Performance Comparison of Different Video Shot Boundary Detection Techniques

R.Gangeswari[1], A. John Pradeep Ebenezer[2]

[1]M.Phil Scholar, [2]Head of the Dept,

[1]PG & Research Department of Computer Science,[2]Department of Computer Applications,

St.Joseph's College of Arts and Science, Cuddalore, Tamil Nadu, India.

Email:[1]gangajaiakash13@gmail.com, [2]bobbyebe@gmail.com

ABSTRACT

Video picture processing is a methodology for manage video files in a method to switch the video data in a compelling as well as proficient way. It is a standout amongst the mainly well known dimensions in a video and image depend on technologies like monitoring. Shot change boundarydetection is one among the real study zones in record signal dispensation. Shot boundary detection (SBD) depends on the process of identifying image dissimilarities because of the transitions. It is highly useful in various scenarios such as indexing inside video database, video confining and so on. In the recent years, a number of SBD approaches were presented in this domain. This paper performs a performance analysis of three existing SBD approaches like pixel comparison, block based comparison and histogram comparison. The working principle of these three methodologies is explained and a comparison is also made based on its experimentalresults. At the end of the experimental analysis, it is reported that histogram comparison method is found to be effective over other methods.

Keywords: *Shot boundary; Pixel; Video; Histogram*

1. INTRODUCTION

In the recent decades, massive amount of digital videos are generated in each and every second in various areas like digital library, healthcare, e-learning, on demand video telecasting, TV and internet. The video files are commonly searched by the use of text or content. Upon retrieving videos by text, the videos areelucidated using words and textual keyword is employed to search videos. But, it consumes more time and it increases with the increase in size of database with higher amount of data. So, effective retrieval process is needed which should have the capability to produce better results than the conventional text based methods. The process of indexing, penetrating and retrieving videos as of hugedatabase such as YouTube, Netflix, hotstarand so onwill be highly effective when we partition the whole video file into segments (shots). A shot represented as anuninterruptedseries of frames captured by camera [1]. For the separation of a video to shots, first it needed to spot the shot limitationspresent within a video. Shot boundary detection (SBD) is anessential to index videos and browsing automatically. It is highly useful in various scenarios such as indexing into video database, video compression and so on. The fundamental block of every video file is a frame and its organization is represented in Fig. 1. The frame series are indexed through frame number. Once the video is broken, the attained frames will be of equal sizes. In general, 25 to 30 frames will be present in each second. A video shot is a series of interconnectedsuccessive frames captured by a distinct camera with no intervention and the shots are integrated to generate a video. An individual view is comprised of a distinct or many shots which represents a yarn unit inside a video.

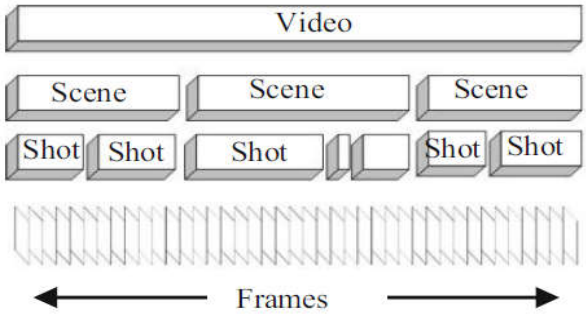

Fig 1. Structure of a Video

SBD is one among the real study zones in video signal processing. SBD depends on the process of identifying visual dissimilarities because of the transitions. The changeover of shots insidecan be classified into two kinds namely abruptin addition to gradual transitions. Abrupt transition also called as hard cuts or cuts which may present inside an individual frame while the camera terminates or restarts. Gradual transitions are tootermed as edit effects otherwise cinematic effects, commonly employed effects be fades, dissolve, wipes. A fade-in is steadyraise in intensity starts from a black frame toward bright frame. In contrast, fade-out is a consistent reduction in intensity starts form a bright frame and alsooutcome will be a black frame. Once one frame gets overlaid on another frame i.e. frames of earlier shot goes dim whereas the next shot will look bright and is termed as dissolve. A wipe is also an kind of distortion where one shot is replaced by a new shot by moving as ofindividual side of the frame to one more or with a special shape such as clock, rectangle, oval and so on [2]. In earlier days, more concentration has been given to cut identification and the concentration is shifted towards gradual transition identification. Generally, abrupt transitions are found to easy to identify when compared to gradual transition. Regardless ofthe numerous proposed methodologies and techniques so far, strong algorithms to detect different kinds of shot limitations i.e. stable quality of exposure for abrupt as well as gradual, have to be developed [3].

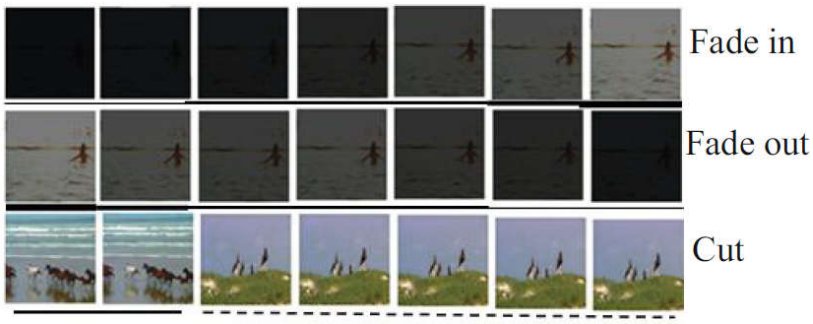

Fig 2.Gradual and Sudden Shot Change

In the recent years, a number of SBD approaches were presented in this domain. This paper performs a performance analysis of three existing SBD approaches like pixel comparison, block based comparison and histogram comparison. The working principle of these three methodologies is explained and a comparison is also made based on its experimental results. At the end of the experimental analysis, it is reported that histogram comparison method is found to be effective over other methods. The formulation of the left over part of the manuscript is given as follows: Section 2 explains the three SBD

techniques in detail. The experimental analysis part is given in Section 3 and the paper is ended with some observations and concluded in Section 4.

2. SHOT BOUNDARY DETECTION TECHNIQUES

The diversemethodologies for shot boundary detection are:

A. Pixel comparison:

Pair wise pixel relationship among two repeated frames examines the dissimilarity in intensity rate of the related pixel.

$$D(f, f+1) = \frac{\sum_{x=1}^{X}\sum_{y=1}^{Y}\left(\left(I_f(x,y)\right)\left(I_{f+1}(x,y)\right)\right)}{XY} \quad (1)$$

where f and $f+1$be two contiguous frames of dimension X×Y. If it is the strength estimation of pixel on coordinate (x, y) of frame f. Pixel wise correlation is restricted to object and camera movement. Because of a little transform in camera or object movement can bring about vast pixel divergence [4].

B. Block based correlation

Every frame is partitioned into n blocks plus every block is compared to the neighboring block of subsequent frame. A transition is affirmed if the quantity of altered blocks involving two back to back frames is more noteworthy than a specified threshold. Each frame is separated into blocks implythat all blocks is in use which brings about statistical image (lessened image). [5], [6] Mean Square Error of relating pixels is figured for neighboring statistical frames to locate the correct wipe transition area. They utilized Hough transform to decide the width of the strips in statistical image (single line or two lines). The estimation of normal inclination and amount of lines decides the wiping pattern [5]. In [7], system in view of pixel astute distinction between back to back frame is figured for wipe transition identification and a few approaches in light of horizontal, vertical as well as box-molded wiping design is measured. The X and Y direction assessment of limit line (horizontal, vertical, slanting, clock) connecting two adjoining frame is figured in [8]. Sugano M. et al [9] utilized methods to find abrupt,dissolve and wipes in compressed space. A frame is measured as abrupt shot boundary if total of forward forecast large scale blocks also intra full scale blocks is more noteworthy than a fixed threshold T_1 and sum of in reverse large scale blocks are lesser than T_2 where $T_1 > T_2$.(P and I frames). If there should be an occurrence of B frames, figure of forward large scale blocks is under T_3, and total of in reverse forecast and intra full scale blocks is more noteworthy than T_4 where $T_4 > T_3$.

C. Histogram examination

For computerized images, a shading histogram speaks to the quantity of pixels to facilitate to have settled hues in shading collection. A color histogram can work for color space similar to RGB, HSV, CMYK, YCbCr [10] and so on [4].

Worldwide histogram examination:

Histogram of two progressive frames is figured and compared to one another. When the variation in histogram is higher than a fixed threshold, then a transition occurs.

$$D(f, f+1) = \sum_{i=1}^{n}\left|H_f(i) - H_{f+1}(i)\right| \quad (2)$$

where $H_f(i)$ is the histogram rate for gray scale of frame f and n is the overall grayscale count.

Local histogram comparison

All frames areseparated into n blocks and histogram of every block undergoes a comparison with histogram of equivalent block of subsequent frame.

$$D(f, f+1) = \sum_{k=1}^{x} \sum_{i=1}^{n} \left(H_f(i,k) - H_{f+1}(i,k) \right) \quad (3)$$

Where $H_f(i,k)$ is the histogram value for grayscale$i$ for block k of frame f and x is the total number of blocks.

Xue L. et al [11] anticipated an algorithm that enhances execution by taking out smooth intervals from video. Highlights, for example, pixel wise variation, HSV color histogram in addition to edge histogram are of the new-fangled video grouping are separated and specified as input vectors to support vector machine. The yields of SVM are arranged into three classifications as abrupt, gradual and so forth. Utilizing HSV color histogram, contrast and adaptive threshold hard cuts are renowned in [12] [9]. They additionally ascertained the nearby histogram contrast and neighborhood adaptive threshold for gradual shot transition location.

Histogram is separated from every video frames and a matrix is made from the histogram value as the segment of the matrix and Singular rate Decomposition is connected to the matrix in this manner decreasing the component vector and gives quick computation. Comparability measures like Euclidean along with cos distance utilized to locate the abrupt and gradual transitions. A reversed triangle pattern matching is also utilized to locate the gradual transitions. An adjustment to straightforward histogram correlation strategy is introduced in [13]. At first, I frames are separated from the MPEG video stream. Intensity, row/horizontal, column/vertical histograms are registered for each and every one ofthe I frames and undergone comparison by chi-square test. The two calculations works in the compacted area, requiring just fractional unraveling of the packed video stream. The initial move towards video scene division and indexing is shot location as depicted in [14]. They utilized lookalike threshold method which means anelevated threshold *Th* and a near to the ground threshold *Tl* is chosen and if dim level histogram contrast, *D* of two adjacent frames is bigger than *Th*, a cut is proclaimed. In the event that D is bigger than *Tl* and not as much as *Th*, they are gathered.Presently, if collective distinction is more prominent than *Th*, a gradual transition exists.

Joyce and Liu[15] exhibited two calculations for distinguishing dissolve and wipes. The former is dissolve discovery technique which is executed both as a basic threshold-based finder and as a parametric indicator by modeling the error properties of the removed insights. The second is a calculation to identify wipes in view of image histogram attributes during transitions. Pardo A. [16] projectedan algorithm for hard cut discovery that gives preferable outcomes over component based, pixel based or basic histogram based methodologies. At first,inter frame histogram variation between frames are computed for an arrangement of bins. Likelihood for inter frame outline distinction to be more noteworthy than a predefined threshold is set and validated. In the event that a shot transform happens, histogram contrast is relied upon to be increasingly while for no shot variation, histogram distinction is required to be not as much of and in concurrence with past histogram distinction.

3. PERFORMANCE EVALUATION

3.1. Shot boundary detection experiment

For the validation of different SBD techniques, a set of test videos from TRECVID 2001 is employed. The details of the test videos such as number of frames, total number of transitions, number of cuts and gradual transitions are tabulated in Table 1. All these video sequences are transformed into the

uncompressed AVI set-up with a resolution of 320×240 pixels and a measurement lengthwise of total 112804 with a total duration of 3780 seconds in total.

3.2. Performance measures

The three SBD approaches are compared to one another interms of dissimilar performance measures. Precision, recall and F-measure are the commonly used metrics to investigate the results of SBD techniques. Recall indicates the performance of the technique on the basis of 'How many wipes were observed manually in the video and how many were precisely identified by the automated algorithm. Precision represents the accuracy of the algorithm when faced with issue of reducing false positives identified by the algorithm. False positives are those identified which are not present in the video but identified as wipes by the algorithm. False positives are the performance hampering detections occurred because of noise, object and camera motion.

Table 1 Description of Test Video Files

Videos	Frames	Transitions			Sources
		Total	Cut	Grad.	
anni009	12307	103	38	65	NASA 25th anniversary show, Segment 09
anni005	11364	65	38	27	NASA 25th anniversary show, Segment 09
NAD53	25783	159	83	76	Report #260
NAD57	12781	67	44	23	Report # 264

The measures recall and precision can be expressed as follows

$$Recall = \frac{No.of\ frames\ correctly\ identiifed}{No.of\ frames\ correctly\ identiifed + No.of\ frames\ for\ missed\ detections} \quad (4)$$

$$Precision = \frac{No.of\ frames\ correctly\ identiifed}{No.of\ frames\ correctly\ identiifed + No.of\ frames\ for\ false\ detections} \quad (5)$$

Additionally, F1 measure is also employed to analyze the performance of the SBD approaches. F1 measure is harmonic mean value that takes care of recall and precision in a similar way. F1 measure can be computed as follows

$$F1 - measure = 2 * \frac{Recall * Precision}{Recall + Precision} \quad (6)$$

3.3. Results and discussion

In this section, the results obtained by the three SBD approaches on the same set of test video files on the basis of recall, precision and F1 measure are analyzed in detail. The comparative results of the cut detection of the SBD approaches are tabulated in Table 2 and the precision values are illustrated in Fig. 3. From the table values, it is clear that the precision, recall and F1 measure values of the pixel difference

method are lower which implies poor SBD performance. At the same time, block based method showed better results over pixel comparison method on all the applied test videos. However, it shows inefficiency to outperform histogram based scheme. The histogram based scheme obtained better results and it implies that the histogram based scheme is the better selection to detection shot boundaries in any videos.

Table 2 Comparative results of cut detection of the SBD approaches

	Precision			Recall			F1 measure		
	Pixel	Block	histogram	Pixel	Block	histogram	Pixel	Block	Histogram
anni009	0.767	0.821	0.931	0.612	0.748	0.894	0.754	0.821	0.912
anni005	0.815	0.846	0.968	0.785	0.793	0.965	0.821	0.864	0.965
NAD53	0.715	0.789	0.914	0.654	0.812	0.831	0.816	0.811	0.955
NAD57	0.856	0.887	0.984	0.814	0.848	0.974	0.856	0.894	0.989

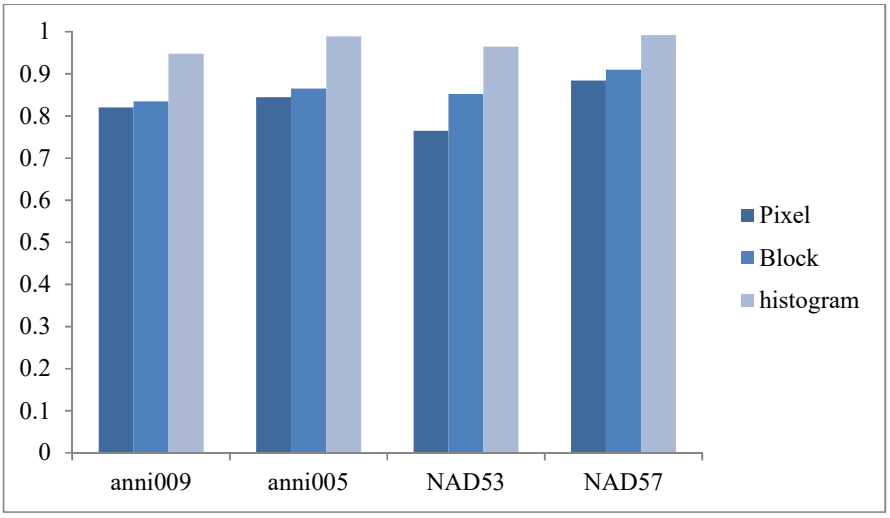

Fig 3. Comparative Results of Cut Detection In terms of Precision

Similarly, the comparative results of the gradual transition detection of the SBD approaches are given in Table 3 and the comparison of precision values are depicted in Fig. 4. From the table values, it is apparent that the pixel difference method showed worst performance interms of precision, recall and F1 measure when compared to block based and histogram based methods. By contrast, block based method attains higher values than pixel comparison method and it fails to show better performance than the histogram based means. On all the applied test videos, the experimental values indicated that histogram based method is the effective techniques for efficient SBD on any videos.

Table 3 Comparative Results of Gradual Transition Detection of The SBD Approaches

	Precision			Recall			F1 measure		
	Pixel	Block	histogram	Pixel	Block	histogram	Pixel	Block	Histogram
anni009	0.821	0.835	0.948	0.765	0.795	0.901	0.789	0.856	0.935
anni005	0.845	0.865	0.989	0.812	0.825	0.991	0.859	0.892	0.978
NAD53	0.765	0.852	0.965	0.795	0.895	0.898	0.868	0.889	0.965
NAD57	0.884	0.910	0.992	0.845	0.886	0.989	0.892	0.908	0.994

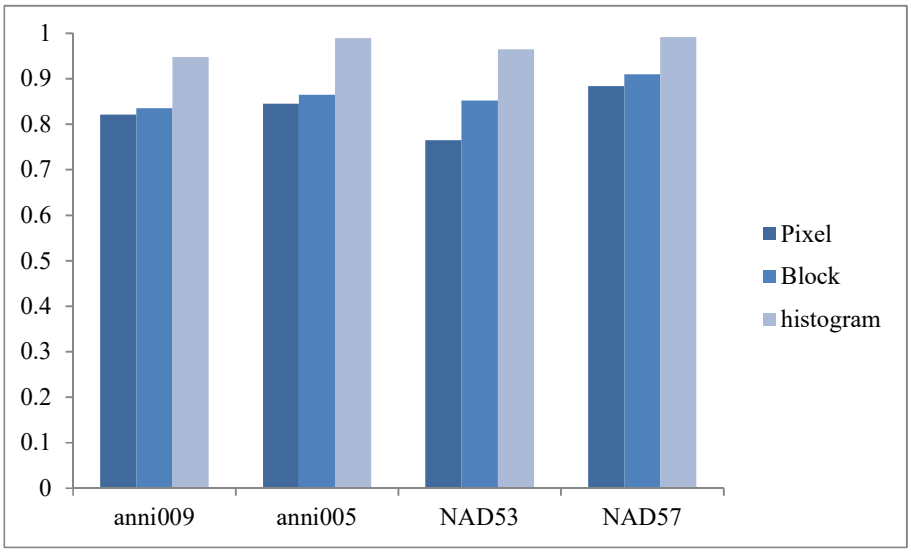

Fig 4. Comparative Results Of Gradual Transition Detection Interms Of Precision

4. CONCLUSION

To segment a video into many shots, it is first needed to find the shot boundaries present in a video. SBD is essential to index videos and browsing automatically. This paper performs a performance analysis of three existing SBD approaches like pixel comparison, block based comparison and histogram comparison. The working principle of these three methodologies is explained and a comparison is also made based on its experimental results. For the validation of different SBD techniques, a set of test videos from TRECVID 2001 is employed. The three SBD approaches are compared to one another interms of different precision, recall and F-measure. From the experimental results, it is evident that the histogram based process obtained superior results and it implies that the histogram based method is the superior choice to detection shot boundaries in any videos.

REFERENCES

[1] Chung-Lin Huang; Bing-Yao Liao, "A robust scene-change detection method for video segmentation," Circuits and Systems for Video Technology, IEEE Transactions on, vol.1l, no.12, pp.1281,1288, Dec 2001.

[2] Koprinska, 1., &Carrato, S. (2001). Temporal video segmentation: A survey. Signal processing: Image communication, 16(5), 477-500.

[3] Hanjalic, A, "Shot-boundary detection: unraveled and resolved?," Circuits and Systems for Video Technology, IEEE Transactions on , vo1.12, no.2, pp.90, I 05, Feb 2002.

[4] Jadon, R. S., Santanu Chaudhury, and K. K. Biswas. "A fuzzy theoretic approach for video segmentation using syntactic features." Pattern Recognition Letters 22.13 (2001): 1359-1369.

[5] Fernando, W.AC.; Canagarajah, C.N.; Bull, D.R., "Wipe scene changedetection in video sequences," Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on, vol.3, no., pp.294,298 vol.3, 1999

[6] Chavan, S., Fanan, S., &Akojwar, S. (2014). An Efficient Method for Detection of Wipes in Presence of Object and Camera Motion

[7] Min Wu; Wolf, W.; Liu, B., "An algorithm for wipe detection," image Processing, 1998. iCiP 98. Proceedings. 1998 international Conference on, voU, no., pp.893,897 voU, 4-7 Oct 1998

[8] Campisi, P.; Neri, A.; Sorgi, L., "Wipe effect detection for video sequences," Multimedia Signal Processing, 2002 iEEE Workshop on, vol., no., pp.161,164, 9-11 Dec. 2002

[9] Sugano, M.; Nakajima, Y.; Yanagihara, H.; Yoneyama, A, "A fast scene change detection on MPEG coding parameter domain," image Processing, 1998. iCiP 98. Proceedings. 1998 international Conference on, voU, no., pp.888,892 vol. I, 4-7 Oct 1998.

[10] Tka1cic, M., &Tasic, J. F. (2003, September). Colour spaces: perceptual, historical and applicational background. In Eurocon.

[11] Xue Ling; Li Chao; Li Huan; Xiong Zhang, "A General Method for Shot Boundary Detection," Multimedia and Ubiquitous Engineering, 2008. MUE 2008. international Conference on, vol., no., pp.394,397, 24-26 April 2008.

[12] Hua Zhang; Ruimin Hu; Lin Song, "A shot boundary detection method based on color feature," Computer Science and Network Technology (ICCSNT), 2011 International Conference on, vol.4, no.,pp.2541,2544, 24-26 Dec. 2011

[13] Patel, N. V., &Sethi, 1. K. (1997). Video shot detection and characterization for video databases. Pattern Recognition, 30(4), 583-592. [14].

[14] Zhang, Hong Jiang, AtreyiKankanhalli, and Stephen W. Smohar. "Automatic partitioning of full-motion video." Multimedia systems 1.1 (1993): 10-28

[15] Hui Chen; Cuihua Li, "A practical method for video scene segmentation," Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Cotiference on, vol.9, no., pp.153,156, 9-11 Ju[y 2010

[16] Pardo, A[varo. "Probabi[istic shot boundary detection using interframe histogram differences." Progress in Pattern Recognition, Image Analysis and Applications. Springer Ber[in Heide[berg, 2006. 726-732