# Shot Boundary Detection Framework For Video Editing Via Adaptive Thresholds And Gradual Curve Point

**Neetish Kumar [a], Deepa Raj [b]**
[a]Research Scholar, Department of Computer Science, BBA University (A Central University), Lucknow, India.
[b]Associate Professor, Department of Computer Science, BBA University (A Central University), Lucknow, India.

**Article History**: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

**Abstract:** Day to day huge volumes of extended videos entrained from documentaries, cinemas, athletics and surveillance cameras are evolving over video databases and in internet. Processing these videos manually is hard, costly and time-consuming. For extracting these long-duration videos an automatic procedure is desperately needed. As a vital factor the Shot boundary detection (SBD) is considered for lot of video analysis tasks, for example video editing, indexing, summarization and action recognition. In the analysis of video content SBD is considered to be one of the vital task. Based on this, we have presented an effective SBD approach. We have used the gradient and color information for abrupt transition detection. For Gradual transition detection the average edge information of the gradual curves in the sequence of frames are obtained. From the optimal edge detector an average edge frame is gained. The computational complexity is reduced by this approach by processing only the transition regions. The proposed approach when compared to the exiting work done have achieved improved results in terms of precision, recall and F1.

**Keywords:** Shot boundary detection, Gradient, Luminance, Abrupt, Gradual, Average edge.

## 1. Introduction

These days, with the rise in digitalization there is a rise in the distribution of multimedia data (particularly video) over the Internet. Owing to this surplus growth, an effective tool is desired for video retrieval and indexing. Properly recognition of the video contends are needed for designing an effective tool therefore the temporal video segmentation is required. SBD of Temporal video is task of video segmentation into significant shots by identifying the transition among sequential frames, and the boundary between two consecutive shots are marked by the transition **(Rashmi, B. S., and H. S. Nagendraswamy 2020)**. Abrupt transition and gradual transition are the two types of transitions. The formerly one is said as an amendment in the video contents in which there is no connection of frames among the margins of two shots. There is a slow change in the content of the frames in the gradual transition, some of the frames intersect in the consecutive shots and the duration of these frames is said as a gradual transition period. It is not easy to detect abrupt transition as it includes lot of complexities such as OCM and fast illumination causing extreme false positive results. The surplus volume of video contends have triggered a considerable desire for effectual schemes that can handle, store and deploy the entire contends **(Liu et al. 2017)**. For attained this, video content are added with their storage (indexing) and then with the help of video structure their basic units are examined. Because of certain video attributes the evaluation of video structure becomes a difficult task. The videos are split into simple elements as per the video structure analysis. There are four levels of video structure i.e. i) frames, ii) shots, iii) scenes, and iv) stories. A shot is considered to be the basic element of video which is believed to be the apt point for content based video indexing and retrieval (CBVIR) **(Youssef et al 2017)**. A hard transition (HT) is formed when two shots are attached directly. In contrast, employing the editing sequences (indirect concatenation) the soft transition is formed.

In videos, HTs are the most prevalent than ST **(Loukas et al. 2016)**. In the past two decades SBD has attracted the attention of many researchers. Which has classes: compressed and uncompressed domains. When compared to compressed domain the latter has gained lot of interest for valuable and fabulous visual information. Though, additional processing time is required by the uncompressed domain-based algorithms owing to the video frames decoding practice **(Singh, Alok et al. 2019)**. Some of the methods employed for feature extraction are Pixel-based, edge-based, histogram-based, and motion-based. Than the time domain, the transform domain is considered for the analysis of SBD. Because the signals are viewed in different domain which is allowed in the transform domain moreover provides a great shift concerning its potent ability for evaluating the components of the signal **(Kar, Tejaswini et al. 2017)**. An image/signal with the help of orthogonal polynomials can be to convert into transform (moment) domain from the time/spatial domain. Scalar quantities such as orthogonal moments and transform coefficients are signified using visual information. The projection of signal are represented by these moments are represent on orthogonal polynomials. The capability of orthogonal polynomials (OPs) are considered by their energy compaction and the localization properties **(Chakraborty, Saptarshi et al. 2019)**. In this work we proposed a SBD in which abrupt transition can be detected using gradient and color information. Luminance distortion and the gradient similarity are evaluated to measure each frames contrast and structural changes. In the abrupt detection phase the adaptive threshold across the videos are used to extract the transition. The Gradual transition is obtained by distribution of average edge information considering the gradual curves. Using the from optimal edge detector the average edge frame is obtained.

The complete organization of the paper is done subsequently: A brief survey about SBD is presented in Sect. 2. Sect. 3 presents the background information of the feature extraction methods employed in this study. Sect. 4 presents the brief description on the proposed SBD system. Analysis and discussion is made in the Sect. 5 which is followed by the Sect. 6 which draws the conclusion.

**2. Literature Survey**

Over the past years lot of Video SBD methods have been developed. The intention of these SBD techniques is to attain precise results rapidly. For cut transition detection these techniques offers pleasing results but until the last decade the Gradual transition detection endured to be a hitch when certain vital inventions in the field of SBD were attained. The SBD or the temporal video segmentation is a wide topic which is examined by TRECVid **Smeaton et al. (2010)**. By two methods the SBD can be evolved (1) functioning in the compressed domain and (2) uncompressed domain. In compressed domain some of the Features such as motion vectors, DCT coefficients are used for SBD. Color Histograms were presented by **Mas J et al. (2013)** for assessing video frame representation. The consecutive frames pixel intensities for Video SBD. A Singular Value Decomposition (SVD) approach was proposed by **Cernekova et al. (2007)** which was used for major feature values extraction from the non- significant values. The execution time was minimized by this approach neglecting the non-significant feature values. **(Ren et al. 2009)** proposed a SBD approach in MPEG videos using the global and local features of the frames. A motion vector was employed to detect shot transitions in the form of motion prediction error. **Adjeroh et al. (2009)** in proposed an edge based method that used multiple multilevel features for SBD. The limitation of this work is the requirement of major computational power. A Candidate Segment selection approach was proposed by **Li et al. (2009)** as a pre-processing step at this time a main development was attained in 2009 in decreasing the execution time of SBD techniques.For SBD **Lu et al (2013)** used Pattern Matching and SVD. A similar candidate segment selection was employed in this technique but with abetter adaptive threshold calculation. For SBD **GG et al. (2014)** proposed a hybrid system using color histogram and Gist. A frame transition parameters was proposed by **Tippaya et al. (2015)** combining global feature and local feature in which to classify the type of transitions a neural network is used. **Mas et al. (2003)** proposed a frame skipping approach for SBD which used an adaptive thresholding adopting a preprocessing technique. Gradual transitions were detected deploying a triangle pattern matching approach. To find out the transitions **Liang et al. (2017)** proposed a SBD using foveation technique calculating the local attention consistency measure. **Bi, Chongke et al. (2018)** to detect SBD energy, used moment and contrast in which using co-occurrence matrix the texture features are computed. **Chawla et al. (2018)** used the entropy and SURF feature for SBD in which the frames entropy is computed using the intensity histogram.

**3. Background Information**

The proposed feature extraction process is explained in this section.

**3.1 Gradient Similarity**

Information from images can be extracted using Image gradients **(Liu, Anmin et al. 2011)**. The change in intensity is measured by every pixel of a gradient image of that same point in a given direction of the original image. The gradient images in the *x* and *y* directions are calculated to attain the absolute range of direction. The gradient similarity presented is well-defined in Eqn 1.

$$G(x,y)=\frac{2(1-R)+T}{1+(1-R)^2+T} \tag{1}$$

Where

$$T=\frac{T'}{MAX\left(G_x,G_y\right)}$$

$$R=\frac{\left|G_x-G_y\right|}{MAX\left(G_x,G_y\right)}$$

The gradient values of the image blocks *x* and *y*, are signified as $G_x$ and $G_y$. *G(x, y)* is the gradient resemblance among *x* and *y* and its values in range of [0, 1]. The Gradient value $G_x$ (similar for $G_y$) is found as the supreme weighted average of difference for every block in an image. The value of T is fixed at 200.

**3.2 Luminance Similarity**

The visible distortion occurs due to the luminance changes moreover the structure also changes they are not so frustrating **(Sharma et al. 2005)**. Using the Eqn 2 the Luminance similarity is described.

$$E(x_i, y_i) = 1 - \left( \frac{x_i - y_i}{L_u} \right)^2 \quad (2)$$

In image blocks x, y the pixels at position i are represented as $x_i, y_i$, the pixel values dynamic range is denoted as $L_u$. Among the image pixels $x_i$ and $y_i$ within the range of [0, 1] the luminance similarity is represented as $E(x_i, y_i)$.

In gradient and luminance similarities the general form of integration for an image pixel pair *xi* and *yi* to obtain the overall quality indicator $Q(x_i, y_i)$ which is defined as:

$$Q(x_i, y_i) = (1 - W(G, E)).G + W(G, E).E \quad (3)$$

The shortened forms of $Q(x_i, y_i)$, $G(x_i, y_i)$ and $E(x_i, y_i)$ are denoted as Q, G, and E respectively. The relative importance of the two components are adjusted using the weighting function W(G, E).Using Eqn 4 W(G, E)is computed.

$$W(G, E) = P.G \quad (4)$$

The positive weighting parameter is denoted as ‘P’. Since P also has to lesser than 0.5 and Gis in the range of [0, 1]. In this paper the value of Pis taken as 0.1.

**3.3 CIEDE 2000 Colour Variance**

The CIELAB colour space is the main facet behind the colour-difference formula of CIEDE2000 [22]. Eqn 5 defines the CIEDE2000 colour difference.

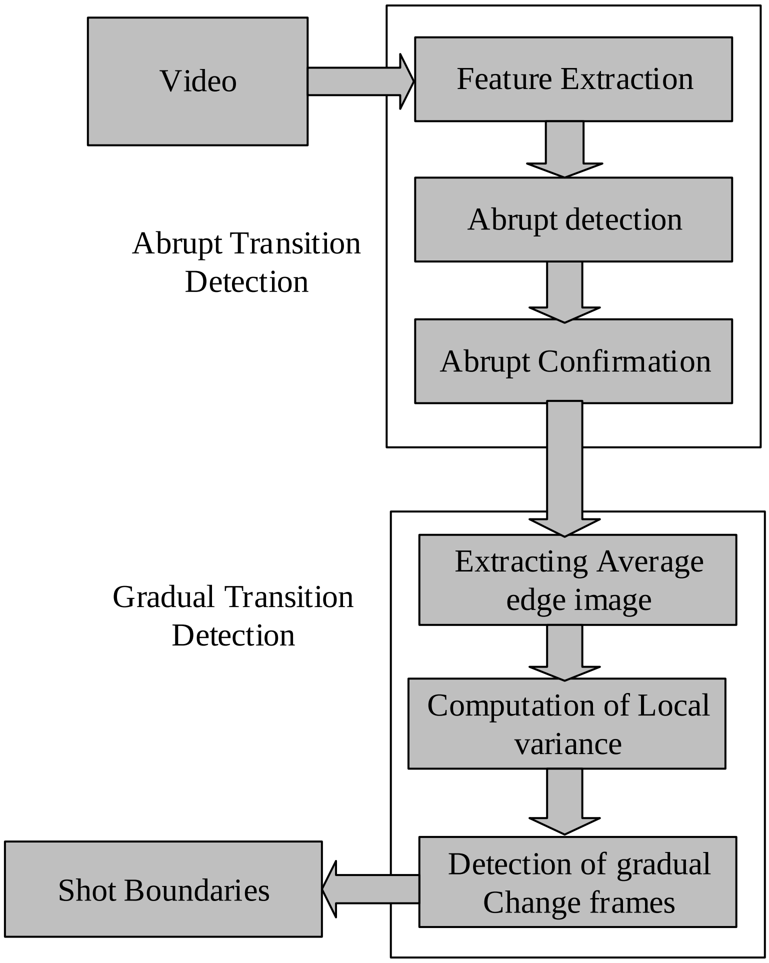
$$\Delta F = \Delta F(l_1^*, a_1^*, b_1^*, l_2^*, a_2^*, b_2^*) = \sqrt{\left(\frac{\Delta l'}{k_L s_L}\right)^2 + \left(\frac{\Delta c'}{k_C s_C}\right)^2 + \left(\frac{\Delta h'}{k_H s_H}\right)^2 + R_T \left(\frac{\Delta c'}{k_C s_C}\right)^2 + \left(\frac{\Delta h'}{k_H s_H}\right)^2} \quad (5)$$

In CIEDE2000 the differences in lightness, chroma, and hue are represented as l, c and h. The rotation function is denoted as $R_T$ and the blue regionis responsible for the interaction between chroma and hue differences. In CIEDE2000 five corrections on CIELAB have been presented and they are lightness ( $s_L$ ), chroma ( $s_c$ ), the hue ( $s_H$ ) weighting functions. The CIEDE 2000 Colour difference is applied in a Video Sequence as a rate to find the relationship across the frames.

**4. Proposed Methodology**

In this section the proposed system is explained in detail. Fig. 1illustrates the steps of the proposed system. The proposed system comprises of the feature extraction, thresholding, gradual and abrupt transition detection.

**Figure.1** The proposed block Diagram




**4.1 Feature extraction**

It is the initial procedure of the proposed system in which the features are extracted from the equivalent frames such as Luminance Similarity Comparison (*E*) and Gradient Similarity Measurement (*G*) using Eqn (1) and (4). Combining these two features the equivalent frames quality similarity Measure (*q*) is found using (3). The features *Q* and *E* are respectively used to find the possible abrupt transition detection stage. The abrupt sections Lab colour difference is used in the confirmation stage.

**4.2 Thresholding**

Among the sequential frames a threshold is used for clarifying with the abrupt or gradual changes. Beyond a certain threshold when there is a distance concerning the sequential frames then a transition is declared. It is vital to choose an apt threshold for ensuring high accuracy and for finding the both changes. A possible abrupt threshold ( $\alpha$ ) and final threshold ( $\lambda$ ) are used for identifying transitions in a video and they are represented in Eqn (6) and (7).

$$\alpha = \mu_q + \left(C_1 \times \sigma_q\right) \tag{6}$$
$$\lambda = \mu_{\Delta E} - \left(C_2 \times \sigma_{\Delta E}\right) \tag{7}$$

The possible abrupt transition is detected using the threshold $\alpha$ and in the abrupt confirmation stage $\lambda$ is used. $C_1$ used in Eqn 6 is a constant and the value lying in between the range of [-3.2, -2.8] is selected likewise in Eqn 7 the constant $C_2$ having a range of [1.6, 2].

**4.3 Abrupt section**

Generally this section is categorized into two phases i.e., Probable Abrupt Detection Stage and Abrupt Affirmation Stage.

**4.3.1 Probable abrupt detection Phase**

All the frames in a video of this phase is classed into two types i.e., probable transition frames and non-transition frames. This phase is involved to confirm that the transition like behaving frames are retained and rest are rejected. The quality indicator (Q) is used for the classification which is the combination of gradient and luminance resemblance and is computed using Eqn 3. As given in Eqn 10 the probable abrupt frames (PA) are found using the probable abrupt threshold ( $\alpha$ ).

The irregular and the former frames of $fPA_i$ are given by $fPA_{i\pm\eta}$. In Lab colour space $fPA_i$ is the probable transition frame. Experimentally $\eta$ is set as 2 which is a user defined constant.

$$PA = \begin{cases} \Pr obable \text{ tran., if } q_i \geq \alpha \\ non-tran., otherwise \end{cases} \quad (8)$$

**4.3.2 Abrupt confirmation phase**

The probable abrupt transition frames $(fPA_i)$ in this stage is stated as actual abrupt transition (A) based on the Eqn 9.

$$A = \begin{cases} True,\ if\ \Delta E\left(fPA_{i-\eta}, fPA_{i+\eta}\right) \geq \lambda \\ False, \qquad otherwise \end{cases} \quad (9)$$

Among the frames fPA(i-η) and fPA(i+η) the CIEDE2000 colour difference is E(fPA(i-η), fPA(i+η)). The $\pm\eta^{th}$ frames from the $i^{th}$ frame is indicated by *η which is a* constant. For evaluation the value of *η* is considered as 2.

**4.4 Gradual transition detection phase**

In this phase using the gradual curves which are considered by the dissemination of average edge information the Gradual transition detection is done

**4.4.1 Average edge image extraction**

An average edge image was built for each frame. It is found that than the average intensity of the pixels contain more intensities. Than the original image these images are smoother and distinct. The average edge image found with some changes was stated from effect average gradient (EAG). The average edge of image is obtained using the steps mentioned below.

1. Conversion of a color image into a gray image.
2. with the threshold 100 attain the edge image using optimal edge detector
3. Using equation (10) calculate the average gradient (AG)

$$A_G = \sum_{x,y} F(Gradient)(X,Y) / \sum_{x,y} R(X,Y) \quad (10)$$

Where $R(X,Y) = 1, if\ \text{F(Gradient)(X,Y)} > 0$

$R(X,Y) = 0, if\ \text{F(Gradient)(X,Y)} = 0$

4. Based on the average gray ( $A_G$ ) value Obtain an average edge image; using equation (11) consider a threshold value which is new.

$$F(X,Y) = \begin{cases} F \text{ Gradient(X,Y)} if \text{ F Gradient(X,Y)} < A_G \\ \quad 0, \qquad \text{if F Gradient(X,Y)} \leq A_G \end{cases} \quad (11)$$

**4.4.2 Computation of gradual point**

Based on the average gradient image detection of Gradual point is done. Initially in every 20 frames the variance was computed. To analyze the gradual change twenty frames were considered. For gradual sequence the variance is almost the same if it resembles as shot sequence. Based on equations (12) and (13) the variance is calculated.

$$VAR(x) = \frac{1}{T-1} \sum_{k=x}^{x+T-1} \left(AGI(k) - mean(k)\right)^2 \quad (12)$$

$$Mean(x) = \frac{1}{T}\sum_{k=i}^{x+T-1} AGI(k) \tag{13}$$

Where x = 1, 2 ...n – T+1

In a window the total number of frames is denoted as T

In the kth window AGI (k) is the variance of the AGI.

**4.4.3 Detection of the gradual change point**

One of the frames in the local parabolic sequence has a least value. The gradual change point was attained by analysing the width and depth of the sequence. A gradual change point is confirmed once the equations (14) and (15) are fulfilled.

$$D_{VAR} = |\omega LocMax[i \pm 1] - \omega LocMin[i] > 0 \tag{14}$$

$$D_{FRAME} = |FrmLocMax[i \pm 1] - FrmLocMax[i] < 20 \tag{15}$$

Where, i = 1, 2, 3...n

The variances of ith frame are denoted as $\omega LocalMin[i]$ and $\omega LocalMax[i]$ having the local minimum and maximum. The has local maximum frame number id denoted as $FrmLocMax[i]$

## 5. Experimental results and discussion

This section provides a detailed explanation regarding the database used and the experimentation of the employed approach.

### 5.1 Dataset

The dataset used for experimentation is the TRECvid 2007 video databases for temporal video segmentation. This benchmark dataset which is provided by *Netherlands Institute for Sound and Vision*. Compressed MPEG videos are used for this work.

### 5.2 Performance analysis

The parameters considered for the analysis of the proposed approach are Precision ($P_{re}$), Recall ($R_{ec}$) and F1 Score ($F_1$) factors, which are computed by means of Equations (16), (17) and (18).

$$R_{ec} = \frac{NC}{NC + NM} \times 100 \tag{16}$$

$$P_{re} = \frac{NC}{NC + NF} \times 100 \tag{17}$$

$$F_1 = \frac{2 * R_{ec} * P_{re}}{R_{ec} + P_{re}} \tag{18}$$

In a video the transitions which are correctly, missed and falsely detected are represented as *NC*, *NM* and *NF*. *Using the* chosen videos of TRECVid 2007 experimentation is conducted.

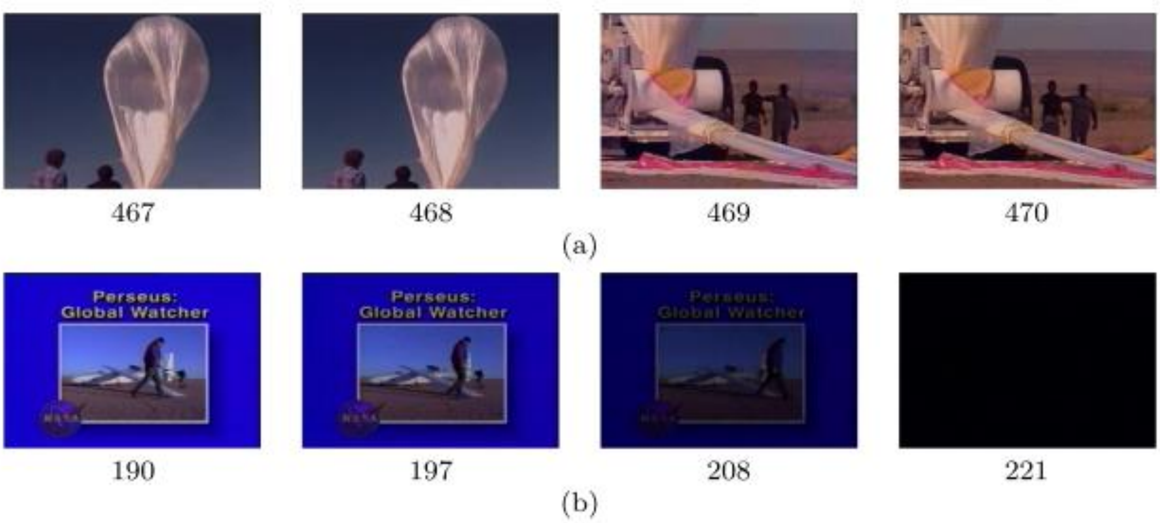
Table 1 provides the performance of some of the videos that are collected. Each videos computation time is provided in the Table 1. The correctly detected transitions from the *NAD*58.*mpg* video is represented in Figure 2.Using the videos the proposed systems average $F_1$ scores are 94.5% and 80.9% for both the transitionswith an average overall performance of 91.6%

**Table. 1.** Results of the Proposed System for TRECVid 2007 Database

| Video | Computation time(sec) | Abrupt | | | Gradual | | | Overall | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **Rec** | **Pre** | **F1** | **Rec** | **Pre** | **F1** | **Rec** | **Pre** | **F1** |
| **BG 3027** | 1630 | 84.3 | 100.0 | 92.2 | 100.0 | 51.0 | 65.7 | 95.9 | 92.2 | 95.0 |
| **BG 3097** | 1535 | 86.9 | 100.0 | 92.7 | - | - | - | 88.9 | 100.0 | 92.6 |
| **BG 3314** | 1155 | 84 | 100.0 | 89.9 | | | | 84 | 100.0 | 90.0 |
| **BG 16336** | 84 | 92.0 | 100.0 | 93.7 | | | | 96.0 | 100.0 | 96.4 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **BG 37309** | 305 | 100.0 | 100.0 | 100.0 | 85.5 | 62.6 | 74.5 | 95.5 | 82 | 88 |
| **BG 37770** | 506 | 100.0 | 100.0 | 100.0 | 91 | 78.4 | 87.3 | 94 | 84 | 89 |
| **BG 22677** | 507 | 97.2 | 100.0 | 99 | 84.4 | 87.3 | 88.8 | 92.3 | 95 | 92 |
| **BG 36658** | 871 | 94.2 | 97.2 | 96 | 86 | 89 | 87 | 92.7 | 95.4 | 94.0 |
| **BG 8947** | 545 | 94.6 | 97.3 | 95.9 | 72.2 | 100.0 | 83.9 | 90.3 | 97.7 | 92.8 |
| **BG 4455** | 772 | 95.2 | 98.6 | 96.9 | 85.3 | 90 | 88 | 93 | 96 | 94 |
| **BG 35153** | 872 | 92.0 | 100.0 | 94.2 | 87 | 82.0 | 81.6 | 87.2 | 91 | 91 |
| **Clip** | 198 | 89.5 | 100.0 | 94.2 | | | | 87.5 | 100.0 | 94 |
| Average | **664.0** | **93** | **97** | **96** | **89.2** | **81.3** | **83.2** | **91.5** | **92.6** | **92.8** |

**Figure.2** illustration of the identified **a** abrupt transition and **b** gradual transition




To validate proposed system, three SBD techniques proposed by the authors in eigen value decomposition and Gaussian transition detection method **(Amiri et al. 2012)**, temporal segmentation method **(E Santos et al. 2017)** and 3D convolutional networks method (**Liu et al. 2017)** are considered.

**Table. 2.** Summary of the assessment of the proposed scheme with the existing works

| **Process** | **Evaluation parameter** | **videos** | | | | **Average** |
|---|---|---|---|---|---|---|
| | | **Anni006** | **Anni009** | **Anni010** | **NAD58** | |
| **Proposed** | $R_{ec}$ | 83.9 | 85.5 | 87.3 | 90.2 | 86.5 |
| | $P_{re}$ | 77.3 | 90.2 | 93.6 | 94.1 | 91.7 |
| | $F_1$ | 83.6 | 87.5 | 92.4 | 92.5 | 89.6 |
| **Eigen value decomposition and Gaussian transition detection method** | $R_{ec}$ | 93.8 | 84.5 | 91.2 | 93 | 92.3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | $P_{re}$ | 85.2 | 82.1 | 78.4 | 92 | 84.5 |
| | $F_1$ | 89 | 81.5 | 83.5 | 91.4 | 87 |
| **Temporal segmentation method** | $R_{ec}$ | 84.3 | 87.9 | 88.5 | 92.8 | 89.1 |
| | $P_{re}$ | 92.2 | 86 | 84.2 | 91.5 | 86.7 |
| | $F_1$ | 91.1 | 84.2 | 85.0 | 93.7 | 89 |
| **3D convolutional networks method** | $R_{ec}$ | 92.8 | 94.3 | 84.5 | 90.2 | 91.7 |
| | $P_{re}$ | 95.5 | 80.6 | 85.6 | 90.6 | 87.3 |
| | $F_1$ | 90.2 | 88.2 | 88.9 | 92.1 | 89.7 |

Moreover, for gradual detection the obtained percentage of **85.7%** is high. The $F_1$ score as per the table 1 endure as high as **91.2%.** Table 2 illustrates the assessment made among the proposed system and the existing methods.

## 6. Conclusion and future suggestions

A SBD approach is employed in this study for the detection of abrupt and gradual transitions. The abrupt transition is detected using adaptive threshold ( $\alpha$ ).In the confirmation phase, the similarity features (gradient and luminance) are used. Along with the adaptive threshold ( $\lambda$ ), the lab features $(\lambda E)$ are used to analyse the frames. For detecting the gradual transition the gradual curve was obtained. For this initially the average edge image was found. Amongst sequences and local variance the minimum difference were considered. The gradual change point finally was identified. The presented high-quality real-time results proves that our proposed system is more enhanced that the Existing shot detection approaches for video editing based SBD which needs quality and where speed.

## References


1. Adjeroh, Don, M. C. Lee, Nagamani Banda, and Uma Kandaswamy. (2009). Adaptive edge-oriented shot boundary detection. EURASIP Journal on Image and Video Processing 2009, no. 1. 859371.
2. Amiri, Ali, and Mahmood Fathy. (2012). Video shot boundary detection using generalized eigenvalue decomposition and Gaussian transition detection. Computing and informatics 30, no. 3: 595-619.
3. Bi, Chongke, Ye Yuan, Jiawan Zhang, Yun Shi, Yiqing Xiang, Yuehuan Wang, and RongHui Zhang. (2018) Dynamic mode decomposition based video shot detection. IEEE Access 6: 21397-21407.
4. Chakraborty, Saptarshi, and Dalton Meitei Thounaojam. (2019). A novel shot boundary detection system using hybrid optimization technique. Applied Intelligence 49, no. 9: 3207-3220.
5. Chawla, Rashmi, Poonam Singal, and Amit Kumar Garg. (2018). A Mamdani fuzzy logic system to enhance solar cell micro-cracks image processing. 3D Research 9, no. 3: 34.
6. E Santos, Anderson Carlos Sousa, and Helio Pedrini. (2017) Shot boundary detection for video temporal segmentation based on the weber local descriptor. In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1310-1315. IEEE, 2017.
7. GG, Lakshmi Priya, and S. Domnic. (2014). Walsh–Hadamard transform kernel-based feature vector for shot boundary detection. IEEE Transactions on Image Processing 23, no. 12 : 5187-5197.
8. Kar, Tejaswini, and Priyadarshi Kanungo. (2017). A motion and illumination resilient framework for automatic shot boundary detection." *Signal, Image and Video Processing* 11, no. 7: 1237-1244.
9. Li, Y-N., Z-M. Lu, and X-M. Niu. (2009). Fast video shot boundary detection framework employing pre-processing techniques. IET image processing 3, no. 3: 121-134.
10. Liang, Rui, Qingxin Zhu, Honglei Wei, and Shujiao Liao. (2017). A video shot boundary detection approach based on CNN feature. In 2017 IEEE International Symposium on Multimedia (ISM), pp. 489-494. IEEE.
11. Liu, Anmin, Weisi Lin, and Manish Narwaria. (2011).Image quality assessment based on gradient similarity. IEEE Transactions on Image Processing 21, no. 4 : 1500-1512.

12. Liu, Tingxi, Yao Lu, Xiaoyu Lei, Lijing Zhang, Haoyu Wang, Wei Huang, and Zijian Wang. (2017). Soccer video event detection using 3D convolutional networks and shot boundary detection via deep feature distance. In International Conference on Neural Information Processing, pp. 440-449. Springer, Cham.
13. Loukas, Constantinos, Nikolaos Nikiteas, Dimitrios Schizas, and Evangelos Georgiou. (2016). Shot boundary detection in endoscopic surgery videos using a variational Bayesian framework. International journal of computer assisted radiology and surgery 11, no. 11, 1937-1949.
14. Lu, Zhe-Ming, and Yong Shi. (2013). Fast video shot boundary detection based on SVD and pattern matching. IEEE Transactions on Image processing 22, no. 12: 5136-5145.
15. Mas J, Fernandez G (2003) Video shot boundary detection based on color histogram, in Proc. TRECVIDCernekova, Zuzana, Constantine L. Kotropoulos, and Ioannis Pitas. Video shot-boundary detection using singular-value decomposition and statistical tests. Journal of Electronic Imaging 16, no. 4 (2007): 043012.
16. Mas, Jordi, and Gabriel Fernandez. "Video Shot Boundary Detection Based on Color Histogram. 2003. In *TRECVID*. .
17. Rashmi, B. S., and H. S. Nagendraswamy. (2020). Video shot boundary detection using block based cumulative approach. Multimedia Tools and Applications : 1-24.
18. Ren, Jinchang, Jianmin Jiang, and Juan Chen. (2009) Shot boundary detection in MPEG videos using local and global indicators. IEEE Transactions on Circuits and Systems for Video Technology 19, no. 8: 1234- 1238.
19. Sharma, G., Wu, W., & Dalal, E. N. (2005). The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur, 30(1), 21-30.
20. Singh, Alok, Dalton Meitei Thounaojam, and Saptarshi Chakraborty. (2019). A novel automatic shot boundary detection algorithm: robust to illumination and motion effect. Signal, Image and Video Processing : 1-9.
21. Smeaton, Alan F., Paul Over, and Aiden R. Doherty. (2010). Video shot boundary detection: Seven years of TRECVid activity. Computer Vision and Image Understanding 114, no. 4: 411-418.
22. Tippaya, Sawitchaya, Suchada Sitjongsataporn, Tele Tan, Kosin Chamnongthai, and Masood Khan. (2015). Video shot boundary detection based on candidate segment selection and transition pattern analysis. In 2015 IEEE International Conference on Digital Signal Processing (DSP), pp. 1025-1029. IEEE.
23. Youssef, Bendraou, Essannouni Fedwa, Aboutajdine Driss, and Salam Ahmed. (2017). Shot boundary detection via adaptive low rank and svd-updating. Computer Vision and Image Understanding 161: 20-28.