Research Inventy: International Journal Of Engineering And Science
Issn: 2278-4721, Vol. 2 Issue 2(January 2013), Pp 17-21
Www.Researchinventy.Com

X^2 (Chi-Square) Based Shot Boundary Detection and Key Frame Extraction for Video

Prajesh V. Kathiriya[1], Dhaval S. Pipalia[2], Gaurav B. Vasani,
Alpesh J. Thesiya[4], Devendra J. Varanva[5]
[1](EC, Balaji Institute, India), [2,3,5](EC, Rk University, India),
[4](EC, LJ Institute, India),

Abstract : *To Extract valid information from video, without any loss of information, much attention is being paid to video processing technology. For this key frame is very useful technique. Extracting a small number of key frames that can abstract the content of video is very important for efficient browsing and retrieval in video databases. Consequently, technologies for video segmentation and key-frame extraction have become crucial for the development of advanced digital video systems. In this paper, Shot Boundary is detected by using very popular method i.e. Histogram method. By using this Shot Boundary the Key frame is extracted from the video.*

Keyword: *Histogram, Key Frame Extraction, Shot Boundary Detection,*

I. INTRODUCTION

In the recent years, use of video bases information in increasing more and more. Due to this many research is done in the area of video. 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. Key frame extraction is very useful in video summarization, searching a salient features from the video online video retrieval and abstract establishing. For finding a key frame from video, first find out the edge detection between two scenes i.e. shot boundary detection and then find a valid frame which contains the salient features or abstract of the video. Detection. Many persons have worked in the area of the video detection, video summarization an searching the main content of the video [3-6]. Among these approaches, shot transition detection is the first step of content-based video analysis and key frame is a simple yet efficient form of video abstract. Figure 1 shows the flow process for key frame extraction. As shown in figure 1, video stream is taken as an input and converts into different scenes and then it is into the different shots. After the detection of the shot boundary the valid key is extracted from the different frames.

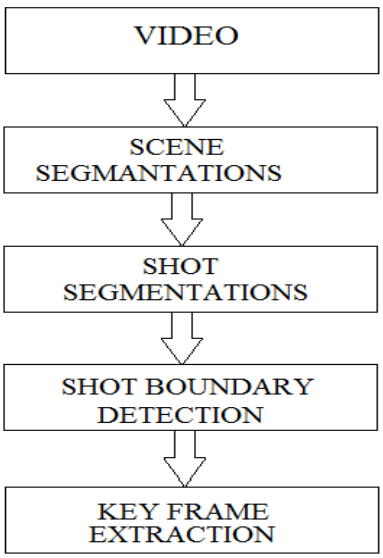

Fig.1 Flow of Process

II. TYPES OF SHOTS

Videos are made using different combination of shots. Many shots together make a different scene and different scenes together make a total video. Shot shows a continuous action in an image sequence. The consecutive frames from the start to the end of recording in a camera are called shot [1]. There are mainly two types of transitions which used in that field.

1. Cut or Abrupt (discontinuous): It is defined as finish first shot and direct start second shot its called Cut Transition.
2. Gradual (continuous): The gradual change occurs over multiple frames. It has four different types of transitions. Those are

A) Fade in: Image gradually change from blank to current image is called fade in.
B) Fade out: Image gradually change from current image to blank is called fade out.
C) Dissolve: Image gradually changes between two distinct frames is called dissolve.
D) Wipe: A wipe occurs when a line moves across the screen, with the new scene appearing behind line.

III. SHOT BOUNDARY DETECTION

3.1. Histogram Method:

There are six kinds of histogram match [8]. However, through comparing several kinds of histogram matching methods, Nagasaka [10] reached on conclusion that x2 histogram outperformed others in Shot Boundary Recognition. Hence, x2 histogram matching method is referred in this paper. Here we also used different approach for finding a difference of two frames because here we do not convert image into block but we just take whole image intensity for three basic colors. For that first of all, frames are converted into three different color i.e. R(Red), G(Green), B(Blue). The steps for histogram methods are as follows:

1) Compute the Histogram of k^{th} 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 (1):

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

Where H(k,i) and H(k+1,i) is stands for histogram of Red, Green and Blue for consecutive frames.

2) Calculate the total difference for the total video and then calculate the mean difference(MD) using (2):

$$MD = \sum_{k=1}^{N-1} \frac{D_f(k,k+1)}{N-1} \qquad (2)$$

Where N is total number of frames.

3) Compute standard variance STD of histogram difference over whole video sequence by using (3):

$$STD = \sqrt{\sum_{k=1}^{N-1} \frac{[D_f(k,k+1) - MD]^2}{N-1}} \qquad (3)$$

4) Calculate the two threshold for two types of shots, i.e. Cut and Gradual Threshold by using (4):

$$T = MD + STD \times A \qquad (4)$$

Where A is pre-specified constant for both Cut and Gradual transitions. So we get two thresholds T_{CUT} and T_{GRD} for cut and gradual transitions respectively.

5) Now if the mean difference of two consecutive frames is greater than Cut threshold (T_{CUT}), then Cut transition is occurred in video sequence. If the mean difference of two consecutive frames is greater than

Gradual threshold (T_{GRD}) and lower than Cut threshold (T_{CUT}), then Gradual transition is occurred in video sequence.

If $T_{GRD} < D_f(k, k+1) < T_{CUT}$, Then Gradual Transition occur.

If $D_f(k, k+1) > T_{CUT}$, Then Cut Transition occur

IV. KEY FRAME EXTRACTION

IV.1 Algorithm:

1) For finding a KEY frame from video, take first frame of each shot is reference frame and all other frames within shots are general frames. Computing the difference between all the general frames and reference frame in each shot with the above algorithm.[2,7]
2) Searching for the maximum difference within a shot: $Max(i) = \{D_f(1, k)\}$, k= 2, 3…N.
3) Now 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]

IV.2 Flow Chart:

The whole process of extracting a KEY frame from the video is as shown in fig.2 which shows the flow chart for the whole process.

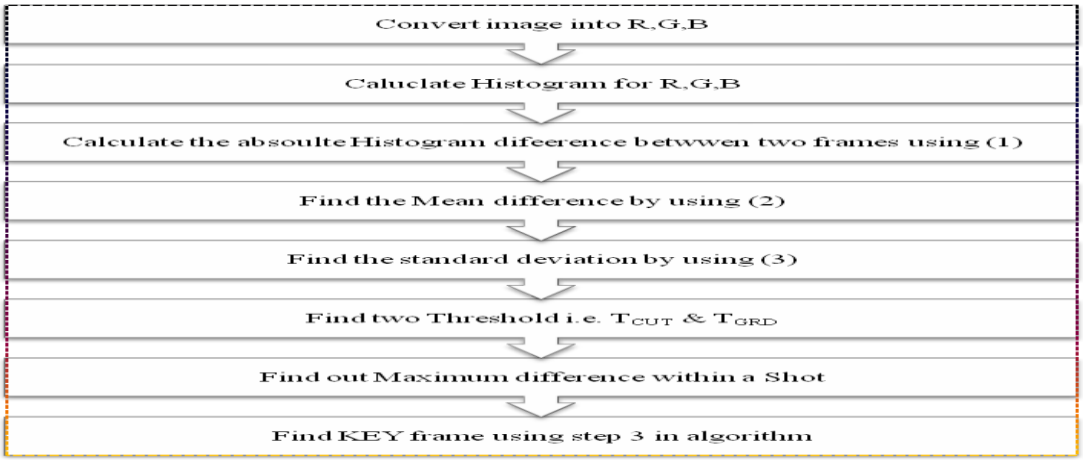

Fig. 2 Flow Chart

V. SIMULATION RESULT

We have implemented this algorithm in different types of uncompressed video stream. Here we show the results of two different videos, one is from cricket match and another one is an action sequence from movie Resident Evil 5.

V.I Cricket Video:

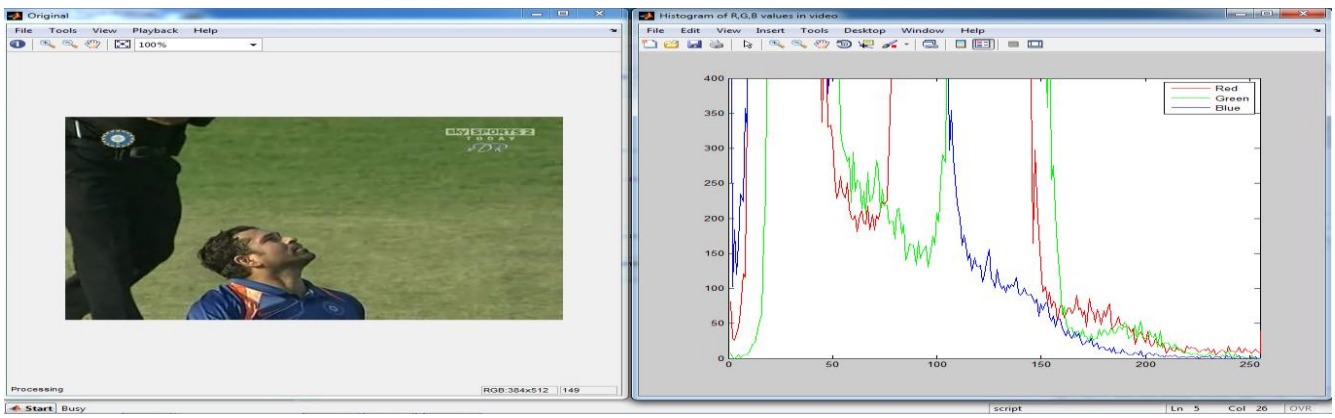

Fig. 3 Left side is showing that video is playing and right side shows Histogram of R,G,B

Fig. 3 shows the video is playing in left side part and corresponding Histogram is seen in right side window. Whereas fig. 4 shows the plot between Histogram differences versus total number of frames. At last fig. 5 shows extracted Key frames. In this video from 527 frames total 11 KEY frames are extracted.

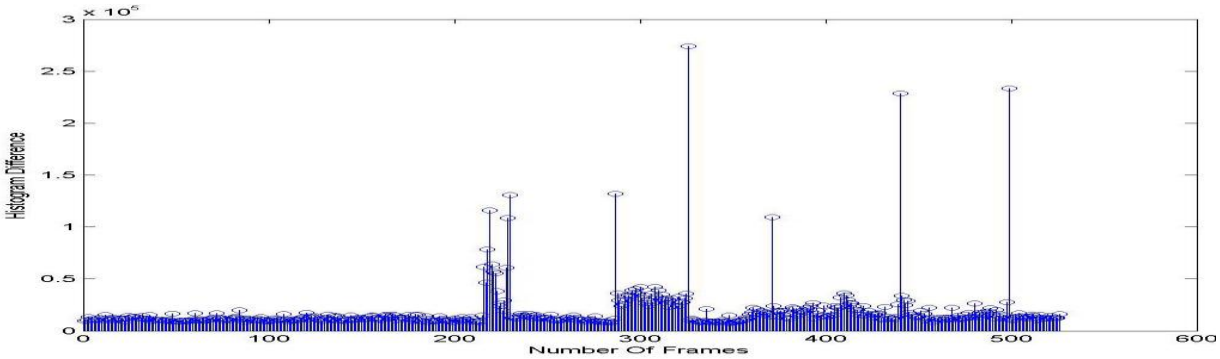

Fig. 4 Plot of Histogram Difference versus Number of Frames

Fig. 5 Extracted KEY Frames

V.II Action Movie:

This video is taken from the movie Resident Evil 5. Fig. 6 shows the plot between Histogram differences versus total number of frames and fig. 7 shows extracted Key frames. In this video from 500 frames total 38 KEY frames are extracted.

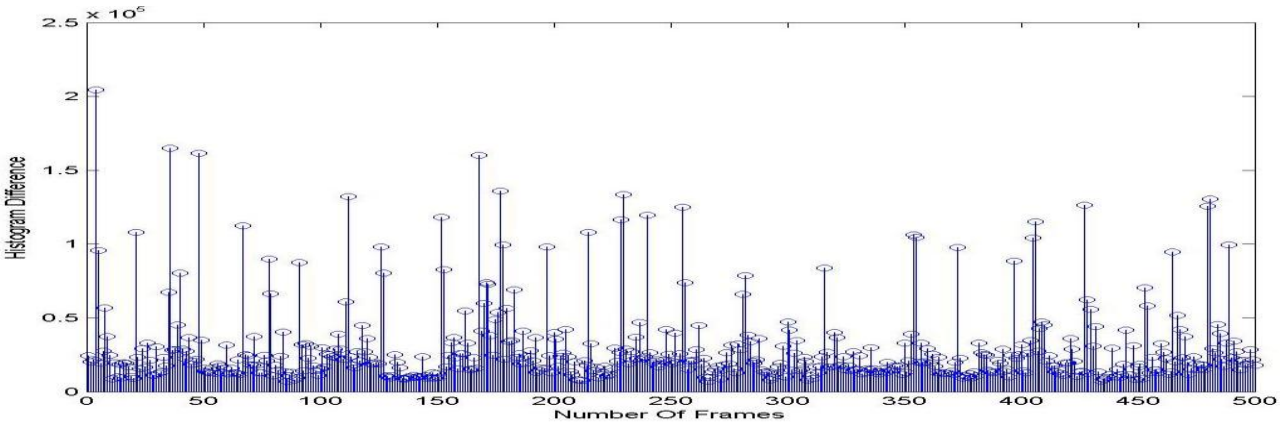

Fig. 6 Plot of Histogram Difference versus Number of Frames

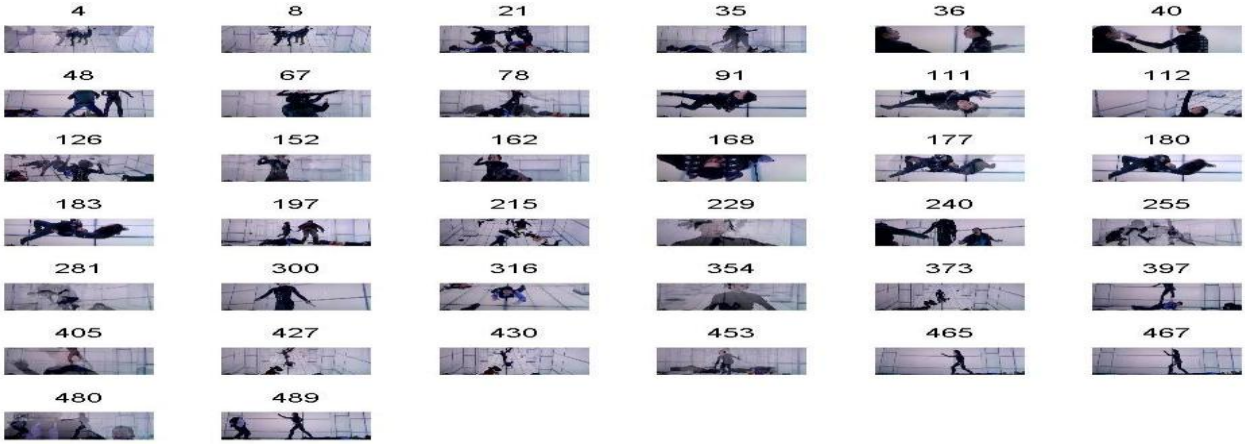

Fig. 7 Extracted KEY Frames

VI. CONCLUSION & FUTURE SCOPE

In this paper, we see Chi Square technique for color Histogram method to find a shot boundary from the video. After detecting a shot boundary we can extract the KEY frame from the video. Histogram method is very time consuming process but the accuracy is higher especially in Gradual transitions. So we can say that there is trade-off between speed and accuracy in Histogram method. This method is well supported to the uncompressed stream. In future this method can extend to all types of videos and also extend for any format of the video.

REFERENCES

[1] Seung-Hoon Han, Kuk-Jin Yoon, and In So Kweon, 2000. A new technique for shot detection and key frames selection in histogram space. Workshop on Image Processing and Image Understanding, pp 475-479.

[2] ZHAO Guang-sheng, "A Novel Approach for Shot Boundary Detection and Key Frames Extraction", 2008 International Conference on Multimedia and Information Technology, IEEE-2008.

[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] Z. Cernekova, I. Pitas, and 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.

[5] N. Babaguchi, Y. Kawai.T. Ogura, and 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.

[6] 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.

[7] Naimish Thakar,Upesh Patel, Pradip Panchal and Hitesh Panchal "Shot Boundary Detection and Key Frame Extraction from Uncompressed Video Stream", International Conference of Signals, Systems and Automation, ICSSA-2011, ISBN 978-1-6123-3002-0.

[8] Y. Cheng, X. Yang, and 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: 582-585.

[9] D. Swanberg, C. F. Shu, and R. Jain, ''Knowledge guided parsing and retrieval in video databases,'' in Storage and Retrieval for Image and Video Databases, Proc. SPIE 1908, 173–187 ~1993!.

[10] F. Arman, A. Hsu, and M-Y. Chiu, ''Image processing on encoded video sequences,'' Multimedia Systems 1~5!, 211–219 ~1994!.

[11] T. D. C. Little, G. Ahanger, R. J. Folz, J. F. Gibbon, F. W. Reeve, D. H. Schelleng, and D. Venkatesh, ''A digital on-demand video service supporting content-based queries,'' Proc. ACM Multimedia 93, pp. 427–436, Anaheim, CA ~1993!.

[12] J. Boreczky, ''Using video and audio data for content extraction and shot and scene boundary determination,'' PhD dissertation, University of California Berkeley, to appear ~1996!.