Graph Networks for Multiple Object Tracking

Jiahe Li, Xu Gao, Tingting Jiang

https://github.com/yinizhizhu/GNMOT.
Motivation

☐ Most graph models are static
  ■ Nodes and edges are fixed

☐ Graph Network
  ■ Has the ability of reasoning
  ■ Nodes and edges will be updated iteratively and reasonably
Contributions

- We propose a new near-online MOT method with an end-to-end graph network framework followed by strategies for handling missing detections.

- The updating mechanism is carefully designed in our graph networks.

- The proposed method achieves encouraging performance.
Graph Network

  - Graph network has the ability of reasoning

  - General graph network framework
  - The node, the edge and the global variable
  - Updating modules for each component
Our 4-step graph network
Our 4-step graph network

Edges

Edge Updating Module
Our 4-step graph network

Nodes

- Node Updating Module

Edges

- Edge Updating Module
Our 4-step graph network

Edges → Nodes

Edge Updating Module I

Node Updating Module

Edges

Edge Updating Module II
Our 4-step graph network

Edges → Nodes → Global → Edges

- Edge Updating Module I
- Node Updating Module
- Global Updating Module
- Edge Updating Module II
The pipeline of our method

- Appearance Graph Network
- Motion Graph Network
Weighted Strategy

\[ S = \alpha AGN + (1 - \alpha) MGN \]

AGN and MGN denote the appearance similarity and the motion similarity respectively.
Missing Detection Handling

1. Single Object Tracker

Detection Recovery Strategy:
- t-2
- t-1
- t

Frame
## Main Results

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Detection</th>
<th>Methods</th>
<th>MOTA</th>
<th>IDF1</th>
<th>MT %</th>
<th>ML %</th>
<th>FP</th>
<th>FN</th>
<th>IDS</th>
<th>FM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOT16</td>
<td>Public</td>
<td>LINF, ECCV 2016</td>
<td>41.0</td>
<td>45.7</td>
<td>11.6%</td>
<td>51.3%</td>
<td>7896</td>
<td>99224</td>
<td>430</td>
<td>963</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MHT_bLSTM*, ECCV 2018</td>
<td>42.1</td>
<td>47.8</td>
<td>14.9%</td>
<td>44.4%</td>
<td>11637</td>
<td>93172</td>
<td>753</td>
<td>1156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NOMT, ICCV 2015</td>
<td>46.4</td>
<td>53.3</td>
<td>18.3%</td>
<td>41.4%</td>
<td>9753</td>
<td>87565</td>
<td>359</td>
<td>504</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ours without SOT</td>
<td>47.4</td>
<td>42.6</td>
<td>14.5%</td>
<td>34.4%</td>
<td>7795</td>
<td>86178</td>
<td>1931</td>
<td>3389</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ours</td>
<td>47.7</td>
<td>43.2</td>
<td>16.1%</td>
<td>34.3%</td>
<td>9518</td>
<td>83875</td>
<td>1907</td>
<td>3376</td>
</tr>
<tr>
<td></td>
<td>Private</td>
<td>Ours without SOT</td>
<td>58.4</td>
<td>54.8</td>
<td>27.3%</td>
<td>23.2%</td>
<td>5731</td>
<td>68630</td>
<td>1454</td>
<td>1730</td>
</tr>
<tr>
<td>MOT17</td>
<td>Public</td>
<td>MHT_bLSTM*, ECCV 2018</td>
<td>47.5</td>
<td>51.9</td>
<td>18.2%</td>
<td>41.7%</td>
<td>25981</td>
<td>268042</td>
<td>2069</td>
<td>3124</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ours without SOT</td>
<td>50.1</td>
<td>46.3</td>
<td>18.6%</td>
<td>33.3%</td>
<td>25210</td>
<td>250761</td>
<td>5470</td>
<td>8113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ours</td>
<td>50.2</td>
<td>47.0</td>
<td>19.3%</td>
<td>32.7%</td>
<td>29316</td>
<td>246200</td>
<td>5273</td>
<td>7850</td>
</tr>
</tbody>
</table>

Table 1. Experiments on MOT16 and MOT17 test set. The best result in each metric is highlighted in bold, and the second best result is underlined. * indicates the use of additional training data.
Thanks