Player Position Estimation by Monocular Camera for Soccer Video Analysis

Hirokatsu Kataoka[1] and Yoshimitsu Aoki[1]

[1]Keio University, Kanagawa, Japan
(Tel: +81-45-566-1796; E-mail: kataoka@aoki-medialab.org)

Abstract: In football contents, there are many services now. For example, player's tactics analysis and production of highlight scenes. In order to put into practice these contents, it is necessary to acquire important information from the football videos. In this paper, we propose a method to track multiple players and ball in a football video which is captured by monocular camera. In the case of capturing the video from a single view, there might be a lot of occluded situations of players. To overcome this problem, we propose a robust tracking method for multiple players by combining Particle Filter and Classifier. In order to track the ball, we try to track applying labeling and nearest neighbor algorithm. And more, we apply perspective transformation to extract player's position on the pitch. We show the experimental result and effectiveness of our proposed method.

Keywords: Object Tracking

1. INTRODUCTION

In football contents, there are many services now[1]-[7]. For example, player's activity analysis and production of highlight scenes. In order to put into practice these contents, it is necessary to acquire important information. In football games, the study about football videos has been increased in the field computer vision, because football player tracking is a challenging issue, due to factors such as occlusion, camera movements and complicated movement of the players. For the analysis of football videos, we must obtain positional information of players and ball. At the beginning, this works were done manually. But, this way is time-consuming. So we'd like to track players and ball automatically.

In the previous method, Template matching and background subtraction were the tracking techniques of football scenes. These methods have some problems. For instance, processing time and tracking accuracy. It is inadequate under complicated situations and occlusion.

These days, tracking methods habe been improved and these techniques were applied for football video analysis. Specially, Particle Filter tracking brought in great progress[8]. This method based on object model and move the center of gravity for the direction where the similarity is high. It is easy to distinguish object from background. But, if there are many objects and overlapping occurrence, tracking is extremely difficult. These techniques don't consider the crowded actions. So, some Particle Filter base methods have been improved for occlusion problem. At first, Particle calculates two types of feature, color and edge[9]. In this calculation, it is hard to lose tracking object. However, this way can't track object we have some similar objects in frames. It is difficult to track football players because of there are many players that wear the same uniform. So, there is a method that track in occlusion. Dearden et al apply SIR Particle Filter[1]. This method tracks Sample Importance Resampling (SIR) Particle Filter and cares occluded situations. However, we need to improve tracking method that care under occluded situations.

On the other hands, there are many studies about football video analysis. Ohta et al obtain 3D position from various videos from different viewpoints, hence player's occlusion decreases in the video and high accuracy tracking becomes possible[2][3]. Beetz et al propose Automated Sports Game Analysis Model(ASPOGAMO)[4][5][6]. They can specify "ball possessing player", "all player's positions", "scene understanding" and so on. Hamid et al proposed the algorithm that displays movement flow of player and offside line after tracking players and ball[7]. But, these techniques have a problem of few facilities to capture football videos.

In this paper, we present a technique to track the football players and ball and specify player's position on the pitch using monocular camera. Moreover, we realize football tactics analysis. The camera pans and brings the players and the ball into the view, enabling to record the whole pitch by monocular camera. Therefore, we must track players and ball considering the camera motion and the occlusion.

2. PROPOSED METHOD

Figure1 shows the flow of our proposed method. At the beginning, we estimate pitch area from color information, Hue and Saturation. We use Particle Filter to track each football player. In order to solve occlusion problem, we detect player and resample the center of gravity considering player's velocity under crowded situations. In minute ball tracking in images, we get the region information by labeling and nearest neighbor. Besides, we project player's point from original image to bird's-eye view. This method is used perspective transformation that can specify player's point on the pitch.

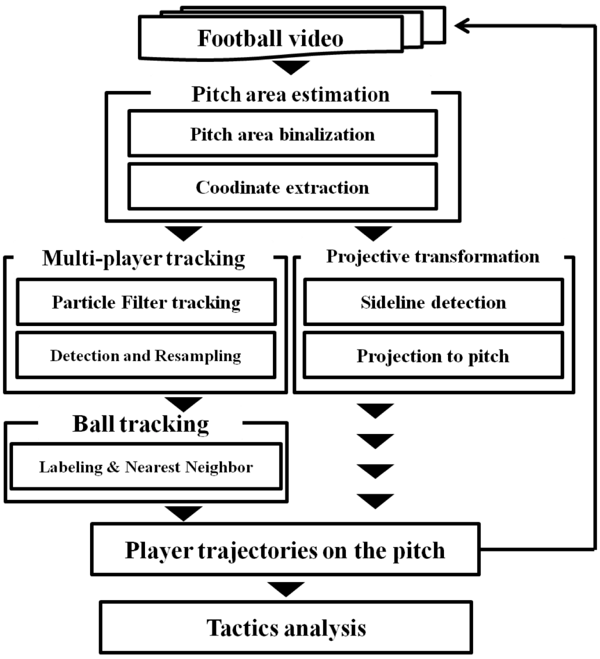

Fig. 1 Proposed method

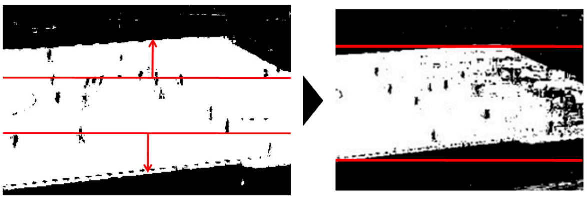

Fig. 2 Pitch area estimation

2.1 Pitch area estimation

To restrict processing region, we estimate pitch area. At first, we convert color space from RGB to HSV. Lawn area is binarized by thresholding Hue and Saturation. And then, we get upper and lower pitch area. Figure2 shows scanning binarized image and extraction upper and lower pitch area coordinates.

2.2 Player tracking

We apply Particle Filter to track each player, and resample the center of gravity by Real AdaBoost when player's overlap.

Particle Filter is an analyzing method based on prediction[8]. This method has particles that consist of observation model and likelihood observation. Particle Filter is effective for the player tracking. The processing steps is shown below.

Step1 Initialization

In this step, arranging particles after specifying player's position by classifier. We'll explain later how to compose classifier.

Step2 State prediction

Linear uniform motion was applied to the observation model of particle. Moreover, add system noise to position and velocity to adjust irregular movement that player's turnabout and horizontal camera motion :

$$x_{t+1} = x_t + u_t + w_x \quad (1)$$
$$y_{t+1} = y_t + v_t + w_y \quad (2)$$
$$u_{t+1} = u_t + w_u \quad (3)$$
$$v_{t+1} = v_t + w_v \quad (4)$$

where x, y are particle's position, u, v are x and y direction's velocity. w_x and w_y are positioning noise, w_u, w_v are velocity noise. The settings of positioning noise is -5 - +5 pixels, velocity noise is -3 - +3 pixels considering camera swing and player's complicated movement.

Step3 Likelihood calculation

Model histogram is acquired from player at first frame. This histogram is compared with color histogram from around of particle. Bhattacharyya cofficient evaluates likelihood shown below :

$$L = \sum_{u=1}^{m} \sqrt{p_u q_u} \quad (5)$$
$$\sum_{u=1}^{m} p_u = \sum_{u=1}^{m} q_u = 1 \quad (6)$$

where L is likelihood, p, q are color histogram and m is number of bins.

Step4 Likelihood evaluation

The center of gravity is calculated this step. We calculate the center of gravity shown below :

$$(g_x, g_y) = (\sum_{i=1}^{n} L_i x_i, \sum_{i=1}^{n} L_i y_i) \quad (7)$$

g is the center of gravity of player. L is each particle's likelihood and x and y are position of particle.

Particle Filter track object iteration of Step2 - 4.

Although this method is effective for tracking individual player, it is difficult to keep tracking when the objects have similar features to tracking object. Especially when the same team's players come close 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. In occluded situations, we apply a Real AdaBoost[10] & Histograms of Oriented Gradients (HOG)[11] classifier in order to detect the players and resample the center of gravities. We extract 1944 dimensions HOG feature from 30 $\times$ 40 pixels local area.

Decision of occlusion is based on between two players distance. We define the distance don't violate each Particle Filter's effective range. Figure3 shows player detection and center of gravities resampling considering player's velocity under occlusion. We obtain player's velocity among multiple frames. The player's center of gravity is decided the nearest estimation point.

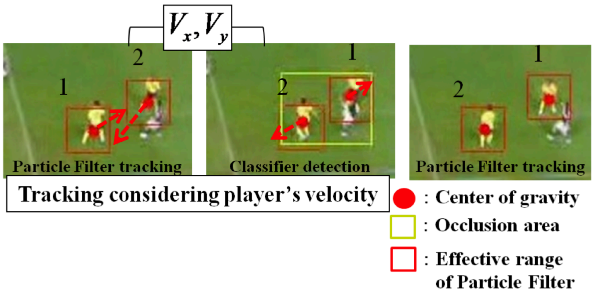

Fig. 3 Flow of player tracking

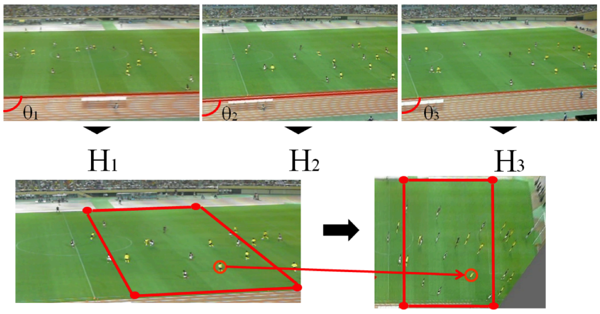

Fig. 4 Projective transformation

2.3 Ball tracking

In this section, we describe the method of minute ball tracking in images. It is difficult to track ball because the size is approximately 3 × 3 - 7 × 7 pixels. Accordingly, we evaluate ball region using labeling and correspond to the center of gravity by nearest neighbor. This situation, extracting number of pixels and center of gravity as region information. The binalization is calculated the same way in section 2.1. The number of pixels are 15 - 70 pixels in this scene. This is the condition of spatio-temporal ball corresponding. We stop ball tracking up to appear again, if this tracking method loses sight of the ball.

2.4 Projective transformation

We project player's point from original image to bird's-eye view (Figure4). Projective transformation consists of preprocessing and projection in each frame.

As preprocessing, we correspond the sideline's angle to transformation matrix. The line detection algorithm is Hough transformation from lower part of pitch area. The average of all lines is the representation angle in the image. The corresponding points between original image and bird's-eye view is acquired manually.

In each frame, we detect sideline in the same way and specify player's position on the pitch applying the corresponding perspective matrix.

3. EXPERIMENT

We verified proposed method, player tracking, ball tracking and perspective transformation in the real world football game scenes. The image size is 640 × 480 pixels. The camera operation is only swing motion. Our proposed method runs Dell Precision T5400, Windows XP 32bit. The PC has 3.25GB RAM.

3.1 Player tracking experiment

We compared proposed method with previous method. This experiment verifies tracking accuracy under occluded situations. Moreover, there are two types of situation, the different team and the same team. These are difference in that the uniform's color is the same or not. The success is the situation that one method don't lose tracking player when overlapping each other. Table 1 shows tracking accuracy, our proposed method and previous methods. "Color" calculates only color histogram, and "Color + Edge" calculates combination[9] color histogram and shape of edge. "SIR Particle Filter" indicates the [1]

Table 1 Player tracking accuracy in occlusion

	different team	same team
Color	98 / 100 (98%)	15 / 100 (15%)
Color + Edge	99 / 100 (99%)	26 / 100 (26%)
SIR Particle Filter	98 / 100 (98%)	68 / 100 (68%)
Proposed method	98 / 100 (98%)	90 / 100 (90%)

In the case of different team occlusion, the tracking rate is high in all method. Color based Particle Filter can distinguish the difference of uniform's color easily.

By contrast, previous methods can't track in the case of the same team occlusion. The method that only calculating color histogram can't track in the case of player distance is close. The method of "Color and Edge" can track player's approach, however, it is difficult to process most part of occlusion. "SIR Particle Filter" tracks considering occlusion. However, this method can't track after the most part of occlusion in this experiment.

Our proposed method can track players successfully, put into practice tracking rate 90%. Figure 5 shows the previous method and proposed method tracking. Resampling the center of gravity is effective for occlusion. Real AdaBoost classifier can give center of gravities when players separate from each other. Furthermore, we enable to identify player number by considering player's velocity. However, as for a failing situation of our proposed method, many players crowd the same place. This is not able to acquire player's edge gradient. To solve this problem, we'll detect the player go out of occlusion and identify player number later.

3.2 Ball tracking experiment

Next experiment is ball tracking. Ball position is manually operated in first frame, and we give positional information in the case of miss tracking. This experiment evaluates ball detection rate. Proposed and previous method apply for 7 football game scenes, 2,000 frames. We used Particle Filter combining color histogram + edge[9] and Zero-mean Normalized Cross-Correlation (ZNCC)[12] as previous method. Table 2 shows the result of ball tracking experiment.

"Color + Edge" tracking occur miss matching in the case of ball and player occlusion. It is difficult to calcu-

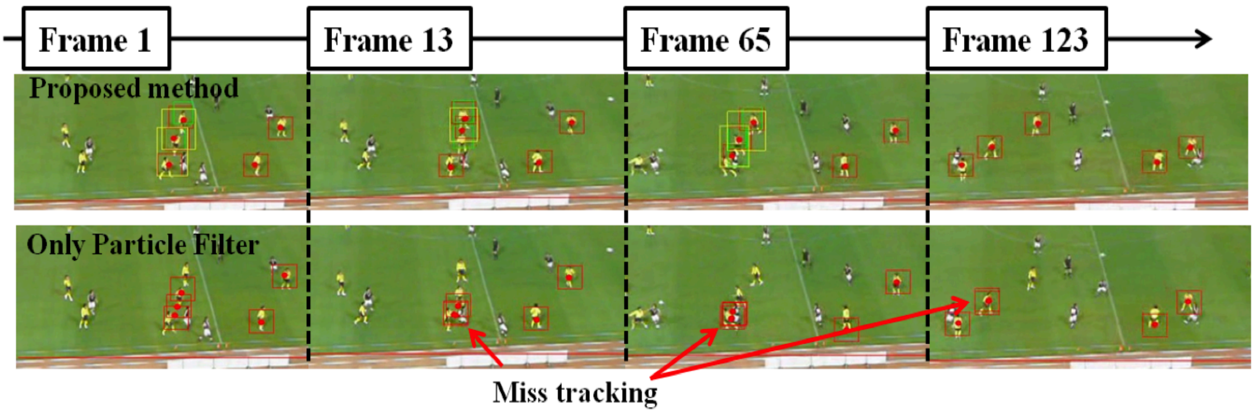

Fig. 5 Player tracking experiment

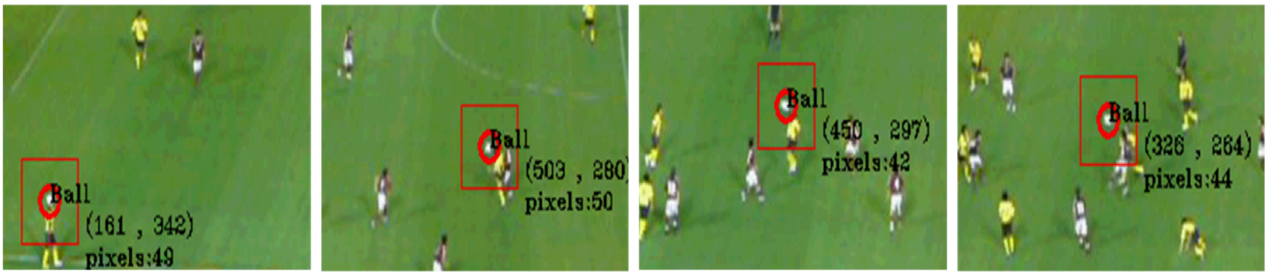

Fig. 6 Ball tracking experiment

Table 2 Ball detection rate

	Success frame	Detection rate
Color + Edge	636 / 2000	31.8%
ZNCC	can't track	—
Proposed method	1664 / 2000	83.2%

late likelihood. The color histogram calculation is unable to extract histogram stably. The edge shape is hard to evaluate object shape under occlusion.

”ZNCC” matching cannot track almost all frames. The ball template is too simple and small to calculate likelihood. ZNCC is not suitable tracking method in this case.

In contrast, proposed method extracts number of pixels and correspond to near labeled area from previous frame (Figure 6). For that reason, Labeling and nearest neighbor algorithm is effective for minute ball tracking. Area restriction works successfly no overlapping situations. Though proposed method is able to match under crowded area, it has difficulty in overlapping most part of the ball. To overcome this problem, we connect ball trajectory when the ball is redetected after overlapping.

3.3 Perspective transformation experiment

We obtain player position on the pitch projecting from original image to bird's-eye view. Figure 7 is shown the one of example trajectory, and Figure 8 is multi-players position on the pitch. As a result of perspective projection, we enabled to extract trajectories on the pitch using monocular camera. The trajectories have no gap in this experiment. And more, Table 3 shows the processing time through pitch area estimation to perspective projection.

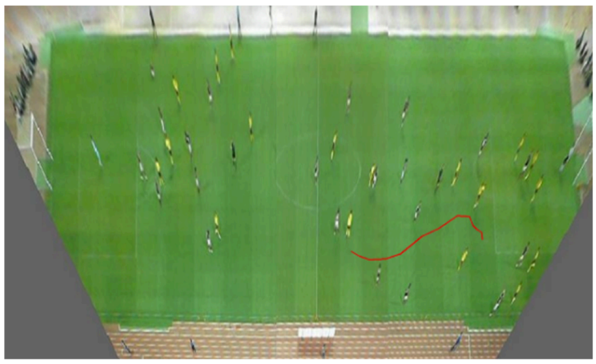

Fig. 7 Player trajectory

Table 3 Processing time

	Processing time(ms/frame)	Frame rate (fps)
4 players	36.39	27.48
11 players	47.96	20.85

From the result, we actualized the high-speed method for football video analysis. This processing time comes pitch area estimation, occlusion area detection and preprocessing of projection. Pitch area estimation enable to restrict processing area. Consequently, we can care not only processing time but also miss detection/tracking. As for occlusion area detection, this step suppress the classifier detection area. It is effective for high-speed detection

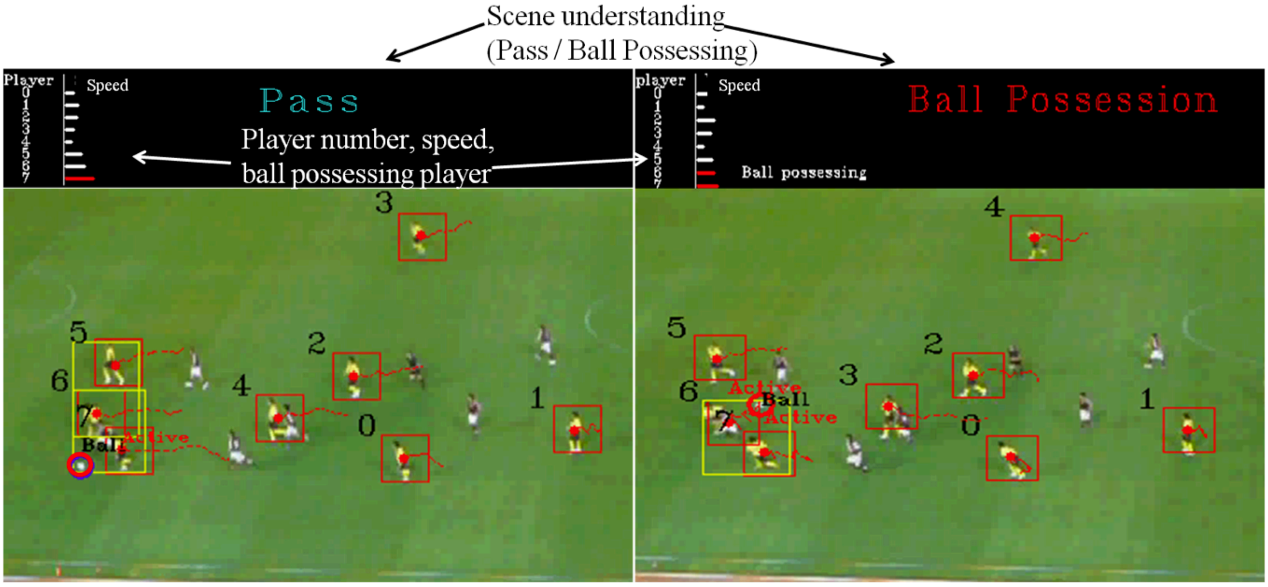

Fig. 9 Tactics analysis

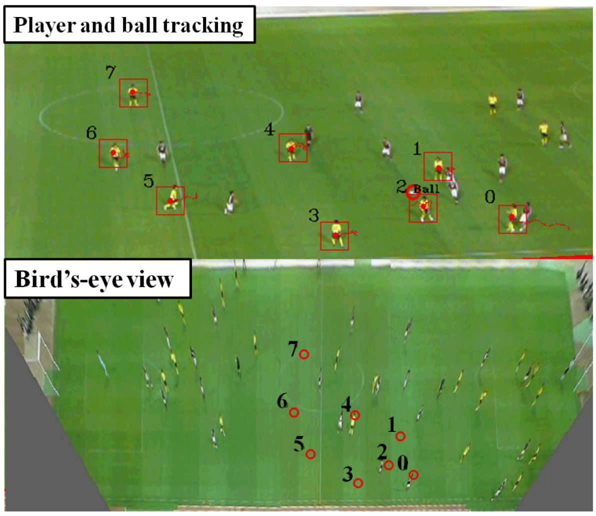

Fig. 8 Player projection

due to the classifier detection is time-consuming. Preprocessing of projection bring into process decreasing. We can get trajectories on the pitch from only angle of sideline.

4. FOOTBALL VIDEO ANALYSIS

This step explains about ”tactics analysis”. In our method, we enabled to acquire player and ball position from football video and player trajectories on the pitch. Figure 9 shows the football tactics analysis based on positional information. The bar graph shows player’s speed. This system displays ”active” player from player’s relative speed. Moreover, our system can understand scenes, ”Ball Possessing” or ”Pass” from player and ball position.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed the players and ball tracking method for football movie analysis using monocular camera. Each player tracking technique is Particle Filter. In occlusion, classifier detects players and resamples the center of gravities. Ball tracking is applied labeling and nearest neighbor. In order to specify player’s position on the pitch, we convert the position in image to bird’-eye view. In addition, we realized ”tactics analysis” with monocular camera.

Future work will be 3D ball tracking to estimate ball position on the pitch. As for the video analysis, we’ll accumulate a number of football scenes to analyze characteristic as a team or player.

REFERENCES

[1] Anthony Dearden, Yiannis Demiris, Oliver Grau, ”Tracking Football Player Movment From a Single Moving Camera Using Particle Filters”, 3rd European Conference on Visual Media Production (CVMP2006), pp.29-37, 2006

[2] Itaru Kitahara, Yuichi Ohta, Hideo Saito, Shinji Akimichi, Tooru Ono, Takeo Kanade, ” Recording of Multiple Videos in Large-scale Space for Large-scale Virtualized Reality ”, Proceedings of International Display Workshops (AD/IDW’01), pp. 1377-1380(2001)

[3] Nozomu Kasuya, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta,”Watching a Player’s View in Real Soccer Scenes by Using Player Trajectories”,The Second Korea-Japan Workshop on Mixed Reality (KJMR09), 4 pages, 2009.

[4] Michael Beetz, Suat Gedikli, Jan Bandouch, Bernhard Kirchlechner, Nico v. Hoyningen-Huene, Alexander Perzylo, ”Visually tracking football

games based on TV broadcasts", Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI), 2007

[5] Michael Beetz , Bernhard Kirchlechner , Martin Lames, "Computerized Real-Time Analysis of Football Games", IEEE Pervasive Computing, 2005

[6] Nicolai v. Hoyningen-Huene, Michael Beetz, " Robust real-time multiple target tracking ", The Ninth Asian Conference on Computer Vision MP1-24, Sep 2009

[7] Raffay Hamid, Ramkrishan Kumar, Matthias Grundmann, Kihwan Kim, Irfan Essa, Jessica Hodgins, "Player Localization Using Multiple Static Cameras for Sports Visualization", 23rd IEEE Conference on Computer Vision and Pattern Recognition(CVPR2010), June 2010

[8] Isard M, Blake A, " Condensation : conditional density propagation for visual tracking ", in Int. J. Computer Vision, vol.28, no.1, pp.5-28, 1998

[9] Ke-Yan Liu, Shan-Qing Li, Liang Tang, Lei Wang Wei Liu, "FAST FACE TRACKING USING PARALLEL PARTICLE FILTER ALGORITHM", ICME2009, pp.1302-1305, 2009

[10] Y. Freund, R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", Computational Learning Theory, Eurocolt, pp.23-27, 1995

[11] N Dalal,B Triggs, "Histograms of oriented gradients for human detection", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.886-893, 2005

[12] Takuro Nishino, Yasuo Ariki, Tetsuya Takiguchi, "Situation Recognition Using 3D Positional Information of Ball from Monocular Soccer Image Sequence", The 2009 International Conference on Multimedia, Information Technology and its Applications (MITA2009), pp.109-112, 2009