Counterattack Detection in Broadcast Soccer Videos using Camera Motion Estimation

Mohamad-Hoseyn Sigari, Hamid Soltanian-Zadeh
Control and Intelligent Processing Center of Excellence
School of Electrical and Computer Engineering
College of Engineering
University of Tehran
Tehran, Iran
hoseyn_sigari@yahoo.com, hszadeh@ut.ac.ir

Vahid Kiani, Hamid-Reza Pourreza
Machine Vision Research Laboratory
Computer Engineering Department
Faculty of Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
vahid.kiani@rocketmail.com, hpourreza@um.ac.ir

***Abstract*—This paper presents a new method for counterattack detection using estimated camera motion and evaluates some classification methods to detect this event. To this end, video is partitioned to shots and view type of each shot is recognized first. Then, relative pan of the camera during far-view and medium-view shots is estimated. After weighting of pan value of each frame according to the type of shots, the video is partitioned to motion segments. Then, motion segments are refined to achieve better results. Finally, the features extracted from consecutive motion segments are investigated for counterattack detection. We propose two methods for counterattack detection: (1) rule-based (heuristic rules) and (2) SVM-based. Experiments show that the SVM classifier with linear or RBF kernel results in the best results.**

***Index Terms*—Broadcast soccer video, counterattack detection, camera motion estimation, event detection, video analysis.**

I. Introduction

Nowadays, a variety of digital videos such as movies, news and sport videos are available on the web, hard disks, and non-volatile memories. Thus, traditional indexing/retrieval methods (tag-based search), manual analysis and summarization of video data are boring and challenging. Thus, we need some automatic or semi-automatic systems to analyze the video and extract content and semantic of video.

Sport video, especially soccer video is one of the interesting video types for analysis and information extraction. Sport video analysis is performed for video classification [1], video summarization [2, 3, 4, 5], video retrieval [6, 7], team/player tactic analysis [8, 9], developing referee-assistant systems [10, 11] and etc. Thus, it is demanded by coaches and sport reporters for technical and tactical analysis [8, 9] or statistical information of a match (e.g. percentage of ball possession [12]). Additionally, TV networks are interested to show advertisements on broadcast sport videos using virtual advertising techniques [13].

Almost all of the mentioned applications of sport video analysis are based on event detection. In fact, event is the basic entity of a video in the semantic level of computational hierarchy. Thus, a robust and high performance analysis is usually performed based on detection of events occurred in the video.

Soccer event detection systems can be categorized into two categories. The first category includes the systems detect events in broadcast videos. These systems usually used for the application of video summarization [2, 3, 5], video indexing/retrieval [6, 7] and virtual advertising [13]. On the other hand, the second category includes the systems developed for event detection using a network of special cameras. The cameras used by such systems are fixed and usually covers the entire field of match. For example, the systems developed for real-time detection of goal vent [10] and offside event [11] are categorized in the second category. Such systems are usually used as referee-assistant systems.

However researchers proposed a lot of methods for event detection in broadcast soccer videos, but they usually consider the most important events including goal [2, 3, 14, 15, 16], attack [2, 7, 14, 16] and corner [2, 14, 16, 17]. There are only a few researches that detect other events such as offside [2, 14], penalty [16], card [2] and counterattack [16] events. This paper presents a new method for counterattack detection.

In sports science, there are two types of soccer counterattacks. One is slow and based on ball controlling that relies on a lot of short soccer passes in all directions. The other one is a quick direct attack by moving the ball forward into the goal of the other team. In general, counterattack refers to the second definition. Thus, in this paper we assume that a counterattack is a quick attack into the opposition half by the defending team after winning the ball from the team previously attacking. Counterattack detection may be used for different applications such as tactical analysis, video indexing/retrieval and video summarization.

There is a main cue to detect a counter attack: fast turnover from one side of the playing field to the other side. This cue is also used for counterattack detection in [16]. The method proposed in [16], uses finite state machine and some heuristic rules to model some basic events (e.g. forward pass and turnover) in the feature space. Then, the system detects a counterattack when two specific basic events, i.e. forward pass and turnover are detected consequently. According to their experimental results, precision and recall rate of the proposed method are 83.3% and 83.3% respectively. This method is based on heuristic rules and may need to be adjusted when input video changes.

In this paper, we present a new method for counterattack detection in broadcast soccer videos. Additionally, we compare performance of two classification approaches for counterattack detection. The first approach is rule-based and the second one is SVM (Support Vector Machine)-based. Remaining of our paper organized as follows. In Section II, details of the proposed method are presented. Experimental results are reported in Section III. Conclusions and future works are explained in Section IV.

II. The Proposed Method

Our proposed method detects counterattack based on camera motion estimation. Block diagram of our proposed method for counterattack detection is depicted in Fig 1. At first, boundaries of shots are detected and video is segmented to the shots. Then, view type of each shot is determined. For the shots shown in far-views (long-shots) or medium-views (medium-shot), we estimate global motion of the camera. Then, the estimated camera pan is weighted and the video is segmented based on weighted camera pan. Finally, according to the estimated camera pan of consecutive motion segments, we detect counterattack events.

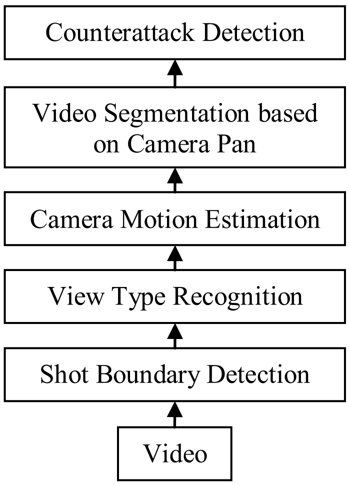

Fig. 1. Block diagram of our proposed method for counterattack detection

A. Shot Boundary Detection

Shot boundary detection is a low-level video processing that partitions the video into low-level parts. There are many methods for shot boundary detection [18, 19]. The main idea for shot boundary detection is based on calculating difference of features in consecutive frames. A fast and efficient method for shot boundary detection is based on the difference of color histograms of the consecutive frames. In this method, gradual shot transitions may be discarded, therefore, in the proposed method, the threshold applied for shot detection is determined adaptively to detect both of the abrupt and gradual shots.

To detect shot transitions by a fast and reliable method, a 16-bin histogram of each frame is computed for each of the RGB channels. Then, the histogram coefficients are normalized based on the size of the video frame. Therefore, there are 48 histogram coefficients for each video frame. The sum of absolute difference of the histogram coefficients for the *i*th frame and its previous frame is denoted by $hist\_diff_i$. The shot boundaries are detected by applying an adaptive threshold on $hist\_diff$. The adaptive threshold is defined according to the following equation for each video.

$$\begin{aligned} &th_{shot\ \mathrm{detection}} = \\ &average(hist\_diff_i \mid hist\_diff_i > mean_{hist\_diff}) \end{aligned} \quad (1)$$

Here, $mean_{hist\_diff}$ is the average value of the vector $hist\_diff$ for a video. Therefore, the threshold for shot detection is equal to the average value of some elements in $hist\_diff$ that are greater than $mean_{hist\_diff}$. By applying this threshold, cut and gradual shot transitions can be detected with a high detection rate. However, this method may detect some video segments with huge motion as shot transition. Since we need to minimize the overall shot transitions detection error, we may tradeoff some false shot transitions with detection of almost all cut and gradual shot transitions.

B. View Type Recognition

View type of soccer video shots is usually divided into four types: (1) far-view (or long-shot), (2) medium-view (or medium-shot), (3) close-up, and (4) out-field [20]. Estimation of camera motion in the shots with far- and medium- view may help us to detect counterattack events. On the other hand, camera motion estimation of the shots with close-up or out-field view is not useful for this purpose. Thus, in the proposed method, we classify the view types into three categories: (1) far-view (long-shots), (2) medium-view (medium-shot) and (3) others (close-up and out-filed views).

The first step of our view type classification is performed by detection of grass in the HSV color space. For this purpose, the grass color is detected using the approach presented in [21]. Then, the ratio of the grass pixels to the total pixels of the frame is calculated as R_{Grass}. In addition, the size of the biggest object in the field is calculated as S_{Obj}. The biggest object is detected by non-grass pixels in the grass field. This object must be surrounded by grass from all sides. In addition, proportion of the height to the width of the biggest object must be less than 10.

In a far- or medium- view shot, we expect to have a high R_{Grass} value, while S_{Obj} of far-view shots is greater than medium-view shots [2, 4]. These features are fed to a cascade classifier including two linear SVM classifiers for view type classification. The proposed cascade classifier rejects the shots with other view types by the first SVM. Then, the second SVM classifies the remaining samples into far- and medium- view types. Fig 2 depicts our proposed cascade classifier for view type recognition.

Processing of all frames in a shot is not necessary for recognition of the view type of that shot, then only 5 frames are selected from 10%, 30%, 50%, 70%, and 90% of shots duration. After classification of the view type of the selected frames, the majority vote method is used to estimate the view type of the shot.

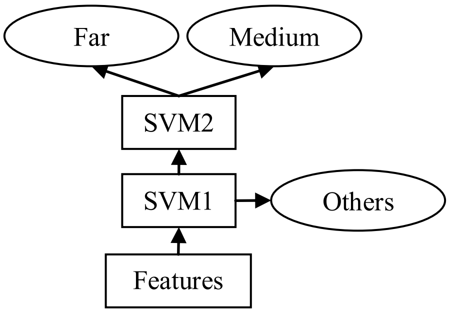

Fig. 2. Our proposed cascade classifier for view type recognition

C. Camera Motion Estimation

For estimation of the camera motion in a shot, we use th method presented in [22]. In this method, a fast block-base approach is used for global motion estimation in MPEG- videos. In fact, in this paper, camera motion estimation i performed by motion vectors extracted from MPEG codes i each shot separately. Thus, it is very fast. Additionally, thi method is robust because it uses a reliable mechanism fo removal of noise and outlier data.

In this method, we assume camera motion is constructed b three basic motions: pan, tilt and zoom. After extraction o reliable motion vectors, a pixel *(x,y)* in the current frame i mapped to the corresponding pixel *(x',y')* in the previou reference frame by:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} z & 0 \\ 0 & z \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} p \\ t \end{bmatrix} \qquad (2$$

where z, p and t stands for zoom (scale), pan and til respectively. Putting start and end point of each motion vecto of a frame in (2) results in a linear equation system. W achieve to the initial values of the parameters of camera motio when this linear equation system is solved. In the next step, th initial parameters of camera motion are refined by an iterativ method [22].

The result values for the parameters of camera motion ar relative amounts. In other words, the value of pan, tilt an zoom of the camera is estimated for each frame of a sho relatively with respect to the first frame of that shot.

In our proposed method for counterattack detection, w assume camera zoom is almost constant during a shot. Thus we need only pan value of camera motion. In fact, pan value o camera motion shows the horizontal movement of camera along the field from on half to the other half.

In a counterattack event, team A attacks to team B, then team B wins the ball and attacks to team A quickly. Therefore, camera moves from left (right) to right (left) when team A attacks (forward movement); and then, camera moves from right (left) to left (right) when team B counterattacks (turnover movement). The forward movement and the turnover movement are two basic events that compose a counterattack event when they occur consequently. Although these two basic events occur consequently, they may be shown in a shot or in some shots with different view types. Thus, we partition the video to the new segments called motion segments.

To this end, pan of the camera is estimated for all far-view and medium-view shots. Also, zero is used as the pan of the camera for the shots with other view types (close-up and out-filed). Finally, a signal denoted by P is generated that shows the pan of the camera for each frame. Fig 3 shows the pan signal of a match (Italy vs. Slovakia, 1st half). P contains both positive and negative values. Positive values show that the movement of the camera is to the right side. In contrary,

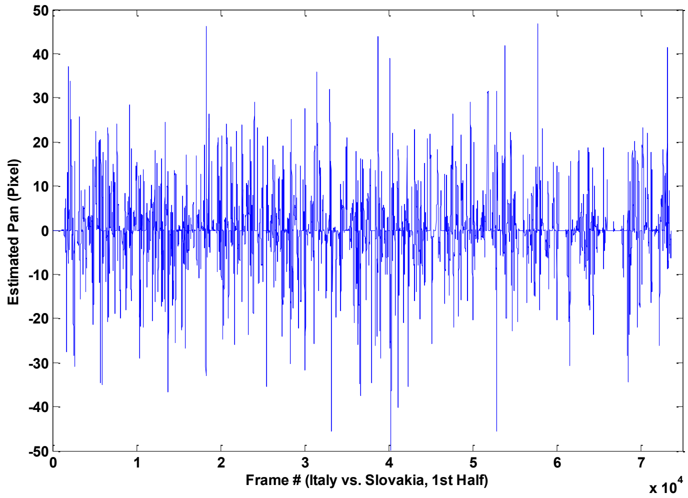

Fig. 3. Estimated pan signal (P) of a sample match (Italy vs. Slovakia, 1st

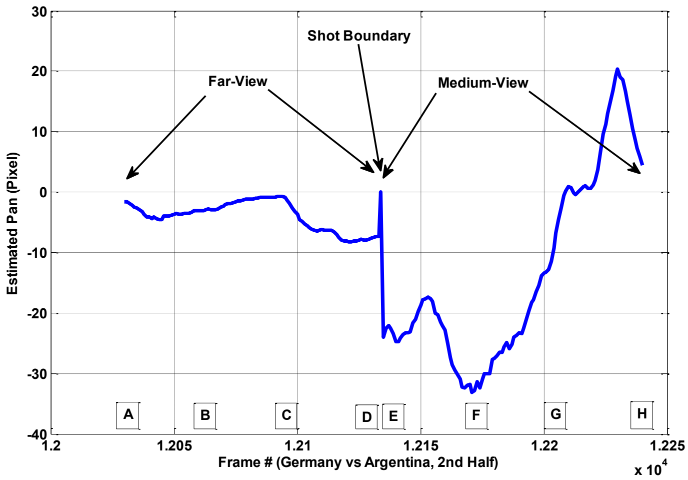

Fig. 4. Estimated camera pan of the frames appeared during two consecutive shots. The first shot has far-view and the second one has medium-view.

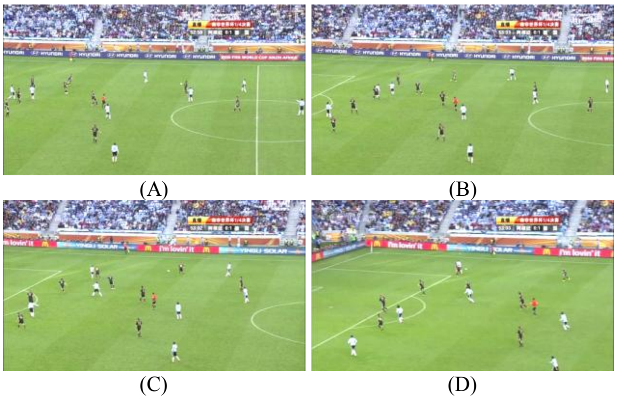

Fig. 5. Sample frames related to the far-view shot mentioned in Fig 4.

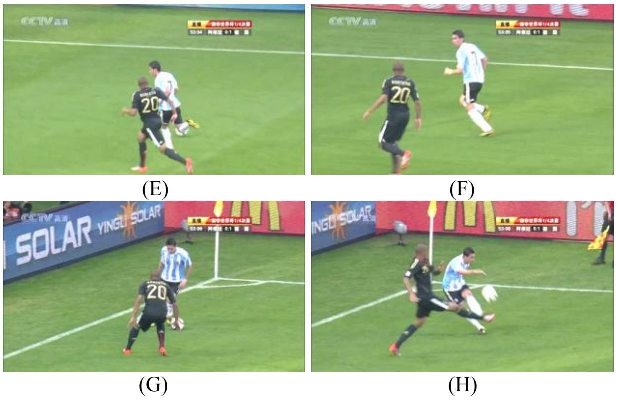

Fig. 6. Sample frames related to the medium-view shot mentioned in Fig 4.

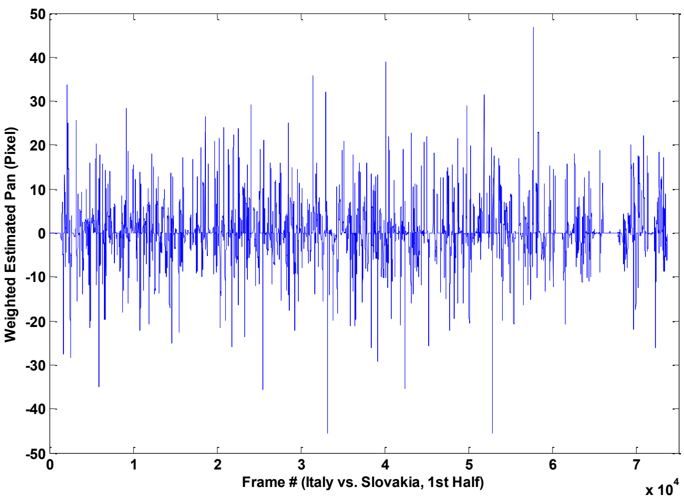

Fig. 7. Weighted pan signal (P) of the sample match (Italy vs. Slovakia, 1st half). Original pan signal of this match is shown in Fig 3.

As mentioned before, the main drawback of our method for camera motion estimation is that we estimate only relative amounts of parameters. On the other hand, pan of the camera in a far-view shot is usually less than in a medium-view shot. In a far-view shot, camera may cover the region of interest with lower motion, but in a medium-view shot, covering the region of interest needs more camera motion, including camera pan. To have a better insight about camera pan in far-view and medium-view shots, we show the changes of pan value during some frames of two different shots in Fig 4. Fig 4 shows the estimated pan value of the frames appeared during two consecutive shots. The first shot has a far-view and the second one has a medium-view. Some sample frames related to these shots are shown in Fig 5 and 6. According to Fig 4, 5 and 6, when camera moves to the left, estimated pan is negative and vice versa. But the absolute value of the estimated pan for the medium-view shot is greater than the far-view shot.

According to the above explanations, camera pan during the frames of a medium-view shot changes faster than a far-view shot. Thus, we use a simple method to weight the *P* signal. Based on the trial and error experiments, we assume the weight 1 and 0.25 for far-view and medium-view shots respectively. In other words, *P* value of the frames related to medium shots is replaced by one fourth of its previous value, while *P* value of the frames of far-view shots is remaining without any change. The weighted pan signal of the previously mentioned match (Italy vs. Slovakia, 1st half) is shown in Fig 7. Thereafter, we substitute *P* signal by weighted *P* signal.

D. Video Segmentation based on Camera Pan

Based on the *P* signal, the video is partitioned to the motion segments. To this end, sign of *P* signal is calculated first. Sign of *P* shows the side of camera pan (-1 for left movement, +1 for right movement and 0 for no movement). The consecutive frames those have a same sign are put in a same segment. Thus, the frame that sign of *P* signal changes in there is detected as boundary of motion segmentation. In other words, in boundary of segments, sign of *P* signal changes. The motion segments are classified into three classes based on the sign of *P* signal: (1) no motion, (2) movement to the right and (3) movement to the left.

In the next step, we refine the result of motion segmentation by a simple algorithm. We know that the camera may move to right (or left) altogether while the camera movement may be stopped for a few seconds. Additionally, during forward passes between players of a team in an attack to the right side, the players may pass the ball together with a short backward pass. This may cause a camera movement to the left for a very short time, while the overall movement of camera is to the right side. Thus, we change the signs of very short motion segments (shorter than 25 frames) to zero value. Then, we concatenate very short segments to the next segment. Also, we replace the *P* values of such very short segments by the average *P* value of the next segment which was merged by the very short segment.

E. Counterattack Detection

After refinement of motion segmentation, some features are extracted from the *i*th motion segment M_i:

1. T_i^1 : start point of motion segment (in frame)
2. T_i^2 : stop points of motion segment (in frame)
3. L_i : length of motion segment (in frame)
4. S_i : sign (side) of movement (-1 ,0, +1)
5. A_i : average of *P* signal during the motion segment

Finally, two motion segments that have different sign (±1) and occur close together may be considered for counterattack event. For counterattack detection, we propose some simple rules applied on a motion segment M_i. At first, we find the nearest motion segment M_j occurred before M_i that $S_i = -S_j$. Then, a segment of video occurred in $[T_j^1 - T_i^2]$ is a counterattack event if:

- $T_i^1 - T_j^2 < Th_{interval}$
- $L_i > Th_{length}$
- $L_j > \frac{1}{2} Th_{length}$
- $A_i > Th_{average}$
- $A_j > \frac{1}{2} Th_{average}$

According to above rules, M_i and M_j must occur close enough together. Also the length and the average of P signal of M_i and M_j must greater than some thresholds. In this case, M_i and M_j are the forward pass and turnover events respectively. The threshold applied for M_j is one half of the threshold applied for M_i, because the attack (forward pass) may occur slower than counterattack (turnover).

In addition to the proposed rule-based method, we use an SVM classifier for counterattack detection. First the features are calculated from the candidate M_i and M_j ($S_i = -S_j$): $T_i^1 - T_j^2$, L_i, L_j, A_i, A_j. Then, the features are normalized (with zero mean and unit variance) and fed to an SVM for counterattack detection. Output of the SVM is 1 if a counterattack event is detected. We evaluate performance of different kernel of SVM for counterattack detection.

III. Experimental Results

Experiments are performed on some broadcast soccer videos related to FIFA World Cup 2010, South Africa. The selected matches are listed in Table I. The total duration of this dataset is about 10 hours. All videos are 720×480 and the video frame rate is 25 fps in MPEG-1 format.

TABLE I. List of videos used for the experimental results.

Video #	Description	Video #	Description
1	Germany vs. Argentina (1st Half)	2	Germany vs. Argentina (2nd Half)
3	Germany vs. England (1st Half)	4	Germany vs. England (2nd Half)
5	Germany vs. Spain (1st Half)	6	Germany vs. Spain (2nd Half)
7	Germany vs. Uruguay (1st Half)	8	Germany vs. Uruguay (2nd Half)
9	Greece vs. Argentina (1st Half)	10	Greece vs. Argentina (2nd Half)
11	Slovakia vs. Italy(1st Half)	12	Slovakia vs. Italy (2nd Half)

We proposed two approaches for counterattack detection: rule-based and SVM-based. The rule-based approach is performed by some heuristic rules defined by an expert. Also, for the second approach, we may use different kernels for the SVM classifier. In our experiments, we evaluate linear, polynomial and RBF (Radial Basis Functions) kernels.

The first approach does not need training, but we have to tune our thresholds based on some samples. Also, SVM must be trained before using as a classifier. For tuning our first approach and training of SVM in the second approach, we use the videos # 1, 3, 5, 7, 9 and 11. Thus, the entire counterattack events occurred in these videos (totally 46 samples) is selected as positive samples. For negative samples, we select 46 samples manually form training videos to have a balanced training data set.

For evaluation of different kernels of SVM, we apply three kernel types: linear, polynomial and RBF. The linear kernel has not any parameter, but for polynomial and RBF kernels, the parameter of kernel is the degree of polynomial function (p) and the radius of radial function (r). For polynomial kernels, we evaluate different values for p: 2, 3, 4 and 5. Also for RBF kernels, we use different values for r: 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5.

According to our experiments, the best values for the parameter of polynomial and RBF kernels are p=2 and r=0.1, respectively. The results achieved for performance evaluation of different methods are shown in Table II.

According to the achieved results, performance of SVM classifiers with linear and RBF (r=0.1) kernels are better than other methods. Although linear and RBF kernels have different precision and recall, the F-measure of both is same.

TABLE II. Results of our experiments to evaluate different approaches for counterattack event detection.

	# True Detected	# False Detected	# False Rejected	Precision	Recall	F-Measure
Rule-Based	31	20	27	60.8%	53.4%	56.9%
SVM Linear	35	21	24	**62.5%**	58.6%	**60.5%**
SVM Polynomial p=2	34	26	24	56.7%	58.6%	57.6%
SVM RBF r=0.1	34	22	23	60.7%	**60.3%**	**60.5%**

IV. Conclusions

This paper presented a new method for counterattack event detection using estimated camera pan. To this end, relative pan of the camera during shots was estimated. Then, the pan values were weighted based on the view type of shots. After weighting of pan signal, the video was partitioned to motion segments based on the camera pan. Then, motion segments were refined to remove very short segments and combine very short segments with its long neighbors. Finally, we proposed two methods for counterattack detection: (1) rule-based (heuristic rules) and (2) SVM-based.

Results showed that the rule-based method has the lowest performance with respect to SVM-based method. Although SVM-based method is better than rule-based method, different kernels of SVM resulted in different precision, recall and F-measure. According to the achieved results, the SVM classifier with linear or RBF (r=0.1) kernel resulted in the best F-measure.

For future works, we propose to remove replay parts as a preprocessing phase. This preprocessing phase may decrease mistakes of the system, especially decrease false detection rate.

Also, ensemble approaches for classification may increase precision and recall rate for counterattack detection. The ensemble approaches resulted in better performance for the previous applications of image and video analysis [1, 23].

Results of counterattack detection may be used for coaches to analyze strategies and tactics of own and opponent teams. Although our proposed automatic system for counterattack detection is not very accurate, we may develop a semi-automatic system as a Decision Support System (DSS) that helps the coaches to detect counterattacks accurately with a little information inserted manually.

REFERENCES

[1] Mohamad Hoseyn Sigari, Samaneh Abbasi Sureshjani, Hamid Soltanian-Zadeh, "Sport Video Classification using an Ensemble Classifier", Iranian Machine Vision and Image Processing, Tehran, Iran, pp. 1-4, November, 2011.

[2] Mostafa Tavassolipour, Mahmood Karimian, Shohreh Kasaei, "Event Detection and Summarization in Soccer Videos Using Bayesian Network and Copula", IEEE Transactions on Circuits And Systems for Video Technology, vol. 24, no. 2, pp. 291-304 2013.

[3] Ahmet Ekin, A. Murat Tekalp, Rajiv Mehrotra, "Automatic Soccer Video Analysis and Summarization", IEEE Transactions on Image Processing, vol. 12, no. 7, pp. 796-807, July, 2003.

[4] Vahid Kianiand, Hamid Reza Pourreza, "Flexible Soccer Video Summarization in Compressed Domain", International eConference on Computer and Knowledge Engineering, Mashhad, Iran, pp. 1-6, 2013.

[5] Mohamad-Hoseyn Sigari, Hamid Soltanian-Zadeh, Hamid-Reza Pourreza, "Fast Highlight Detection and Scoring for Broadcast Soccer Video Summarization using On-Demand Feature Extraction and Fuzzy Inference", International Journal of Computer Graphics, vol. 6, no. 1, 2015.

[6] Shu-Ching Chen, Mei-Ling Shyu, Na Zhao, "An Enhanced Query Model for Soccer Video Retrieval using Temporal Relationships", International Conference on Data Engineering, Tokyo, Japan, pp. 1133-1134 2005.

[7] Maheshkumar H. Kolekar, Kannappan Palaniappan, Somnath Sengupta, Gunasekaran Seetharaman, "Semantic Concept Mining Based on Hierarchical Event Detection for Soccer Video Indexing", Journal of Multimedia, vol. 4, no. 5, pp. 298-312, October, 2009.

[8] Guangyu Zhu, Changsheng Xu, Qingming Huang, Yong Rui, Shuqiang Jiang, Wen Gao, Hongxun Yao, "Event Tactic Analysis Based on Broadcast Sports Video", IEEE Transactions on Multimedia, vol. 11, no. 1, pp. 49-67, January, 2009.

[9] Pascual J. Figueroa, Neucimar J. Leite, Ricardo M. L. Barros, "Tracking Soccer Players Aiming Their Kinematical Motion Analysis", Computer Vision and Image Understanding, vol. 101, no. 2, pp. 122-135, 2006.

[10] T. D'Orazio, M. Leo, P. Spagnolo, M. Nitti, N. Mosca, "A Visual System for Real Time Detection of Goal Events During Soccer Matches", Computer Vision and Image Understanding, vol. 113, no. 5, pp. 622-632, May, 2009.

[11] T. D'Orazio, M. Leo, P. Spagnolo, P. L. Mazzeo, N. Mosca, M. Nitti, A. Distante, "An Investigation into the Feasibility of Real-Time Soccer Offside Detection From a Multiple Camera System", IEEE Transactions on Circuits And Systems for Video Technology, vol. 19, no. 12, pp. 1804-1818, December, 2009.

[12] Xinguo Yu, Leong Hon Wai, Joo-Hwee Lim, Qi Tian, Zhenyan Jiang, "Team Possession Analysis for Broadcast Soccer Video based on Ball Trajectory ", Joint Conference of Int'l Conference on Information, Communications and Signal Processing, and Pacific Rim Conference on Multimedia, pp. 1811-1815, 2003.

[13] Vahid Babaee-Kashany, Hamid Reza Pourreza, "Camera Pose Estimation in Soccer Scenes Based on Vanishing Points and its Application in Virtual Advertisement Insertion", Iranian Conf. on Machine Vision and Image Processing, Isfahan, Iran, 2010.

[14] Monireh-Sadat Hosseini, Amir-Masoud Eftekhari Moghadam, "Fuzzy Rule-based Reasoning Approach for Event Detection and Annotation of Broadcast Soccer", Applied Soft Computing, vol. 13, no. 2, pp. 846-866, 2013.

[15] Yina Han, Guizhong Liu, Gerard Chollet, "Goal Event Detection in Broadcast Soccer Videos by Combining Heuristic Rules with Unsupervised Fuzzy C-Means Algorithm", International Conference on Control, Automation, Robotics and Vision, Hanoi, Vietnam, pp. 888-891, December, 2008.

[16] Jurgen Assfalg, Marco Bertini, Carlo Colombo, Alberto Del Bimbo, Walter Nunziati, "Semantic Annotation of Soccer Videos: Automatic Highlights Identification", Computer Vision and Image Understanding, vol. 92, no. 2-3, pp. 285-305, 2003.

[17] Jila Hosseinkhani, Hamid Soltanian-Zadeh, Mahmoud Kamarei, Oliver Staadt, "Ball Detection with the Aim of Corner Event Detection in Soccer Video", International Symposium on Parallel and Distributed Processing with Applications Workshops, Busan, South Korea, pp. 147-152, May, 2011.

[18] Alan F. Smeaton, Paul Over, Aiden R. Doherty, "Video Shot Boundary Detection: Seven Years of TRECVid Activity", Computer Vision and Image Understanding, vol. 114, no. 4, pp. 411-418, 2010.

[19] John S. Boreczky, Lawrence A. Rowe, "Comparison of Video Shot Boundary Detection Techniques", Journal of Electronic Imaging, vol. 5, no. 2, pp. 122–128, April, 1996.

[20] T. D'Orazio, M. Leo, "A Review of Vision-based Systems for Soccer Video Analysis", Pattern Recognition, vol. 43, no. 8, pp. 2911–2926, August, 2010.

[21] Lei Wang, Boyi Zeng, Steve Lin, Guangyou Xu, Heung-Yeung Shum, "Automatic Extraction of Semantic Colors in Sports Video", International Conference on Acoustics, Speech, and Signal Processing, pp. 617-620, May, 2004.

[22] Vahid Kianiand, Hamid Reza Pourreza, "Robust GME in Encoded MPEG Video", International Conference on Advances in Mobile Computing and Multimedia, pp. 147-154, 2011.

[23] Fatemeh Ahmadi, Mohamad-Hoseyn Sigari, Mohamad-Ebrahim Shiri, "A Rank based Ensemble Classifier for Image Classification using Color and Texture Features", Iranian Conference on Machine Vision and Image Processing, Zanjan, Iran, pp. 1063-1068, 2013.