Interactive Multi-camera Soccer Video Analysis System

Yunjin Wu, Ziyuan Zhao, Shengqiang Zhang, Lulu Yao, Yan Yang, Tom Z. J. Fu, and Stefan Winkler*
Advanced Digital Sciences Center (ADSC), University of Illinois at Urbana-Champaign, Singapore

ABSTRACT

Automatic sports video analysis is an active field of research, and accurate player & ball tracking is essential for soccer video analysis and visualization. However, the variations over frames and the scarceness of large-scale well-annotated datasets make it difficult to perform supervised learning using pre-trained models on different soccer video analysis, especially for Multi-Camera Multi-Target Tracking (MCMT). In this paper, we introduce an end-to-end system for multi-camera soccer video analysis that makes heavy use of parallel processing for optimization of the processing workflow. The proposed thread-level parallelism speeds up our system by more than 15 times while maintaining the level of accuracy. The system tracks the trajectories of the ball and the players in a world coordinate system based on soccer videos captured by a set of synchronized cameras. Based on these trajectories, various player-, ball-, and team-related statistics are computed, and the resulting data and visualizations can be interactively explored by the user.

CCS CONCEPTS

• **Computing methodologies** → *Activity recognition and understanding*.

KEYWORDS

multi camera multi people tracking, soccer video analysis, parallel computing

ACM Reference Format:
Yunjin Wu, Ziyuan Zhao, Shengqiang Zhang, Lulu Yao, Yan Yang, Tom Z. J. Fu, and Stefan Winkler. 2019. Interactive Multi-camera Soccer Video Analysis System. In *Nice '19: ACM Multimedia, 21–25 Oct, 2019, Nice, France.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

In the filed of soccer video analysis, Multi-Camera Multi-Target Tracking (MCMT) algorithms are commonly used to track people in complex environments. Multi-camera (MC) tracking, as known as Multi-view object tracking, is regarded as a data association problem across cameras. And in multi-target tracking (MT) over frames, tracking-by-detection [2], in which object detection and data association play important roles, achieves impressive results.

Traditionally, background subtraction, geometric & sparsity constraints, and occlusion reasoning are performed in MCMT and work well when the number of people in the scene is limited [1, 3, 6]. However, MCMT is by nature a composition optimization problem, which is a time-consuming process.

While deep learning methods have achieved great performance for object detection on single images [7], such methods cannot be leveraged well for the data association problem across cameras. Furthermore, a large amount of well-labeled data is necessary to train a deep learning model.

In this work, we present an end-to-end multi-camera player & ball tracking system for soccer video analysis and visualization. In our system, with the camera parameters provided, the image planes of different cameras can be projected onto a 3D location in the world coordinate system of the soccer field. In this world coordinate system, we estimate a probabilistic occupancy map for players, which forms the basis for the subsequent trajectory search. Shape matching is used for ball tracking. In addition, to speed up our system, a carefully designed multi-threading strategy for parallel processing is implemented. Based on the extracted player and ball trajectories, various statistics about the players, teams, and the ball can be explored through an interactive web-based user interface.

2 PROPOSED ARCHITECTURE

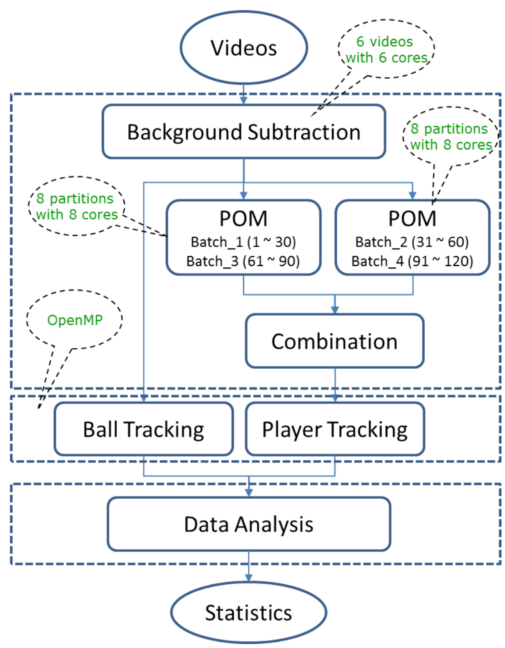

Figure 1: System architecture diagram.

The proposed system architecture is shown in Fig. 1. We assume video captures from six synchronized cameras (three on each side

---

*Tom Z. J. Fu is now with BIGO Technology Pte. Ltd., Singapore.
Stefan Winkler is now with the National University of Singapore (NUS).

Nice '19, 21–25 Oct, 2019, Nice, France

ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

of the soccer field). Frames undergo background subtraction using OpenCV-based implementation of Gaussian mixture models [8], from which foreground masks are extracted. Meanwhile, the soccer field is partitioned into eight sub-divisions; player projections are generated for each sub-division separately. Following that, we estimate the probabilities of the locations occupied by a player according to probabilistic occupancy maps (POMs) [6]. POMs are calculated for each sub-division and then combined. Player tracking is formulated as a multiple trajectory search problem over a sequence of POMs. Appearance modeling helps to further improve robustness to occlusions. Ball tracking follows a simpler shape matching paradigm; triangulation of the ball positions in opposing camera views provides the location in 3D.

To accelerate the system, parallel processing is implemented via multi-threading of different modules. In background subtraction, data from six cameras can be processed simultaneously using 6 threads. We use 2 threads per sub-division in the POM computation, which work together to process alternating batches of frames. Finally, we implement an OpenMP [4] architecture to accelerate the trajectory search and tracking process using shared-memory programming.

3 INTERACTIVE DEMO SYSTEM

Our interactive demo system allows the user to visualize various data and statistics of the soccer game; some examples are shown in Fig. 2 and Table. 1. This includes individual player trajectories on the soccer field or the distribution of positions of players from each team. We can further analyze player speed and distance run, ball possession, pass accuracy, and so on. This allows users to compare individual player performance as well as study team formation, tactics, and strategies, which can be helpful for review, feedback, and training purposes.

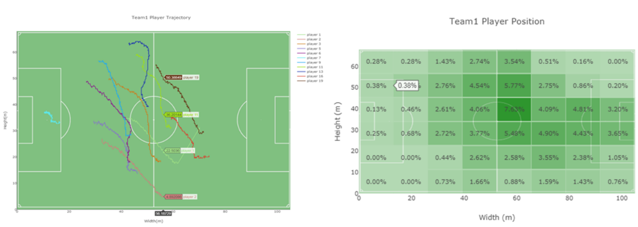

Figure 2: Player trajectories (left) & Distribution of player positions (right)

Table 1: Player statistics

Player	Distance	Max speed	Ball possession	Passes
10	2.3km	18km/h	188sec	25
14	1.8km	14km/h	231sec	18
3	1.5km	19km/h	18sec	5
7	1.2km	12km/h	40sec	4
8	1.1km	16km/h	113sec	20

4 EVALUATION PROTOCOL AND RESULTS

We test our architecture on the ISSIA dataset [5], which provides annotations for player & ball locations, with each player being assigned a unique ID. The tracking results are projected to the frames of the six videos. If a player projection in a given frame has an intersection over union (IOU) with a player annotation higher than a certain threshold, we consider it a true positive (TP), otherwise it is a false positive (FP). If a player annotation is not matched by a player projection, it is counted as a false negative (FN). For ball tracking, we use the distance between a ball projection and the ball annotation to determine its accuracy. If the distance is within a certain range, we consider it as a true positive, outside that range it is counted as a false positive. If a ball annotation is not matched, it is counted as a false negative. For player & ball tracking evaluation, we use two metrics, false positive rate (FPR) and false negative rate (FNR). For players, we also consider ID switch rate (ISR); if two consecutive projections of the trajectory have different IDs, it counts as one ID switch.

We implement our system using different video resolutions on a server running Ubuntu 14.04 with a Intel(R) Core(TM) i7-3770 CPU @ 3.40GHz x 8 and 8GB of RAM. For player tracking, with an IOU threshold of 0.3, both FPR and FNR are small (0.02 and 0.04 respectively), which shows that the obtained player trajectories well match actual player positions in the videos. The ISR is very low (about 0.005), indicating that ID switching seldom occurs in the obtained player trajectories.

For ball tracking, we achieve an FPR in the range of $[0.06; 0.12]$ and corresponding FNR in the range of $[0.16; 0.10]$. Ball tracking is clearly more difficult than player tracking, because the ball is small, frequently occluded by players, and its color is similar to the clothes of one team.

With parallel computing, the processing speed of our system can be improved by more than 15 times in our implementation, without affecting overall tracking accuracy.

5 CONCLUSIONS

This paper introduces an interactive system for multi-camera soccer video analysis. Without manual annotation, this system can accurately track players and the ball for further analysis and visualization. Multi-threading parallelism in different modules is used heavily to improve processing speed by more than 15 times. An interactive demo system allows users to explore various player & ball data and statistics through different visualizations.

ACKNOWLEDGMENTS

This Project is funded by the Singapore Sports Science & Technology Research Grant (SSSTRG) established by the Singapore Sports Institute, a division of Sport Singapore. More information on the SSSTRG is available at http://tiny.cc/ssstrg

REFERENCES

[1] Alexandre Alahi, Laurent Jacques, Yannick Boursier, and Pierre Vandergheynst. 2011. Sparsity driven people localization with a heterogeneous network of cameras. *Journal of Mathematical Imaging and Vision* 41, 1-2 (2011), 39–58.

[2] Mykhaylo Andriluka, Stefan Roth, and Bernt Schiele. 2008. People-tracking-by-detection and people-detection-by-tracking. In *2008 IEEE Conference on computer vision and pattern recognition*. IEEE, 1–8.

[3] Jerome Berclaz, Francois Fleuret, Engin Turetken, and Pascal Fua. 2011. Multiple object tracking using k-shortest paths optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 9 (2011), 1806–1819.

[4] Leonardo Dagum and Ramesh Menon. 1998. OpenMP: An industry-standard API for shared-memory programming. *Computing in Science & Engineering* 5, 1 (1998), 46–55.

[5] Tiziana D'Orazio, Marco Leo, Nicola Mosca, Paolo Spagnolo, and Pier Luigi Mazzeo. 2009. A semi-automatic system for ground truth generation of soccer video sequences. In *Proc. 6th IEEE International Conference on Advanced Video and Signal Based Surveillance*. 559–564.

[6] Francois Fleuret, Jerome Berclaz, Richard Lengagne, and Pascal Fua. 2008. Multi-camera people tracking with a probabilistic occupancy map. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 2 (2008), 267–282.

[7] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems*. 91–99.

[8] Zoran Zivkovic et al. 2004. Improved adaptive Gaussian mixture model for background subtraction.. In *Proc. ICPR*. 28–31.