Proposal of a modular system for tracking indoor and outdoor sports

VLADIMIR PLEŠTINA¹, HRVOJE DUJMIĆ², VLADAN PAPIĆ¹
Department of polytechnics¹, Department of electronics²
University of Split
Teslina 12, Split
CROATIA
vladimir.plestina@pmfst.hr, hrvoje.dujmic@fesb.hr, vladan.papic@pmfst.hr

Abstract: Use of computers and computer vision to detect and track players became the topic of interest of various research groups. Regardless of whether a custom set of static cameras is used or the images are acquired from directed TV coverage, player detection and tracking is a complex process. As a sport in general is popular and commercially potent area, commercial, quite expensive programs for analysis and track players in different environments have been developed. In this paper, after an overview of various methods presented in the literature, a proposal of a modular system for tracking indoor and outdoor team games using computer vision is given.

Key-Words: - modular system, tracking players, computer vision

1 Introduction
For many years people are trying to find the best way for tracking players. Various methods were applied like placing the microphone on the edge of the field, using GPS transmitters and application of computer vision. Regardless of the sport, camera and computer appliance provides a wide range of possibilities. Systems with computer vision can be divided into three parts. The first relates to camera system, second is dedicated to the detection and identification of players, and third is oriented towards methods and algorithms used for players tracking.

It can be used specialized camera system for detection and monitoring, or, like some authors do, it can be used directed TV images. Sullivan [1] uses a system with four wide screen Wide angle camera with high resolution. Iwase [2] placed eight cameras, four on each side of the playground. Xu [4] also uses eight cameras, but they are positioned differently. Four cameras are placed on one side of the field, and two behind every goal. Gedikli [3] uses video clips and calculate players orbit. Captured image is needed to be processed in order to get information, in this case, information about the location of a particular player. The easiest way to find players is subtraction picture and known background [1] i [2]. Figueroa [9] uses a player separation algorithm without background subtraction. Gedikli [3] segment players with known models of colour and then with templates and known colour distribution calculate the most similar model. After that he uses prediction while separation and labeling of players is done using the previously calculated data.

For tracking players, authors [1], [2], [4] and [9] calculates orbit of earlier detected and marked objects. Each one of them has its own algorithm for tracking which is linked with player detection algorithm. Neither of these systems is fully automated. In any case human is supervisor which takes initiative when determines that system is wrong. In this work it will be shown methods provided by different authors and based on these methods it will be proposed modular system that can independently track players during the match.

2 Methods overview
As it has already been stated in the introduction, a computer vision system for tracking players generally consists of three main steps: camera system, players detection and players tracking (Figure 1).

![Fig. 1 Main steps of computer vision system for tracking players](image Reference)

More details on each of these steps will be explained in the following subsections.
2.1 Camera position
Camera position depends on system needs. With own camera set it is easier to track players as reference points are known. The camera position is well known, there is no zoom and camera movements, so calculations are much easier to make. On the other hand, it is necessary every time mount and calibrate the camera before the game. The system is not standardized so it is necessary to have special equipment.

Using pre-existing system simplifies use for user (sport clubs, TV networks), but requires different ways of player detection which complicates calculations.

Directed match images can be used for a particular type of analysis, but cannot be used to track all players during the entire game.

Sullivan's [1] system with four wide angle wide screen high resolution camera can locate players when they are isolated and easily monitor their movement. Problems occur when player occlusions appear. Then path length measure stops. Iwase [2] solve occlusion problem with more cameras around the field. This author placed eight cameras, four on each side of the playground. According to him, the first step is to track every player with each camera. When occlusion occurs other camera is used to determine identity of the player. As camera position is known it is easy to get position of each player.

Iwase mentioned two terms, “Inner-camera Operation” and “Inter-camera Operation”. “Inner-camera Operation” is done separately for each camera for tracking players. When occlusion occurs “Inter-camera Operation” is processed. Figure 2 represent flow chart of these two methods.

Due to large number of cameras, tracking player can be done from all angles what makes tracked player analysis easier. The problem of this system is heavy equipment which is necessary for monitoring the large number of input parameters (pictures from eight cameras) that need to be processed.

Xu [4] also uses eight cameras: four cameras are placed on one side of the field, and two behind each goal. Each camera observes a certain part of field. Described system [4] is not used to track and analyze individual player, but to recognize teams to help future tracking. Same author in his other work [5] use this system for tracking.

Tracking players also can be done using directed TV broadcast. Some authors use TV images or direct TV broadcast from a playing field to track player(s). Gedikli [3] presents ASPOGAMO system which count players path based on video recordings. This system can give information about a particular player based on recorded TV broadcast. Although the system finds player on the field, problem is that TV transmission cover only part of the field. Also, with this system it is possible to provide information based on individual sequences, but it is not possible to constantly monitor all players. The advantage of this system is that captured image can be used for recognition methods which could not be applied when the camera is far away. Example is face recognition or player number recognition. In addition, it should be mentioned that broadcast cameras are requiring calibration [8], [13].

To decide what kind of camera system will use, it is necessary to know what will be monitored. System with multiple cameras is suitable for monitoring whole game in order to obtain characteristics of all players.

Systems with existing cameras on the ground have some limitations. They cannot monitor all players throughout the game. Such cameras are mainly focused in the area where is a ball. This cameras as well as directed recordings are used when it is needed to get certain data from sequences, like players when performing free strokes or corner.

Qu [7] and Iwase [2] use camera collaboration. Such way of cooperation between cameras provides the ability to more accurately monitor individual player when occlusion occur.

2.2 Player detection
Most authors are looking for the simplest way to find object in an image. Knowledge about playfield background is one of the methods. Sullivan [1] uses algorithm that includes known background and player jersey. After applying Gaussian distribution, algorithm
finds objects (players) on image. Such player labelling as objects provides the possibility of tracking them, but does not define a specific player. Iwase [2] uses background subtraction. Then he converts an image into a binary image, filters out noise and detects certain regions. Problem occurs with player shadow and overlaps so it is necessary to have an additional algorithm that divides overlapped players. Same authors like Iwase, Figueroa [9] separate players and track them. These separated segments Figureoa label as “blob”. In his algorithm he uses intensity vertical distribution of “blob” to determine to which team recognized player belongs.

Gedikli [3] uses player segmentation with known colour types, then with known template and colour distribution calculates model. After that, using the predicted positions and previously calculated data separates players and labels them. Xu [4] and [5] use Gaussian distribution for background separation. Authors which use their own camera systems usually use background separation methods which cannot be used on images accrued from TV broadcast. Beetz [10] presents ASPOGAMO system that use special colour model and separate green soccer field. He also uses Gedikli method [3] for labelling players.

2.3 Tracking players

Our aim is to create system that independently and precisely can track players. Some authors [1], [2], [4] and [9] calculate path of detected and labelled objects. All algorithms include tracking and labelling of particular player. Path is calculated between two frames and depends on player location on each.


Lefevre [16] presents “Fast Snake – based” algorithm for tracking using the non-stationary camera. This is contour based algorithm that does not use any prediction. Similar algorithm is explained in Hsiao’s [17] work. Okuma [18] presents Boosted particle filter for detecting and tracking hockey players. This method has very good results and it can be used as a base for tracking in any sport.

3 Modular tracking system

Basic idea of our modular system is to provide model based system with independent and also mutually compatible modules. Every module provides result which is part of the global system solution. The system is divided into separable modules which independently perform a specific operation and produces result. Functionality of the system is possible without individual modules, but it is also possible to add new. In this work we combine methods explained in earlier papers. Modular approach allow user to simply decide what part of system will be used without losing system functionality.

Added modules improve system functionality and each part of the system can work without a specific module. The system is divided into three levels: camera system level, player detection and labelling level and player tracking level.

Common part is auto-correction module which represents feedback and correct wrongly marked or tracked players. Central database contains all necessary data for system functionality. Time-space calibration module joins and calibrates connected modules and mark player label and position on the field based on reference points. Figure 3 shows preliminary solution with two separate systems for monitoring (system A and system B).

System A includes fixed cameras located along the ground. Idea is to get image of entire field and players. As a reference points are known we can determine objects and their positions on the field. The system and connected modules are connected with the central database which contains all the necessary information that individual module uses.

- Module to determine location of players on the field provides information about player position based on knowledge about the team, tactics and player location likelihood.

- Occlusion detection module is activated when player overlap with another player. According to previous
knowledge about position, speed and player movement direction, system split and label overlapped players.

Figure 3 Modular tracking system

Figure 4 shows the application of boosted particle filter to the soccer game between professors on Faculty of Electrical Engineering Mechanical Engineering and Naval Architecture. Okuma [18] presented this method to track hockey players. This game is filmed with static camera (system A). In Figure 4, trajectory for one player with 100 particles during 20 seconds is outlined. Player tracking analysis has been processed through 550 frames. To obtain world coordinates of the tracked player(s), two main approaches can be used. First one is calculation of coordinates based on the image coordinates of the tracked object from only one camera. Image object coordinates are trasformed using the projective transformations knowing world coordinates of some characteristic objects in the image (field dimensions, goal position, etc…). Another approach is based on the tracked coordinates in images obtained by two or more static cameras and implementation of 3D
algorithms for determining spatial position of object such as DLT method [19]

System B receives TV broadcast camera or already existing cameras on playfield information. Unlike system A which uses calibrated and fixed cameras, images from system B cameras are varying. They display only one part of the field and image can be zoom-in or zoom-out. As there are more of such cameras and TV broadcast is directed, image can be shown from different angles. Such system can not be used to track all players on the field, but as part of complete system proposed in this paper. It can be used to improve player labelling in cases which are unclear when system A is applied. System B is also compatible with different modules.
- Face recognition module provides information about the player by recognizing players face and comparing it with the face of players in the database.
- Number recognition module recognizes number on a jersey and compares it with information in database (Figure 5).
- Anthropometry recognition module recognizes player characteristics.

![Image of a soccer player with a jersey number highlighted]

Fig. 5 Number recognition from close-up shots

Figure 5 shows jersey number detection as a result of number recognition module. Recognition algorithm uses HSV colour model [20] and module gets directed image from TV broadcast camera (system B). The information acquired with system B could be obtained with system A if high-resolution cameras are used. However, processing time would be slower and it would be impossible to track players in real time.

System A and system B information are connected and calibrated through time and space calibration module. It compares system reference points and provides information about the player with certain probability. Player is labelled in labelling module which creates a basic prerequisite for tracking players. Auto-correction module checks the selected player according to data of system A. If a certain player has low probability, correction module corrects player label. After player is marked, tracking module tracks player and writes data records in the database.

4 Conclusion

This article is the result of studying a large number of articles related to detection, separation and monitoring of people and sports activities. Shown methods provide overview and base to create modular system for tracking players. Proposed and explained modular system is result of weak-points analysis of present systems. The main idea is to create a system with independent modules that can be connected together to provide better and more robust results. Presented approach provides flexibility and openness for further expansions, inclusion of different algorithms and their activation/deactivation as well as possibility of different system configurations without losing main functionality. Further development of the system will be focused on improving player detection and tracking module and adding new, game-strategy type knowledge to the system’s database.

Acknowledgment

This work was supported by the Ministry of Science and Technology of the Republic Croatia under projects: ICT systems and services based on integration of information (023-0231924-1661) and Computer Vision in Identification of Sport Activities Kinematics (177-0232006-1662)

References:


