Soccer Ball Detection in Occluded Situations for Single Static Camera Systems

Josef Halbinger and Jürgen Metzler
Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB,
Fraunhoferstr. 1, 76131 Karlsruhe, Germany

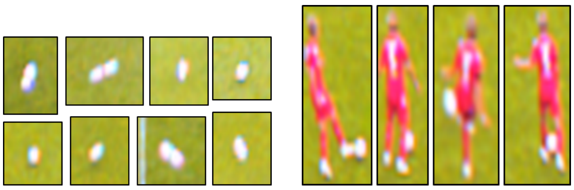
Keywords: Soccer, Sport Analysis, Ball Tracking.

Abstract: The interest on acquiring player and ball data during soccer games is increasing in several domains such as media. Consequently, tracking systems are becoming widely used for live data gathering. Due to costs, stadium infrastructure, media rights etc. there is a trend for stand-alone mobile low-cost soccer tracking systems. The drawback of such systems is that generally only low-resolution images of the players are available which strongly exacerbates the problem of detection and tracking the soccer ball. Besides difficulties that arise from the appearance of the ball by itself, the detection of the ball in situations where it is occluded by the players is very challenging. This paper presents a tracking framework for the reconstruction of the ball trajectory from monocular low-resolution soccer image sequences. The focus of this paper is the detection of the ball in occluded situations. The approach is tested and evaluated on Bundesliga data sets.

1 INTRODUCTION

The increasing professionalization of soccer is accompanied by a growing media attention as well as game analysis and professional training. Especially the automation of live analysis of soccer games is interesting for several domains such as media. However, the automation requires a robust acquisition of player and ball data that still relies heavily on the interaction of operators (so-called scouts) in current systems. Live acquisition of quantitative motion data such as distances covered by players, distances between players or ball possession can only be done by sophisticated automation. Our overall two-camera tracking system provides this kind of quantitative data for supporting a scout and for the automated acquisition of the relevant data (Herrmann et al., 2011). It automatically detects, classifies and tracks the ball, the 22 soccer players, the referee and the two linesmen in one image sequence of double Full HD resolution.

The main contribution of this work is the detection of the ball in situations in which the ball is close to a player or even partially covered by one. Detection of the ball in image sequences generally is a difficult task as the appearance of the ball varies from image to image. For instance, the high accelerations occurring at the ball may cause motion blur so that the shape of the ball is then more of an ellipse than a circle (see Fig. 1 (a)). Also, the color of the ball may be vary from image to image because of changes of the illumination conditions or it has the same color as the lines of the pitch which exacerbates the ball detection task.



(a) (b)

Figure 1: (a) Variety of the appearance of the ball extracted from one image sequence and (b) examples for partially occluded situations.

Another challenge is the image resolution of the ball which is usually very small so that also confusions with body parts may occur. Depending on the camera perspective, the ball is in front of a complex image background such as the audience which exacerbates its detection as well. Besides difficulties that arise from the appearance of the ball by itself, the detection of the ball is very challenging in situations where it is occluded by the players (see Fig. 1 (b)). Every time a player touches the ball there is a chance that the ball is not fully visible for a short time as parts of the player's body can move between ball and the cam-

Halbinger J. and Metzler J..
Soccer Ball Detection in Occluded Situations for Single Static Camera Systems.
DOI: 10.5220/0004586001110115
In *Proceedings of the International Congress on Sports Science Research and Technology Support* (icSPORTS-2013), pages 111-115
ISBN: 978-989-8565-79-2

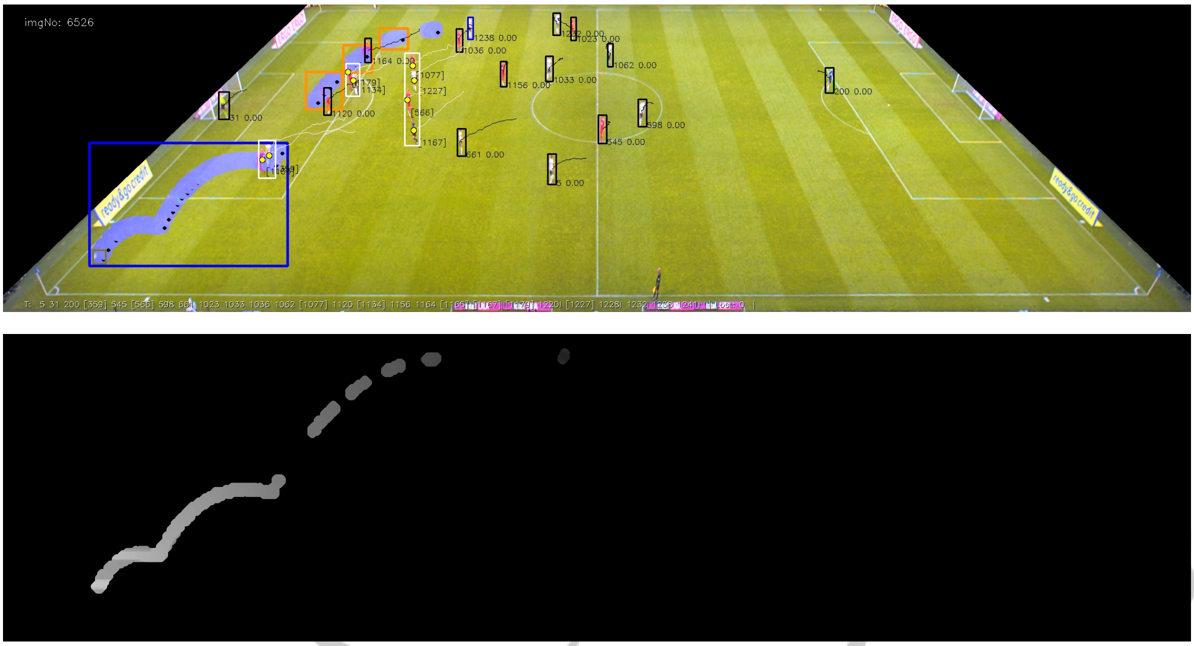


Figure 2: Sample snapshot of the processed input image sequence including detected ball tracklets marked by rectangles (top) and one for the corresponding motion history image (bottom).

era. However, as long as the ball is at least partially visible, there is an opportunity to identify the ball in the image.

The motivation of this work is to find a solution which can help to detect the ball in such cases. At the first stage, a circle detection method is applied before, at the second stage, the detected circles are evaluated by examining the Freeman chain code (Freeman, 1961) of the found contours. There are a couple publications for detecting and tracking the soccer ball as seen in (D'Orazi and Leo, 2010). However, most of these approaches focus on tracking the ball in broadcast soccer videos and require a high image resolution. The approach presented in this contribution, is applicable to static camera systems, even for low-resolution cameras. So it can be used in huge tracking systems consisting of several cameras usually fixed installed in stadiums as well as for low-cost tracking systems that generally consist of 1-3 cameras capturing the entire pitch.

The contribution is structured as follows: In Section 2, the module for the ball detection is described. It is a two-stage approach: At the first stage, the ball is detected in situations in which the ball is generally not occluded. Also, robust partial ball trajectories (tracklets) are extracted. In order to get also detection hypotheses in occluded situations (within the gaps between the tracklets), a ball detector specialized for occluded situations is used at the second stage. Then, in Section 3, results of an evaluation on data sets of a Bundesliga match are presented.

2 BALL DETECTION

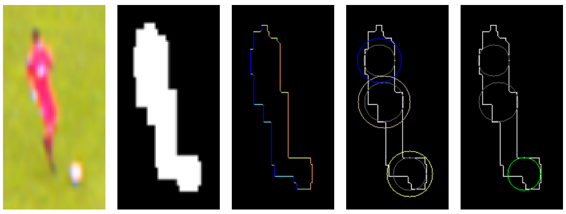
The reconstruction of the ball trajectory requires a reliable soccer ball detection. However, this is challenging as there are usually a lot of occlusions. Furthermore, a high detection rate should be achieved at a low false alarm rate.In order to be able to fulfill this requirement, we follow a detection approach that has been widely established (see e.g. (D'Orazi and Leo, 2010)): we divide the detection task in two steps. First, ball candidates are extracted followed by an add-on verification of them.

2.1 Not Occluded Situations

At the first stage, the soccer ball has to be detected. Due to the real-time constraint for live applications, a feature-based detection with e.g. a sliding windows approach cannot be used. Instead, as we capture the images from a static camera, we first apply the foreground/background segmentation from Kim et al. (Kim et al., 2004). Temporal static background like the pitch and marking lines are segmented as background, whereas moving objects generate changing appearance and are therefore segmented as foreground. During the ball candidate extraction, all foreground regions are extracted and checked for their size using calibration information of the cameras. Foreground regions that are no candidates for the soccer ball due to their size are removed. Out of the remaining regions, the external contours are extracted as a sequence of points and analyzed afterward. If the number of the contour pixels is higher than a bias, an ellipse is fitted to it and the mean squared error between every sequence point and the ellipse is calculated. Ball candidates with a high mean squared error are removed. The remaining candidates are kept as verified foreground regions in the foreground/background segmented image. Then, a dilatation is applied and the last n binary images are accumulated to a so-called Motion History Image (MHI) of verified ball candidates. Finally, ball tracklets (robust partial ball trajectories) are finally extracted from the MHI. Fig. 2 shows a MHI and some results of detected/extracted ball tracklets.

2.2 Partially Occluded Situations

The foreground/background segmentation often merges ball and player into a single silhouette if they are either close to each other or partially occlude each other. Thus, the ball is not a singular object and only appears as a bump poking out of the player's silhouette in the resulting image (Fig. 3 (a) and (b)).



(a) (b) (c) (d) (e)

Figure 3: (a) Input image, (b) foreground/background segmentation of the input image, (c) chain code representation of the outer contour: *left*- and *right*-values (yellow and light blue/horizontal lines) of the chain code are of particular interest, (d) detected Hough circles (gray/smaller circles) and circular RoI RoI_{bc} in which the CCH is calculated (blue/upper, brown/central and yellow/lower circle), (e) detected Hough circles and identified ball (green/lower cirlce).

At the second stage, the goal is to identify these bumps. In order to achieve this, we apply a two-step approach again. In the first step, circles (or at least parts of circles) in the image are detected via Hough transform (Kimme et al., 1975). In the second step, the Freeman chain code (Freeman, 1961) is considered to decide if a detected circle is a soccer ball.

In the following, the details of the procedure are given: At the beginning of the first step, all the player silhouettes of the foreground/background segmented image are extracted into separate images. On each of these silhouette images, a Hough transform for circle detection is applied. All detected circles and circular arcs that approximately match the predefined ball dimensions are determined as ball candidates. A resulting Hough circle c is characterized by the center coordinates x and y as well as the radius $r : c = (x, y, r)$.

The Hough transform variant chosen in this work is called the Hough gradient method (Bradski and Kaehler, 2008). Unlike comparable methods, this variant only uses a two-dimensional accumulator instead of a three dimensional one. This is achieved by incrementing only accumulator cells along the gradient direction of each non zero pixel of the edge map instead of incrementing a complete circle and therefore keeping a separate accumulator for every predefined possible circle radius. This is beneficial to the running time of the algorithm. The downside is a lower recognition rate of circles with a concentric counterpart. But this flaw is acceptable since concentric circles do not occur in the segmented image material.

At the beginning of the second step, the outer contour of the silhouette image is calculated. Then the Freeman chain code of the contour is determined (Fig. 3 (c)). Now, in a circular Region of Interest (RoI) around the ball candidates that were identified before, the Chain Code Histogram (CCH) is computed (Iivarinen and Visa, 1996).

The circular RoI is constructed around the center coordinates x and y of the ball candidate, adding a small Δ to the radius r (Fig. 3 (d)). The Δ is added to encounter the problem that the detected circles of the Hough gradient method tend to be slightly smaller than they actually are. As a result, the considered RoI around the ball candidate RoI_{bc} is defined as $RoI_{bc} = (x, y, r + \Delta)$.

As described in (Iivarinen and Visa, 1996), the CCH is a discrete function

$$p(k) = \frac{n_k}{n}, k = 0, 1, ..., K - 1 \,, \qquad (1)$$

where n_k is the number of chain code values k in a chain code, and n is the number of links in a chain code. In case of the Freeman chain code there are $K = 8$ possible directions.

Generally, a bump has a high amount of *left* and *right*-values of the chain code at the same time, while the *left*-values are on the upper side of the bump and *right*-values on the lower side of the bump. As a consequence, a RoI_{bc} with a CCH that provides certain frequencies of occurrence of *left*- and *right*-values λ and ρ is defined to indicate a bump in the silhouette. If this frequency lies beyond a certain threshold τ (and RoI_{bc} originates from inside the silhouette), it is assumed that a bump exists in this area. As this bump also matches the dimensions of the soccer ball, the examined RoI_{bc} is identified to have the soccer ball in it (Fig. 3 (e)):

$$Ball = \begin{cases} 1 : & \lambda \geq \tau \quad \text{and} \quad \rho \geq \tau \\ 0 : & else \end{cases} . \quad (2)$$

3 EXPERIMENTAL RESULTS

We tested the first-stage of our approach - the tracklet extraction - on a data set of a Bundesliga match consisting of an image sequence with about 140.000 images of double Full HD resolution (see Fig. 2 for an example). There are 1428 tracklets to detect, in situations where the ball is neither occluded nor merged with a player. In these situations the approach detected 1343 tracklets and missed 85. There was no false alarm, i.e. all detected ball tracklets were correctly detected as such.

The second-stage - ball detection in occluded situations - was tested on two data sets of a Bundesliga match. Both sets consist of image regions that were extracted from the same Bundesliga sequence as in the first stage test. The first data set consists of 1408 non-consecutive images with a resolution of 50×100 pixels, 704 of them showing the ball close to a player or partially occluded by a player. The other 704 images don't show a ball. The second data set consists of 9634 consecutive images with a resolution of 64×128 pixels, 111 of them showing the ball close to a player or partially occluded by a player. 9323 images don't show the ball.

As mentioned in Section 2.2, the threshold τ describes the required frequencies of occurrence λ and ρ of $left$ and *right* chain code values inside of RoI_{bc}. In order to determine the optimal threshold, τ is iterated from a specified minimum to a specified maximum in both data sets. The range is set in a way that all possible cases are covered: it starts with a configuration that identifies every Hough circle as a ball and ends with a configuration that detects no single Hough circle as a ball. The results are displayed in a ROC (Receiver Operating Characteristics) curve that puts the true positive rate of a data set in relationship with its false positive rate (see Fig. 4).

In the second data set, a true positive respectively false positive rate of 1.0 could not be reached. The reason for this is that no Hough circles matching the predefined ball dimensions were found. This leads to scenarios where varying the τ-threshold has no effect.

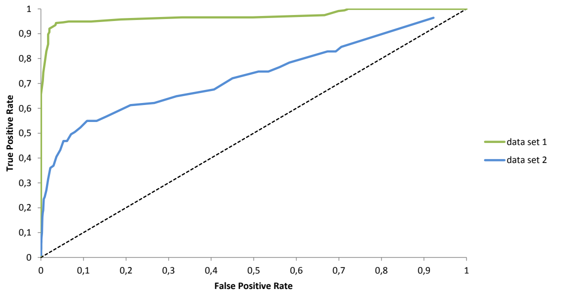
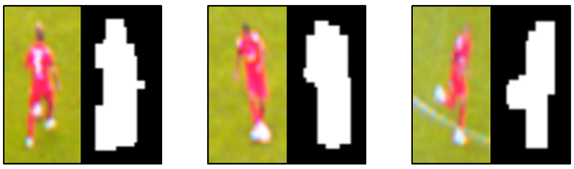


Figure 4: ROC curve of the tested second-stage approach: the ball detection approach for occluded situations. "data set 1" consists of 1408 non-consecutive images: 704 of them showing the ball close to a player or partially occluded by a player and the other 704 images don't show a ball. "data set 2" consists of 9634 consecutive images: 111 of them showing the ball close to a player or partially occluded by a player and 9323 images don't show the ball.

The results also differ because the second data set has more difficult cases: On the one hand, there are several images in which the ball is between the player's legs as illustrated in Fig. 5 (a) or right in front of the foot as shown in Fig. 5 (b). As a result, the ball does not appear as a bump in the segmented image. On the other hand, there are segmented images that have a strong bump, although there is no ball on the input image as shown in Fig. 5 (c).



(a) (b) (c)

Figure 5: Difficult cases for the ball detector: (a) Ball between player's legs, (b) ball right in front of a player and (c) player without a ball, although there is a bump in the segmented image.

4 CONCLUSIONS

In this paper, a two-stage approach for the detection of the soccer ball with the focus on occluded situations where the ball is partially occluded or merged with a player has been presented. We could yield a reliable extraction of ball tracklets in not occluded situations. Also, the ball detector for occluded situations is able to reliably detect balls in cases where the ball is partially occluded. With the exception of the delay in the output of the ball coordinates, which depends on the length of the motion history, the proposed approach is real-time capable.

REFERENCES

Bradski, G. and Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV Library*. O'Reilly Media, Sebastopol.

D'Orazi, T. and Leo, M. (2010). A review of vision-based systems for soccer video analysis. *Pattern Recognition*, 43(8):2911–2926.

Freeman, H. (1961). On the encoding of arbitrary geometric configurations. *IRE Transactions on Electronic Computers EC*, 10(2):260–268.

Herrmann, C., Manger, D., and Metzler, J. (2011). Feature-based localization refinement of players in soccer using plausibility maps. In *Proc. of the IPCV (WORLDCOMP)*, volume 2.

Iivarinen, J. and Visa, A. (1996). Shape recognition of irregular objects. In *Proc. SPIE 2904*, pages 25–32.

Kim, K., Chalidabhongse, T., Harwood, D., and Davis, L. (2004). Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11(3):172–185.

Kimme, C., Ballard, D., and Sklansky, J. (1975). Finding circles by an array of accumulators. *Communications of the ACM*, 18(2):120–122.