Tracking Soccer Ball in TV Broadcast Video

Conference Paper *in* Lecture Notes in Computer Science · September 2005

DOI: 10.1007/11553595_81 · Source: DBLP

Kyuhyoung Choi
Maxst Ltd.

Yongduek Seo
Sogang University

Tracking Soccer Ball in TV Broadcast Video

Kyuhyoung Choi and Yongduek Seo

Dept. of Media Technology, Sogang University,
Seoul, Korea 121-742
{kyu, yndk}@sogang.ac.kr
http://vglab.sogang.ac.kr/yndk_htdocs/htdocs/research_tracking.html

Abstract. This paper focuses on soccer ball tracking which is known to be more difficult than that of players due to its small size in an image and abrupt changes in its motion. Suggested is an effective soccer ball tracking algorithm which estimates ball position by exploiting the background image and player tracking results. In other words, the trajectory of ball is derived as image blobs by eliminating player blobs and the background parts from an image sequence. This algorithm performed well on a pretty long TV broadcast sequence in which the ball is frequently occluded by players.

1 Introduction

Soccer has been titled the most popular sport worldwide and even today World-cup is considered as a global festival with far-reaching effects. Analysis of soccer video sequences has been an interesting application in computer vision and image analysis as more and more related papers are published recently.

Tracking players and ball must be a necessary step before an higher level analysis. There have been some researches on tracking players [1–9]. Among them, the papers such as [2, 3, 9] have dealt with the ball tracking as well.

However ball tracking has not been thoroughly studied yet and that is the focus of this paper. Even though ball tracking belongs to single object tracking while player tracking falls within multi-object tracking, ball tracking is not easier than players tracking due to following aspects. Usually ball blobs in images are very small, which makes it difficult to derive features from and to be characterized. Sudden changes in its motion is another factor to make it challenging. In addition, occlusion and overlapping with players causes a severe problem in tracking the ball continuously; The ball becomes invisible and appears at places where a continuous prediction could not reach. In [10], it is evaluated whether a candidate trajectory, which is generated from the candidate feature image by a candidate verification procedure based on Kalman filter, is a ball trajectory instead of whether a sole object is a ball. In [11], an indirect ball detection strategy based on non-ball elimination is applied and CONDENSATION algorithm, a simple version of particle filters, is used to track ball.

Our approach is based on the work of [12] where an image of ball trajectory blob is derived to be used as the proposal density in particle filtering frame work. The soccer sequences were taken from a fixed camera while ours a moving camera.

The ball tracking as well as the players tracking in this paper is done by using particle filters, or equivalently, by SMC (Sequential Monte Carlo) methods [13–17]. In tracking multiple blobs of the players, we utilized the method proposed in [5] to address the problem of particle migration during occlusion between the same team players by probabilistic weighting of the likelihood of a particle according to the distance to its neighbors. This paper then concentrates on tracking the ball in a soccer video sequence. We utilize the result of players tracking in order to obtain measurement images that do not have players' blobs.

As mentioned above, two major problems we consider in this paper are 1) the image portion of the ball in a frame is as small as 3×3 in pixels and the color is almost white but blurred due to its motion, and 2) the interaction with players causes overlapping or occlusion and makes it almost impossible to detect and predict the ball area in the sequence by a simple usage of a particle filter.

To solve the first problem, we remove the image blobs of the players using the result of the players' tracking, segment out the ground field using a lower threshold, and finally accumulate the image blobs through the sequence. After an image filtering, this procedure results in a ball blobs connected continuously. Based on this accumulation image, particles are randomly generated only from those areas that have some blobs, which could be a noise blob, too, due to incomplete segmentation. Then, the particle filter evaluates each of the random particles to produce a tracking result.

However, when occlusion or overlapping happens the accumulation does not provide meaningful ball blobs any more. In this case, our tracker changes the ball tracking mode to *invisible* from *visible*, finds and marks players near the location where the ball have disappeared, and chases the players instead of trying to estimate the ball location. This mode transition is done on the basis of the number of meaningful pixels in the accumulation image. For each player who is suspected (marked) to have the ball, searching for the ball is done in a predetermined area with the player position as the center. When a player comes close enough to the marked, it also becomes enlisted. After a detection of the reappearance of the ball by counting the meaningful pixels, the proposed algorithm resumes ball tracking.

Temporary occlusion by a player causes the ball to appear to be stopped and kicked by him even though he never touches it. Excluding those spurious cuts from the ball trajectory completes the event detection by identifying real kickers and receivers.
Sequential Monte-Carlo method is explained in Section 2. Section 3 deals with pre-image processing and player tracking. The method of ball tracking is discussed in 4. Section 5 provides experimental results and finally Section 6 concludes this paper.

2 Sequential Monte-Carlo Algorithm

Particle filtering or sequential Monte-Carlo (SMC) algorithm estimates the posterior distribution $p(x_t|z_t)$ sequentially, where x_t is the state and z_t is the measurement at time t, given a sequential dynamic equation with Gauss-Markov process.

The posterior is represented by random particles or samples from the posterior distribution. When it is not possible to sample directly from the posterior distribution, a proposal distribution q of known random sampler can be adopted to compute the posterior, and in this case the posterior at time t is represented by the set of pairs of particle s and its weight w updated sequentially:

$$w_t = w_{t-1} \frac{p(x_t|z_t)p(x_t|x_{t-1})}{q(x_t|x_{0:t-1}, z_{1:t})} \tag{1}$$

After computation of w_t's for the particles generated from q and normalization $\sum_1^N w_t^i = 1$, where N is the number of particles, the set of particles comes to represent the posterior distribution. Particles have the same weight $1/N$ after re-sampling based on the weights or the posterior distribution.

Taking the proposal distribution as $q = p(x_t|z_{t-1})$ results in $w_t = w_{t-1}p(z_t|x_t)$, saying that the posterior can be estimated by evaluating the likelihoods at each time using the particles generated from the prediction process of system dynamics. Incorporated with resampling, the weight update equation can be further reduced to $w_t = p(x_t|z_t)$, where weight normalization is implied afterwards. This is the method of *condensation* algorithm [13, 14].

To solve the problem at hand by the condensation algorithm, one needs design appropriately the likelihood model $p(z|x)$ and state dynamic model $p(x|x_{t-1})$. In this paper, the random proposal particles are not generated from $p(x|x_{t-1})$ in the ball tracking, but from a novel proposal distribution taking account of the accumulated measurements. Therefore, we use Equation 1 for updating the weights for the posterior density.

3 Pre-image Processing & Player Tracking

The field part of original soccer image, I_k^{ogn} at frame k is subtracted to yield field-free image I_k^{sub} using histogram as in Figure 1. In I_k^{sub}, the pixels of field parts are marked as black. Via CCL (connected component labeling) I_k^{ccl} is obtained. Size filtering deletes colored blobs that have either bigger or smaller enough size not to be considered as those of people.

Player tracking is done in the way of [12]. For the image I_k^{sub} of k th frame, state estimates of players are done by the particle filter assigned respectively.

4 Ball Tracking

The basic idea in ball tracking is that the image consists of the players, ball and static background. So we may get I_t^{ball}, the image of ball only at the frame

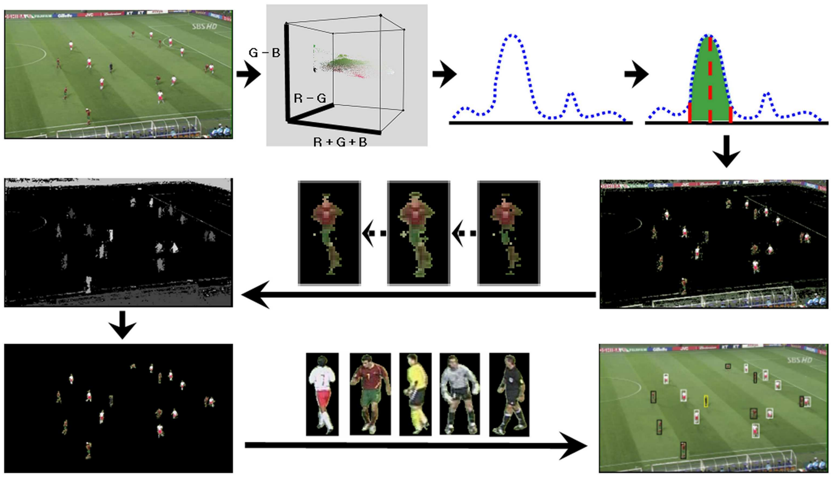


Fig. 1. Image processing

number t if we remove the portions of the background and players from the image.

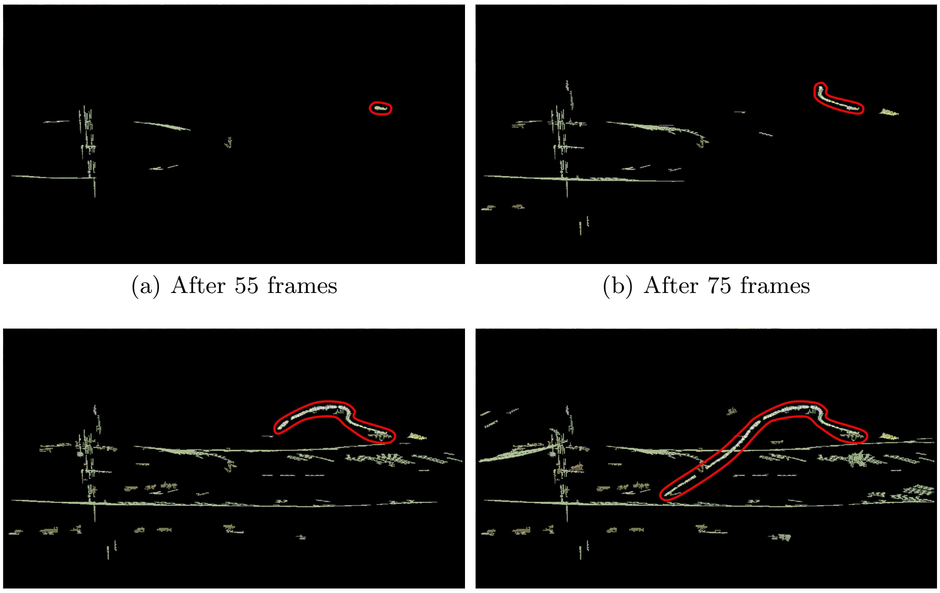
While player tracking is done at every single frame, ball tracking is batch processed at every m-th frame, where the interval of ball tracking is to produce a long enough accumulated area of the ball blobs. Examples of the accumulation are shown in Figure 2 and in our experiments the ball tracking interval m was 20 frames. If the blobs of players are deleted completely from the background-free image, we can get an accumulation image of I^{ball}s that is supposed to contain white pixels only from the ball area. However, notice that it contains noise pixels, too, due to incomplete background removal and players' blob detection. One could see that the ball has been in *visible* mode through the sequence since there are white accumulated areas (the linear structure in the accumulation image). The discontinuity means that the ball has been *invisible* during a period due to some reasons such as occlusion and overlapping. During the visible mode, we use a first order dynamic model for the ball motion perturbed by Gaussian random noise η:

$$\mathbf{x_t} = \mathbf{2x_{t-1}} - \mathbf{x_{t-2}} + \eta, \tag{2}$$

where $\mathbf{x} = (\mathbf{x}, \mathbf{y})$ is the location of the ball. The shape of the ball is modelled simply to be 3×3 rectangular. We measure the color values on the pixels in the 3×3 rectangle whose center is given by $\mathbf{x}$ - the state of the ball motion. Hence, our observation model for a ball particle is defined to be:

$$p(\mathbf{z_t}|\mathbf{x_t}) = \prod_{\mathbf{i}} \prod_{\mathbf{c}} \exp\left(-\frac{(\mathbf{c_i} - \mu_{\mathbf{c}})^{\mathbf{2}}}{\sigma_{\mathbf{c}}^{\mathbf{2}}}\right), \tag{3}$$

where i denotes a pixel location i in the 3×3 rectangle, c_i the value in RGB color space at the pixel location, and μ_c and σ_c the mean and standard deviation

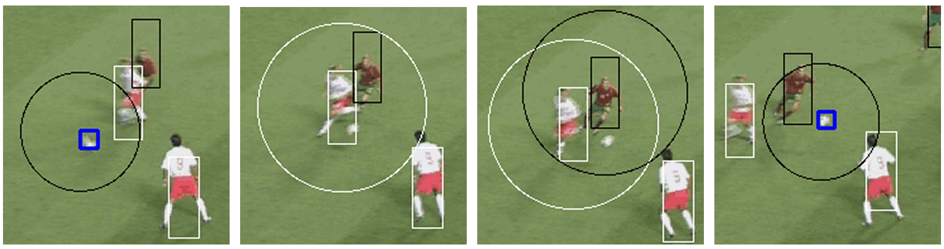


(a) After 55 frames (b) After 75 frames

(c) After 115 frames (d) After 135 frames

Fig. 2. Accumulation images for the ball blobs.

calculated based on the pixel values around the ball area in a few video frames. Particles for the tracking is generated in the image region detected as the ball area after removing the players' blob and the background. Those pixels are designed to have equal probability and hence a uniform random sampler is utilized. The likelihood is evaluated using Equation 1, and the ball location is given by the weighted average of the particles. When the ball is in the mode of *invisible*,



(a) frame 353 (b) frame 357 (c) frame 359 (d) frame 362

Fig. 3. Sub-images of some frames of interest

we stop tracking the ball. In this case, the ball is assumed to be possessed by players near the place where the ball has become invisible. As shown in Figure 3, for each of the players who are suspected to have the ball, ball searching is done in the circled area with the player position as the center. Any player who comes close enough to the suspects also becomes enlisted. After the ball reappears and is detected through the accumulation, that is, one end of another ball blob trajectory (e.g. Figure 2) is found, the proposed algorithm resumes normal ball tracking as in the early part of this section. In order to determine the ball tracking mode, we observe the number of pixels of the ball area in the accumulation image. At the frame number t ($t \neq 0$ and $(k-1)m \leq t < km$ for a natural number k), this value is given as the sum:

$$S_t = \sum_{j \in \{t-1,t,t+1\}} \sum_{l \in W_t} C_j(\mathbf{x_l}), \tag{4}$$

where $\mathbf{x_l}$ denotes an l-th pixel location in the search window W_t whose center is given by the estimated ball position at the frame number t, and C_j is an indication function:

$$C_j(\mathbf{x}) = \begin{cases} 0 & \text{if the color at } I_j^{ball}(\mathbf{x}) \text{ is black} \\ 1 & \text{otherwise} \end{cases} \tag{5}$$

Note that we incorporate the three consecutive image measurements in Equation 4 for a robust computation. Mode change is done simply by thresholding. When S_t is smaller than a threshold Th then the tracking mode changes to *invisible*, and as we explained before, the players are kept traced until our tracker finds the re-appearance of the ball pixels, that is, $S_t \geq Th$. At the most frames of *ivisible* mode S_t is zero and over 50 for the *visible*. When $S_t < 15$ in the real experiment, the mode changed to *invisible* and nearby players were traced to find the initiation of the ball blobs.

5 Experiments

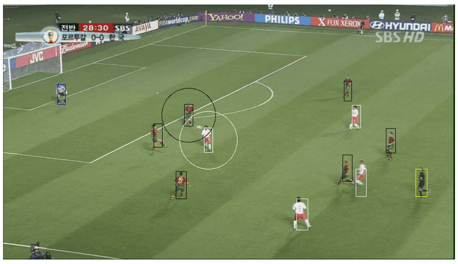
Experiments were carried out on a video sequence of 600 images whose size is 960×540 pixels. Figure 4 shows some frames of the results of which the detail is contained in accompanying video clip. The rectangle around each player is colored to show his class: ordinary players of each team, goal keeper of each team, and referee. A black circle around the ball means that the ball is not occupied by any player and thus the tracking mode is *visible*, and a colored circle shows the search area whose center is given by the location of the player, who is marked as a candidate having the ball. Notice that the color of the circle and the rectangle of the player are the same. The interval m was 20 and the threshold, Th for the mode transition was set to 15.

6 Conclusion

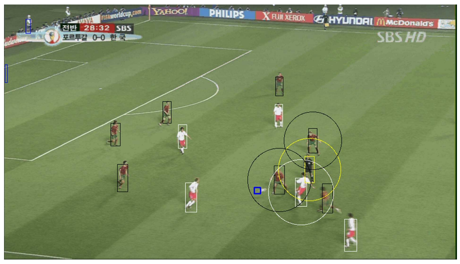
The algorithm presented in this paper have focused on an effective way of tracking the ball in a soccer match video broadcast on TV. The result of multiple

(a) Frame 55

(b) Frame 145

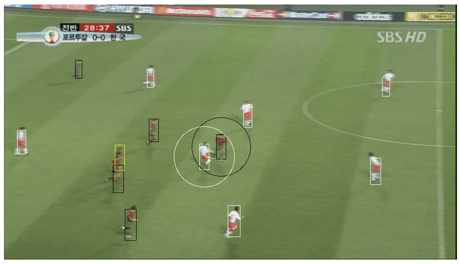


(c) Frame 333



(d) Frame 392

(e) Frame 465



(f) Frame 541

Fig. 4. Examples of result images

player tracking was made use of in order to obtain a robust measurement for the ball tracking. By removing the blobs of players, we could obtain an accumulation image of the ball blobs. This accumulation image provided us not only a proposal density for the particle filtering but also a clue to deciding whether the ball was visible or invisible in the video frames. Basically, the ball tracking was done by particle filtering. However, the performance was highly improved by two ingredients: first, taking the accumulation image as the proposal density, and second, mode change by counting the meaningful ball pixels. When the ball was invisible, we pursued every nearby players until the ball pixel came out again. Since the ball pixels were accumulated in time, the tracking algorithm showed in the real experiment a very robust ball tracking results, that was not shown by other studies. By excluding cuts in ball trajectory blob due to temporary occlusion, pairs of kicker and receiver are decided to extract events in the sequence.

Acknowledgment This research was performed for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy of Korea.

References

1. Intille, S., Bobick, A.: Closed-world tracking. In: Proc. Int. Conf. on Computer Vision. (1995)
2. Yow, D., Yeo, B., Yeung, M., Liu, B.: Analysis and presentation of soccer highlights from digital video. In: Proc. Asian Conf. on Computer Vision. (1995)
3. Seo, Y., Choi, S., Kim, H., Hong, K.: 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, Florence, Italy. (1997)
4. Iwase, S., Saito, H.: Tracking soccer player using multiple views. In: IAPR Workshop on Machine Vision Applications. (2002)
5. Ok, H., Seo, Y., Hong, K.: Multiple soccer players tracking by condensation with occlusion alarm probability. In: Int. Workshop on Statistically Motivated Vision Processing, in conjunction with ECCV 2002, Copenhagen, Denmark. (2002)
6. Yoon, H., Bae, Y., Yang, Y.: A soccer image mosaicking and analysis method using line and advertisement board detection. ETRI Journal **24** (2002)
7. Utsumi, O., Miura, K., IDE, I., Sakai, S., Tanaka, H.: An object detection method for describing soccer games from video. In: IEEE International Conference on Multimedia and Expo (ICME). (2002)
8. Kang, J., Cohen, I., Medioni, G.: Soccer player tracking across uncalibrated camera streams. In: Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS). (2003)
9. Yamada, A., Shirai, Y., Miura, J.: Tracking players and a ball in video image sequence and estimating camera parameters for 3d interpretation of soccer games. In: Proc. International Conference on Pattern Recognition. (2002)
10. Yu, X., Xu, C., Leong, H., Tian, Q., Tang, Q., Wan, K.: Trajectory-based ball detection and tracking with applications to semantic analysis of broadcast soccer video. In: ACM MM03, Berkeley. (2003) 11–20
11. Tong, X.F., Lu, H.Q., Liu, Q.S.: An effective and fast soccer ball detection and tracking method. In: ICPR (4). (2004) 795–798
12. Choi, K., Seo, Y.: Probabilistic tracking of the soccer ball. In: Int. Workshop on Statistical Methods in Video Processing, in conjunction with ECCV 2004, Prague, Czech Republic. (2004)
13. Kitagawa, G.: Monte-carlo filter and smoother for non-gaussian nonlinear state space model. Journal of Computational and Graphical Statistics (1996)
14. Blake, A., Isard, M.: Active Contours. Springer-Verlag (1997)
15. Liu, J., Chen, R.: Sequential Monte Carlo methods for dynamic systems. (1998)
16. Doucet, A., Godsill, S., Andrieu, C.: On sequential monte-carlo sampling methods for bayesian filtering. (2000)
17. Doucet, A., Freitas, N.D., Gordon, N., eds.: Sequential Monte Carlo Methods in Practice. Springer-Verlag (2001)