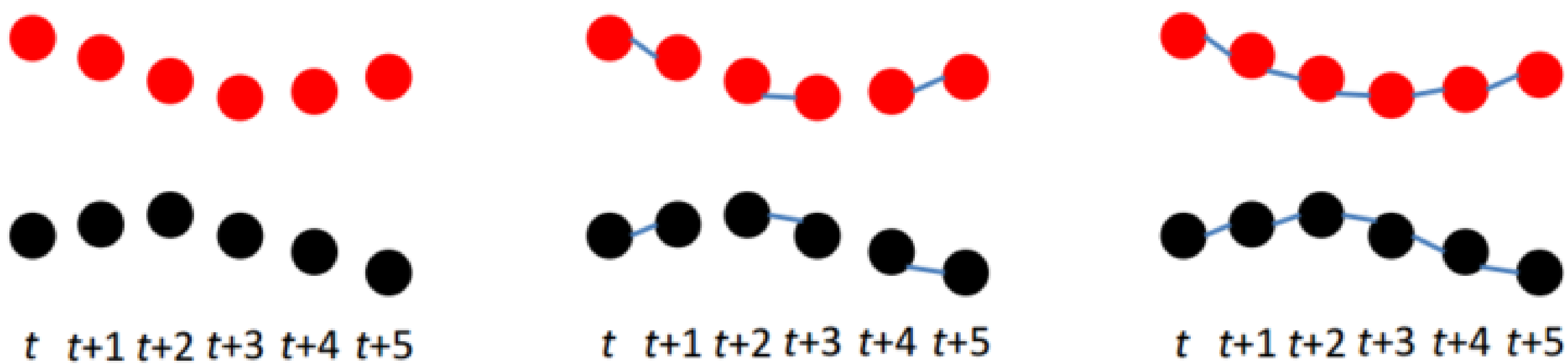
Shop Assistant 2cor 3368 Three PastShop 1cor 0837 Two Enter Shop 1cor 0268 One Stop Move Enter 1cor 0813

Zhang, S., Wang, J., Wang, Z., Gong, Y. and Liu, Y., 2015. Multi-target tracking by learning local-to-global trajectory models. *Pattern Recognition*, *48*(2), pp.580-590.

Background

Core problem of multiple object tracking based on “tracking-by-detection”:

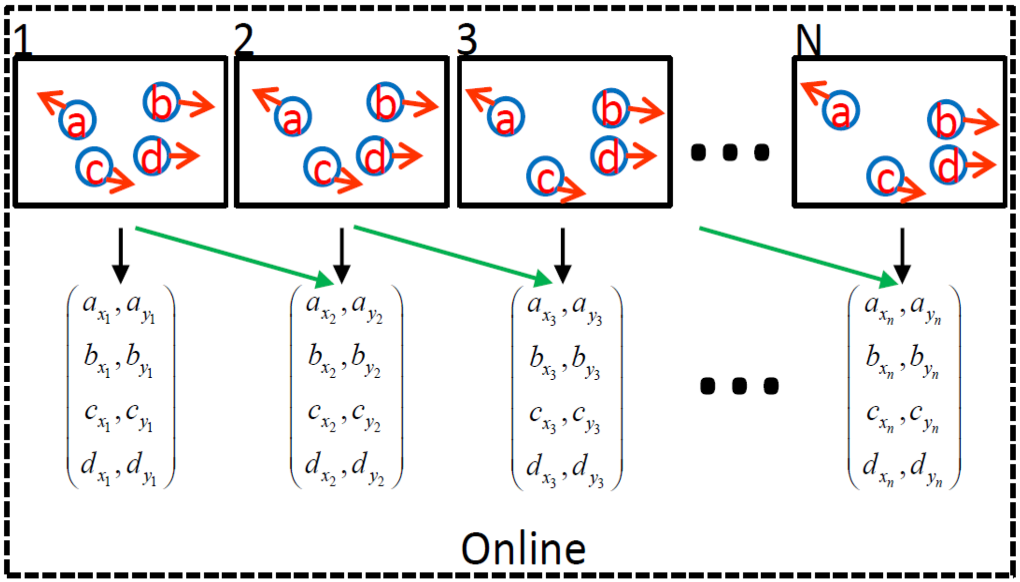
How to determine the relationship among object detections in different frames? (Data association)



An illustration of the *data association* process in the “tracking-by-detection” framework.

Luo, W., Xing, J., Zhang, X., Zhao, X. and Kim, T.K., 2014. Multiple object tracking: A literature review. *arXiv preprint arXiv:1409.7618*.

Background



Multiple object tracking based on online data association.

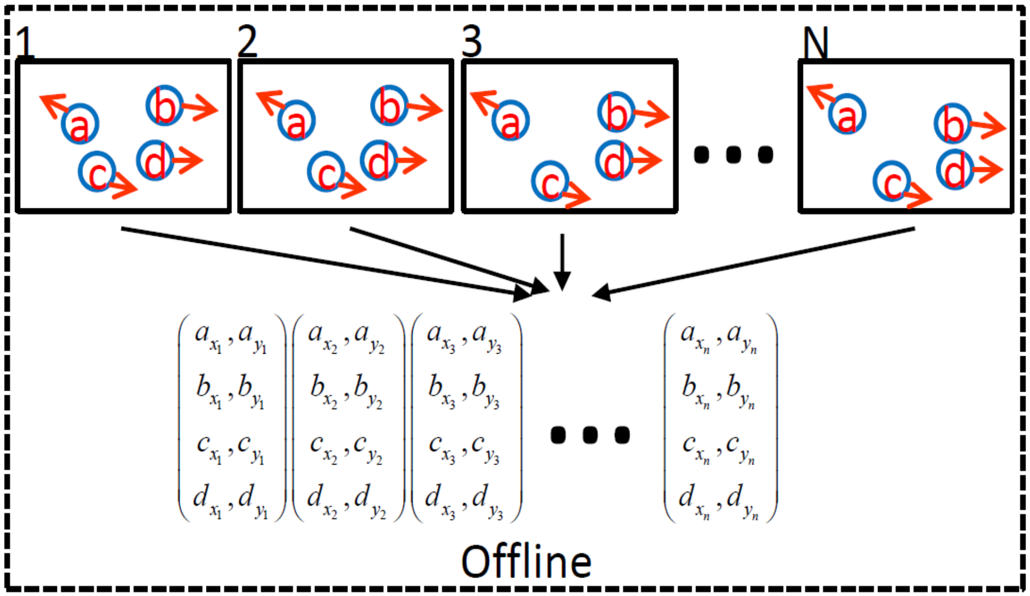
Online style data association:

- Handle tracking targets frame-by-frame;
- Only associate object detections in present frame with previous generated trajectories;
- Capable of handling online and real-time video data;
- Usually based on efficient probabilistic/deterministic optimization models;
- Tracklet drifting and ID-switching may occur when handing long-term video data.

Similarity between object detection and tracklet： $sim(o_i^t, d_k^t) = w^T \Phi(o_i^t, d_k^t) + b$

Luo, W., Xing, J., Zhang, X., Zhao, X. and Kim, T.K., 2014. Multiple object tracking: A literature review. *arXiv preprint arXiv:1409.7618*.

Background



Multiple object tracking based on offline data association.

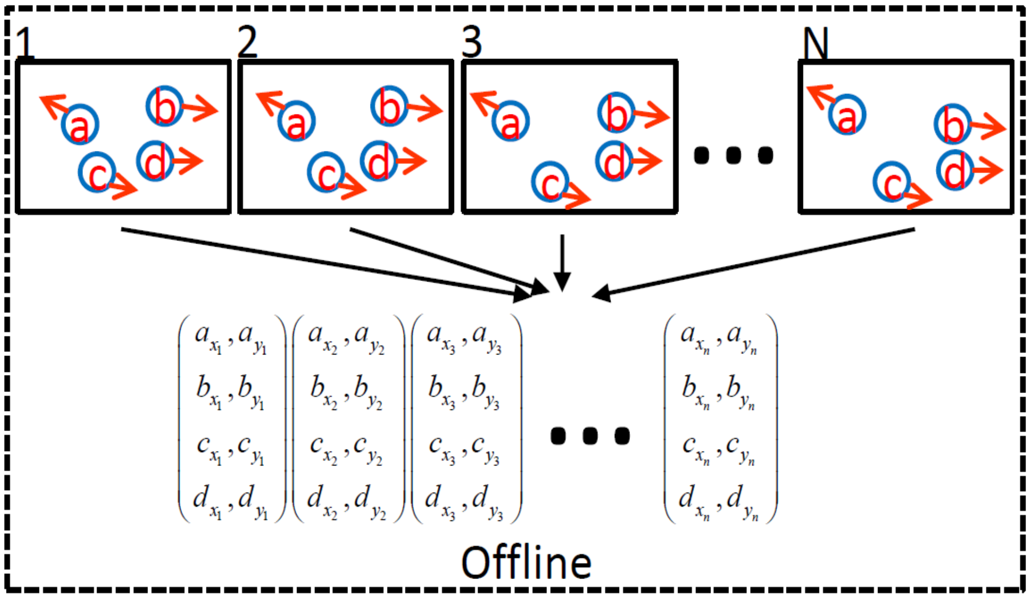
Offline style data association:

- Handle object detections from all frames in a batch manner;
- Trajectories are more robust with the observations from future frames;
- Only capable of handling offline video data;
- Usually formulated as min-cost or max-flow problem in graph;
- Seeking hierarchy solution for long-term videos: the error may also accumulated.

Luo, W., Xing, J., Zhang, X., Zhao, X. and Kim, T.K., 2014. Multiple object tracking: A literature review. *arXiv preprint arXiv:1409.7618*.

Background

Intuition: Could we adapt an efficient online mode tracker to handle offline video data, while still preserving the tracking accuracy?



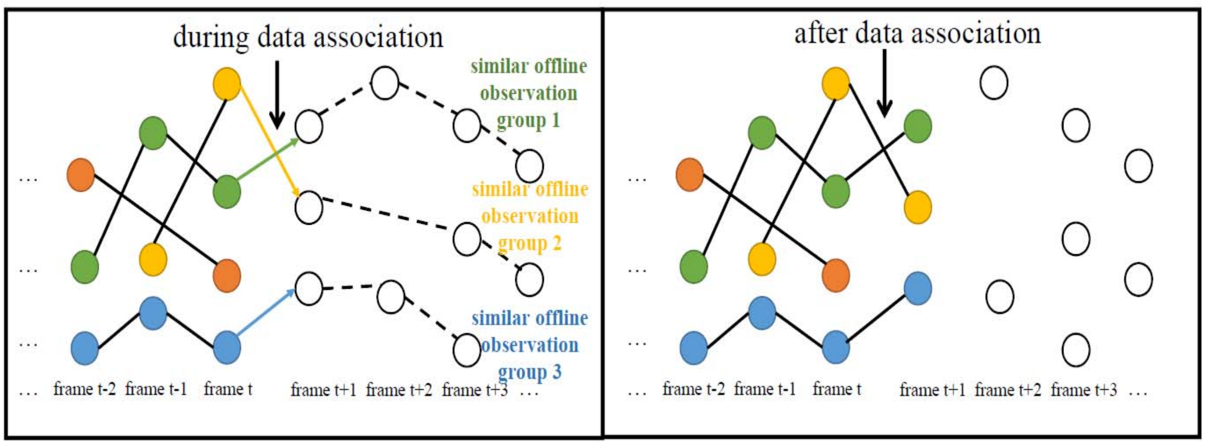
Multiple object tracking based on offline data association.

Offline style data association:

- Handle object detections from all frames in a batch manner;
- Trajectories are more robust with the observations from future frames;
- Only capable of handling offline video data;
- Usually formulated as min-cost or max-flow problem in graph;
- Seeking hierarchy solution for long-term videos: the error may also accumulated.

Luo, W., Xing, J., Zhang, X., Zhao, X. and Kim, T.K., 2014. Multiple object tracking: A literature review. *arXiv preprint arXiv:1409.7618*.

Proposed methodology



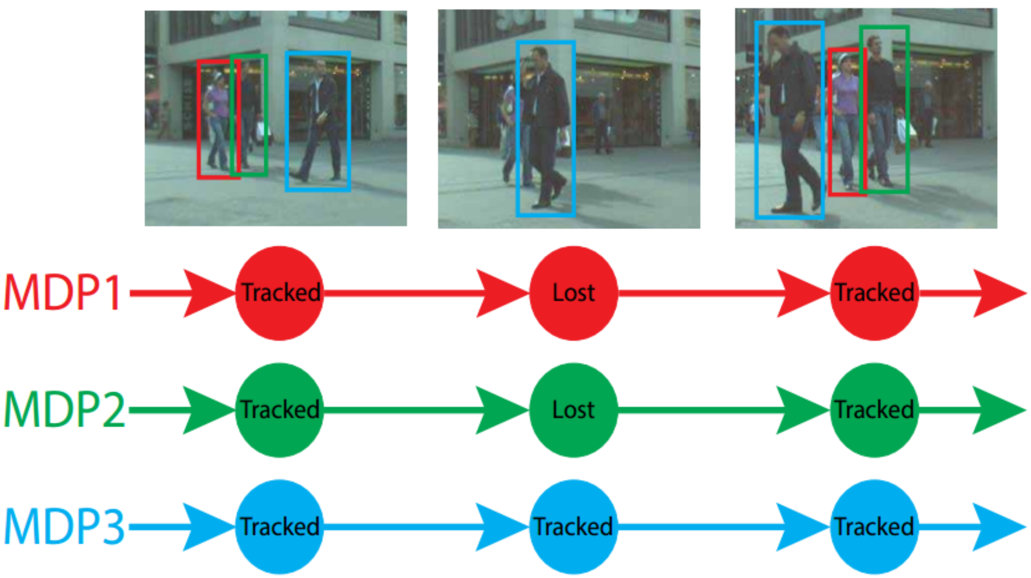
An illustration of the data association process of our proposed mixed style tracker.

Mixed style data association:

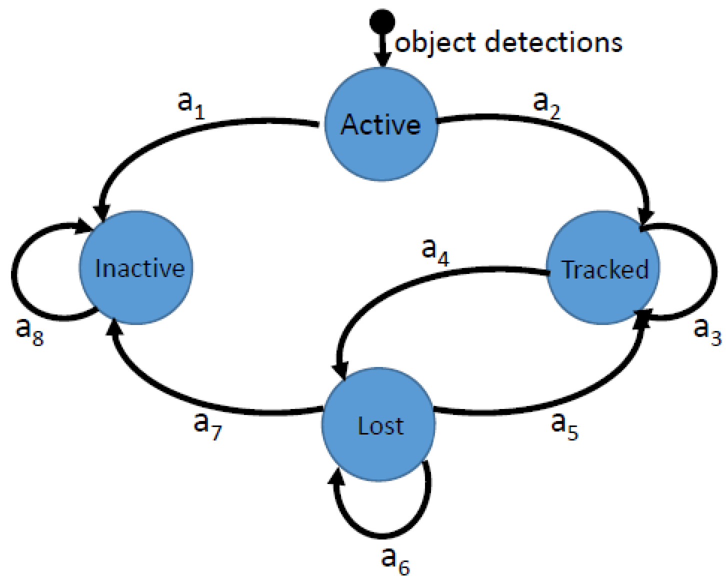
- The “Rewind to track” strategy: proposed to generate backward multiple object tracklets;
- The “similar offline observation group”(dotted line), outputted by the “Rewind to track”, is employed for robust similarity measurement;
- The final trajectories are still formulated in an online manner to preserve the efficiency;
- Only associate detections in present frame ：error will not be accumulated.

Similarity between object detection and tracklet： $sim(o_i^t, d_k^t) = \sum_{q=0}^{t'} w^T \Phi\left(o_i^t, d_k^{t,q}\right) + b,$

Proposed methodology



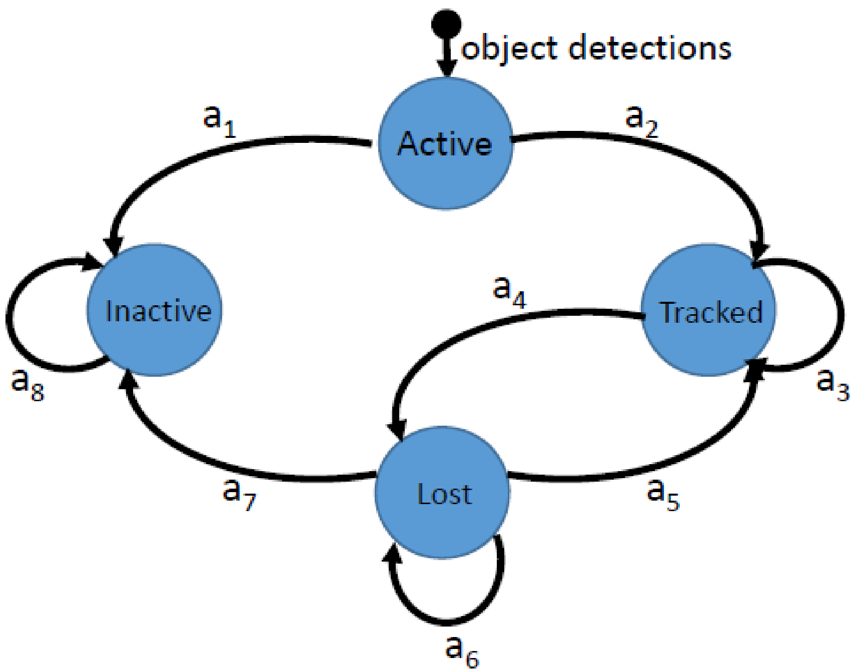
Multiple object tracking based on Markov Decision Process (MDP).



The agent's state transition map of a tracking object.

Xiang, Y., Alahi, A. and Savarese, S., 2015. Learning to track: Online multi-object tracking by decision making. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4705-4713).

Proposed methodology

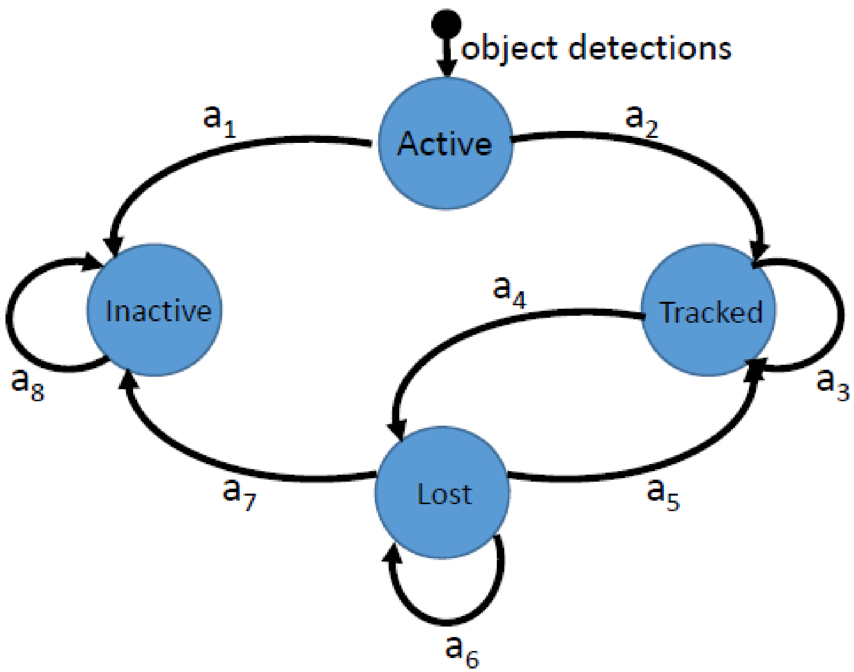


The agent's state transition map of a tracking object.

An agent's of a particular tracking object could be represented with a tuple $\big(s, a, \pi, R(s, a)\big)$.

- $s \in S$: **state,** an object's status in a particular frame, generated according to tracklets;
- $a \in A$: **action,** transit an object from one state to another;
- $\pi(s)$: **policy function,** determine a mapping from the state space *S* to the action space *A*: $\pi(s) \rightarrow a$, via maximizing the reward function;
- $R(s, a)$: real-valued **reward function** $R(s, a) : S \times A \rightarrow R$, define a reward value by executing action *a* in state *s*.

Proposed methodology



The agent's state transition map of a tracking object.

States description:

- Active：any newly appeared object detection is initialized with this state;
- Tracked：the agent will be kept in this state, if and only if it historical tracklets could be extended to present frame (based on TLD tracking assumption);
- Lost：object is disappeared or occluded. Next state may be: (1) back to Tracked state; (2) keep Lost state; (3) transfer to Inactive state (equivalent to solving the data association problem) ;
- Inactive：represents invalid object detections or permanent lost objects.

Proposed methodology

- Given the agent feature $\Phi(s)$ in state s, the reward function could be represented by a linear mapping from the feature:

$$R(s,a) = w \cdot \Phi(s)$$

- at frame t_0, the tracker adapts policy $\pi\left(s_{t_0}\right) \rightarrow a_{t_0}$. The corresponding value expectation $E\left[V^{\pi}\left(s_{t_0}\right)\right]$ (the afterwards reward by adapting a_{t_0}) is:

$$E\left[V^{\pi}\left(s_{t_0}\right)\right] = w \cdot \mu(\pi) \ ,$$

where as $\mu(\pi) = E\left[\sum_{t=t_0}^{\infty} \gamma^t \, \Phi(s) | \pi\right]$, which is the feature expectation of the agent (γ is the decay factor, 0≤γ≤1).

Proposed methodology

reward function is unknown

- Given the agent feature $\Phi(s)$ i tate s, the reward function could be represented by a linear ma policy function is unknown the feature:

$$R(s, \quad \cdot \Phi(s)$$

- at frame t_0, the tracker adapts policy $\pi(s_{t_0}) \rightarrow a_{t_0}$. The corresponding value expectation $E[V^\pi(s_{t_0})]$ (the afterwards reward by adapting a_{t_0}) is:

$$E[V^\pi(s_{t_0})] = w \cdot \mu(\pi) \ ,$$

where as $\mu(\pi) = E\left[\sum_{t=t_0}^{\infty} \gamma^t \, \Phi(s) | \pi\right]$, which is the feature expectation of the agent (γ is the decay factor, 0≤γ≤1).

Proposed methodology

- Unknown: both the reward function and the policy function
- Known: labelled groundtruth objects' trajectories on the training set, i.e., the expert's state-action sequences:

$$D = \{s_{t_1}, a_{t_1}, s_{t_2}, a_{t_2}, \ldots, s_{t_n}, a_{t_n}\}$$

- Objective: minimizing difference between expert's and algorithm's reward expectation:

$$\min||E^*[V^\pi(s_{t_i})] - \tilde{E}[V^\pi(s_{t_i})]||$$

  - solve the optimal policy function parameter: $\tilde{\pi}$ (Reinforcement Learning)
  - solve the optimal reward function parameter: $\tilde{R}(s,a)$ (Inverse Reinforcement Learning)

*Apprenticeship Learning: Reinforcement Learning + Inverse Reinforcement Learning

*Abbeel, P., & Ng, A. Y. (2004, July). Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning* (p. 1). ACM.

Proposed methodology

Q: How to training multiple agents in a particular training video?
A1: Sequentially (polling variant of AL); A2: Parallelly (parallel variant of AL)

Proposed methodology

Q: How to training multiple agents in a particular training video?

A1: Sequentially (polling variant of AL); A2: Parallelly (parallel variant of AL)

- RL phase, parallelly learning the reward function:

$$w^{(p)} = arg\max_{w} \min \sum_{j=1}^{N} w^{T}(\mu_j(\pi^{(p-1)}) - \mu_{E,j}(\pi^{(p-1)})),$$

- IRL phase, parallelly updating policy function parameters：

$$\pi^{(p)} = arg\max_{\pi} \sum_{j=1}^{N} E[V_j^{\pi}(s_t)] \,|_{\tilde{R}^{(p)}(s,a)=w^{(p)}\cdot\Phi(s)},$$

- Multiple agents feature updating：

$$\forall j, \mu_j^{(p)} = \mu_j(\pi^{(p)}).$$

Algorithm 1 Parallelized apprenticeship learning for *lost* state with backward tracklets utilization.

Input: Video sequences $V = \{v_i\}_{i=1}^{N}$, ground truth trajectories $O_i = \{o_{i,j}\}_{j=1}^{N_i}$ and object detections $D_i = \{d_j^i\}_{j=1}^{N_i'}$;

Output: reward function parameters (w_{lost}, b_{lost}) for lost status data association;

Initialization of reward function: $w_{lost}^0 \leftarrow w_0, b_{lost}^0 \leftarrow b_0, S \leftarrow \emptyset$
Initialization for each target $o_{i,j}$ in each v_i: set MDP of $o_{i,j}$ in *tracked* after $t_{start}(i,j) \leftarrow$ index of the first frame where $o_{i,j}$ correctly detected
$p \leftarrow 0$
repeat
    $p \leftarrow p + 1$
    for each video v_i in V **do**
        $t \leftarrow 1$
        while $t \leq$ last frame of v_i **do**
            for target $o_{i,j}$ in v_i which $t_{start}(i,j) \geq t$ **do**
                Follow policy π^{p-1}, compute $\mu_{i,j}^{\pi}$ as Eq.6, choose action a
                Compute ground truth action: a_{gt}
                if state is *lost* and $a \neq a_{gt}$ **then**
                    $S \leftarrow S \cup \{\phi(o_{i,j}^t, d_k^t), y_k\}$
                    $S \leftarrow S \cup \{\phi(o_{i,j}^t, d_k^{t,q}); y_k\}, 1 \leq q \leq t'$
                    Save failure position: $t_{start}(i,j) \leftarrow t$
                else
                    State transfer: Execute action a
                end if
                If all targets failed **then** break;
            end for
        end while
    end for
    Obtain new reward function parameters (w_{lost}^p, b_{lost}^p): solve Eq.4 with S
    Obtain new policy π^p: solve Eq.5 with (w_{lost}^p, b_{lost}^p)
until all targets are successfully tracked.

Proposed methodology

Q: How to training multiple agents in a particular training video?
A1: Sequentially (polling variant of AL); A2: Parallelly (parallel variant of AL)

- RL phase, parallelly learning the reward function:

$$w^{(p)} = arg\max_{w} \min \sum_{j=1}^{N} w^T\left(\mu_j\left(\pi^{(p-1)}\right) - \mu_{E,j}\left(\pi^{(p-1)}\right)\right),$$

- IRL phase, parallelly updating policy function parameters：

$$\pi^{(p)} = arg\max_{\pi} \sum_{j=1}^{N} E\left[V_j^{\pi}(s_t)\right] |_{\tilde{R}^{(p)}(s,a)=w^{(p)}\cdot\Phi(s)},$$

- Multiple agents feature updating：

$$\forall j, \mu_j^{(p)} = \mu_j\left(\pi^{(p)}\right).$$

Parallelized apprenticeship learning strategy:

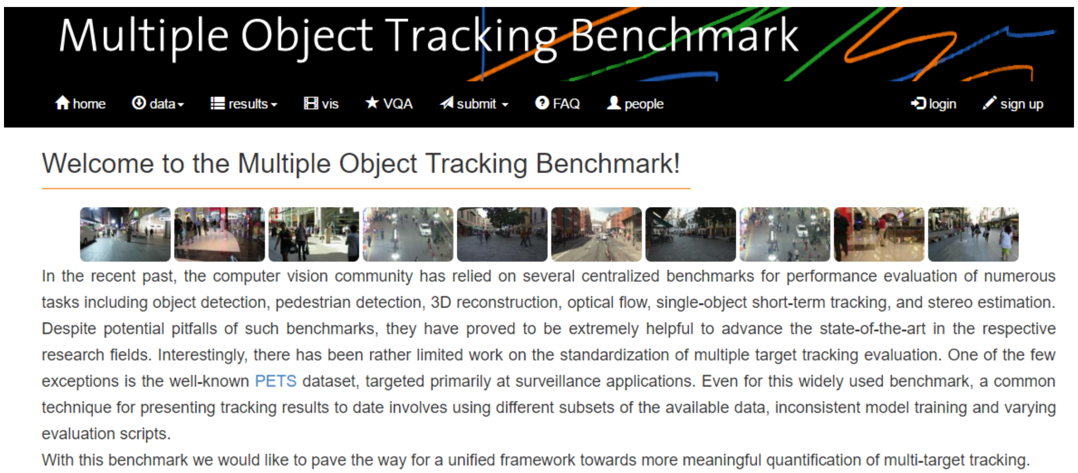
- Simultaneously maintaining the statuses of all tracking objects on the training set;
- Updating the reward function parameters with all the objects on the training video , so that the convergence speed is faster;
- Resuming the training from the last failure point for an agent. O(n) training time complexity for a video with *n* frames and *k* objects (polling variant of AL: O(*n*k*)).

Algorithm 1 Parallelized apprenticeship learning for *lost* state with backward tracklets utilization.

Input: Video sequences $V = \{v_i\}_{i=1}^{N}$, ground truth trajectories $O_i = \{o_{i,j}\}_{j=1}^{N_i}$ and object detections $D_i = \{d_j^i\}_{j=1}^{N_i'}$;
Output: reward function parameters (w_{lost}, b_{lost}) for lost status data association;

Initialization of reward function: $w_{lost}^0 \leftarrow w_0, b_{lost}^0 \leftarrow b_0, S \leftarrow \emptyset$
Initialization for each target $o_{i,j}$ in each v_i: set MDP of $o_{i,j}$ in *tracked* after $t_{start}(i,j) \leftarrow$ index of the first frame where $o_{i,j}$ correctly detected
$p \leftarrow 0$
repeat
    $p \leftarrow p+1$
    for each video v_i in V **do**
        $t \leftarrow 1$
        while $t \leq$ last frame of v_i **do**
            for target $o_{i,j}$ in v_i which $t_{start}(i,j) \geq t$ **do**
                Follow policy π^{p-1}, compute $\mu_{i,j}^{\pi}$ as Eq.6, choose action a
                Compute ground truth action: a_{gt}
                if state is *lost* and $a \neq a_{gt}$ **then**
                    $S \leftarrow S \cup \{\phi(o_{i,j}^t, d_k^t), y_k\}$
                    $S \leftarrow S \cup \{\phi(o_{i,j}^t, d_k^{t,q}); y_k\}, 1 \leq q \leq t'$
                    Save failure position: $t_{start}(i,j) \leftarrow t$
                else
                    State transfer: Execute action a
                end if
                If all targets failed **then** break;
            end for
        end while
    end for
    Obtain new reward function parameters (w_{lost}^p, b_{lost}^p): solve Eq.4 with S
    Obtain new policy π^p: solve Eq.5 with (w_{lost}^p, b_{lost}^p)
until all targets are successfully tracked.

Experiment evaluation



The MOT Challenge 2015 Multiple object tracking benchmark*

- 22 video sequences (11 for training and 11 for testing);
- overall contains 61440 object detections generated by the ACF detector;
- over 10 minutes tracking data annotations;
- having lots of variations in camera perspective, shaking and weather conditions, etc.;
- The evaluation results on test set must be obtained via the official evaluation server.

*https://motchallenge.net/

Experiment evaluation

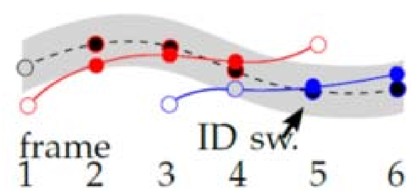
- The CLEAR metric for multiple object tracking evaluation:
  - Multiple Object Tracking Accuracy (MOTA, the higher the better)

$$MOTA = 1 - \frac{\sum_t (FN_t + FP_t + IDS_t)}{\sum_t GT_t}.$$

  - Multiple Object Tracking Precision (MOTP, evaluating object detector performance, the higher the better)

$$MOTP = \frac{\sum_{t,i} d_{t,i}}{\sum_t c_t}.$$

  - Mostly Tracked trajectories (MT, the higher the better)
  - Partially Tracked trajectories (PT)
  - Mostly Lost trajectories (ML, the lower the better)
  - Tracklet ID Switches (IDS, the lower the better)



Experiment evaluation

Dataset	Method	MOTA	MOTP	MT	PT	ML	IDS
TUD-Campus	Online MDP	51.53	72.02	1	7	0	13
	LP2D	32.00	72.50	0	6	2	10
	AL-poll-ReID	54.92	**72.68**	3	5	0	6
	AL-parallel-online	55.71	72.36	3	4	1	11
	AL-parallel-mixed	**57.61**	71.55	**3**	5	**0**	**5**
ETH-Sunnyday	Online MDP	35.79	**77.38**	5	13	12	59
	LP2D	32.10	77.00	2	13	15	34
	AL-poll-ReID	47.69	76.67	8	12	10	33
	AL-parallel-online	49.09	76.34	5	13	13	17
	AL-parallel-mixed	**51.08**	76.67	**8**	12	**10**	**16**
ETH-Pedcross2	Online MDP	9.13	71.98	2	24	107	80
	LP2D	4.40	72.80	0	16	117	214
	AL-parallel-online	12.22	71.52	3	23	107	67
	AL-poll-ReID	11.34	71.26	4	31	98	79
	AL-parallel-mixed	**13.40**	**72.51**	**5**	30	**97**	**64**
ADL-Rundle-8	Online MDP	**19.49**	72.74	6	13	9	**28**
	LP2D	1.80	73.10	2	17	9	194
	AL-poll-ReID	14.82	72.58	5	14	9	44
	AL-parallel-online	15.18	72.08	6	14	8	114
	AL-parallel-mixed	16.03	**72.75**	**6**	13	**9**	42
Venice-2	Online MDP	32.21	74.15	6	15	5	50
	LP2D	4.30	74.20	2	19	5	493
	AL-poll-ReID	31.17	**74.59**	4	17	5	38
	AL-parallel-online	33.19	74.06	6	14	6	40
	AL-parallel-mixed	**34.90**	74.39	**7**	15	**4**	**37**
KITTI-17	Online MDP	62.23	72.00	1	8	0	2
	LP2D	33.10	73.20	0	4	5	9
	AL-poll-ReID	62.87	71.67	1	8	0	3
	AL-parallel-online	62.91	71.78	1	8	0	3
	AL-parallel-mixed	**63.91**	**72.78**	**2**	6	**0**	**1**

- OnlineMDP: The original online MDP-based multiple object tracking algorithm. Reward function and policy function is learned via polling variant of AL;
- LP2D: The baseline method provided by the MOT Challenge 2015;
- AL-poll-ReID: Add person ReID module on the OnlineMDP;
- AL-parallel-online: Parallelized apprenticeship learning process over the AL-poll-ReID;
- *AL-parallel-mixed*: Add mixed style data association strategy on the basis of AL-Parallel-online.

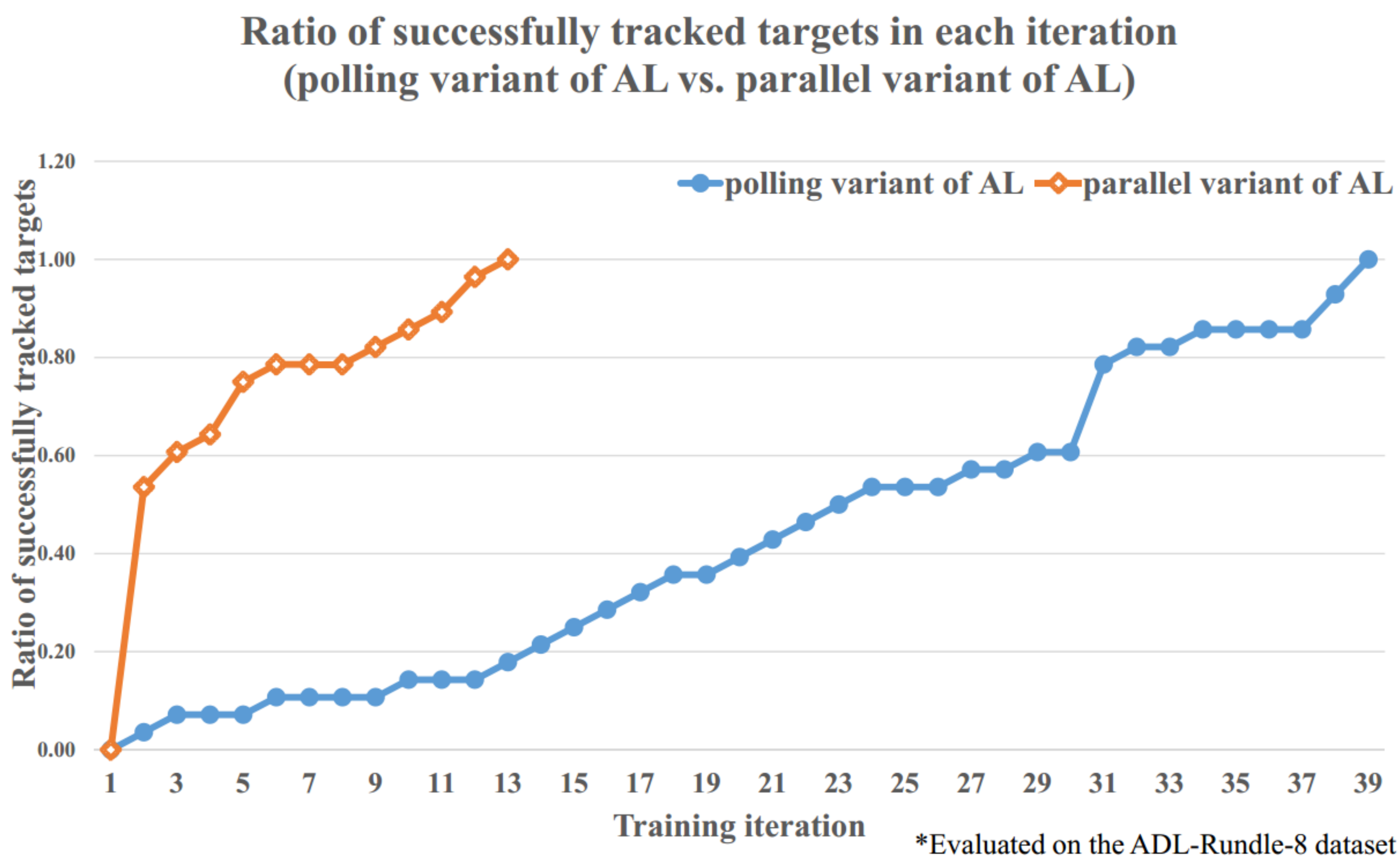
Experiment evaluation

- Obtains the state-of-the-art performance on MOT Challenge 2015 using public person detection.

Method	MOTA	MOTP	MT(%)	PT(%)	ML(%)	IDS
LP2D[1]	19.80	71.20	6.70%	52.10%	41.20%	1649
MotiCon[20]	23.10	70.90	10.40%	48.30%	41.30%	1018
LINF1[2]	24.50	71.30	5.50%	29.90%	64.60%	744
LP_SSVM[3]	25.20	**71.70**	5.80%	41.20%	53.00%	646
SCEA[4]	29.10	71.10	8.90%	43.80%	47.30%	604
OnlineMDP[15]	30.30	71.50	13.00%	48.60%	38.40%	690
Ours(AL-parallel-mixed)	**32.60**	71.30	**16.00%**	**49.60%**	**34.40%**	**580**

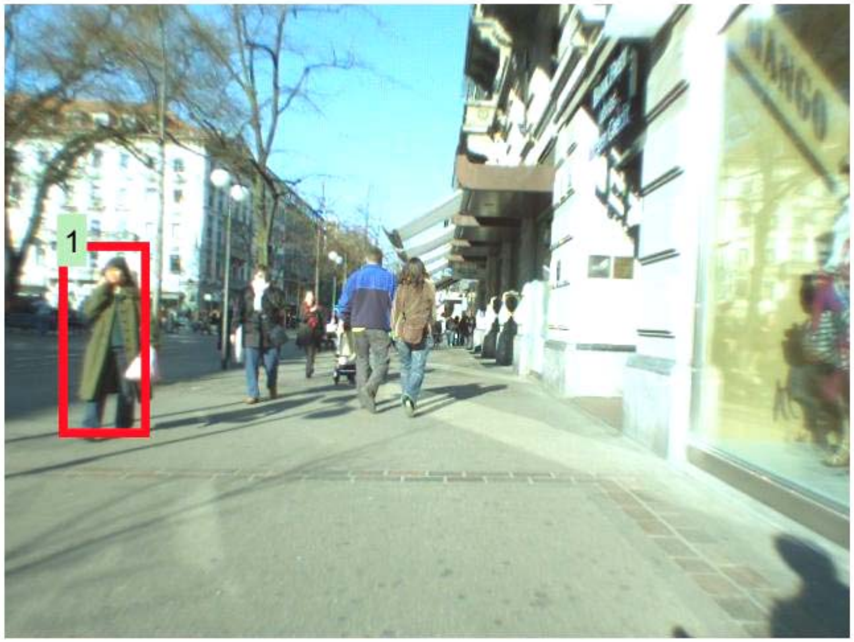
MOTA: multi-object tracking accuracy; MOTP: multi-object tracking precision; MT: mostly tracked;
PT: partially tracked; ML: mostly lost; IDS: ID switches

Experiment evaluation



Experiment evaluation

OnlineMDP



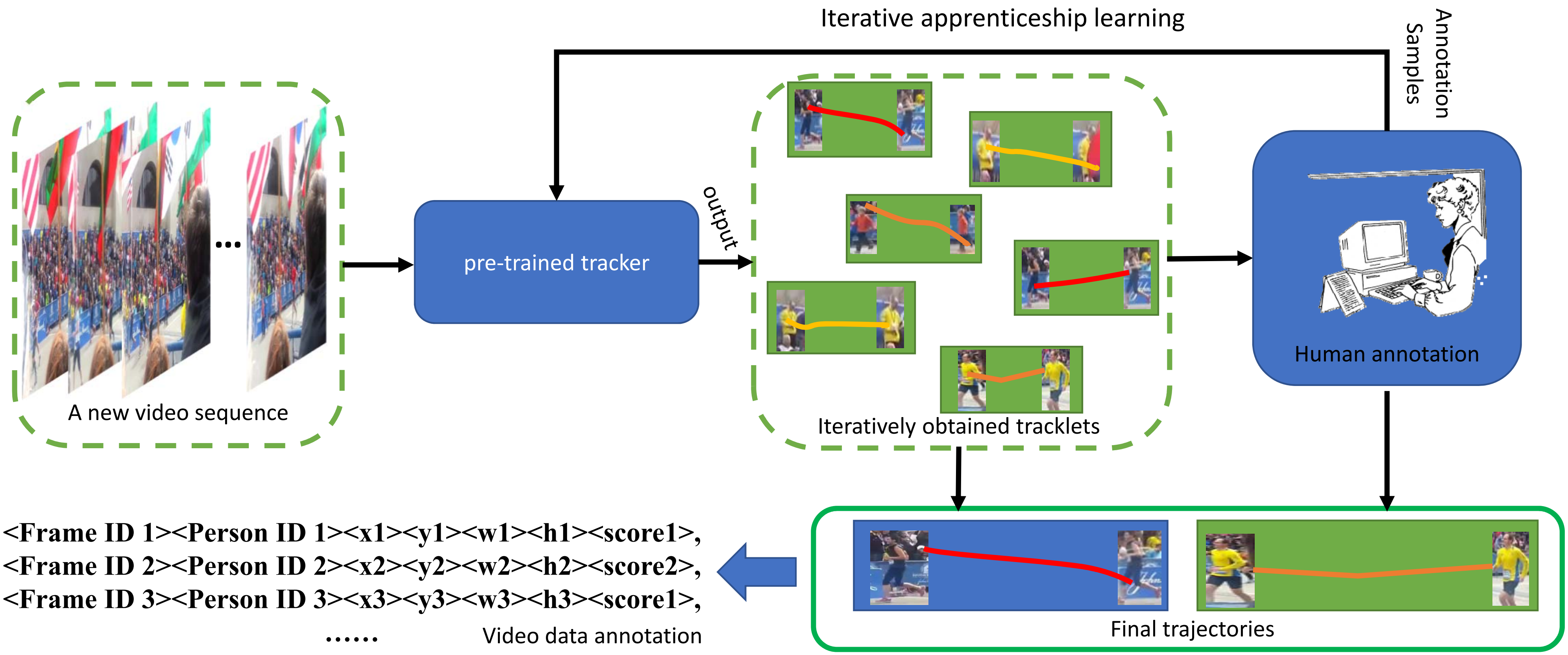
Ours (AL-parallel-mixed)

Xiang, Y., Alahi, A. and Savarese, S., 2015. Learning to track: Online multi-object tracking by decision making. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4705-4713).

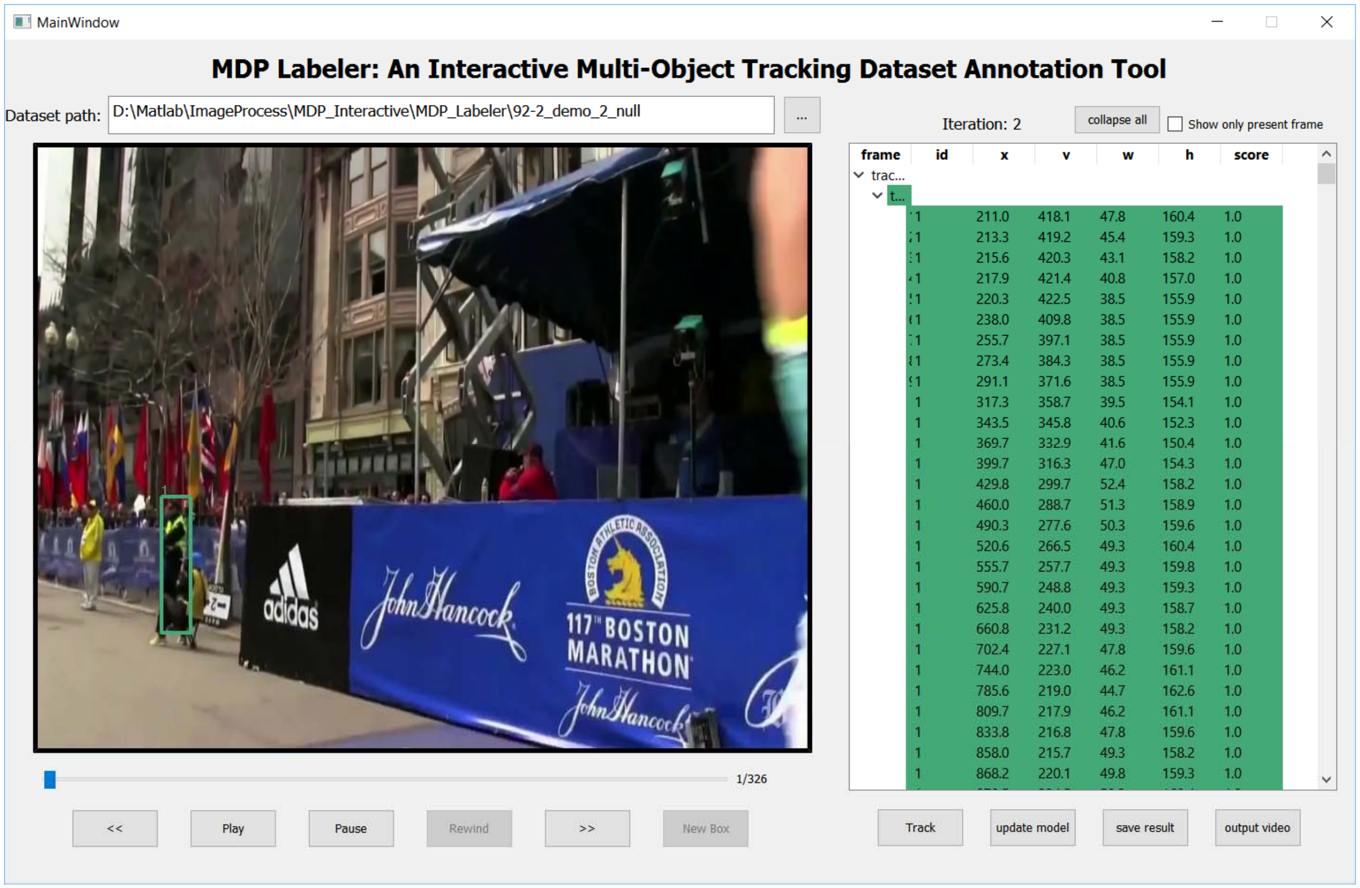
Interactive dataset annotation tool

The variation between training videos in MOT Challenge 2015 dataset and CMU Human Rights dataset.

Interactive dataset annotation tool



Interactive dataset annotation tool



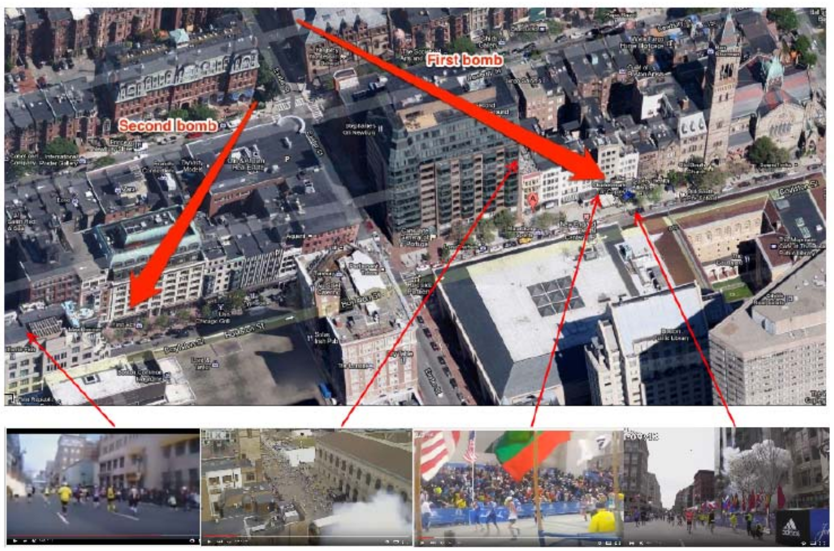
- Enable user to visualize tracking results;
- Easy to tune, merge, split existing trajectories or even add new bounding boxes and object trajectories;

Combine with apprenticeship learning

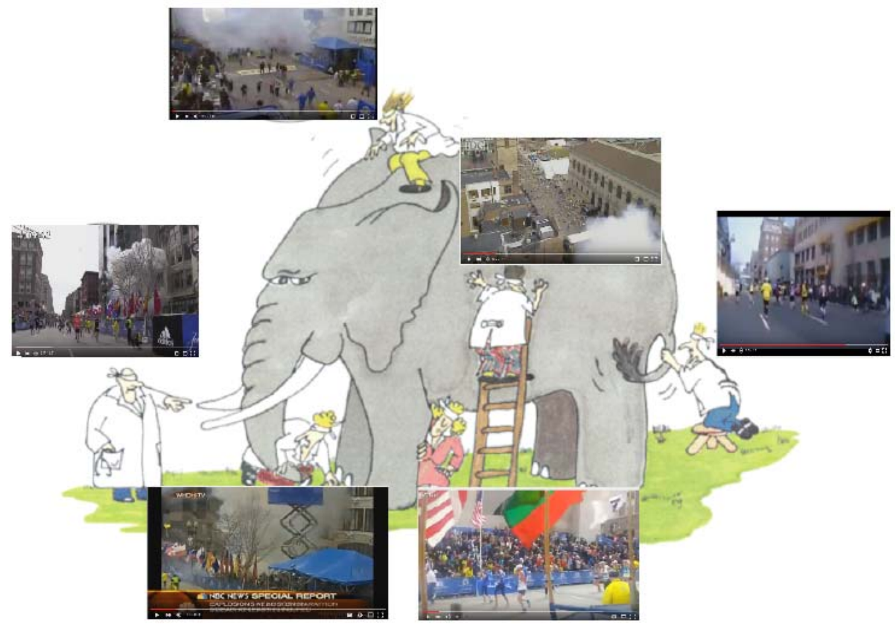
- The users modifications are recorded and serve as training data for the new MDP based tracker

https://github.com/grantlj/CMU_MDP_Interactive

3D event reconstruction demo



Boston Marathon 2013: event reconstruction based on large scale video data.

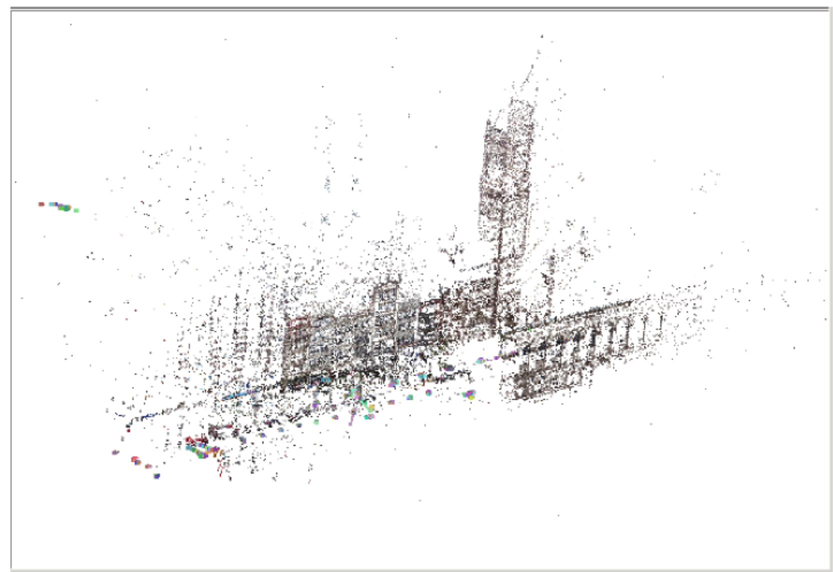


Blind men and an elephant: the metaphor for event reconstruction.

Chen, J., Liang, J., Lu, H., Yu, S.I. and Hauptmann, A., 2016. Videos from the 2013 Boston Marathon: An Event Reconstruction Dataset for Synchronization and Localization.

3D event reconstruction demo

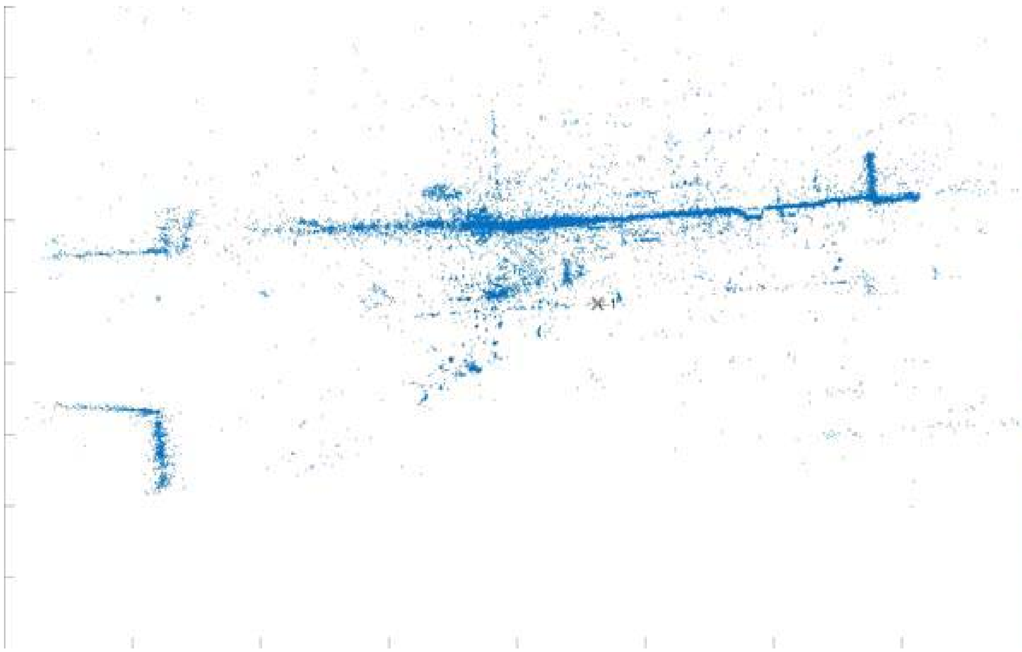
(a) gopro (b) cell phone (c) surveillance camera (d) professional camera



The application of multiple object tracking in event reconstruction: exhibiting person trajectories in 3D point clouds.

Chen, J., Liang, J., Lu, H., Yu, S.I. and Hauptmann, A., 2016. Videos from the 2013 Boston Marathon: An Event Reconstruction Dataset for Synchronization and Localization.

3D event reconstruction demo



Thank you!

Q&A

Carnegie Mellon University
Language Technologies Institute