Tracking soccer players using the graph representation

Pascual Figueroa & Neucimar Leite
Institute of Computing
State University of Campinas
C.P. 6176, 13083-970 Campinas, Brasil
{pascual,neucimar}@ic.unicamp.br

Ricardo M. L. Barros
Physical Education Faculty
State University of Campinas
C.P. 6134, 13083-851 Campinas, Brasil
ricardo@fef.unicamp.br

Isaac Cohen & Gerard Medioni
Institute for Robotics and Intelligent Systems
University of Southern California, USA
{medioni,cohen}@usc.edu

Abstract

In this work, we consider the problem of tracking soccer players during a game by using multiple cameras. The main goal consists in finding the position of the players on the pitch at each instance of time. The occlusion is treated by splitting segmented blobs and the tracking is performed using a graph representation, where nodes correspond to the blobs obtained by image segmentation and edges represent the distance between the blobs.

1. Introduction

The movement of the soccer players on the field, as a function of time, is a useful information that can contribute for improving the performance of players at different positions. The measured values may be associated to physiological variables as well as to technical and tactical information [3].

One of the first studies about the players movement using video cameras was done by Mayhew and Wenger [3]. They filmed one player during a game and estimated their position and spent time for each activity, such as walking, running, jogging, staying, etc.

Aiming a better quantification of the kinematical variables of the soccer player movement, Erdman [5] filmed a soccer game with one stationary TV camera (using wide-angle lens of 130°) and a transparent squared sheet adapted to the monitor of the screen.

Recently advances in video technology and computer processing performance have motivated the interest of researchers in using computer vision and image processing techniques for the automatic analysis of the sport games by videogrammetry. Taki et al. [2] presented a method for a quantitative evaluation of the team work in soccer games. For this purpose, they used many static cameras and isolated players were tracked using template matching. Choi et al. [1] used the TV broadcasted images and the players were tracked by template matching and, in case of occlusions, they used the technique of histogram backprojection.

One of the most difficult problem in tracking soccer players concerns the occlusion and the players congestions which occur, especially, in cases of free kick or corner. Iwase and Saito in [4] used 8 cameras covering the region of the goal in order to treat this problem. However, this kind of solution is expensive and the occlusion problem are not totally resolved.

In this work, we propose a method for tracking soccer players using many static cameras which, together, should cover the whole playing field. As we will see elsewhere, occlusions and congestions are treated by splitting blobs containing two or more players using a simple player model and morphological filters.

This papers is organized as follows: the next section introduces the segmentation algorithm. Section three describes the tracking procedure. Section four shows some results of the method applied to real sequences and, finally, the last section gives some conclusions.

2. Segmentation of the players

Background subtraction is a simple and very common method used for segmenting moving objects which consists of the difference between a set of images and its background model. To consider changes of the environment such as illumination, shadows, background objects etc, these methods

Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)
1051-4651/04 $ 20.00 IEEE

need to regularly update the background representation. For this purpose, some statistical adaptive methods [6] can be used. These methods update the background model considering the background pixel as a Gaussian distribution and work well in relatively simple scenes. The segmentation algorithm considered in this work consists of the following basic steps:

- Background extraction using statistical methods,
- Difference between the current frame and the image corresponding to the extracted background,
- Morphological filtering (opening and closing) to eliminate noise,
- Image binarization,
- Labeling of connected pixels and extraction of regions as *blobs*.

3. Tracking of the players

As the main difficulty of the tracking process concerns the temporal occlusions of the objects, the splitting of the blobs aiming at separating or isolating the players is an important step in this direction. The splitting takes into consideration the spatio-temporal information of the image sequence. The spatial information is explored by considering the size, shape and colors of the blobs, while the temporal information explores the relation between blobs of different frames. In this work, we use a graph representation to define this temporal dependence.

The graph is constructed from the set of blobs obtained during the segmentation step in such a way that nodes represent blobs, and edges represent the distance between these blobs. This representation model allows us to approach better the correspondence problem of the objects which can help not only in the splitting of the blobs but also to track them along the sequence. The tracking of each player is performed by searching a optimal path in the graph.

3.1. Number of components definition

The correct determination of the number of components in a blob is important for the blob splitting and making correct decisions during the tracking. This number is difficult to determine, specially when a blob containing more than one player in one frame is split or joined with other blobs in the next frame.

To solve this problem we group the nodes considering the edges between them, as it is shown in Fig. 1. Two nodes v_1 and v_2 in frame t belong to the same group if there is any node u, in frame $t+1$, so that there exist edges (v_1, u) and (v_2, u) in the graph.

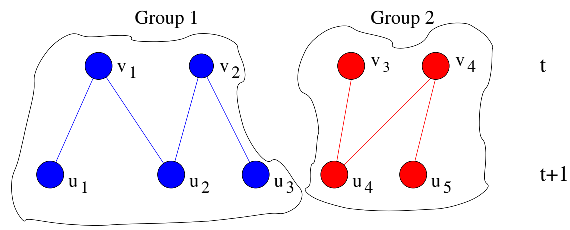

Figure 1. Nodes grouped by common edges.

The area of the blobs is the parameter used here to define the number of components of each node. For each group of nodes in frame t, belonging to the same group, it is defined a subdivision process of the objects in frame $t+1$.

3.2. Blob color definition

In soccer, the uniform of the teams have different colors. This information is used in the team's identification and in the solution of some occlusion problems. Choi et al. [1] used the histogram backprojection method to identify the teams for full color player uniform. With our camera positioning setup it is difficult to discriminate colors and, thus, we work only with the intensity information of the players uniform. Generally, a player can be modeled as a group of many regions, each region having some predominant colors.

In this work, we try to divide the model of the player in two or more regions, so that each region represents a part of the team's uniform, i.e., t-shirt, short, socks, etc. For each region, we define a filtering based on the vertical intensity distribution of the blobs. The two vertical lines in Fig. 2 represent the values, T_1 and T_2, defining the limits used to determine the significant intensity values of each interest region. These limit values is defined as the minimal and maximal mean value of the vertical distribution. For each region R_i, we count the number of pixels, p, of a blob that are outside the thresholds T_1 and T_2, defining the following values

$$S1_{R_i} = \#\{p\} \quad \forall p \in R_i \wedge R_i(p) < T1$$
$$S2_{R_i} = \#\{p\} \quad \forall p \in R_i \wedge R_i(p) > T2$$

These values are associated with the more discriminat intensity values of each region based on the considered vertical distribution of the blobs. Thus, the color or intensity of each region, R_i, can be defined as

$$C_{R_i} = \begin{cases} 1, & \text{if } \frac{S1_{R_i}}{S1_{R_i}+S2_{R_i}} > 0.6 \\ 2, & \text{if } \frac{S1_{R_i}}{S1_{R_i}+S2_{R_i}} < 0.4 \\ 0, & \text{otherwise} \end{cases}$$

Finally, the color of the blob, representing the team identification, can be defined from the color information of a region or by a combination of colors of different regions.

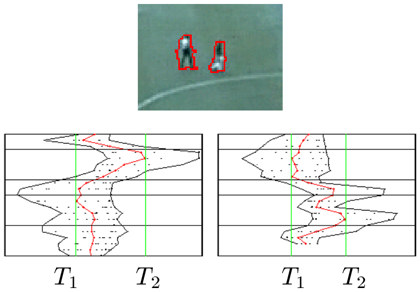

Figure 2. Vertical intensity distribution

3.3. Splitting the blobs

During the segmentation step, some objects can be grouped into one blob due to the small distance between them or by effects such as shadows and noises. Our first approach to the splitting problems tries to isolate players which are linked by short connections, as it is shown in Fig. 4. Further, a blob can be split by considering the constructed graph and the model of the blobs.

To illustrate the splitting process, we selected some regions of a video sequence (Fig. 3) presenting some occlusions. Fig. 4 shows the blobs obtained after the segmentation step, and Fig. 5 shows the corresponding graph and the number of components in each node.

Figure 3. Some cases of occlusions.

Figure 4. The blobs related to the sequence.

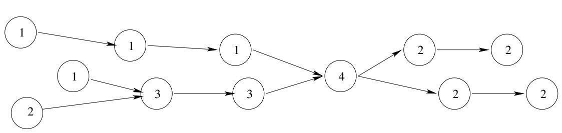

Figure 5. The graph representation.

3.3.1 Splitting using morphological operators

In order to eliminate short connections, we consider the blobs as regions of the original images and apply a sequence of morphological operations for pos-processing and splitting these blobs. First, the original image (Fig. 4) is eroded vertically n times (n depends on the players position on the field). The next step consists of a conditional thickening of these shaded regions, limited to the corresponding blob contours, thus resulting in Fig. 6.

3.3.2 Splitting using the blobs model

A blob in the current frame can be split in two or more blobs depending on the number of components of the corresponding node in the graph and on its configuration in the previous frame. From the isolated blobs (before the occlusion) is taken the model of each blob and these are fitted in the joined blob for searching the best matching.

One blob can be split vertically or horizontally depending on the position of the related blobs in the previous frame and the number of its components. As the splitting of more than three players becomes more complex, here we consider only two or three player splitting at same time. If the size of the blobs is not big enough to split horizontally or vertically then the objects are considered to be in total occlusion and they share the same position.

An example of the splitting by model is shown in the figure 7.

Figure 6. after splitting by segmentation

Figure 7. after splitting by using blob model

3.4. Tracking each player

In order to start tracking a player, its initial blob is defined and the corresponding node is found in the graph. At each step, we traverse the graph by considering a minimal path, using the distance information between the blobs. This method represents an easy way to track isolated players since there is only one edge at each step to be considered.

The tracking of the players in case of contact or occlusions by other players is more difficult. Although the splitting of the blobs is supposed to eliminate this problem, there are still situations in which one node may contain more than one player.

Player	Number of solved occlusions	Number of non-solved occlusions
defender	38	6
mid-fielder	36	3
forwarder	57	9
mean	43.6 (87.9%)	6 (12.1%)

Table 1. Evaluation of the tracking algorithm for 3 players during 10 minutes of game.

In this work we consider mainly cases of short temporal occlusions or contacts of more than two players. A special attention is given here to the case of two player occlusions since it is one of the most common situation in soccer. The tracking, in such a case, is performed by considering that the trajectory is the same for both of them. To maintain the right path of the players during the tracking, when the players are separated after occlusions, we consider the color of the blob together with the distance information between blobs conveyed by the graph. During an occlusion of two players, we also consider the direction of their trajectory in order to decide the correct paths, specially, in the case when players in contact belong to the same team.

At each step of the tracking method, the real 2D coordinates of the players in the field is reconstructed, using the calibration parameters and the image coordinates representing the players location in the image. Initially, it is also defined the field of view of each camera. With this information, we can determine exactly from which cameras a tracked player is visible and choose the one that better focus on this player.

4. Applications and some results

To illustrate the method proposed here, we consider a game of a Brazilian league filmed by four digital cameras, each of them placed at one side of the pitch and at the highest location of the stadium. The position and the distance between the cameras were chosen arbitrarily, in such a way that one camera covered at least one fourth of the playing field. The game was filmed in the afternoon of a sunny day with a partially clouded sky. A 10 minute game was processed and the tracking was performed for the 3 players of one team, one defender, one mid-fielder and one forwarder.

Table 1 illustrates some numerical results of the tracking test. The second column shows, for each player, the number of occlusions or contacts correctly solved, i.e., the cases in which the blobs were correctly tracked during the occlusions. The third column shows the situation in which the tracking failed because of two or more player occlusions. The last line of the table shows the mean for each case, considering all the three players.

The quality of the image is very important in the solution of occlusions, and when farther is the player against the camera, their size become smaller, and consequently it is more difficult to recognize. One way to improve the tracking may be using more cameras positioned in the other side.

5. Conclusion

In this work we presented a method for tracking soccer players using many static video cameras. In order to cover whole playing field and still conserve good players feature in the image, we defined experimentally, a minimal configuration of four cameras. The tracking algorithm is based on searching paths in a graph and the situations of occlusion or being in contact were treated by splitting the blobs and using the blob color information. The obtained results show the effectiveness of the proposed method which can be used to track all the players during a whole game.

Acknowledgement

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

References

[1] S. Choi, Y. Seo, H. Kim and K.S. Hong, ”Where are the ball and players? Soccer game analysis with color-based tracking and image mosaick,” *In Proc. Int. Conf. on Image Analysis and Processing*, pp. 83-91, 1997.

[2] T. Taki, J. Hasegawa, and T. Fukumura, ”Development of motion analysis system for quantitative evaluation of team work in soccer game”, *In Proc. Int. Conf. on Image Processing”*, pp. 815-818, 1996

[3] S. Mayhew and H. Wenger, ”Time-motion analysis of professional soccer”. *Journal of Human Movements Studies*, v. 11 pp. 49-52, 1985.

[4] S. Iwase and H. Saito, ”Tracking soccer players based on homography among multiple views”, *In Proc. of SPIE 2003*, v. 5150, pp. 283-292, 2003.

[5] W. Erdmann, ”Quantification of games - preliminary kinematic investigation in soccer”, *Science and Football II*, pp. 174-165, 1991.

[6] C. Stauffer and W.E. Grimson, ”Adaptive background mixture for real-time tracking”, *In Proc. Int. Conf. on Computer Vision and Pattern Recognition*, pp. 246-252, 1999.