Intelligent Data Analysis 13 (2009) 833–857
DOI 10.3233/IDA-2009-0395
IOS Press

A novel multimedia data mining framework for information extraction of a soccer video stream

Ehsan Pazouki* and Mohammad Rahmati
Computer Engineering Department, Amirkabir University of Technology, Hafez Street, Tehran, Iran

Abstract. A video stream is usually massive in terms of data content with abundant information. In the past, extracting explicit semantic information from a video stream; i.e. object detection, object tracking and information extraction; has been extensively investigated. However, little work has been devoted on the problem of discovering global or implicit information from huge video streams. In this paper, a framework has been presented for extracting information for a specified player from soccer video broadcast by data mining techniques. Concepts and information which exist in a soccer video broadcast are useful for team coaches. But, due to various reasons; i.e. wide field of view of a video stream, huge data, existence of great number of important objects in the play field of a soccer match and the occurrence of number of important events, manual extraction of information from soccer video broadcast is difficult and time consuming task. In this paper, a set of techniques is presented that automatically extract some useful information of a player, i.e. velocity and traversed distance, from a soccer video broadcast. Processing of video sequence under change of lighting conditions, fast camera movement and player's occlusion is a challenging task. Our proposed framework comprise of 3 stages, player segmentation, player tracking and information extraction. All three stages must be robust under various challenges. The performance of our proposed system has been evaluated using a variety of soccer video broadcast having different characteristics in term of lighting conditions. The experiments showed that the efficiency of our system is satisfactory.

Keywords: Multimedia data mining, soccer video analysis, information extraction, player tracking

1. Introduction

Introduction of different techniques for understanding digital video content has become a hot research topic in recent years. Their main focus is on structure analysis [12] and event detection [5]. Since video contents normally take a long period of time to play and occupy a large storage space, mining useful information from the whole video content help users to have a better understanding, however, it is usually a difficult task and it has not been fully explored yet.

Data mining deals with the process of data for finding useful patterns or extracting previously unknown knowledge from a massive set of data [15]. In contrast to the textual information, which has been studied for a long time before the information of video contents are different. Two special characteristics of video content are 1) they are continuous sequences with temporal relations among them, 2) each video segment normally contains abundant information itself. Due to the semantic gap between human perception and computer based low-level features, many video understanding techniques such as shot boundary

*Corresponding author. E-mail: ehsanp61@aut.ac.ir.

1088-467X/09/$17.00 © 2009 – IOS Press and the authors. All rights reserved

detection and event extraction remain open problems. Thus, investigation on video mining is at its early stage and it normally requires extracting semantic information. The main motivation of video mining is to find and discover knowledge from the stream based on visual and audio cues. The knowledge may typically include structural information within a video clip, association information among various clips, and trend information based on the analysis for a massive size of video set.

Recently, mining information in sports video data, especially soccer videos, has become an active research topic. In the data mining technique, some algorithm such as, clustering and classification are performed on the large scale data set, and useful information is extracted. This information may be used for analysis of primary data set in a better way. For example analyzer and coaches are interested to attain useful information of a player from the video data. Extracting these information involves problems such as players detection, players tracking and player occlusion.

Furthermore, data mining from human motion has become an important subject in machine vision community. Among many aspects of this area, the sports video analyzing especially, in the ball games such as soccer has attracted much attention. Motion analysis in this area involves many problems such as, rapid and unpredictable motion, sudden direction changing, occlusion and collision of target objects by other objects and non-rigid target objects such as players and ball. Many researches [1,3,4,14,16,18, 21–29] have been introduced to handle part and/or all of these problems.

An evaluation model was proposed in [4] to quantitatively express the performance of soccer players, using as input the relationships between the trajectories of 22 players and a ball and having as output the performance evaluation of several players in a quantitative way. Ekin et al. [9], have assumed that the presence of soccer highlights can be inferred from the occurrence of one or several slow motion shots and from the presence of shots where the referee and/or the goal post is framed. Particle filtering, also known as the Condensation algorithm [19] and bootstrap filter [8], has recently been proven to be a powerful and reliable tool for tracking application. Choi et al. [24] used the TV broadcasted images and tracked players by template matching and, in case of occlusions, they used the technique of histogram back projection. The position and bounding box of ball at the starting time was manually initialized, and then Kalman filtering and template matching were performed for ball tracking around the nearby player. In [26] they also adopt multiple cameras to track players and apply the results in free-view visualization generation. In their system, the players are tracked in a virtual ground image instead of using original image.

In general, Different limitations have been assumed by other systems for soccer video analysis, these are,

1. Multi cameras, which are placed in determined positions, are used to cover all play field.
2. Single and fixed camera, for whole field coverage area was used in indoor sports.
3. Extraction of information are event based and not based on players information.
4. Almost all systems required manual initializations.
5. The information extraction was limited to part of play field and not the whole play field, for example to goal region.
6. Some manual interaction during the learning phase are needed for modeling of important object.

In contrast, our system have important features which are,

a- Cramp initializations are not necessary.
b- The soccer video broadcast data that we have used have placed no conditions on the camera parameter.
c- Information could be extracted along any particular part of play field as long as they are bounded within the grass field.

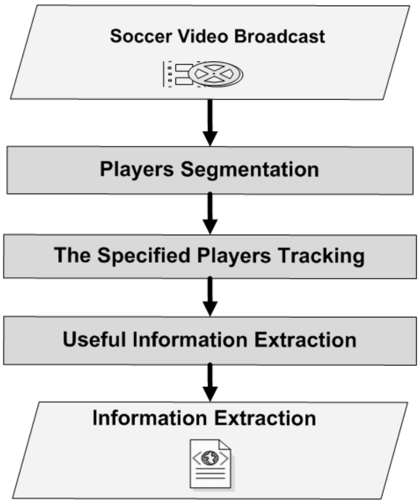


Fig. 1. Global block diagram of our proposed system.

Information extraction of a player in our proposed system involves several stages. The players are segmented in the field region and they are tracked on segmented frames using outfit model of two teams and some knowledge for players. Then, the information is extracted from obtained players' route. The block diagram of proposed algorithm is presented in Fig. 1.

The remainder of this paper is organized as follows: the proposed algorithm for player segmentation is presented in Section 2. The algorithm for extracting the uniform team model is presented in Section 3. In the Section 4, the player tracking mechanism is presented. Finally, in the Section 5 an approach for obtaining the information from extracted players' trajectory is presented. Experimental evolution of our proposed system is in Sections 6 and 7 is the concluding remarks

2. Player segmentation

In the first step of our data mining framework, input video sequence data are clustered into two; clusters of players and non-players. Clustering is performed based on the coordinate of pixels on the frame boundary and color value of pixels by image processing techniques.

The segmentation results have an important effect on the next steps. If objects are not detected in this step, they cannot easily be recovered in the later steps. In other words, an effective approach in the segmentation approach would have an important effect on the other steps. Therefore, this section presents an effective, accurate and general approach for segmentation section of our system.

Many researches have been introduced to cope with different aspects of this topic. Khatonabadi et al. [25] proposed a technique for soccer analysis in goal scenes. Also, they used a field model to provide panoramic view of the scenes. They used background color approximation and some geometric and structural parameters for segmenting the field area. Pascual et al. [23] proposed a technique for players tracking in soccer. They used background subtraction by dynamic modeling for player segmentation. Vandenbroucke et al. [21] proposed a new technique for finding hybrid color space, and then used this color space for accurate segmentation. Utsumi et al. [22] proposed a novel approach for detection and tracking of players by means of kernel creation. They used local edge and color rarity for segmentation.

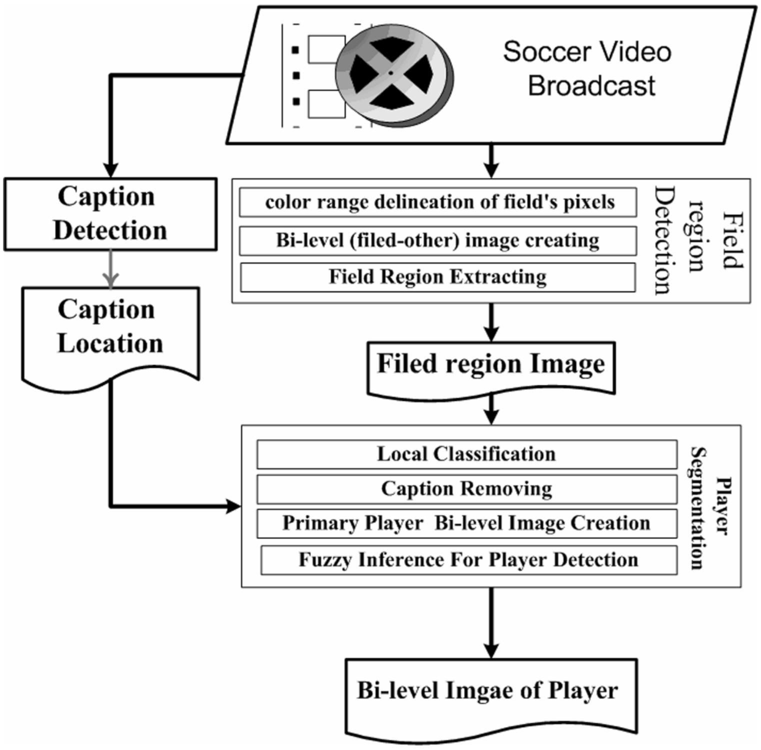


Fig. 2. Block diagram of proposed approach for player segmentation.

Gong et al. [29] proposed an approach that players, ball and important lines features and motion are used for analyzing soccer video broadcast.

In this section, an innovative approach is proposed for segmentation of players in the field of soccer game. The general block diagram of our approach is shown in Fig. 2. This stage is comprised of several steps such as, caption detection, field region detection, player segmentation.

2.1. *The caption detection in soccer video broadcast*

Removal of annotation and captions that have been inserted in soccer video must be performed in the first step. These captions may have an important effect on the other sections of our system. Therefore, in the first step, location of captions should be detected and removed. Here, captions are detected by using their essential features which is their fixed locations. In this method, the standard deviation of each pixel across the consecutive frames is calculated. Then, the candidate caption pixels are marked using a threshold on the standard deviations. The Bi-level image of caption is generated using the candidate pixels followed by improvement stage employing noise reduction operators such as 5×5 median filter and closing and dilation morphological operator with a 5×5 ellipsoidal structural element. The type and feature of filter and structural element are obtained under different situations [10]. In Fig. 3, the Bi-level image of captions is shown where the detected captions may be removed for other operations.

2.2. *Field region detection*

The field region extraction is important because by bounding the remaing processes within the field region lead to decreasing complexity and have critical effect on achieving accurate results. Field region

(a) (b)

Fig. 3. Bi-level image of caption, (a) A frame of soccer game. (b) Bi-level image of caption.

localization is performed in three steps. First, color range of field's pixels is delineated. Second, Bi-level field or non-field image is segmented from color range. Then, field region are extracted from created Bi-level image.

2.2.1. Color range delineation of field's pixels

In our proposed approach, we use RGB color space. Since the video scenes that we used are long shot view of main camera, hence, most pixels of these scenes belong to the field region. To delineate this range, the pixel values at the maximum point of histogram for each color's channel are used. However, the histograms of these scenes are skewed. Therefore, to delineate range accurately, the approach proposed by Khatonabadi in [25] is used. In this approach, first, histogram is thresholded on bin's value, then, centroid of achieved histogram is calculated using,

$$i'_{\text{peak}} = \frac{\sum\limits_{H(i) \geqslant \beta \cdot i_{\text{peak}}} i \cdot H(i)}{\sum\limits_{H(i) \geqslant \beta \cdot i_{\text{peak}}} H(i)} \tag{1}$$

where i'_{peak} is centroid value, i_{peak} is peak point of histogram, i is histogram index and $H(i)$ is bin value on i index. The β is threshold bin value, that in our case 10% are selected. This value is used as median of range. Finally, variance of achieved centroid is calculated, and is used as variance of rang.

2.2.2. Bi-level field creation

A membership function which is used as a measure to determine the pixel's field membership, is defined, which is,

$$\mu(x) = e^{-\left(\frac{x - i'_{\text{peak}}}{2\times(\alpha\times\sigma_{\text{peak}})}\right)^2}, \tag{2}$$

where σ_{peak} is variance of centroid value and α is variance bias. The Gaussian membership function is used for each channel of color space, where the calculated centroid value is used as its mean and a percentile of calculated variance of centroid is used as its variance. Obviously, the non-field membership

Fig. 4. Selected field region image of Fig. 3(a) using our algorithm.

measure value is $1 - \mu(x)$. The primary Bi-level image of the field is generated using the following two steps. First, using the threshold value for the membership measure, candidate pixels are detected. In the second step, an extra condition, which ensures that the green channel value of color space is greater than other two channels is used.

2.2.3. Field region extraction

The field region extraction is performed in two steps. In the first step, an image area reduction technique is performed on the primary Bi-level image of field obtained in Subsection 2-2-2. The technique is started by assigning primary image area with same size as the primary Bi-level image. The image area size is decreased iteratively. Decreasing process is supervised by counting deleted field's pixels and a threshold is used on this counted value. Finally, the image area is obtained, and by masking this image area on the primary Bi-level image of field, parts of non-field objects are removed without removing field's parts.

In the second step, the obtained image area is applied on the primary Bi-level image of field and a new image is generated. Then, connected components of field's pixel are detected from the new image. Finally, the largest connected component is selected and the bounded rectangle on this connected component is extracted as field region and is used on the next steps. An example is shown in Fig. 4 where the field region is extracted by using this algorithm.

2.3. Player segmentation

In this step, combination of local classification and fuzzy inference are used for extracting the players. The proposed approach does not use any special kernel for players. Only, the global information for player; general specification of player and field's color, are used for player segmentation.

2.3.1. Local classification

The local classification is performed by using a rectangular shape kernel on image of the field. The origin of the kernel is placed on each pixel of field image. At each step, the pixels of image within the kernel are classified using the value of non-field membership measure. This value is calculated similar to the technique introduced in Section 2-2-2. In order to make it robust, modifications are made. In Section 2-2-2, the global i'_{peak} and σ_{peak} are used on the whole image. Here, in the first step, image is

(a) (b)

Fig. 5. Bi-level image of players for Fig. 3. (a) Primary bi-level image of players. (b) Final bi-level image of players.

partitioned into four blocks of equal sizes, then i'_{peak} and σ_{peak} are calculated for each block separately. The value of non-field measure of each pixel is calculated using i'_{peak} and σ_{peak} corresponding to the block of this pixel. Then, pixels are classified using a threshold value on calculated value of non-field measure.

A matrix is created with the same size as the image of field, such that each entry of the matrix represents the membership of a player corresponding to pixel on the image of the field. The corresponding center point of kernel to the matrix is assigned with a value γ which is between zero and one and is obtained from,

$$\gamma = \frac{(KA + NPC - FPC)}{2 \times KA}, \tag{3}$$

where *KA* is kernel area, *NPC* is number of non-fields' pixel in classification and *FPC* is number of fields' pixel in classification. Finally, the primary Bi-level image of players (Fig. 5-a) is provided by applying a threshold value on the value of the resulting matrix.

2.3.2. Fuzzy inference

In this step, connected components are generated from the primary Bi-level image of the players. For each of the resulting connected component, parameters such as area, the amount of compactness and the ratio of a height to width are computed. The compactness is calculated using,

$$\varsigma = \frac{P^2}{4\pi \times A^2}, \tag{4}$$

where P is perimeter of bounded rectangular on connected components and A is area of bounded rectangular on connected components.

For each of these parameters the membership functions have been defined, then three fuzzy rules have been established according to these parameters. Finally, using fuzzy inference, the defined fuzzy rules and the values of the connected components parameters of the Bi-level image of the players is generated. In order to remove noise and obtaining better result a 5 $\times$ 5 median filter [10] and morphological operator with a 5 $\times$ 5 rectangular shape structural element has been used for the resulting image (Fig. 5-b).

3. Uniform color model for team extraction

In the second stage of our data mining framework, the cluster of players, found in previous phase, is clustered into two representing the competing teams and kernels of each cluster are used as uniform

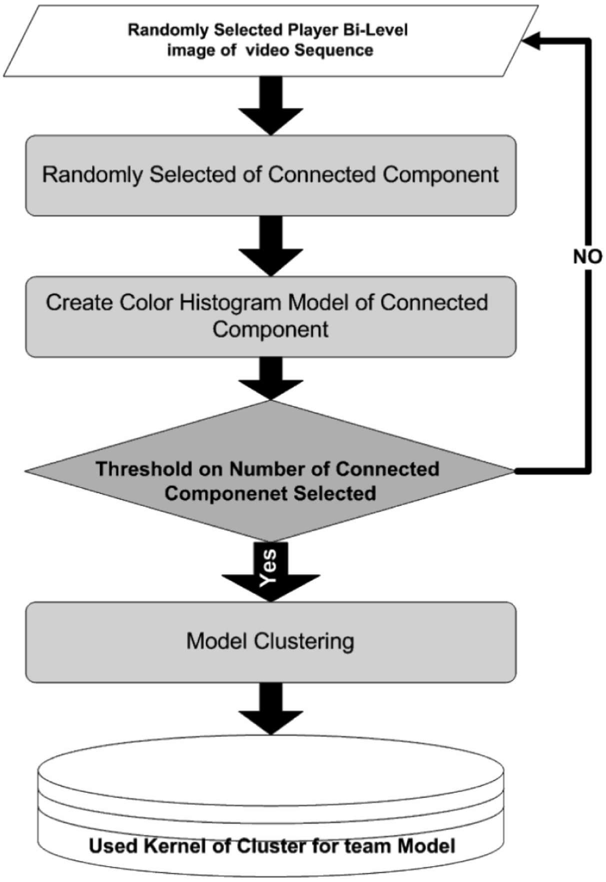


Fig. 6. Team model creation block diagram.

model of these teams.

The rules set by FIFA in a soccer game ensure players of a team wear similar uniform in a match which is different from the opponent team. Therefore, uniform of each team is an appropriate feature for tracking the players of a team. Usually, in the past, uniform of each team are modeled manually by a convenient GUI [1,21]. In this paper, we have proposed an automatic approach for extracting uniform model. In this section, the detail of proposed approach is presented. The block diagram of proposed technique is presented in Fig. 6.

3.1. Structure of the player model

Selecting a model for uniform of each team is an important decision. The color model of uniform wear by each team, which is extracted, is used to track the player, hence, the model must be proper representative of players. The players have dynamic contour, texture and size on the sequence of frames. Hence, different researchers have used variety of features for non-rigid object modeling and tracking. Since, a soccer player in a soccer video broadcast is a non-rigid object, RGB color histogram is used for player and uniform color clothes of each team modeling. Tracking techniques that use color histogram model are adequate for tracking of non-rigid object even in the presence of variation of sizes. The color

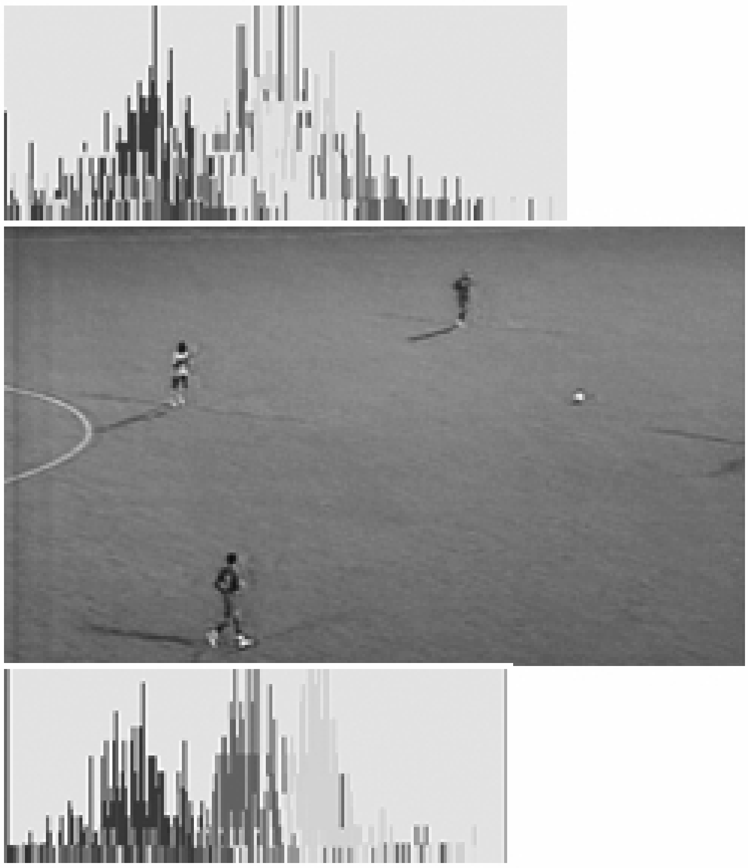


Fig. 7. RGB color histogram of players of two different teams.

histogram model is calculated by,

$$\hat{q}_u = \frac{1}{n} \sum_{i=1}^{n} \delta[b(x_i) - u], \tag{5}$$

where u is histogram bin; x_i is location of th pixel; $b(x_i)$ is associated bin of the pixel at location x_i in the quantized feature space; δ is the kronecker delta function; n is the number of pixels. Sample of RGB color histogram is presented on Fig. 7.

3.2. Uniform team model extraction

The uniform team model is automatically extracted in three steps. First, the constant numbers of connected components are randomly selected from the Bi-level image of players. Second, the RGB color histogram of each selected connected component is created. Third, the created models are clustered into two clusters. The kernel of each cluster is used as the model for the uniform of a team.

Selection of the number of clusters is a critical decision. In the literatures, five clusters are usually used for people present in a soccer match; i.e., players of team A, players of team B, goalkeeper of both teams and referee. In this paper, we only extract players of teams A and B and other persons in the match aren't important for our objective. Therefore, two clusters are spotted for clustering model.

A technique that is used for clustering should be robust in the presence of noise and outliers. The fuzzy possibilistic c-means technique [20] (FPCM) is used for clustering. This technique is fast and handles noise and outliers. Kernels of clusters, obtained by FPCM technique are treated as uniform team model. Since, RGB color histogram obtained by Eq. (5) is a probability distribution, hence, the distance measure of clustering technique is the Bhattacharyya coefficient that is obtained by,

$$\rho_k[\hat{p}_k, \hat{q}_k] = \sum_{u=1}^{m} \sqrt{\hat{p}_{k(u)} \times \hat{q}_{k(u)}}, \textit{ where } k = (R, G, B) \tag{6}$$

where m is a number of bins on histogram model; $\hat{p}$ and $\hat{q}$ are the two histogram models that are compared. Then, the following relation,

$$d = \sqrt{1 - \rho_R[\hat{p}_R, \hat{q}_R]} \times \sqrt{1 - \rho_G[\hat{p}_G, \hat{q}_G]} \times \sqrt{1 - \rho_B[\hat{p}_B, \hat{q}_B]} \tag{7}$$

is used as the distance measure for clustering process.

4. Players tracking

The players are modeled by the same structure that is presented in Section 3-1. Tracking phase of our algorithm consist of two steps. The first, players are tracked in the consecutive frames by RGB color histogram model. In the second step, the specified player is tracked, on the overall match, using uniform color model for the team of the specified player.

4.1. Blind tracking of players

In the third step of our data mining framework, sequences of repeated of templates are extracted by sequence mining technique. The sequences of repetition are searched on the clustered video data by bounded number of frames. The consecutive sequences are extracted by clustering data of players for two consecutive bounded number of frames. This section is performed by tracking technique.

In this section, the players are tracked in the consecutive frames. However, the players are non-rigid in the soccer video and occlusions between players are occurred frequently; therefore, the tracking method that is used must be suitable under these conditions. The players are tracked on consecutive frames using merge histogram back projection technique and CAMShift [3] method. Tracking algorithm has three main steps.

1) The players that enter the camera field of view with delay, is tracked. Input of this step is the Bi-level image of the players for first frame or Bi-level image of the players modified at step 2. For each connected component in the Bi-level image of players, an ellipsoidal template is assumed and a model of this connected component is created which is,

$$CC_{i_t}(ID_i, x_{i_t}, y_{i_t}, AA_{i_t}, AI_{i_t}, H_{i_t}, T_{i_t} = T_{\text{default}}), \tag{8}$$

where ID_i is an unique identifier for the connected component model, x_{i_t} and y_{i_t} are the center coordinate of ellipsoidal model, AA_{i_t} and Al_{i_t} are the main and minor axis of the ellipsoidal model, H_{i_t} is the RGB color histogram model, as presented in Section 3-1, T_{i_t} is the imperceptible counter and T_{default} is default value for T_{i_t}.

2) The whole sequence is processed frame by frame and the Bi-level image of the players for each frame are extracted, then, Bi-level image of the players are corrected and are used in the tracking process. In this step, all model denoted by $CC_{i_{(t-1)}}$, that was tracked on the previous frames are tracked on the current frame. The model $CC_{i_{(t-1)}}$ is tracked in the current frame as explained in the following. New center of coordinate for $CC_{i_{(t-1)}}$ is estimated by Kalman filter by using,

$$\begin{aligned} \hat{x} &= \mathit{kalmanFilter}(x_{(t-1)}), \\ \hat{y} &= \mathit{kalmanFilter}(y_{(t-1)}) \end{aligned} \tag{9}$$

where $\hat{x}_t$ and $\hat{y}_t$ are the new center of coordinate for $CC_{i_{(t-1)}}$. The $H_{i_{(t-1)}}$ that is the RGB color histogram for $CC_{i_{(t-1)}}$ is applied to the current frame by Histogram Back Projection (HBP) approach [6]. Result of this technique is a matrix that its elements are determined by,

$$M[x, y] = q_R(b_R(x, y)) \times q_G(b_G(x, y)) \times q_B(b_B(x, y)), \tag{10}$$

where q_R, q_G and q_B are the RGB components of $H_{i_{(t-1)}}$ and $b_R(x, y), b_G(x, y)$ and $b_B(x, y)$ are the component values of RGB which are extracted at coordinate (x, y) space.

The CAMShift technique is used for finding the best ellipsoid fitted on the calculated matrix of Eq. (10) and from primary ellipsoid model by center coordinate ($\hat{x}_t, \hat{y}_t$) and axis ($AA_{i_{(t-1)}}, AI_{i_{(t-1)}}$, then, the color RGB histogram model from output of CAMShift ellipsoidal model H_{i_t} by center coordinate (x_t, y_t) and axis (AA_{i_t}, AI_{i_t}) is calculated. The similarity measure between new model H_{i_t} and $H_{i_{(t-1)}}$ is calculated by Eq. (6). If the similarity is greater than a threshold value, the $CC_{i_{(t-1)}}$ is tracked in the current frame and the corresponding refreshed model is,

$$CC_{i_t}(ID_i, x_{i_t}, y_{i_t}, AA_{i_t}, AI_{i_t}, H_{i_t}, T_{i_t} = T_{\text{default}}), \tag{11}$$

and, its corresponding ellipsoidal model in the Bi-level image of players is removed. If the similarity is smaller than the threshold value, the $CC_{i_{(t-1)}}$ has not been found on the current frame, and therefore the CC_{i_t} is refreshed as,

$$CC_{i_t}(ID_i, \hat{x}_{i_t}, \hat{y}_{i_t}, AA_{i_{(t-1)}}, AI_{i_{(t-1)}}, H_{i_{(t-1)}}, T_{i_{(t-1)}} - 1) \tag{12}$$

After tracking all $CC_{i_{(t-1)}}$ in the current frame, the modified Bi-level image of the players are processed by step (1) and new players are entered on the current frame are detected.

3) The players that have been moved away from the camera field of view are detected. If the $T_{i_{(t-1)}}$ is equal to zero then $CC_{i_{(t-1)}}$ is removed from tracking list of the players.

The proposed algorithm is useful for non-rigid objects tracking. Also, it has adequate robustness in the partially occlusion between players. The performance of the algorithm is presented in the Section 6 .

4.2. Specified player tracking

In the fourth stage of our data mining framework, the non-consecutive sequence of specified template of a player is extracted based on using the consecutive sequence of previous section, the model as explained in Section 3, and features explained below.

In this subsection, we explain how a player, belonging to a specified team, and within a play zone, is tracked. The diagram of proposed approach for specified player tracking is presented in Fig. 8. The specified player is tracked by three features, which are:

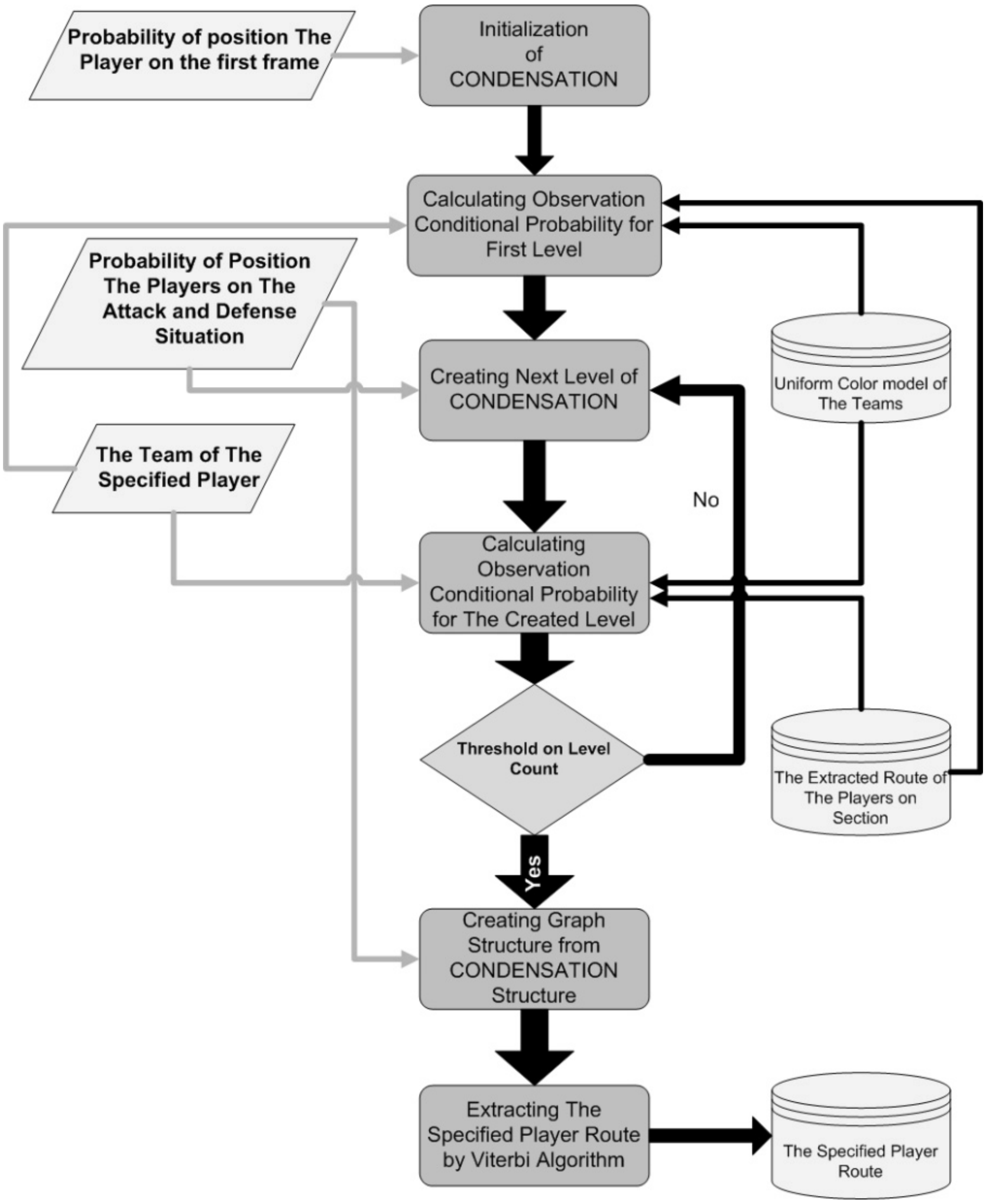


Fig. 8. The diagram of proposed approach for special player tracking.

1) The probability distribution of the specified player position in the first frame
2) The probability distribution of the specified player position in the soccer field, taking into account attacking and defending situations.
3) The specified player team

The specified player is tracked by the CONDENSATION (CONditional DENsity propagaTION) [19] algorithm.

4.2.1. The soccer field model

The specified player trajectory must be extracted within the soccer field and mapped into a field model. Hence, the soccer field must be simulated by a digital model in the system and the frames of the soccer video broadcast must be mapped to the soccer field model correctly. Therefore, the mapping parameters must be obtained, in order to map the frames of a soccer video broadcast to the soccer field model. The parameters may obtained either by using some electronic equipments on the cameras or by using a semiautomatic software and GUI on the soccer field lines.

Afterwards, the position of each extracted connected component, as explained in Section 4-1 could be mapped to the soccer field model using the mapping parameters. In the specified player tracking section, all coordinates are related to the soccer field model. So, all distance measures are calculated in the soccer field model coordinate.

4.2.2. The Structure of modeling using CONDENSATION algorithm

The CONDENSATION algorithm is a probabilistic algorithm. It is an extension of particle filter algorithm, which is used for tracking in non-Gaussian systems. The algorithm has three layers of hierarchical structure. The lowest layer is a sample layer. The samples are structures from tracking system model. In the proposed approach the tracking system is the specified player on the field model. The samples have structure represented by five parameters as,

$$\textit{sample}_i(x_i, y_i, ID_i, Len_i, P_i), \tag{13}$$

where x_i and y_i are the position of $\textit{sample}_i$ in the soccer field model, ID_i is the unique identifier for the corresponding connected component model of $\textit{sample}_i$, $\textit{Len}_i$ is the tracking length of $\textit{sample}_i$ and the P_i is the probability of correctness of $\textit{sample}_i$. The probability of presence of a specified player at a specified position of the soccer field model is presented by P_i.

The second layer is a sample set. In the proposed approach each sample set correspond to one frame and it represents the probability of a specified player‘s position in that frame. The third layer is called super sample set which its elements are sample sets related to sequence of soccer video corresponding to the specified player tracking trajectory. The algorithm consist of the following steps:

1) Initialization of algorithm by creating the first sample set corresponding to the first frame and calculating the probability of correctness for each sample in the sample set
2) Creation of other layers comprising of three iterative steps: Drift, Diffuse, Measure.

The details of these two steps are given in the following,

4.2.3. Initialization

The first sample set is created by the first feature of the specified player which was explained in Section 4-2. The N points are randomly selected in the soccer field model which probability of presence of a specified player at these points is greater than a threshold value. The probability of presence of a specified player for each point is obtained by the first feature of the specified player. Therefore, the first sample set is created by N samples. Thus, for each sample its x_i and y_i parameters are determined.

For each sample of first sample set, the shortest distance of a connected component to this sample in the corresponding frame is selected and the unique identifier for that connected component is assigned to ID_i. The distance measure between the sample and a connected component is calculated by Euclidean distance between that position of the sample and the center of ellipsoidal connected component model in the soccer field model.

The P_i is obtained by calculating the similarity measure, that is Bhattacharyya coefficient, between the ellipsoidal model of ID_i and the uniform color model of the specified player team. Then, initialization step is performed and the first sample set is obtained.

4.2.4. creating the other sample sets

The rest of sample sets are created by two steps.

1. Drifting and diffusing
2. Measuring

The drifting and diffusing are explained first. For each sample in the previous sample set, four samples of a neighbor are created to form the new child samples. Then, for each child sample the corresponding connected component is obtained. If the connected component for the new sample is equal to the corresponding connected component of its parent, the tracking length parameter of the new sample is increased by one. Then, the correctness probability parameter of the new sample is calculated by equation,

$$p(z|x) = p_1(z|x) \times p_2(z|x) \times p_3(z|x). \tag{14}$$

where $p(z|x)$ is the conditional probability of observation given sample $x, p_1(z|x)$ is the similarity measure given by Eq. (6) between the histogram model of corresponding connected components of the sample x and the uniform color histogram of the specified player team, $p_2(z|x)$ is the correctness of tracking for the corresponding connected component of the sample x which is explained in Section 4-1. The $p_3(z|x)$ is calculated by the tracking length parameter of the sample x as,

$$p_3(z|x) = \begin{cases} \frac{(1-\mathit{offset})}{\mathit{threshold}} \times \mathit{Len}_x + \mathit{offset} & \mathit{Len}_x < \mathit{threshold} \\ 1 & \mathit{Len}_x \geqslant \mathit{threshold} \end{cases} \tag{15}$$

Hence, in drifting process the new sample set is obtained by previous sample set and is quadruplicated samples greater in numbers. In measuring step, the next sample set is obtained by selecting quarter samples from the new sample set using the conditional probability of observation.

4.2.5. Creating the graph structure from the CONDENSATION structure

The CONDENSATION algorithm is performed when the sequence of the soccer video is processed completely. Therefore, the sample set is created by some of the frame of the soccer video. Then, the weighted graph is created from the obtained CONDENSATION structure as following.

a) The nodes of graph are the samples of the sample sets
b) Each sample set is one layer of the graph
c) The edges of the graph connect the samples of the sample set to the samples of the next sample set in the CONDENSATION structure.
d) The weight of each node is the correctness probability of the corresponding sample in the CONDENSATION structure.
e) The weight of the edges are obtained by equation,

$$w_{i,j} = \begin{cases} \mathit{similarity}(H_i, H_j) & d_{i,j} \leqslant \mathit{threshold} \\ 0 & d_{i,j} > \mathit{threshold} \end{cases} \tag{16}$$

where $d_{i,j}$ is the Euclidean distance between two nodes, H_i and H_j are the color histogram models of the nodes and *similarity*() is the similarity measure between two models by Eq. (6). Then, the weighted graph is created from the CONDENSATION structure.

4.2.6. Route Extraction of the specified player

The specified player route is extracted from weighted graph using Viterbi algorithm. By repeating the algorithm for all players, the routes of all players on the soccer field model are obtained and the information is extracted.

5. Information extraction from specified player route

In the fifth stage of our data mining framework, useful information of a player is extracted from obtained non-consecutive sequence of specified template. The information about the specified player performance could be obtained from the extracted the trajectory. Some useful data that could be extracted are:

a) The existence percentile of a specified player in different zones of soccer field is determined and is defined by,

$$R_{i,j} = \frac{1}{n}\sum_{k=1}^{n} \mathit{Exs}(i,j,k) \times 100, i = 1,\ldots,12, j = 1,\ldots,22, \tag{17}$$

where,

$$\mathit{Exs}(i,j,k) = \begin{cases} 1 \textit{ if player } j \textit{ exist in region } i \textit{ at frame } k \\ 0 \qquad\qquad \textit{otherwise} \end{cases} \tag{18}$$

where j is the specified player number, i is the region of the soccer field index and k is the frame number.

b) The traversed distance of the specified player is estimated by,

$$\mathit{Distance}_j(k) = \sum_{f=2}^{k} D_j(f, f-1), j = 1,\ldots,20, \tag{19}$$

where k is the frame number and $D_j(f, f-1)$ is the function that calculates the traversed distance by the specified player between frames f and $f-1$ and 20 is total number of players that we specified within a soccer field.

c) Estimation of average velocity by the specified player is determined by,

$$\mathit{Vavg}_j = \frac{1}{n-1} t_{fr} \times \sum_{k=2}^{n} \frac{\mathit{Distance}_j(k)}{k}, j = 1,\ldots,2, \tag{20}$$

where t_{fr} is the time between two consecutive frames.

d) Measuring the obedience of a specified player from the team plan using the percentage of presence within the prescribed soccer field zone.

6. Experimental results

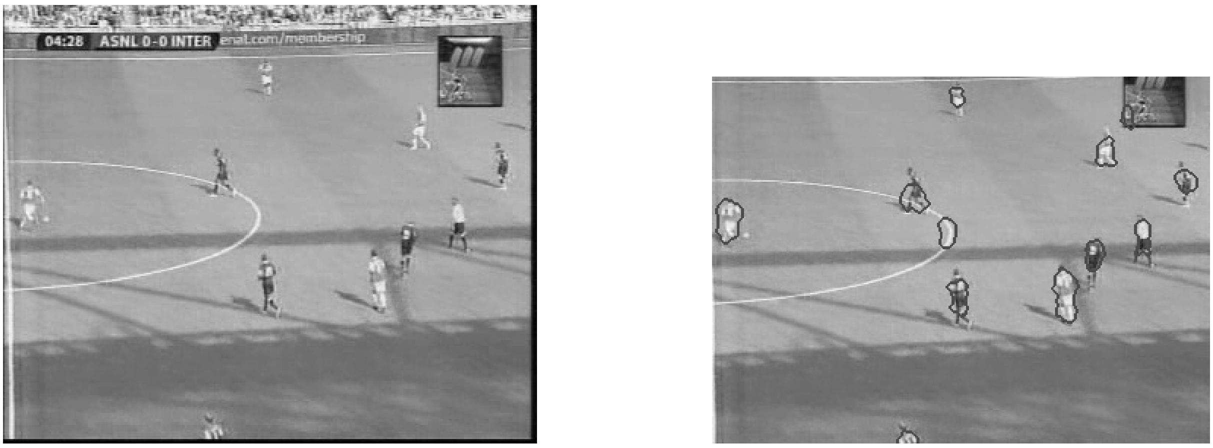
The performance of our proposed system is evaluated by conducting several experiments on different soccer broadcast video data. In this section, results for different part of our system are separately presented. The experiments are performed on three different conditions; dummy lighting condition, natural uniformed lighting condition and natural non-uniformed condition.

Table 1
Value of parameters used in our experiments

Parameter	Value
The threshold value on the edge image (line segmentation)	0.6*Maximum Value of the edge image
The threshold value on grayscale intensity (line segmentation)	110
α	0.85
The threshold value on pixels' player classification (player local classification)	0.98
The threshold value on pixels' ball classification (ball local classification)	200

Table 2
Average value of measurements for player and line segmentation on 1000 nonconsecutive frames and 3 situations

Dummy light		Natural uniformed light		Natural non-uniformed light	
Precision	Recall	Precision	Recall	Precision	Recall
94.4%	91.5%	100%	93.3%	96.5%	97%

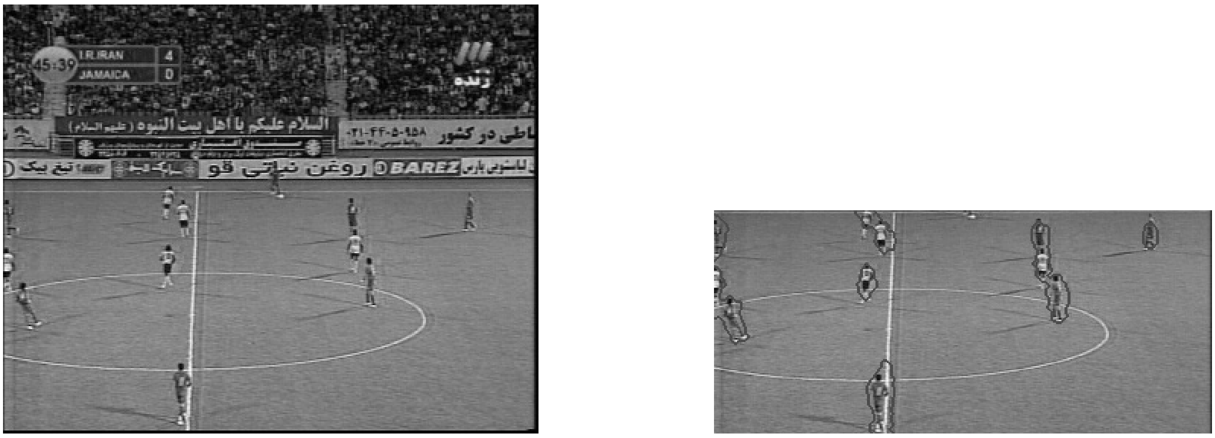


(a) (b)

Fig. 9. An image with severe non-uniform natural lighting condition. (a) Original image. (b) Obtained image by applying Bi-level image of players on original image.

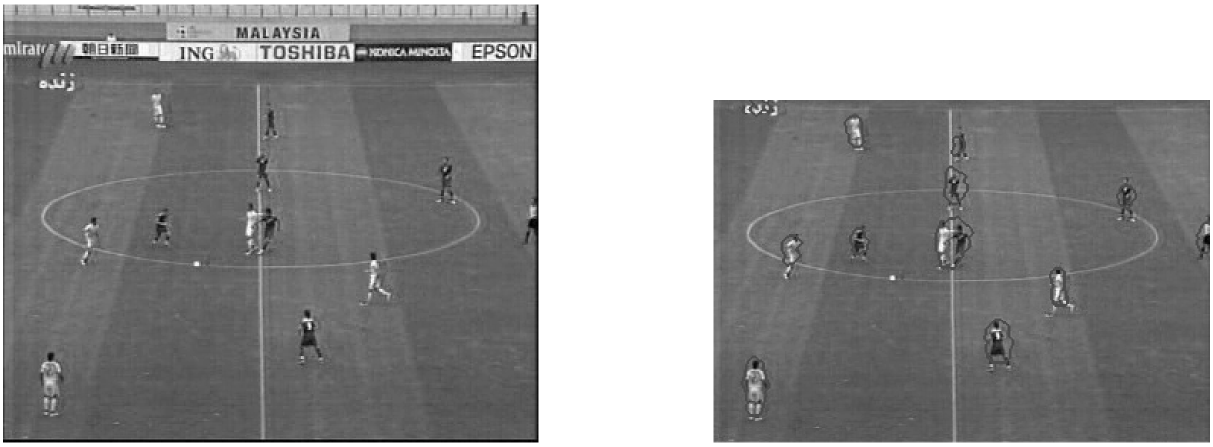
6.1. Segmentation

The segmentation stage of our proposed approach has been tested using soccer video broadcasts under variety of situations. In Figs 9, 10 and 11, the final results of segmentation under three different conditions are presented. These conditions are natural and uniform lighting condition, natural and non-uniform lighting and the presence of shadow due to building and artificial lighting. The performance of the proposed approach is evaluated using 1000 nonconsecutive frames of 3 soccer videos broadcast. Two measures are calculated for each frame. The mean values of these measures are presented in Table 1. The values of the parameters which are used in experiments are given in Table 2.



(a) (b)

Fig. 10. An image with artificial lighting condition. (a) Original image notice the shadows created by the players. (b) Obtained image by applying Bi-level image of players on original image.



(a) (b)

Fig. 11. An image with severe uniform natural lighting condition. (a) Original image. (b) Obtained image by applying Bi-level image of players on original image.

6.2. Uniform color model extraction

The soccer video with artificial lighting condition has been used for evaluating the performance of uniform color model extraction. The uniform color models of two teams are automatically obtained as well as, they are extracted manually. The quality of the uniform color model obtained automatically is evaluated by calculating the similarity measure between manual extraction model and automatic extraction model. The player classification using automatic extraction model is presented in Fig. 12. Also, the calculated similarity measure between two models is presented in Table 3.

Table 3
The similarity measure between two models

The Bhattacharyya similarity coefficients for the automatic extracted uniform color model of team A and team B	The Bhattacharyya similarity coefficients for team B	The Bhattacharyya similarity coefficients for team A
0.6	**0.9**	**0.9**

(a) (b)

Fig. 12. The players classification by using the uniform color model extracted automatically from 600 consecutive frames. (a) The players of team A. (b) The players of team B.

(a) (b) (c)

(d) (e) (f)

Fig. 13. The result of tracking the players on dummy light condition a. frame 1 b. frame 10 c. frame 31 d. frame 50 e. frame 67 f. frame 100.

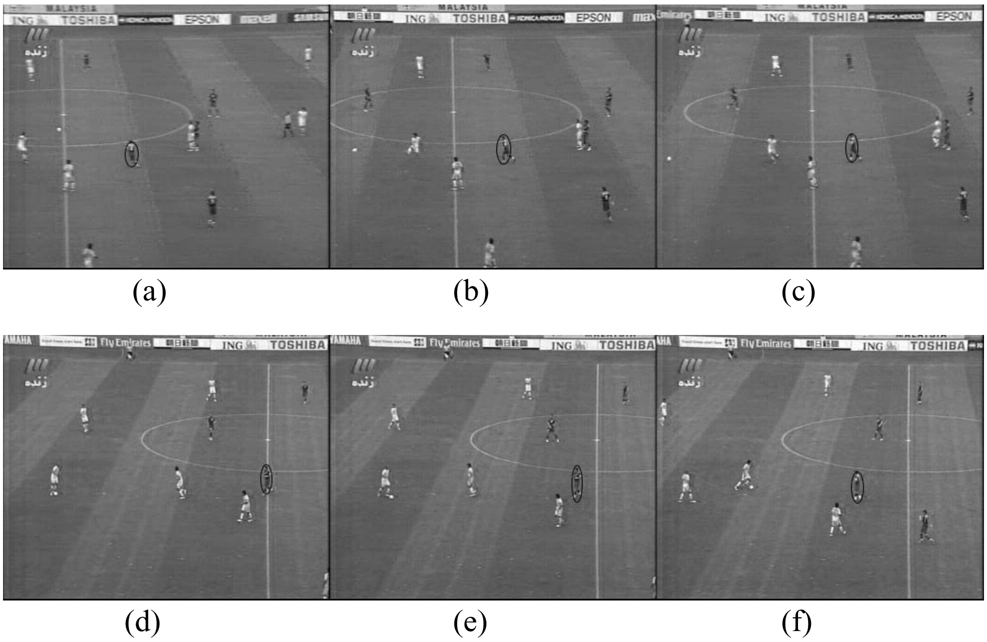


Fig. 14. The result of tracking the players on uniform natural light condition a. frame 1 b. frame 13 c. frame 27 d. frame 50 e. frame 67 f. frame 100.

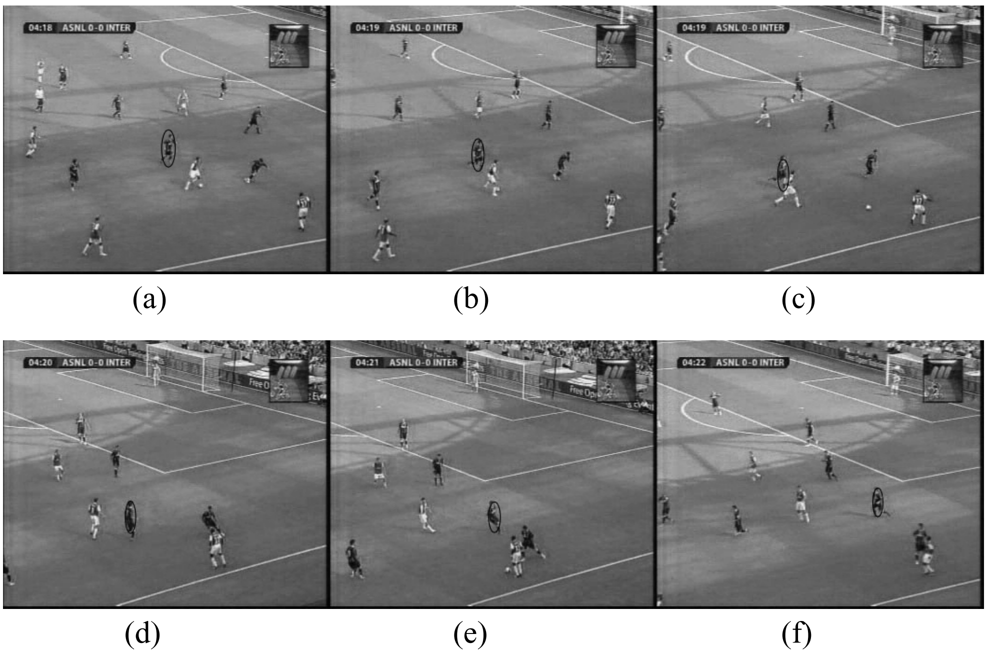


Fig. 15. The result of tracking the players on sever nonuniform natural light condition a. frame 1 b. frame 13 c. frame 20 d. frame 50 e. frame 67 f. frame 100.

Table 4
Four obtained metrics and statistics comparison for the performance evaluation the proposed tracking technique

Statistics	artificial Lighting Condition	Uniform Natural Lighting Condition	Non-Uniform natural Lighting Condition
The mean of the distances between positions of player on the automatic extracted route and manual extracted route	7.76	20.3	3.24
The mean of the distances between positions of player on the automatic extracted route and manual extracted route by spatially separated	7.76	20.3	3.24
The mean of the distances between positions of player on the automatic extracted route and manual extracted route by temporally separated	5.3	12.4	3.23
The mean of the distances between positions of player on the automatic extracted route and manual extracted route by temporally and spatially separated	5.3	12.4	3.23

6.3. *Blind tracking of players*

In this section, the results of the blind tracking of players are presented under three different lighting conditions. Another characteristics, that the tracking sequence has is that the players are partially occluded by other players. The results of tracking of the players are presented in Figs 13, 14 and 15 for 100 consecutive frames. In these figures, the specified player is encircled and tracking results are shown for six different frames.

6.4. *Specified player tracking*

The results of our tracking algorithm for a specified player are presented in this section. The performance of the proposed tracking approach is measured by four statistics that are presented in Table 4. The route extracted for a specified player is presented in Figs 16, 17 and 18 for 1000 consecutive frames under three different lighting conditions. For evaluating the performance of proposed approach, the specified player route is extracted both manually and automatically. Then, the four metrics are applied and comparisons are made for both extracted trajectories. The results of our comparison are presented in Table 4.

6.5. *Extraction of the data of the specified player trajectory*

Data has been extracted for the specified players which have been tracked using the algorithm of Section 6-4 are presented in Table 5 and Figs 19–21. In Figs 19–21 the presence percentile of a specified player which are determined in 12 zones of a soccer field under three different lighting conditions are presented. Qualification evaluation of the results is performed by manually extracting the data of the specified players and compared them with the automatic extraction.

7. Conclusion

In this paper, a system is introduced in which the information of each players in a soccer match video broadcast could be extracted. In the proposed system, the input video images are processed, and all players in the field are detected and tracked separately. Tracking of players provide some useful information such as the average velocity and the distance traversed by specified player. In order to gather information from a video soccer game, players are first detected from non-players. Then players are

Table 5
The information is extracted for the specified players on three verity light conditions

The data extracted	Artificial lighting condition (1000 frame)		Uniform natural lighting condition (600 frame)		Non-uniform natural lighting (500 frame)	
	Automatic	manual	Automatic	manual	Automatic	manual
the specified player average velocity(km/h)	23.2	22.6	46.4	43.2	37.8	34.9
the traversed distance by the specified player(m)	217	214	258	240	207	194

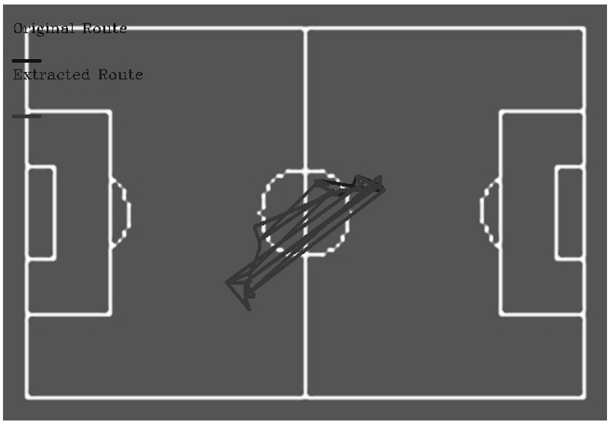


Fig. 16. The specified player route is extracted on 1000 consecutive frames on the dummy light condition.

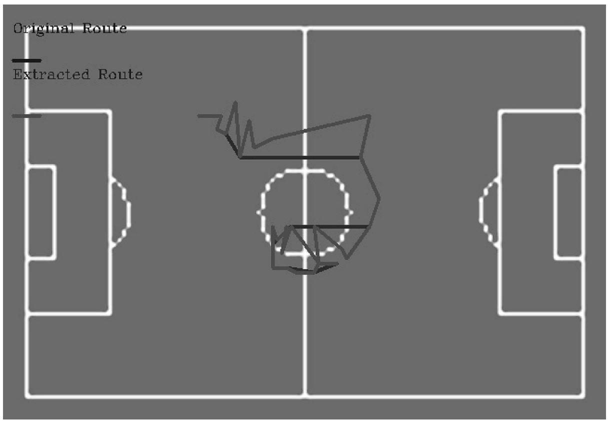


Fig. 17. The specified player route is extracted on 600 consecutive frames on the uniform natural light condition.

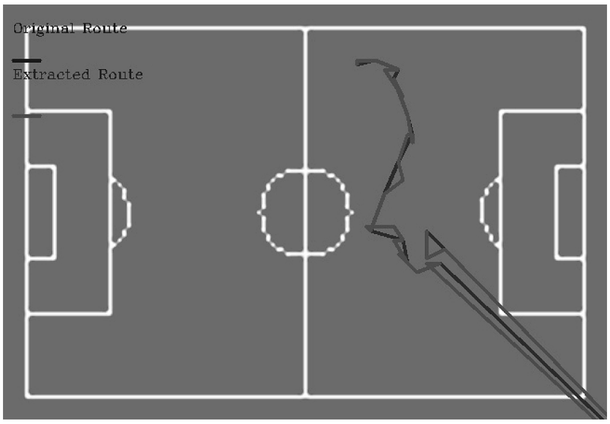


Fig. 18. The specified player route is extracted on 500 consecutive frames on the nonuniform natural light condition.

(a)

(b)

Fig. 19. The specified players existence percent on soccer field zones for dummy light condition. a. The automatic extracted results b. The manual extracted results.

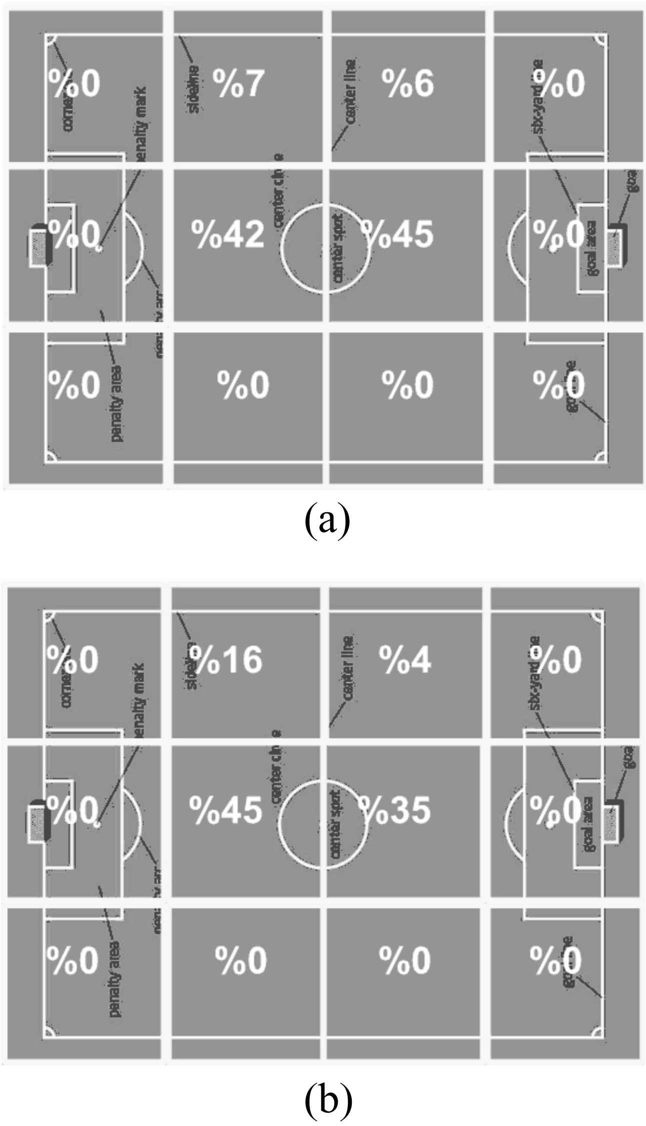
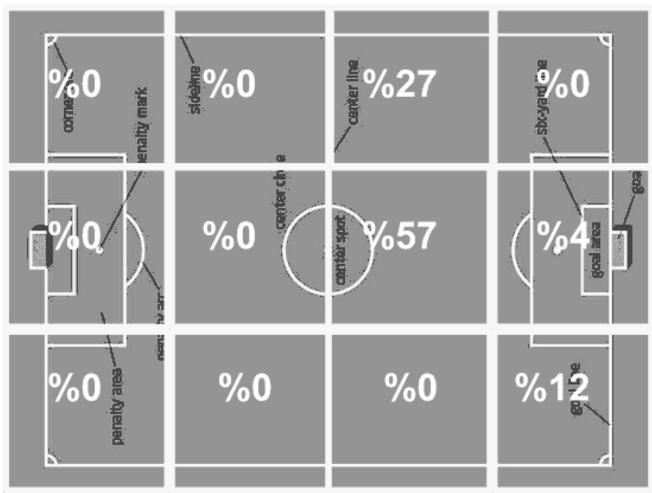


Fig. 20. The specified players existence percent on soccer field zones for uniform natural light condition. a. The automatic extracted results b. The manual extracted results.

tracked and necessary parameters are evaluated. Performance of each soccer player is evaluated based on extracted information, i.e. velocity and traversed distance, etc. The extracted information is useful for soccer analyzers and coaches. In our proposed system, problem that is usually categorized in image processing and machine vision system is solved by data mining prospect.

Experimental results of our system are presented in section 6 and its capability for extracting useful information from the soccer video broadcast is demonstrated. Our approach is robust under various conditions such as players occluded with lines, player's shadows, and severe non-uniform field's luminance. Our proposed approach yield accurate results under different situations without employing any special offline modeling of objects of interest.

(a)



(b)

Fig. 21. The specified players existence percent on soccer field zones for non-uniform natural light condition. a. The automatic extracted results b. The manual extracted results.

References

[1] A. Ekin, Sports Video Processing for Description, Summarization, and Search, *Doctor of Philosophy, Department of Electrical and Computer Engineering The College School of Engineering and Applied Sciences*, University of Rochester, 2003.
[2] A.J. Viterbi, A Personal History of the Viterbi Algorithm, *IEEE Signal Processing Magazine [120]*, (July 2006).
[3] A. Yilmaz, O. Javed and M. Shah, Object Tracking: A Survey, ACM Comput Surv 38, 4, Article 13 (Dec. 2006), 45 pages.
[4] C. Kang, J. Hwang and K. Li, Trajectory Analysis for Soccer Players, *International Conference on Data Mining – Workshops* (*ICDMW'06*), IEEE, 2006.

[5] D.A. Sadlier and N.E. O'Connor, Event detection in field sports video using audio-visual features and a support vector Machine, *IEEE Transactions on Circuits and Systems for Video Technology* **15**(10) (Oct 2005), 1225–1233.
[6] D. Comaniciu, V. Ramesh and P. Meer, Kernel-Based Object Tracking, *IEEE Transactions on Pattern Analysis and Machine Intelligent* **25**(5) (May 2003).
[7] D. Comaniciu, V. Ramesh and P. Meer, Real-Time Tracking of Non-Rigid Objects using Mean Shift, *IEEE CVPR 2000*.
[8] D. Tweed and A. Calway, Tracking Many Objects Using Subordinated CONDENSATION, *BMVC, 2002*.
[9] E.A. Murat Tekalp and R. Mehrotra, Automatic soccer video analysis and summarization, *IEEE Transactions on Image Process* **12**(7) (July 2003).
[10] E. Pazouki, Analysis and Evaluation of Technical Performance of a Specific Player during a Soccer game based Video Sequence, MS thesis, Amirkabir University of Technology, Computer Engineering Department 2008.
[11] G. Welch and G. Bishop, An Introduction to the Kalman Filter, *Department of Computer Science*, University of North Carolina at Chapel Hill.
[12] H. Lu, B.C. Ooi, H.T. Shen and X. Xue, Hierarchical Indexing Structure for Efficient Similarity Search in Video Retrieval, *IEEE Transactions on Knowledge and Data Engineering* **18**(11) (Nov 2006), 1544–1559.
[13] I.F. A. Board, Laws of the Game, *Fédération Internationale de Football Association 11 Hitzigweg*, 8030 Zurich, Switzerland.
[14] J. Needham, Tracking and Modeling of Team Game Interactions, *degree of Doctor of Philosophy*, School of Computing, The University of Leeds, 2003.
[15] J. Oh, J. Lee and S. Hwang, Video Data Mining: Current Status and Challenges. Encyclopedia of Data Warehousing and Mining, (A book edited by Dr. John Wang). Idea Group Inc. and IRM Press. 2005.
[16] J. Pers and S. Kovacic, Computer Vision System for Tracking Players in Sports Games, *First Int'l Workshop on Image and Signal Processing and Analysis*, June 14–15, 2000, Pula, Croatia.
[17] K. Okuma, A. Taleghani, N.D. Freitas, J.J. Little and David G. Lowe, A Boosted Particle Filter: Multitarget Detection and Tracking, *University of British Columbia*.
[18] M. Beetz, J. Bandouch and S. Gedikli, Camerabased Observation of Football Games for Analyzing Multiagent Activities, ACM 2006 159593-303-4/06/0005.
[19] M. Isard and A. Blake, CONDENSATION – Conditional Density Propagation for Visual Tracking, *Int J Computer Vision* (1998), in press.
[20] N.R. Pal, K. Pal and J.C. Bezdek, Amixed c-means clustering model, in IEEE Int Conf. Fuzzy Systems, Spain, 1997, pp. 11–21.
[21] N. Vandenbroucke, L. Macaire and J. Postaire, Color image segmentation by pixel classification in an adapted hybrid color space, Application to soccer image analysis, *Computer Vision and Image Understanding* **90** (2003), 190–216.
[22] O. UTSUMI, K. MIURA, I. IDE, S. SAKAI and H. TANAKA, an object detection method for describing soccer games from video, *IEEE*, (2002).
[23] P.J. Figueroa, N.J. Leite and R.M.L. Barros, Tracking soccer players aiming their kinematical motion analysis, *Computer Vision and Image Understanding* **101** (2006), 122–135.
[24] S. Choi, Y. Seo, H. Kim and K.S. Hong, Where are the ball and players? Soccer game analysis with colorbased tracking and image mosaick, In Proc. Int. Conf. on Image Analysis and Processing, 1997, pp. 83–91.
[25] S.H. Khatoonabadi and M. Rahmati, Automatic soccer players tracking in goal scenes by camera motion elimination, Image and Vision Computing, In Press, Corrected Proof, Available online 10 July 2008.
[26] S. Iwase and H. Saito, Parallel Tracking of All Soccer Players by Integrating Detected Positions in Multiple View Images, 17th International Conference on Pattern Recognition (ICPR'04) 1051-4651/04.
[27] S.S. Intille and A.F. Bobick, Tracking Using a Local Closed-World Assumption: Tracking in the Football Domain, Master of Science in Media Arts and Sciences, School of Architecture and Planning, Submitted to the Program in Media Arts and Sciences, on August 5, 1994.
[28] W.C. Naidoo and J.R. Tapamo, Soccer Video Analysis by Ball, Player and Referee Tracking, SAICSIT 2006, Pages 51–60.
[29] Y. Gong, T.S. Lim and H.C. Chua, Automatic Parsing of TV Soccer Programs, *IEEE International Conference on Multimedia Computing and Systems*, May, 1995.

Copyright of Intelligent Data Analysis is the property of IOS Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.