Analysis of Player Motion in Sport Matches

Janez Perš, Matej Kristan, Matej Perše and Stanislav Kovačič

Faculty of Electrical Engineering, University of Ljubljana
Trzaska 25, SI-1000 Slovenia
e-mail: janez.pers@fe.uni-lj.si

Abstract

The system for analysis of player motion during sport matches, developed at University of Ljubljana is presented. The system allows for non-intrusive measurement of positions of all players in indoor sports through whole match, using only inexpensive video equipment - cameras mounted on the ceiling of the sports hall. Tracking process (obtaining trajectories from videos) is automatic and only supervised by operator, to initialize player positions at the beginning and correct the mistakes during the tracking. The software provides means for user friendly calibration of video data - using court markings of each supported sport as reference coordinates. The system has been tested for European handball, basketball, squash and tennis. Manual annotations can be added, to complement the quantitative data. Software keeps synchronization between annotations and trajectory data and provides means to use custom annotation dictionaries. Due to calibration, the results are provided in court coordinates (meters, centimeters) and can be exported (synchronized with annotations in same file) for further analysis with any application (e.g. Microsoft Excel, SPSS). Software itself supports several types of graphical data presentation.

1 Introduction

Most sports include complex motion, which can be studied at different levels of detail. This is especially true for team sports (e.g. soccer, handball, basketball). Some aspects of an athlete's movement can be studied in detail using commercially available high-speed, high-accuracy biomechanical measurement systems [1, 2]. These systems provide high-quality, high-resolution data on human movement, however, they also impose severe limitations, which make them inappropriate for studying motion covering large areas for long intervals of time.

In team sports, action is often spread across the whole playing field (which, for example in European handball, measures 40×20 meters), and matches can last for an hour or more. Most of the biomechanical measurement systems provide extremely accurate data on human movement, but cannot effectively cover such a large area. These systems provide high temporal resolution (100 Hz or more), but over relatively short intervals of time, when the duration of a typical match is considered. Many of these systems require some kind of markers to be attached to the body of the athlete for automatic operation. However, the markers are distracting and not acceptable during regular league or championship matches.

However, even coarse information about movement of the players participating in the match can be used in many ways, such as tactical analysis and performance analysis. Such data could be obtained with technology that sacrifices some of the accuracy, but provides means for much easier

Dagstuhl Seminar Proceedings 08372
Computer Science in Sport - Mission and Methods
http://drops.dagstuhl.de/opus/volltexte/2008/1689

acquisition of data, covering large areas. Years ago, researchers used self-made, video-based systems with resolution of 1 meter for successful analysis of a soccer match [3].

The system for analysis of player motion during sport matches, developed at University of Ljubljana is based on similar principles. The key features of the system are:

- Low hardware cost - we use analog surveillance cameras, coupled with off-the-shelf DVD video recorders. DVD recorders can be placed at an easily accessible location. Camcorders can be used as well, but we find this option inconvenient, as it requires frequent access to the camcorder itself.
- Flexibility - surveillance cameras can be fitted with many brands and types of inexpensive lenses. It is therefore possible to adapt the system to a different sport hall at a relatively little cost.
- Off-line processing - videos are recorded during the match, then transferred to personal computer after the recording. Off-line processing provides a way for obtaining accurate and reliable data, as the processing results can be reviewed and corrected.
- Automated processing - advanced computer vision methods are employed for tracking the players, and the tracking is automatic.
- Operator supervision - the tracking process is initialized by human operator, and then supervised. Tracking errors are fixed by operator as well.

In this paper, we describe the workflow for a typical use of such a system, and its individual components. Some technical aspects are covered as well.

2 Video acquisition

The design of the software requires that the video is captured from the bird's eye perspective. This significantly reduces the complexity of the tracking and camera calibration. The cameras do not need to be *exactly* perpendicular to court surface, but have to cover whole area of the court. Figure 1 provides illustration of the camera placement for a typical basketball or handball match.

When using multiple cameras, they are positioned to provide even court coverage. Some overlapping of video is required, but software does not depend on correspondence between views of the same player from different cameras. Due to bird's eye perspective, occlusions between the players are rare. Since each of the cameras covers its own area of the court, the calibration is relatively simple from the user's standpoint - the boundaries of the court can be used as a reference. On the other hand, the system can only be successfully deployed in the indoor environments, as there are no reliable (and stable) means to position camera above the court in outdoor settings. Figure 2 shows few frames from videos from different sports.

Currently, analog surveillance cameras are used for video acquisition. Although this is not state-of-the-art technology, it offers several advantages in comparison to more advanced digital solutions:

- Inexpensive cabling with little restriction in length. Analog composite video cables and S-Video cables can be up to 100 meters long with very little degradation in video signal. This is extremely important when locations with cameras (on the ceiling) are not easily accessed. Due

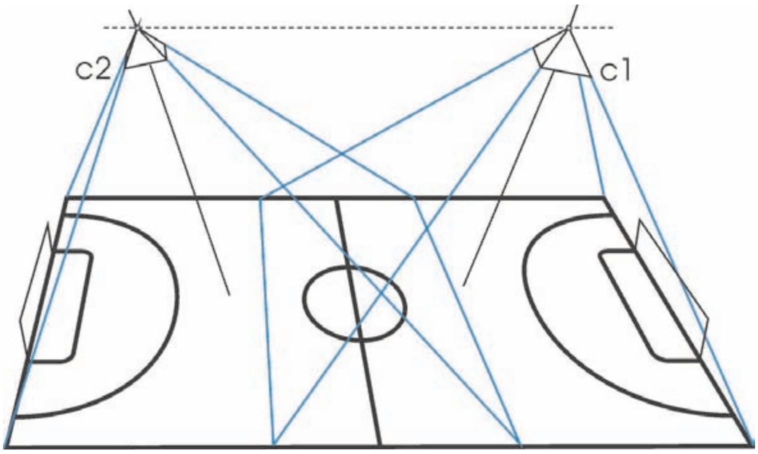

Figure 1: Illustration of the camera placement for a typical basketball or handball match - each camera covers one half of the court.

to security reasons, such locations are sometimes off-limits for researchers during important matches with large audience, and in many cases there is extra equipment (cranes) needed to access such locations. Long cabling enables recorders to be in an entirely different location, where recording media can be easily changed, when needed.

- Inexpensive recording equipment. Off-the-shelf DVD video recorders can be used to record standard analog video signals, but for higher resolutions, a PC or specialized equipment has to be used. Based on our experience, PCs are poor choice to record valuable matches due to their somewhat lower stability. There are also chances that the recording is disrupted by the other activities of the PC's operating system during the recording that lasts two hours or more. On the other hand, even off-the-shelf DVD recorders will record a two hour match without a glitch.

- Low storage demands for captured video. Although there are cameras that can provide resolution in multiples of the standard PAL video resolution, they also generate correspondingly huge amounts of video data, which require powerful hardware for real-time compression at high frame rates, adding to the overall cost of such system.

In our experience, at the time of the writing of this paper (October 2008) the advantages of the analog technology still outweigh its disadvantages, especially when recording matches of the high importance, which usually require tight schedules for hardware installation and testing, and place significant constraints on access to the recording equipment. However, with the further development of the digital video, the system can be used with digital recordings with no modifications to its software components.

3 Workflow

After the videos are recorded, they are transferred to personal computer's hard drive using DVD media. If full resolution video is desired, the content of the DVD can be transferred to the hard

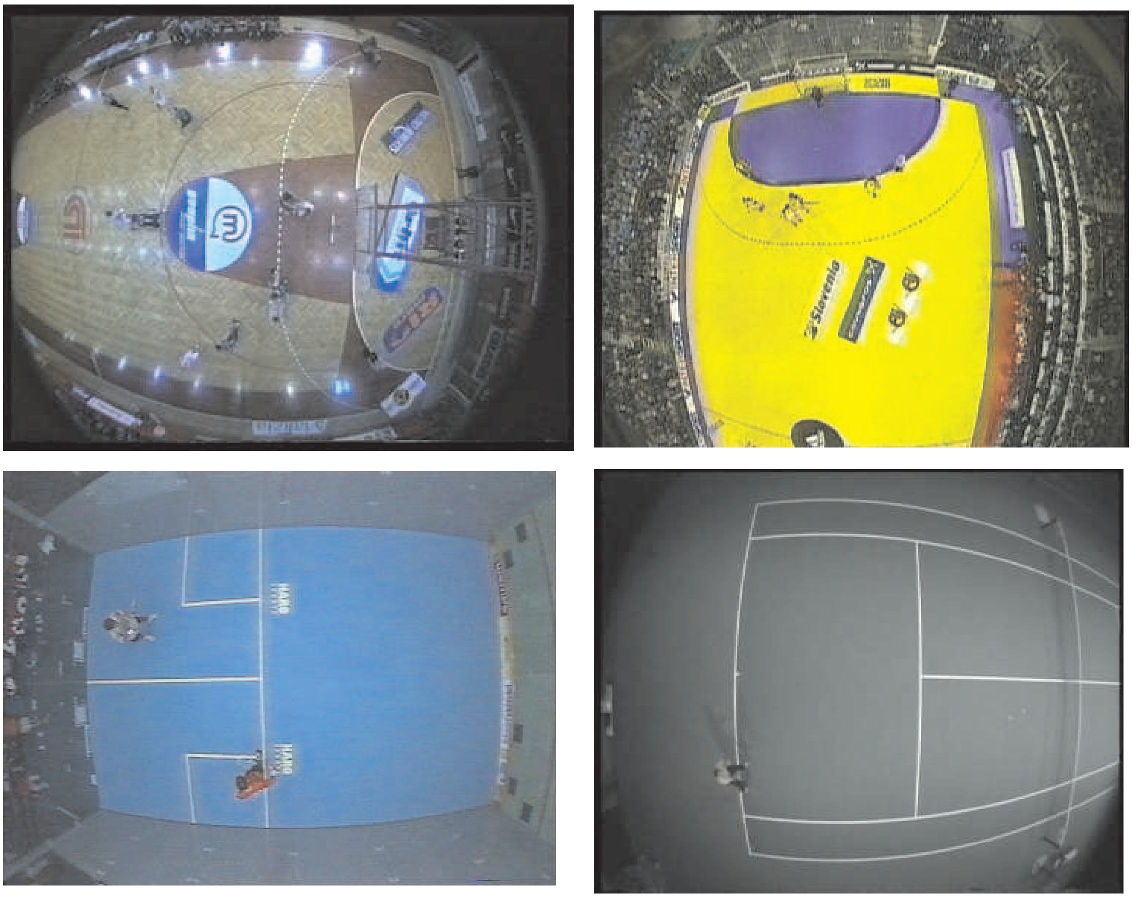

Figure 2: Few frames from videos from different sports - from top-left to bottom-right: basketball, handball, squash, tennis.

drive using the utility, which is part of the software package, or simply by copying video files from the DVD (*.VOB files) and renaming them appropriately. To conserve disk space and speed up the processing, videos may be downsampled to lower resolution using one of the many commercial or free applications, developed specifically for such task (called "DVD ripping"). The software will accept videos in either MPEG1 or MPEG2 format.

Then, the next steps of video processing follow: video calibration, tracking, manual annotation (optional), export of trajectory data and annotations to the file in tab separated format or export of graphical diagrams of trajectories, velocity and acceleration.

3.1 Calibration

Calibration procedure has two parts. It is done separately for each of the videos in case of multiple cameras. First step is temporal calibration - videos need to be aligned in time. The frame, that contains an event near the beginning of the match, visible in videos from all cameras, is selected. Such event is for example initial throw in handball or referee's throw in basketball. If a single camera is used, any frame can be chosen as a starting point. This frame represents the temporal

Figure 3: User interface of the calibration module during the spatial calibration. User is marking the rightmost boundary of the area covered by one of the cameras. This boundary is the center line of the court, as seen on the schematic representation on the right side of the user interface.

reference - the time instant $t = 0$ for the match. Then, for each video, ending frame of the match is chosen.

After that step, the software scans through the video and prepares the image of the empty court - by randomly choosing a handful of frames and applying median operator along temporal axis for every pixel. The process is quick and transparent for the user. The end result, image of the empty court, is used for spatial calibration. User is asked to mark boundaries of the court. Using predefined templates for supported sports, the software computes relations between pixels in the video and coordinates of the court. In that process, several parameters are calculated: the radial distortion of the lens, the projective distortion of the image due to imperfect camera placement (tilt), the scaling factors and the offsets. Fixed height of the players is assumed - the best tracking accuracy is achieved, if player's center of gravity is tracked, which is true most of the time.

The second step - the spatial calibration - can be skipped if user provides a video from the same camera, which has been already calibrated (e.g. from the previous match, if the cameras were not moved). In that case, software simply copies the existing calibration parameters.

Figure 3 shows the user interface of the calibration module.

3.2 Tracking

After calibration for videos from all cameras has been performed, the main tracking module can be started. Software searches for the videos from the same match in the project folder, chosen for the match. Simple video naming scheme allows software to automatically find all videos from all cameras for the match in question, and automatically associates them with their respective calibration files. Player data (names, surnames and team affiliations) are provided as a separate text file, which is automatically loaded as well.

The videos are displayed as thumbnails on the right side of the user interface, and by clicking of one of those thumbnails, video is shown in the main display window of the interface. List of players is displayed on the right side of the interface.

First, user has to choose a few players to track. This is done by selecting a player in the list and then clicking on his position in the main image. This action assigns player to the group of players that are tracked. Usually, operators select only a few (not more than 3 or 4) players for tracking at the same time. This way, operator has a good overview of how the tracking is progressing, and, a chance to correct any inaccuracies or mistakes in tracking without stopping the tracking process.

If major problems arise, operators can slow down tracking significantly (to about 1 frame per second) to give themselves more time for any interventions or corrections. If the problems are significant (such as a group of players colliding, lying on a floor, or similar), they can stop the tracking, sort the problem out, and reinitialize tracking, similarly to the process used at the beginning of the tracking.

Keyboard shortcuts are provided for quick selection of players, so the interventions can be done without much physical effort.

If the player ends it participation in the play, it can be "placed" at fixed position anywhere on or around the court (e.g. on the bench). Software will update its position with constant value of position, regardless of the visual information coming from the placement position. This way, whole team can be analyzed, even though only a minority of players are on the court at the same time (e.g. in team sports with frequent substitutions, such as basketball and handball). Players that are "placed" on the bench do not have to be checked for correctness of the tracking, since their positions are unchanged until they are explicitly returned to court by the operator.

When the match is re-analyzed to track the next batch of players, the existing tracking data for already tracked players is kept untouched - players are by default set to "do not touch" mode by the software to prevent accidental overwriting of data, unless operator explicitly chooses them to be re-tracked.

During the tracking, software handles transitions between different cameras automatically. Since all cameras are calibrated, the position of the player in one video can be recalculated to the position in other video, when players cross the boundaries between the cameras, and usually, no reinitialization of player position is needed.

Interface provides a few settings, which adjust much larger number of parameters, but are hidden from the user. For example, users can choose higher processing speed on the expense of accuracy - this adjusts the number of particles in the underlying particle filter, which forms the core of the tracking engine [4]. They can adjust the apparent (visual) size of the players to account for variations in camera setting, and they can adjust the dynamics of the tracker. Dynamics setting adjusts the parameters of the built-in Kalman filter [5] and therefore provides higher or lower tolerance for fast and sudden player movements. Rest of the parameters are not available for adjustment. User interface of the tracking module is shown in Figure 4.

Figure 4: User interface of the tracking module. Three players are marked for tracking, one of them is "in focus" - selected in the player list on the right side and marked with green circle.

3.3 Manual annotations

Not all data can be automatically extracted from videos in a reliable way. Therefore, software provides a means to add manual annotations to the automatically extracted position/trajectory data. The main benefit of this combination is the common timeline for both the annotation and the player positions. Annotation module interface is similar to the interface of the tracking module. A set of *dictionary* files, one for each sport, is used to provide operator with available choices for annotations. Software supports two types of annotations - an "instant" annotations, which are used to mark the events that happen at certain time instant. Such annotations are used for example to mark shots and fouls, and can be accompanied with position on the court, which is marked by the operator in the video. They are visible as single mark in the exported data file. Calibration is used to transform the mouse click coordinates in the video to the coordinates in the court. The other type of annotations is used to mark larger intervals of time, such as active and passive phases, offense or defense phase of certain team, and similar. Those annotations are visible as a separate columns in the exported data file.

3.4 Data export and display

Presentations module provides means to draw various types of graphical displays, some of them shown in 5. After the module is loaded, it loads the trajectories of all players, and performs post-processing, which includes smoothing of positions with a Gauss-shaped kernel and calculation of the

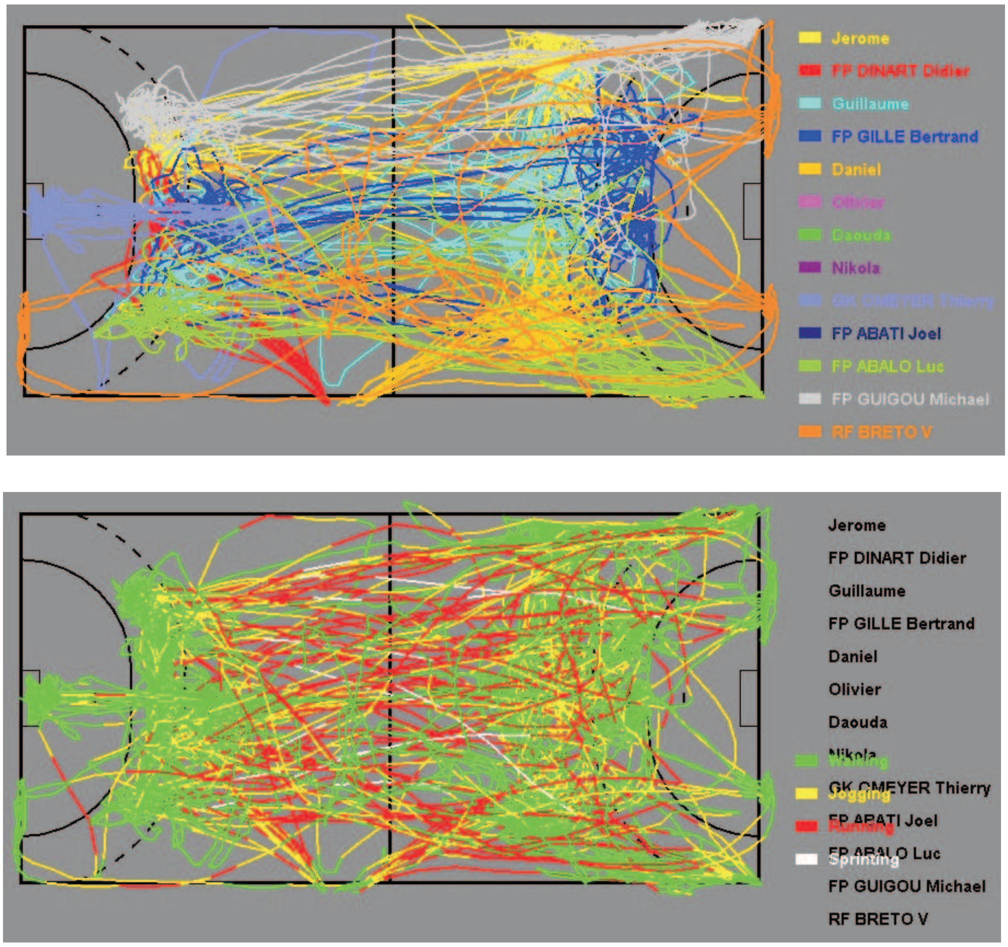

Figure 5: Two displays of the trajectories, generated by the presentation module. The upper display shows trajectories of different players in different color. The lower display shows trajectories encoded with the velocity of player at the certain point. The trajectories are from handball match and the fast transitions across the middle are clearly visible.

derived parameters, such as velocity and acceleration. Data can be exported to the tab separated text file, where positions, velocities and accelerations of every player are provided, along with the properly aligned manual annotations.

4 The technology

The software is written in C++ and Java. Both the interface and the processing engine can be compiled under Windows and Linux without modifications. User interface was done in Java, however, to achieve best performance, the image processing and analysis part was written in C++.

Software employs state-of-the-art algorithms for object tracking. The core of the tracking algorithm is the particle filter [4], designed to work with relatively low resolution videos, which are easily obtainable with off-the shelf equipment. Players are modelled as ellipsoid blobs, and color histograms, sampled inside the player ellipse are used to describe player appearance throughout the tracking. Such model is wrapped in the powerful envelope of the particle filter, which takes the care of adapting the model according to the changes in the player's appearance. The algorithms are described in detail in [4, 5].

The RMS error in position of the tracking is around 30 cm. This may seem high, but the main sources of errors are not tracking methods, but the various simplifications that are introduced to the concept of large-scale tracking to make it feasible for practical use. For detailed analysis of the factors that influence the accuracy of such measurements, see [6].

5 Current work

There are many reasons researchers from the field of computer vision are interested in analysis of sport videos. Of course, development of the new tools and methods for use in sport science and sport analysis is just one of them.

The most attractive property of sport for the computer vision researchers is that it provides well documented environment for human motion and human activity. Therefore, it is extremely appropriate setting for development and testing of various computer vision methods and algorithms, that have much wider applicability than sport domain alone. Algorithms for people tracking, for analysis of human motion and activity and algorithms for analysis of group (multi-agent) activity are difficult to test in real-world environments. Expert knowledge about motion, activity and goals of observed persons is rarely available, even in surveillance applications, for which there is significant commercial interest.

On other hand, sport provides clear rules and goals for the persons involved. Expert knowledge is readily available, and the motion, tactics and strategy for many sports have been studied in detail.

Our motivation for collaboration with sport expert lies in access to sport environments, where we can obtain videos, test our algorithms, and, even more importantly, get the insight into the laws that govern the motion and activity on the videos we observe and process using the computer vision methods. For example of recent work, see [7].

6 Conclusion

We presented a system for automatic analysis of player motion in indoor sport matches. The system is result of 10 years of collaboration between researchers from sport and computer vision domain at University of Ljubljana. The additional presentation of the software can be seen at the following address: `http://www.youtube.com/watch?v=Lnrd4YeGGhI&fmt=18`

References

[1] J. G. Richards. The measurement of human motion: A comparison of commercially available systems. *Human Movement Science*, 18(5):589–602, October 1999.

[2] A. Gruen. Fundamentals of videogrammetry - a review. *Human Movement Science*, 16(2-3):155–187, April 1997.

[3] W. S. Erdmann. Gathering of kinematic data of sport event by televising the whole pitch and track. In *Proceedings of 10th ISBS symposium*, pages 159–162, Rome, 1992.

[4] M. Kristan, M. Perse, J. Pers, and S. Kovacic. Closed-world tracking of multiple interacting targets for indoor-sports applications. *Computer Vision and Image Understading*, 2008 - in press.

[5] M. Perse, J. Pers, M. Kristan, G. Vuckovic, and S. Kovacic. Physics-based modelling of human motion using kalman filter and collision avoidance algorithm. In *4th International Symposium on Image and Signal Processing and Analysis ISPA*, pages 328–333, September 2005.

[6] J. Pers, M. Bon, S. Kovacic, M. Sibila, and B. Dezman. Observation and analysis of large-scale human motion. *Human Movement Science*, 21(2):295–311, 2002.

[7] M. Perse, M. Kristan, S. Kovacic, Vuckovic G., and J. Pers. A trajectory-based analysis of coordinated team activity in a basketball game. *Computer Vision and Image Understading*, 2008 - in press.