Hybrid histogram approach for key frame extraction for different video stream

Jignesh P. Siroya1 Chetan R. Chauhan2
[1]PG Student, [2]Assitant Professor
[1, 2] Department of Computer Engineering
[1, 2] Noble Group of Institution Junagadh

Abstract--- Now a day, there are different kinds of videos available on internet like sports video, entertainment video, news video etc. Many researchers have done research on the video processing. Video abstraction, content searching, video summarization becomes crucial term for the development of the advanced digital video system. Key frame is the frame which contains the salient feature of the video so it is very useful for the video abstraction, summarization and content searching. In this paper, three techniques are discussed for the detection of shot boundary and key frame extraction for uncompressed video stream briefly and also compare these techniques. The three techniques here discussed are X^2 based Histogram Method, Edge Change Ratio and Pixel Based Differences.

Keywords:- Edge Change Ratio, Histogram (X^2 Based) Method, Key Frame Extraction, Pixel Difference Method, Shot boundary Detection

I. INTRODUCTION

Owing to the decreasing cost of storage devices, higher transmission rates, and improved compression techniques, digital video is becoming available at an ever increasing rate [7]. With the wide usage of video information, video abstraction and summarization becomes more important research for the researchers. Shot Boundary Detection (SBD) and Key frame extraction is very useful term for the video abstraction, summarization and content searching from the video. Key frame is a frame which gives the information about the whole shot and shot is a continuous frames captured by a single camera in single event. To find out the key frame from the video stream, first video is converted into different scenes. Those different scenes are converted into different shots and from that shot boundaries are detected and from that key frames are extracted for the input video stream.

There are many techniques for find out the shot boundary detection and from that shot boundary key frame is extracted. In this paper, three different techniques are discussed for the extraction of key frame. The three techniques include (1) X^2 based Histogram Method, which is based on the color histogram difference i.e. RGB differences (2) Edge Change Ratio, which is based on the difference between the edges of two pixels and (3) Pixel Based Differences, which is based on pair-wise difference between pixel intensities of consecutive frame.

The remainder of this paper is organized as follows. Section II, describes the types of shot i.e. cut and gradual. Section III, describes the algorithms for the three different techniques. Section IV, gives the experimental result for the different techniques and also gives the comparative analysis for the same. Section V, gives the conclusion drawn from the experiment performed.

II. TYPES OF SHOT TRANSITIONS

As earlier said, Video is made from the different scenes and different scenes are combination of the different shots. Shot can be defined as a frames captured by a single camera in single event. As shown in fig. 1, shot has two types of transitions i.e.Cut and Gradual.

A. Cut or Abrupt (discontinuous)

It is defined as when first shot is finished and directly start the second shot it is called Cut Transition.

B. Gradual (Continuous)

The gradual change occurs over multiple frames. It has four different types of transitions. Those are as follows:

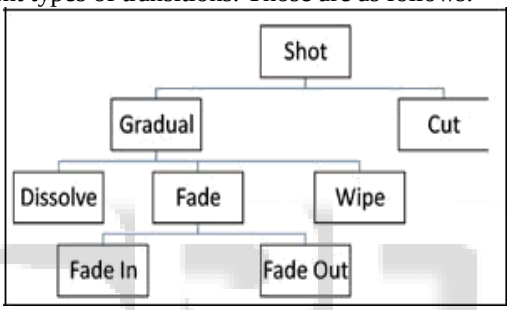

Fig. 1: Types of Shot Transitions

1. Fade in: Image gradually change from blank to current image is called fade in.
2. Fade out: Image gradually change from current image to blank is called fade out.
3. Dissolve: Image gradually changes between two distinct frames is called dissolve.
4. Wipe: A wipe occurs when a line moves across the screen, with the new scene appearing behind line [10].

Fig. 2: shows the different types of shot transitions.

Fig. 2(a): Cut Transition

Fig. 2(b): Fade in and Fade Out

Fig. 2(c): Dissolve

III. KEY FRAME EXTRACTION TECHNIQUES

A. Histogram(X^2) Based Method

There are six kinds of histogram match [4]. However, through comparing several kinds of histogram matching methods, Nagasaka [9] reached on conclusion that X^2 histogram outperformed others in Shot Boundary Recognition. Hence, X^2 histogram matching method is referred in this paper. In this method, different approach is used for the detection. Here the frame is converted into three different colors: Red, Green and Blue.

If there are total N number of frames then compute the Histogram of kth and (k+1)th frames for different three colors Hr, Hg and Hb, where Hr, Hg, and Hb are histogram of red, green and blue respectively. Now calculate the difference between two frames using equation (1).

$$D_f(k,k+1) = \sum_{i=1}^{3} \frac{[H(k,i) - H(k+1,i)]^2}{H(k,i)} \quad (1)$$

By using equation 1, the mean and standard deviation is calculated and from that threshold (T) is calculated by using equation 2. Here two thresholds are calculated for cut (TCUT) and gradual
(TGRD).

$$T = MD + STD \times A \quad (2)$$

Now find out the maximum difference between reference frame and all other frame within a one shot i.e. Max (i) = {D_f (1, k)}, where k=2,3,4…N. If the Max(i) > MD, then the frame with the maximum difference is called a key frame and otherwise with respect to the odd number of a shot's frames, the frame in the middle of shot is chose as key frame; in the case of the even number, any one frame between the two frames in the middle of shot can be chose as key frame [7].

B. Edge Change Ratio

Edge change ratio depends on the number of edges change between two successive frames. There are total six edges are there. Those are Canny, Sobel, Prewitt, Roberts, Laplacian and Zero cross. Among all of them, Canny edge is the best one. In this paper,Canny edge is used for finding out the shot detection.

For finding the edge change in video stream, first count the white point for two successive frames i.e. total edge data. After that, calculate the matched data between the two successive frames with respect to threshold distance r. Now calculate the entering and exiting edge pixel for the successive frames.

For two images, entering edge is the fraction of edge pixels in the second image that are greater than a distance r from the edges in the first image i.e. Eenter and exiting edge is the fraction of edge pixel in the first image that are greater than a distance r from the edges of the second image i.e. Eexit. From this value, calculate the maximum value as per the equation (3).

$$EC(n, n+1) = \max(E_{enter}, E_{exit}) \quad (3)$$

By using equation (3), the shot boundary is detected if the edge change ratio is greater than a threshold value. From the detected shot the key frame is extracted.

C. Pixel Difference Method

This is a very simple method to find out the shot transitions. Calculate the total number of pixel that differ between two successive frames and by using that calculate the mean and standard deviation for the same. Shahraray divided the images into 12 different regions and got the best match for each region in a neighborhood around the other image [12]. The formula for finding the difference between two frames is as follow in equation (4).

$$D_f(k,k+1) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[P_k(i,j) - P_{k+1}(i,j)]}{MN} \quad (4)$$

Where MxN is dimension of the image and Pk(i,j) is the intensity of the first frame and Pk+1(i,j) is the pixel intensity of the second frame. Then find out the mean of the total difference and also find out standard deviation of the difference. By using these two values,as per equation (2) find the two threshold i.e. high threshold (T_{HIGH}) for cut detection and low threshold (T_{LOW}) for gradual detection.

Now if the difference between two frames is greater than the mean difference then key frame is extracted. This method is good for cut or abrupt transition but this not as good in the gradual transition. This method is sensitive to camera motion or object. So this method is not work well in gradual transition.

IV. EXPERIMENTAL RESULT

All three algorithms are implemented on the MATLAB software. For the comparison, three different video stream is taken one is sports video, second one is news reader video and third one is the action scene from the movie. Comparison between all these three methods is done by the Recall and Precision value. Recall is the percentage of the desired items that are retrieved whereas Precision is the percentage of retrieved items that are desired items. Recall and Precision can be calculated by using the equation (5) and (6) respectively [2].

$$\text{Recall} = \frac{\text{Correct}}{\text{Correct + Missed}} \quad (5)$$

$$\text{Precision} = \frac{\text{Correct}}{\text{Correct + False positive}} \quad (6)$$

Table 1 shows the total number of frames, actual shot and key frame in the video stream for the all three kinds of video. Table 2 shows the comparative results between all three algorithms. Table 2 includes experimentally detected shot and key frame and also includes precision and recall value for all the methods.

Table 1: Actual Data

Types of Video	Number of Frames	Actual Shot Present	Actual Key frame Present
Sports Video	402	18	11
College Video	451	6	4
Action Movie	501	26	32

Notations used in the Table 2 are as follows:

MU : Method Used
VT : Video Type
SD : Shot Detected
KFD : Key Frame detected
CKF : Correct Key Frame
MKF : Missed Key Frame
FKF : False Key Frame
RKF : Recall Key Frame
PKF : Precision Key Frame
XH : X2Based Histogram
ECR : Edge Change Ratio
PD : Pixel Differences
S : Sports Video
C : College Video
M : Movie Video

Table 2: Experimental Data of Recall & Precision

MU	VT	SD	KFD	CKF	MKF	FKF	RKF	PKF
	S	25	9	8	3	1	72	88
XH	C	6	2	2	2	0	50	100
	M	38	33	31	1	2	97	94
	S	27	21	10	1	11	90	47
ECR	C	7	2	2	2	0	50	100
	M	40	34	30	2	4	93	88
	S	30	3	3	8	0	27	100
PD	C	8	2	2	2	0	50	100
	M	43	37	29	3	8	90	78

V. CONCLUSION & FUTURE SCOPE

All the methods are implemented on MATLB software and corresponding results are summarized in the Table 2. Edge Change Ratio method is good for the gradual detection. X^2 based Histogram method is very good for the detection of the cut transition. It is also good for the gradual transition but not as precise as the hard cut detection. Pixel difference method is simple to implement but it is sensitive to camera motion so it is not good as other two methods. The false positive is more in the pixel difference method and because of that the precision ratio is low for pixel method. Histogram and edge change ratio method is good compare to pixel difference method.

The key frame extraction can be done by other method like neural network which gives the less false positives and so that the result becomes more precise. Other method like graph method which used the Support Vector Machine (SVM) can also improve the result of X^2 based Histogram method.

REFERENCES

[1] ZHAO Guang-Sheng,"A Novel Approach for Shot Boundary Detection and Key Frames Extraction", 2008 International Conference on Multimedia and Information Technology, IEEE-2008.
[2] John S. Boreczky, Lawrence A. Rowe,"Comparison of video shot boundary detection techniques", Journal of Electronic Imaging, Vol. 5, No. 2, April 1996, pp. 122-128.
[3] A. Hanjalic,"Shot-boundary detection: Unraveled and resolved?", IEEE Circuits and System for Video Technology, Vol. 12, No. 2, February, 2002, pp. 90-105.
[4] Y. Cheng, X. Yang, D. Xu,"A method for shot boundary detection with automatic threshold", TENCON'02. Proceedings. 2002 IEEE Region 10 Conference on Computers, Communication, Control and Power Engineering[C], Vol. 1, October 2002, pp. 582-585.
[1] N. Babaguchi, Y. Kawai.T. Ogura, T. Kitahashi, "Personalized abstraction of broadcasted American football video by highlight selection", IEEE Transaction On Multimedia, Vol. 6, No. 4, August 2004, pp. 575-586.
[2] A. Hanjalic,"Multimodal approach to measuring excitement in video", Proceedings of International Conference on Multimedia and Expo ICME 03[C]. Vol. 2, July 2003, pp. 289-292.
[3] Naimish Thakar, Upesh Patel, Pradip Panchal, Hitesh Panchal,"Shot Boundary Detection and Key Frame Extraction from Uncompressed Video Stream", International Conference of Signals, Systems and Automation, ICSSA-2011.
[4] Z. Cernekova, I. Pitas, C Nikou,"Information Theory-based shot cut/fade detection and Video Summarization", IEEE Transaction on Circuits and System for Video Technology, Vol. 16, No. 1, January 2006, pp. 82-91.
[9] Nagasaka, Y. Tanaka, E. Knuth, L. Wegner,''Automatic video indexing and full-video search for object appearances", InVisual Database Systems II, pp. 113–127, Elsevier Science Pub-lishers ~1992.
[10] Seung-Hoon Han, Kuk-Jin Yoon, So Kweon,"A new technique for shot detection and key frames selection in histogram space", Workshop on Image Processing and Image Understanding, pp. 475-479, 2000.
[11] J. Boreczky,''Using video and audio data for content extraction and shot and scene boundary determination", Ph.D dissertation, University of California Berkeley, to appear ~1996.
[12] B. Shahraray,''Scene change detection and content-based sampling of video sequences", In Digital Video Compression: Algorithms and Technologies, Proc. SPIE2419, 2–13 ~1995. -

