A Player-Possession Acquisition System for Broadcast Soccer Video

♥ Xinguo Yu, ♦ Tze Sen Hay, ♥ Xin Yan, ♦ Engsiong Chng
♥ Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613
{xinguo, yanxin}@i2r.a-star.edu.sg
♦ School of Computer Engineering, Nanyang Technological University, Singapore 639798
TzeSen@pmail.ntu.edu.sg, aseschng@ntu.edu.sg

Abstract

A semi-auto system is developed to acquire player-possession *for broadcast soccer video, whose objective is to minimize the manual work. This research is important because acquiring player-possession by pure manual work is very time-consuming. For completeness, this system integrates the ball detection-and-tracking algorithm, view classification algorithm, and play/break analysis algorithm. First, it produces the ball locations, play/break structure, and the view classes of frames. Then it finds the* touching-points *based on ball locations and player detection. Next it estimates the* touching-place *in the field for each touching-point based on the view-class of the touching frame. Last, for each touching-point it acquires the* touching-player candidates *based on the touching-place and the roles of players. The system provides the graphical user interfaces to verify touching-points and finalize the touching-player for each touching-point. Experimental results show that the proposed system can obtain good results in touching-point detection and touching-player candidate inference, which save a lot of time compared with the pure manual way.*

1. Introduction

With the skill advance of playing soccer game, coaches and viewers do not rely only on their subjective observations and feelings to evaluate the team/player performance. Instead, they more and more rely on objective analysis results of a game such as team-possession, player-possession, quality of ball control and pass, team formats, etc [3-5, 8-9]. However, it is very difficult to acquire these results as these analyses entail the ball detection and tracking, player identification and recognition, etc. So far, there have not done enough work on such data extractions.

This paper presents a semi-auto system to acquire player-possession for broadcast soccer video. Since we cannot acquire player-possession in a fully automatic way the system targets to minimize the manual work. The system automatically finds the *touching-point* candidates (a touching-point is a frame in which a player touches the ball). Then for each touching-point it estimates the *touching-place*, where the player touches the ball in the soccer field. For each touching-point the system also automatically produces *touching-player candidates*, which are the players who probably touch the ball at the point. For facilitating the operators of the system, the system provides a GUI to verify the touching-points. It also has a GUI to visually finalize the touching-players, which shows touching-player candidates with their portraits. The user can assign the touching-player of a touching-point by just selections and clicks.

The related work in the literature mainly lies in general soccer video analysis and performance evaluation. For general soccer video analysis, prior work has done on play/break structure analysis [7], ball detection and tracking [9], view tracking and classification [2, 6]. For performance evaluation, prior work has done on motion analysis [5], group-behavior analysis [3], and team-possession analysis [8]. This paper is a follow-up of our previous work in [8], which not only improves the existing components and integrates the relevant algorithms, but also adds the GUI and touching-player candidate analysis for each touching-point.

2. Outline of the System

The proposed system in this paper comprises the graphical user interface (GUI) and the background analysis two portions, whose block diagram is given in Fig 1. The background portion is to acquire the touching-point candidates of the given video and the touching-player candidates for each touching-point. The foreground portion of the system provides GUI to verify the touching-points and to finalize the touching-players. The background and foreground portions works together to reduce the time of acquiring the player-possession.

For acquiring the touching-points, the system integrates the ball-detection-and-tracking algorithm to obtain the ball locations and the structure analysis to obtain the play segments. In a play segment, we can find all the ball-player points. In the other aspect, we form the ball-motion curves from the ball locations, in which the pivots indicate some kinds of force have acted on the ball. The main kind of forces is that players have touched the ball. A ball-player point with a significant pivot is high probably to be a touching point. Thus we can remove the ball-player points very close to it as player cannot have another touching in a very short time. The remaining ball-player points are called the touching-point candidates. Then the operator of the system can use the GUI to verify them to produce the touching-points.

For acquiring player-possession, the remaining task is to acquire the touching-players. Of course, the player closest to the ball is the touching-player. We want to find who the closest person is. It is impossible to recognize who the closest person is through recognizing the image of the player due to the low resolution. Hence, we design an inference

0-7803-9332-5/05/$20.00 ©2005 IEEE

procedure to acquire the touching-player candidates. First, we use Support Vector Machine (SVM) to recognize the team of the touching-player. At the same time, SVM also tell whether he is a goalkeeper. Second, we use the view information (view class) and the player roles to produce the touching-player candidates of the touching-point. The GUI of our system displays all the players from the same team in the two groups: the touching-player candidates and the others. Thus, the operator of the system can assign the touching-player using just selections and clicks.

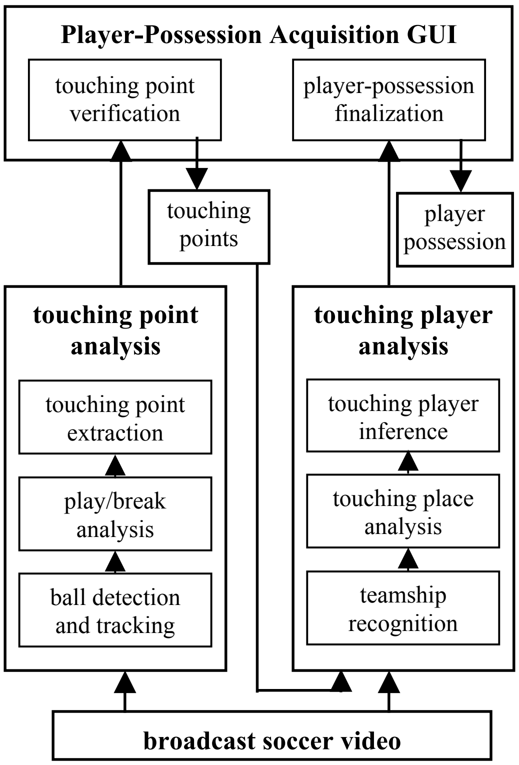

Fig1. Block diagram of the player-possession acquisition system for broadcast soccer video.

3. Touching-Point Acquisition

We assume that the ball locations of all the frames are known as we have already developed the ball-detection-and-tracking algorithm [9]. We also assume that the play/break structure is known as the structure analysis has achieved its good results [7]. With the obtained ball locations of all the frames in [9], this section introduces ball-player points and the pivots. Then, we infer touching-point candidates. Last, based on these candidates the operator uses the GUI to acquire the touching-points.

3.1. Touching-point candidate extraction

A touching-point is a frame where a person (a player or a goal keeper) touches the ball. Obviously, the touching-player must contact the ball when he touches the ball. Hence, a touching-point must be a ball-player point, which the ball has a zero distance with the touching player. Thus, we can take the ball-player points as the basis to further find touching points. Fig 2 shows a sample ball motion curve with the found ball-player points.

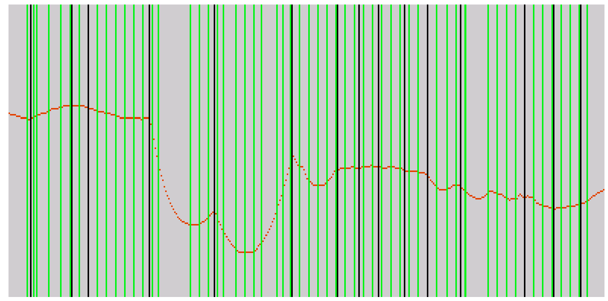

Fig 2. Ball motion curve with the detected ball-player points. The dark color vertical lines indicate the groundtruth; all the vertical lines are the detected ball-player points, meaning no missing.

On the other hand, the ball-speed curve probably forms a local minimum or an acceleration starting point when a person touches the ball. Let V = {p | p is a local speed minima or an acceleration starting point}. Let $f_1(p) = r$ and $f_2(p) = c$ be the functions, representing Y and X curve of the ball location over frames, where r and c are the row and column of the ball center in the frame p respectively. Let S_1 = {p | p is a local trajectory maximum of function f_1} and S_2 = {p | p is a local trajectory maximum or minimum of function f_2}. Then set $S = V + S_2 - S_1$ is called a set of pivots. The motion curve of the ball must form a pivot when a player touches the ball with some force. A ball-player point with a significant pivot is claimed to be a touching point. Then we remove the ball-player point around it. The remaining ball-player points are considered as the touching-point candidates.

3.2. Touching-point verification

Among the touching-point candidates acquired by the automatic procedure, some are false-alarm, though there is no missing. Hence, the system provides a graphical user interface (GUI) to verify the acquired touching-point candidates, which is illustrated in Fig 3.

Fig 3. The GUI of touching-point verification.

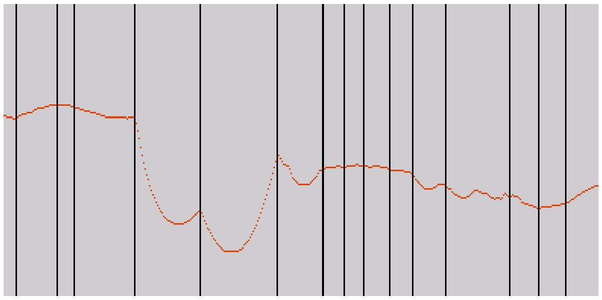

Fig 4. Ball motion curve with the touching points after the verification. They also are the groundtruth.

4. Touching-Player Analysis

For a touching-point, we can find the closest person to the ball, who is the touching-player. This section produces the touching-player candidates using the relations between the touching-place and the roles of players.

4.1. Touching-player extraction

Field and object extraction: For a given frame, we find its dominant color, which is the field color if its dominant color is in the field color range of the video that is obtained in advance by a statistical procedure. The field color range of the frame is made from its dominant color. Every pixel with color in the field color range is colored in the field color. A region growing procedure is used to find the field boundary by finding the largest field color region. "Non-field" object is assigned a field color too. Now a seed-growing procedure is used to find the connected components that are in the ball neighborhood in the field.

Evaluation method: A method is developed to evaluate whether an object is a person. Then, we discard objects that are not persons. Subsequently, a SVM will recognize their team further.

Let $f1, f2, \cdots, fk$ be all features that we use to evaluate these objects. Assume that $K(F)$ is the set of all detected objects in the neighborhood of the ball in the field of frame F. Let $o \in K(F)$ be an object and $P(o \mid f_i)$ $(i = 1, 2, \cdots, k)$ be the probability that indicates how likely o is a person with respect to the feature fi. The choice of features allows us to assume that they are independent with relatively small error. With this assumption the probability that o is a person has a simple formula by the multiplication rule of probability.

$$P(o) = \prod_{i=1}^{k} P(o \mid f_i) \qquad (1)$$

The object o is removed if probability P(o(is small. Only the person who is nearest to the ball is kept for further judging which team touches the ball. The cases that the nearest persons are not the touching-players are very few. Hence, this can achieve very accurate result.

4.2. Teamship recognition

For the person extracted for each touching point (frame) as described above, we use a Support Vector Machine (SVM) on the color histogram of the person to recognize his team.

Color histogram: In a soccer game, the people in the soccer field are in five categories: the players in Team A and B, goalkeepers in Team A and B, and the referee. The jerseys of these five categories of people are in five different colors. Hence, the color histogram of a person can determine which team he belongs to.

For each type of people, we manually identify a color that differentiates them from other people in advance. For each such color, we build several color bins which span a range of color around it. Then for each person we calculate the distribution of pixels which has colors falling into the pre-built bins. This distribution forms a color histogram of a person, which is used to evaluate his team through a SVM.

Fig 5 shows the ball motion curve over frames in which black vertical bars indicate the verified touching-points; As and Bs indicate the teams that touches the ball at that frame. Here A and B represent Senegal team and Turkey team respectively. They wear white and red jerseys respectively in the tested video.

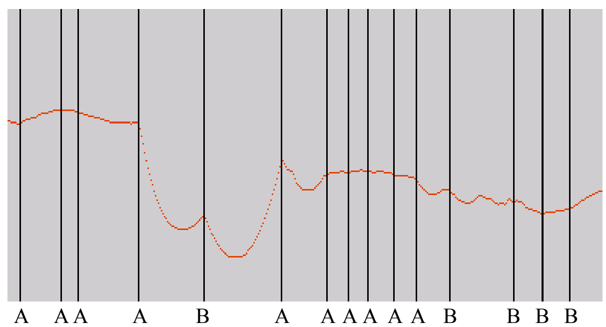

Fig 5. Ball motion curve with touching points and their team labels. As indicate Senegal, Bs Turkey.

4.3. Touching-player candidate inference

We can know which players are in the playfield according to the record of player substitution. In a team, each player plays a role and he has her own action area, as illustrated in Fig 7. Inversely, we can narrow down the candidates of touching-players if we know the touching-place and the roles of all the players in the playfield of a team. The soccer field has a fixed line setting, which form various several objects in video such as ellipse, straight lines, and goalmouth. These objects can be detected using the developed algorithms in several papers [2, 6]. Based on these detected objects we can classify the frames into five classes for the left team: Strike Zone, Midfield Zone, Defense Zone, the border of Strike Zone and Midfield Zone, and the border of Midfield Zone and Defense Zone. When the touching takes place in a zone, the corresponding players are considered as the touching-player candidates.

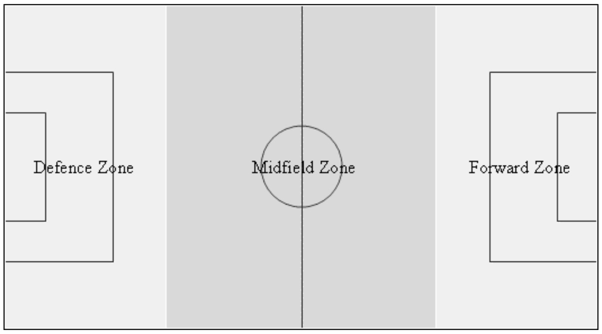

Fig 6. The three areas of soccer field in the viewpoint of the left team.

4.4. Touching-player finalization

We display each acquired touching-point, its touching-player candidates, and the corresponding frame. Thus, the operator can identify the touching-player and assign the touching-player to the touching-point, as is illustrated in Fig 7. The system records the touching-players and obtains the possession of each player. The operator can play the mini-video around the touching point when he cannot recognize the player from the touching frame.

Fig 7. The GUI of touching-player finalization.

5. Evaluation Criteria and Performance

We cut a play segment at middle of two adjacent touching-points. A possession segment is defined as the segment between two adjacent cutting points. We say that the touching-player possesses this segment. The *player-possession* of a player is the assembly of all his possession segments. Tests have been conducted on the sequences of the video of the first half game of the match of 2002 FIFA World Cup Quarter Finals----Senegal vs Turkey, which was held in Osaka, Japan. Table 1 shows the T-p (touching-point) candidate acquisition performance, in which "B-p", "T-p-c" are the abbreviations of "ball-player" and "touching-point candidates". The smaller "T-p-c", the more time is saved in acquiring the touching points.

Table 1. T-p candidate acquisition performance.

	groundtruth		T-p candidate acquisition		
Seq.	frame	T-p	B-P	T-p-c	miss
27785-29256	1472	41	218	197	0
45533-47004	1472	30	223	188	0

Table 2 shows the performance of touching-player candidate inference, in which column "correct" ("*wrong*") gives the numbers of the touching points that the touching-players are (*are not*) in the given candidates. The average number of candidates is about 4.

Table 2. Touching-player candidate inference.

Seq.	# T-p	# correct	# wrong	%
67748-70744	72	56	16	77.78%
41174_42450	33	23	10	69.70%

6. Conclusions and Discussions

We have presented a semi-auto system for acquiring player-possession for broadcast soccer video, which significantly reduce the manual work, compared with pure manual work. This system has facilitated the player-possession acquisition in three aspects. First, it can automatically extract the touching-point candidates without missing. Thus, it significantly reduces the work because the number of touching-points is much fewer than the number of the total frames. Second, it provides the touching-player candidates and GUI so that operator can determine the touching-player through selections and picks. Thus, the operator can save a lot of time, comparing without the system.

The system leaves room for improvement and extension: (1) there is a room to further save time by improving the algorithm for detecting the touching-point candidates; (2) the results in tracking players may facilitate to tell who is closest to the ball in 3D space; (3) the text from video, comments, and internet will help the player-possession analysis.

7. References

[1] H. Byun and S. W. Lee. A survey on pattern recognition applications of support vector machines, Int'l J. of Pat. Recog. and Artificial Intel., vol 17, no 3, pp 459-486, 2003.

[2] Y. Gong, T. S. Lim, H. C. Chua, H. J. Zhang, and M. Sakauchi. Automatic parsing of TV soccer programs, 2nd Int. C. on Multimedia Comp. and Sys. , 1995, 167-174.

[3] T. Kawashiima, K. Yoshino and Y. Aok. Qualitative image analysis of group behavior, CVPR 94, pp. , June 1994.

[4] S. Mayhew and H. Wenger. Time-motion analysis of professional soccer, *J. of Human Movements Studies*, vol. 11, pp. 49-52, 1985.

[5] T. Taki, J. Hasegawa, and T. Fukumura. Development of motion analysis system for quantitative evaluation of teamwork in soccer games, ICIP 96, pp. 815--818, 1996.

[6] K. W. Wan, J. H. Lim, C. Xu, X. Yu. Real-time camera field view tracking in soccer video, ICASSP 03, 185-188, 2003.

[7] L. Xie, P. Xu, S. F. Chang, A. Divakaran, and H. Sun. Structure analysis of soccer video with domain knowledge and hidden Markov models, Pattern Rec. Letter, vol 25, pp767-775, 2004.

[8] X. Yu, H. W. Leong, J. H. Lim, Q. Tian, and Z. Jiang. Team-possession analysis for broadcast soccer video based on ball trajectory, PCM 03, pp. 1811-1815, 2003.

[9] X. Yu, C. Xu, H. W. Leong, Q. Tian, Q. Tang, and K. W. Wan. Trajectory-based ball detection and tracking with applications to semantic analysis of broadcast soccer video, *ACM MM 03*, pp11-20, 2003.