

## RESEARCH ARTICLE

# Gradual transition detection in shot boundary using gradual curve point

Asirvatham Kethsy Prabavathy<sup>1\*</sup> and J. Devi Shree<sup>2</sup>

<sup>1</sup> Department of Computer Science and Technology, Faculty of Computer Science and Engineering, Karunya University, Coimbatore, India.

<sup>2</sup> Department of Electrical and Electronic Engineering, Faculty of Electrical and Electronic Engineering, Coimbatore Institute of Technology, Coimbatore, India.

Revised: 09 January 2018; Accepted: 23 February 2018

**Abstract:** Video shot boundary detection (SBD) is a fundamental step in automatic video content analysis. Gradual transition is based on the gradual curves characterised by distribution of average edge information in the sequence of frames. An average edge frame is obtained from optimal edge detector. Features are extracted based on the minimum difference between the sequences of frames. The minimum variance in the local frame sequence is chosen as a gradual curve point. Finally, the performance is computed based on the precision (quality) and recall (quantity) value.

**Keywords:** Average edge, gradual curve, gradual transition, precision and recall, shot boundary detection.

## INTRODUCTION

Shot boundary detection (SBD) is a fundamental step in automatic video content analysis that can be used in the field of video indexing, video compression and video access. The objective of SBD is to detect the visual discontinuities along with the time domain. During this detection, the extracted visual features are used to measure the degree of similarity between frames in a given shot.

A video is hierarchically organising the structure of frames, shots, scenes and stories. A frame is at the lowest level in the hierarchy; it is the basic unit of a video, and also it is representing a still image. SBD techniques are used to group the set of frames into shots, which means a shot designates a continuous sequence of view frames

recorded by an uninterrupted camera operation. A scene is a collection of shots which presents different views of the same object of interest. A story is a collection of scenes that defines an unbroken event. The structure of a video is shown in Figure 1.

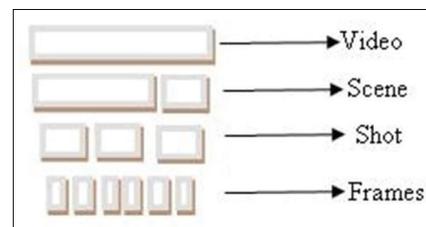


Figure 1: Video structure

The transitions between shots are divided into two types: abrupt and gradual, while shot boundaries can be categorised into two types: cut transition (CT) and gradual transition (GT). GT can be further classified into dissolve, wipe, fade (fade out and fade in), etc. (Chavan & Akojwar, 2013). SBD, also known as temporal video segmentation, is the process of identifying the transitions between the adjacent frames. In the perspective of visual aspect, video is a kind of three-dimensional signal in which two of them reveal the visual content in the horizontal and vertical frame direction, and the third one reveals the variations of the visual content over the time axis. SBD is mainly based on the representation

\* Corresponding author (kethsy\_cse@karunya.edu;  <https://orcid.org/0000-0002-6925-6480>)



of visual content, construction of continuity signal and classification of continuity values (Thounaojam *et al.*, 2016; Tippaya *et al.*, 2017).

GT techniques are generally complex methods to build a general model to detect the type and location of these transitions due to various types of editing techniques and various lengths of GTs (Moeglein *et al.*, 2017). Dissolve transition means the last few frames of the disappearing shot temporally overlap with the first few frames of the appearing shot. Here, the intensity of the disappearing shot reduces from normal to zero, which is known as fade out. When the intensity of the appearing shot rises from zero to normal it is known as fade in. In fade transition, the disappearing shot fades out into the blank frame and then the blank frame fades into the appearing shot (Liu & Dai, 2016).

Wipe transition is a group of techniques for changing the shot in which the appearing and disappearing shots exist at the same time in various spatial areas of the intermediary video frames, and the region taken up by the former develops until it completely takes the place of the latter (Fan *et al.*, 2017).

A large number of SBD techniques have been developed, however, most of the researchers have done so based on the hard cut or cut detection (Heng & Ngan, 2002; Bescos *et al.*, 2005; Yuan *et al.*, 2007; Li *et al.*, 2009). Gradual transition is a more challenging technique than cut transition. Many researchers have developed techniques related to gradual transition, but still it is a challenging one.

A B-spline interpolation curve filtering technique of gradual transition in video sequence was used to find the number of wipes and dissolve automatically (Li *et al.*, 2010). SIFT key-point matching algorithm was used to capture the changing statistics of different kinds of gradual transitions such as fade, wipe and dissolve accordingly. This SIFT method was mainly developed to reduce the computational cost and to improve the accuracy (Qin *et al.*, 2010).

A fast k-step slipped window method was used to select the candidates of the shot boundary using k-step slipped window and adaptive thresholds. Then it detects hard cuts and flash exclusion and gradual transition in an orderly manner (Liu *et al.*, 2003). A moving query window technique was developed to detect the gradual transition. The query window method was used to find the inter-frame distance between the number of frames in an input video sequence (Yufeng *et al.*, 2010). Wipe transition is one of the important types of gradual

transition detection. A colour sub-image and edge sub-image of each frame were decomposed by DB-4 wavelet transition. To detect the wipe transition, the image is divided into the sub-image of 8x8 pixel blocks and Gaussian mode is applied to adjust the threshold value of the input image automatically (Truong *et al.*, 2000).

An adaptive threshold method was used to reduce the artefacts that are created by noise and motion in scene cut detection (Mohanta *et al.*, 2012). A frame estimation scheme was used to find the previous and next frame in the video sequence. The scatter matrix of edge strength and motion matrix were calculated using frame transition parameters and frame estimation errors (Tsamoura *et al.*, 2008). The colour coherence change technique exhibits less sensitivity to local or global motion. Instead of threshold selection a machine learning technique was used here. To detect the gradual transition instead of threshold selection, a combined meta-segmentation scheme with local and global motion of the video was used (Su *et al.*, 2005).

Dissolve type transition algorithm are used to accurately distinguish the dissolve from disturbance caused by motion. The dissolve was modelled based on its nature and filtered out by possible confusion caused by the effect of motion *via* statistics (Won *et al.*, 2003). In this method, dissolve detection was based on the analysis of a dissolve modelling error. This was calculated based on the difference between an ideally modelled curve without any correlation and an actual variance curve with a correlation (Yu *et al.*, 2011).

A twin-comparison approach was utilised to detect the dissolve candidates. A special curve was filtered according to the inherent characteristic of dissolves indicated by two parameters: (i) percentage of dissolve-supporters in a frame; (ii) variance of coefficients in the dissolve model. Using this curve function, non-dissolve transitions are removed (Ling *et al.*, 2005). Gradual transitions are detected based on the variance projection function. Using this variance projection function, the distance between the video frame and the feature vector of changes in the distance was calculated. Finally using support vector machine (SVM) it was determined whether the changes of the distances are caused by gradual transition or not (Lu *et al.*, 1999). The detection technique utilise the properties of QR-decomposition and extracts a block-wise probability function that illustrates the probability of video frames to be in shot transitions. The technique detects the shot boundary transition by analysing this probability function (Amiri & Fathy, 2009).

**METHODOLOGY**

**Extracting the average edge image**

For each frame, an average edge image was constructed based on the pixels. It is identified that the pixels have more intensities than the average intensity. A block diagram of the proposed gradual transition detection is given in Figure 2. These images are distinct and smoother than the original image. Calculating the average edge image with some changes was referred from effect average gradient (EAG) (Volkmer & Tahaghoghi, 2004). The following steps were used to find the average edge of image.

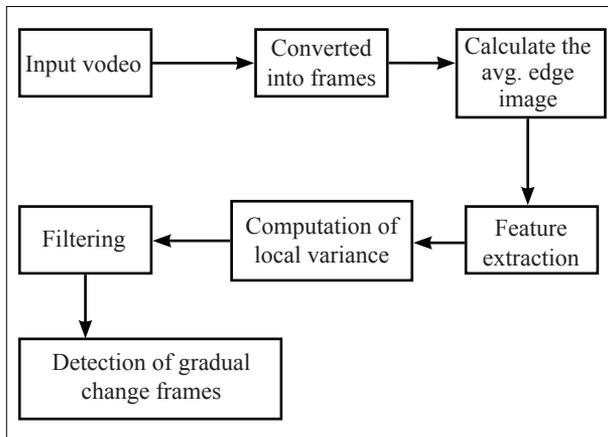
1. Convert a colour image (frame) to a gray image (frame). The distribution of a gray image is shown in Figure 3.
2. Obtain the edge image by using an optimal edge detector for the original image with threshold 100. The distribution of an average edge image is shown in Figure 4.
3. Compute an average gradient (AG) using equation (1).

$$AG = \sum_{x,y} f(\text{Gradient})(x,y) / \sum_{x,y} P(x,y) \quad \dots(1)$$

where  $P(x,y) = 1$ , if  $f(\text{Gradient})(x,y) > 0$   
 $P(x,y) = 0$ , if  $f(\text{Gradient})(x,y) = 0$

4. Find an average edge image based on the average gray (AG) value; consider as a new threshold value using equation (2).

$$f(x,y) = \begin{cases} f \text{ Gradient}(x,y), & \text{if } f \text{ Gradient}(x,y) < AG \\ 0, & \text{if } f \text{ Gradient}(x,y) \leq AG \end{cases} \quad \dots(2)$$

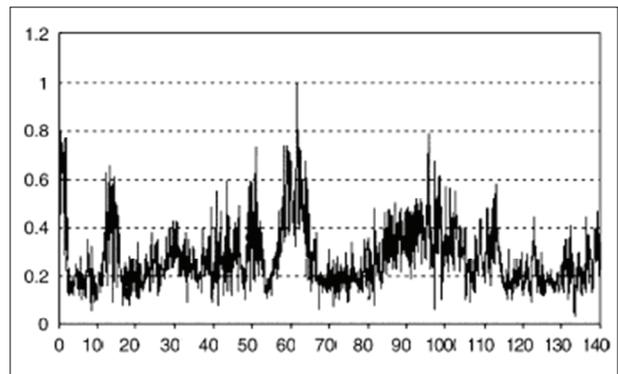


**Figure 2:** Block diagram for gradual transition detection

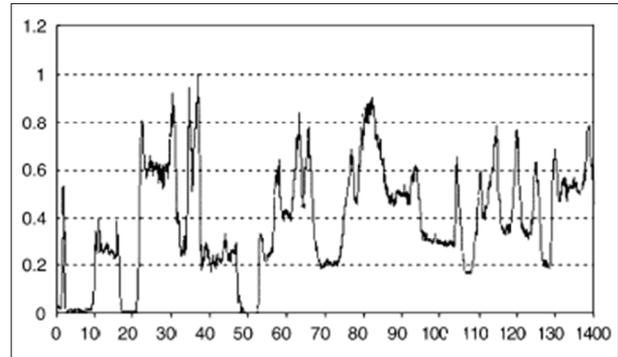
**Finding the minimum distance between sequences**

To extract nine different types of variances from nine equal sized non-overlapping blocks as shown in Figure 5 an average edge image was used to maximise the property of gradual transition.

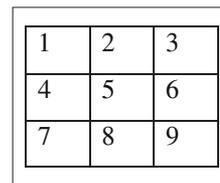
To search the blocks the maximum depth and width of parabolic curve of the image was taken. Gradual sequences were obtained by computing the maximum and minimum difference sequences between the overall block and each block. These gradual sequences show more distinct parabolic shapes compared with variance sequence frames (Thounaojam *et al.*, 2016).



**Figure 3:** Distribution of a gray image



**Figure 4:** Distribution of an average edge image



**Figure 5:** Equal sized blocks

### Computation of gradual point

Gradual point is detected based on the average gradient image. First, the variance in every 25 frames was calculated. Twenty five frames were taken as a sliding local window to examine the gradual change. If it is a similar shot sequence, the variance is almost the same for gradual sequence. Variance is computed based on equations (3) and (4).

$$\text{Var}(i) = \frac{1}{L-1} \sum_{k=i}^{i+L-1} (\text{AGI}(k) - \text{mean}(k))^2 \quad \dots(3)$$

$$\text{mean}(i) = \frac{1}{L} \sum_{k=i}^{i+L-1} \text{AGI}(k) \quad \dots(4)$$

where  $i = 1, 2 \dots n - L + 1$

$L$  - total number of frames in a window (25)

$\text{AGI}(k)$  - variance of the AGI in  $k^{\text{th}}$  window.

### Reducing distortions

To obtain a more accurate detection, an additional filtering procedure was implemented for reducing the distortion within the curve. This was achieved by using a morphological opening function in the curve. Then the curve became smooth.

### Detection of the gradual change point

In a gradual change, one of the frames has a minimum value in the local parabolic sequence. The gradual change point was detected based on the sequence width and depth. If equations (5) and (6) are satisfied, then it is decided as a gradual change point.

$$D_{var} = |\sigma_{locmax}[i \pm 1] - \sigma_{locmin}[i]| > 0 \quad \dots(5)$$

$$D_{frame} = |Frm_{locmax}[i \pm 1] - Frm_{locmax}[i]| < 25 \quad \dots(6)$$

where,

- $i = 1, 2, 3 \dots n$
- $\sigma_{localmin}[i]$  and  $\sigma_{localmax}[i]$  are variances of  $i^{\text{th}}$  frame that have the local minimum and maximum, respectively,
- $Frm_{locmax}[i]$  - a frame number that has local maximum.

Table 1 shows the precision and recall values for gradual transition detection using gradual curve method.

**Table 1:** Results for gradual shot boundary detection

Programme	Precision	Recall
News	96	96.8
Cartoon	91.2	94.3
Sports	96.3	98.3
Movie	90.3	94.9

## RESULTS AND DISCUSSION

Experiments were performed using MatLab (2010). The video clips were obtained from the Internet and various television programmes, and the complete video database was made available upon request. The precision in equation (7) and recall in equation (8) were computed to evaluate the gradual transition.

Precision, as a measure of quality, which is the percentage of correct detection:

$$\text{Precision} = \frac{\text{Number of true positive detections}}{\text{Number of tests result in as positive}} \times 100 \quad \dots(7)$$

Recall, as a measure of quantity, which is the percentage of true detection

$$\text{Recall} = \frac{\text{Number of true positive detections}}{\text{Number of tests result in as positive}} \times 100 \quad \dots(8)$$

Table 2 shows the comparison of the gradual curve method with QR decomposition and Gaussian transition detection.

## CONCLUSION

A gradual transition detection using gradual curve was obtained based on the proposed method from various videos (news, cartoon, sports and movie). First, the average edge image was identified. Then minimum difference between sequences and local variance were calculated. Finally, gradual change point was identified. The gradual curve method was compared with QR decomposition and Gaussian transition. Evaluation

**Table 2:** Comparison with QR decomposition techniques

Programme	Average of our techniques		QR decomposition techniques	
	Precision	Recall	Precision	Recall
News (Avg)	96	96.8	94	98
Cartoon (Avg)	91.2	94.3	89	92
Sports (Avg)	96.3	98.3	95	97
Movie (Avg)	90.3	94.9	91	95

proved that the gradual curve method was more reliable than the QR decomposition method. In future, applying fuzzy rules will improve the quality and quantity, that is, it will increase the precision and recall values.

## REFERENCES

- Amiri A. & Fathy M. (2009). Video shot boundary detection using QR-decomposition and gaussian transition detection. *EURASIP Journal on Advances in Signal Processing* **9**: 12 – 15.  
DOI: <https://doi.org/10.1155/2009/509438>
- Bescos J., Cisneros G., Martinez J.M., Menendez J.M. & Cabrera J. (2005). A unified model for techniques on video-shot transition detection. *IEEE Transactions on Multimedia* **7**(2): 293 – 307.  
DOI: <https://doi.org/10.1109/TMM.2004.840598>
- Chavan S.A. & Akojwar S.G. (2013). Gradual transition detection algorithms in video segmentation: a survey. *International Journal of Scientific and Engineering Research* **4**(2): 1 – 5.
- Fan J., Zhou S. & Siddique M.A. (2017). Fuzzy color distribution chart-based shot boundary detection. *Multimedia Tools and Applications* **76**(7): 10169 – 10190.
- Heng W.J. & Ngan K.N. (2002). Shot boundary refinement for long transition in digital video sequence. *IEEE Transaction on Multimedia* **4**(4): 434 – 445.  
DOI: <https://doi.org/10.1109/TMM.2002.806532>
- Li J., Ding Y., Shi Y. & Li W. (2010). A divide-and-rule scheme for shot boundary detection based on SIFT. *International Journal of Digital Content Technology and its Applications* **4**(3): 202 – 214.
- Li Y.N., Lu Z.M. & Niu X.M. (2009). Fast video shot boundary detection framework employing pre-processing techniques. *IET Image Processing* **3**(3): 121 – 134.  
DOI: <https://doi.org/10.1049/iet-ipr.2007.0193>
- Ling J., Lian Y.Q. & Zhuang Y.T. (2005). A new method for shot gradual transition detection using support vector machine. *Proceedings of the 4<sup>th</sup> International Conference on Machine Learning and Cybernetics*, Guangzhou, China, pp. 5599 – 5604.  
DOI: <https://doi.org/10.1109/ICMLC.2005.1527934>
- Liu T., Zhang H. & Qi F. (2003). A novel video key-frame-extraction algorithm based on perceived motion energy model. *IEEE Transactions on Circuits and Systems for Video Technology*, volume 13, no. 10, pp. 1006 – 1013.
- Liu X. & Dai J. (2016). A method of video shot-boundary detection based on grey modeling for histogram sequence. *International Journal of Signal Processing, Image Processing and Pattern Recognition* **9**(4): 265 – 280.  
DOI: <https://doi.org/10.14257/ijcip.2016.9.4.25>
- Lu H.B., Zhang Y.J. & Yao Y.R. (1999). Robust gradual scene change detection. *International Conference on Image Processing* **3**: 304 – 308.  
DOI: <https://doi.org/10.1109/ICIP.1999.817122>
- Moeglein W.A., Griswold R., Mehdi B.L., Browning N.D. & Teuton J. (2017). Applying shot boundary detection for automated crystal growth analysis during *in situ* transmission electron microscope experiments. *Advanced Structural and Chemical Imaging* **3**(2).  
DOI: <http://doi.org/10.1186/s40679-016-0034-x>
- Mohanta P.P., Saha S.K. & Chanda B. (2012). A model-based shot boundary detection technique using frame transition parameters. *IEEE Transactions on Multimedia* **14**(1): 223 – 233.  
DOI: <https://doi.org/10.1109/TMM.2011.2170963>
- Qin T., Gu J., Chen H. & Tang Z. (2010). A fast shot-boundary detection based on K-step slipped window. *2<sup>nd</sup> IEEE International Conference on Network Infrastructure and Digital Content*, 24 – 26 September, Beijing, China, pp. 190 – 195.  
DOI: <https://doi.org/10.1109/ICNIDC.2010.5657841>
- Su C.W., Liao H.Y.M., Tyan H.R., Fan K.C. & Chen L.H. (2005). A motion-tolerant dissolve detection algorithm. *IEEE Transactions on Multimedia* **7**(6): 1106 – 1113.  
DOI: <https://doi.org/10.1109/TMM.2005.858394>
- Thounaojam D.M., Khelchandra T., Singh K.M. & Roy S. (2016). A genetic algorithm and fuzzy logic approach for video shot boundary detection. *Computational Intelligence and Neuroscience* **2016**: Article ID 8469428.  
DOI: <http://doi.org/10.1155/2016/8469428>
- Tippaya S., Sitjongsataporn S., Tan T., Khan M.M. & Chamnongthai K. (2017). Multi-modal visual features-based video shot boundary detection **5**: 12563 – 12575.  
DOI: <https://doi.org/10.1109/ACCESS.2017.2717998>
- Truong B.T., Dorai C. & Venkatesh S. (2000). New enhancements to cut, fade, and dissolve detection processes in video segmentation. *Proceedings of the 8<sup>th</sup> ACM International Conference on Multimedia*, Marina Del Ray, USA, pp. 219 – 227.  
DOI: <https://doi.org/10.1145/354384.354481>
- Tsamoura E., Mezaris V. & Kompatsiaris L. (2008). Gradual transition detection using color coherence and other criteria in a video shot meta-segmentation framework. *2008 15<sup>th</sup> International Conference in Image Processing (ICIP – 2008)*, 12 – 15 October, San Diego, USA, pp. 45 – 48.  
DOI: <http://doi.org/10.1109/ICIP.2008.4711687>
- Volkmer T., Tahaghoghi S.M.M. & Williams H.E. (2004). Gradual transition detection using average frame similarity. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)*, 27 June – 02 July, Washington DC, USA, p. 139.  
DOI: <https://doi.org/10.1109/CVPR.2004.357>

- Won J.U., Chung Y.S., Kim I.S., Choi J.G. & Park K.H. (2003). Correlation based video-dissolve detection. *Proceedings of the International Conference on Information Technology Research and Education (ITRE2003)* (eds. G. Noreddine & J. Kacprzyk), pp. 104 – 107.
- Yu F., Lu Z. & Li Y. (2011). Dissolve detection based on twin-comparison with curve fitting. *International Journal of Innovative Computing, Information and Control* 7(5): 2417 – 2426.
- Yuan J.H., Wang H., Xiao Y.L., Zheng W.J., Li J.M., Lin F.Z. & Zhang B. (2007). A formal study of shot boundary detection. *IEEE Transactions on Circuits and System for Video Technology* 17(2): 168 – 186.  
DOI: <https://doi.org/10.1109/TCSVT.2006.888023>
- Yufeng L., Yinghua Y. & Guiju L. (2010). A novel wipe transition detection method based on Multi-Feature. *3<sup>rd</sup> International Conference on Knowledge Discovery and Data Mining*, 9 – 10 January, Phuket, Thailand, pp. 451 – 454.  
DOI: <https://doi.org/10.1109/WKDD.2010.99>