An SVM-Based Soccer Video Shot Classification Scheme Using Projection Histograms

Nan Nan, Guizhong Liu, Xueming Qian, and Chen Wang

School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China, 710049
nannan@mail.xjtu.edu.cn, liugz@xjtu.edu.cn,
qianxm@mail.xjtu.edu.cn, wangchen615@gmail.com

Abstract. In this paper, we propose an SVM based shot classification scheme, which categorizes shots of soccer videos into 4 kinds, namely global shots, medium shots, close-up shots and audience shots. The proposed scheme consists in a new adaptive dominant color detection algorithm, as well as a novel feature, the projection histogram. Experiments show that our scheme performs competitively and our new feature is effective.

Keywords: SVM, dominant color, projection histogram, shot classification.

1 Introduction

Shot classification is undoubtedly one of the critical techniques for analyzing and retrieving digital videos. A number of shot classification algorithms for sports videos have been proposed recently ([1, 2]). All these methods, however, rely on the proper thresholds or the accurate shot information.

Considering the limitations of the above schemes, here we propose an SVM-based method, which uses frame features rather than shot features to achieve better flexibility. A new adaptive dominant color detection algorithm and a new feature, namely the projection histogram, are proposed in this scheme. Three SVMs are designed and utilized to classify shots in a serial order. Experiments show that the new feature is effective and our method performs competitively.

The rest of this paper is organized as follows. Section 2 introduces the detailed approaches of the whole scheme. In Section 3, experiments and a result analysis are presented. A conclusion is drawn in Section 4.

2 An SVM-Based Shot Classification Scheme

As illustrated in Fig.1, our system utilizes three SVMs, which uses different combinations of the respective features, to classify shots into four main categories, namely the global, medium, close-up and audience in a serial order.

Y.-M.R. Huang et al. (Eds.): PCM 2008, LNCS 5353, pp. 883–886, 2008.

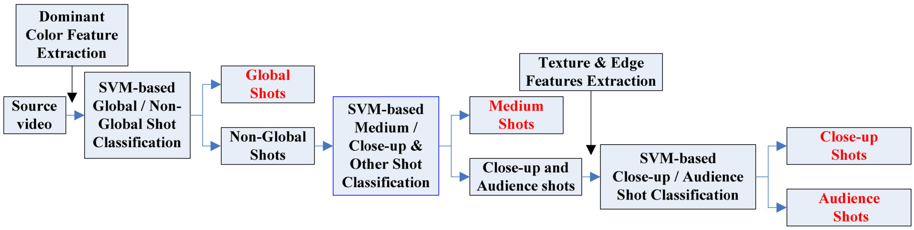


Fig. 1. Structure of the overall shot classification scheme

2.1 Color Feature Extraction

The dominant color for a soccer video is essentially the color of the grass field and its size and distribution offer significant cues for shot classification. The dominant color ratio and projection histograms are proposed to describe them respectively.

Our scheme follows 3 steps to adaptively detect the dominant color: **Step1. Initial dominant color modeling**: we construct the accumulative histograms of the HSI components from hundreds of sample frames selected form several soccer videos, and calculate the initial dominant color using the method in [3]; **Step2. Initial dominant color region determination**: the initial dominant color and the cylindrical metric (described in [3]) are then used to classify which pixels in the current frame are initial dominant colored; **Step3. Adaptive dominant color refinement**: for the pixels labeled as initial dominant colored, their accumulative histograms of HSI are constructed again and the same operations as in the first step are utilized to obtain the refined dominant color of the current frame.

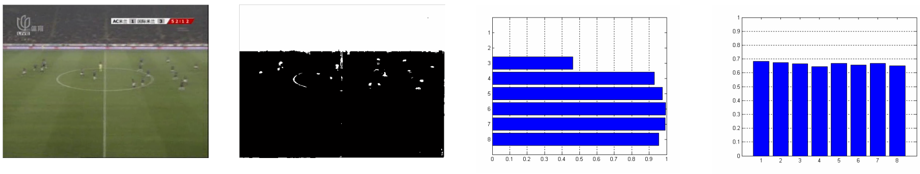
To obtain the dominant color ratio, we simply count the number of dominant colored pixels of the frame and divide it by the total number of pixels in that frame. To obtain the dominant color projection histograms, a binary image of the current frame is generated at first by labeling the dominant colored pixels as 0 and the other pixels as 1, and then projected onto both the vertical and horizontal axes, through which two 8-bin histograms are constructed (as shown in Fig. 2), respectively.

2.2 Texture and Edge Feature Extraction

Due to the inherent limitations of color features, such as the sensitivity to lighting, weather, shade and the incapacity of describing the details, the texture and edge information are utilized as the supplements.

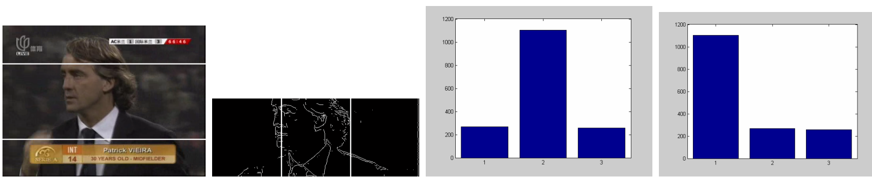
The texture intensity (TI) of a pixel is the variance of the intensity component of the pixels in the sliding window centered at it. Here we use ATI (the average of TI) and VTI (the variance of TI) to characterize the statistical attributes of the texture for the whole frame.

For edge features, a 3-bin horizontal projection histogram of the edge information is utilized, and the extraction follows the steps as Fig.3 show. First, a frame is partitioned along the vertical direction into 3 slices by a proportion of 1: 2: 1, then the central slice is extracted to maintain the most part of the frame’s content and avoid the influence caused by the logos and text lines. Second, we equally partition the slice



(a) (b) (c) (d)

Fig. 2. Dominant color projection histograms (a) The original image. (b) The binary image. (c) The histogram of vertical axis projection. (d) The histogram of horizontal axis projection.



(a) (b) (c) (d)

Fig. 3. Edge feature extraction (a) Partitioning the frame. (b) Partitioning the central slice and detecting the edges. (c) Obtaining the edge projection histogram. (d) Sorting the histogram.

into three blocks along the horizontal direction and use the sobel operator to detect the edge information of them respectively. Third, we calculate the edge pixel number in each of the blocks and obtain the edge projection histogram.

2.3 The Selection of Features for the Three SVMs

For the global shot versus non-global shot classification, we use the dominant color ratio of the frame and the two middle values of both the vertical and horizontal axis projection histograms. The 5-dimension feature vector is defined as follows:

$$v1 = \{domClrRatio, verHist(4), verHist(5), horHist(4), horHist(5)\} \quad (1)$$

For the medium shot versus close-up and other shot classification, we sort the horizontal axis projection histogram in descending order, and pick up its 6 middle values plus the dominant color ratio to form the 7-dimension feature vector:

$$v2 = \{domClrRatio, sorted\_horHist(2:7)\} \quad (2)$$

For the close-up shot versus the audience shot classification, the dominant color ratio, ATI, VTI, and the edge projection histogram sorted in descending order are utilized to from the 6-dimension feature vector:

$$v3 = \{domClrRatio, ATI, VTI, sorted\_EdgeHist(1:3)\} \quad (3)$$

3 Experimental Results

The soccer video sequences we selected for the experiment are from multiple sources, whose sizes vary from 512*384 to 480*360. All together 4000 frames are extracted from over 1000 different shots. The radial basis function is chosen as the kernel function for the SVMs and the threshold for the sobel detector is fixed at 0.03. A quarter of the total frames are used to train the SVMs, and the rest frames are used for test. Recall and Precision are used for evaluating the experimental results. The classification results of both our scheme and the one in [2] are demonstrated in Tab.1, from which we can see that our scheme obtains better classification performance whatever the shot type is. Moreover, our method can distinguish the close-up shots from the audience shots, while the method in [2] can not.

Table 1. The results of the classifications for different shots

Shot Type	Recall		Precision	
	Our Method	[2]	Our Method	[2]
Global	0.965	0.693	0.965	0.847
Medium	0.947	0.764	0.930	0.6753
Close-up	0.920	0.7390	0.940	0.939
Audience	0.940		0.920	

4 Conclusion

This paper addresses an SVM-based shot classification scheme, which utilizes multiple features and three different SVMs to accomplish the whole classification process, and its novelty is incarnated in the new adaptive dominant color detection algorithm and the usage of the projection histograms in shot classification. Future work will focus on the extension of the scheme to different kinds of sports videos.

Acknowledgments. The authors would like to thank the anonymous reviewers for their valuable and constructive comments that helped improve the presentation of this paper substantially. This work is supported in part by National 973 Project No.2007CB311002, National Natural Science Foundation of China (CNSF) Project No.60572045, and the Ministry of Education of China Ph.D. Program Foundation Project No.20050698033.

References

1. Ekin, A., Tekalp, A.M.: Shot type classification by dominant color for sports video segmentation and summarization. In: ICASSP, pp. 173–176 (2003)
2. Ekin, A., Tekalp, A.M.: Automatic Soccer Video Analysis and Summarization. IEEE Transactions on Image Processing, 796–807 (2003)