Physics-based 3D Position Analysis of a Soccer Ball from Monocular Image Sequences

Taeone Kim, Yongduek Seo, Ki-Sang Hong
Dept. of EE, POSTECH
San 31 Hyoja Dong, Pohang, 790-784, Republic of Korea

Abstract

In this paper, we propose a method for locating 3D position of a soccer ball from monocular image sequence of soccer games. Toward this goal, we adopted ground-model-to-image transformation together with physics-based approach that a ball follows the parabolic trajectory in the air. By using the transformation the heights of a ball can be easily calculated using simple triangular geometric relations given the start and the end position of the ball on the ground. Here the heights of a ball are determined in terms of a player's height. Even if the end position of a ball is not given on the ground due to kicking or heading of a falling ball before it touches the ground, the most probable trajectory can be determined by searching based on the physical fact that the ball follows a parabolic trajectory in the air. We have tested and experimented with a real image sequence the results of which seem promising.

1 Introduction

Recently, computer vision techniques have been widely applied to sports games for automatic extraction and analysis of various information. Intille and Bobick[1] track players using the concept of closed-world in scenes of American football games. They incorporated contextual knowledge such as the number and the type of objects on the ground into low-level tracking. Kawashima and Yoshino[2] proposed a qualitative analysis method of group behavior. They extract a hierarchical representation of the group structure in the sequence. Reid and Zisserman[3] investigated the problem of accurate metrology from uncalibrated video sequences, showing that an accurate position of a goal-bounded soccer ball could be obtained by constructing overhead views using two video sequences captured from different view-points. For the high-level description system, there are SOCCER[4], REPLAI[5] which automatically generate natural language descriptions of the analyzed image sequence. These systems translate visual information obtained from vision systems into verbal information.

In this paper, we present a method of determining 3-dimensional position of a soccer ball from monocular image sequences of soccer games, which is a part of automatic analysis of soccer games. A ball flying in the air plays an important role in a soccer game analysis because of its influence on the players' behavior. The knowledge of the height and the direction of the ball is useful for determining whether the ball is possessed or is being passed by a player, and for predicting who will have the ball. The 3-dimensional information of a soccer ball position may be obtained from binocular image pairs. In this case, however, additional hardware and stereo camera calibration is needed to analyze the images. Our approach is based on the analysis of a monocular TV image sequence, thus eliminating the need for any special setup. Our observation is that there exists a simple triangular geometric relations between the heights of the ball and a player. In section 2, it is shown that the height of a ball from the ground can be easily calculated using the triangular relations given the start and the end positions of the ball on the ground when transforming image sequence into the ground model. Here the heights of a ball are determined in terms of a player's height. In section 3, it deals with the case when the end position of a ball is not given on the ground due to kicking or heading of a falling ball before it touches the ground. It is shown that even in this case the most probable trajectory can be determined by searching based on the physical fact that the ball follows a parabolic trajectory in the air. We have tested and experimented with a real image sequence the results of which are shown in section 4.

2 Finding the height of the ball

2.1 General description

Figure 1 shows a situation on the ground model of a soccer game. A ball is flying on a trajectory from a starting point S to an ending point E through a point H. We assume that the ball trajectory is confined in a plane, which means the line on the ground connecting

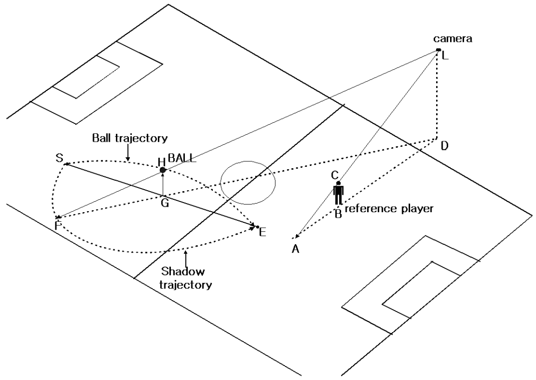


Figure 1: Ball Trajectory, shadow trajectory and a reference player

the points S and E is a straight line and the plane formed by the line and the ball trajectory$(S-H-E)$ is perpendicular to the ground. The shadow trajectory denotes a trajectory of the ball projected by camera, which can be seen in images. From the geometry shown in the figure, the following triangular geometric relations can be established between the height of the ball$(\overline{GH})$ and the height of the reference player$(\overline{BC})$ as given by

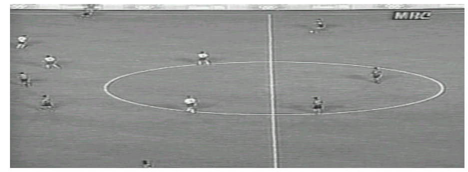
$$\frac{\overline{FG}}{\overline{FD}} = \frac{\overline{GH}}{\overline{DL}} \tag{1}$$

$$\frac{\overline{AB}}{\overline{AD}} = \frac{\overline{BC}}{\overline{DL}} \tag{2}$$

Eliminating $\overline{DL}$ from equations (1) and (2), the ball height is given by

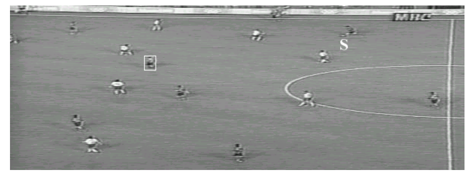
$$\overline{GH} = \frac{\overline{FG}}{\overline{FD}} \frac{\overline{AD}}{\overline{AB}} \overline{BC} \tag{3}$$

From the equation (3), it can be noticed that the ball height $(\overline{GH})$ is given in terms of the height of a player $(\overline{BC})$, which we call the reference player. If a goal post is visible in a frame, the accurate height information of the goal post can also be used instead of the inaccurate height of the reference player. In order to compute the ball height, the ground-projected camera position D should be determined together with the position of points F, A and B, which represent the projected ball position, projected head and foot position of the reference player, respectively. Once D is determined, projected ball position G on the ground can be easily determined by intersecting $\overline{SE}$ and $\overline{FD}$. Note that these positions should be given in the coordinate of ground model, which can be obtained from each frame by transforming images to ground model.

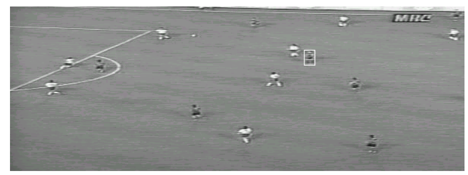


1st frame

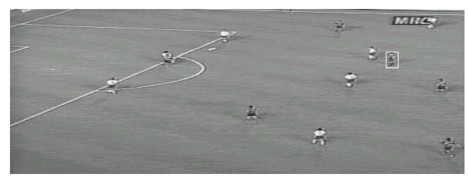
48th frame



81th frame



132th frame



151th frame

Figure 2: Exemplary frames of soccer sequence

2.2 Transformation of the image sequence

Figure 2 shows several frames of a soccer sequence. The image-to-model transformation is obtained by a simple four-point homographic planar transformation. We selected as feature points four points (three points on the center circle, one on the side line) from the ground model as indicated in Figure 3. If the center circle is visible in the image, we can easily determine the 3×3 homographic transformation matrix. An example of transformed image is shown in Figure 4. Note that the center circle in the image looks circular. In order to transform every frame to ground model whether the center circle is visible or not, mosaicking technique is applied([6][7][8]). Figure 5 shows a mosaic image from a total of 150 frames given in Figure 2. Note that the trajectory of the flying ball

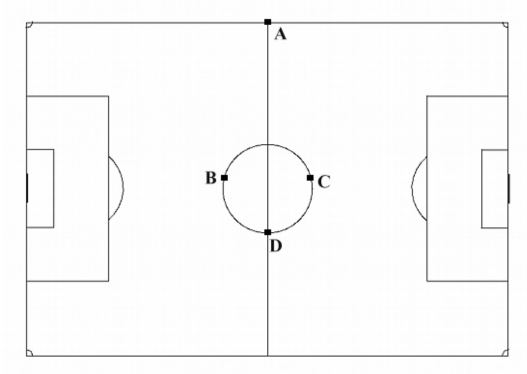


Figure 3: Ground model with four feature points indicated as black dots(B,C are tangential points from A to the center circle)

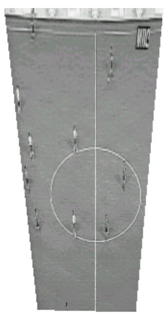


Figure 4: An example of transformed image (1st frame)

is also shown as black dots in Figure 5, which is obtained from each image in the sequence by tracking the ball. This trajectory corresponds to the shadow trajectory ($S-H-E$) in Figure 1. By transforming the mosaic image the shadow trajectory is determined on the ground model as shown in Figure 6.

2.3 Calculation of ball height

In order to calculate the ball height, the ground-projected camera position D should be determined first. Suppose two objects which are perpendicular to the ground and have nearly same depth from camera. They must look parallel in an image. If we find this kind of two parallel lines in an image and transform them to ground model, they must intersect at the position D. For example, lines $\overline{FG}$ and $\overline{AB}$ in Figure 1 look parallel in an image. We choose a player as an object perpendicular to the ground and draw a line manually that seems perpendicular to ground. Another line is obtained by drawing a parallel line passing through the ball as shown in Figure 7. Transformation of these two lines to ground model is also shown in the Figure 7. The position D is determined by the intersection of these two transformed lines. Although we have used a player as an object perpendicular to the ground, which may cause some inaccuracy, a goal post is a much better candidate for it if it is visible in an image. As mentioned earlier, the line $\overline{SE}$ denotes the ball path vertically projected on the ground. At each frame, the point G is determined by the intersection of lines $\overline{FD}$ and $\overline{SE}$,which enables calculation of the length of FG. Now that all the terms in equation 3 are prepared, the ball height can be calculated.

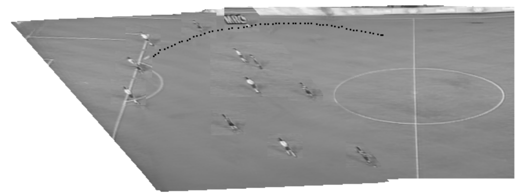


Figure 5: Mosaic image of 150 frames : Trajectory of a flying ball is shown in dots.(Beginning at the 81th frame and ending at the 143th frame)

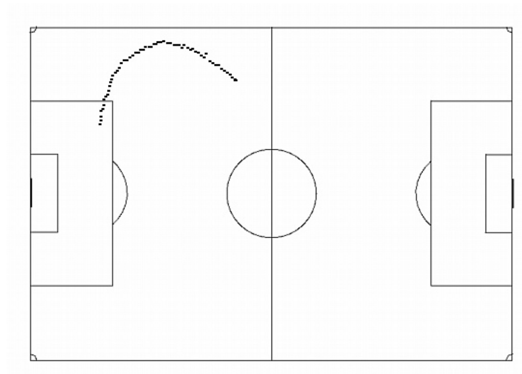
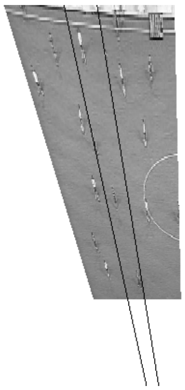


Figure 6: The shadow trajectory on the ground model

3 Physics-based ball direction search

In the above derivation, it is assumed that both the starting and the ending points are on the ground. However, it is not always the case. If a player kicks the ball before it reaches the ground, ground line would have different direction. In this section we describe a physics-based searching algorithm for finding the direction of the ball when the end position is not available on the ground. It is reasonable to assume that

(a)Two parallel vertical lines



(b)Transformed vertical lines

Figure 7: The ground-projected camera position **D** is produced by the intersection points of two transformed vertical lines in (b)

unless the ball has a lot of spin or there is a strong wind, a flying ball approximately follows the motion of a freely falling body, thus moves along a parabolic trajectory in a vertical plane that passes through a direction line. Based on this physical fact, it is possible to find out the correct trajectory that best gives a parabolic curve. In order to find out the best parabolic curve, it is necessary to back-project the shadow ball trajectory on the vertical plane in every direction possible. Several direction lines are indicated in Figure 8, which are intersecting lines of vertical planes with the ground. The ball trajectories on the several vertical planes are illustrated in Figure 9.

In Figure 9, points P', P'', P''' represent back-projected points of a point F on the vertical planes E', E'', E''' , respectively. In this way back-projecting the points of the shadow trajectory will make new ball trajectories on the resulting vertical planes. We refer to these new ball trajectories as *virtual ball trajectories* and their new direction lines as *virtual direction lines.* Once given virtual ball trajectories on the planes, the next step is to search one with the best parabolic curves, which is the one with minimum er-

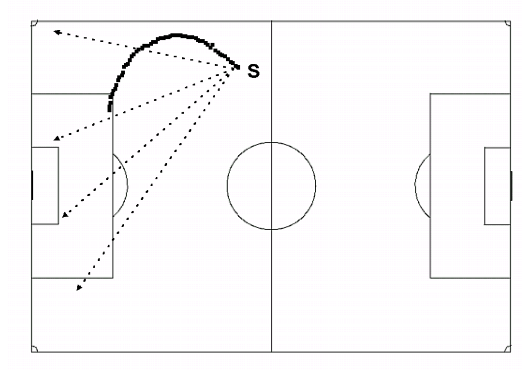


Figure 8: Several searching direction lines

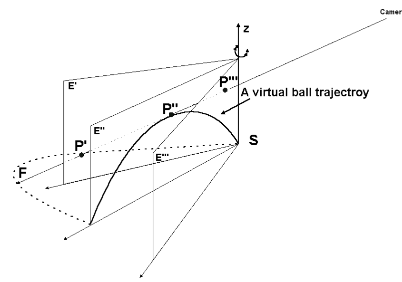


Figure 9: vertical planes and a back-projected ball trajectory on the planes

ror when the back-projected trajectories are fitted to parabolic curve. This procedure is explained in detail in the following. Let T_j be a set of back-projected points $p_i^j = (x_i^j, z_i^j)$, where j denotes the searching index, i denotes the frame number, x_i^j is the projected ball position on the j-th virtual direction line and z_i^j is the height computed using equation 3 along x_i^j. T_j is fitted to the parabolic curve $z = ax^2 + bx + c$. The parameter (a, b, c) is computed using the least mean square method,

$$\begin{bmatrix} (x_1^j)^2 & x_1^j & 1 \\ (x_2^j)^2 & x_2^j & 1 \\ \vdots & \vdots & \\ (x_N^j)^2 & x_N^j & 1 \end{bmatrix} \quad \text{or} \quad \mathbf{X^j A = Z^j} \tag{4}$$

where N is the total number of frames or number of data points. Using pseudo-inverse, **A** is given by

$$\mathbf{A} = ((\mathbf{X^j})^{\mathbf{T}}\mathbf{X^j})^{-1}(\mathbf{X^j})^{\mathbf{T}}\mathbf{Z} \tag{5}$$

The fitting error of T_j is evaluated as RMS of the

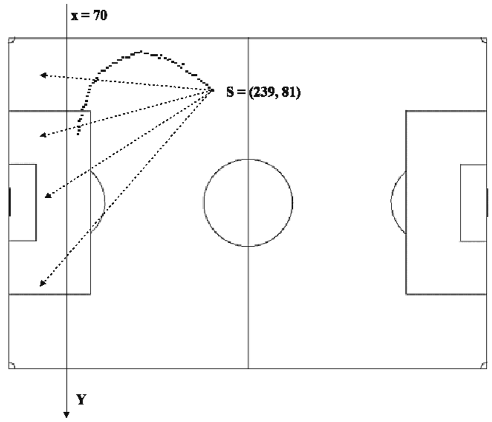


Figure 10: Search around the start position **S**

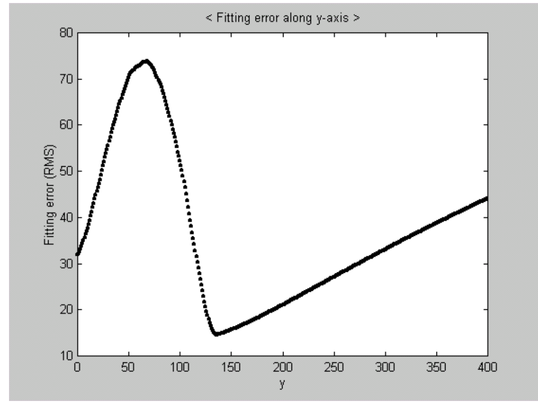


Figure 11: Fitting errors along y direction at $x = 70$

differences between fitted curve points and real data points,

$$E_j = \sqrt{\frac{1}{N}\sum_i^N(\mathbf{A^T X_i^j} - \mathbf{z_i^j})^2}, \tag{6}$$

$$\text{where} \quad \mathbf{X_i^j} = ((\mathbf{x_i^j})^2, \mathbf{x_i^j}, \mathbf{1})^\mathbf{T}$$

The direction we want to find is the one with minimum E_j. It is interesting to note that the end position on the ground can be predicted by computing the intersecting point of the parabola $z = ax^2 + bx + c$ with $z = 0$.

4 Experimental results

As an experiment we tested the proposed algorithm with a real soccer image sequence shown in Figure 2. The ball in this sequence begins flying at the 81th frame. The end position of this ball is not on the ground because a player kicks the ball before it reaches the ground. A reference player is marked with a rectangle from 81th frame in Figure 2. For the trajectory of the ball shown in Figure 5 we find a correct direction of the ball and a predicted end position using physics-based search. Several direction lines passing through the line at $x = 70$ are shown in Figure 10. Figure 11 shows the resulting fitting errors plotted along the y direction at $x = 70$. The minimum occurs at point $y = 136$. The searched direction line $\overline{SE}$ and predicted end position E are shown in Figure 12. Calculated heights of the ball are shown in Figure 13 with the reference player's height set to *1.8m*. The maximum height of the ball calculated is about *3.6m*.

5 Conclusions

In this work we propose a method for measuring 3-dimensional position of a flying soccer ball from a monocular image sequence. Given a ground model and an object perpendicular to the ground, we show that the height of the soccer ball from ground can be calculated in terms of the height of a reference player using simple triangular geometric relations. In this case both the start and the end position of a ball trajectory should be located on the ground. If this is not the case due to kicking or heading the ball before it touches the ground, a physics-based search is applied to estimate the ball's ground hitting position. Instead of a player a goal post can be used to get more accurate results if it is visible in the image sequence. In further study, will be carried out error analysis and extension to spinning ball which is not confined in a vertical plane. Note that the method we propose in this paper can be applied to the measurement of 3-dimensional position of any objects flying in a trajectory ruled by a physical law, given ground model and an object perpendicular to the ground.

References

[1] S. S. Intille and A.F. Bobick, "Tracking using a local closed-world assumption: Tracking in the football domain," *MIT Media Lab Perceptual Computing Group Tech. Report No. 296*, August 1994.

[2] T. Kawashima, K. Yoshino and Y. Aok, "Qualitative image analysis of group behavior", *In Proc. Computer Vision and Pattern Recognition*, pp. 690-693, June 1994.

[3] Ian Reid and Andrew Zisserman, "Goal-directed Video Metrology", *In Proc. European Conf. on Computer Vision*, 1996.

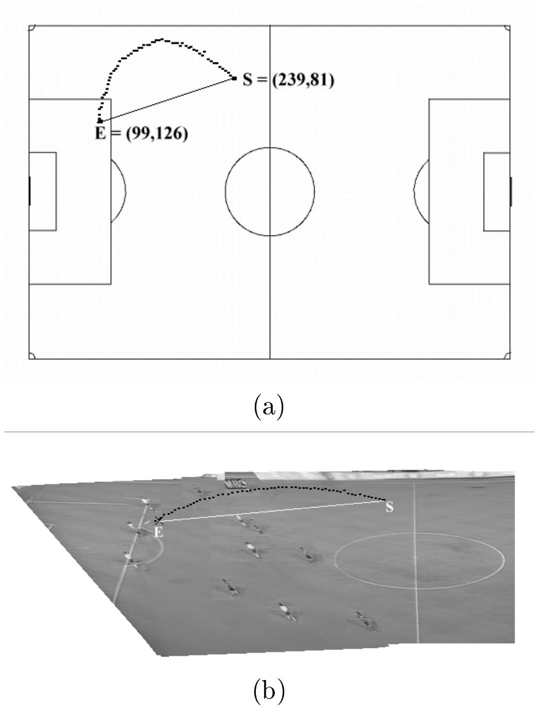


Figure 12: The computed direction line **SE** and the end position **E**. : (a) on the ground model. (b) on the mosaic image.

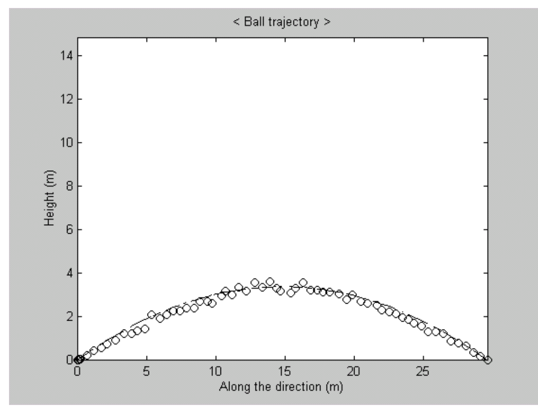


Figure 13: A computed ball trajectory with the parabolic curve fitted in '−−'

[4] E. Andre, G. Gerzog and T. Rist, "On the Simultaneous Interpretation of Real World Image Sequences and Their Natural Language Descriptions: the system SOCCER", *In Proc. European Conf. AI*, pp. 449-454, August 1988.

[5] G. Retz-Schmidt, "Recognizing Intentions in the Domain of Soccer Games", *In Proc. European Conf. AI*, pp. 455-457, August 1988.

[6] M. Irani, D. Anandan and S. Hsu, "Mosaic Based Representation of Video Sequences and Their Applications", *In Proc. of International Conference on Computer Vision*, pp. 605-611, 1995.

[7] H. Sawhney, S.Ayer and M. Gorkani, "Model-based 2D&3D Dominant Motion Estimation for Mosaicking and Video Representation", *In Proc. of International Conference on Computer Vision*, 1995.

[8] R. Szeliski and S. B. Kang, "Direct Methods for Visual Scene Reconstruction", *In Workshop on Representations of Visual Scenes*, 1995.