International Journal of Computer Applications (0975 – 8887)
Volume 60– No.3, December 2012

Analysis of Popular Video Shot Boundary Detection Techniques in Uncompressed Domain

Swati D. Bendale,
Student Of ME Computer Engineering,
GTU, Gujrat, India.

Bijal.J.Talati,
Computer Engg Department,
SVIT, Vasad, Gujrat, India.

ABSTRACT
Now days there are tremendous amount of videos available on internet. Entertainment video, news video, sports video are accessed by users to fulfill their different needs. Our daily routine systems are also producing huge amount of videos for example surveillance system, shopping malls, home videos etc. These videos need to be accessed for different purposes. Current research topics on video includes video abstraction or summarization, video classification, video annotation, content based video retrieval. In nearly all these application one needs to identify shots and key frames in video which will correctly and briefly indicate the contents of video. This paper compares some of the popular shot boundary detection techniques in uncompressed domain. The merits and demerits of each of the techniques are also discussed. Some experiment done are also discussed.

General Terms
Popular shot detection techniques

Keywords
shot detection, cut; pixel based, block based, histogram based techniques,

1. INTRODUCTION
A video can be broken down in scene, shot and frames. A scene is a set of contiguous shots having a common semantic significance. A shot is a set of contiguous frames all acquired through a continuous camera recording. Each shot can be represented by one or more key frames. Key frames works as highlights, so that instead of searching whole video software can only search through these key frames to determine its relevance according to user's need or to briefly represent what the video is all about.

Gargi et al.[1] has evaluated 6 different color spaces using bin-to-bin, chi-square, histogram intersection, average color etc. frame difference method. Pascal et al.[2] has proposed method to detect hard cut, fade and dissolve transition from web video. They used color histogram to detect hard cut. They detected the centre of fades by thresh holding the strictly monotonic decreasing first derivative of the luminance variance. A dissolve is detected by variance of luminance. Zuzana and Ioannis [3] has used Mutual Information(MI) and (JE) between two successive frames for detecting shot break, calculated separately for each of the RGB components. Zhang et al. proposed a method called twin-comparison to detect gradual transitions using the color histogram difference [4].

Section 2 reviews different shot detection types. Section 3 reviews some popular shot detection techniques. Section 4 discusses the experiments performed on three different videos using some of the techniques discussed below. Section 5 gives conclusion drawn based on experiments performed.

2. SHOT TRANSITION TYPES
Fig. 1 shows different shot transition types. Shot transition is of basic two types: abrupt and gradual. Gradual transition is further classified as fade, dissolve and wipe. Abrupt transition occurs in single frame. A fade out is a slow decrease in brightness resulting in a black frame; a fade in is a gradual increase in intensity starting from a black image. Dissolves show one image superimposed on the other as the frames of the first shot get dimmer and those of the second one get brighter. The Hard cut is shown in Fig. 2 below, Fade is shown in Fig. 3 and dissolve is shown in Fig. 4 below. Detection of gradual cuts are more difficult than detection of hard cuts.

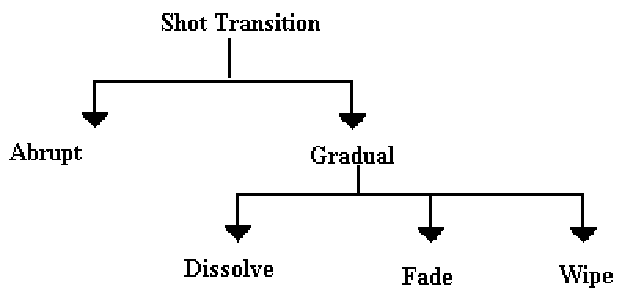

Fig. 1: Types Of shot transitions

Fig. 2: Hard Cut

Fig. 3: Fade

Fig. 4: Dissolve

3. SHOT DETECTION TECHNIQUES

3.1 Pixel based shot boundary detection

This method finds out pair-wise difference between pixel intensities of consecutive frame. Absolute sum of pixel difference can be used to find out difference between consecutive frames. If difference between frames is greater than particular threshold value then cut is detected. Following equation can be used to find out pixel based shot boundary detection, where absolute sum of pixel difference is used to find out difference between frame i and frame i+1 [5]. The formula (1) is used for pixel based shot boundary detection.

$$D(i,i+1)=\frac{\sum_{x=1}^{X}\sum_{y=1}^{Y}\left|P_i(x,y)-P_{i+1}(x,y)\right|}{XY} \qquad (1)$$

Where $P_i(x,y)$ is intensity of pixel at location (x,y) of frame i and $P_{i+1}(x,y)$ is intensity of pixel at location (x,y) of next consecutive frame, where image is of dimension X x Y. In case of RGB image difference between all three components at each position (x,y) have to be found out and then summation of their absolute difference can be taken, as in case of gray scale or binary image as above. If the difference D(I,i+1) is above some threshold then cut is detected. The threshold value can be selected either statically or dynamically. Dynamic threshold can be selected based on mean or standard deviation of video frames or combination of both [6]. WeimingHu et al. [7] and Alan Hanjalic [8] has discussed static and adaptive thresholding used in shot boundary detection techniques .This methods works well for abrupt cut detection but unable to detect gradual cut in videos. This technique is simple, easy to implement and sensitive to object or camera motion.

3.2 Block based shot boundary detection

In this technique each frame is divided in fixed number of blocks and difference between blocks at consecutive position in frame i and i+1 has been used to find out difference between frames. If this frame difference is greater than particular threshold value then break is detected. In [9] this method is described as statistical differences. The formula (2) is used for block based shot boundary detection.

$$D(i,i+1)=\sum_{k=1}^{b}C_k\left|DP(i,i+1,k)\right| \qquad (2)$$

Where each frame is divided into b blocks and DP(i,i+1,k) indicated difference of k^{th} block between i^{th} and $(i+1)^{th}$ frame and C_k is predetermined coefficient for block k. The absolute difference between all the blocks of two consecutive frames is added to find out the difference between the two frames. In [5] this method is described as likelihood ratio. Equation (1) is for calculating likelihood ratio.

$$LHR(i)=\frac{\left[\left(\frac{s_i+s_{i+1}}{2}\right)+\left(\frac{(\mu_i+\mu_{i+1})^2}{2}\right)\right]^2}{\sigma_i\times\sigma_{i+1}} \qquad (3)$$

Where,

μ_i , μ_{i+1} is the mean intensity and σ_i , σ_{i+1} is the variance of the frame i and i+1 successively.

This methods is relatively slow due to complexity of formulas. This methods can't identify dissolve, fade or fast moving objects. But computationally better than pixel based shot detection.

3.3 Histogram based shot boundary detection

This method computes gray or color histograms of the two consecutive frames of video. If the difference between the two histograms is above a threshold, a shot boundary is assumed.

$$D(i,i+1)=\sum_{j=1}^{n}\left|H_i(j)-H_{i+1}(j)\right| \qquad (4)$$

Equation (4) is used for histogram based shot boundary detection. Where j indicates the gray level value. Hi(j) is the histogram for the gray level j in the frame i and n is the total number of gray levels. This method is less sensitive to object an camera motion. Computationally less heavy. This method detects hard-cut, fade and dissolve and fails when there is large amount of motion. The histogram can be calculated in different color spaces such as RGB , HSV ,YCbCr , YIQ , XYZ , L*A*B*. Also we can use different measure to calculate the frame difference. Jen-Hao et al. [10] has used histogram based techniques to identify cuts in news video.

3.3.1 Bin to Bin Difference

The bin to bin difference is calculated with following formula (5),

$$fd_{b2b}=\frac{1}{2N}\sum_i\left|\mathrm{h}(i)-\mathrm{h}(i+1)]\right| \qquad (5)$$

The sum of the absolute difference of values in corresponding bins of histograms is used to calculate the shot break. If this sum is greater than some particular threshold then break is detected. In above formula N is the number of pixels in a frame [1].

3.3.2 *Chi-Square Difference*

The chi-Square difference is calculated with following formula (6),

$$fd_{chi} = \frac{1}{N^2}\sum_{i}\frac{|h_1(i) - h_2(i+1)|}{h_2(i)}\ h_2(i)\ ,\ 0 \qquad (6)$$

Here again N is number of pixels in frame. If this sum is greater than some particular threshold then break is detected [1].

3.3.3 *Histogram Intersection*

The histogram intersection is calculated using the formula (7),

$$fd_{\text{int}} = \frac{\sum_{i} MIN(h(i) - h(i+1))}{N} \qquad (7)$$

Where minimum value between each bin is selected for frame i and i+1 and the summation of this minimum values is normalized by total number of pixels. If this sum is greater than some particular threshold then break is detected [1].

Shimna and Kalpana [11] has identified shot first using color histogram and then identified key frames which are represented using feature vector. Swati and Ravi has [12] has discussed video shot boundary detection techniques using color histogram.

4. EXPERIMENT RESULT

Experiment done in MATLAB 2010 using some of above mentioned techniques. The result of the experiment is discussed here. When applied pixel to pixel method the difference between sequential frames is almost same for the below dataset, see Table I. Hence the method detected either too many shots which include lots of false shot break if low threshold value is kept or if threshold value is increased by some little extent no shots are detected. Even the time required to execute the method is too high. On the 3 GHz i5 processor with 3 GB RAM the method took 63.772 sec to execute for serial video. This is too high as compared to other methods. The method took near about 11 times more time to execute as compared to other methods. For all the below experiments except pixel to pixel method the threshold value used is $T = \mu + a \times \sigma$. Where μ is the mean and σ is the standard variance of difference of all video frames of video [6]. The video details used for experiment is as presented in Table I.

TABLE I.

VIDEO TEST DATA

Type of video	#number of frames	Actual shot present
cartoon	1836	30
commercial	1836	36
serial	1836	19

The value of a=3 is taken for experiments above, in all of the color spaces except for HSV. For bin-to-bin histogram using HSV method a=2 is taken. For experiments in histogram methods above the color space is divided in 32 bins except for HSV where color space is divided in 256 bins. These values are selected on basis of different experiments performed on above dataset. They are giving more accurate results for above dataset.

TABLE II

EXPERIMENT RESULTS

Method Used	Distance Measure Used	Color Space Used	Video	Cut Detec-ted
Block Comparion Method	Chi-Square	Gray Scale	cartoon	28
			commercial	38
			serial	14
Histogra Method	bin-to-bin	Gray Scale	cartoon	48
			commercial	35
			serial	9
Histogram Method	bin-to-bin	RGB	cartoon	41
			commercial	37
			serial	18
Histogram Method	bin-to-bin	HSV	cartoon	36
			commercial	38
			serial	18
Histogram Method	bin-to-bin	YCbCr	cartoon	40
			commercial	37
			serial	17

6. CONCLUSIONS AND FUTURE SCOPE

Based on above experiments conclusion drawn is that pixel to pixel method is computationally very heavy. Bin-to-bin histogram method in gray scale space and chi-square block comparison method have detected hard cuts. Bin-to-bin histogram methods in RGB and HSV color space have detected hard as well as dissolve cuts. Whereas bin-to-bin histogram method in YCbCr color space have detected hard, dissolve as well as fade cuts.

In case of hard cut the false hits are caused by dynamic scenes with strong object motion, blasts or fast camera pans. Block color histogram can is computationally heavier than global color histogram but it can be used when one has to give more weight age to some area of image as compared to others. A perfect shot boundary detection algorithm can be determines by analyzing the contents of video.

New techniques such as shot detection based on motion histogram, edge change ratio, clustering, fuzzy color histograms, neural network based shot detection techniques, mutual information based shot boundary techniques can be studied to check their performance against fast motion, dynamic camera pans .

7. REFERENCES

[1] R. Kasturi, S.H. Strayer, U. Gargi, S. Antani, "An Evaluation of Color Histogram Based Methods in Video Indexing", Department of Computer Science and Engineering Technical Report CSE-96-053, September 15, 1996.

[2] Pascal Kelm, Sebastian Schmiedeke, and Thomas Sikora, "Feature-Based Video Key Frame Extraction For Low Quality Video Sequences", WIAMIS 2009.

[3] Zuzana Cernekova, Ioannis Pitas, "Information Theory-Based Shot Cut/Fade Detection and Video Summarization", supported by the Commission of the European Communities in the framework of the Methods for Unified Multimedia Information Retrieval (MOUMIR) Project RTN-1999-00177 MOUMIR, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 16, NO. 1, JANUARY 2006

[4] H. J. Zhang, A. Kankanhalli, and S. Smoliar, "Automatic Partitioning of Full-Motion Video," MultimediaSystems, vol. 1, pp. 10-28, 1993.

[5] Krishna K. Warhade, Shabbier N. Merchant and U. B. Desai , " Performance Evalution of Shot Boundary Detection Metrics in the Presence of Object and Camera Motion" , IETE JOURNAL OF RESEARCH, vol 57 , issues 5 , Sep-oct 2011.

[6] Mr. Sandip T. Dhagdi, Dr. P.R. Deshmukh, " Keyframe Based Video Summarization Using Automatic Threshold & Edge Matching Rate**",** International Journal of Scientific and Research Publications, Volume 2, Issue 7, July 2012 , ISSN 2250-3153.

[7] WeimingHu, Li LI, Nianhua Xie , Xianglin Zeng, and Stephen Maybank, "A survey on Visual Content Based Video Indexing and Retrival", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-APPLICATIONS AND REVIEWS, Vol 41,No. 6, November 2011.

[8] Alan Hanjalic, "Shot-Boundary Detection: Unraveled and Resolved?", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 12, NO. 2, FEBRUARY 2002.

[9] John S. Boreczky, Lawrence A. Rowe, "Comparison of video shot boundary detection techniques", Journal of Electronic Imaging 5(2) , 122–128(April1996).

[10] Jen-Hao Yeh, Jun-Cheng Chen, Jin-Hau Kuo and Ja-Ling Wu, "TV Commercial Detection in News Program Video", Circuits and Systems, ISCAS, IEEE International Symposium, Vol. 5 pp. 4594 – 4597, 2005.

[11] Shimna Balakrishnan, Kalpana S. Thakre, "VIDEO MATCH ANALYSIS: A Comprehensive Content based Video Retrival System", International Journal of Computer Science and Application,issue 2010, ISSN 0974-0767.

[12] P. Swati Sowjanya , Ravi Mishra, "VIDEO SHOT BOUNDARY DETECTION- COMPASRISON OF COLOR HISTOGRAM GIST METHOD", IJREAS, Volume 2, Issue 2, February 2012 , ISSN 2249-3905.