

Where Are the Ball and Players? Soccer Game Analysis with Color-Based Tracking and Image Mosaick

Yongduek Seo, Sunghoon Choi, Hyunwoo Kim, Ki-Sang Hong
e-mail : {dragon, shoon, pccc, hongks}@indy.postech.ac.kr

Dept. of EE, Pohang University of Science and Technology
San 31 Hyoja Dong, Pohang, 790-784, Republic of Korea

Abstract. *Knowing the locations of the players and the ball on a ground field is important for soccer game analysis. Given an image sequence, we address three main problems: 1) ground field extraction, 2) player and ball tracking and team identification and 3) absolute player positioning. The region of ground field is extracted on the basis of color information, within which all the other processing is restricted. Players are tracked by template matching and Kalman filtering. Occlusion reasoning is done by color histogram back-projection. To find the location of a player, a field model is constructed and a transformation between the input image and the field model is computed using feature points. Otherwise, an image-based mosaicking technique is applied. Using this image-to-model transformation, the absolute positions and the trajectories of players on the field model are determined. We tested our method on real image sequences and the experimental results are given.*

1 Introduction

In this work we are to solve the problem of tracking the players and the ball in an image sequence of a soccer game which is captured from a broadcasted TV signal. Our task is to track all the players and the ball presented in the sequence and to find their positions on the ground field. Because of inter-field motion difference, we use only the even field of a frame. Sample images of the sequence are shown in Figure 1. Three main problems for automatic soccer game analysis are addressed here:

1. The ground field should be extracted in order to track players, to find the half line, a side line and the center circle, and to make mosaic image for computing image-to-model transformation. We use a color histogram information under the assumption that the region of the ground is nearly green and occupies most areas of the images
2. Each of the players and the ball should be tracked and identified in the image sequence. Players move non-rigidly, frequently collide with each other and are occluded by other players. Template matching and the Kalman filtering are applied for player tracking. Occlusion reasoning is done by the color histogram back-projection method[7]. Players of the same team are grouped together (team identification). This is done by comparing the spatial color distribution of the player templates. A similar method is used in ball tracking. However, since the ball is too small to track alone, a method is devised to reason the location of it.

3. The absolute positions of the players should be known. Because there are a small number of features in soccer sequence, we cannot easily find the absolute locations of players on the field. A field model is constructed and image-to-model transformations are computed to attack the problem. When features are found, the transformations are obtained directly by computing 3×3 homography matrices. Otherwise image mosaicking technique [5] is used to find out the image-to-model transformation. Then the trajectory of each player is computed. These trajectories show movement pattern of a player or a group of players. Also it can be used in game analysis after the game or video annotation.

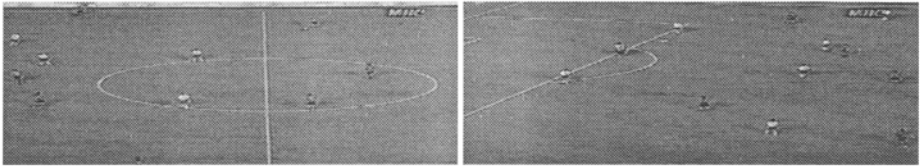


Fig. 1. Two 640x240 images(fields) of a soccer video, which are sampled at thirty frames per second and deinteraced

A comparable work to ours is that of Intille and Bobick [4] in which a method is proposed to track players in images of an American football game. There are lots of yard lines and other features in their images which help to compute the image-to-model transformation. When players are occluded, they select distinctive features based upon the objects in the closed-world (ground field) and track them. In our case, however, only a small number of features are visible, which makes it difficult to compute the image-to-model transformation. We use the color histogram back-projection algorithm to solve the occlusion problem. Kawashima[6] tried to analyze the group behavior of soccer players using the color histogram projection method. Taki[3] developed a motion analysis system to evaluate teamwork qualitatively in soccer games given images from multiple cameras. Yow[8] proposed a method to show a big panoramic highlight scene by applying a mosaicking technique to soccer image sequences. These systems mainly aimed at obtaining qualitative information or just presenting a special view. Our system track each player and the ball, identifies the players' team and computes the absolute locations of the players.

In section 2 the field extraction method is presented. Player and ball tracking and team identification are in section 3, and the field model and the method for computing absolute player position in section 4. Finally concluding remarks and future research topics are given in section 5.

2 Field Extraction

Because all the players are on the field, we extract the region of the ground field first by segmenting out non-field regions like advertisements, and then the players are extracted

on this field. It is assumed that the field has a uniform color of green and occupies large area in the image. So we calculate histogram of each color, R/G/B, and find peak values R_{peak} , G_{peak} , and B_{peak} , respectively. Because the field occupies the largest area in the image, we can assume this color represents the field.

After making a binary image using the histogram and peak values, we apply morphological filtering and extract the largest connected component, the field mask, that represents the ground field and has many holes caused by the colors of the players. Then, the boundary of the ground field region is obtained by using a boundary-following algorithm [1].

3 Player and Ball Tracking

3.1 Player Tracking

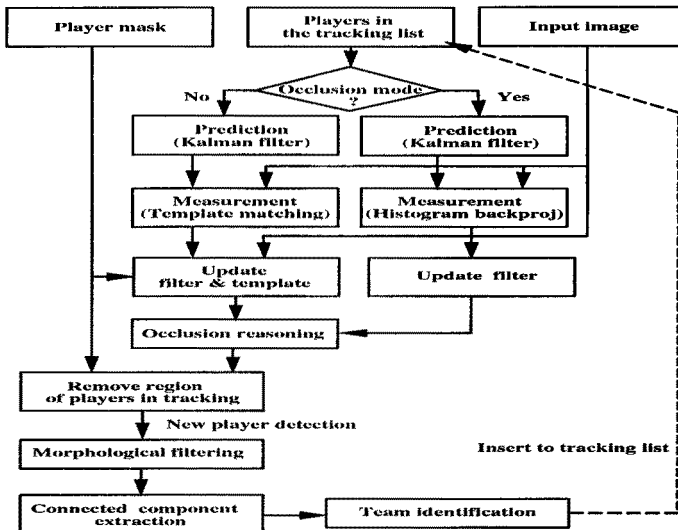


Fig. 2. Procedure for tracking players

For player tracking, we use Kalman filter-based template matching. Initial templates of the players are extracted from player masks using connected component extraction at initialization time. Examples of the templates are shown in Figure 3. In the sequence, new players that do not significantly overlap with the bounding box of a player already tracked are found out and they are inserted into the tracking list. Locations of players in the next frame are predicted by Kalman filtering, and template matching at that location is performed. Finally, the player templates are updated. Figure 2 shows the whole procedure for tracking players, and Figure 3 shows an example of tracking a player based on Kalman filtering. The main problem of player tracking is resolving occlusions. In

this paper we only consider occlusions between different teams, and we use *Histogram Backprojection* [7] for occlusion reasoning. The ratio histogram R_i is computed from the histogram of the player template and image histogram: $R_i = \min \left[\frac{M_i}{I_i}, 1 \right]$ where, M_i is the histogram of player template, I_i is the histogram of input image and i denotes the index of each bin of the color histogram. Then the histogram R_i is back-projected onto the back-projection image, which is then convolved by a mask. The peak in the convolved image is the expected location of the player. In the case that a player is not seen due to occlusion and is not detected, the position of the player is taken as the predicted location from the previous frame.

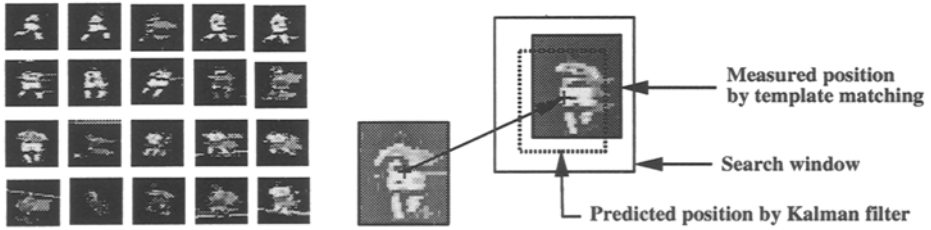


Fig. 3. Left: player templates. Right: player tracking scheme

3.2 Team Identification

Kawashima [6] has analyzed the group behavior of soccer players using color histogram backprojection to isolate players on each team. But the histogram is not appropriate for team identification since different teams can have similar histograms. In this paper, we use vertical distribution of colors.

1. *Compute vertical distribution of R, G, and B, respectively:* Project player templates horizontally and compute average values of R, G, B for each row. The vertical length of the distribution is then normalized.
2. *Compare this distribution with each team's model distribution:* The similarity measure is computed by convolution within a small range because the vertical location of a player in his template can vary slightly.

Figure 4 shows examples of vertical distributions of color values for some players. Similar distributions are obtained when players are on the same team. Figure 5 shows experimental results of team identification using different colors of bounding box.

3.3 Ball Tracking

The method for ball tracking is similar to that of player tracking. However ball tracking is more difficult than player tracking since 1) automatic detection of the ball at the initial time is not easy because it is very small in the image and has a white color which cannot

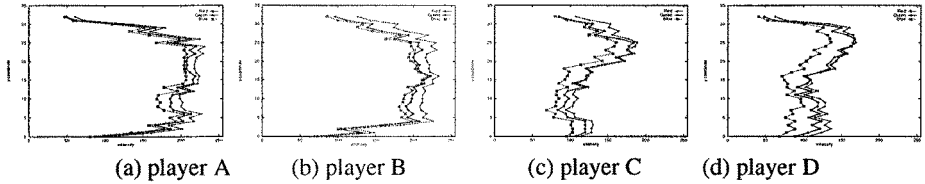


Fig. 4. Vertical color distributions of templates of players. x -axis is the horizontal average of RGB, respectively and y -axis is the normalized height of the player template. Player A and player B (C and D) have a similar distribution because they are the same team.

be discriminated from other noisy parts, and 2) during tracking, if a player has the ball, the ball is occluded by the player. To solve these problems:

1. The position and bounding box of the ball are manually initialized at the starting time.
2. If a player is sufficiently near the ball or he is overlapped with the ball, the player is marked "has ball". Then, ball tracking is stopped and the ball is searched for around the player who has the ball. Search windows for various templates are shown in Figure 5. If the ball is found again in the search window, ball tracking is restarted.

In Figure 5, the ball position is indicated by an arrow and the image at the second row on the right is an example of when a player "has the ball."

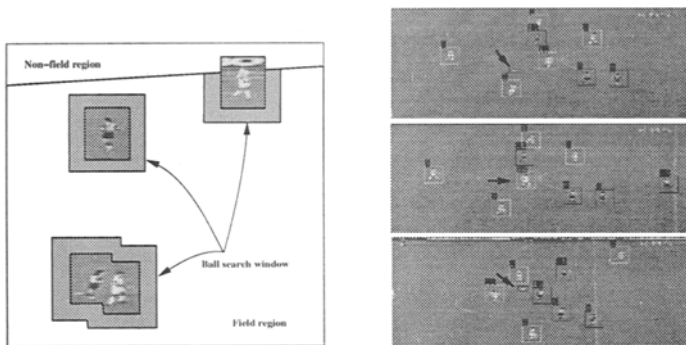


Fig. 5. Ball tracking and players tracking. Left: Ball searching area around the bounding box of players is depicted. Right: The arrow indicates the ball. In the second image the arrow denotes a player who has the ball. Different colors of bounding boxes of the players are due to team identification.

4 The Field Model and Player Position

In this section we describe the field model and how to get a transformation between the image and the model.

4.1 The Field Model

Figure 6 shows the field model. Players are located at certain positions on the field. Since the camera is capable of panning and zooming, the locations of players will change from frame to frame. To find out the positions of the players on the field model, the transformation between the image and the field model must be known. When features such as lines, corner points or the center circle are found in the image, the transformation can be found directly. A simple homographic planar projective transformation¹ can be used to map the image to the field model. From this transformation the positions of the players on the field model are obtained. Figure 6 shows examples of feature points, lines and the center circle. The features are detected automatically to compute image-to-model transformation. The right image of Figure 6 shows computed player positions on the field model.

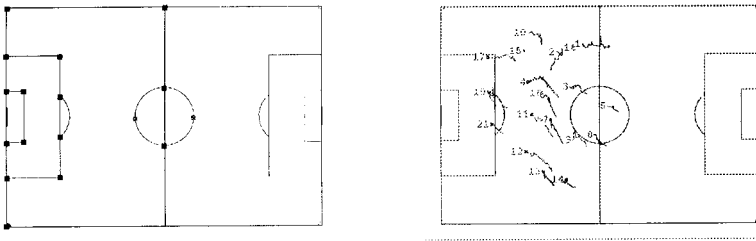


Fig. 6. Left: Field model. The depicted features are detected and used to find the image-to-model transformation. Right: Absolute trajectories of tracked players on the field model for 150 frames. The image-to-model transformation allows us to determine the absolute positions of the players.

4.2 Soccer Image Mosaicking

In Section 4.1, we have described how to compute the projective transformation between an input image and the field model when features are visible. But, if they are not visible, we cannot compute the image-to-model transformation. In this case we compute a transformation between a reference frame and the field model using the image mosaicking method.

To align two images we use the hierarchical direct registration technique [5]. First, pyramids are constructed from the two processed images I_p^1 and I_p^2 of two input images,

¹ This transformation is a 3×3 matrix and has 8 essential parameters, since it is defined up to scale.

and then the motion parameters are estimated in a coarse-fine manner. Within each level the match measure is given by $E(\mathbf{u}) = \sum_{\mathbf{x}} (I_p^1(\mathbf{x}, t) - I_p^2(\mathbf{x} - \mathbf{u}(\mathbf{x}), t - 1))^2$ where $\mathbf{x} = (x, y)$ is the image position, I_p is the processed image and $\mathbf{u}(\mathbf{x}) = (u(x, y), v(x, y))$ is the image motion defined by

$$u(x, y) = \frac{p_1x + p_2y + p_5}{p_7x + p_8y + 1}, \quad v(x, y) = \frac{p_3x + p_4y + p_6}{p_7x + p_8y + 1} \quad (1)$$

$E(\mathbf{u})$ is minimized with respect to the transformation parameters p_i via the Levenberg-Marquardt method. The processed image I_p is given through the following steps:

1. Make an edge image I_e using Sobel algorithm.
2. Compute the histogram of the edge image and select threshold I_{th} at the 85% position of histogram
3. Threshold the input image.

$$I_p(x, y) = \begin{cases} 0 & : I_e(x, y) < I_{th} \\ I_e(x, y) & : I_e(x, y) \geq I_{th} \end{cases} \quad (2)$$

4. Remove the regions of the players in the image.

Figure 7 shows the mosaic image for 150 image frames.

4.3 Player Position

In section 4.1, we found the transformation between the image and the model, and in section 4.2 we constructed the mosaic image. Using the transformations obtained from these procedures we can find the absolute positions of players in the field model. Figure 6 shows absolute locations of the players for 150 frames.

5 Conclusion

In this paper, we have shown how to track players and the ball, and have computed their absolute positions on the real ground field model. The three main problems addressed are ground field extraction, player and ball tracking and player positioning. Using real soccer image sequences, we tested our method and the results are presented.² In future work, we will develop algorithms which can deal with more complex problems as follows:

1. Occlusion reasoning for more complex case
2. Computing the absolute position of the ball during play
3. Automatic scaling of the size of the player
4. More accurate computation of the image-to-model transformation
5. View morphing or graphic synthesis at other viewpoints.
6. Exchanging Advertisements.

² More details can be found at <http://cafe.postech.ac.kr>

References

1. D. H. Ballard, C. M. Brown, *Computer Vision*, Prentice Hall, pp.143-146, 1982.
2. R. M. Haralick, L. G. Shapiro, *Computer and Robot Vision*, Addison Wesley, pp. 622-624, 1992.
3. T. Taki, J. Hasegawa and T. Fukumura, *Development of Motion Analysis System for Quantitative Evaluation of Teamwork in Soccer Games*, In Proc. of IEEE International Conference on Image Processing, Vol. III, pp. 815-818, 1996.
4. S. S. Intille, A. F. Bobick, *Tracking using a local closed-world assumption: Tracking in the football domain*, MIT Media Lab Perceptual Computing Group TR No.296, MIT, August, 1994.
5. M. Irani, D. Anandan, S. Hsu, *Mosaic Based Representation of Video Sequences and Their Applications*, ICCV'95, pp. 605-611.
6. T. Kawashima, K. Yoshino, Y. Aok. *Qualitative image analysis of group behavior*, In. Proc. Comp. Vis. and Pattern Rec., June 1994.
7. M. J. Swain, D. H. Ballard, *Color indexing*, Int. Journal of Computer Vision, 7-1, pp. 11-32, 1991.
8. D. Yow, B.L. Yeo, M. Yeung and B. Liu, *Analysis and Presentation of Soccer Highlight from Digital Video*, In Proc. of Second Asian Conference on Computer Vision, Vol. II, pp. 499-503, 1995.

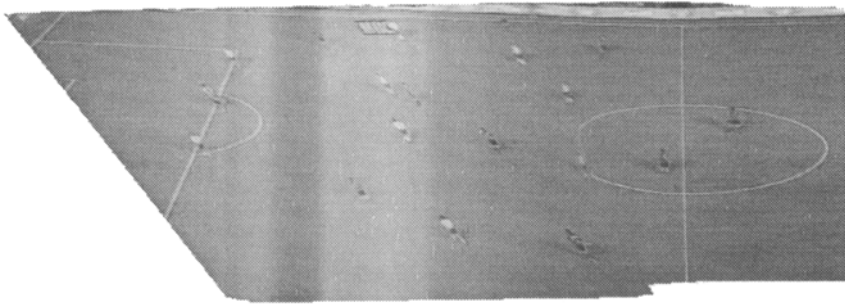


Fig. 7. The mosaic image of 150 frames