Multi-target tracking in team-sports videos via multi-level context-conditioned latent behaviour models

Jingjing Xiao[1]
shine636363@sina.com
Rustam Stolkin[2]
r.stolkin@bham.ac.uk
Ales Leonardis[3]
a.leonardis@cs.bham.ac.uk

[1] School of Electronics, Electrical and Computer Engineering, University of Birmingham, Birmingham, UK
[2] School of Mechanical Engineering,University of Birmingham, Birmingham, UK
[3] School of Computer Science,University of Birmingham, Birmingham, UK

Sports team tracking poses challenges not present in conventional pedestrian tracking: motion is erratic and players wear similar uniforms with frequent inter-player occlusions. We propose a multi-level multitarget sports-team tracker, which overcomes these problems by modelling latent behaviours at both individual and player-pair levels, informed by team-level context dynamics Fig.1.

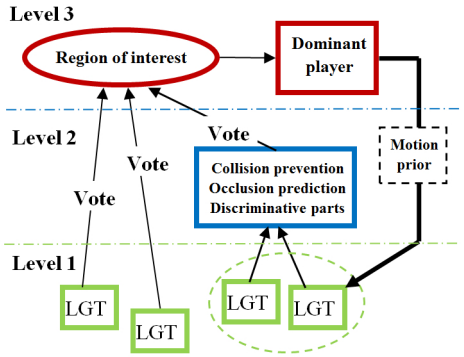

Figure 1: Multi-level tracking algorithm. Level 1: each player tracked by [1]. Level 2: player-player occlusions handled by player-pair behaviour model. Level 3: group or team-level context-dynamics gives dominant player trajectory prediction.

1 Individual player level (Level 1)

At the lowest level (Level 1), we track individual players using the state-of-the-art LGT "Local-Global" tracker [1].This, itself involves two "layers" of tracking: a parts-based set of "local" patches (based on intensity distributions), and a "global" target model (incorporating motion, shape and colour distributions). These local and global layers each provide constraints for re-learning the other, which enables stable adaptation, shown in Fig. 2.

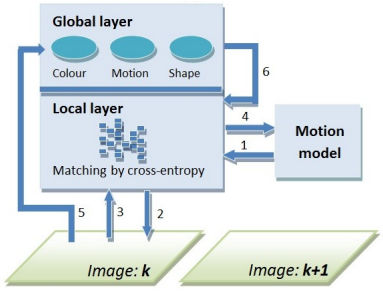

Figure 2: Single-target tracking steps at each frame. 1-spatiotemporal prediction, 2-match local layer, 3-update patches, 4-update motion model, 5-update global layer, 6- add new patches. Adapted from [1]

2 Local group-level (Level 2)

The LGT player models (Level 1) are next augmented by an additional model at the local group-level (Level 2), which encodes the motion preferences of two or more players in close proximity, in the form of a probability distribution representing their tendency to avoid collisions. The pair-wise collision-avoidance model is used to modify the local patch models and global target models of a target pair: the global motion model is modified by the collision avoidance model, providing a stronger motion prior; a prediction is made about which local patches will be occluded during the pair-wise player interaction; and remaining patches are weighted according to their predicted discriminative power during such interactions.

3 Global group-level (Level 3)

We next examine the motion of multiple players at the global group-level (Level 3). Based on player positions, provided by the lower tracking levels, we propose an adaptive approach to meshing the playing area in which the mesh resolution scales appropriately with player density. A player-voting method is then proposed which computes a region of interest (ROI), based on the distribution of player locations and their individual velocities Fig 3.

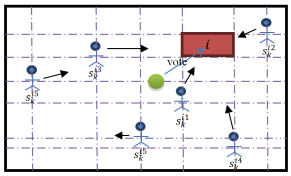

Figure 3: forming mesh according to players' distribution. Green circle: centre of players' distribution; Red region: potential region of interest.

The region of interest does not necessarily indicate the ball position, but may equally indicate the future ball position, or some other position of strategic importance, as predicted by the players. Using this information, it is possible to select one or more "dominant" players, who tend to move with a clearly identifiable trajectory towards the ROI, with a high degree of confidence.

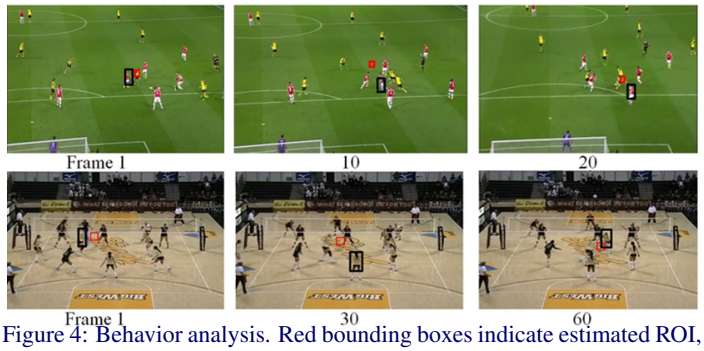

Figure 4: Behavior analysis. Red bounding boxes indicate estimated ROI, Black bounding boxes show a dominant player.

In Fig.5, the group-level models enable successful tracking of interacting/occluding player-pairs where LGT fails (see the right-most player-pair in the right-most image).

Figure 5: Frames 34, 81 of volleyball sequence: LGT (left pair) and our multi-level tracker (right pair). Green/red bounding boxes denote correct/erroneous tracking respectively.

[1] Luka Cehovin, Matej Kristan, and Ales Leonardis. Robust visual tracking using an adaptive coupled-layer visual model. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(4):941–953, 2013.