A Novel Method of Shot Boundary Detection using Center Symmetric Local Binary Pattern

T. Kar
School of Electronics Engineering
KIIT University,
Bhubaneswar

P. Kanungo
Dept of Electronics &Telecommunication Engg.
C.V. Raman College of Engineering,
Bhubaneswar

Abstract— **Impulsive growth of multimedia technology and popularity of social sites facilitates increasing application of video information which includes video indexing, browsing, retrieval and classification. Most of the videos, available in internet for public access are non-edited videos. Efficient way of searching and storage need an efficient method of annotation. Automatic cut detection is the first stage of automatic annotation process. In this paper we addressed the problem of video segmentation of only non-edited videos by classifying the boundary and non-boundary frames. The efficiency of intensity based cut detection methods decrease with variation of intensity of the scene. The centre symmetric local binary pattern is one of the powerful texture descriptor which provides a strong spatial correlation among the neighboring pixels, which is also invariant to light variation. Therefore in the proposed method, center symmetric local binary pattern histogram feature is used to detect abrupt shot boundaries in a video. The absolute sum CSLBP histogram difference between two consecutive frames is chosen as the similarity measure which is compared with a threshold value to detect the hard cuts in a non-edited video. The proposed algorithm is experimented with six test videos and its efficacy is validated with few existing popular approaches.**

Index Terms— Shot Boundary Detection, Histogram, Texture feature, Centre symmetriuc Local binary Pattern, Video segmentation.

I. INTRODUCTION

Video indexing and content based image retrieval is a well studied problem. A lot of work has been reported in literature in past two decades. The inevitable step of any video indexing and retrieval operation is the temporal segmentation [1] of video into small meaningful units which is known as shot [2]. A shot in general can be defined as a sequence of frames that has been captured (or appears to be captured) in uninterrupted fashion and continuous run of a single camera. The two consecutive shots are separated either by a single frame known as cut or by a multiple frames that changes in a mild way known as gradual transition. Cuts naturally occurs during creation of any video (non-edited video) but gradual transition is the result of numerous types of editing effects that is fabricated to make the video visually more fascinating. For example a dissolve effect is created by super imposing the boundary frames of two consecutive shots for a specific duration.Eexhaustive reviews of video shot boundary detection algorithms and exclusive comparison of their performances are addressed [2-5] in the literature.

Any shot boundary detection algorithm moves through three essential steps namely feature extraction, development of a similarity measure and making a decision about presence or absence of a cut based on an empirical automatic threshold.

Histogram feature based method [7,8], pixel intensity based method [9,10], edge feature based method [11,12], SIFT and SURF feature based method [13] are some cut detection methods that have been reported in literature. Instead of considering a single feature many authors have chosen a combination of features [14-15] to handle the issue of shot boundary detection.

Lakshmipriya et al.[12], has addressed shot boundary detection method based on edge strength using block based orthogonal features of a frame which takes care of the fast motion and lighting effects whereas failed to handle sudden flash light or explosion events in a scene. Amel et al. [16] used the motion activity to detect the shot boundaries which can handle the camera motion as well as the object motion in a video. Jialei et al.[17] used the mutual information as a feature to evaluate the distance between two consecutive frames and used SVM to detect a threshold for localizing the shot boundary. Jiang et al.[18] proposed a dual detection model based on pre- detection and re-detection process. Uneven block colour histogram difference and pixel value difference are used as a distance feature whereas an adaptive binary search is used to locate the shot boundary from the distance features in the pre-detection step. In the re-detection process a scale invariant feature is used to refine the detected boundaries to reduce the false detection which improves the precision rate. Threshold differentiates between two shots. Therefore threshold selection is one of the crucial and challenging task for the efficiency[3] of any shot detection algorithm. Among all these methods histogram based cut detections are the simplest and fastest whereas most of the time the histogram based methods fail to eliminate false cuts. False cuts increase because of having variations of the histogram of two similar scene or two consecutive frames of same shot.

Histogram difference methods and pixel difference methods are the popular techniques of hard cut detection in non edited videos. The non edited videos indicate videos without any gradual transitions (GT), fade-in or fade-out

effects. These two methods that is absolute sum histogram difference (ASHD) and absolute sum intensity difference (ASID) are based on the intensity values of the pixels. ASHD and ASID are based on the intensity features, which are the global features between two consecutive frames of a video, that may be similar for two different scenes or dissimilar for two same scenes due to the effect of light variation. Hence, under light variation conditions the efficacy of ASHD and ASID methods decrease due to the increase in missed cuts and false cuts. This motivates us to think about a spatial correlation parameter (local feature) instead of the intensity parameter to construct the global feature for the detection of hard cuts in a non-edited video. One of the spatial correlation features is the LBP texture feature which has a very low effect on the light variation in a scene[19]. But it has higher calculation complexity and fails for flat image areas which has less texture information. Solution is to represent the image using centre symmetric local binary pattern(CSLBP) [19] which captures texture information ,has illumination invariant property and works well for flat image areas.

In the current work a new feature based cut detection method is proposed to increase the efficiency of cut detection process.

II. RELATED WORK

Absolute sum histogram difference based method and pixel wise intensity difference method are the most popular cut detection algorithms due to their simplicity and efficacy. Therefore in this section these methods are explained and the performances of these methods are also compared with the proposed method in section V.

A. Pixel-wise intensity difference method

The sum of absolute intensity difference [9] (ASID) between two corresponding pixels of two consecutive frames is calculated as in (1). This ASID is compared against a threshold for detecting a cut between two consecutive frames. This method fails for large object motion and camera motion [1,12,15].

$$ASID_{t,t+1} = \sum_{i=0}^{R}\sum_{j=0}^{C}\left| f_t(i,j) - f_{t+1}(i,j)\right| \quad (1)$$

Where f_t and f_{t+1} are the t^{th} and $(t+1)^{th}$ frame of a video , R and C are the number of rows and columns of a frame.

B. Absolute sum based histogram difference method.

It is one of the simplest method which uses the histogram feature for the cut detection in a video. This histogram based method [7] computes the gray level histogram difference of the two images known as absolute sum of histogram difference (ASHD). The ASHD between two consecutive frames is evaluated as in (2).

$$ASHD_{t,t+1} = \sum_{i=0}^{255}\left|h_t(i) - h_{t+1}(i)\right| \quad (2)$$

Where, h_t *and* h_{t+1} are the histograms of t^{th} and $(t+1)^{th}$ frames respectively. If the $ASHD_{t,t+1}$ in (2) is above a threshold value then a shot boundary is detected in between t^{th} and $(t+1)^{th}$ frame. This method also fails for the fact that two different images can have same histogram as histogram representation ignores the spatial correlation between the pixels in the image.

III. ILLUSTRATION OF CSLBP FEATURE

The normal LBP[19] operator generates a histogram of length 256 which is rather longer and hence impractical to be used for region description. This issue of LBP histogram representation is addressed by modifying the scheme of pixel comparison in the local neighborhood. The new scheme is known as CSLBP [20,23] representation. In this method of region description only the symmetric pairs of pixels about the centre is compared. So the number of comparisons are halved and thus generates a histogram of length 16 , making it more simpler and appropriate for region description. To create robustness on flat image areas, the gray level difference is compared against a small threshold value to limit the unusual gradient magnitudes.

Evaluation of CSLBP feature (CSLBPF) of a pixel g_c, centered at (x_c,y_c) over a 3x3 neighbourhood is illustrated as follows:

$$CSLBPF\ (x_c, y_c) = \sum_{p=0}^{3} S(p)2^p \quad (3)$$

Where, p is the position of a pixel in a 3x3 window as shown in Fig. 1. For p=0, the corresponding gray value is denoted as g(0).

$$\text{Where, } S(p)=\begin{cases} 1 & g(p) - g(p+4) > 0.01 \\ 0 & otherwise \end{cases} \quad (4)$$

Using this CSLBP feature a cut detection algorithm known as BBCSLBPCD algorithm is proposed in section IV.

g(0)	g(1)	g(2)
g(7)	g_c	g(3)
g(6)	g(5)	g(4)

Fig. 1. Example of CSLBP in 3X3 neighborhood

(a)

(b)

Fig. 2. 1368th frame of Video V2 in Table I, (b) corresponding CSLBP feature image.

IV. PROPOSED BBCSLBPCD ALGORITHM

The efficiency of most of the methods discussed in section II decreases with high speed object motion or camera motion or with illumination variation. The most simplest and efficient method is the ASHD based cut detection but histogram information does not take care of any neighborhood or spatial relation. Considering the spatial correlation property of CSLBP, a new method known as "Absolute Sum centre symmetric Local Binary Pattern Histogram Difference" (ASLBPHD) is proposed to detect the cuts in non edited video more efficiently.

A. ASCSLBPHD based cut detection

The proposed CSLBP shot boundary detection algorithm can be summarized as follows:

- Step 1: Convert the RGB color frames of the video into gray scale.
- Step 2: Apply 5X5 wiener filter on every frame.
- Step 3: Generate CSLBP feature image of each frame.
- Step 4: Generate CSLBP feature histogram (CSLBPFH) from the LBP feature frames of step3.
- Step 5: Evaluate the normalised absolute sum CSLBPFH difference of consecutive frames using (6) and evaluate C using (7) for all the frames.
- Step6: If for the t^{th} frame, C(t)=1 then declare a cut at the t^{th} frame.

V. SIMULATIONS AND DISCUSSIONS

In the current work we have experimented on TRECVid[21] data set and publicly available dataset to validate our proposed method. "NAD-58" is from the trecvid 2001 dataset. "Before Sunrise" and "2_Brother" are two english movie having high speed object motion. "The Big Bang Theory" is a sitcom video which has been used by many literature earlier. "Masoom" is a hindi film song clipping. "Littlemiss Sunshine" is an english movie which has strong light variation in it. This list of videos along with the number of frames and ground truth cuts are tabulated in Table 1. All these videos are uncompressed .avi format with a spatial resolution of 320x240 and gray level resolution of 256. The proposed method is compared with basic absolute sum local binary pattern histogram difference method (ASLBPHD)[22] ,ASHD and ASID based method using three major performance criteria i.e recall (R), precision (P) and F1 measure [3]. Performance measures R, P and F1 are defined in (5), (6) and (7) respectively.

$$\text{Re}call(R) = \frac{D_{TC}}{T_C} \quad (5)$$

Where D_{TC} is the number of detected true cut by the algorithm . T_C is the actual number of true cuts in a video which is equal to the ground truth cuts tabulated in Table 1.

$$\text{Pr}ecision(P) = \frac{D_{TC}}{D_C} = \frac{D_{TC}}{D_{TC} + F_C} \quad (6)$$

Where D_C is the number of cuts detected and F_C is the number of false cut.

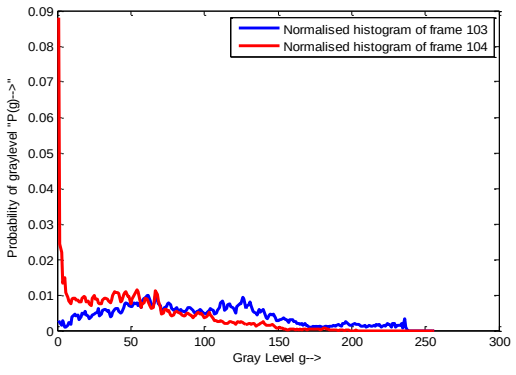
$$F_1 = \frac{2*R*P}{R+P} \quad (7)$$

Where F_1 is the harmonic average of R and P.

For illustration purpose 1368th gray scale frame and corresponding CSLBP feature frame from the video Before sunrise is shown in Fig. 2 (b). The feature image does not reflect intensity information as in case of gray level image

(a)

(b)



(c)

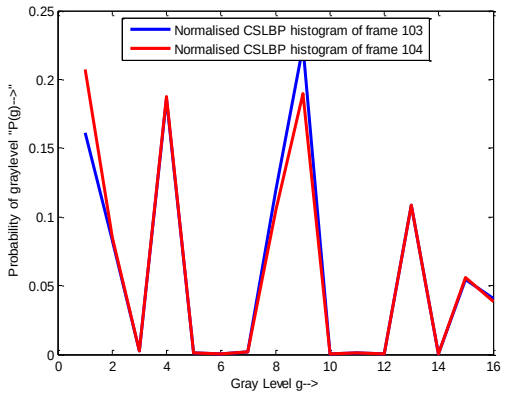


Fig. 3. (a) Frame no 103 (b) frame no 104 of the video Littlemiss sunshine.(c)Normalised histogramof two frames (d)CSLBP histogram of two frames

TABLE I. TEST VIDEOS AND GROUND TRUTH DATA

Video no	Video Name	Number of frames	Ground Truth cuts
V1	NAD-58	2000	9
V2	BS	2000	12
V3	2_B1	3000	19
V4	TBBT(Sitcom)	3000	32
V5	LM	4000	30
V6	Masoom(hindi song)	2000	9

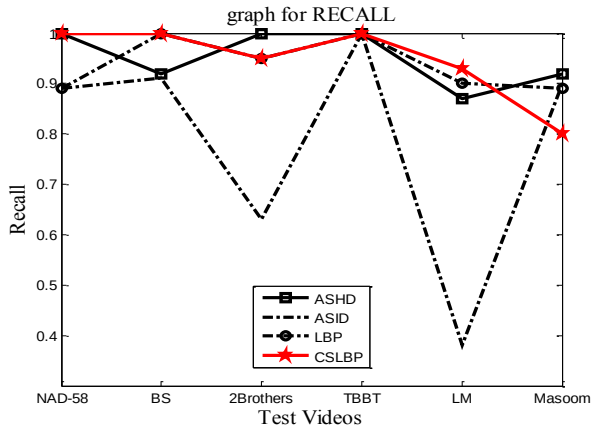


Fig. 4. Recall measure

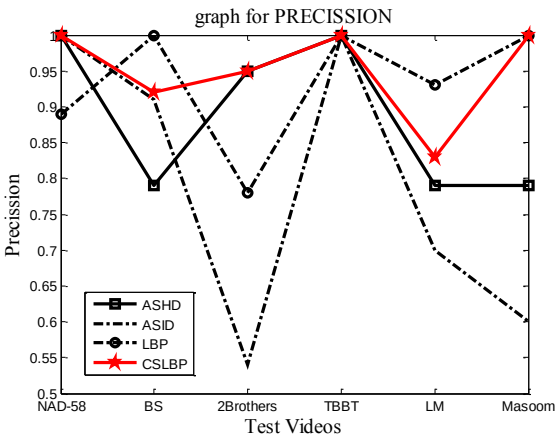


Fig. 5. Precision measure

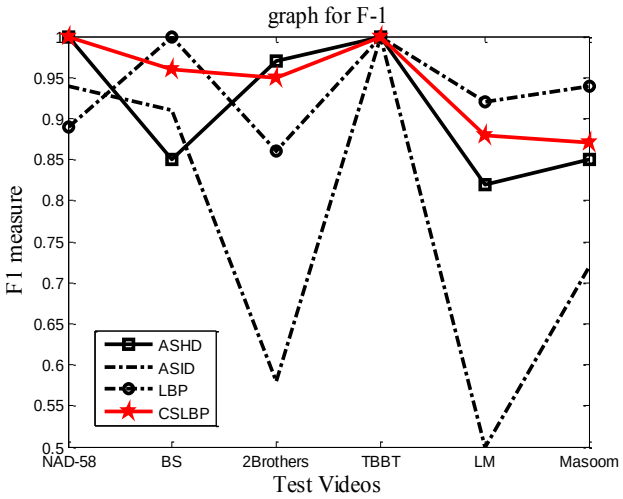


Fig. 6. F1 measure

The performance of the proposed absolute sum CSLBP feature histogram difference method is evaluated based on the (5), (6) and (7) respectively.

For validation purpose, two popular methods like ASHD, ASID along with ASLBPHD[22] method are also applied on the same test videos .Individual performance measures for different videos are also plotted in Fig. 4, 5 and 6 respectively.

It is observed from the results that the proposed method is superior in terms of the recall measure, precision measure and F1 measure for most of the videos. This shows the superiority of our algorithm and the number of detected true cuts are almost close to the ground trough cuts for most of the cases.

VI. CONCLUSIONS

In this paper a new shot boundary detection scheme has been proposed based on texture feature extracted from center symmetric local binary pattern. From performance analysis, it is found that the proposed ASCSLBP based method is able to detect hard cuts efficiently. The proposed CSLBP feature is capable to handle sudden illumination changes and works better on flat image areas as compared to LBP feature .

In terms of complexity the proposed method works better than the LBP feature based cut detection (ASCSLBPHD) [22] while more complex than ASHD and ASID methods. The proposed method is able to detect almost all hard cuts in non-edited videos. Our future work will focus on detecting gradual transition using CSLBP texture feature in combination with motion feature.

REFERENCES

[1] H. Zhang, A. Kankanhalli, and S. W. Smoliar, "Automatic partitioning of full-motion video," *Multimedia Syst.*, vol. 1, pp. 10–28, 1993.

[2] J. S. Boreczky and L. A. Rowe, "Comparison of video shot boundary detection techniques," *Proc. IS&T/SPIE Conf. Storage and Retrieval for Image and Video Databases IV*, vol. SPIE 2670, pp. 170–179, 1996.

[3] U. Gargi, R. Kasturi, and S. H. Strayer, "Performance characterization of video-shot-change detection methods," *IEEE Trans. Circuits Syst. VideoTechnol.*, vol. 10, no. 1, pp. 1–13, Feb. 2000.

[4] C. Cotsaces, N. Nikolaidis, and I. Pitas,'Video Shot Detection and Condensed Representation, a review',2006.
[5] P. Panchal, S. Merchant, 'Performance Evaluation of fade and dissolve transition shot boundary detection in presence of motion in video,' 1st International Conference on Emerging Technology Trends in Electronics, Communication and Networking,2012.
[6] Krishna K. Warhade, Shabbier N. Merchant and U. B. Desai,"Performance Evaluation of Shot Boundary Detection Metrics in the presence of Object and Camera Motion," IETE journal of research,vol 57(5),pp.461-466,2011.
[7] Rainer Lienhart, "Comparison of automatic shot boundary detection algorithm,"Image and video processing VII, in proc. of SPIE 3656-29.
[8] G.G. Lakshmi Priya, S. Domnic, "Video Cut Detection using Block based Histogram Differences in RGB Colour Space," International conference on Signal and Image Processing, pp.29-33,2010.
[9] A. Hanjalic, "Shot-boundary detection: Unraveled and resolved?" *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 2, pp. 90–105, Feb. 2002.
[10] C. W. Su, H. Y. M. Liao, H. R. Tyan, K. C. Fan, and L. H. Chen, "A motion-tolerant dissolve detection algorithm," *IEEE Trans. Multimedia*, vol. 7, no. 6, pp. 1106–1113, Dec.
[11] Yoo H.W, Ryoo H.J. 'Gradual shot boundary detection using localised edge blocks', Multimed.Tools Appl.,2006,28,pp. 283-300.
[12] G.G. Lakshmipriya,S. Domnic,"Edge Strength Extraction using Orthogonal vector for shot Boundary Detecton,"Procedia Technology(6),pp.247-254,2012.
[13] M. Birinci,S.Kiranyaz,"A Perceptual Scheme for Fully Automatic Video Shot Boundary detection,"signal processing:image communication,pp.410-423,vol.29,2014.
[14] R.A. Jyoce, B.D. Liu: 'Temporal segmentation of video using frame and histogram space', IEEE Trans. Multimed., 2006, 8, (1), pp. 130–140.
[15] G. G. Lakshmipriya, S. Domnic' Walsh-Hadamard transform kernel based feature vector for shot boundary detection' IEEE trans on image processing. Vol.23,no.12,Dec 2014.,pp. 5187-5197.
[16] Abdelati Malek Amel, Ben Abdelali Abdessalem and Mtibaa Abde llatif "Video Shot boundary detection using motion activity Descriptor, Journal of Telecommunications,"vol.2(1),pp.54-59,2010.
[17] Jialei Bi, Xianglong Liu, Bo Lang, "A Novel Shot Boundary Detection Based On Information Theory using SVM" 4th International Congress on Image and Signal Processing,pp.512-516, 2011.
[18] Xinghao Jiang, Tanfeng Sun, Jin Liu,Juan Chao,Wensheng Zhang. "An Adaptive Video Shot Segmentation Scheme Based on Dual-detection model," Nurocomputing,vol.116,pp.102-111,2013.
[19] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns"IEEE Trans. Pattern Analysis and Machine Intelligence. Vol. 24, no. 7, pp.971-987, July 2002.
[20] M. Heikkila,M. Pietikainen, C. Schimid, "Description of interest regions with center symmetric local binary patterns", in:5th Indian Conference on computer vision ,Graphics and image processing, vol.4338,2006,pp.58-69.
[21] TRECVID dataset available on www.open-video.org.
[22] T.Kar, P.Kanungo, "A Texture Based Method for Scene Change Detection", Conference on Power, Communication and Information Technology, 15-17 Oct. ,2015.IEEE explore.
[23] T. Kar, P. Kanungo, "Cut Detection Using Block based Centre Symmetric Local Binary Pattern", International conference on Man and Machine Interfacing (MAMI), Dec 17-19,2015.IEEE explore.