Segmentation in Soccer Video Based on Fusion Algorithm

Qiao ZHANG, Ping SHI[1]

Abstract. According to the practical application of soccer video, the characteristics of the data structure and the shot conversion methods of soccer videos are described. A method of shot segmentation in soccer video based on fusion algorithm is proposed. In this method, the χ^2 histogram match based on gray histogram, sliding window algorithm with adaptive threshold and double sliding windows algorithm are combined to detect and segment the shot boundary. Experiment results show that the average recall rate of cut shot is 96.6% and the average precision rate is 94.9%; The average recall rate of gradual shot is 79.5% and the average precision rate is 81.4%.

Keywords: shot boundary detection • shot segmentation • sliding window • histogram • dual thresholds.

1 Introduction

With the rapid development in web technology and multimedia technology, sports video, especially soccer video, is becoming more and more important in our daily life. The realization of soccer video shot segmentation is the basis of many advanced applications such as video content retrieval and video resource management. So, how to efficiently and accurately segment soccer video shot has become a hot spot in recent years.

Now, there are many studies on video segmentation, which also achieve fruitful results both at home and abroad. Ye[1] proposed to use the sub-block histogram as

[1] Qiao ZHANG (✉)

School of information engineering, Communication University of China, Beijing, China
e-mail: zhangqiaochn@163.com

Ping SHI (✉)

School of information engineering, Communication University of China, Beijing, China
e-mail: shiping@cuc.edu.cn

the characteristic value, together with the introduction of multi-scale wavelet to analyze the similar degree of two frame texture before video segmentation. Reference [2] suggests using fuzzy clustering algorithm for shot segmentation, and this method can obtain a relatively good result. Reference [3] extends the traditional dominant color to several semantic colors, and defines the color ratio feature, then an adaptive semantic color extraction algorithm is proposed, considering the temporal variations in the semantic color due to environment. Reference [4] presents a novel mechanism which capable of perform semantic shot segmentation without additional computational and description cost for semantic video content depiction process. Reference [5] proposes to use the geo-colorimetric descriptor of the points to characterize the frames, and this algorithm can avoid the quadratic calculation of conventional methods of scenes segmentation. Chen[6] proposes an adaptive threshold method, and in this method, wavelet transform algorithm is used to extract frame differences. As for the soccer video, Ekin[7] realizes the segmentation according to color ratio of soccer video, and proposed a fully automatic and computationally framework for analysis and summarization of soccer videos using cinematic object-based features, but this method has a high computational complexity.

In this paper, a fusion algorithm is introduced to segment video shot in soccer videos. The χ^2 histogram match based on gray histogram is used to calculate frame differences, and then sliding window algorithm with adaptive threshold and double sliding windows algorithm are combined to detect and segment the shot boundary. The theory of this algorithm is introduced in section 2. Then in section 3, soccer videos are chosen to evaluate the algorithm.

2 A Method of Shot Segmentation in Soccer Video Based on Fusion Algorithm

Soccer video shots are usually can be classified into long shots, mid-shots, close-up shots and other shots. The most transition types of the video shot are Cut and Gradual Transition. From the above introduction about video shot segmentation, we can see that algorithm based on a single criterion does not make it to achieve sufficient accuracy. Therefore, in this paper, the following three algorithm are combined together to detect the shot boundary and then segment the video: "χ^2 histogram match based on gray histogram", "sliding window algorithm with adaptive threshold", and "double sliding windows algorithm". The first step of this algorithm is to calculate frame difference, the second step is to determine the adaptive threshold based on the average frame difference, finally we detect and segment the Cut and Gradual Transition boundary according to the two sliding windows and two adaptive thresholds.

2.1 Characteristic Extraction

Frame difference in the same shot is usually similar, but it changes quickly in the shot boundary. So for the sake of a better accuracy, an adaptive threshold, which is calculated by local frame difference, is essential before shot segmentation. Gray histogram is one of the characteristics that can describe the distribution of the whole grayscale of one image pixel, and it is a reflection of the statistical characteristics of frames in a video. Gray histogram calculation can save the computing time, and can effectively improve the speed of shot segmentation, so we first calculate the gray information, which is defined as:

$$Gray = 0.3 \times R + 0.59 \times G + 0.11 \times B \qquad (1)$$

In the above formula R, G, B represent three color channels of a pixel in a frame. And then calculate the gray histogram according to the following formula:

$$D = \sum_{k=0}^{K-1} H[f(x,y,t),k] - H[f(x,y,t+1),k] \qquad (2)$$

And the χ^2 distance of two gray histogram of two adjacent frame is chose as the frame difference between two frames. The χ^2 distance is originated from "χ^2 histogram match based on gray histogram", which defined as following:

$$D = \sum_{k=0}^{k-1} \frac{\{H[f(x,y,t),k] - H[f(x,y,t+1),k]\}^2}{H[f(x,y,t+1),k]} \qquad (3)$$

Determine the similarity of two frames according to distance D, if D is greater than a certain threshold, a cut or gradual transition may be happened. χ^2 histogram match bases on gray histogram is a relatively perfect method in soccer video segmentation, and the result of it is better than the other methods.

2.2 Calculate the Threshold Based on Frame Difference

Liu[8] implement the adaptive thresholds in a fast coarse-to-fine video shot segmentation to decrease both the computational complexity and error detections caused by the camera motion effectively. Reference [9] proposes to use adaptive algorithm to compute the threshold, in this paper, the threshold is calculated based on the average frame difference of the entire lenses, when there is a large frame difference in a smoothly changed lens, it may contributes to a false detection. In order to adjust to the needs of soccer video segmentation, the local adaptive threshold must be adopted in the algorithm, which is determined by the change law of the frame difference. The principle of the algorithm is in following:

In a shot $S(S_1,S_2,S_3,....,S_n,...,S_m)$, S_n(n=1,2,3,…,m) represents the frame in the shot, and $(A_1,A_2,A_3,....,A_n,....,A_{m-1})$ stands for the frame difference between every two adjacent frames, and $A_n=S_{n+1}-S_n$. The frame difference between S_n and S_{n-1} is calculated as:

$$A_{n-1} = S_n - S_{n-1} \tag{4}$$

To improve the precision of the algorithm, two sliding windows are put into use, and the bigger one is used to determine the average value of the frame difference within the bigger window, while the smaller window is used to detect the shot segmentation point. The bigger window contains M frames, while the smaller windows contains N frames (M is greater than N), and the smaller window N is inside the bigger window M, and the ending frame of the two windows is the same. Calculate the average value of frames in the bigger window M:

$$mean = \frac{A_n + A_{n+1} + ... + A_{n+M}}{M} \tag{5}$$

Two adaptive thresholds are defined to detect the shot segmentation: T1 (5 times of the average value mean) and T2 (1.5 times of the average value mean) are used to detect the Cut and Gradual Transition.

2.3 Shot Segmentation in Soccer Video

When there is a shot change, frame difference A_n will change a lot. So after get the threshold T1 and T2, the maximum frame difference A_{max} in the smaller window N should be found. If A_{max} is greater than or equal to T1, then it is a cut shot boundary. And if A_{max} is smaller than T1 but greater than threshold T2, there may be the start frame or a frame in the middle of a Gradual Transition. If there are large frame differences that all of them are greater than T2, between the consecutive frames and the algebraic sum of the frame differences A, thus can be sure that the gradual shot boundary exists in this place. As shown in Figure 1:
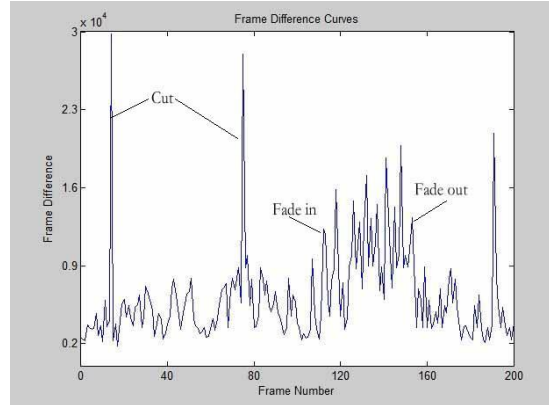

Fig. 1 The Change Curve of the Frame Difference

3 Experiment

In order to testify the effectiveness of the soccer video segmentation algorithm proposed in this paper, experiments are performed base on randomly selected soccer video material, which is obtained from the Internet, disk and other storage medium. Four video clips are chosen from two football competitions to compose the experimental data sets of the system. Each frame contains 720 x 576 pixels.

This paper makes a contrastive analysis of these two algorithms: Algorithm one is the traditional dual-threshold comparative method, which need to manually set the global detection threshold T1 (Cut detection threshold) and T2 (Gradient detection threshold) to segment the video. Algorithm two is our method that had been mentioned in this paper. Table 1 shows the comparison result of the testing between the two methods.

Table 1 Experimental Result

Video No.	Frame No.	Cut	Gradual	Algorithm One(Tradition)		Algorithm Two(New)	
				Precision Cut(Gradual)	Recall Cut(Gradual)	Precision Cut(Gradual)	Recall Cut(Gradual)
Clip1	6000	34	8	80.0% (71.4%)	82.4% (62.5%)	94.4% (87.5%)	100.0% (87.5%)
Clip2	6210	42	9	78.1% (75.0%)	76.2% (66.7%)	95.3% (77.8%)	97.6% (77.8%)
Clip3	11302	61	7	84.3% (60.0%)	81.7% (42.9%)	93.5% (71.4%)	95.1% (71.4%)
Clip4	24820	131	20	87.5% (68.4%)	80.2% (65.0%)	94.7% (84.2%)	96.2% (80.0%)
AVR	48332	268	44	84.3% (69.2%)	80.0% (61.4%)	94.9% (81.4%)	96.6% (79.5%)

The above statistical data of the experiment shows that the overall effect of the algorithm mentioned in this paper is better than the traditional dual-threshold comparison method. Experiment shows that the average recall rate of cut shot is 96.6% and the average precision is 94.9%; while the average recall rate of gradual shot is 79.5% and the average precision is 81.4%. This algorithm is relatively simple and easy to implement, and it can obtain significant effect with relatively low overhead

4 Conclusion

This paper introduced the basic algorithm on the typical shot segmentation algorithm first. And based on the analysis of the existing algorithm and the feature of soccer video, a fusion algorithm is put forward in this paper in order to avoid

the weakness of global threshold and the difficulty of choosing proper thresholds. First of all, utilizing the χ^2 histogram match method based on gray histogram to extract the frame difference, then calculate the adaptive threshold according to the mean frame differences within the big sliding window, and then using the small sliding window to detect whether this frame is the shot boundary of cut or a gradual transition. And then segment the shot after shot boundary detected.

On the whole, this shot segmentation method has a better robustness to light changes, scene changes, camera movement and so on. Almost all cut and most of the gradual transition can be detect using this fusion method.

References

1. Jun Ye, E Wang, Jianliang Li. (2007). Similarity Based Algorithm to Detect Video Shot. *Journal of Hebei Institute of Technology*, 29(2), 75~80.
2. Xinbo Gao, Xiaou Tang. (2002). Unsupervised video-shot segmentation and model-free anchorperson detection for news video story parsing. *IEEE Trans on CSVT*, 12(9), 765~776.
3. Zhengxin Niu, Xinbo Gao, Dacheng Tao and Xuelong Li. (2008). Semantic Video Shot Segmentation Based on Color Ratio Feature and SVM. *IEEE International Conference on Cyberworlds*, 157~162.
4. L. Ranathunga, R. Zainuddin and N. A. Adbullah. (2011). Conventional video shot segmentation to semantic shot segmentation. *IEEE International Conference on Industrial and Information Systems*, 186~191.
5. Chergui A., Bekkhoucha A., Sabbar W.. (2012). Video Scene Segmentation Using the Shot Transition Detection by Local Characterization of the Points of Interest. *IEEE International Conference on Sciences of Electronis, Techonologies of Information and Telecommunications*, 404~411.
6. Zhuoyi Chen, Mingji Zhou and Wanning Ding. (2011). Efficient algorithm on video shot segmentation of the adaptive threshold. *IEEE International Conference on Computer Science and Automation Engineering*, 4, 634~637.
7. Ekin A., Tekalp A., Mehrotra R. (2003). Automatic soccer video analysis and summarization. *IEEE Transactions on Image Processing*, 12(7), 796~ 807.
8. Liu Liu, Jian-Xun Li. (2010). A novel shot segmentation algorithm based on motion edge feature. *Symposium on Photonics and Optoelectronic*, 1~5.
9. WEN Jun, XIE yuxiang, LAO Songyang. (2004). Method of Highlights Analysis from Soccer Video. *Computer Engineering*, 30(6), 159~161.