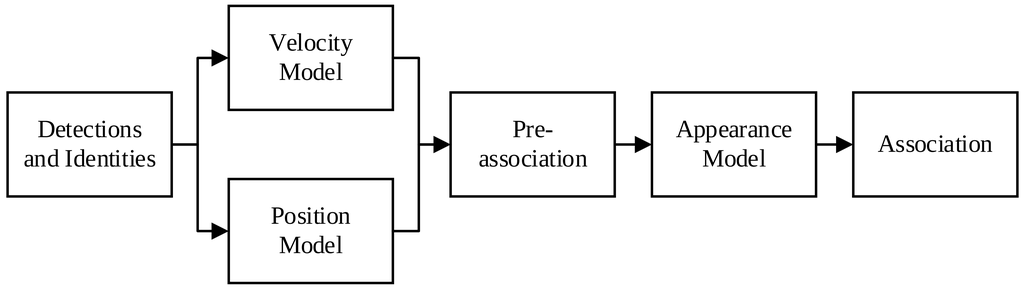
LSTM MULTIPLE OBJECT TRACKER COMBINING MULTIPLE CUES

Yiming Liang, Yue Zhou
Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai, China

Outline

Shortcomings of traditional multi-object tracking methods:
Either time-consuming or not accurate enough
Advantages of proposed multi-object tracking methods:
Give consideration to both speed and tracking accuracy by LSTM networks

Overall Architecture



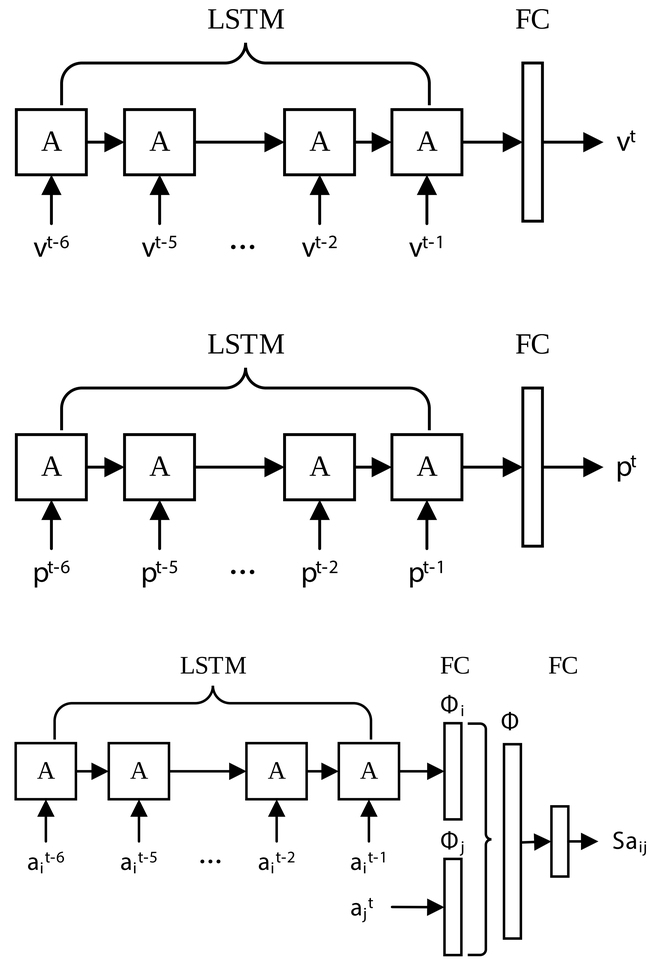
Match detections and identities only by velocity and position models to save time. If necessary, run the appearance model to make the final decision. LSTM networks are used to learn long term and short term features of trajectories.

Experimental results

Tracker	MOTA	IDF1	MT	ML	FP	FN	ID Sw.	Frag	Hz
GMPHD_KCF	39.6	36.6	8.8	43.3	0,903	284,228	5,811	7,414	3.3
EAMTT	42.6	41.8	12.7	42.7	30,711	288,474	4,488	5,720	**11.5**
IOU17	45.5	39.4	15.7	40.5	**19,993**	281,643	5,988	7,404	**1,522.9**
EDMT17	50.0	**51.3**	**21.6**	**36.3**	32,279	**247,297**	**2,264**	3,260	0.6
MHT DAM	**50.7**	47.2	20.8	36.9	22,875	252,889	**2,314**	**2,865**	0.9
Ours	**50.3**	**47.9**	**21.8**	**36.2**	**22,204**	**249,342**	3,243	**3,155**	1.9

Compare with recently multi-object tracking methods: GMPHD_KCF (AVSS2017), EAMTT (ECCV2016), IOU17 (AVSS2017), EDMT17 (CVPR2017), MHT DAM (ICCV2015)

Velocity/Position/Appearance Model



- Velocity model input: the differences of coordinates between two frames.
- Position model input: the IOU values of two bounding boxes.
- Appearance model input: the feature vectors Extracted from a CNN.