❏   979

Goal Detection in Soccer Video: Role-Based Events Detection Approach

Farshad Bayat*, Mohammad Shahram Moin*, Farhad Bayat**

* Departement of Electrical and Computer Engineering, Islamic Azad University, Qazvin Branch, Qazvin, Iran
** Departement of Electrical Engineering, University of Zanjan, Zanjan, Iran.

Article Info	ABSTRACT
	Soccer video processing and analysis to find critical events such as occurrences of goal event have been one of the important issues and topics of active researches in recent years. In this paper, a new role-based framework is proposed for goal event detection in which the semantic structure of soccer game is used. Usually after a goal scene, the audiences' and reporters' sound intensity is increased, ball is sent back to the center and the camera may: zoom on Player, show audiences' delighting, repeat the goal scene or display a combination of them. Thus, the occurrence of goal event will be detectable by analysis of sequences of above roles. The proposed framework in this paper consists of four main procedures: 1- detection of game's critical events by using audio channel, 2- detection of shot boundary and shots classification, 3- selection of candidate events according to the type of shot and existence of goalmouth in the shot, 4- detection of restarting the game from the center of the field. A new method for shot classification is also presented in this framework. Finally, by applying the proposed method it was shown that the goal events detection has a good accuracy and the percentage of detection failure is also very low.

Corresponding Author:

Farshad Bayat,
Department of Electrical and Computer Engineering,
Islamic Azad University, Qazvin Branch,
Nokhbegan Blvd. Qazvin. Iran.
Email: farshad.bayat@qiau.ac.ir

1. INTRODUCTION

In recent years, the development of multimedia applications in different fields, increase of the internet impact and development of communications, led to the development of multimedia storage technology. Therefore, producing information management and retrieval tools for handling these large data became necessary. In this regard, sport competitions are one of the most popular multimedia and on top of them, watching soccer is the most exciting and popular sport videos. Thus, high volumes of soccer videos are produced and this popularity has led researchers to be interested in creating new ideas and ways to manage and retrieve information from soccer videos [1, 2 and 3]. Soccer video processing and information retrieval techniques are divided into two online and offline general structures. Online techniques extract the information of a live game and thereby enhance visitors' information of a game being displayed. For example, the number of team's offside, the distance traveled by a player, the time of ball possession by teams. In [4] some cameras with high frame rate (e.g.200fsp) were placed in the stadium and used to find online goal. The purpose of using cameras with high frame rate is that in those shots that goal event occurs the ball speed is often very high; therefore, using the video obtained from conventional systems with about 30 frames per second recording speed, makes detection of goal occurrence very difficult and perhaps impossible. In [4] two techniques have been used for detecting goal event. In the first method, the algorithm of circle detection (circular Hough transform algorithm) is used to find candidates. In the second technique,

the produced candidates are evaluated and compared using a supervised neural network, and the proper candidate is selected. By tracking the position of the ball, the occurrence of goal event is determined. Since watching the whole soccer game video is time consuming so, offline information retrieval and management tools help soccer enthusiasts and researchers in this field. The offline extraction and analysis tools allow us to extract the necessary information from the games in the lowest time. An example of such important information is detection of goal events in soccer games. Utilizing these management and information retrieval tools summarize the video game and result in saving the soccer experts' time when analyzing the tactics of the game. Goal detection based on cinematic features is presented in [5]. This paper points out that just using video processing techniques for goal detection is difficult. Therefore, they have used the advantage of the natural events that occur after a goal event. The authors have proposed following procedures for goal detection: considering the interruptions during the game, considering the existence of at least one shot out of the field (e.g. audience shot) or a close-up shot, considering the presence of at least one replay shot. In [6, 7] the scoreboard information is used to find a goal occurrence in soccer video. In this method, goal occurrences are determined by detection of scoreboard location during the video display and tracking changes in it. A disadvantage of this method is that the texture of most of the scoreboards is homogeneous with the texture of the audience, which leads to difficulty in scoreboard detection and match result finding. Another disadvantage of this method is the occurrence of image interference on the location of scoreboard (e.g. the TV channels' logos, advertising etc.), that make the detection of changes on the scoreboard difficult.Using multi-clues is another strategy to find the goal occurrence [8]. In this method the candidate shots are chosen by using the clues such as increase of sound intensity and occurrence of goal bar in the shot. Next, the goal events are determined by reviewing the next shots and identifying the visual clues of goal scoring such as the audiences' shot and replay shot. This paper provides a framework to detect the goal events in a soccer game video. For this purpose, we extract video events using both visual and audio channels. This framework is composed of several main stages that in each stage we relate a number of features to higher semantic concepts. In the first step we use basic features of sound and determine the important events of video. In the next stage, we use dissimilarity features for shot boundary detection and then classify these shots with basic visual features and in the next stage we find the moments when the audiences' shot and close shot exist which is the indicator of important scenes of video. In the final step we use the created meaningful units and detect goalmouth and game start framework. By applying these steps we can find goal events and summarize the video. This framework is illustrated in Figure 1.
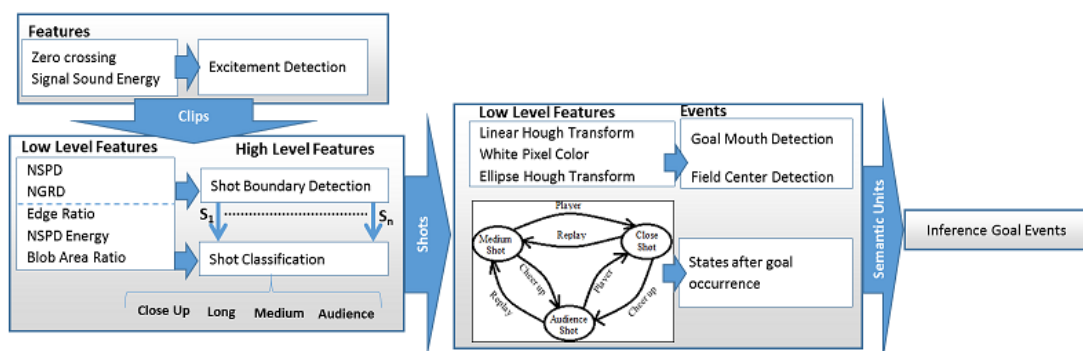

Figure 1. Block diagram of the proposed method.

2. EXTRACTING THE EXISTING EVENTS FROM SOCCER VIDEO

Undoubtedly the most popular sport in the world is soccer. The main idea is that as an important event occurs in a soccer game, the audiences' and reporters' sound intensity increases. Thus, the audio channel has high semantic information and this feature can be widely used for detection of important events. Using this feature leads to increase of detection accuracy and naturally results in deletion of unimportant parts and decrease of processing complexity of calculations.In most of the existing important events extraction approaches, the basic audio signal energy level feature is used. In [9], the above feature along with the detection of some keywords including the words such as "goal" and "penalty" is used for finding important events of soccer game. Sometimes in the game, the audiences loudly and continuously encourage their favorite team, so in these cases the audio signal energy level feature won't have the required accuracy for the detection of important events. In [10], in addition to the audio signal energy level feature, the basic

feature of zero crossing rate is used to enhance the accuracy of important events detection and the results show that the accuracy has increased to a large extent. In this paper, we have used the proposed method in [10]; and after identifying the exciting parts of the match, we segment the video in the periods of 50 seconds before and 85 seconds after the detected points, in order to perform subsequent processing. It is worth noting that the segmentation length has been obtained after repeated experiments and simulations.

3. SHOT BOUNDARY DETECTION

The most important step in the series of videos processing and retrieving is shot boundary detection. By shot boundary detection, the basic semantic structure of video will be ready for the next steps of video processing. Shot boundary is determined by significant changes in the composition of color or pixel location [11]. The dominant color in the soccer video is green; however its amount is changed in some parts of the video. For example, the amount of green color where the camera is located in the middle of the field or far view is significantly different from close or audience shot. These changes can be used as a measure of dissimilarity to separate the shot boundaries. Another dissimilarity criterion that is also used in this paper is the difference in the color intensity.

3.1. The Feature of Green Color Ratio Difference

In this paper we have used the feature of green color ratio difference that is specified as follows.

$$GRD(i) = |G_r(i + 1) - G_r(i)| \tag{1}$$

$$G_r(i) = \frac{Pg\ (i)}{M \times N} \tag{2}$$

In this formula, $M \times N$ shows the total number of pixels and Pg (i) shows the number of green pixels in the ith frame that is characterized as follows:

$$Pg\ (i) = \sum_{y=1}^{M} \sum_{x=1}^{N} G_{(x,y)}(i) \tag{3}$$

$$G_{(x,y)}(i) = \begin{cases} 1, & T < ED_{(x,y)}(i) \\ 0, & O.W \end{cases} \tag{4}$$

$$\text{Where} ED_{(x,y)}(i) = \sqrt{\left(1 - P(x,y)_g\right)^2 + P(x,y)_r^{\ 2} + P(x,y)_b^{\ 2}}.$$

The P (x, y) variable represents the values of RGB color in the position of (x, y) in the ith frame. The $ED_{(x,y)}(i)$ variable shows the Euclidean distance between the colors of the pixel (x, y) in the ith frame relative to the pure green color. Certainly, the choice of the threshold value in different periods and lighting conditions of stadium would be different. To find the appropriate threshold value, we use the following four steps:

1. Select number of frames from video randomly and call these frames F_i.
2. For each frame, calculate the Euclidean distance relative to green color and arrange them in descending order.
3. To find the appropriate threshold value, P frames are selected from the beginning of the arranged set and from each F_i؛ $1 < i < P$; calculate Euclidean distance for each pixel, and then arrange these values in ascending order.
4. Remove $\frac{3}{5}$ of the end elements and $\frac{1}{5}$ of the first elements of the set and average the rest of the values of the set and consider it as the threshold value.

Removal of the $\frac{3}{5}$ of the end values of the set leads to deletion of inaccurate information and also removal of the $\frac{1}{5}$ of the data from the beginning of the set makes the Euclidean distance from the pure green greater and provides conditions to choose darker colors. Threshold value is usually in the [0.6, 0.8] range.

The criterion of green color ratio changes, determines the proper boundary for the shot from the shot boundary in the frames that green color combinations change dramatically. But in this criterion, the spatial information of pixels is not considered. To boost the detection accuracy of shot boundary, we use a criterion that considers the spatial information of pixels.

3.2. The Feature of Color Intensity Difference

If $I^c_{(x,y)}(i)$ is cth color channel intensity ($c \in \{Red, Green, Blue\}$ in (x, y) position and ith frame, criterion of dissimilarity, difference of spatial intensity of colors, is defined as follows:

$$SPD(i) = Min(FD(i, \alpha, \beta)) \quad \forall \alpha, \beta \in \{-1, 0, 1\} \tag{5}$$

$$FD(i, \alpha, \beta) = \frac{1}{KMN} \left(\sum_{c=1}^{K} \sum_{y=Max(1,1-\beta)}^{Min(M,M-\beta)} \sum_{x=Max(1,1-\alpha)}^{Min(N,N-\alpha)} \left| I^c_{(x+\alpha,y+\beta)}(i) - I^c_{(x,y)}(i+1) \right|^q \right)^{1/q}, q \in [1, \infty) \tag{6}$$

In which q defines the type of distance, so that q=1 and q=2 that shows the city distance and Euclidean distance, respectively. Parameters of α and β determine the sliding amount of the first frame over the second frame in the x and y axes, respectively. The $SPD(i)$ equation determines the minimum distance between two consecutive frames in the 9 directions specified in Figure 2.
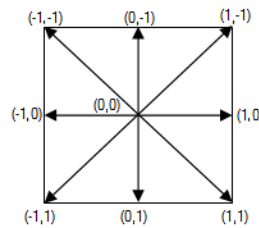

Figure 2. Different values of α and β in 9 directions

The GRD feature produces values in [0, 1] range and SPD feature produces values in [0, Imax] range that Imax is the maximum lighting intensity in the red and blue channels. For ease of work and classification process, it is better to normalize the values in the range of zero and one and we should notice that the GRD values are usually in the range of [0, 0.86] before normalization. Following formulas can be used for normalization:

$$NSPD(i) = \frac{SPD(i)}{Max(SPD)} \tag{7}$$

$$NGRD(i) = \frac{GRD(i)}{Max(GRD)} \tag{8}$$

In Figure 3-a you can see the normalized SPD change rate and in Figure 3-b the normalized GRD change rate in terms of the frame number. We use two NSPD and NGRD features to determine shot boundary. So, feature space is as follows:

$$Feature = \{(NSPD(i), NGRD(i)) \mid 0 \leq i \leq T\} \tag{9}$$

The best candidate for shot boundary is the frames in which the values of NSPD or NGRD features or both values are in local maximum. In order to find the local maximum, the value of feature should be maximum in comparison with it's before and after. Therefore, we use following equations in order to find the difference between values:

$$PDiff(F, i) = F(i) - F(i - 1) \tag{10}$$

$$NDiff(F, i) = F(i) - F(i + 1) \tag{11}$$

In the above equation F are the same features that are specified in equation (9). In order to find local maximum, we use the following equation:

$$LM(F,i) = \begin{cases} Max(PDiff(F,i), NDiff(F,i)), \\ \qquad PDiff > Norm(F) \, and \\ \qquad NDiff > Norm(F) \\ \\ 0, \qquad o.w \end{cases}$$

(12)

$$Norm(F) = \begin{cases} 0.01 \qquad\qquad if\, fisNGRD \\ Mean(NSPD) if\, fisNSPD \end{cases}$$

The $Norm(F)$ shows the difference compared to neighbors whose values are selected empirically.
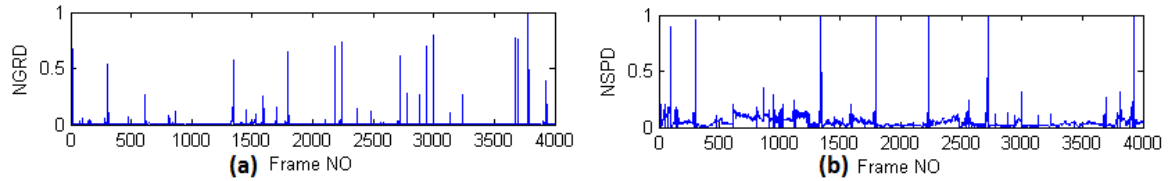

Figure 3. Frame-based features diagram: (a) The feature of green color ratio changes in terms of frame number, (b) The feature of spatial changes of pixels' color intensity in terms of frame number

And given that the shot length is at least 40 frames, we use windowing operator and for each window select a local maximum. To do this, the following equation is used to identify the shot boundary:

$$SB(F,i) = \begin{cases} 0 & \exists\, j \in [i+1, i+l] \mid LM(F,j) > LM(F,i) \\ 1, & o.w. \end{cases}$$

(13)

In which l indicates the length of the window, here we have considered 40 frames.

When the SB value in ith frame equals 1, the frame will be the shot boundary, indeed in this frame the values of features are local maximum. In Figure 4.a shot boundary extraction based on NSPD feature and in Figure 4.b the shot boundary extraction by NGRD is shown.

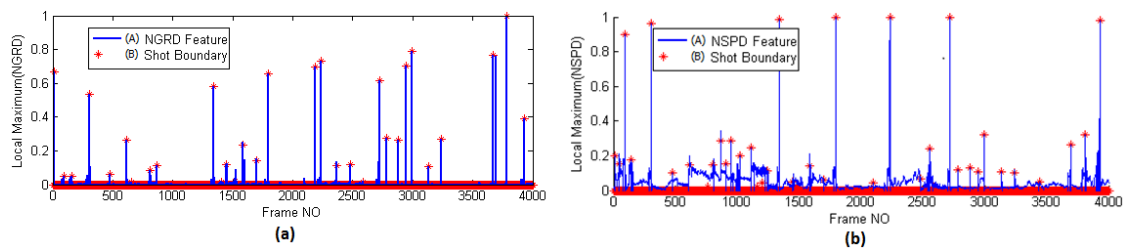Figure 4. (a) Shot boundary extraction by NSPD feature (A) NSPD feature curve (B) Extracted boundaries by local maximum. (b) Shotboundary extraction by NGRD feature (A) NGRD feature curve (B) Extracted boundaries by local maximum.

In order to find the final shot boundaries, we use the combination of shot boundaries found based on NSPD and NGRD features, as follows:

$$FSB(i) = Max(SB(NSPD, i), SB(NGRD, i))$$

(14)

We have used windowing technique in order to decrease error in finding the shot boundary. Given that the shot length is at least 40 frames, we can use a window with a length of 40 frames; and we select the main shot boundary from among the set of shot boundaries with the most SPD in each window. Figure 5 shows the result of combining the features.
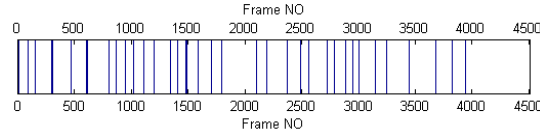
Figure 5. Combination of shot boundaries that are specified by NSPD and NGRD features.

4. CLASSIFICATION OF SHOTS

There are generally four different shots in soccer games. These four shots are shown in Figure 6. The far shot is usually used to show the game situation in which the number of players and great amount of playing field's grass is in the image. The middle shot is used to show the movements of the players and the camera zooms to the position of the player who has the ball so that the number of players with relatively large size is included in the image. Close shot usually occurs when stop has occurred during the game which is usually associated with important events of the play; and in this mode, the camera will zoom on a player to the full extent so that the majority of the image is related to the player. Audience shot is related to the view that camera images the audience present in the stadium that is usually after the goal occurrence. After identifying the shot boundaries, a key frame is extracted from each shot and the algorithm which is shown in Figure 7 is performed as a hierarchy. The features of color ratio, holes' area ratio and edge ratio are used for shot classification.

Figure 6. Shot Kinds, respectively (A) Far shot (B) Middle Shot (C) Close Shot (D) Audience Shot.

To calculate holes' area ratio, first we extract grass area from the main frame and detect the blobs in the grass area and calculate the blob area ratio as follows:

$$HoleRatio(i) = \frac{\sum_{b=1}^{C} Area(Hole^i(b))}{M \times N} \tag{15}$$

In which $Hole^i$ is the set of holes in the ith frame and C shows the number of holes.
To calculate the edge ratio we canny edge detection on gray level frame.

$$EdgeRatio(i) = \frac{Cany(i)}{M \times N} \tag{16}$$

We use the following equation to calculate Energy of NSPD in each shot:

$$NSPD_{Energy}(s, e) = \frac{\sum_{i=s}^{e} |NSDP(i)|^2}{e - s} \tag{17}$$

Parameters s and e specify the beginning and end of the shot, respectively.
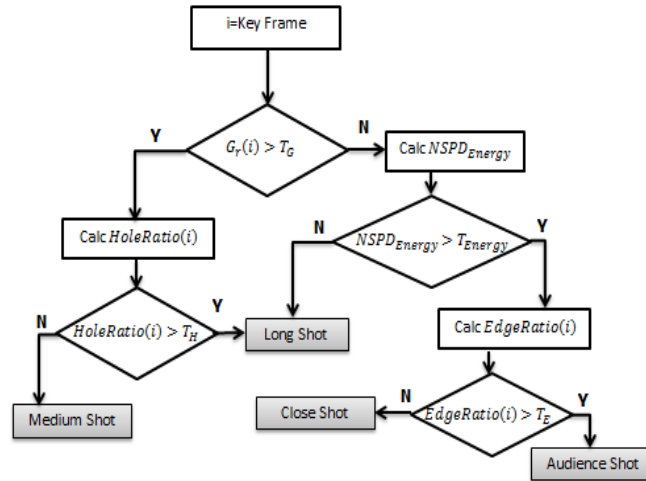
Figure 7. Diagram of Shot Classification

5.   DETECTION OF GOAL EVENT

We detect goal events using semantic units. These units are obtained hierarchically and by using image processing techniques. The high-level units are described below.

5.1 Detection of Goalmouth

This paper has used the presented ways in [12] for detection of goalmouth which is composed of the following three stages:

1. First, we apply top hat transformation (THT) on the frame. This transformation leads to increase of white color intensity and therefore the edges of goalmouth will be more specific and for a binary image, a self-adaptive threshold can be used.
2. The white vertical lines in the image are extracted as the proper candidate for the goalmouth as follows. V indicates vertical white lines.

$$PGP_l(x,y) = \begin{cases} 1 & ,\forall(x,y) \in VL_l(x) , if\, PP(x,y) = 1 \\ 0 & , \qquad\qquad O.W \end{cases} \tag{18}$$

$$PP(x,y) = \begin{cases} 1 & ,\exists\, z \in [y-10, y+10], \qquad \forall i \in [-10,+10], \\ & if\, Color(x, z+i) = White \\ 0 & , \qquad\qquad O.W \end{cases} \tag{19}$$

In the above equation, $VL_l(x)$ indicates the vertical lines.
3. Two lines that have the following terms are selected as the lines of goalmouth from among the vertical lines extracted in step 2.

$$rule = \begin{cases} 20 \le L_i \le 0.3 * H \\ 20 \le D \le 0.4 * W \\ 0.5 \le \dfrac{D}{Max(L_1, L_2)} \le 20 \\ 0 \le |(T_i - B_i) - L_i| \le 20 \\ 1 \le O_v \le 0.3 * H \\ 0.5 * H \le Y_c \le H \end{cases} \quad i = 1, 2 \tag{20}$$

H and W, respectively, determine the width and height pixels and Li shows the amount of pixels in each line, D shows the horizontal distance between the two lines, B_i and T_i, respectively, presents the highest and lowest position of line, and O_v represents the overlap of the two lines in the vertical direction. By forming the smallest rectangle by the two vertical lines, C (X_c, Y_c) is the center of the rectangle that Y_c represents the width of the center of the rectangle. The process of goalmouth detection is illustrated in Figure 8.

5.2 Detection of Center of the Field

Shots related to the center of the field are composed of two main components:
1. Existence of ellipse in the image
2. Ellipse cut by a line

There are different ways to survey ellipse such as Least-Squares Fitting (LSF) [13, 14], Invariant Pattern Filter (IPF) [15] and Ellipse Hough transform (EHT) [16]. In this paper we have used the Hough transform method, and have made binary image as follows, to obtain better results.

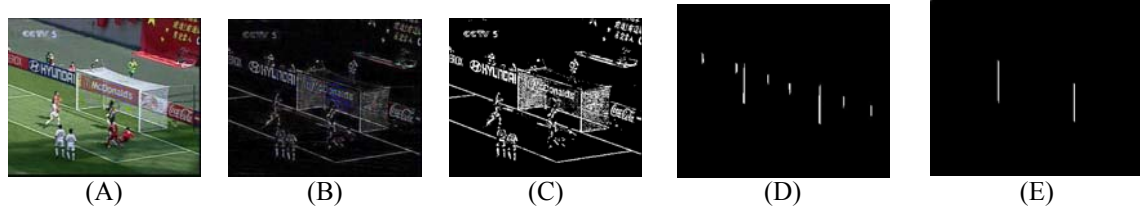| (A) | (B) | (C) | (D) | (E) |

Figure 8. Stages of goal bar extraction: (A) the main image, (B) application of EHT, (C) making image binary, (D) candidates of vertical lines, (E) goal bar [12].

First, we calculate the Euclidean distance of the corresponding frame relative to green color, then, apply canny edge detection algorithm on it.
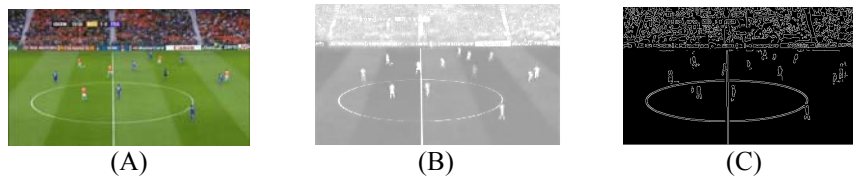

| (A) | (B) | (C) |

Figure 9. Process of preparing a binary image (A) the main image (B) Euclidean distance relative to pure green color (C) application of canny edge detection.

Eliminate parts with dense edge and apply EHT algorithm on it, after creating a binary image. This process is shown in Figure 9. Apply linear Hough transformations on the binary image created in last stage and find the lines' candidates. If there is a line in the candidates that its angle with the horizontal angle of the ellipse is about 90, and the distance between the center of the ellipse and the perpendicular line to the candidate line is low, the shot is considered as the game start. When a goal occurs, game will be interrupted for a few minutes. This interruption is used to show the happiness of audiences and replay of goal event. Thus, after a goal shot there is usually a close-up shot, audience shot and replay shot. The states after goal occurrence are illustrated in Figure 10. To detect goal event, we find the parts that at least one of the above modes has been in the shot change. In the case the goalmouth is detected in the middle shot and game start is detected after modes in Figure 11, we denote that middle shot as the goal event.

6. EXPERIMENTAL RESULTS

We have used several soccer games' videos available on internet to test this approach. These videos are recorded from TV channel and with the rate of 25 frames per second. The experiment is composed of three parts. In the first part, we have evaluated detection of shot boundary. We have separated video pieces with approximately 3,375 frames manually and have compared with approaches presented in the paper. The results are shown in Table 1.
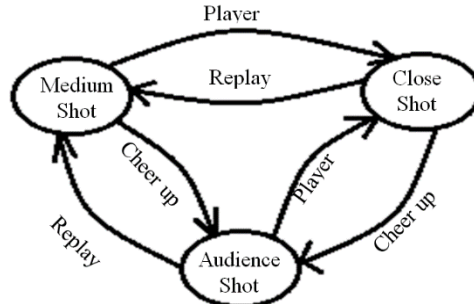
Figure 11. Different shot type states after goal occurrence.

The evaluation criteria are defined as follows:

$$Recall = \frac{N_c}{N_c + N_m} \ , \quad Precision = \frac{N_c}{N_c + N_f} \tag{21}$$

In which N_c shows the number of correct cases, N_m the number of missed cases and N_f indicates the number of false cases. As you can see, the recall criterion and precision criterion are acceptable. The results of evaluating shots' classification into four classes of far view, middle view, close and audience view are presented in Table 2. The results are desirable and satisfactory. The final results of goal detection are shown in Table 3. As it can be seen, the precision criterion is almost 92%, and the recall measure is 78%.

Table 1. Evaluation of shot boundary detection.

Match NO	Clips	N_c	N_m	N_f	Recall	Precision
1	21	287	17	31	0.94	0.90
2	19	215	12	26	0.94	0.89
3	17	183	14	26	0.92	0.87
4	16	167	21	10	0.89	0.94
5	16	161	15	19	0.91	0.89
6	15	143	14	11	0.91	0.93

Table 2. Evaluation of shot classification

Shot Type	N_c	N_m	N_f	Recall	Precision
Long	350	59	61	0.86	0.82
Medium	242	36	34	0.87	0.88
Close	198	21	29	0.90	0.87
Audience	84	11	5	0.88	0.94

Table 3. Evaluation of goal detection

Match NO	1	2	3	4	5	6
Total Goal	7	3	5	3	4	2
Correct	6	3	4	3	2	1
Miss	1	0	1	0	2	1
False	1	0	0	0	1	0
Recall	0.86	1	0.8	1	0.5	0.5
Precision	0.86	1	1	1	0.67	1

7. CONCLUSION

A new approach for shot boundary separating and classifying was presented in this paper. We identify the goal shots based on high-class semantic units extracted from video. As it is shown in the result evaluation section, the algorithm based on two features of dissimilarity of green color ratio changes and color intensity changes and with the help of self-adaptive threshold, the results of shot boundary have the adequate precision and recall. We can also detect goal events with proper precision by creating semantic units and by using semantic structure of the game. The goal of obtained results is accepted.

REFERENCES

[1] S.F. de Sousa Júnior, A. de A Araújo, D. Menotti, "An overview of automatic event detection in soccer matches", Applications of Computer Vision (WACV), 2011 IEEE Workshop on, pp. 31, 38, 5-7 Jan. 2011.
[2] T.T. Pham, T.T. Trinh, V.H. Vo, N.Q. Ly, D.A. Duong, "Event Retrieval in Soccer Video from Coarse to Fine Based on Multi-Modal Approach", Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF), 2010 IEEE RIVF International Conference on , pp. 1, 6, 1-4 Nov. 2010.

[3] J. Yu, Y. He, K. Sun, Z. Wang, X. Wu, "Semantic Analysis and Retrieval of Sports Video", Frontier of Computer Science and Technology, 2006. FCST '06. Japan-China Joint Workshop on, pp. 97, 108, Nov. 2006.

[4] T. D'Orazio, M. Leo, P. Spagnolo, M. Nitti, N. Mosca, A. Distante, "A visual system for real time detection of goal events during soccer matches", Computer Vision and Image Understanding, 113(5): 622–632, 2009.

[5] A. Ekin, A. M. Tekalp, R. Mehrotra, "Automatic soccer video analysis and summarization", IEEE Transactions on Image Processing, 12(7): 796–807, 2003.

[6] S. Yang, W. Xiangming, S. Yong, Z. Liang, Y. Lelin, L. Haitao, "A Scoreboard Based Method for Goal Events Detecting in Football Videos", Digital Media and Digital Content Management (DMDCM), 2011 Workshop on , pp. 248, 251, 15-16 May 2011.

[7] M. Hung, C. Hsieh, "Event Detection of Broadcast Baseball Videos", Circuits and Systems for Video Technology, IEEE Transactions on , vol. 18, no. 12, pp. 1713, 1726, Dec. 2008.

[8] P. Shi, Y. Xiao-qing, "Goal Event Detection in Soccer Videos Using Multi-Clues Detection Rules", Management and Service Science, 2009. MASS '09. International Conference on, vol., no., pp. 1, 4, 20-22 Sept. 2009.

[9] L.H. Yong, H. Tingting, "Integrating Multiple Feature Fusion for Semantic Event Detection in Soccer Video", Artificial Intelligence, 2009. JCAI '09. International Joint Conference on, pp. 128, 131, 25-26 April 2009.

[10] M.H. Kolekar, K. Palaniappan, "A hierarchical framework for semantic scene classification in soccer sports video", TENCON 2008 - 2008 IEEE Region 10 Conference, pp. 1, 6, 19-21 Nov. 2008.

[11] X. Gao, X. Tang, "Unsupervised video-shot segmentation and model-free anchorperson detection for news video story parsing", Circuits and Systems for Video Technology, IEEE Transactions on , vol. 12, no. 9, pp. 765, 776, Sep. 2002.

[12] Y. Yang, S. Lin, Y. Zhang, S. Tang, "Highlights extraction in soccer videos based on goal-mouth detection", Signal Processing and Its Applications, 2007. ISSPA 2007. 9th International Symposium on, pp. 1, 4, 12-15 Feb. 2007.

[13] A. Fitzgibbon, M. Pilu, R.B. Fisher, "Direct least square fitting of ellipses", Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol. 21, no. 5, pp. 476, 480, May 1999.

[14] W. Jianping, "Robust Real-Time Ellipse Detection by Direct Least-Square-Fitting", Computer Science and Software Engineering, 2008 International Conference on , vol. 1, pp. 923, 927, 12-14 Dec. 2008.

[15] T. J. Atherton, D. J. Kerbyson, "Size invariant circle detection", Image and Vision Computing, 17(11): 795–803, 1999.

[16] X. Yu, H.W. Leon., C. Xu, Q. Tian, "A robust Hough-based algorithm for partial ellipse detection in broadcast soccer video", Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on, vol. 3, no., pp. 1555, 1558, Vol. 3, 30-30 June 2004.