Online Multi-target Tracking using Recurrent Neural Networks - Milan et al.

Seminar Biomedical Image Analysis
Freiburg 01.08.16
1. Introduction
   a. Motivation
   b. Problem - Challenges - Ideas
   c. Digression: Bayes Filter

2. Related work, Contribution

3. Approach
   a. Overview
   b. RNN as Bayes Filter + Track Manager
      i. Digression to RNN
   c. LSTM for Data-Association
      i. Digression to LSTM

4. Experiments and Implementation

5. Conclusion
Introduction: Motivation

2D MOT (Multi Object Tracking Challenge) 2015

Source: https://motchallenge.net/
Introduction: Problem

Problem: Locate multiple targets of interest in a video sequence over time

Source: https://motchallenge.net/
Introduction: Problem - Challenges

Challenges: Dynamic number of object detections, Data association, State Estimation

Image source: https://motchallenge.net/
Challenges: Dynamic number of object detections, Data association, State Estimation

Solution

● Sequences and RNN, RNN as Bayes Filter!
● Data association can also be “learned”!
● Existence probability
What is $x_t$ (current estimate) given $x_{0:t-1}$ (previous estimates) and $z_t$ (measurement)??
Predict using previous estimate and motion model!
What to trust more: The prediction or measurement?
Correct (update) using the measurement!
Which observation belongs to which track?
• Variants of MHT, JPDA used
  • Various simplified models (Linear programs etc.)
  • Numerous numerical optimization techniques
Related Work

- Variants of MHT, JPDA used
  - Various simplified models (Linear programs etc.)
  - Numerous numerical optimization techniques
- Little work on using Deep Learning to Multi-Object tracking: chiefly due to **unavailability of training data**
Related Work

- Variants of MHT, JPDA used
  - Various simplified models (Linear programs etc.)
  - Numerous numerical optimization techniques
- Little work on using Deep Learning to Multi-Object tracking: chiefly due to **unavailability of training data**
- RNN promising, mainly used for language processing
- Issues:
  - Multi dimensional space
  - Includes continuous and discrete variables
  - Multiple outputs possible
Main Contributions

- **Bayesian filter** by unified RNN approach
  - Model - free approach
  - Linear, nonlinear or higher order dependencies
- **Data Association** learned from data
- Generated synthetic training data
- Qualitative and quantitative result presented
Approach: Outline

Source: Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
Source: Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
Approach: RNN as Bayes Filter, Track manager

Current state: $x_t$
Measurements: $z_t$
Existence probability: $\epsilon_t$
Data Association: $A_t$

RNN block

Prediction

Update

Birth / Death

Updated state: $x_{t+1}$
Updated Existence probability: $\epsilon_{t+1}$
Approach: RNN as Bayes Filter, Track manager

- Current state: $x_t$
- Measurements: $z_t$
- Existence probability: $\epsilon_t$
- Data Association: $A_t$

**RNN block**

- Prediction
- Update
- Birth / Death

- Updated state: $x_{t+1}$
- Updated Existence probability: $\epsilon_{t+1}$
Approach: RNN as Bayes Filter, Track manager

Current state $x_t$

Measurements $z_t$

Existence probability $\epsilon_t$

Data Association $A_t$

RNN block

Prediction

Update

Birth / Death

Updated state $x_{t+1}$

Updated Existence probability $\epsilon_{t+1}$
RNNs are neural networks with loops!

Source: colah.github.io/posts/2015-08-Understanding-LSTMs
Digression: RNN

More intuitive way of seeing an RNN!

Source: colah.github.io/posts/2015-08-Understanding-LSTMs

Akhil Thomas
Digression: RNN

A simple RNN!

Source: colah.github.io/posts/2015-08-Understanding-LSTMs
Approach: RNN as Bayes Filter, Track manager ...

RNN in their approach

Source: Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
Approach: RNN as Bayes Filter, Track manager ...

- Current state: $x_t$
- Measurements: $z_t$
- Existential probability: $\epsilon_t$
- Data Association: $A_t$
- Prediction: $x^*_t+1$
- Update: $x_{t+1}$
- Birth / Death
- Updated Existential probability: $\epsilon_{t+1}$
- Absolute difference to $\epsilon_t$: $\epsilon^*_t+1$

RNN block

Akhil Thomas
Approach: RNN as Bayes Filter, Track manager ...

RNN block

Current state $x_t$

Hidden state $h_t$

Prediction

Predicted state $x_{t+1}^*$

Updated hidden state $h_{t+1}$
Approach: RNN as Bayes Filter, Track manager ...

- Current state: $x_t$
- Prediction: $\hat{x}_t$
- Measurements: $z_t$
- Update: $\tilde{x}_{t+1}$
- Data Association: $A_t$
- Updated Existential probability: $\epsilon_{t+1}$
- Updated state: $x_{t+1}$
Approach: RNN as Bayes Filter, Track manager ...

RNN block

Birth / Death

Existential probability $\epsilon_t$

Absolute difference to $\epsilon_t$

$\epsilon^*_{t+1}$
\[ \mathcal{L}(x^*, x, \mathcal{E}, \tilde{x}, \tilde{\mathcal{E}}) = \frac{\lambda}{ND} \sum_{\text{prediction}} \|x^* - \tilde{x}\|^2 + \frac{\kappa}{ND} \|x - \tilde{x}\|^2 + \nu \mathcal{L}_\mathcal{E} + \xi \mathcal{E}^*, \]

\text{Source: Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.}
Approach: Outline ...

Source: Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
Approach: LSTM for Data Association

LSTM block for an individual track $i$

- Current cell state: $c_t$
- Feature vector: $C_t = \|z_t - x_t\|_2$
- Hidden state: $h_t$
- Data Association: $A_{t+1}^i$
- Updated cell state: $c_{t+1}$
- Updated hidden state: $h_{t+1}$
Long term dependency!

Source: colah.github.io/posts/2015-08-Understanding-LSTMs
Digression: LSTM

- A simple LSTM
- Cell state -- Conveyor belt!

Source: colah.github.io/posts/2015-08-Understanding-LSTMs
Approach: LSTM for Data Association ...

Source: Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
Experiments: Training data generation

- Few training data available
- Perturbation of real data
  - Mirroring
  - Translation
  - Rotation
- Sampling from a generative model
  - Learn trajectory model from each available training sequence
- Physically-based trajectory generation
  - Simulating real world motion and cameras
Platform: Lua and Torch7

Network size:
- RNN: 1 layer, 300 hidden units
- LSTM: 2 layers, 500 hidden units

Optimization:
- RMSprop
- Convergence under 200,000 iterations

Data:
- 100,000 20-frame long sequences
- Mini batches of 10 samples per batch
- Normalized to [-0.5, 0.5]
Experiments and Results

Source: Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
Experiments and Results

Source: https://motchallenge.net/
## Experiments and Results

**Table 1.** Tracking results on the MOTChallenge training dataset. *Denotes offline post-processing.

<table>
<thead>
<tr>
<th>Method</th>
<th>Rcll</th>
<th>Prcn</th>
<th>MT</th>
<th>ML</th>
<th>FP</th>
<th>FN</th>
<th>IDs</th>
<th>FM</th>
<th>MOTA</th>
<th>MOTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalman-HA</td>
<td>28.5</td>
<td>79.0</td>
<td>32</td>
<td>334</td>
<td>3,031</td>
<td>28,520</td>
<td>685</td>
<td>837</td>
<td>19.2</td>
<td>69.9</td>
</tr>
<tr>
<td>Kalman-HA2*</td>
<td>28.3</td>
<td>83.4</td>
<td>39</td>
<td>354</td>
<td>2,245</td>
<td>28,626</td>
<td>105</td>
<td>342</td>
<td>22.4</td>
<td>69.4</td>
</tr>
<tr>
<td>JPDA$_m^*$</td>
<td>30.6</td>
<td>81.7</td>
<td>38</td>
<td>348</td>
<td>2,728</td>
<td>27,707</td>
<td>109</td>
<td>380</td>
<td>23.5</td>
<td>69.0</td>
</tr>
<tr>
<td>RNN_HA</td>
<td>37.8</td>
<td>75.2</td>
<td>50</td>
<td>267</td>
<td>4,984</td>
<td>24,832</td>
<td>518</td>
<td>963</td>
<td>24.0</td>
<td>68.7</td>
</tr>
<tr>
<td>RNN_LSTM</td>
<td>37.1</td>
<td>73.5</td>
<td>50</td>
<td>260</td>
<td>5,327</td>
<td>25,094</td>
<td>572</td>
<td>983</td>
<td>22.3</td>
<td>69.0</td>
</tr>
</tbody>
</table>

*Source:* Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
### Table 2. Tracking results on the MOTChallenge test dataset. *Denotes an offline (or delayed) method.

<table>
<thead>
<tr>
<th>Method</th>
<th>MOTA</th>
<th>MOTP</th>
<th>FAR</th>
<th>MT%</th>
<th>ML%</th>
<th>FP</th>
<th>FN</th>
<th>IDs</th>
<th>Frag.</th>
<th>FPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDP [48]</td>
<td>30.3%</td>
<td>71.3%</td>
<td>1.7</td>
<td>13.0</td>
<td>38.4</td>
<td>9,717</td>
<td>32,422</td>
<td>680</td>
<td>1,500</td>
<td>1.1</td>
</tr>
<tr>
<td>JPDA(_m)* [13]</td>
<td>23.8%</td>
<td>68.2%</td>
<td>1.1</td>
<td>5.0</td>
<td>58.1</td>
<td>6,373</td>
<td>40,084</td>
<td>365</td>
<td>869</td>
<td>32.6</td>
</tr>
<tr>
<td>TC_ODAL [49]</td>
<td>15.1%</td>
<td>70.5%</td>
<td>2.2</td>
<td>3.2</td>
<td>55.8</td>
<td>12,970</td>
<td>38,538</td>
<td>637</td>
<td>1,716</td>
<td>1.7</td>
</tr>
<tr>
<td>RNN_LSTM</td>
<td>19.0%</td>
<td>71.0%</td>
<td>2.0</td>
<td>5.5</td>
<td>45.6</td>
<td>11,578</td>
<td>36,706</td>
<td>1,490</td>
<td>2,081</td>
<td>165.2</td>
</tr>
</tbody>
</table>

**Source:** Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
Conclusions

- RNN’s and its relation to sequences makes it look promising
- Their approach showed that RNN can be utilized to design a Bayes Filter
- LSTM for Data Association not a trivial task
- First approach that employs end-to-end training for multi-target tracking
1. Online Multi-target Tracking using Recurrent Neural Networks, A. Milan et al.
2. https://motchallenge.net/
3. colah.github.io/posts/2015-08-Understanding-LSTMs
Thanks :)