Robust Tracking of Multiple Football Players Considering Occlusion and Crowded Situations

Hirokatsu Kataoka, Yoshimitsu Aoki
Graduate School of Science and Technology

Keio University
Kanagawa, Japan
E-mail: kataoka@aoki-medialab.org, aoki@elec.keio.ac.jp

ABSTRACT

In this paper, we propose a method to track multiple players in a soccer game video which captured by a single camera. The camera pans and brings the players and the ball into the view, enabling to record the whole pitch. The players' trajectories in frames and camera movement are obtained to estimate the trajectories on the pitch. Moreover, we propose an effective tracking method for solving occlusion problem by combining Particle Filter and Real AdaBoost Classifier. Experimental results show the effectiveness of the proposed method.

Keywords: Particle Filter, Real AdaBoost, Object Tracking

1. Introduction

In sports contents, there are many services now. For example, internet live news flash service systems [1] and player activity analysis in sports videos. In order to put into practice these contents, it is necessary to acquire important information from the sports videos. In sports games, the study in soccer videos has increased [2], because soccer player tracking is a challenging issue, due to factors such as occlusion, camera movements and complicated movement of the players. For the analysis of soccer videos, we must obtain positional information of players and soccer ball. At the beginning, it was done manually. But, this method costs huge labor and a great amount of money. So acquire tracks automatically is required. The method of tracking players based on template matching and background subtraction [3].But it is inevitable to be inaccurate in tracking. So, various cameras are installed on the roof of a stadium and a video billboard [2] [4]. We can obtain different angle's image, hence player's occlusion decreases and high accuracy tracking becomes possible. But the cost and the limited facilities this technique can be applied on represent a problem for this method. Besides, tracking method is developing and those techniques apply for sports video analysis. Specially, great progress tracking accuracy by Mean-Shift Tracker [5] and Particle Filter [6]. These methods based on object model and center of gravity is moved in the direction where the similarity is high. So, it is easy to distinguish between object and background. But, if there are many objects and occur overlapping, tracking is extremely difficult. These techniques are not considered the action of crowd.

To solve these problems, we propose a technique to track football players using a single camera set up at the stadium. The camera pans and brings the players into the view, enabling to record the whole pitch. Player tracks on the pitch can be acquired in high accuracy by horizontally swinging the camera.

In tracking method, we use Particle Filter. It is time when crowd some objects that has resemble features are difficult to track. So, we resample the center of gravity when the distance between players become near and attempt to occlusion problem.

2. Proposed Method

Fig.1 shows the flow of the proposed method. Firstly, we estimate pitch area by using color information. The result of pitch area estimation is utilized for tracking players on the pitch and camera motion estimation.

Secondly, target players on the pitch area are tracked by applying a particle filter technique based on color histogram information. Although this method is effective for tracking the target players, it is difficult to keep the tracking when the object have similar color features. Especially when the players come near or occluded each other, the color-based particle filter may perform poorly. As a solution against this problem, we check the positions of the target players by referring the center of gravity for each player, which is calculated by a particle filter, during tracking process. Then, we detect situations that the players move closer and might be occluded each other. In such situations, we apply a Real AdaBoost technique in order to detect the players and resample the center of gravities of a combination of the particle filter and the Real AdaBoost.

In addition to tracking local positions of the players in each image, global positions of the players on the ground are required for analyzing tactical issues. The global positions of the players can be mapped by using those of local positions and camera motion estimated by optical flows obtained from feature points around the pitch.

Finally, the trajectories of the target players are obtained on the pitch as shown in the bottom of Fig.1.

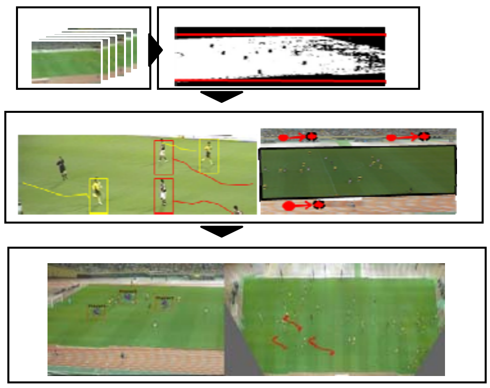

Fig.1 Proposed method

2.1 Pitch area estimation

In order to restrict pitch area and extract player information from the video sequence, we estimate pitch area. To extract pitch area and binarization, convert color space RGB to HSV. Lawn area is extracted by thresholding Hue($70<H<120$), Saturation($0.35<S<0.72$)(Fig.2). From the center of image to up/down scan horizontally. It can get upper and lower y-coodinates.

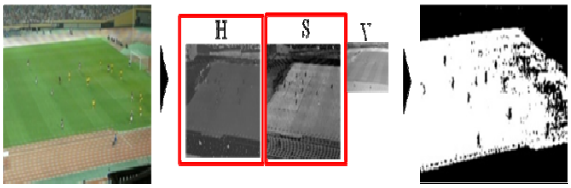

Fig.2 Lawn area extraction

2.2 Player tracking in the video

We apply Particle Filter [7] to track players and resample center of gravities by Real AdaBoost [8], when overlap player.

Particle Filter is a method of analyzing the time series based on presumption, and the particle with the movement model and likelihood used. The method of observing the movement model and the likelihood can be freely set and the state of the object can be observed and be presumed by averaging a lot of particles. Therefore, it is thought there is an advantage that it is not easy to lose sight of the pursuit object, and it is effective for the soccer player's pursuit. The procedure is shown below.

Step1.Initialization

In Initialization step, arrange particles after specify player's position. The movement model used in step2 and the color histogram used in step3 were set here. Moreover, the number of particles to tracks player was provided 70.

Step2. State presumption

The following state is presumed and the particles are moved. Linear uniform motion was applied to the movement model of particle. Moreover, add noise player position and velocity to adjust irregular movement that player's turnabout and horizontal camera motion. It shows movement model below.

$$x_{t+1} = x_t + u_t \Delta t + w_x$$
$$y_{t+1} = y_t + v_t \Delta t + w_y$$
$$u_{t+1} = u_t + w_u$$
$$v_{t+1} = v_t + w_v$$

x, y : Player position
u, v : Velocity
w : Noise
Δ : Delta

Step3. Likelihood calculation

Color histogram acquired from player image prepared first is compared with color histogram acquired from surrounding of particle (Fig.3). In order to calculate likelihood, use the Bhattacharyya coefficient.

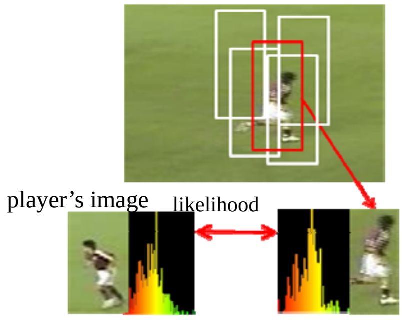

Fig.3 Likelihood calculation

$$L = \sum_{u=1}^{m} \sqrt{p_u q_u}$$
$$\sum_{u=1}^{m} p_u = \sum_{u=1}^{m} q_u = 1$$

L : Likelihood
p, q : Histogram
m : number of histogram bins

Step4. Likelihood evaluation

The center of gravity of the object is calculated here. The center of gravity is calculated below.

$$(g_x, g_y) = (\sum_{i=1}^{n} w_i x_i, \sum_{i=1}^{n} w_i y_i)$$

g : center of gravity
w : likelihood
x, y : player position

After calculate the center of gravity, it transfers to the state presumption, and presumes the following player position. The presumption area of the next state is in a rectangular 30×30pixels, pay attention to the player's local movement (Fig.4). Likelihood was adjusted to 0 on the rectangular outside.

It is set initial state with Step1, and track players repeat Step2~4.

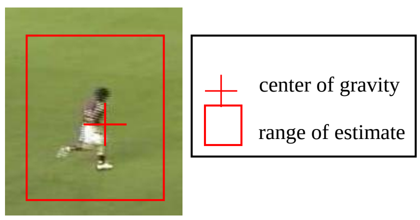

Fig.4 Range of player position

Next, we propose to detect a complex movement of player minutely by Real AdaBoost classifier. Real AdaBoost Algorithm is a learning method expansion of AdaBoost [9]. Real AdaBoost is known high accuracy detector because of generates many weak classifiers and combine them into a strong classifier. It shows as follows learning method.

[FlowOfMachineLearning]

Preparations

- DatasetNumberN
- DatasetHasFeature$(X_1, X_2, \cdots, Xn)$
- Positive / NegativeLabel $+1/-1$

1.InitializeSampleDistribution

$$W_i = \frac{1}{N} \quad i = 1,2,\cdots N$$

2.(1)CalculateWeakClassifier'sDistribution

$$W_+^j = \sum_{i:u_i=+1} D(i)$$

$$W_-^j = \sum_{i:u_i=-1} D(i)$$

(2)OutputOfWeakClassifier

$$h(x) = \frac{1}{2}\ln\left(\frac{W_+}{W_-}\right)$$

(3)EvaluateOutputOfWeakClassifier

$$Z = 2\sum_j \sqrt{W_+^j W_-^j}$$

3.SelectWeakClassiier h MinimizingZ

$$h_t = \arg\min Z$$

4.UpdateSampleDistribution

$$D_{t+1}(i) = D_t(i)\exp[-y_i h_t(x_i)]$$

Normalize D_t

TheFinalStrongClassifier H is

$$H(x) = \mathrm{sign}\left[\sum_{t=1}^{T} h_t(x)\right]$$

It is time when learning Real AdaBoost must set the feature. Here, we apply HOG (Histograms of Oriented Gradients) often used human detection [10]. HOG is acquire brightness gradient and make histogram, it is possible to detect target or not by roughly shape and contour. Players are detected in learning HOG by Real AdaBoost. To obtain HOG, divide “cell” and “block”. It is an area, cell is 5x5pixels, and block is 3x3cells. Brightness gradient and intensity are calculated from size uniformed image. Here, I(x , y) is image brightness, m(x , y) shows intensity and g(x , y) is gradient.

$$m(x,y) = \sqrt{fx(x,y)^2 + fy(x,y)^2}$$

$$g(x,y) = \tan^{-1}(fy(x,y)/fx(x,y))$$

$$fx(x,y) = I(x+1,y) - I(x-1,y)$$

$$fy(x,y) = I(x,y+1) - I(x,y-1)$$

To make histogram, calculate gradient and intensity each pixel. Here, it is not need to consider direction of gradient, we adopted 0~180° gradient as the orientation range. We get each pixel, then vote them into 9 bins in 0~180°. After calculate histogram, normalize histogram and obtain HOG feature. Normalize 3x3cell (1block), shift one cell and normalize HOG.

$$h' = \frac{h}{\sqrt{\sum_{i=0}^{k} h_i^2 + \varepsilon}} \quad (\varepsilon = 1)$$

HOG feature is h, k is number of dimension. Here, obtain HOG from 30x40pixels image. We extracted 1944 dimensions HOG, divided 6x8cell. (histogram bins 9× block cells 9 × width number of normalization 6-2 × height number of normalization 8-2) Fig.5 shows the result of player’s detection by Real AdaBoost.

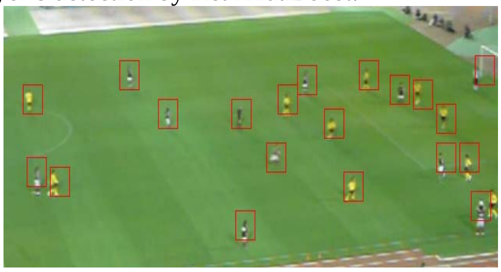

Fig.5 Player detection by HOG + Real AdaBoost

Decision of occlusion is the between player’s distances. This distance set don’t violate each particle’s effective area. It is judged occlusion area, detect soccer players by Real AdaBoost classifier. Moreover, to leave over detection, calculate likelihood in detected window. If detected number of players equal tracked number of players, resample the center of gravities (Fig.6).

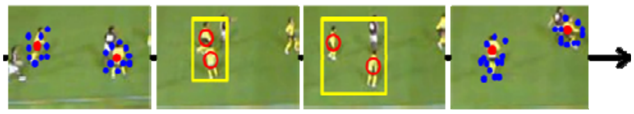

Fig.6 Resampling center of gravity in occluded area

2.3 Movement of camera

The amount of the movement of camera between frames is calculated by movement of feature points in the frame [11]. Optical flow is applied for the pursuit of feature points. Various methods are proposed to optical flow. Accuracy excellent, high-speed Lucas-Kanade method to process was applied here. The movement of x and y is calculated as a numerical value. Moreover, to acquire movement of camera in the frame accurately, the flow was acquired excluding the pitch area (Fig.7).

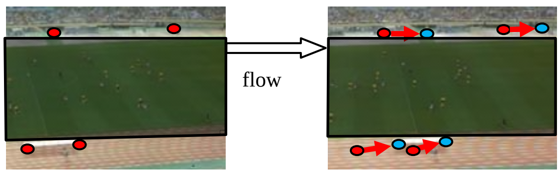

Fig.7 Camera movement between frames

2.4 Player tracks on the pitch

The player tracks on the pitch were obtained from the player tracks in the soccer game video and the movement of camera. Here, player tracks can be acquired by requesting the difference between the player tracks in the soccer game and the amount of the movement of camera to swing the camera horizontally. The player tracks on the pitch were acquired, and it draws to bird’s eye view (Fig.8).

Left figure is an input image, and it draws to the bird's eye view of right figure in tracks obtained by pursuing the player on a rectangular inside. Player tracks on the pitch can be acquired in the video that the camera pans horizontally, because movement of camera is acquired here and it reflects it in the tracks acquisition. Figure 9 shows the change in tracks by the presence of the movement of camera. The gap is caused in player's movement locus if it draws in the image as it is. So, the track on the pitch can be acquired by acquiring the movement of camera, and subtracting it from the player tracks in the video.

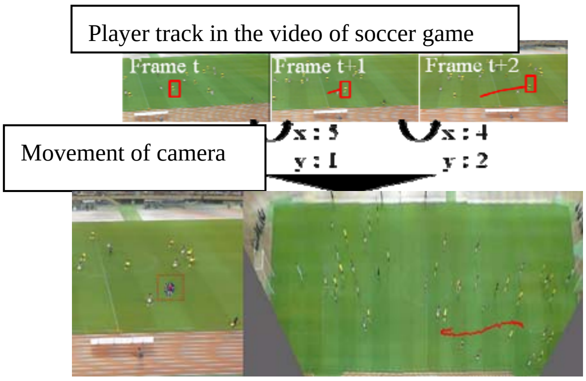

Fig.8 Player track on the pitch

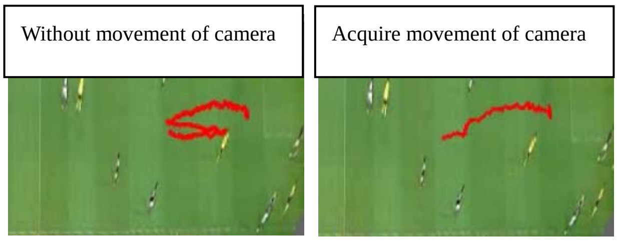

Fig.9 Presence of the movement of camera

3. Experiment

We conducted experiments for verifying tracking accuracy using the proposed method in the soccer game video. The resolution of the video is 640×480 pixels, and the frame rate is 30fps. The color histogram of the player was prepared before the experiment (Fig.10). For learning procedure, we prepared 3,500 positive images, 7,600 negative images, respectively (Fig.16). A number of weak classifires is 50. The performance of the computer was Quad Core Intel Xeon 2.66GHz CPU, and 3.25GB Memory.

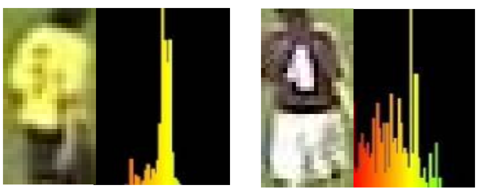

Fig.10 Player images and color histogram

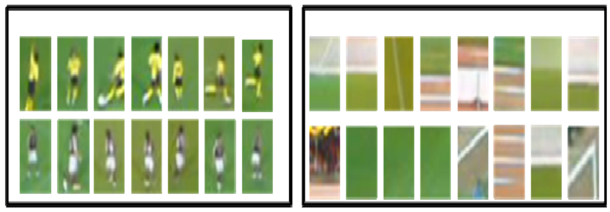

Fig.11 Learnig dataset

Table.1 shows the results of tracking rate. The performances of the proposed method and two conventional methods were compared in Table1. The Conventional methods are Particle Filter with color histogram feature, and with both of color histogram and edge features. We selected 50 scenes that include overlapping of multiple players. As the result, the successful tracking rate of the proposed method is over 80%, which scores the best performance rate among the methods. This indicates the proposed method is much robust than the other methods.

Table.1 The result of tracking experiment

	Success	Successful rate
Histogram	15 / 50	30%
Histogram +Edge	25 / 50	50%
Proposed method	41 / 50	82%

By setting the movement model that considering noise, it is able to track under the situations that camera movement and player's turnabout. In addition, our proposed method can track under complicated situations, for example, overlap of multiple players, because of calculated likelihood by color histogram and resampling the center of gravity after player detection.

As failure situations, many players gathered the same place and all the tracking methods did not work well. This is not able to acquire HOG and detect player by Real AdaBoost. To solve this problem, it is thought that change how to give occlusion area.

4. Conclusion

In this paper, we propose a method to track multiple players in a soccer game video which captured by a single camera. The camera pans and brings the players and the ball into the view, enabling to record the whole pitch. The players' trajectories in frames and camera movement are obtained to estimate the trajectories on the pitch. Moreover, we propose an effective tracking method for solving occlusion problem by combining Particle Filter and Real AdaBoost classifier. Experimental results show the effectiveness of the proposed method.

When we acquire the movement of camera, the estimation errors of motion vectors accumulates because the movement of camera is relatively acquired from the adjacent frames. Therefore, it is necessary to reduce these errors for obtaining more accurate players' trajectories on the pitch. We also consider soccer ball tracking, which is important for tactical analysis of soccer games.

References

[1] Mei Han, Wei Hua, Terrence Chen, Yihong Gong, ”Feature Design in Soccer Video Indexing”, ICICS-PCM2003, pp950-954, vol2, December 2003

[2] Naho Inamoto, Hideo Saito, “Free Viewpoint Video Synthesis and Presentation from Multiple Sporting Videos”, ICME2005, 4pp, July 2005

[3] Thomas Bebie, Hanspeter Bieri, “SoccerMan－reconstructing soccer games from video Sequences”, ICIP1998, pp898-902, Oct 1998

[4] Takayoshi Koyama, Itaru Kitahara, Yuichi Ohta, “Live Mixed-Reality 3D Video in Soccer Stadium”, ISMAR’03, pp.178-187

[5] Gael Jaffre, Alain Crouzil, “NON-RIGID OBJECT LOCALIZATION FROM COLOR MODEL USING MEAN SHIFT”, ICIP2003, pp.III-317-20 vol.2 September 2003

[6] Yu Huang, Joan Llach, “VARIABLE NUMBER OF “INFORMATIVE” PARTICLES FOR OBJECT TRACKING”,ICME2007, pp.1926-1929, July 2007

[7] Michael Isard, Andrew Blake, “CONDENSATION – Conditional Density Propagation for Visual Tracking”, International Journal of Computer Vision 29(1),5-28(1998)

[8] Wei-Chao Lin, Michael Oakes, John Tait, “Real AdaBoost for Large Vocabulary Image Classification”, CBMI 2008, pp.192-199, June 2008

[9] Paul Viola, Michael Jones, “Robust Real-time Object Detection”, Proc. of the 2nd International Workshop on Statistical and Computational Theories of Vision - Modeling,Learning, Computing, and Sampling (2001)

[10] Zhipeng Li, Yun Sun, Fuqiang Liu, Wenhuan Shi, “An Effective and Robust Pedestrians Detecting Algorithm & Symposia”, Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems Beijing, China, pp545-549, October 2008

[11] Bruce D Lucas, Takeo Kanade, “An Iterative Image Registration Technique with an Application to Stereo Vision”, Proceedings of Imaging Understanding Workshop, pp.121-130(1981)

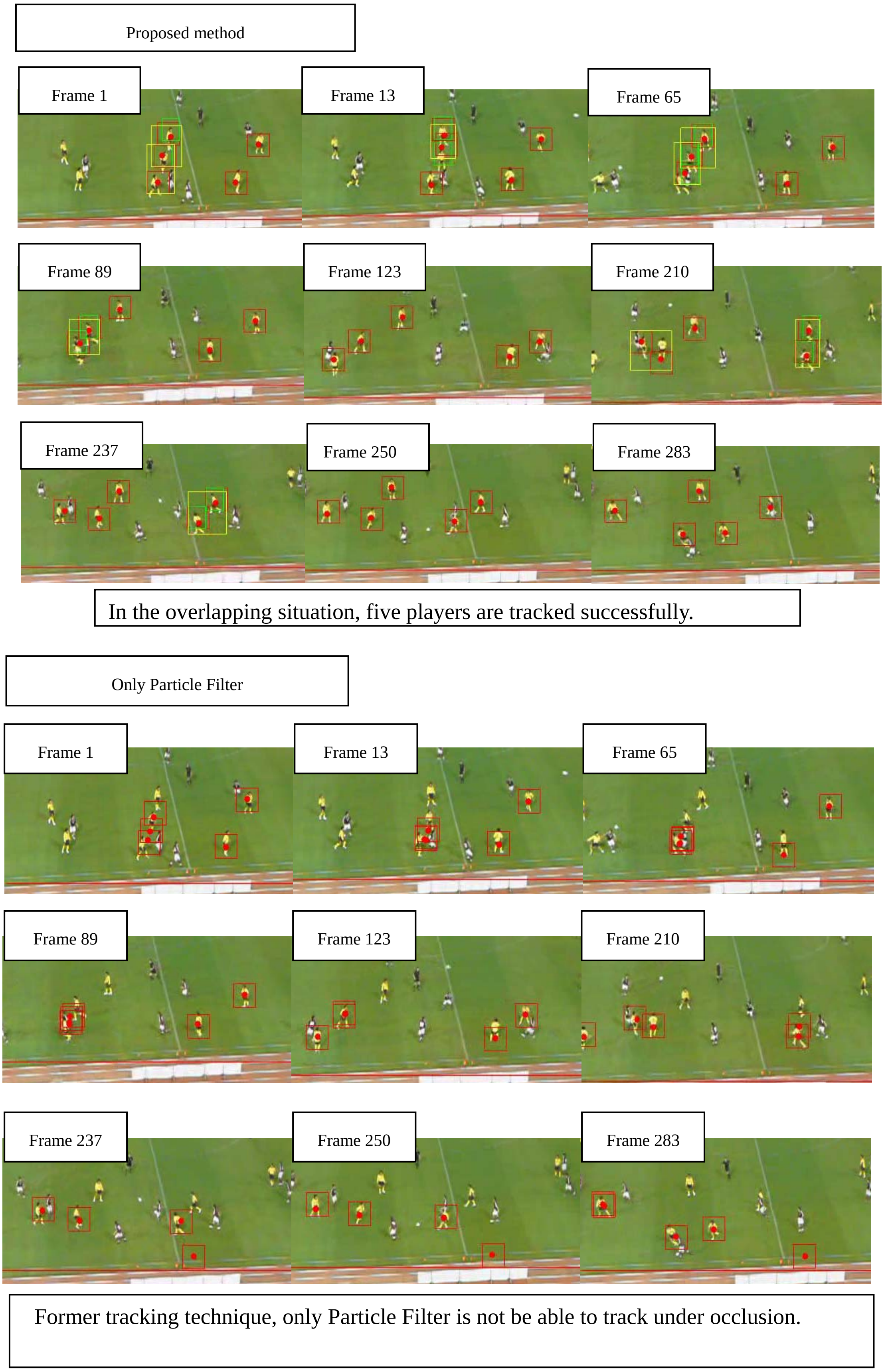

Fig.12 Compare proposed method with only Particle Filter method