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

Yu Xiang$^{1,2}$, Alexandre Alahi$^1$, and Silvio Savarese$^1$

$^1$Stanford University, $^2$University of Michigan

ICCV 2015
Multi-Object Tracking

Visual surveillance

Sport Analysis

Robot navigation

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

- 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.

Ours

Simple similarity measure + Powerful optimization
Learning to Track

Different features/cues between targets and detections

\[ \text{Similarity} = w_1 \phi_1(\text{appearance}, \text{rectangle}) + \cdots + w_n \phi_n(\text{appearance}, \text{rectangle}) \]

Weights to combine different cues (to be learned)

- Appearance
- Location
- Motion
- Etc.
Offline-learning vs. Online-learning
Offline-learning vs. Online-learning

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

Ground truth trajectory

Markov Decision Process (MDP)

Supervision

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>x</td>
</tr>
<tr>
<td>During Tracking</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td><strong>With supervision</strong></td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td><strong>Use history of the target</strong></td>
<td>x</td>
<td>✓</td>
</tr>
</tbody>
</table>

- **Offline-learning**
  - Before Tracking: ✓
  - During Tracking: x

- **Online-learning**
  - Before Tracking: x
  - During Tracking: ✓
## 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

19
Markov Decision Process for a Single Target
Markov Decision Process for a Single Target

Active

Tracked

Inactive

Lost

object detection
Markov Decision Process for a Single Target
Markov Decision Process for a Single Target

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
Template Tracking in Tracked States
Template Tracking in Tracked States

Frame 50

Frame 57

31
Markov Decision Process for a Single Target

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

\[ \text{Similarity} = w_1 \phi_1(\text{\textbullet}, \text{\textcolor{green}{\square}}) + \cdots + w_n \phi_n(\text{\textbullet}, \text{\textcolor{green}{\square}}) + b \]

Inverse reinforcement learning: tracking objects in training videos!

Hard positive examples

Hard negative examples
Inverse Reinforcement Learning

Ground truth trajectory

Supervision

tracked

lost

time axis

t-2
t-1
t
Inverse Reinforcement Learning

Ground truth trajectory

Supervision

1

2

3

4

t-2

t-1

t

time axis
Inverse Reinforcement Learning

Ground truth trajectory

Supervision

Wrong decision!
Update your weights!

Negative example
Inverse Reinforcement Learning

- t-2
- t-1
- t

Ground truth trajectory

Try it again

.Get it wrong!

Association to this one!

No association

Update your weights!

Positive example

Supervision

Try it again

Wrong decision!

Positive example

Trace track

lost

tracked
Inverse Reinforcement Learning

Try it again

Ground truth trajectory

1. Supervision
2. Good job!
3. Keep going!
4. No update of the weights

Try it again

tracked

lost

t-2 t-1 t

time axis

Good job!
Keep going!
No update of the weights
Markov Decision Process for a Single Target

object detection

Active

Tracked

Inactive

Lost
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

t-2 t-1 t

time axis
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

MDP1
MDP2
MDP3

Initialize new targets
Terminate detection

t-2 t-1 t

time axis
Online Multi-Object Tracking with MDPs

MDP1 → Tracked → Lost → Tracked

MDP2 → Tracked → Lost → Tracked

MDP3 → Tracked → Tracked → Tracked
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

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

• Contribution of different components

object detection

Framework Analysis

Active

Tracked

Inactive

Lost

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

- Contribution of different components

![Diagram showing object detection process]

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

• Contribution of different components

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

\(\phi\): Similarity

\(w_i\): Weight

\(\phi_i\): Function

\(b\): Bias

\textbf{Framework Analysis}

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

- Contribution of different components

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

*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
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
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>TUD-Campus</td>
</tr>
<tr>
<td>ETH-Bahnhof</td>
<td>ETH-Sunnyday</td>
</tr>
<tr>
<td>ADL-Rundle-6</td>
<td>ADL-Rundle-8</td>
</tr>
<tr>
<td>KITTI-13</td>
<td>Venice-2</td>
</tr>
<tr>
<td>PETS09-S2L1</td>
<td>KITTI-17</td>
</tr>
</tbody>
</table>

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

- MOTA values for testing sequences:
  - TUD-Campus: 46.8, 14.0, 20.0, 30.8, 60.8
  - ETH-Sunnyday: 43.4, 13.3, 22.6, 30.8, 60.3
  - ADL-Rundle-8: 48.2, 11.5, 26.1, 29.8, 57.8
  - Venice-2: 47.5, 13.9, 20.9, 32.1, 59.9
  - KITTI-17: 42.1, 11.5, 22.1, 29.4, 61.2
## 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**
- **Tracked**
- **Inactive**
- **Lost**

Object Detection

Data Association
Target Re-identification
## Code

### GitHub: yuxng / MDP_Tracking

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

- **112 commits**
- **1 branch**
- **0 releases**
- **1 contributor**

### Branch: master

<table>
<thead>
<tr>
<th>File</th>
<th>Content</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd_party</td>
<td>hungarian</td>
<td>7 months ago</td>
</tr>
<tr>
<td>qsub</td>
<td>qsub</td>
<td>8 months ago</td>
</tr>
<tr>
<td>.gitignore</td>
<td>remove files</td>
<td>2 months ago</td>
</tr>
<tr>
<td>LICENSE</td>
<td>Initial commit</td>
<td>10 months ago</td>
</tr>
<tr>
<td>LICENSE_TLD</td>
<td>add TLD license</td>
<td>2 months ago</td>
</tr>
<tr>
<td>LK.m</td>
<td>clean up</td>
<td>8 months ago</td>
</tr>
<tr>
<td>LKassociate.m</td>
<td>add comments</td>
<td>2 months ago</td>
</tr>
<tr>
<td>LK_crop_image_box.m</td>
<td>add comments</td>
<td>2 months ago</td>
</tr>
<tr>
<td>LK_initialize.m</td>
<td>add comments</td>
<td>2 months ago</td>
</tr>
</tbody>
</table>

### HTTPS clone URL

https://github.com

You can clone with HTTPS or Subversion.
Thank you!