HYBRID APPROACH FOR SHOT BOUNDARY DETECTION FROM UNCOMPRESSED VIDEO STREAM

[1]Ketan Chaudhari, [2]Santosh Sangada, [3]Upesh Patel, [4]Prof.J.P. Chaudhari, [5]Prof.P.I.Panchal
[1,2,3]PG Student, [4,5]Associate Professor: V.T.Patel Department of Electronics & Communication Engineering, CSPIT, CHARUSAT, Changa, India
[1]ketanec_esd2008@yahoo.com,[2]sangada_santosh@yahoo.com,[3]upeshpatel.ec@ecchanga.ac.in,[4]jitendrachaudhari.ec@ecchanga.ac.in,[5]pradippanchal.ec@ecchanga.ac.in

***Abstract*— To elaborate a video in terms of its content, it needs to be partitioned into its smallest visual unit called shot. To segment a video into shots, shot boundary is needed. Here hybrid approach adapted for uncompressed video shot transition detection. Color histogram is used for the cut and dissolve boundary. For cut boundary, traditional threshold technique works well but it fails in some cases like camera flash, fast zooming etc. To improve its performance 2nd derivative method is accepted for cut detection and for the dissolve boundary histogram difference with twin comparison method is used. Pixel intensity based approach is utilized for fade detection.**

Keywords*— *color histogram, gradual boundary detection, hard cut detection, shot boundary detection, twin comparison.

I. INTRODUCTION

In today's advance world, the rapid advance of multimedia and web technologies, video data in various formats are available. To enable efficient browsing, searching and retrieval with these huge video data resources, the video database systems are needed. The traditional method is time-consuming, lacks the speed because it uses human beings to manually explain the videos with text keywords. Therefore, more advanced methods are needed to support automatic indexing and retrieval directly based on videos content, which provide information related to video without consuming the time and with higher speed.

Shot boundary detection is one of the key techniques for digital video analysis. Video shot boundary detection is usually the first and important step for content-based video retrieval, which helps to segment a video by detecting boundaries between camera shots. A digital video sequence consists of group of scene. A scene is a collection of one or more shots focusing on one or more objects of interest. In other words it is a set of images (frames) taken from a single camera. A shot boundary separates two consecutive shots when one shot changes to another shot. It means that if *n* number of shots taken from different cameras with different angles and describing the same object or event, then such collection of *n* shots form a scene. Fig.1 illustrates the structure of video sequence.

Video shot boundary detection has various applications in different domains like video indexing, video compression, video access and others. Many techniques have been developed [4-6] and compared [8] to detect frame transitions in video sequences. One of the simplest ways of detecting shot transition is to compare the corresponding pixels between two consecutive frames. Another way is by using grayscale or color histograms of two frames. There are several other methods like edge changes and some predefined models, objects, regions to detect shot changes. Hybrid of these techniques has also been investigated [7].

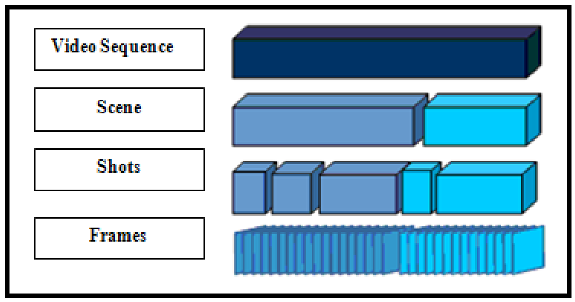

Fig.1 Illustration of video sequence

In this paper hybrid approach is presented to achieve shot boundary. Color histogram difference [7] for abrupt cut boundary and color histogram difference with twin comparison method [8] for dissolve boundary detection are presented. The simplicity of the method relies on the low complexity of the computation of the color histogram difference. For fade (out/in) pixel intensity [11] based method is presented because fade contain at least single monochrome frame and standard deviation of a monochrome frame is near about zero which make easy to identify a monochrome image from others.

II. SHOT BOUNDARY DETECTION

A. Introduction

Shot boundaries can be broadly classified into two types: abrupt transition and gradual transitions [1]. Abrupt transition is instantaneous transition from one shot to the subsequent shot. Gradual transition occurs over multiple frames, which is generated via the application of more elaborated editing effects involving several frames, so that F_k frame belongs to one shot, frame F_{k+1} to the second, and the k-1 frames in between represent a gradual transformation of F_k into F_{k+1}. Gradual transition can be further classified into fade out/in (FOI) transition; dissolve transition, wipe transition, and others transition, according to the characteristics of the different editing effects. As given in [1], they can be defined as follows:

- Cut transition: This is instantaneous transition in which frame F_k belongs to one shot and F_{k+1} to the next shot, a clear discontinuity therefore existing.
- Fade transition: This is a shot transition with the first shot gradually disappearing (fade out) before the second shot gradually appears (fade in). During the FOI, two shots are spatially and temporally well separated by some monochrome frames.
- Dissolve transition: This is a shot transition with the first shot gradually disappearing while the second shot gradually appears. In this case, the last few frames of the disappearing shot temporally overlap with the first few frames of the appearing shot.
- Wipe transition: This is actually a set of shot change techniques, where the appearing and disappearing shots coexist in different spatial regions of the intermediate video frames. One scene gradually enters across the view while another gradually leaves.
- Other transition types: There is a multitude of inventive special effects techniques used in motion pictures. They are very rare and difficult to detect.

B. Color Histogram Method

In case of reduce the sensitivity to camera and object movements; the histogram comparison method is preferred because if two frames with unchanging background and unchanging objects will have little difference in their histograms. Color histogram method is based on the computation of difference of color histogram between two consecutive frames as a measure of discontinuity. The sum of absolute difference between the bin values can be computed using following equation,

$$HD_{RGB}(X,Y)=\sum_{i=1}^{N}|h_x(i)-h_y(i)| \qquad (1)$$

where h_x is the color histogram of image X having N different bins. The difference between color histograms of frames belonging to a video sequence can be computed as

$$HistDiff[i]=\sum_{j=1}^{N}|h_i(j)-h_{i-1}(j)| \qquad (2)$$

Where, h_i is the color histogram with N bins of frame i corresponding to the video sequence and *HistDiff[i]* is the histogram difference of two consecutive frames.

C. Abrupt Cut Boundary Detection

For abrupt cut detection color histogram is calculated. Then threshold is determined using mean (μ) and standard deviation (σ) of calculated histogram difference. If the difference is greater than a threshold T_{cut} gives cut boundary. This scheme can achieve relatively good performance. But some time it can fail to detect correct boundary and can give false result. So here simple threshold method is used to get cut candidates only and second order derivative method to determine correct cut boundaries. Threshold can be determined using following equations.

$$\mu=\frac{1}{N-1}\sum_{i=1}^{N-1}D(i,i+1) \qquad (3)$$

$$\sigma=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}[D(i,i+1)-\mu]^2} \qquad (4)$$

$$T_{cut}=\mu+k\sigma \qquad (5)$$

Where, μ and σ are mean and standard deviation respectively and k is pre-specified constant. N is number of frames in video sequence.

Second Order Derivative Method

Due to illumination change or object/camera motion, a lot of feature variations within the same shot frequently exceed T_{cut}. Therefore, it causes many false alarms. To overcome this drawback second order derivative method [9] is adopted and can be determine using (6).

$$HD2=d(x_{k+1},x_{k+2})-d(x_k,x_{k+1}) \qquad (6)$$

where $HD2$ denotes the 2[nd] derivative of the histogram difference and $d(x_k,x_{k+1})$ denote the feature variation between the k^{th} and $(k+1)^{th}$ frame. In traditional method, feature variation $d(x_k,x_{k+1})$ is directly compared with T_{cut} . In second order derivative, instead of $d(x_k,x_{k+1})$, $(d(x_{k+1},x_{k+2})$ - $d(x_k,x_{k+1}))$ is compared with T_{cut}. This scheme not only works well but also can effectively eliminate the false positives.

D. Dissolve Boundary Detection

There are two passes for the dissolve detection. In first pass possible cuts are eliminated using high $T_{cut.}$ In the second pass starting frame F_s of the dissolve transition is detected using the lower threshold T_l. Once F_s is detected then compared it to subsequent frames called an accumulated comparison because during a gradual transition this difference value increases. By adding differences accumulated histogram can be determined. The end frame of the transition is detected if two constraints are satisfied. First is the difference between successive frames falls below T_l and second is the accumulated difference is greater than T_{cut} at the same time.

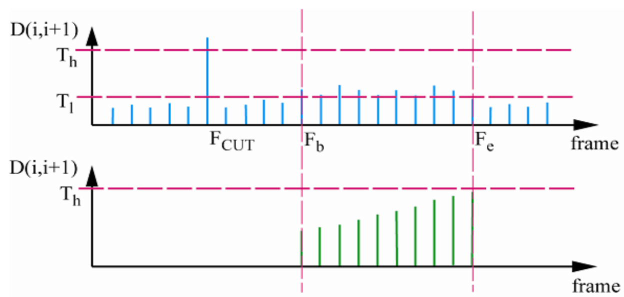

Fig.2 a) consecutive and b) accumulated histogram differences[8]

If consecutive difference falls below T_l before accumulated difference reaches T_{cut} start frame F_s will dropped.

E. Fade Boundary Detection

Monochrome frame is key feature of the fade transition. So the key problem of detecting FOI transition is to identify that monochrome frame. By using color histogram method it is difficult to identify the monochrome image because it produces wrong results during the night scenes. Because monochrome frame seldom appears elsewhere except within the FOI transition.

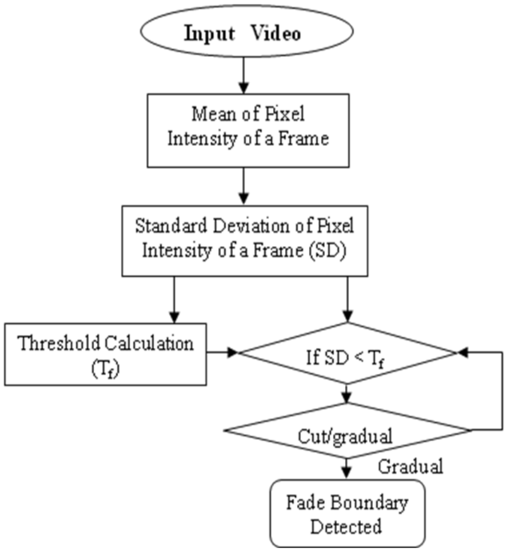

Fig.3 Flow chart for fade detection

There is at least one monochrome frame within the FOI transition. One of the most dominant characteristic of monochrome frame is its low standard deviation of pixel intensities. So here standard deviation feature is utilized in FOI detection process. Flow chart is as shown in fig(7).Threshold is based on standard deviation of all the frames. It can be calculated using (7):

$$T_f = \frac{\text{average of standard deviation} \times \text{scaling factor}}{100} \qquad (7)$$

Scaling factor is a variable according to the video. If SD of any frame finds below the threshold T_f, then that frame will be declare as monochrome frame and proceed for fade boundary detection. Fig.9 shows the results of detected monochrome frames using this method.

III. Experimental Results and Analysis

The performances of the implemented algorithms are evaluated based on the *recall* and *precision* criteria. *Recall* is defined as the percentage of desired items that are retrieved. *Precision* is defined as the percentage of retrieved items that are desired items [3]:

$$\text{Recall} = \frac{\text{Correct}}{\text{Correct} + \text{Missed}} \qquad (8)$$

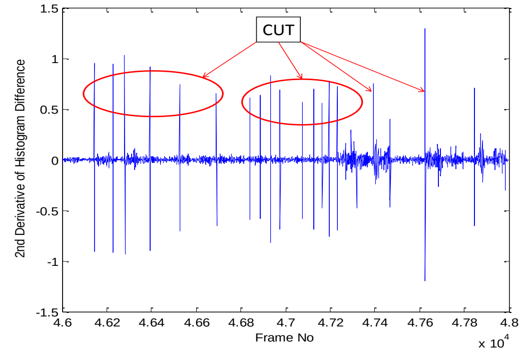

Fig.4 2nd order color histogram difference of a frame sequence with cut boundaries for uncompressed video

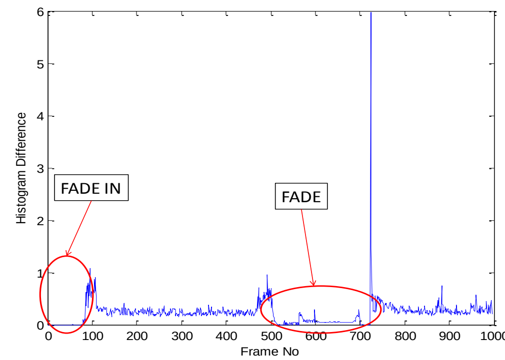

Fig.5 Color histogram difference of a frame sequence with Fade boundaries for uncompressed video

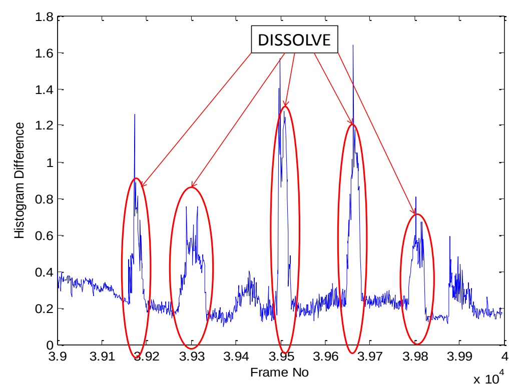

Fig.6 Color histogram difference of a frame sequence with Dissolve boundaries for uncompressed video

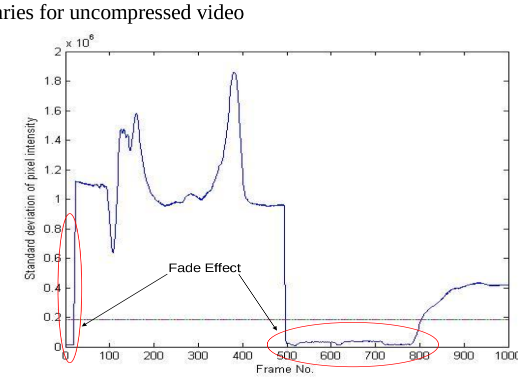

Fig.7 Fade transition detected using SD of pixel intensity for uncompressed video.

Table 1. Analysis of uncompressed video sequences

Movie	Frame Range	Effect	Uncompressed Video		
			R	P	F_1
Shrek 1	1 to 90000	Cut	0.96	0.94	0.95
		Gradual	0.80	0.84	0.82
The Last Airbender	1 to 90000	Cut	0.90	0.92	0.91
		Gradual	0.79	0.76	0.77
Die Hard 1	1 to 84000	Cut	0.90	0.87	0.88
		Gradual	0.8	0.66	0.72

$$\text{Precision} = \frac{\text{Correct}}{\text{Correct} + \text{FalsePositives}} \quad (9)$$

In order compare the overall performance of the algorithms, F_1 measure, which combines recall and precision results with equal weight, is adopted [3]:

$$F_1(recall, precision) = \frac{2 \times recall \times precision}{recall + precision} \quad (10)$$

The adapted video shot transition detection technique is evaluated by using 3 different kinds of videos for different duration video set. Two of them are normal movies and one is animated cartoon movie. All videos have been segmented and manually identified hard cuts and gradual transitions. The video clips were obtained mainly from the Internet and various television programs, and included various movie formats, such as AVI, MPEG. Different effects (eg. Cut, dissolve and fade) of shot boundaries are shown in figures from Fig.4-Fig.7 for uncompressed videos.

In case of all video sequence, total 2256 shot transitions existed; of which, 2196 were hard-cut transitions, and 60 were gradual shot boundaries. Table 1 show the detail analysis of uncompressed video sequences in terms of recall, precision and F_1 measure. For cartoon movie this approach performed well as compared to other two movies.

IV. Conclusions

In the color histogram difference based approach for the cut boundary detection is simple and robust in front of camera and moving object for uncompressed videos.

As compared to color histogram based approach pixel intensity based approach performed well for fades detection. Performance of twin comparison method for finding dissolve boundaries is also good for uncompressed videos.

In future performance of this algorithm can be improved and can be modified for compressed video.

Acknowledgment

We would like to thank specially to Prof. Pradip Panchal, Prof. J.P.Chaudhari and Prof. Hiren Mewada for their helpful guidance and insightful suggestions. We also thank to Mr.Y.A.Talati (Librarian) for providing us contents related to project, without which this work might not have been possible.

References

[1] Jun Li, Youdong Ding,Wei Li and Yunyu Shi *"DWT-based Shot Boundary Detection Using Support Vector Machine ",*International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM), ISSN: 2150-7988 Vol.1, pp.214-221, 2009.

[2] J.S.Boreczky and L.A.Rowe, *"Comparison of video shot boundary detection techniques",* Journal of Electronic Imaging, 5(2):122–128, April 1996.

[3] Emrah Asan,*"Video shot boundary detection by graph theoretic approaches"*, september 2008.

[4] C.Huang, H.Lee,C.Chen,*"Shot change detection via local keypoint matching",* IEEE Transactions on Multimedia 10 (6) (2008) 1097-108.

[5] J.Yuan, H.Wang, L.Xiao, W.Zheng, J.Li, F.Lin and B.Zhang, *"Aformal study of shot boundary detection",* IEEE Transactions on Circuits and Systems for Video Technology 17 (2) (2007) 168-186.

[6] O.Kucuktunc, U.Gudukbay, O.Ulusoy, *"Fuzzy color histogram based video segmentation",* Sep-2009

[7] Jordi Mas and Gabriel Fernandez, *"Video shot boundary detection based on color histogram"*, Digital television center, ramonLlull university Barcelona, Spain, Avaailable:http://www.nlpir.nist.gov/projects/tvpubs /tvpapers 03/ramonlull.paper.pdf.

[8] Matko Saric, Hrvoje Dujmic and Domagoj Baricevic,*" Shot Boundary Detection in Soccer Video using Twin-comparison Algorithm and Dominant Color Region"*, FESB, University of Split, Split, Available:http://www.foi.hr/CMS_home/znan_strucni_rad/konferencije/IIS/2007/papers/T09_05.pdf.

[9] Maheshkumar H. Kolekar and Somnath Sengupta, *"video shot boundary detection: a comparative study of three popular approaches",* electronics and Electrical Communication Engineering Department, Indian Institute of Technology, Kharagpur721302,INDIA,Available:*http://www.ncc.org.in/download.php?f=NCC2004/s17_05.pdf*.

[10] H.J. Zhang, A. Kankanhalli, S.W. Smoliar, "Automatic partitioning of full-motion video", Multimedia Systems 1(1) (1993) 10-28.

[11] Jinhui Yuan, Wujie Zheng, Le Chen, Dayong Ding, Dong Wang, Zijian Tong, Huiyi Wang, Jun Wu Jianmin Li, Fuzong Lin, Bo Zhang,*" Tsinghua University at TRECVID 2004: Shot Boundary Detection and High-level Feature Extraction"*, State Key Laboratory of Intelligent Technology and System Department of Computer Science and Technology Tsinghua University Beijing 100084, P. R. China.