International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 22 (2017) pp. 12616-12624

Shot Boundary Detection using Logarithmic Intensity Histogram: An Application for Video Retrieval

Kalpesh M. Shah
Research Scholar, Department of Electronics & Communication Engineering,
Charotar University of Science and Technology – CHARUSAT Campus-Changa,
Off. Nadiad-Pe tlad Highway, Gujarat -388 42, India.
Orcid Id: 0000-0003-0539-3838

Ramji M. Makwana
Chief Executive Officer (CEO), AI *eSmart Solutions, Rajkot-36000, Gujarat, India.*
Orcid: 0000-0002-7270-5371

Abstract

This paper presents automated shot boundary detection for video retrieval applications that exploits the logarithmic histogram of intensity to extract intensity histogram features from video frames. The proposed algorithm has the main purpose of identifying appropriate boundaries between shots of a video. In the first step, logarithmic histogram of the video frame is obtained with mapping of pixel intensity to an appropriate logarithmic level. This quantization process maps the image features into logarithmic histogram. To find shot boundaries between video frames, video frames are compared sequentially using logarithmic histogram and absolute difference. Extreme change between content of two successive frames considered if higher value of disparity is obtained. In lieu of statistical threshold, frame rate is used for identifying shot boundaries. If detected video shot having frames less than frame rate then boundary is not considered as a shot boundary and frames are clubbed with previously detected shot. Videos from various categories like sports, news and cartoon are used to evaluate the performance of proposed technique. It shows significant detection of boundaries having sharp transition. Method evaluated with 30 video sets and it is observed method performs very well with good rate of Hit and lesser number of Miss. As threshold is frame rate, it performs very well even in case of gradual shot boundaries. Proposed algorithm has also been evaluated on standard database TRECVID 2001 with good figure of Recall and Precision. Its performance is evaluated using on video retrieval application with 24Video sets from three different categories sports, Cartoons and News.

Keywords: Shot Boundary detection; Logarithmic histogram; Difference vector; Precision; Recall

INTRODUCTION

Developments in technology during last decades have extremely increased the accessibility and popularity of the Internet, digital broadcasting, and digital library services. It has tremendously raised the volume of multimedia data and so demand of one touch access to relevant data. These needs of society have involved the researchers to develop efficient methods of video parsing, indexing, and retrieval. Sometimes it is tedious task to describe every video sequence using few keywords especially when video content changes rapidly. Besides of it , this annotation procedure demands tremendous human power; keywords, its selectivity and its effectiveness also affects the outcome. Therefore, video retrieval and indexing should be based on video content analysis. Video content analysis starts with video parsing; it is the segmentation of video sequences into video shots, which allows efficient execution of the subsequent operations such as feature extraction, shot detection, and key frame selection. To represent a video content efficiently, it is desirable that video parsing can correctly detect the presence of shot transitions.

Shot boundary detection and key-frame extraction are primary tasks of video content analysis. Shot boundary detection becomes the groundwork for video retrieval and management. Shot boundary detection allows the computer to discover the editing positions and produce the original shot sequences. The detected shots will become the basic query units in video retrieval systems.

Video is a space and time varying structure; can be divided as video scenes, video shots, video frames and video key frames. All play a major role for video indexing and video retrieval applications. In (Thompson, 1998) (Vasileios T. Chasanis, Jan. 2009) (C. Cotsaces, Mar. 2006), defined a shot; it's a unit of visual information captured in single time interval by a camera that may show a certain action or event is called video shot. Hence, in order to collect the entire visual details of video content properly and to achieve an entire detailed coverage of the video, shot detection is a fundamental step of video content analysis. A video stream normally contains many shots, and the transitions between shots probably occur in different ways. The transitions can be generally categorized into two types: abrupt transitions and gradual transitions. In an abrupt transition, a frame from one shot, is followed by a frame from a different shot and is usually accompanied by a sharp change in luminance chrominance and texture. A gradual transition observes on sequence of several frames to complete, and the change between frames is gradual and continuous. Apparently, detecting gradual transitions is more challenging than detecting abrupt transitions.

Shot boundary detection aims to segment a given video sequence into its constituent shots, and to identify and classify the different shot transitions (Don Adjeroh, May. 2009). Different algorithms have been proposed, based on color histograms (Ballard, 1991; Tanaka, 1992), pixel color differences(H. Zhang, 1993), color ratio histograms (Lee, 1997), edges (R. Zabih, 1999), and motion. In this proposed work, we studied the problem of video partitioning using features of logarithmic intensity histogram without computing statistical threshold. Manual change of threshold may vary based on video content and is not a practical solution.

In this method intensity distribution is plotted on logarithmic scale. It scales pixel distribution logarithmically on fixed number of bins unlikely linear histogram distribution. Absolute difference is calculated sequentially for whole video sequence. This method does not rely on statistical threshold value; it depends on frame rate of a video. Rest of the paper covers State of the Art, Proposed Method. Subsequently, section includes how threshold decision simplified followed by experimental results and discussions.

STATE OF THE ART

Major Studies on shot boundary detection are related with visual features extraction color, edge, texture, and motion and comparing them among successive frames. This study shows transition between two shots can be either abrupt or gradual. The first is generated as a result of a cut effect, in which frame belongs to one shot and f_{i+1} to the next shot have clear discontinuity i.e. disparity value is very high. Second is generated using the application of video editing effects covering several sequential frames or because of motion spread through the range of few frames, so that frame f_i belongs to one shot, frame f_{i+N} to the second, and the $N-1$ frames in between represent a gradual transformation. Algorithms detecting shot transitions are based on the calculation of the difference between values of a certain measured characteristics like color, texture, edge, motion etc. of two consecutive frames; this value is later processed via a threshold decision in order to determine the existence of a boundary or not (Jesús Bescós, Apr. 2005). Shot boundary mainly has two basic steps.

Difference vector is calculated to find the dissimilarity between two frames using features extracted from each image. It includes selection of features, function to evaluate inter-frame disparity like Euclidian distance, Chi-square difference, and Absolute difference. For deciding shot boundary, threshold level is calculated using certain statistics based on extracted feature vectors.

ABRUPT SHOT CHANGES (CUTS)

Sharp transition detection method using 3×3 averaging filter had been implemented in (Zhang, 1994) to reduce camera motion and noise effects. The simplest approach to detect two frames are significantly different is to count the number of pixels that change in value more than some threshold. But, this method is very sensitive to camera motion.

Pixel to pixel matching with block-based comparisons was introduced by (R Kasturi, 1991) using local mage characteristic unlike pixel to pixel comparison mentioned ahead. This approached had increased the robustness to camera and object movement. Each frame of a video is divided into b blocks that are compared with their subsequent blocks in next frame. These blocks comparison is calculated using likelihood ratio. To reduce sensitivity to camera and object movements (A. Nagasaka, 1995) used simplest approach of gray level histograms. An abrupt transition is confirmed if the absolute sum of histogram differences between two succeeding frames f_i and f_{+1} is greater than a predefined threshold level. One more easy and very effective approach proposed by (H. Zhang, 1993) was to compare color histograms and used absolute difference to find dissimilarity.

In (Vasileios T. Chasanis, Jan. 2009) two features of the histograms difference and the pair-wise comparisons of pixels in the clustering method have been combined they end obtained very good results in finding existing shots with higher accuracy. In (Tekalp, 1998), the method proposed was called DC- images and was created and compared. DC-images, are the declined images of the original images regarding the location: the (i,j) pixel of the DC image is equal to the average block (i,j) of the original image. In (F. Arman A. H., 1993), shot boundaries are detected by the comparison of colored histograms of DC-images of sequential frames. These DC images are generated using DC terms of discrete cosine transform coefficients for a frame.

Gargi et al. (U. Gargi, 2000) listed a performance analysis few color histogram based SBD algorithms, where shot boundaries are detected using histogram differences of consecutive frames in various color spaces using various difference similarity measures. They observed that the histogram intersection method (M. Swain, 1991) performed very well but gradual transitions are not detection using it. Though DCT coefficients and motion vectors (F. Arman A. H.-Y., 1993)(H. Zhang, 1993)(J. Meng, 1995) (H. C. Liu, 1995)(B.-L. Yeo, 1995)(K. Shen, October 1995)(Patel, 1995) (Patel, 1995) approaches are computationally efficient but it’s performance levels deteriorated compare to histogram based methods

(Hampapur, 1995) had used model based comparison which considered the video production system as a template. (Zabih, 1995) used an edge features for detection of video shot boundary where it looks for entering and exiting edge pixels.

GRADUAL SHOT CHANGES

Accumulated differences between two frames of the gradual shot transition were calculated by (H. Zhang, 1993) using twin-comparison method. In an initial phase a high level of threshold is used to identify sharp cuts while in the second phase a lower value threshold is used to detect the first frame f_s of a gradual transition. This frame is then compared to subsequent frames and its difference value is accumulated. This accumulated comparison increases during a gradual transition. When the difference between consecutive frames decreases less than T_l , or when the accumulated comparison has increased to a value higher than T_h then end frame f_e of the gradual transition is detected. If the consecutive difference falls below T_l before the accumulated difference exceeds T_h , then the start frame f_s of gradual transition is dropped and the search of gradual transition remains continue.

The thresholds are highly depends on the content of a video. To overcome this problem, (B. Gunsel, 1998) proposed an application of unsupervised clustering algorithm by considering video segmentation as a 2-class clustering problem; shot change and no shot change using the well known K-means algorithm explained by (Pappas, 1992) is used to cluster the frame dissimilarities. Two similarity measures based on color histograms were used: s2 statistics and the histogram difference. (Zeinab Zeinalpour-Tabrizi, 2010) Proposed a efficient technique for gradual shots detection and classification, using fractal analysis and AIS-based classifier. They computed features of “vertical intercept” and “fractal dimension” of each frame which are calculated using Fourier transform coefficients. Classifier based on Clonal Selection Algorithm was used. They evaluated their algorithm using TRECVID2006 benchmark dataset.

DISPARITY MEASURES

(Mostafa Tavassolipour, Feb. 2014) extracted two types of features: spatiotemporal and compressed domain features. They used the Bhattacharyya measure defined by Eq. (1)

$$d(H1, H2) = \sqrt{1 - \frac{1}{N\sqrt{\overline{H}_1\,\overline{H}_2}} \sum_i \sqrt{H1(i)H2(i)}} \quad (1)$$

Where, $\overline{H}_k = \frac{1}{N}\sum_j H_k(j)$ and N denotes the number of histogram bins.

In order to evaluate inter-frame disparity (Jesús Bescós, Apr. 2005) have used RGB color bands and applied a criterion consisting of a variant of the Pearson’s Test. Color bands contain a similar amount of information for the computation of the objective similarity between two frames it could be merged into a single distance function Eq. (2)

$$d_1^n[i] = \frac{d_1^n R[i] + d_1^n G[i] + d_1^n B[i]}{3} \quad (2)$$

Where $d_1^n Band[i]$ is the aforementioned variant of the Pearson’s Test applied to the values that each color band takes for two consecutive frames.

(Colin O’Toole, 1999) used the dissimilarity analogue of the cosine measure for comparing the histograms of adjacent frames. The two N-dimensional vectors, A and B, represent the color signatures of the frames. The distance $D_{cos}(A, B)$ between vectors A and B is given by Eq. (3):

$$D_{cos}(A, B) = 1 - \frac{\sum_i^N a_i b_i}{\sum_i^N {a_i}^2 \, \sum_i^N {b_i}^2} \quad (3)$$

Where a_i is one bin in A and b_i is the corresponding bin in B.

THRESHOLD DECISION

Most of the algorithms comprise calculation of threshold after a disparity measure is computed, to identify video shot boundaries. At the point, where a frames difference value exceeds the threshold value is labeled as shot boundaries. In, (Srinath, 2001) proposed a simple threshold which was automatically computed on the histogram difference measure presented by (H. Zhang, 1993). This computation is based on entropy maximization of the distribution of the ID measure. In, (H. Zhang, 1993) Proposed a twin-comparison technique to identify gradual boundaries. Instead of only one, two thresholds are used. A low threshold is used to identify possible frames at the start of a gradual transition, and a high threshold is used to identify shot boundaries. In (Cheong, 2004) four thresholds were used. That is, one distinguishes pixels with high optical flow one determines pixels violating the smoothness constraint, one threshold shot changes, and, finally, the last one is used to remove false alarms.

INTENSITY HISTOGRAM

Frame features can be represented through distribution of intensities throughout the frame. To extract this feature intensity histogram is used. This histogram represents occurrence of each intensity value in an image. Y axes represents occurrence of intensity and x axes shows intensity span in an image, generally it is considered from 0 to 255. Proposed method spread up the bins on a log scale an find the occurrence of pixels of each intensity. Gray level pixels intensity varies from 0 to 255. If we scale it on a logarithmic base of 2, we may represent it using total 8 bins only for intensity spread from 0 to 255 as shown in Table I and Fig.1.

Table 1: exponential intensity distribution with 8 bins

Intensity spread	Corresponding bin
0	0
1	1
2 to 4	2
5 to 8	3
9 to 16	4
17 to 32	5
33 to 64	6
65 to 128	7
128 to 255	8

If intensity spread is narrow down, more number of bins can be defined to represent the intensity distribution on a log scale. In the proposed method, motivation to use 64 quantized levels is one can have better one to one matching and it is shown in Table 2. Few bins are not mentioned because it does not belong to any intensity level. Figure 2 shows a mapping between 256 intensity levels with 64 bins using Logarithmic scale. Higher the quantization level, higher the number of bins and so histogram comparison can be minimized and smooth one to one bin wise comparison can be obtained.

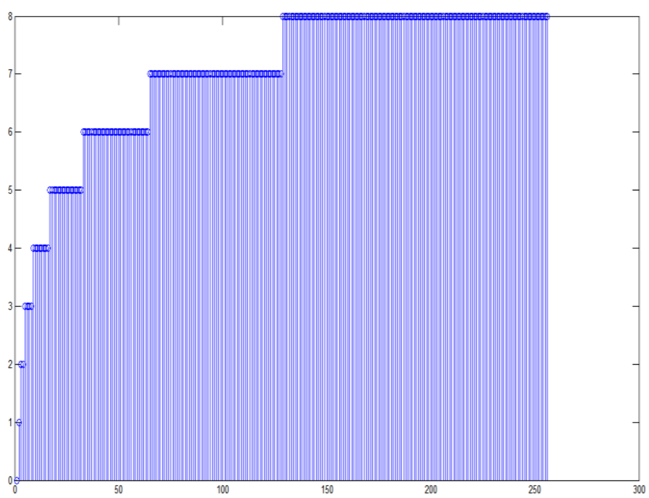

Figure 1: Exponential distribution with 8 quantized levels/bins s shown in Table II

Table 2: Exponential intensity distribution with 64 bins

Assigned Bin	Intensity spread	Assigned Bin	Intensity spread
1	1	42	35 to 38
9	2	43	39 to 41
13	3	44	42 to 45
17	4	45	46 to 49
19	5	46	50 to 53
21	6	47	54 to 58
23	7	48	59 to 63
25	8	49	64 to 69
26	9	50	70 to 76
27	10	51	77 to 82
28	11	52	83 to 90
29	12	53	91 to 98
30	13	54	99 to 107
31	14	55	108 to 117
32	15	56	118 to 127
33	16 to 17	57	128 to 139
34	18 to 19	58	140 to 152
35	20	59	153 to 165
36	21 to 22	60	166 to 180
37	23 to 24	61	181 to 197
38	25 to 26	62	198 to 215
39	27 to 29	63	216 to 234
40	30 to 31	64	235 to 255
41	32 to 34		

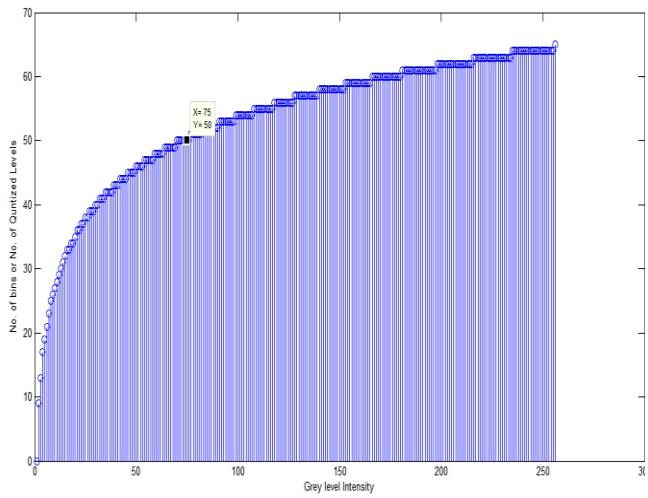

Figure 2: Exponential distribution with 64 quantized levels/bins as shown in Table 2

Video is made up of sequence of images. Its frame rate is defined as number of images per second given in Eq. (4).

$$N = T \times R \tag{4}$$

N= Total number of frame in a video

R= Frame rate per second

T=Length of video in seconds

Shot boundary is extracted using histogram features which has been calculated and plotted using logarithmic scale. Histogram is a graph of occurrence of specific pixel intensity range against a number of intensity spreads which is called as a bin. In this proposed method frames are separated from video as '.bmp' image format with its actual size.

Figure 3 shows a sample image of a news video. Figure 4 is gray level histogram of Figure 3. Figure 5 represents same histogram with 64 quantized levels shown in Table II. All these histograms are plotted on linear scale. Shot boundary detection using histograms comparisons were already proposed in (U. Gargi, 2000)(Vasileios T. Chasanis, Jan. 2009).

Figure 3: Sample Image

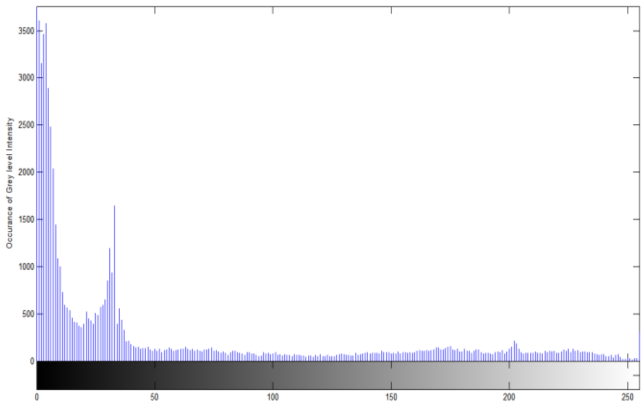

Figure 4: Histogram of Fig. 1 intensity spread 1: 256

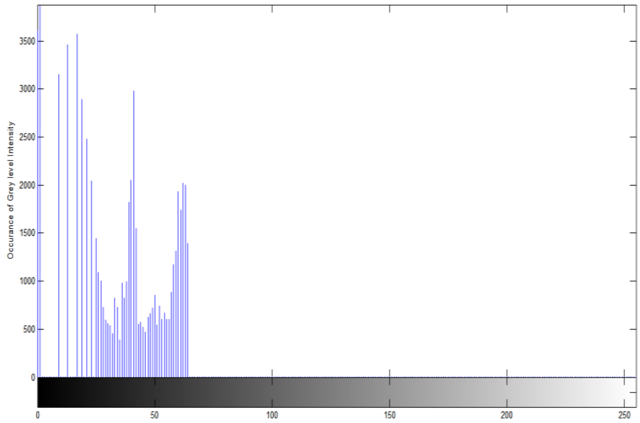

Figure 5: Exponential distribution of Intensity 1:64

```
for i=1:r
for j=1:c
I1_log(i,j)=ceil(log(double(I1(i,j))) / log(1.090504470));
end
end
for i=1:r
for j=1:c
if I1_log(i,j)==-inf
I1_log(i,j)=1;
end
end
end
```

Figure 6: Algorithmic steps to find Log of Image

Fig.6 shows mathematical steps to generate log image having 64 intensities only. Histogram of I1_log is plotted in Fig. 7. These 64bins histogram is compared with the histogram of all the frames of a video sequentially. Absolute difference is used to find this distance vector. Here sa2.avi video from sport category is used which contains 598 images of dimensions 240 × 352. Difference vector of 597 × 1 is calculated and plotted as shown in Figure 8. This difference vector is sorted in descending order as shown in Figure 9. It gives an idea about how much sharp cut boundaries are covered in a video, high difference values indicate sharp transition. Technically, It is not a good idea to change threshold manually depending upon video content density. To filter out high transition difference, frame rate is proposed as a threshold. This threshold is independent of content of video.

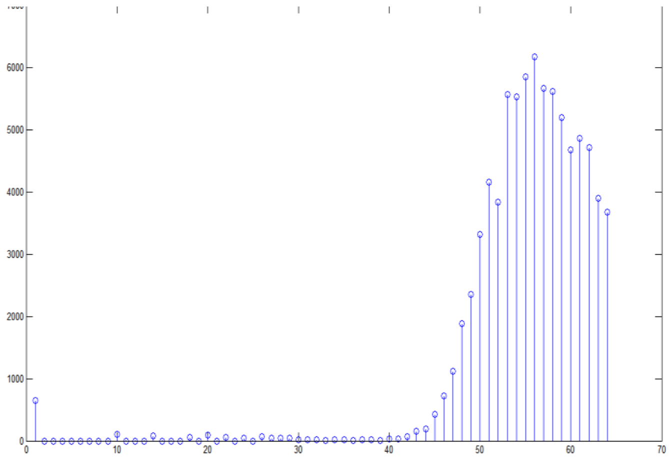

Figure 7: Histogram of Image with 64 bins

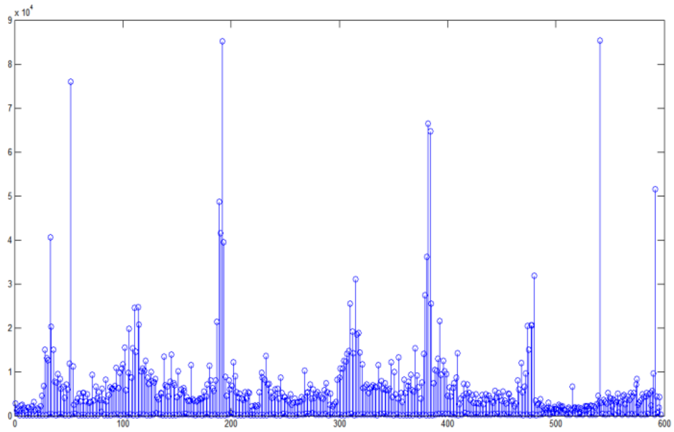

Figure 8: Frame Number Vs Frame Difference for sa2avi

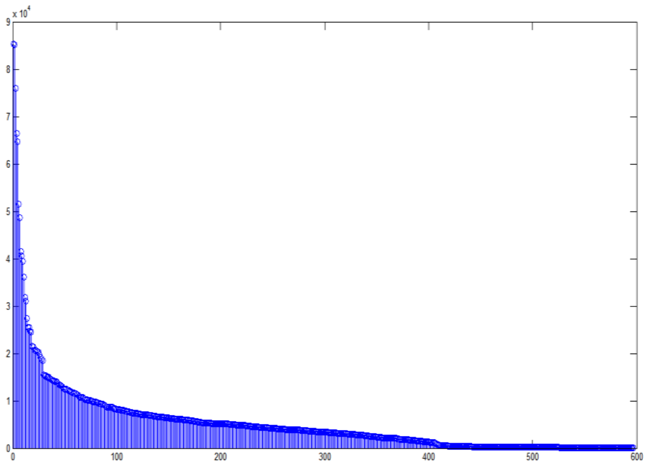

Figure 9: Sorted Frame Number Vs Frame Difference for sa2.avi video

THRESHOLD DECISION SIMPLIFIED

Threshold plays a major role to take decision for shot boundary detection. (A. Nagasaka, 1995) had used first order and second order statistical threshold for shot boundary detection and (H. Zhang, 1993) had used twin threshold for detecting gradual shot boundary. Threshold always remain a dynamic parameter which depends on content of a video. Adaptive threshold proposed by (T. Qin, Sep. 2010) also could not solve the problem significantly. So, Frame rate is considered as threshold, provided that each shot must be of minimum of one second i.e. boundary must be detected between two shots whenever frames between two consecutive shots must be equal to frame rate or more. If it is less video frame rate, then it cannot be considered as shot transitions i.e. it must be a continue sequence of frames with previously detected shot. It is a part of previous shot, does not indicates beginning of a new shot.

Because of this consideration gradual shot boundary detection task gets simplified without using statistical threshold. Fig. 10 is frame numbers derived from Fig. 9. To find these frame numbers sorted difference vector is divided by frame rate and highest difference vector frame number is generated which is a vector of 20 ×1 in a given example It can be calculated as shown in equation (5), FR=30, Difference Vector=597 × 1,

$$HDV=Ceil\ (DV/FR) = ceil(597/30) = 20 \quad (5)$$

DV Difference Vector size and

FR Frame rate.

This HDV termination number is calculated using formula (5) and these frames are shown in Figure 10. These highest difference frame numbers are extracted from above calculated difference vector and it is compared sequentially to verify difference between these two frame numbers. If these frame number difference is higher than FR, it is considered as an abrupt transition. It also finds gradual transition very effectively. Figure 11 is a top seven highest difference frame number where boundary exists. These seven frame numbers are obtained after clubbing difference vector having frame difference less than FR.

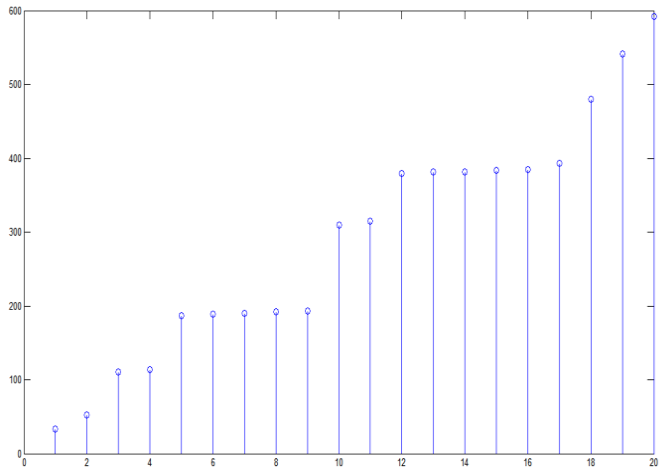

Figure 10. Frame Number having difference higher than frame rate for sport video sa2.avi

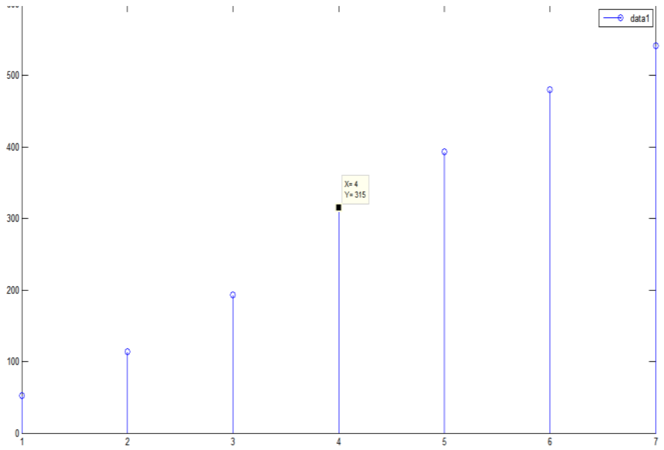

Figure 11. Seven shot boundaries for sa2.avi video

EXPERIMENTAL RESULTS & DISCUSSION

The performance of the proposed abrupt shot boundary detection technique was evaluated on various video sequence downloaded from www.youtube.com is shown in table III. Experiments were carried out on 30 videos from various categories containing news, sports and cartoon. The test video sequences have lot of video effects camera and object motion, like zoom-ins/outs, pans, sharp change of illumination due to strong camera flashes. All videos are of different durations and extracted image frames are of different dimensions. Figure 12 shows first seven transitions detected as shot boundary for a video sa2.avi. As we can see during sharp transitions content are changing drastically. Figure 13 and Figure 14 are shot boundary detected in News category video tn2.avi and cartoon category video Test5.1.avi respectively. There are certain shots in a video were shots are changing gradually. We observed method is finding out the sharp transitions as well performing well in gradual transition also. Most of the research has considered threshold for practical evaluation but threshold is changing depends on type and content of the videos. So in this method frame rate is used as a threshold to find shot boundary. If shot boundaries are detected as actual it is defined is considered as a Hit, If existed boundary could not be identified then it is defined as Miss. If any shot boundary is detected wrongly i.e. boundary does not exist but it is falsely detected then it is called false hit. Results are shown in tabular and graphical form in Fig. 15 and Table III. In some videos, shot boundaries are missed due to frame rate as a threshold that can be simplified for future work. These are mainly observed in low frame rate videos where few shot boundaries are missed to be identified as a shot.

For comparative evaluation the proposed algorithm Data sets were selected from TRECVID 2001 database. Precision (P) and Recall (R), and combined measure of it (F1) were used measure its rate of retrieval. The recall represents the detection ability of the proposed algorithm and the precision represents its accuracy. The equations for recall, precision and combined measure is:

$$\text{Recall}=\frac{N_c}{N_c+N_m}$$

$$\text{Precision}=\frac{N_c}{N_c+N_f}$$

$$F1=\frac{2\times P\times R}{P+R}$$

Where N_C stands for the number of correctly detected shot boundaries, N_m is the number of missed ones and N_f the number of false detections.

The proposed shot boundary detection method is compared with existing methods Adaptive Oriented and localized edge block (Hun-Woo Yoo). The results of these algorithms on similar video sequence are shown in Table. It is observed that precision rate and combined measure of the proposed method is better than the other methods which are 81%, 88% and 80%, 85% respectively. This shot boundary algorithm is evaluated for video retrieval application using 24 videos datasets covering categories of sports, cartoons and news. Performance is shown in Figure 16. Performance is measured in precision which chows good sets of figure is obtained in retrieving similar videos from database within few seconds only. This work can be extended for more number of video retrievals along with key frame detection.

ACKNOWLEDGMENT

Authors would like to thank Charutar Vidya Mandal and A.D. Patel Institute of Technology for supporting and providing resources for research work. Authors express heartfelt gratitude for Electronics & Communication Engineering department of A.D. Patel Institute of Technology.

REFERENCES

[1] Nagasaka, Y. T. (1995). Automatic video indexing and full video search for object appearances. 113-127.

[2] Gunsel, A. F. (1998). Temporal video segmentation using unsupervised clustering and semantic object tracking. 592-604.

[3] B.-L. Yeo, B. L. (1995). A unified approach to temporal segmentation of motion JPEG and MPEG compressed video. *Proceedings of 2nd International Conference on Multimedia Computing and Systems*, (pp. 81-83).

[4] Ballard, M. J. (1991). Color indexing. 7 (1), 11-32.

[5] Cotsaces, N. N. (Mar. 2006). Video shot detection and condensed representation. A review. vol. 23, (2), pp. 28-37.

[6] Cheong, L.-F. (2004). Computer Vision and Image Understanding.

[7] Colin O'Toole, A. S. (1999). Evaluation of Automatic Shot Boundary Detection on a Large Video Test Suite .

[8] Don Adjeroh, M. C. (May. 2009). Adaptive Edge-Oriented Shot Boundary Detection. 2009 (Article ID 859371).

[9] F. Arman, A. H. (1993). Image processing on compressed data for large video databases. *Proceedings of the first ACM international conference on Multimedia* , 267-272.

[10] F. Arman, A. H.-Y. (1993). Feature management for large video databases. Proceedings *of IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases I,* vol. SPIE 1908 (pp. 2-12).

[11] H. C. Liu, G. Z. (1995). Automatic determination of scene changes in MPEG compressed video. *Proceedings of ISCAS-IEEE International Symposium on Circuits and Systems*, (pp. 764-767).

[12] H. Zhang, A. K. (1993). Automatic partitioning of full motion video. 1 (1), 10-28.

[13] Hampapur, A. J. (1995). Production model based digital video segmentation. 9-46.

[14] Hun-Woo Yoo, H.-J. R.-S. Gradual shot boundary detection using localized edge blocks. *Multimedia tools applications*, 283-300.

[15] J. Meng, Y. J. (1995). Scene change detection in a MPEG compressed video sequence. *Proceedings of SPIE/IS&T Symposium on Electronic Imaging Science and Technology: Digital Video Compression: Algorithms and Technologies*, vol. 2419.

[16] Jesús Bescós, G. C. (Apr. 2005). A Unified Model for Techniques on Video-Shot Transition Detection. 7 (2), pp. 293-307.

[17] K. Shen, E. D. (October 1995). A fast algorithm for video parsing using MPEG compressed sequences. *IEEE International Conference Image Processing,* (pp. 252-255).

[18] Lee, D. A. (1997). Robust and efficient transform domain video sequence analysis: an approach from the generalized color ratio model. 8 (2), 182-207.

[19] M. Swain, D. B. (1991). Color indexing. *Int. J. Comput. Vis*. 7 (1) , 11-3.

[20] Mostafa Tavassolipour, M. K. (Feb. 2014). Event Detection and Summarization in Soccer Videos Using Bayesian Network and Copula. 24 (2).

[21] Pappas, T. (1992). An adaptive clustering algorithm for image segmentation. 901-914.

[22] Patel, I. S. (1995). A statistical approach to scene change detection. *Proceedings of IS&T/SPIE Conference Storage and Retrieval for Image and Video Databases III*, vol. SPIE 2420, (pp. 329–338).

[23] R Kasturi, R. J. (1991). Computer Vision: Principles. 469-480.

[24] R. Zabih, J. M. (1999). A feature-based algorithm for detecting and classifying production effects. 7 (2), 119-128.

[25] Srinath, J. Y. (2001). Pattern Recognition Letters .

[26] T. Qin, J. G. (Sep. 2010). A fast shot-boundary detection based on K-step slipped window. *2nd IEEE International Conference on Network Infrastructure and Digital Content*, (pp. 190-195).

[27] Tanaka, A. N. (1992). Automatic video indexing and full video search for object appearances.

[28] Tekalp, A. F. (1998). Efficient Filtering and Clustering Methods for Temporal Video Segmentation and Visual Summarization. *Journal of Visual Communication and Image Representation*, 1998, 336-351.

[29] Thompson, R. (1998). Grammar of the Shot, *Focal Press*.

[30] U. Gargi, R. K. (2000). Performance characterization of video-shot-change detection methods. Circuits Syst. Video Technol. 10 (1) , 1-13.

[31] Vasileios T. Chasanis, A. C. (Jan. 2009). Scene Detection in Videos Using Shot Clustering and Sequence Alignment. 11 (1).

[32] Zabih, R. M. (1995). A feature-based algorithm for detecting and classifying scene breaks. International Multimedia Conference and Exhibition. California: Multimedia Systems.

[33] Zeinab Zeinalpour Tabrizi, F. A.-M. (2010). Gradual Shot Boundary Detection and Classification Based on Fractal Analysis. 4 (8).

[34] Zhang, H. (1994). Video parsing using compressed data. *Proceedings of SPIE Symp. Electronic Imaging Science and Technology: Image and Video Processing II*, (pp. 142-149).