Video Shot Boundary Detection: A Review

Ananya SenGupta
Department of Computer Science & Engineering
National Institute of Technology Silchar
Assam, India
ananyasengupta1991@gmail.com

Dalton Meitei Thounaojam
Department of Computer Science & Engineering
National Institute of Technology Silchar
Assam, India
dalton.meitei@gmail.com

Kh. Manglem Singh
Department of Computer Science and Engineering
National Institute of Technology Manipur
Manipur, India
manglem@gmail.com

Sudipta Roy
Department of Information Technology
Assam University Silchar
Assam India
sudipta.it@gmail.com

Abstract— **Due to huge growth in multimedia and technology, it is very important to go through the point of interest rather than accessing the entire video. For efficient indexing and retrieving the interest points, content based video retrieval is used. The first step toward CBVR is shot boundary detection. It is necessary to partition the video into shots for easy indexing and retrieval of video. Therefore, segmentation plays an important role in digital media processing, pattern recognition, and computer vision. In this paper, we present different approaches to shot boundary detection problem.**

Keywords: Shot boundary detection, cut, dissolve, fades, wipes.

I. INTRODUCTION

Now a days, in many areas of digital libraries, hospitals, distance learning, video-on-demand, digital video broadcast, interactive TV, multimedia information large collections of digital videos are being created. Searching of videos can be text based or content based. In text based video retrieval, videos are annotated with text and a textual keyword is used for searching. Text based approach is more time consuming as well as it populates the database with a lot of data. Therefore efficient retrieval process is content based retrieval which gives more appropriate result than traditional text based approach.

Indexing, searching and retrieval of videos from large databases like YouTube, daily-motion etc. can be more efficient if we divide the entire video sequence into segments (shots). A shot is unbroken sequence of frames taken from a camera [1]. To segment the video into shots we need to locate the shot boundaries in a video. Transition of shots within a video can be of two types: abrupt and gradual transitions. Abrupt transition also known as hard cuts or cuts occur within a single frame when stopping and restarting of camera. Gradual transitions are also known as edit effects or cinematic effects, often used effects are fades, dissolve, wipes. A fade-in is a gradual increase in intensity starting from a black frame to a bright frame where as a fade-out is a gradual decrease in intensity starting from a bright frame and results to a black frame. When one frame gets superimposed on other frame i.e. frames of previous shot gets dimmer while those of second shot gets brighter, are called dissolve. A wipe is a type of transition where one shot replaces another by travelling from one side of the frame to another or with a special shape like clock, rectangle, oval etc [2]. Early work focused on cut detection, while more recent techniques deal with gradual transition detection. Abrupt transitions are easier to detect than gradual transition. Despite countless proposed approaches and techniques so far, robust algorithms for detecting various types of shot boundaries i.e. constant quality of detection for abrupt as well as gradual, have not been found yet. [3]

II. SHOT BOUNDARY DETECTION

The different approaches for shot boundary detection are:

A. Pixel comparison:

Pair wise pixel comparison between two consecutive frames evaluates the difference in intensity value of the corresponding pixel.

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

Where f and $f+1$ are two adjacent frames of size X×Y, I_f is the intensity value of pixel at coordinate (x, y) of frame f. Pixel wise comparison is limited to object and camera motion. Due to a small change in camera or object motion can result in large pixel difference [4].

B. Block based comparison:

Each frame is divided into n blocks and each block is compared with corresponding block of next frame. A transition is declared if the number of changed blocks between two consecutive frames is greater than a given threshold.

Each frame is divided into blocks mean of each block is taken which results in statistical image (reduced image). [5][6]

978-1-4799-6085-9/15/$31.00©2015

Mean Square Error of corresponding pixels is calculated for adjacent statistical frames to find the exact wipe transition region. They used Hough transform to determine the thickness of the strips in statistical image (single line or two lines). Value of average gradient and number of lines determines the wiping pattern [5]. In [7] technique based on pixel wise difference between consecutive frame is calculated for wipe transition detection and some methodologies based on horizontal, vertical and box- shaped wiping pattern is considered. The X and Y trajectory estimation of boundary line (horizontal, vertical, diagonal, clock) between two adjacent frame is calculated in [8].

Sugano M. et al [9] used techniques to locate abrupt, dissolve and wipes in compressed domain. A frame is considered as abrupt shot boundary if sum of forward prediction macro blocks and intra macro blocks is greater than a predefined threshold T_1 and number of backward macro blocks are smaller than T_2 where $T_1 > T_2$.(P and I frames). In case of B frames, number of forward macro blocks is less than T_3 , and sum of backward prediction and intra macro blocks is greater than T_4 where $T_4 > T_3$.

C. Histogram comparison:

For digital images, a color histogram represents the number of pixels that have fixed colors in color range. A color histogram can be build for color space like RGB, HSV, CMYK, YC_bC_r [10] etc. [4]

i. *Global histogram comparison*:

Histogram of two successive frames is computed and compared. If the histogram difference is greater than some predefined threshold then a transition happens.

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

Where $H_f(i)$ is the histogram value for gray level i of frame f and n is the total number of gray levels.

ii. Local histogram comparison:

Each frame is divided into n blocks and histogram of each block is compared with histogram of corresponding block of next frame.

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

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

Xue L. et al [11] proposed an algorithm that improves performance by eliminating smooth intervals from video. Features such as pixel wise difference, HSV color histogram and edge histogram are of the new video sequence are extracted and given as input vectors to support vector machine. The outputs of SVM are classified into three categories as abrupt, gradual and etc.

Using HSV color histogram difference and adaptive threshold hard cuts are detected in [12] [9]. They also calculated the local histogram difference and local adaptive threshold for gradual shot transition detection.

Histogram is extracted from each video frames and a matrix is created from the histogram values as the column of the matrix and Singular Value Decomposition is applied to the matrix thereby reducing the feature vector and provides fast computation [13][14][15]. Similarity measures like Euclidian [13] and cos distance [14] [15] distance are used to find the abrupt and gradual transitions. In [15], an inverted triangle pattern matching is used to find the gradual transitions.

[16] Deals with the false shot boundary detection due to flash light. If the histogram difference is greater that a predefined threshold then flash light detector is invoked. In flashlight model two models are introduced namely cut model (for cut detection) and flash model (flash light detection). The model for flash light is average intensity changes from one level to another on occurrence of flashlight and comes back to the original level after one frame.

A modification to simple histogram comparison technique is presented in [17]. Initially I frames are extracted from the MPEG video stream. Intensity, row/horizontal, column/vertical histograms are computed for all the I frames and compared by chi-square test [18]. Both algorithms operate in the compressed domain, requiring only partial decoding of the compressed video stream.

The first step towards video scene segmentation and indexing is shot detection as described in [19] [20]. They used twin threshold approach which means a high threshold *Th* and a low threshold *Tl* is selected and if gray level histogram difference, *D* of two adjacent frames is larger than *Th*, a cut is declared. If *D* is larger than *Tl* and less than *Th*, they are accumulated. Now if accumulated difference is greater than *Th*, a gradual transition exists.

Joyce and Liu [21] presented two algorithms for detecting dissolve and wipes. The first is a dissolve detection algorithm which is implemented both as a simple threshold-based detector and as a parametric detector by modelling the error properties of the extracted statistics. The second is an algorithm to detect wipes based on image histogram characteristics during transitions.

Pardo A. [22] proposed an algorithm for hard cut detection that gives better results than feature based, pixel based or simple histogram based approaches. Initially inter frame histogram difference between frames are calculated for a set of bins. Probability for inter frame difference to be greater than a predefined threshold is set and tested. If a shot change occurs, histogram difference is expected to be more whereas for no shot change, histogram difference is expected to be less and in agreement with previous histogram difference.

D. Feature based video segmentation:

a. Local features based segmentation:

SIFT [23], SURF [24] and MSER [25] [26] are local feature descriptor. Local features of successive frames are computed

and compared. If the numbers of matched features are less than some predefined threshold, a transition is declared.

To reduce computational cost and improve performance, non shot boundary frames are removed from the original video. SIFT features [23] are extracted from the frames belonging to only shot boundaries to identify the abrupt and gradual transition [27].

Deepak, C. R., et al [28] proposed an algorithm for shot boundary detection using color correlogram and Gauge Speeded-up robust features [29]. Initially linear frame comparison of color correlogram and G-SURF features are extracted for transition detection. To improve performance same method is applied by extracting key frames of the video sequence.

In [30], feature vector of each frame is computed by applying local feature transform (LFT) on each frame. After applying LFT on each frame of the video sequence, features are extracted and first and second moments are computed for channels of colour space to compute the feature vector.

A frame is considered as abrupt boundary if it has less/ no matched SURF features [24] with its successive frame. For fade detection, entropy of all the frames are computed and compared with adjacent frames. [31] Uses both local features (SURF) and global features (entropy) for video segmentation.

b. Global feature based segmentation:

Colour, texture, shape, edge, text, audio [32] and motion features of a frame are computed and compared with the next frame.

Counting the entering and exiting edge pixels between two consecutive frames shot boundary can be detected. A part from counting the number of entering and exiting edge pixels, cuts, fades, wipes and dissolve is located in presence of camera and image motion by computing global motion between frames [33] using image registration technique described in [34].

Hauptman et al [35] used text features to segment video which was implemented in Informedia Digital Video Library Project at Carnegie Mellon University.

Shot boundary detection in the presence of illumination change, fast object motion, and fast camera motion Mishra, R., Singhai, S. K., & Sharma M. [36] proposed an algorithm that extracts structure feature of each frame and similarity is computed between adjacent frames. Structure feature is extracted by using dual tree complex wavelet transform (CWT) [37].

In [38] mutual information and joint entropy is calculated between all pairs of frames. Object motion and shot transition is detected with canny edge detector.

Uncompressed sports video consists of huge camera movement, similar background and motion of dissimilar objects. [39] Describes an algorithm based on feature correlation, histogram difference and running average difference for hard cut detection in this type of videos. Also a motion based algorithm to identify shots cut is proposed in [40]. By calculating the normalized correlation between blocks of frames and locating the maximum correlation coefficient, an inter frame difference metric is generated.

An object based shot boundary detection is described in [41] and [42] for abrupt and gradual transition detection. A time stamp is attached with each object to locate the number of frames in which that particular object appears. Change is object appearance is considered as shot transition. Object tracking using some modification to Canny's egde detection algorithm is shown in [41]. The limitations of object based detection [41] [42] are large object movements, if objects suddenly disappear in the frames, flashlight light scenes [41].

Mutual information and joint entropy are calculated for all the frames of video. The amount of information transported to the next frame is mutual information which is sufficient for detecting cuts where intensity changes abruptly. For fade in/out, joint entropy between two consecutive frames are calculated for each components of RGB. [43] [44]

Cuts and wipes are detected using Markov energy model based on color and texture discontinuities in shot boundaries. Texture and color features and are computed by Gabor decomposition [45] and RGB histogram respectively. All computations are performed on spatio-temporal slices (collection of 1D image in sequence). For dissolve detection, within a specific time interval, mean intensity and variance of a slice is calculated. The periods having constant mean values and concave upward parabola curves are considered as dissolve. [46]

X. Gao, J. Li, and Y. Shi [47] proposed an algorithm that extracts a set of corner-points (features) from the first frame of a shot. Using Kalman Filtering [48] these features are matched with the features of the subsequent frames, accordingly with the changing pattern of pixel intensity shot boundary is detected.

An algorithm for video segmentation in uncompressed domain is proposed in [28]. For abrupt transition detection only I frames are uncompressed to DC images and characters of these DC images are extracted and compared with the adjacent DC image. For gradual transition, the range of shot boundary is detected by CDDC algorithm and located by the change in amount of intra coding macro blocks in P frames.

DC images are extracted from MPEG video and features are extracted and selected through Ada Boost [49] for cut detection [50]. A fade out followed by fade in is detected by considering the change in luminance, where the intensity values shows a V-shaped graph, detection of dissolve is identification of downward parabola / U- shaped pattern [50].

E. Clustering based approach:

In this approach, n frames are selected at random as initial cluster centres. The distance between each frame and cluster centres are calculated and the frames with small distance are clustered together. The common clustering algorithms are fuzzy clustering [51], mean sift clustering [52], K means clustering [53] etc.

B.Han, X.Gao, and H.Ji [54] used techniques Fuzzy c-means clustering (abrupt detection), Gaussian Weighted

Housdorff Distance and edge count ratio (fade transition), method based on similarity of color distribution (dissolve transition) and motion vector based on the three-dimension wavelet transformation (wipe transition).

A pre-processing technique [15][55] or frame skipping technique [56] is used to reduce the computing time thereby finding cluster/group of frames where a possible abrupt or gradual transitions may present by using thresholding technique.

Table 1: List of survey

Year	Reference	Author
2001	[57]	Koprinska I. et al.
2001	[58]	Lienhart, R
2002	[3]	Hanjalic A.
2006	[59]	Cotsaces, C
2007	[60]	Jinhui Yuan et al
2008	[61]	Geetha P. et al.
2014	[62]	Sao Nikita et al.
2014	[32]	Thounaojam D. M et al

Table 2: Results of some existing algorithm

Year	Reference	Abrupt	Fade	Dissolve	Wipes
1995	Zahib R. et al[33]	✓	✓	✓	✓
1997	Patel N.V. et al [17]	✓	×	×	×
1998	Sugano M. et al [9]	✓	✓	×	✓
1999	C. W. Ngo et al [34]	✓	×	✓	✓
2000	Gong Y. et al [13]	✓	✓	✓	✓
2001	D. Zhang et al [16]	✓	×	×	×
2001	W.J. Heng et al [41]	✓	✓	✓	✓
2003	Porter et al [40]	✓	×	×	×
2006	Joyce et al [21]	×	×	✓	✓
2006	Zuzana C. et al[43]	✓	✓	×	×
2006	Pardo A. [22]	✓	×	×	×
2008	Huan et al [38]	✓	✓	✓	×
2009	A.Hameed [39]	✓	×	×	×
2009	Ren J. eta al [45]	✓	✓	✓	×
2010	Chen H. et al [20]	✓	✓	✓	✓
2011	Hua Z. et al [12]	✓	✓	✓	×
2011	B.H. Shekar et al[30]	✓	×	×	×
2013	Baber J. et al [31]	✓	✓	×	×

III. COMPARISON OF SOME EXISTING TECHNIQUES

Table 3 shows some of the simple video shot boundary detection techniques. In this, a threshold is selected experimentally and no specific thresholding technique is applied. Only abrupt transitions are considered for the comparison.

Table 3: Results for abrupt transition

Video	anni003.mpg	hcil2001_01.mpg	indi002.mpg
Frame no.	4297	4202	845
No. of Cut	24	20	6
Color Histogram Difference [12]	21	20	6
Pixel Difference [4]	18	18	4
Mean and Standard Deviation	22	19	5
MSER [26]	22	20	6

The videos are downloaded from OPEN VIDEO PROJECT. The threshold is selected through experiments and the result can be met more efficient if we apply some verification system [31].

IV. CONCLUSION

Video shot segmentation is the first step towards automatic annotation and indexing of digital video for efficient browsing and retrieval. In this paper we presented different approaches for shot boundary detection and comparison of results among some algorithm. The system can be improved, if the output is verified further whether it is actual abrupt or gradual transitions. In almost all the algorithms, abrupt transitions are effectively detected than gradual transitions. Lighting effect (flash), large object motion in front of camera, fast camera motion, etc., are some of the challenges in detecting gradual transitions as it produces false detection.

ACKNOWLEDGEMENT

We are thankful to the OPEN VIDEO PROJECT for providing the video data free and for the use of it in the research purpose.

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.11, no.12, pp.1281,1288, Dec 2001

[2] Koprinska, I., &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* , vol.12, no.2, pp.90,105, 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.A.C.; Canagarajah, C.N.; Bull, D.R., "Wipe scene change detection 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* , vol.1, no., pp.893,897 vol.1, 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* , vol.1, no., pp.888,892 vol.1, 4-7 Oct 1998.

[10] Tkalcic, 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] Gong, Yihong, and Xin Liu. "Video shot segmentation and classification." In*Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 1, pp. 860-863. IEEE, 2000.

[14] Cernekova, Zuzana, Constantine Kotropoulos, and Ioannis Pitas. "Video shot segmentation using singular value decomposition." In *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*, vol. 3, pp. III-181. IEEE, 2003.

[15] Zhe-Ming Lu; Yong Shi, "Fast Video Shot Boundary Detection Based on SVD and Pattern Matching," *Image Processing, IEEE Transactions on* , vol.22, no.12, pp.5136,5145, Dec. 2013.

[16] Zhang, Dong, Wei Qi, and Hong Jiang Zhang. "A new shot boundary detection algorithm." In *Advances in Multimedia Information Processing—PCM 2001*, pp. 63-70. Springer Berlin Heidelberg, 2001.

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

[18] Chi-square tests. Defense Technical Information Center, 1976

[19] Zhang, Hong Jiang, Atreyi Kankanhalli, and Stephen W. Smoliar. "Automatic partitioning of full-motion video." *Multimedia systems* 1.1 (1993): 10-28

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

[21] Joyce, R.A.; Liu, B., "Temporal segmentation of video using frame and histogram space," *Multimedia, IEEE Transactions on* , vol.8, no.1, pp.130,140, Feb. 2006

[22] Pardo, Alvaro. "Probabilistic shot boundary detection using interframe histogram differences." *Progress in Pattern Recognition, Image Analysis and Applications.* Springer Berlin Heidelberg, 2006. 726-732

[23] Lindeberg, T. (2012). Scale invariant feature transform. Scholarpedia, 7(5), 10491.

[24] Bay, Herbert, TinneTuytelaars, and Luc Van Gool. "Surf: Speeded up robust features." *Computer Vision–ECCV 2006*. Springer Berlin Heidelberg, 2006. 404-417.

[25] Donoser, M.; Bischof, H., "Efficient Maximally Stable Extremal Region (MSER) Tracking," *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on* , vol.1, no., pp.553,560, 17-22 June 2006.

[26] Anjulan, A.; Canagarajah, N., "Video Scene Retrieval Based on Local Region Features," *Image Processing, 2006 IEEE International Conference on* , vol., no., pp.3177,3180, 8-11 Oct. 2006.

[27] Jun Li; Youdong Ding; Yunyu Shi; Wei Li, "Efficient Shot Boundary Detection Based on Scale Invariant Features," *Image and Graphics, 2009. ICIG '09. Fifth International Conference on* , vol., no., pp.952,957, 20-23 Sept. 2009

[28] Deepak, C.R.; Babu, R.U.; Kumar, K.B.; Krishnan, C.M.R., "Shot boundary detection using color correlogram and Gauge-SURF descriptors," *Computing, Communications and Networking Technologies (ICCCNT),2013 Fourth International Conference on* , vol., no., pp.1,5, 4-6 July 2013

[29] Alcantarilla, Pablo F., Luis M. Bergasa, and Andrew J. Davison. "Gauge-SURF descriptors." *Image and Vision Computing* 31.1 (2013): 103-116.

[30] B. H. Shekar, M. SharmilaKumari, and RaghuramHolla. "Shot boundary detection algorithm based on color texture moments." *Computer Networks and Information Technologies.* Springer Berlin Heidelberg, 2011. 591-594.

[31] Baber, J., Afzulpurkar, N., & Satoh, S. I. (2013). A framework for video segmentation using global and local features. *International Journal of Pattern Recognition and Artificial Intelligence*, *27*(05).

[32] Thounaojam, D. M., Trivedi, A., Singh, K. M., & Roy, S. (2014). A Survey on Video Segmentation. In *Intelligent Computing, Networking, and Informatics* (pp. 903-912). Springer India.

[33] Zabih, R., Miller, J., & Mai, K. (1999). A feature-based algorithm for detecting and classifying production effects. *Multimedia systems*, *7*(2), 119-128.

[34] Ngo, C.W.; Pong, T.C.; Chin, R.T., "Detection of gradual transitions through temporal slice analysis," *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* , vol.1, no., pp.,41 Vol. 1, 1999

[35] Hauptmann, A., and M. Smith. "Text, Speech, and Vision for Video Segmentation: The Informedia TM Project." *AAAI fall symposium, computational models for integrating language and vision.* 1995.

[36] Mishra, R., Singhai, S. K., & Sharma, M. (2013, February). Video shot boundary detection using dual-tree complex wavelet transform. In *Advance Computing Conference (IACC), 2013 IEEE 3rd International* (pp. 1201-1206). IEEE.

[37] Selesnick, I.W.; Baraniuk, R.G.; Kingsbury, N.C., "The dual-tree complex wavelet transform," *Signal Processing Magazine, IEEE* , vol.22, no.6, pp.123,151, Nov. 2005.

[38] Zhao Huan; Li Xiuhuan; Yu Lilei, "Shot Boundary Detection Based on Mutual Information and Canny Edge Detector," *Computer Science and Software Engineering, 2008 International Conference on* , vol.2, no., pp.1124,1128, 12-14 Dec. 2008

[39] Hameed, A., "A novel framework of shot boundary detection for uncompressed videos," *Emerging Technologies, 2009. ICET 2009. International Conference on* , vol., no., pp.274,279, 19-20 Oct. 2009

[40] Porter, S., Mirmehdi, M., Thosmas, B.: Temporal video segmentation and classification ofedit effects. Image Vis. Comput. 21(13), 1097–1106 (2003).

[41] Heng, W. J., &Ngan, K. N. (2001). An object-based shot boundary detection using edge tracing and tracking. *Journal of Visual Communication and Image Representation*, *12*(3), 217-239.

[42] Wei Jyh Heng; Ngan, K.N., "Integrated shot boundary detection using object-based technique," *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on* , vol.3, no., pp.289,293 vol.3, 1999

[43] Cernekova, Z.; Pitas, I.; Nikou, C., "Information theory-based shot cut/fade detection and video summarization," *Circuits and Systems for Video Technology, IEEE Transactions on* , vol.16, no.1, pp.82,91, Jan. 2006

[44] Cernekova, Z., Nikou, C., & Pitas, I. (2002, June). Shot detection in video sequences using entropy based metrics. In *Image Processing. 2002. Proceedings. 2002 International Conference on* (Vol. 3, pp. III-421). IEEE.

[45] Jain, A.K.; Farrokhnia, F., "Unsupervised texture segmentation using Gabor filters," *Systems, Man and Cybernetics, 1990. Conference Proceedings., IEEE International Conference on* , vol., no., pp.14,19, 4-7 Nov 1990

[46] Ngo, C.W.; Pong, T.C.; Chin, R.T., "Detection of gradual transitions through temporal slice analysis," *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* , vol.1, no., pp.,41 Vol. 1, 1999

[47] Gao, Xinbo, Jie Li, and Yang Shi. "A video shot boundary detection algorithm based on feature tracking." *Rough Sets and Knowledge Technology*. Springer Berlin Heidelberg, 2006. 651-658.

[48] Haykin, S. S. (Ed.). (2001). *Kalman filtering and neural networks* (pp. 123-174). New York: Wiley.

[49] Freund, Yoav, and Robert E. Schapire. "A desicion-theoretic generalization of on-line learning and an application to boosting." *Computational learning theory*. Springer Berlin Heidelberg, 1995

[50] Jinchang Ren; Jianmin Jiang; Juan Chen, "Shot Boundary Detection in MPEG Videos Using Local and Global Indicators," *Circuits and Systems for Video Technology, IEEE Transactions on* , vol.19, no.8, pp.1234,1238, Aug. 2009

[51] Chi-Chun Lo; Shuenn-Jyi Wang, "Video segmentation using a histogram-based fuzzy c-means clustering algorithm," *Fuzzy Systems, 2001. The 10th IEEE International Conference on* , vol.2, no., pp.920,923 vol.3, 2-5 Dec. 2001

[52] Lu Bei, Li Qiuhong. Video key-frame-extraction based on block localfeatures and mean shift clustering. Wireless Communications Networking and Mobile Computing,2010, 1-4

[53] B Gunsel, AFerman,A M Tekalp. Temporal video segmentation usingunsupervised clustering and semantic object tracking[J].Journal of Electronic Imaging, 1998;7(3):592-604.

[54] Han, Bing, XinboGao, and HongbingJi. "A unified framework for shot boundary detection." *Computational Intelligence and Security*. Springer Berlin Heidelberg, 2005. 997-1002.

[55] Y. N. Li, Z. M. Lu, and X. M. Niu, "Fast video shot boundary detection framework employing pre-processing techniques," *IET Image Process.*, vol. 3, no. 3, pp. 121–134, Jun. 2009.

[56] Gao, Yue, Jun-Hai Yong, and Fuhua Frank Cheng. "Video shot boundary detection using frame-skipping technique." *2011-02-18. http://www. cs. uky. edu/-cheng/PUBL/Paper BD 1. pdf.*

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

[58] Lienhart, R. (2001). Reliable transition detection in videos: A survey and practitioner's guide. International Journal of Image and Graphics, 1(03), 469-486.

[59] Cotsaces, C.; Nikolaidis, N.; Pitas, I., "Video shot detection and condensed representation. a review," *Signal Processing Magazine, IEEE* , vol.23, no.2, pp.28,37, March 2006

[60] Jinhui Yuan; Huiyi Wang; Lan Xiao; Wujie Zheng; Jianmin Li; Fuzong Lin; Bo Zhang, "A Formal Study of Shot Boundary Detection," Circuits and Systems for Video Technology, IEEE Transactions on , vol.17, no.2, pp.168,186, Feb. 2007.

[61] Geetha, P., and Vasumathi Narayanan. "A survey of content-based video retrieval." Journal of Computer Science 4.6 (2008): 474.

[62] Sao, Nikita, and Ravi Mishra. "A survey based on Video Shot Boundary Detection techniques."