International Journal of Advances in Electronics Engineering
Vol:1 Issue:1 ISSN 2278 - 215X

Analysis and Verification of Shot Boundary Detection in Video using Block Based χ^2 Histogram Method

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

***Abstract*— video shot boundary detection, which segments a video by detecting boundaries between camera shots, is usually the first and important step for content-based video retrieval. This paper present block based χ^2 histogram method which is effective in detecting abrupt transitions and gradual transitions, respectively. A new pixel intensity based method is proposed for Fade transition detection.**

***Keywords*— Abrupt transition,Automatic Thresold,Gradual transition,Shot boundary detection, χ^2 histogram.**

I. INTRODUCTION

With the rapid advance of multimedia and Web technologies, video data in various formats are becoming available at an explosive rate. For example, based on a Yahoo! Answers post dated on June 2009, there were over 240,000,000 videos on YouTube, which is the most popular online video sharing Website. The time required to view all these videos was over 800 years. More amazingly, around 500,000 new videos were uploaded to YouTube everyday!

With such enormous video data resources, sophisticated video database systems are highly demanded to enable efficient browsing, searching and retrieval [1-4]. However, the traditional video indexing method, which uses human beings to manually annotate or tag videos with text keywords, is time-consuming, lacks the speed of automation and is hindered by too much human subjectivity. Therefore, more advanced approaches such as content-based video retrieval are needed to support automatic indexing and retrieval directly based on videos content, which provide efficient search with satisfactory responses to the scenes and objects that the user seeks.

Video shot boundary detection, which segments a video by detecting boundaries between camera shots, is usually the first and important step for content-based video retrieval. A video consists of a sequence of images (often being called frames), which can be played consecutively at the speed of around 20 to 30 frames per second in order to view smooth motion. To index and retrieval a video, shot boundary detection is usually conducted to segments the video into shots by detecting boundaries between camera shots.

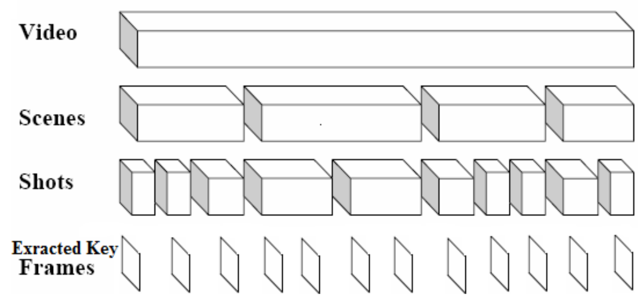

Figure.1. Overview of Shot boundary detection.[9]

A shot is defined as the consecutive frames from the start to the end of recording in a camera. It shows a continuous action in an image sequence [5]. There are two different types of transitions that can occur between shots, abrupt (discontinuous) also referred as cut, or gradual (continuous) such as fades, dissolves and wipes. The cut boundaries show an abrupt change in image intensity or color, while those of fades or dissolves show gradual changes between frames.

- A cut is an instantaneous transition from one scene to the next and it occurs over two frames.
- A fade is a gradual transition between a scene and a constant image (fade out) or between a constant image and a scene (fade in).
- A dissolve is a gradual transition from one scene to another, in which the first scene fades out and the second fades in.
- A wipe occurs as a line moves across the screen, with the new scene appearing behind the line.

II. SHOT BOUNDARY DETECTION

A. *Image Segmentation*

First, each frame is divided into nine blocks, B(1,1), B(1,2), B(1,3), B(2,1), B(2,2), B(2,3), B(3,1), B(3,2), B (3,3) . Then the difference of the corresponding blocks between two consecutive frames is computed. Finally, the final difference

Vol:1 Issue:1 ISSN 2278 - 215X

of two frames is obtained by adding up all the differences through different weights. Different position's pixels have different contribution to shot boundary detection: pixels on the edge are more important than others. Thus, different weights are given to blocks of different position. So more weights are assigned to corner blocks compared to other blocks.

B. Matching Difference

There are six kinds of histogram match [6]. Color histogram was used in computing the matching difference in most literatures. However, through comparing several kinds of histogram matching methods, Nagasaka reached a conclusion that χ^2 histogram outperformed others in shot boundary Recognition [10]. Hence, χ^2 histogram matching method is proposed in this paper.

III. ALGORITHMS DESCRIPTION

A. Shot boundary detection [7]

Let F (k) be the k th frame in video sequence, k =1, 2,.., Fv (Fv denotes the total number of frames in video）

$$D_B(k,k+1,i,j)=\sum_{l=0}^{L-1}\frac{[H(i,j,k)-H(i,j,k+1)]^2}{H(i,j,k)} \qquad (1)$$

Where, H (i, j, k) and H (i, j, k □1) stand for the histogram of blocks at (i, j) in the kth and (k □+1) th frame respectively and L is the number of gray in an image.

$$D(k,k+1)=\sum_{i=1}^{m}\sum_{j=1}^{n}w_{ij}D_B(k,k+1,i,j) \qquad (2)$$

Where m=n=3, w11=2, w12=1, w13=2, w21=1, w22=1, w23=1, w31=2,w32=1, w33=2.

$$MD=\sum_{k=1}^{F_v-1}\frac{D(k,k+1)}{F_v-1} \qquad (3)$$

$$STD=\sqrt{\frac{\sum_{k=1}^{F_v-1}(D(k,k+1)-MD)^2}{F_v-1}} \qquad (4)$$

Calculate Threshold based on Mean deviation and Standard deviation.

Threshold: T=MD +a ×STD (5)

Where **a** is weight factor that vary based on videos.

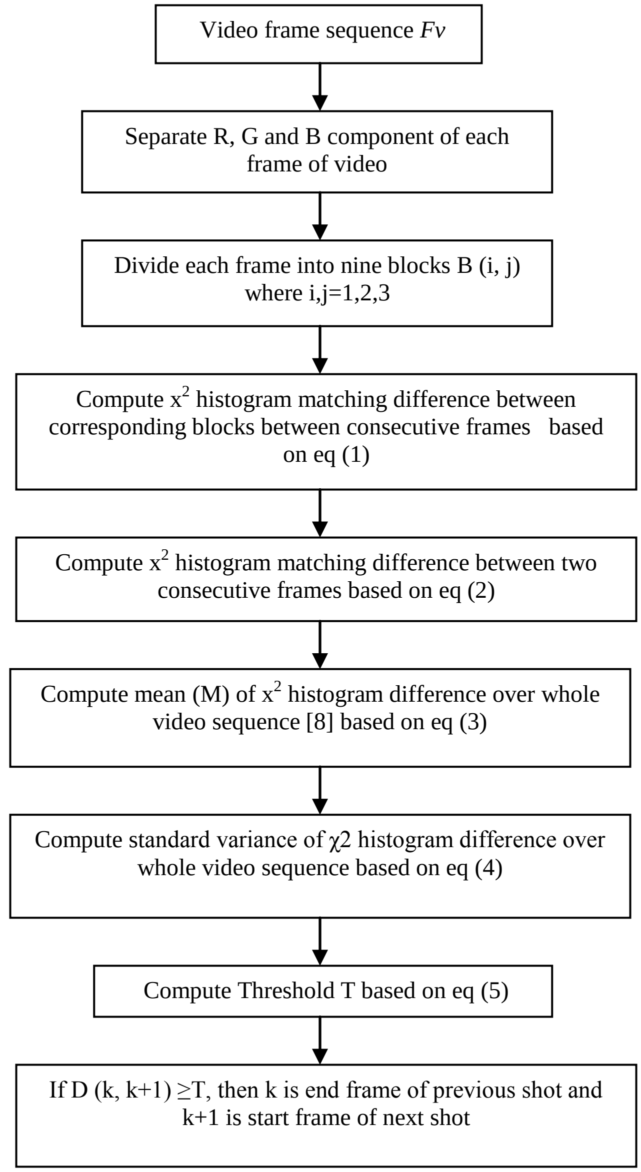

Figure.2. Flowchart of shot boundary detection algorithm.

B. Fade Transition Detection

In above procedure Cut, Fade and Dissolve transitions detected but fade transition is not accurately detected. So a new approach based on Pixel intensity of frame is proposed here.

In case of fade out/in (FOI), the frames of the first shot gradually get darken and disappear and then, the frames of the second shot gradually appear.

The key problem of FOI detection is the recognition of monochrome frame, since there is at least one monochrome frame within the FOI transition but monochrome frame seldom appears elsewhere. One dominant characteristic of monochrome frame is its low standard deviation of pixel intensities. Thus standard deviation feature is utilized in FOI detection process. FOI detection process is described as follows.

Step 1. Detect monochrome frame. A monochrome frame is declared once its standard deviation feature is below a given threshold, which is heuristically determined.

Step 2. Judge the type of entering transition, abrupt or gradual. If it is gradual, search the fade out boundary of the previous shot.

Step 3. Judge the type of exiting transition, abrupt or gradual. If it is gradual, track the fade in boundary of the next shot.

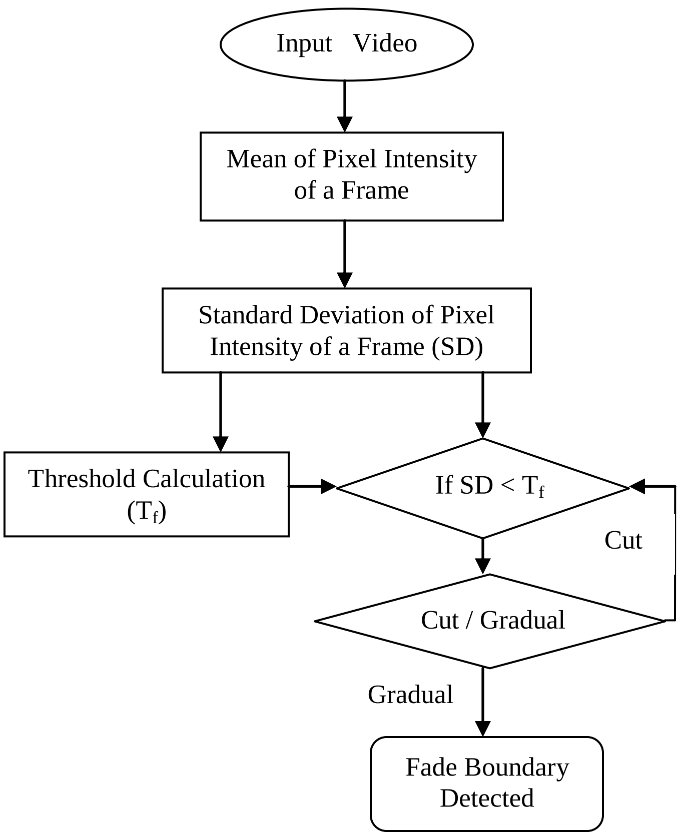

Figure.3. Flowchart of Fade transition detection algorithm.

Threshold is based on standard deviation of all the frames. It can be calculated using Equation (6):

$$T_f = \frac{\text{average of standard deviation} \times \text{scaling factor}}{100} \qquad (6)$$

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.

IV. SIMULATION RESULTS

The proposed video shot boundary detection algorithm is evaluated by using 3 different videos (movies) named Shrek, Die hard 1, Black hawk dawn for different duration. The video clips were obtained mainly from the Internet and various television programs, and included various movie formats, such as AVI (Audio Video Interleaved), MPEG (Motion Picture Export Group).

Following simulation results are plot of frame number versus χ^2 histogram matching difference between consecutive frame. Plots represent variation in χ^2 histogram matching difference for various effects like cut, fade, dissolve etc. Simulation results are obtained from different videos (movies) of Uncompressed and compressed type.

A. Shot boundary detection from uncompressed Videos.

Case 1: Video:: Shrek.avi (Uncompressed)

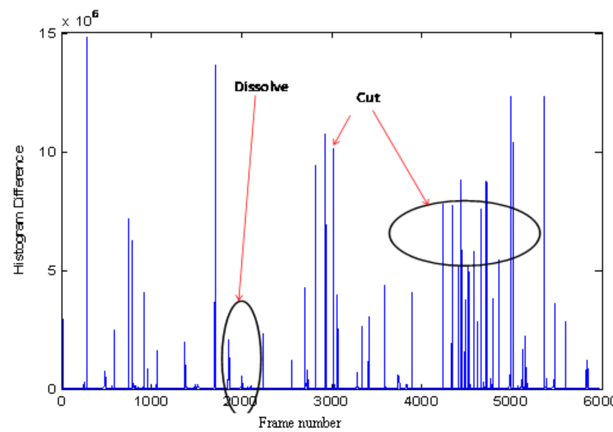

Figure.4: Cut boundary detection

Case 2: Video:: Shrek.avi (Uncompressed)

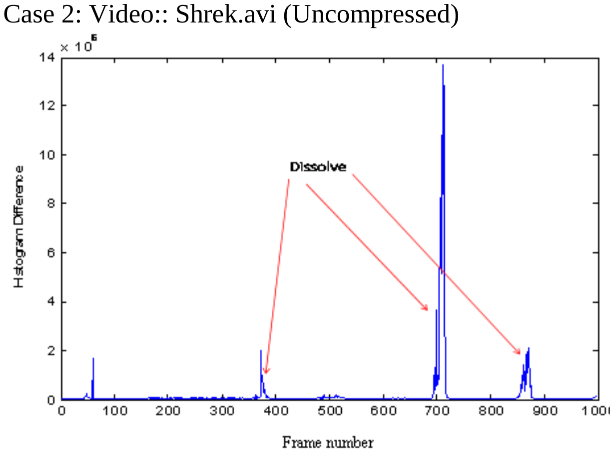

Figure. 5: Gradual boundary detection (Dissolve)

Case 3: Video:: Die hard 1.avi (Uncompressed)

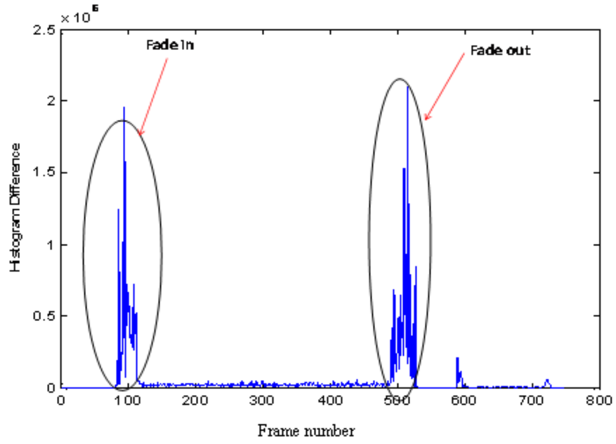

Figure.6: Gradual boundary detection (Fade).

B. Shot boundary detection from compressed Videos

Case 4: Video:: shrek.mpeg (compressed)

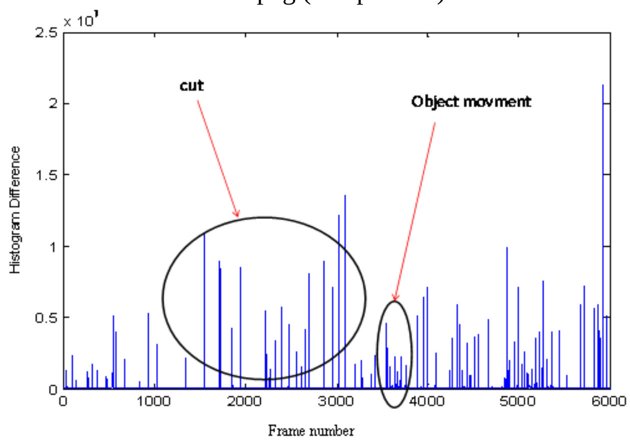

Figure.7: Cut boundary detection.

Case 5: Video:: Black dawn.mpeg(compressed)

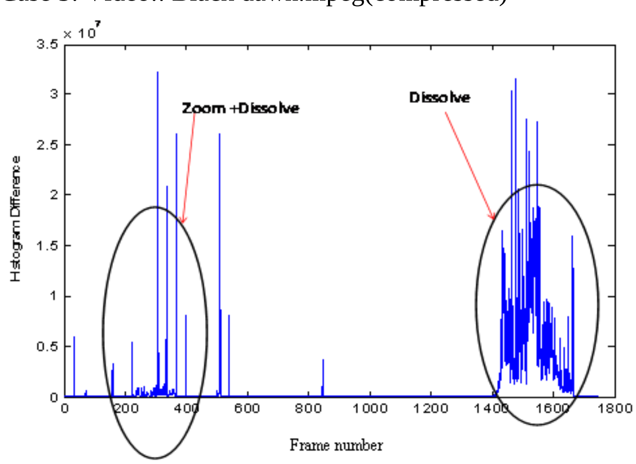

Figure 8: Gradual boundary detection (Dissolve)

Case 6: Video:: Die hard 1.mpeg (compressed)

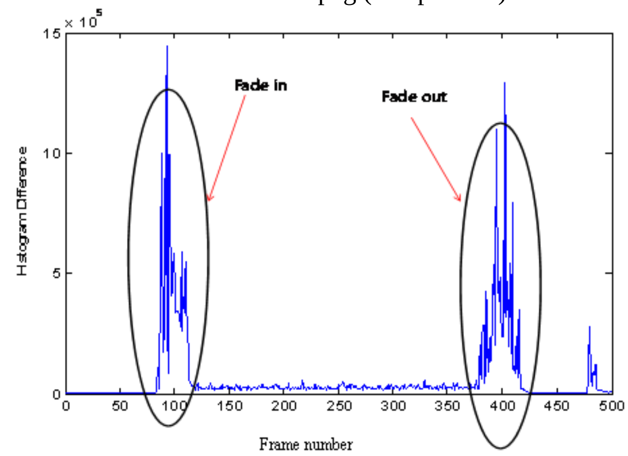

Figure.9: Gradual boundary detection (Fade)

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 [9].

$$\text{Recall} = \frac{\text{Correct}}{\text{Correct} + \text{Missed}} \qquad (7)$$

$$\text{Precision} = \frac{\text{Correct}}{\text{Correct} + \text{FalsePositives}} \qquad (8)$$

In order compare the overall performance of the algorithms, *F* measure, which combines recall and precision results with equal weight, is adopted [9].

$$F_1(\text{recall}, \text{precision}) = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \qquad (9)$$

In following table 12000 frames of three different movies are analyzed and Recall, Precision and F1 measure are calculated and compared [9].

TABLE 1 ANALYSIS OF BLOCK BASED χ^2 HISTOGRAM

Movie	Frame Range	Effect	Desired	Correct	Miss.	F. P.	Recall	Precision	F1 measure
Shrek	1 To 6000	Cut	57	52	5	4	0.91	0.92	0.92
		Dissolve	1	0	1	0	0	0	0
		Fade	4	3	1	0	0.75	0.75	0.75
	37800 To 45300	Cut	53	45	8	2	0.84	0.95	0.9
		Dissolve	9	7	2	1	0.77	0.87	0.85
The Last Air bender	31000 To 37000	Cut	28	28	0	4	1	0.87	0.93
		Dissolve	2	1	1	0	0.5	1	0.66
		Fade	3	3	0	0	1	1	1
	46000 To 52000	Cut	25	23	2	4	0.92	0.85	0.88
		Dissolve	6	4	2	0	0.66	1	0.8
Die Hard	1 To 6000	Cut	22	20	2	5	0.90	0.8	0.85
		Fade	5	2	3	1	0.4	0.66	0.5
	96000 To 102000	Cut	69	65	4	3	0.94	0.95	0.94

CONCLUSION

In this paper, we proposed Block based χ2 Histogram algorithm for shot boundary detection. We detect various shot boundaries like Cut, Fade and Dissolve with the help of χ2 histogram matching difference between consecutive frames and automatic threshold. For accurate fade in/out detection a pixel intensity based approach was used. Experimental results show that the proposed algorithm gives satisfactory performance for shot boundary detection.

ACKNOWLEDGEMENT

We would like to thank specially to Prof. Pradip Panchal, 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] A. Hanjalic, "Shot-boundary detection: Unraveled and resolved?", IEEE Transaction on Circuits and System for Video Technology., Vol.12, No.2, February, 2002, pp. 90-105.

[2] 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.

[3] N. Babaguchi, Y. Kawai,T. Ogura, and T. Kitahashi, "Personalized abstraction of broadcasted American football video by highlight

[4] 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.

[5] Seung-Hoon Han, Kuk-Jin Yoon, and In So Kweon, 2000. "A new technique for shot detection and key frames selection in histogram space."*12th Workshop on Image Processing and Image Understanding,* pp 475-479.

[6] 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.

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

[8] Y. Zhuang, Y. Rui, T.S. Huang, and S. Mehrotra, "Adaptive key frame extraction using unsupervised clustering", *Proceeding.ICIP'98[C], Chicago, IL*,1998, Vol.1, pp. 866-870.

[9] Emrah Asan,"Video shot boundary detection by graph theoretic approaches", September 2008.

selection", IEEE Transaction On Multimedia, Vol.6,No.4, August 2004, pp. 575-586.