Tracking Players and Estimation of the 3D Position of a Ball in Soccer Games *

Yoshinori Ohno, Jun Miura, and Yoshiaki Shirai
Dept. of Computer-Controlled Mechanical Systems, Osaka University
2-1, Yamadaoka, Suita, Osaka 565-0871, JAPAN
{y-ohno,jun,shirai}@cv.mech.eng.osaka-u.ac.jp

Abstract

In soccer games, understanding the movement of players and a ball is essential for the analysis of matches or tactics. In this paper, we present a system to track players and a ball and to estimate their positions from video images. Our system tracks players by extracting shirt and pants regions and can cope with the posture change and occlusion by considering their colors, positions, and velocities in the image. The system extracts ball candidates by using the color and motion information, and determines the ball among them based on motion continuity. To determine the player who is holding the ball, the position of players on the field and the 3D position of the ball are estimated. The ball position is estimated by fitting a physical model of movement in the 3D space to the observed ball trajectory. Experimental results on real image sequences show the effectiveness of the system.

1 Introduction

Recently in the sports domain, especially in soccer, the demand for analyzing matches or tactics is increasing. Viewers, on the other hand, may want to watch only exciting scenes or the digest from a long game. To meet all these requirements, it is essential to know the movement of players and a ball. Most researches track players by using template matching[1, 2, 3]. In these method, however, the operator often has to specify the position of players manually during occlusion. Moreover, most methods do not track a ball or track a ball only in easy cases[4, 5]. The position of a ball is necessary for analyzing matches or tactics.

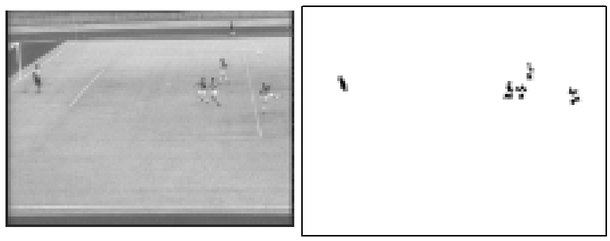
This paper presents a system to automatically track players and a ball in soccer games in the images taken by fixed cameras. The system can cope with occlusion and the posture change and can calculate the position of the players on the field and the position of the ball in the 3D space.

*Proc. ICPR-2000, pp. 145–148.

2 Finding players

2.1 Position in the image

Our system first finds the position of players in the image. We model the color distribution of the uniform by a right-angled parallelepiped in the YIQ-space in advance. The system extracts shirt and pants regions from images by using this model. Since these extracted regions include non-uniform regions whose color is similar to that of the uniform, the system determines player regions among them using the positional relationship between those regions. The system searches for pairs of shirt and pants regions which align vertically in the image, based on the assumption that players stand upright at the beginning of games. Fig. 1 shows an original image and the result of player extraction.



(a) Original image (b) Result of extraction

Figure 1. Extraction of players

2.2 Position on the field

To estimate the players' position on the field, a camera needs to be calibrated. But it is difficult to artificially set marks on the field for calibration. We therefore use the crossing points of the lines because the length of the lines is determined by the rule. Assuming that we use an ideal pin hole camera, the transformation between the image and the field is obtained. So we can estimate the position of players on the field by detecting footing points of them. By using the position on the field and the height of players, we can calibrate a camera in the 3D space.

3 Tracking isolated players

Since it is not efficient to search the whole image for players every frame, the system limits the search area by predicting the movement of players. Since a player moves at an almost constant speed during a short period of time, the next position of the player can be predicted by using a simple linear extrapolation. The system searches only a neighboring area of the predicted position for a pair of region. The system thus avoids mismatch of players between frames.

It is desirable that players are always tracked by extracting pairs of regions, but the system sometimes fails to detect a pair of regions due to noise in the image or the posture change of players. To cope with such cases, the position of a player can be determined from either a shirt region or a pants region as long as there are no other players around the predicted position.

4 Tracking overlapping players

4.1 Two overlapping players

Since two players in different teams wear differently colored uniforms, we can distinguish them by color. The system, therefore, detects occlusion when the regions of either of the players cannot be extracted. The occluding player is determined by the color of extracted regions, and the other is considered to be occluded. The system then searches an area around the occluding player for a pair of regions of the occluded player, and regards newly appearing regions as the occluded player.

On the other hand, when two players in the same team overlap in the image, we cannot distinguish them only by color. In this case, we use the position of the players in the image. Considering that we take images from a high position out of the field, a player who is closer to the camera is detected at a lower position in the image. The system, therefore, detects occlusion when regions of the two players merge into one. A lower player in the image is considered to be an occluding player. When the region splits into two, the lower region is considered to correspond to the occluding player.

4.2 More than two overlapping players

Tracking more than two players is difficult because one image may correspond to various cases. Fig. 2

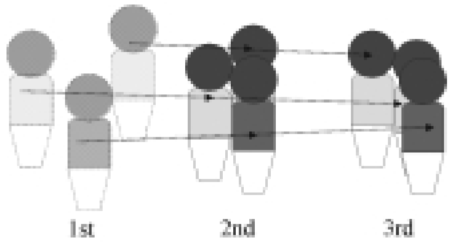


Figure 2. Determination of position

shows an example of overlapping three players in three successive frames. The system detects the same two pairs of regions in the 2nd and the 3rd frame, although the players change places actually. To resolve this ambiguity, we use the average velocity of the players. The system predicts the next position of each player by using his average velocity, and determines the corresponding region by considering the positional relationship in the image and color of regions. Fig. 3 shows the result of tracking overlapping players. The system successfully tracks three players even when two white-shirt players changed their positions behind a dark-shirt player.

Though we only show the occlusion among three players here, this method, which uses the color, the vertical position in the image, and the velocity of players, is easily extended to cope with the occlusion among more than three players.

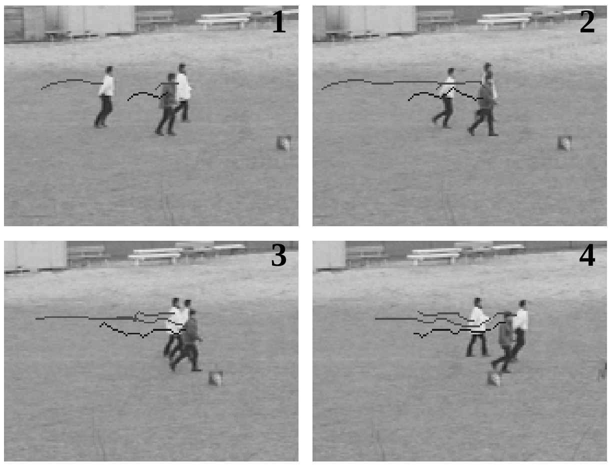


Figure 3. Tracking players close in the image

5 Tracking a ball

5.1 Utilizing motion information

Tracking a ball is difficult because it is small in the image and sometimes moves very fast. Though we can relatively easily recognize a ball in a stationary state, it is difficult to recognize a fast moving ball only by color. To solve this problem, we use motion information. By examining the difference between consecutive frames,

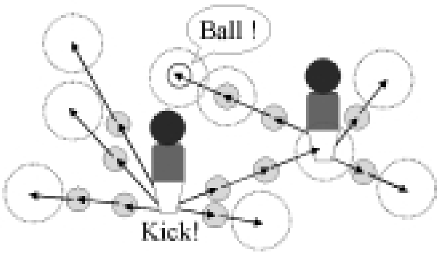


Figure 4. Search for regions in radial directions

the system extracts motion regions as ball candidates. It is, however, difficult to determine a ball region from only one image. Since a ball kicked by a player moves away from him almost in the radial direction, the system searches for motion regions in the radial direction from the player in the subsequent frames. A candidate is retained if the corresponding motion region is found and a candidate is deleted if it is not found. This process is repeated until one candidate is selected as the ball region. The system considers a group of detected regions in consecutive frames as the trajectory of the ball.

In actual games, however, many players usually gather around the ball. The kicked ball may be occluded by another player. To cope with this situation, the system also searches for motion regions in all radial directions from the player as shown in Fig. 4. In this way, the ball can be tracked in a complicated situation. Fig. 5 shows a result of tracking a ball. Circles denote the ball candidates in the present frame, and small squares denote the ball candidates in the past frames. The left hand player in the image kicks the ball and the ball gets into a crowd of players. Another player kicks a ball again, and the ball comes out of the crowd. The system successfully tracks a ball in such a complicated situation.

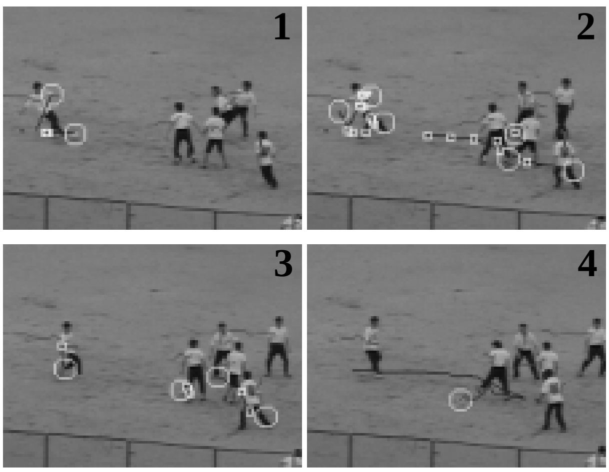


Figure 5. Tracking a ball kicked by a player

5.2 Determining whether a player holds the ball or not

Another problem in tracking a ball is that the ball is likely to overlap with or to be occluded by a player while he holds the ball. To solve this problem, the system stops tracking a ball when a player and a ball overlap in the image, and temporarily considers that the player holds the ball. The system then keeps searching a neighboring area of the player for the ball and tracks again when the ball region is found.

But the player does not necessarily hold the ball even when a player and a ball overlap in the image. To judge whether a player has touched the ball or not in the image, the system approximates a ball trajectory in the image to a parabola. If a newly found ball region is on this trajectory, the system considers that the player did not touch the ball.

5.3 Estimating 3D position of the ball

Even though we know the ball is held, it is sometimes difficult to know a player holding the ball from the image. In Fig. 6, for example, it is difficult to judge whether the player in the back kicks the ball or the player in the front heads the ball only from the position in the image. To resolve this ambiguity, we use the ball position in the 3D space.

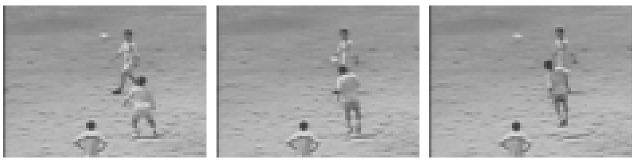


Figure 6. Ambiguous scene

In order to estimate the ball position in the 3D space, we estimate the 3D trajectory of the ball by assuming that the ball motion is determined by the gravity, and the air friction, which is expressed as:

$$\begin{aligned} x(t) &= x(0) + t v_x(0) - \alpha \int_0^t t\, v_x(t)\, dt, \\ y(t) &= y(0) + t v_y(0) - \alpha \int_0^t t\, v_y(t)\, dt, \\ z(t) &= z(0) + t v_z(0) - \int_0^t t\, \{g + \alpha v_z(t)\}\, dt, \end{aligned} \tag{1}$$

where $(x(t), y(t), z(t))$ and $(v_x(t), v_y(t), v_z(t))$ denote the position and the velocity of the ball at time t, g denotes the acceleration of gravity, and α denotes the friction coefficient. Since the transformation between the position in the 3D space and the position in the image is known, the estimated position in the image at each moment depends on the initial position and the initial velocity in the 3D space. We therefore estimate their values which minimize the following sum of squared difference between the estimated position and the observed position in the image:

$$\sum_t \left[\{x_p(t) - x_c(t)\}^2 + \{y_p(t) - y_c(t)\}^2 \right], \qquad (2)$$

where $(x_p(t), y_p(t))$ denotes the projected position of $(x(t), y(t), z(t))$ in eq. (2) and $(x_c(t), y_c(t))$ denotes the observed position in the image. The ball, however, does not always move in the air while it is separated from players in the image. The ball sometimes rolls on the ground. The ball is considered to be on the ground if the residual (eq. (2)) is too large.

We show the result of tracking and estimation in Fig. 7. In this figure, we can recognize that the player at the lower side of the ball in the image is nearest to the ball in the 3D space, though the player at the left side of the ball is nearest to the ball in the image.

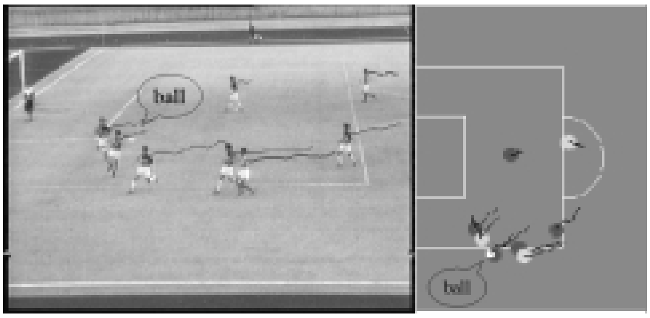


Figure 7. Tracking result left and top view right

6 Tracking by using multiple cameras

We tested our system on real image sequences of a soccer game using multiple cameras. Fig. 8 shows the result of tracking players and a ball, and the top view of the field in the 12th frame. The system successfully tracks players and a ball, and estimates thier positions.

7 Conclusion

We have described a system to track players and a ball in soccer games. The system can cope with occlusion by using the color, the vertical position in the image, and the velocity of players. The system can also estimate the 3D position of the ball. Experimental results using real image sequences show the effectiveness of the system. The system can basically be applied to other team sports in which players wear uniforms. As a future work, we are now planning to analyze matches or tactics by using the position of players and a ball.

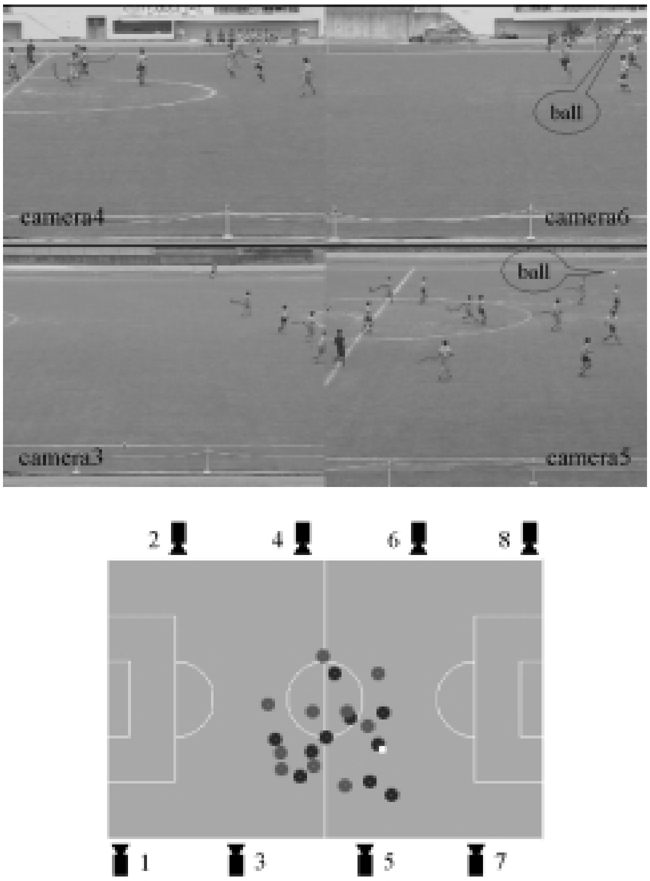


Figure 8. Result and camera arrangement

Acknowledgment

The authors would like to thank Prof. Hasegawa and Mr. Taki of Chukyo University for providing us some of the video sequences used in this paper.

References

[1] Y. Gong, C. Hock-Chuan and L. T. Sin. "An Automatic Video Parser for TV Soccer Games", Proc. of ACCV, pp. 509-513, 1995.

[2] D. Yow, B. L. Yeo, M. Yeung and B. Liu, "Analysis and Presentation of Soccer Highlights from Digital Video", Proc. of ACCV, pp. 499-503, 1995.

[3] T. Taki, J. Hasegawa and T. Fukumura. "Development of Motion Analysis System for Quantitative Evaluation of Teamwork in Soccer Games", Proc. of ICIP, pp. 815-818, 1996.

[4] Y. Seo, S. Choi, H. Kim and K. S. Hong. "Where Are the Ball and Players? Soccer Game Analysis with Color-Based Tracking and Image Mosaick", Proc. of ICIAP, pp. 196-203, 1997.

[5] Y. Ohno, J. Miura and Y. Shirai. "Tracking Players and a Ball in Soccer Games", Proc. of Int. Conf. on MFI, pp. 147-152, 1999.