Attributed Graphs for Tracking Multiple Objects in Structured Sports Videos

Henrique Morimitsu*, Roberto M. Cesar-Jr.*, Isabelle Bloch**

* University of São Paulo
** LTCI, CNRS, Télécom ParisTech, Université Paris - Saclay

CVsports - ICCV - 2015
Multi-object tracking

- Follow many objects simultaneously
- Applications:
  - Surveillance
  - Activity analysis
  - Automatic navigation

(Shu et al., 2012)
Object tracking in sports

Tracking challenges

- Similar appearance
- Occlusion
- Abrupt motion
- Camera cut
Tracking with structural properties

- Structural properties may solve the problem
  - Distance
  - Angle

- Some scenes present structural patterns
Attributed Relational Graphs

Track objects with particle filter

\[ M_A = \begin{pmatrix}
0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 0
\end{pmatrix} \]
Proposed framework

Training

Learn model graph Gm

First frame

Generate new object candidates

Choose best subset of trackers

Generate scene graphs Gs and compute scores

Remove bad trackers
Proposed framework

Learn model graph $G_m$

Histograms
PDF
Color model
Proposed framework

- Sample distances and angles from model PDF
- Generate candidate positions
Proposed framework

- Score candidates according to the model graph
- Greedy approach
- Choose best candidates

\[ f(i, G_s, G_m) = x \cdot a(i, G_s) + (1 - x) \cdot s(i, G_s, G_m) \]

Scene graphs \( G_s \)

\( i \Rightarrow \) frame index

score function

appearance

structure
Results - Datasets

- Sports videos with structure

Youtube table tennis

ACASVA* badminton

* (de Campos et al, 2011)
Results - Structure weight

CERR = Center error

Euclidean distance between ground truth and tracking results

\[ f(i, Gs, Gm) = \chi a(i, Gs) + (1 - \chi) s(i, Gs, Gm) \]
Results

Demo video
Results - Baseline

Comparison with other methods

- Particle Filter + Graph (Ours)
- Particle Filter
- SPOT (Zhang and van der Maaten, 2014)
- STRUCK (Hare et al, 2011)

HITT = Hit team ratio

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Method</th>
<th>CERR</th>
<th>HITT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Youtube</td>
<td>PF + Graph</td>
<td>63</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>PF</td>
<td>107</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>SPOT</td>
<td>83</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>STRUCK</td>
<td>159</td>
<td>0.42</td>
</tr>
<tr>
<td>ACASVA</td>
<td>PF + Graph</td>
<td>50</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>PF</td>
<td>58</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>SPOT</td>
<td>59</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>STRUCK</td>
<td>74</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Results - Baseline
Conclusions

- Structural information improves tracking
  - Recover tracking after occlusion or abrupt motion
- Structure + appearance can be coupled for better results
- Significantly reduces drift on longer videos
- Assumes structure exists
Future work

- Change PF to more robust tracker
- Make the method more self-adaptive
- Model learned online
Acknowledgements

We would like to thank:

- FAPESP grants:
  - 2011/50761-2
  - 2012/09741-0
  - 2013/08258-7
- CNRS/FAPESP 227332
- CNPq
- CAPES
- NAP eScience - PRP - USP
THANK YOU

Questions?
References


References


References
