Rewind to Track: Parallelized Apprenticeship Learning with Backward Tracklets

Jiang Liu\textsuperscript{1,2}, Jia Chen\textsuperscript{2}, De Cheng\textsuperscript{2}, Chenqiang Gao\textsuperscript{1}, Alexander G. Hauptmann\textsuperscript{2}

\textsuperscript{1}Chongqing University of Posts and Telecommunications
\textsuperscript{2}Carnegie Mellon University
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.

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.

Background

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.

Multiple object tracking based on online data association.

\[
\text{sim}(o_i^t, d_k^t) = w^T \Phi(o_i^t, d_k^t) + b
\]

Background

**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.

Multiple object tracking based on offline data association.

Background

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

**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.

Multiple object tracking based on offline data association.

Proposed methodology

**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: \[ \text{sim}(o_i^t, d_k^r) = \sum_{q=0}^{t'} w^T \Phi(o_i^t, d_{k,q}^t) + b, \]
Proposed methodology

Multiple object tracking based on Markov Decision Process (MDP). The agent’s state transition map of a tracking object.

Proposed methodology

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

- \(s \in S\) : 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 \(A\): \(\pi(s) \to a\), via maximizing the reward function;
- \(R(s, a)\) : real-valued reward function \(R(s, a) : S \times A \to R\), define a reward value by executing action \(a\) in state \(s\).
Proposed methodology

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.

The agent’s state transition map of a tracking object.
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 $\pi(s_{t_0}) \rightarrow a_{t_0}$. 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 $\mu(\pi) = E[\sum_{t=t_0}^\infty \gamma^t \Phi(s) | \pi]$, which is the feature expectation of the agent ($\gamma$ is the decay factor, $0 \leq \gamma \leq 1$).
Proposed methodology

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

• 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 $\mu(\pi) = E[\sum_{t=t_0}^{\infty} \gamma^t \Phi(s)|\pi]$, which is the feature expectation of the agent ($\gamma$ is the decay factor, $0 \leq \gamma \leq 1$).
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:

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

• 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

Proposed methodology

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

Q: How to train 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:
  \[ w(p) = \arg \max_w \sum_{j=1}^{N} w^T (\mu_j(\pi^{(p-1)}) - \mu_{E,j}(\pi^{(p-1)})) , \]

- IRL phase, parallelly updating policy function parameters:
  \[ \pi(p) = \arg \max_{\pi} \sum_{j=1}^{N} E[V_j(\pi(s))] |_{R(p)(s,a) = w(p) \cdot \phi(s)} , \]

- 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 \( O_t = \{ o_{i,j} \}_{j=1}^{N} \) and object detections \( D_t = \{ d_{i,j}^{(t)} \}_{j=1}^{N} \);

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

1: Initialization of reward function: \( w_{\text{lost}}^0 \leftarrow w_0, b_{\text{lost}}^0 \leftarrow b_0, S \leftarrow \emptyset \);
2: Initialization for each target \( o_{i,j} \) in each \( v_i \); set MDP of \( o_{i,j} \) in tracked after \( t_{\text{start}}(i,j) \) \( \leftarrow \) index of the first frame where \( o_{i,j} \) correctly detected;
3: \( p \leftarrow 0 \);
4: repeat
5: \( p \leftarrow p + 1 \);
6: for each video \( v_i \) in \( V \) do
7: \( t \leftarrow 1 \);
8: while \( t \leq \) last frame of \( v_i \) do
9: for target \( o_{i,j} \) in \( v_i \) which \( t_{\text{start}}(i,j) \geq t \) do
10: Follow policy \( \pi^{(p-1)} \), compute \( \mu_{t,q}^{(p)} \) as Eq.6, choose action \( a \);
11: Compute ground truth action \( o_{qt} \);
12: if state is lost and \( a \neq o_{qt} \) then
13: \( S \leftarrow S \cup \{ (o_{i,j}', d_{i,j}'^{(t')}; y_{i,j}), 1 \leq q \leq t' \} \);
14: else
15: Save failure position: \( t_{\text{start}}(i,j) \leftarrow t \);
17: State transfer: Execute action \( a \);
18: end if
19: end if
20: end for
21: end while
22: end for
23: Obtain new reward function parameters \( (w_{\text{lost}, b_{\text{lost}}}) \); solve Eq.4 with \( S \);
24: Obtain new policy \( \pi^{(p)} \); solve Eq.5 with \( (w_{\text{lost}, b_{\text{lost}}}) \);
25: until all targets are successfully tracked.
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:
  \[ w(p) = \arg \max \min \sum_{j=1}^{N} w^T (\mu_j (\pi^{(p-1)}) - \mu_{E,j} (\pi^{(p-1)})), \]
- IRL phase, parallelly updating policy function parameters:
  \[ \pi(p) = \arg \max_{\pi} \sum_{j=1}^{N} E[\pi_j (s_t)] | \tilde{r}(p)(s,a) = w(p) \phi(s)] \]
- Multiple agents feature updating:
  \[ \forall j, \mu_j^{(p)} = \mu_j (\pi(p)) \]

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)).
Experiment evaluation

- 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.

The MOT Challenge 2015 Multiple object tracking benchmark*

*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)
    \[ MOTP = \frac{\sum_t d_{t,i}}{\sum_t c_t}. \]
  - 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

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Method</th>
<th>MOTA</th>
<th>MOTP</th>
<th>MT</th>
<th>PT</th>
<th>ML</th>
<th>IDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUD-Campus</td>
<td>Online MDP</td>
<td>51.53</td>
<td>72.02</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>LP2D</td>
<td>32.00</td>
<td>72.50</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>AL-poll-ReID</td>
<td>54.92</td>
<td><strong>72.68</strong></td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-online</td>
<td>55.71</td>
<td>72.36</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-mixed</td>
<td>57.61</td>
<td>71.55</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>ETH-Sunnyday</td>
<td>Online MDP</td>
<td>35.79</td>
<td><strong>77.38</strong></td>
<td>5</td>
<td>13</td>
<td>12</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>LP2D</td>
<td>32.10</td>
<td>77.00</td>
<td>2</td>
<td>13</td>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>AL-poll-ReID</td>
<td>47.69</td>
<td>76.67</td>
<td>8</td>
<td>12</td>
<td>10</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-online</td>
<td>49.09</td>
<td>76.34</td>
<td>5</td>
<td>13</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-mixed</td>
<td>51.08</td>
<td>76.67</td>
<td>8</td>
<td>12</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>ETH-Pedcross2</td>
<td>Online MDP</td>
<td>9.13</td>
<td>71.98</td>
<td>2</td>
<td>24</td>
<td>107</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>LP2D</td>
<td>4.40</td>
<td>72.80</td>
<td>0</td>
<td>16</td>
<td>117</td>
<td>214</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-online</td>
<td>12.22</td>
<td>71.52</td>
<td>3</td>
<td>23</td>
<td>107</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>AL-poll-ReID</td>
<td>11.34</td>
<td>71.26</td>
<td>4</td>
<td>31</td>
<td>98</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-mixed</td>
<td>13.40</td>
<td>72.51</td>
<td>5</td>
<td>30</td>
<td>97</td>
<td>64</td>
</tr>
<tr>
<td>ADL-Rundle-8</td>
<td>Online MDP</td>
<td>19.49</td>
<td>72.74</td>
<td>6</td>
<td>13</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>LP2D</td>
<td>1.80</td>
<td>73.10</td>
<td>2</td>
<td>17</td>
<td>9</td>
<td>194</td>
</tr>
<tr>
<td></td>
<td>AL-poll-ReID</td>
<td>14.82</td>
<td>72.58</td>
<td>5</td>
<td>14</td>
<td>9</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-online</td>
<td>15.28</td>
<td>72.08</td>
<td>6</td>
<td>14</td>
<td>8</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-mixed</td>
<td>16.03</td>
<td><strong>72.75</strong></td>
<td>6</td>
<td>13</td>
<td>9</td>
<td>42</td>
</tr>
<tr>
<td>Venice-2</td>
<td>Online MDP</td>
<td>32.21</td>
<td>74.15</td>
<td>6</td>
<td>15</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>LP2D</td>
<td>4.30</td>
<td>74.20</td>
<td>2</td>
<td>19</td>
<td>5</td>
<td>493</td>
</tr>
<tr>
<td></td>
<td>AL-poll-ReID</td>
<td>31.71</td>
<td><strong>74.59</strong></td>
<td>4</td>
<td>17</td>
<td>5</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-online</td>
<td>33.19</td>
<td>74.06</td>
<td>6</td>
<td>14</td>
<td>6</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-mixed</td>
<td>34.90</td>
<td>74.39</td>
<td>7</td>
<td>15</td>
<td>4</td>
<td>37</td>
</tr>
<tr>
<td>KITTI-17</td>
<td>Online MDP</td>
<td>62.23</td>
<td>72.00</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>LP2D</td>
<td>33.10</td>
<td>73.20</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>AL-poll-ReID</td>
<td>62.87</td>
<td>71.67</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-online</td>
<td>62.91</td>
<td>71.78</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>AL-parallel-mixed</td>
<td>63.91</td>
<td><strong>72.78</strong></td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

- **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.

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

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

Ratio of successfully tracked targets in each iteration (polling variant of AL vs. parallel variant of AL)

*Evaluated on the ADL-Rundle-8 dataset
Experiment evaluation

Interactive dataset annotation tool

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

A new video sequence

pre-trained tracker

Iterative apprenticeship learning

Iteratively obtained tracklets

Human annotation

Annotation Samples

Video data annotation

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

Combine with apprenticeship learning

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

3D event reconstruction demo
Thank you!

Q&A