Learning to Track: Online Multi-Object Tracking by Decision Making

Yu Xiang\textsuperscript{1}, Alexandre Alahi\textsuperscript{2}, and Silvio Savarese\textsuperscript{2}
\textsuperscript{1}University of Washington, \textsuperscript{2}Stanford University
ICCV 2015
Multi-Object Tracking

Visual surveillance

Robot navigation

Sport Analysis

Autonomous driving
Batch Mode vs. Online Mode

• Batch Mode

• Online Mode
Tracking by Detection
Data Association

Tracks at time t-1

Detections at time t

time axis
Challenges

Noisy detection: false alarms and missing detections
Challenges

Occlusion
Similarity Function for Data Association

- Tracks at time t-1
- Detections at time t

<table>
<thead>
<tr>
<th>Time Axis</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
</tr>
<tr>
<td>0.8</td>
</tr>
</tbody>
</table>

- Ours
- Simple similarity measure + Powerful optimization

- Zhang et al., CVPR’08
- Berclaz et al., TPAMI’11
- Breitenstein et al., TPAMI’11
- Pirsiavash et al., CVPR’11
- Butt & Collins, CVPR’13
- Milan et al., TPAMI’14
- Etc.
Learning to Track

Different features/cues between targets and detections

Similarity = \( w_1 \phi_1(\text{Appearance}, \text{Location}) + \cdots + w_n \phi_n(\text{Motion}, \text{Etc.}) \)

Weights to combine different cues (to be learned)
Offline-learning vs. Online-learning
## Offline-learning vs. Online-learning

<table>
<thead>
<tr>
<th></th>
<th>Offline-learning</th>
<th>Online-learning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training time</strong></td>
<td>Before Tracking</td>
<td></td>
</tr>
<tr>
<td><strong>With supervision</strong></td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td><strong>Use history of the target</strong></td>
<td>×</td>
<td></td>
</tr>
</tbody>
</table>

- Li et al., CVPR’09
- Kim et al., ACCV’12
- Etc.
## Offline-learning vs. Online-learning

<table>
<thead>
<tr>
<th></th>
<th>Offline-learning</th>
<th>Online-learning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training time</strong></td>
<td>Before Tracking</td>
<td>During Tracking</td>
</tr>
<tr>
<td><strong>With supervision</strong></td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td><strong>Use history of the target</strong></td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>

- Song et al., ECCV’08
- Kuo et al., CVPR’10
- Bae et al., CVPR’14
- Etc.
Our Solution: Tracking by Decision Making

The target is tracked

The target is occluded

The target is tracked again
Inverse Reinforcement Learning

Markov Decision Process (MDP)

Supervision

Ground truth trajectory

tracked

lost

tracked
## Comparison between Different Learning Strategies

<table>
<thead>
<tr>
<th></th>
<th>Offline-learning</th>
<th>Online-learning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training time</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before Tracking</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>During Tracking</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td><strong>With supervision</strong></td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td><strong>Use history of the target</strong></td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>
Comparison between Different Learning Strategies

<table>
<thead>
<tr>
<th></th>
<th>Offline-learning</th>
<th>Online-learning</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training time</td>
<td>Before Tracking</td>
<td>During Tracking</td>
<td>Before Tracking</td>
</tr>
<tr>
<td>With supervision</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Use history of the target</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Outline

• Markov Decision Process (MDP) for a Single Target
• Online Multi-Object Tracking with MDPs
• Experiments
• Conclusion
Outline

• Markov Decision Process (MDP) for a Single Target

• Online Multi-Object Tracking with MDPs

• Experiments

• Conclusion
Markov Decision Process for a Single Target

Active

Tracked

Inactive

Lost

object detection
Markov Decision Process for a Single Target

object detection

Active

Tracked

Inactive

Lost
Markov Decision Process for a Single Target

object detection

Active

Tracked

Inactive

Lost
Markov Decision Process for a Single Target

- Active
  - Object detection
  - Tracked
  - Inactive
Markov Decision Process for a Single Target

Object detection

Active

Tracked

Inactive

Lost

Single object tracking

Template Tracking in Tracked States
Template Tracking in Tracked States
Template Tracking in Tracked States
Template Tracking in Tracked States
Template Tracking in Tracked States
Template Tracking in Tracked States

Frame 50

Frame 57
Template Tracking in Tracked States
Template Tracking in Tracked States

Frame 50

Frame 57

Tracked

Lost
Markov Decision Process for a Single Target

If lost for more than $T$ frames

Object detection

Active

Tracked

Inactive

Lost

If lost for more than $T$ frames
Data Association in Lost States
Learning the Similarity Function

$$\text{Similarity} = w_1 \phi_1(\text{blue}, \text{green}) + \cdots + w_n \phi_n(\text{blue}, \text{green}) + b$$

Inverse reinforcement learning: tracking objects in training videos!

Hard positive examples

Hard negative examples
Inverse Reinforcement Learning

Ground truth trajectory

Supervision

1

2

3

4

tracked

lost

time axis

t

t-1

t-2
Inverse Reinforcement Learning

Ground truth trajectory

Supervision

1
2
3
4

time axis
Inverse Reinforcement Learning

Ground truth trajectory

Supervision

Wrong decision!
Update your weights!

Negative example

time axis
Inverse Reinforcement Learning

Try it again

Ground truth trajectory

Supervision

1. Wrong decision!
2. Association to this one!
3. Update your weights!
4. Positive example

Ground truth trajectory

(t, 2)

tracked, lost, No association

time axis
Inverse Reinforcement Learning

Try it again

Ground truth trajectory

1

2

3

4

Good job!
Keep going!
No update of the weights

tracked
lost

time axis

Supervision

t-2 t-1 t
Markov Decision Process for a Single Target
Outline

• Markov Decision Process (MDP) for a Single Target

• Online Multi-Object Tracking with MDPs

• Experiments

• Conclusion
Ensemble MDPs for Online Multi-Object Tracking

MDP1
MDP2
MDP3

time axis

t-2 t-1 t
Step 1: Process tracked targets

MDP1
MDP2
MDP3

Time axis

t-2  t-1  t
Step 2: Process lost targets

Hungarian algorithm for lost targets
Step 3: Initialize new targets

Initialize new targets
Terminate detection

MDP1
MDP2
MDP3
Online Multi-Object Tracking with MDPs
Outline

• Markov Decision Process (MDP) for a Single Target

• Online Multi-Object Tracking with MDPs

• Experiments

• Conclusion
Experiments: Dataset

• Multiple Object Tracking Benchmark [1]
  • 11 training sequences
  • 11 test sequences
  • Object detections from the ACF detector [2]

Experiments: Analysis on Validation Set

• Contribution of different components
Experiments: Analysis on Validation Set

• Contribution of different components

object detection

Active

Tracked

Inactive

Lost

MOTA: multiple object tracking accuracy
Experiments: Analysis on Validation Set

• Contribution of different components

object detection

Active → Tracked → Lost → Inactive

MOTA: multiple object tracking accuracy
Experiments: Analysis on Validation Set

• Contribution of different components

- MOTA: multiple object tracking accuracy
Experiments: Analysis on Validation Set

- Contribution of different components

\[
\text{Similarity} = w_1 \phi_1(\text{similarity}) + \ldots + w_n \phi_n(\text{similarity}) + b
\]

\text{MOTA: multiple object tracking accuracy}
Experiments: Analysis on Validation Set

• Contribution of different components

\[
\text{Similarity} = w_1 \phi_1 \left( \begin{array}{c} \text{\textcircled{c}} \\ \text{\textcircled{c}} \\ \text{\textcircled{c}} \end{array} \right) + \ldots + w_n \phi_n \left( \begin{array}{c} \text{\textcircled{c}} \\ \text{\textcircled{c}} \\ \text{\textcircled{c}} \end{array} \right) + b
\]

\[\phi_1, \ldots, \phi_n\]

\[w_1, w_2, \ldots, w_n, b\]

\[\text{MOTA: multiple object tracking accuracy}\]
Experiments: Analysis on Validation Set

• Cross-domain tracking

**MOTA**: multiple object tracking accuracy

- TUD-Stadtmitte: 56.0
- ETH-Bahnhof: 44.8
- ADL-Rundle-6: 47.9
- KITTI-13: 53.2
- PETS09-S2L1: 49.0

Training Sequences

Testing sequences

TUD-Campus
Experiments: Analysis on Validation Set

- Cross-domain tracking

MOTA: multiple object tracking accuracy

<table>
<thead>
<tr>
<th>Training Sequences</th>
<th>Testing Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUD-Stadtmitte</td>
<td>50.0</td>
</tr>
<tr>
<td>ETH-Bahnhof</td>
<td>44.8</td>
</tr>
<tr>
<td>ADL-Rundle-6</td>
<td>47.9</td>
</tr>
<tr>
<td>KITTI-13</td>
<td>53.2</td>
</tr>
<tr>
<td>PETS09-S2L1</td>
<td>49.0</td>
</tr>
</tbody>
</table>
Experiments: Analysis on Validation Set

• Cross-domain tracking

<table>
<thead>
<tr>
<th>Training Sequences</th>
<th>Testing Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUD-Stadtmitte</td>
<td>50.0</td>
</tr>
<tr>
<td>ETH-Bahnhof</td>
<td>44.8</td>
</tr>
<tr>
<td>ADL-Rundle-6</td>
<td>47.9</td>
</tr>
<tr>
<td>KITTI-13</td>
<td>53.2</td>
</tr>
<tr>
<td>PETS09-S2L1</td>
<td>49.0</td>
</tr>
</tbody>
</table>

**MOTA**: multiple object tracking accuracy

<table>
<thead>
<tr>
<th>TUD-Campus</th>
<th>ETH-Sunnyday</th>
<th>ETH-Pedcross2</th>
<th>ADL-Rundle-8</th>
<th>Venice-2</th>
<th>KITTI-17</th>
</tr>
</thead>
<tbody>
<tr>
<td>40.8</td>
<td>14.0</td>
<td>20.0</td>
<td>30.8</td>
<td>60.8</td>
<td></td>
</tr>
<tr>
<td>43.4</td>
<td>13.3</td>
<td>22.6</td>
<td>30.8</td>
<td>60.3</td>
<td></td>
</tr>
<tr>
<td>48.2</td>
<td>11.5</td>
<td>26.1</td>
<td>29.8</td>
<td>57.8</td>
<td></td>
</tr>
<tr>
<td>47.5</td>
<td>13.9</td>
<td>20.9</td>
<td>32.1</td>
<td>59.9</td>
<td></td>
</tr>
<tr>
<td>42.1</td>
<td>11.5</td>
<td>22.1</td>
<td>29.4</td>
<td>61.2</td>
<td></td>
</tr>
</tbody>
</table>
## Experiments: Evaluation on Test Set

<table>
<thead>
<tr>
<th>Tracker</th>
<th>Tracking</th>
<th>Learning</th>
<th>MOTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP_NMS [1]</td>
<td>Batch</td>
<td>N/A</td>
<td>14.5</td>
</tr>
<tr>
<td>TC_ODAL [2]</td>
<td>Online</td>
<td>Online</td>
<td>15.1</td>
</tr>
<tr>
<td>TBD [3]</td>
<td>Batch</td>
<td>Offline</td>
<td>15.9</td>
</tr>
<tr>
<td>SMOT [4]</td>
<td>Batch</td>
<td>N/A</td>
<td>18.2</td>
</tr>
<tr>
<td>RMOT [5]</td>
<td>Online</td>
<td>N/A</td>
<td>18.6</td>
</tr>
<tr>
<td>CEM [6]</td>
<td>Online</td>
<td>N/A</td>
<td>19.3</td>
</tr>
<tr>
<td>SegTrack [7]</td>
<td>Batch</td>
<td>Offline</td>
<td>22.5</td>
</tr>
<tr>
<td>MotiCon [8]</td>
<td>Batch</td>
<td>Offline</td>
<td>23.1</td>
</tr>
<tr>
<td><strong>MDP (Ours)</strong></td>
<td>Online</td>
<td>Online</td>
<td><strong>30.3</strong></td>
</tr>
</tbody>
</table>

**MOTA**: multiple object tracking accuracy

---

[1] Pirsiavash et al., CVPR’11  
[2] Bae et al., CVPR’14  
[3] Geiger et al., TPAMI’14  
[4] Dicle et al., ICCV’13  
[5] Yoon et al., WACV’15  
[6] Milan et al., TPAMI’14  
[7] Milan et al., CVPR’15  
[8] Leal-Taixé et al., CVPR’14
Tracking Results
MDP [Ours]

MotiCon [Leal-Taixé et al., CVPR’14]
MDP [Ours]

MotiCon [Leal-Taixé et al., CVPR’14]
MDP [Ours]  
MotiCon [Leal-Taixé et al., CVPR’14]
Outline

• Markov Decision Process (MDP) for a Single Target
• Online Multi-Object Tracking with MDPs
• Experiments

• Conclusion
Conclusion

Single Object Tracking

- Active
- Inactive
- Lost

Object Detection
Data Association
Target Re-identification
Code

Learning to Track: Online Multi-Object Tracking by Decision Making

Yu Xiang make video 1 to N

3rd_party
qsub
.gitignore
LICENSE
LICENSE_TLD
LK.m
LK_associate.m
LK_crop_image_box.m
LK_initialize.m

hunagrian
qsub
remove files
Initial commit
add TLD license
clean up
add comments
add comments
add comments

7 months ago
8 months ago
2 months ago
10 months ago
2 months ago
8 months ago
2 months ago
2 months ago
2 months ago

Latest commit 744796b on Oct 6

HTTPS clone URL
https://github.com
You can use HTTPS or Subversion.

Clone to Desktop
Download ZIP
Thank you!