A NEW METHOD FOR VIDEO SOCCER SHOT CLASSIFICATION

Youness Tabii, Mohamed Ould Djibril, Youssef Hadi and Rachid Oulad Haj Thami*
Laboratoire SI2M, Equipe WiM, ENSIAS B.P 713, Université Mohamed V - Souissi, RABAT - Maroc

Keywords: Video soccer, shot classification, binary image, golden section.

Abstract: A shot is often used as the basic unit for both video analysis and indexing. In this paper we present a new method for soccer shot classification on the basis of playfield segmentation. First, we detect the dominant color component, by supposing that playfield pixels are green (dominant color). Second, the segmentation process begins by dividing frames into a 3:5:3 format and then classifying them. The experimental results of our method are very promising, and improve the performance of shot detection.

1 INTRODUCTION

Since football is the most popular game in the world, the analysis of its videos has become an important research field that attracts a great number of researchers. The video document presents audio/visual information, thus, making it possible to analyze well this type of documents and to extract the semantics from the videos by making use of several algorithms.

The objective of football video analysis is: (1) to extract events or objects in the scene; (2) to produce general summaries or summaries for the most important moments in which TV viewers may be interested. The segmentation of playfields, events and objects detection play an important role in achieving the above described aims. The analysis of football video is very useful for this game's professionals because it enables them to see which team is better in terms of ball possession or to detect which strategy is useful for each team in a specific moment.

A number of related works which deal with the extraction of the semantics of soccer videos are available in the literature. In (D. Yow and Liu, 1995), the object colour and texture features are employed to generate highlights and to parse TV soccer programs (Y. Gong and Sakauchi, 1995). In (J. Assfalg and Nunziati, 2003), the authors use playfield zone classification, camera motion analysis and players' position to extract highlights. In (S.C. Chen and Chen, 2003), a framework for the detection of soccer goal shots by using combined audio/visual features was presented. It employs soccer domain knowledge and the PRISM approach so as to extract soccer video data. However, (L.Y. Duan and Xu, 2003) propose a mid-level framework that can be used to detect events, extract highlights as well as to summarize and personalize sports video. The information they employed include low-level features, mid-level representations and high-level events. And in (Y. Qixiang and Shuqiang, 2005), the authors present a framework based on mid-level descriptors after the segmentation of the playfield with GMM (Gaussian Mixture Models). In (J. Assfalg and Pala, 2002), camera motion and some object-based features are employed to detect certain events in soccer video. In other works, the authors extract replays, highlights, goals and positions of players and referee.

In the present paper, the standard RGB colour representation is not convenient. The RGB values of decoded frames are transformed into corresponding coefficients in the HSV colour space, before analysis. HSV presents three different parameters: hue, saturation and brightness. Hue determines the dominant wavelength of the colour with values ranging from 0 to 360 degrees. Brightness illustrates the level of white light (0 - 100) while Saturation describes the

*This work has been supported by Maroc-Telecom.

Tabii Y., Ould Djibril M., Hadi Y. and Oulad Haj Thami R. (2007).
A NEW METHOD FOR VIDEO SOCCER SHOT CLASSIFICATION.
In *Proceedings of the Second International Conference on Computer Vision Theory and Applications - IFP/IA*, pages 221-224

proportion of chromatic element in a colour. Values range from 0 to 100, where low values indicate that the colour has much ”greyness” and will appear faded. As humans are much more sensitive to hue than to saturation and brightness, one parameter becomes far more important than the others and the HSV representation is, therefore, excellent for colour analysis.

In this paper, our algorithm for shot classification is presented. It exploits the spatial features of binarization as well as the frame partition (W. Kongwah and Changsheng, 2003) (A. Ekin and Mehrotra, 2003). This algorithm is able to detect a variety of shot types with a high percentage.

The remainder of this paper is organized as follows. Section 2 introduces the algorithm for playfield segmentation. Section 3 describes the new proposed algorithm for the classification of shots in soccer video. Section 4 presents the experimental results. Finally, Section 5 gives a brief conclusion.

2 PLAYFIELD SEGMENTATION

In this section, we present the statistical computation of the dominant colour and binarization (playfield segmentation).

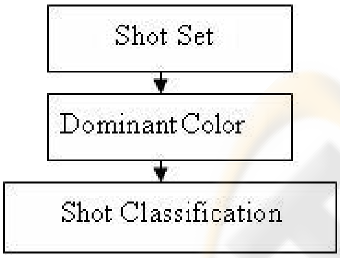
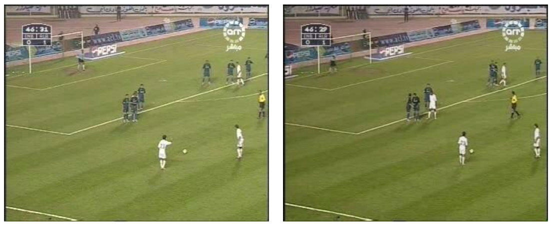


Figure 1: stages for playfield segmentation.

Figure 1 shows the different stages of our procedure for shot classification. The algorithm is composed of three steps : 1) Firstly, we manually extract the shots that are to be classified. 2) Secondly, once the shots are extracted, we compute the dominant colour. The latter allows us to characterize the shots better. 3) Thirdly, to classify the shots , we make used of binarized frames to distinguish different shots.

2.1 Color Dominant Extraction

The playfield usually has a distinct tone of green that may vary from stadium to stadium. But in the same stadium, this green colour may also change due to weather and lighting conditions (see Figure 2). Therefore, we do not assume any specific value for the dominant colour of the field.



(a) (b)

Figure 2: weather and lighting conditions.

We compute the statistics of the dominant field colour in the HSV space by taking the mean value of each colour component around its respective histogram peaks, i_{peak}. An interval $[i_{min}, i_{max}]$ is defined around each i_{peak}. The same method is adopted in (W. Kongwah and Changsheng, 2003):

$$\sum_{i=i_{min}}^{i_{peak}} H[i] <= 2H[i_{peak}] \;\; and \sum_{i=i_{min}-1}^{i_{peak}} H[i] > 2H[i_{peak}] \quad (1)$$

$$\sum_{i=i_{peak}}^{i_{max}} H[i] <= 2H[i_{peak}] \;\; and \sum_{i=i_{peak}}^{i_{max}+1} H[i] > 2H[i_{peak}] \quad (2)$$

$$colomean = \frac{\sum_{i=i_{min}}^{i_{max}} H[i]*i}{\sum_{i=i_{min}}^{i_{max}} H[i]} \quad (3)$$

Using the following quantization factor: 64 hue, 64 saturation, 128 intensity, H is the histogram for each colour component (H,S,V). Finally, the colour mean is then converted into (R_{mean}, G_{mean},B_{mean}) space so as to determine the playfield surface :

$$G(x,y) = \begin{cases} 1 \quad \text{if} \begin{cases} I_G(x,y) > I_R(x,y) + K(G_{peak} - Rpeak) \\ I_G(x,y) > I_B(x,y) + K(G_{peak} - Bpeak) \\ |I_R - Rpeak| < R_t \\ |I_G - Gpeak| < G_t \\ |I_B - Bpeak| < R_t \\ I_G > Gth \end{cases} \\ 0 \quad \text{otherwise} \end{cases} \quad (4)$$

G(x,y) is the binarized image frame in the field colour. In our system, the new thresholds after a number of tests are : $R_t = 12$, $Gt = 18$, $B_t = 10$, $K = 0.9$ and $G_{th} = 85$. the Eq 1-3 are computed for every I/P frames.

4 EXPERIMENTAL RESULTS

The video sequences used for the evaluation of the shot classification algorithm are MPEG compressed movies. This will allow us to test our algorithm on MPEG artificats due to the compression. The sequences also contain objects and camera motion.

About 6 hours video sequences of various soccer matches in different champions leagues transcoded into MPEG 352x288, 1150kbps are used.

These video files are first parsed by using a manual shot detection. The two visual features (i.e. binarization and Golden section) are computed and normalized for each video shot. In total, we have 435 Long shots, 162 Medium shots and 220 close-up shots.

Table 1: Results of shot classification algorithm.

Semantic	True Shot	False Shot	Precision(%)
LS	423	12	97.2
MS	125	10	93.8
CpS	207	13	94.0
Total	782	35	95.0

Where LS is Long Shot, MS is Medium shot and CpS is Close-up Shot.

Table 1 shows the result we obtained for shot classification. The classification rate of LS is high, but for MS and CpS it is relatively low. This is maybe due to the pre-fixed thresholds of *Linemean* and *Colunmmean*. In other words, the features are less discriminant for those types of shots. However our algorithm works satisfactorily.

5 CONCLUSION

In this paper we presented a new method for the classification of video soccer shots on the basis of spatial analysis. The main contribution of the presented work an algorithm for shots classification.

The advantage of our algorithm is clearly seen in its simplicity and effectiveness in providing better results for the classification of the majority of football matches. Besides, the analysis of soccer video on the basis of playfield segmentation is very promising.

REFERENCES

A. Ekin, A. T. and Mehrotra, R. (2003). Automatic soccer video analysis and summarizartion. IEEE, Symp.

D. Yow, B.L. Yeo, M. Y. and Liu, B. (1995). Analysis and presentation of soccer highlights from digital video. ACCV.

J. Assfalg, M. Bertini, A. B. W. N. and Pala, P. (2002). Soccer highlights detection and recongnition using hmms. IEEE.

J. Assfalg, M. Bertini, C. C. A. B. and Nunziati, W. (2003). Semantic annotation of soccer videos: automatic highlights identification. Computer Vision and Image Understanding.

L.Y. Duan, M. Xu, T. C. Q. T. and Xu, C. (2003). Amid-level representation framwork for semantic sports video analysis. ACM.

S.C. Chen, M.L. Shyu, C. L. L. and Chen, M. (2003). Detection of soccer goal shots using joint multimedia features and classification rules. Proceedings of the Fourth International Workshop on Multimedia Data Minig, ACM.

W. Kongwah, Y. Xin, Y. X. and Changsheng, X. (2003). Real-time goal-mouth detection in mpeg soccer video. ACM, Berkley, California, USA.

Y. Gong, L.T. Sin, C. S. H. Z. and Sakauchi, M. (1995). Automatic parsing of soccer programs. IEEE, Syst.

Y. Qixiang, H. Qingming, G. W. and Shuqiang, J. (2005). Exiting event detection in braodcast soccer video with mid-level description and incremental learning. Technical report, MM05,,Singapore, ACM.