The Journal of Applied Sciences Research. 2(2): 103-110, 2015

T.J.A.S.R

The Journal of Applied Sciences Research

Journal homepage: http://www.journals.wsrpublishing.com/index.php/tjasr

Online ISSN: 2383-2215 Print ISSN: 2423-4400

Original Article

Scene Change Detection using Block Processing Method

Dolley Shukla[1,*], Manisha Sharma[2], Chandra Shekhar Mithlesh[3]

[1] Associate Professor, SSCET, Bhilai, India
[2] Professor & H.OD (E & Tc) BIT, Durg, India
[3] ME Scholar, SSCET, Bhilai, India

ARTICLE INFO

Corresponding Author:
Dolley Shukla
dolley020375@gmail.com

How to Cite this Article:
Shukla, D., Sharma, M., & Mithlesh, C.S. (2015). Scene Change Detection using Block Processing Method. *The Journal of Applied Sciences Research,* 2(2): 103-110.

Article History:
Received: 29 March 2015
Revised: 10 July 2015
Accepted: 13 July 2015

ABSTRACT

A various number of Scene Change Detection (SCD) methods are proposed in recent years. Today's research topic on video is video summarization or abstraction, video classification, video annotation and content based video retrieval. Applications of scene change detection are video indexing, semantic features extraction, multimedia information systems, Video on demand (VOD), digital TV, online processing (networking) neural network mobile applications services and technologies, cryptography, and in watermarking. Therefore, in this paper we provide block processing scene change detection techniques & algorithms.

Keywords: Video sequences, scene change, detection, scene change detection, recall, precision, threshold, F-measure, frame rate and sample time.

INTRODUCTION

Multimedia data gets larger day by day with the extensive usage of digital technology that becomes inexpensive and popular. Moreover, video data holds an important part of it and, searching such a tremendous media and finding relevant Information requires an efficient organization and deploying efficient tools (Sakarya and Telatar, 2010).With the development of various multimedia compression standards coupled with significant increases in desktop computer performance and storage, the widespread exchange of multimedia information is becoming a reality. Audio-visual information is becoming available in digital form in various places around the world (Fernando, Canagarajah and Bull, 2001).Internet packet(IP) video is emerging as an important area of multimedia application, especially with increasing bandwidths available over the network today. IP multicast is an effective means for data dissemination and sharing among a large user community (Zhou1 and Vellaikal, 2001). With rapid advances in multimedia communication and due to those applications for digital video like Video telephony, streaming media delivery on internet, CD, DVD Cellular media, educational purpose etc. requires large storage space. Video Compression is necessary to eliminate picture redundancy, allowing video information to be transmitted and stored in a compact and efficient manner. In efficient way of video processing, segmentation of video is necessary. Segmentation is done with the help of scene change detection. In multimedia communication, digital video consists of frames, scene and shots. In any shot one continuous action is going on, captured by a camera, and is a sequence of several frames. For scene change detection two consecutive frames gets matched with each other. Due to the large amount of data contained by video, it is often compressed for efficient storage and transmission. One of the most efficient and popular technique through which we can reduce the temporal redundancy between adjacent frames of a video sequence is called motion estimation (ME) (P. Chauhan *et al*., 2013). Typically scene changes can be detected by inspecting the differences of successive frames. Due to the nature of MPEG compression, the encoded DCT values and motion vectors of a frame already contain information about these differences. Thus, when working in the compressed domain, we can reuse this existing information to decide about the existence of a scene change on a frame by frame basis. When analyzing the DCT values as well as the motion information we can also detect special movements like rotations or zooms within the video stream. For the adaptation of compressed video, detection of scene changes and special movements can provide helpful information to determine suitable adaptation parameters as well as to decide whether a certain frame should be skipped (Brandt, Trotzky and Wolf, 2008).The H.264 advanced video coding (H.264/AVC) standard provides several advanced features such as improved coding efficiency and error robustness for video storage and transmission. In order to improve the coding performance of H.264/AVC, coding control parameters such as group-of-pictures (GOP) sizes should be adaptively adjusted according to different video content variations (VCVs), which can be extracted from temporal deviation between two consecutive frames (Ding and Yang, 2008). The H.264 is the newest compression digital video codec standard issued by the ITU-T Video Coding Experts Group (VCEG) together with the ISO/IEC Moving Picture Experts Group (MPEG). Because of its many new features and excellent performance, H.264 achieves higher compression efficiency than previous video standards. This makes H.264 well suited for transmission over bandwidth limited systems, such as mobile wireless systems (Ding and Yang, 2008). Images may be two-dimensional, such as photograph, screen display, and as well as a three-dimensional, such as hologram. They may be *captured* by optical such as cameras, mirror, lenses, telescope etc. The pixel (a word invented from "picture element") is the basic unit of programmable color on a computer display or in a computer image. A typical video sequence it's organized into a sequence of group of pictures (GOP) (fig. 1). A video can be broken down in scene, shot and frames. Each GOP consists of one I (intra-coded) and a few P (predicted) and B (bi-directionally interpolated) frames. A scene is a logical grouping of shots into a semantic unit. A shot is a sequence of frames captured by a single camera in a single continuous action. To create indexed video databases, video sequences are first segmented into scenes, a process that is referred to as scene change detection (M. Ford *et al*., 1997). Automatic annotation of the input video needs to be developed. Content-based temporal video segmentation is mostly achieved by detection and classification of scene changes (transitions). Basically, transitions can be divided into two categories: abrupt transitions and gradual transitions. Gradual transitions include camera movements such as panning, tilting and zooming, and video editing special effects such as fade-in, fade-out, dissolving and wiping. Abrupt transitions are very easy to detect, as the two frames are completely uncorrelated (Fernando, Canagarajah and Bull, 1999). A shot boundary is the transition between two shots .A digital video consists of frames that are a single frame consists of pixels (Adhikari *et al*., 2008).A video shot is a sequence of successive frames with similar visual content in the same physical location. Therefore, abrupt scene change can be identified when there is a large amount of visual content change across two successive frames while gradual transition is much difficult to detect since the successive frame difference during transition is substantially reduced (Chau *et al*., 2005).The segmentation of a video sequence into individual shots by scene change detection is fundamental for automatic video indexing, scene browsing and retrievals. Once individual video shots are identified, we can index and query video contents by using content-based indexing. It can be expected that large amount of video materiel will be available only in compressed form very soon. Thus it is necessary to develop techniques for scene change detection with direct manipulation of compressed video sequences on the compression domain(Shin et al., 1998).The segmentation process is generally referred to as shot boundary detection or scene change detection. A shot is sequence of frames generated during a continuous camera operation and represents a continuous action in time and space. Video editing procedures produce abrupt and gradual scene changes. A cut is an abrupt scene change that occurs in a single frame. Gradual change occurs over multiple frames and is the product of fade-ins, fade-outs, or dissolves (where two shots are superposed (Lee *et al.*, 2006).A variety of data in the form of image and video are being generated, stored, transmitted, analyzed, and accessed with advances in computer technology and communication network. To make use of these data, an efficient and effective technique needs to be developed for the retrieval of multimedia information. A video stream consists of a sequences of images accompanied with an optional audio stream. The video stream can have arbitrary length. The illusion of moving picture is generated by showing rapidly a sequential series of still images. The human eye is too slow to see the individual still images, and interprets the series of images as continuous movement. Typically the frame rates of videos are about 24 orm30 images per second, which is enough to generate the illusion of continuous movement. Edited video typically consists of multiple shots. A shot is a series of sequential frames that have been filmed in a single camera run. Shots are bound together using several kinds of cuts or transition effects. In movie terminology one or more shots form a scene, which is usually defined to be a set of consecutive shots that form an entity that seems to be continuous in time, or is filmed in a single location (Chhasatia, Trivedi and Shah, 2013). A fade is a gradual transition between a scene and a constant image (fade out) or between a constant image and a scene (fade in). A dissolve is a gradual transition from one scene to another, in which the first scene fades out and the second scene fades in (Huang and Liao, 2001). For the copy protection of video using watermarking, scene change detection is important step (Shukla and Sharma, 2012). To detect scene change, a dissimilarity measure between consecutive frames is often used. The measure must distinguish true scene changes from those that are not. A simple measure is mean absolute frame differences (MAFD) for the *n*th frame (Yi and Ling, 2005).

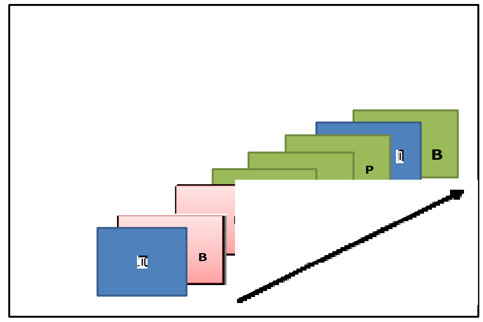


Figure 1: Group of Picture (GOP)

Scene Detection

a) ***Color Histograms.*** frame-to-frame similarities based on colors which appeared within them and positions of those colors in the frame. After computing the inter-frame similarities, a threshold can be used to indicate shot boundaries.

b) ***Edge Detection*****:** Edge Detection is based on detecting edges in two neighboring images and comparing these images (Adhikari *et al.*, 2008).

c) Macroblocks**:** Besides, they investigated the shot boundary detection using macroblocks. Depending on the types of the macroblock the MPEG pictures have different attributes corresponding to the macroblock. Macroblock types can be divided into forward prediction, backward prediction or no prediction at all. The classification of different blocks happens while encoding the video file based on the motion estimation and efficiency of the encoding. If a frame contains backward predicted blocks and suddenly does not have any, it could mean that the following frame has changed drastically which would point to a cut. This approach, however, becomes difficult to implement when there is a shot change, and the frame in the next shot contains similar blocks as the frame before.

PROPOSED METHOD

In the method proposed, a video is used as input signal. The video is first converted into image frames. These input image frames is pass through the feature extraction block set. Feature Extraction Block set convert input image frame into edge image signal. This image signal is input of the Edge Comparison block set, compared Edge is greater the some specific threshold value then scene change occurs, Number of change block is displayed in time offset parameter.

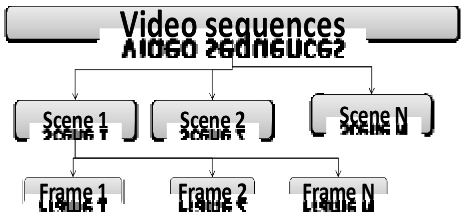


Figure 2: General Video Structure

Boundaries Approaches

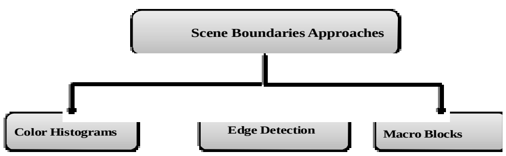


Figure 3: Scene Boundaries Approaches

A. Input unit

The Input Block is From Multimedia File block that reads audio frames, video frames, or both from a multimedia file. The block imports data from the file into a Simulink model. The block support following video files (.qt, .mov, .avi, .asf, .asx, .wmv, .mpg, .mpeg, .mp2, .mp4).

Sample Rate: It is rate of video in frames per second (FPS) i.e.

$$Sample\ Time = \frac{1}{FPS}$$

RBG Component**:** Matrix $(m \times n)$ that represents one plane of the RGB video stream. Outputs from the R, G, or B ports must have same dimensions $(m \times n)$.

B. Feature Extraction unit

One of the RGB component is chosen as input of the feature extraction block.we select different Edge Detection method like Sobel, Prewitt, or Roberts.

The Edge Detection method finds the edges in an input image by approximating the gradient magnitude of the image. The block convolves the input matrix with the Sobel, Prewitt, or Roberts kernel. The outputs two gradient components of the image, which are the result of this convolution operation. Alternatively, the block can perform a thresholding operation on the gradient magnitudes and output a binary image, which is a matrix of Boolean values. If a pixel value is 1, it is an edge. The Edge Detection block computes the automatic threshold values using an approximation of the number of weak and non edge image pixels.

Binary output image is converted into image signal with the help of Data type conversion block. This output is an input of Edge Comparison method and edge of video output block.

C. Edge Comparison Unit

a) ***Block Processing***: This block extracts sub matrices of a user-specified size from the input matrix. It sends each sub matrix to a subsystem for processing, and then reassembles each subsystem output into the output matrix. We are using Block size 32×32.

b) Edge Difference: This Subsystem block represents a subsystem of edge difference that contains it, that is absolute value of current and previous frames.

c) Edge Matching: This Subsystem block represents a edge matching. In this unit the of mean value is taken as input and comparing to a constant value, reshape the dimensions of a vector or matrix input signal in two dimensional Column vector. The sum is performs addition on its inputs vector. This block can add vector, or matrix inputs.

d) Thresholding: comparing specific threshold value, if changes occurs scene change detected and change block is detected. One of the output is connected to the number of change block and video display unit.

D. Number of Change Block

This unit is nothing but a Scope block that work to displays its input with respect to simulation time. The Scope block can have multiple axes (one per port) and all axes have a common time range with independent y-axes. The Scope block allows to adjust the amount of time and the range of input values displayed. we can move and resize the Scope window and you can modify the Scope's parameter values during the simulation.

E**. Output unit**

The output Block is **to Multimedia File block** that reads audio frames, video frames, or both from a multimedia file. The block display data from the file into a Simulink model. The block support following video files (.qt, .mov, .avi, .asf, .asx, .wmv, .mpg, .mpeg, .mp2, .mp4).input of this unit is connected with Feature Extraction unit output and Edge Comparison Unit output.

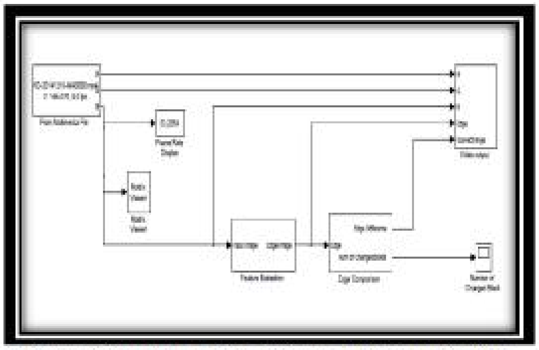


Figure 4: Simulink block of Scene Change Detection.

EXPERIMENTAL RESULT

In this experiment, we observed the video scene change detection method under MATLAB 7.10.0(R210a) environment**.** The size of the video files (.qt, .mov, .avi, .asf, .asx, .wmv, .mpg, .mpeg, .mp2, .mp4) used is 270×1280 and Block Processing in each frames 32×32 block size. The scene Change Detection is shown in figure.

For the evaluation of the proposed method, the video data base which are generally with in ITU-T study Group 12 for standardization of the quality models are used. The resolution of the video sequences varied from SD (720×576), to HD-720p (720×1280) and HD-1080p (1920×1080), in interlaced and progressive formats and frame rates ranging from 25 FPS to 60 FPS. We are used HD-720p (720× 1280) size of video sequences. Table 1 shows the various parameter which contain the video sequences. This parameters are File Type, Size, Length, Frame Per Second (FPS) and Sample Time (T_s).

We are taken the three different video sequence and pass through the Frame rate display unit within 0 to 10 sec offset time interval and analyzing the frame rate with each 2 sec interval that is shown in figure 2.

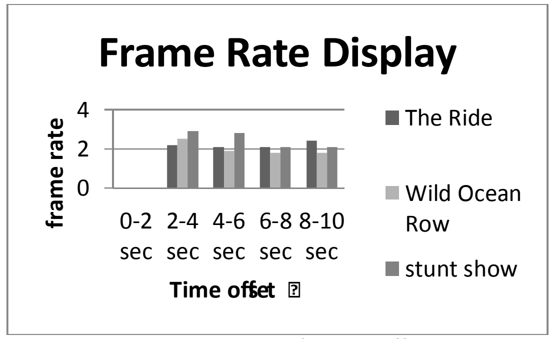


Figure 5: Frame Rate graph of Three Different Video Sequences

The result are reported based on Precision and Recall score, which are widely used for evaluating classification of a system.**"Precision**" defines how reliable the detected by the algorithm is while "**Recall"** defines the overall performance of the algorithm. Result are reported using F-measure which is combine precision and recall in a single measure. These measure are computed by

$$Recall = \left(\frac{correct}{correct + false}\right)$$

$$Precision = \left(\frac{correct}{correct + missed}\right)$$

$$F = 2 \times \left(\frac{precision \times recall}{precision + recall}\right)$$

The selected parameters for the Scene Change Detection algorithm mention in section- III in a table 3 with two different video sequences.

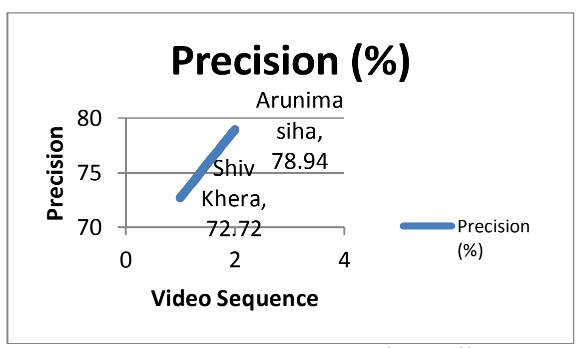


Figure 6: Precision percentage value of two different video sequence

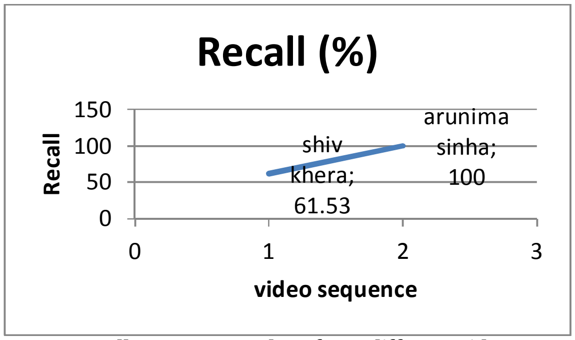


Figure 7: Recall percentage value of two different video sequence

Table 1: HD-720p (720×1280) Video Sequence Parameters

S N	Video Sequence	File Type	Size (MB) (PIX)	Lengh	Fraes Per Second (FPS)	Sample Time (T_S) Sec
1	The Ride	.MKV	74MB (720 × 1280)	00:01:00	30	0.0333
2	Wild Ocean Row	MP4	48.6MB (720 × 1280)	00:01:36	24	0.025
3	Team No Limit Motorcycle Stunt Show	MP4	91.4MB 720 × 1280)	00:13:32	30	0.0333

Table 2: Average Frame Rate of Video Sequences

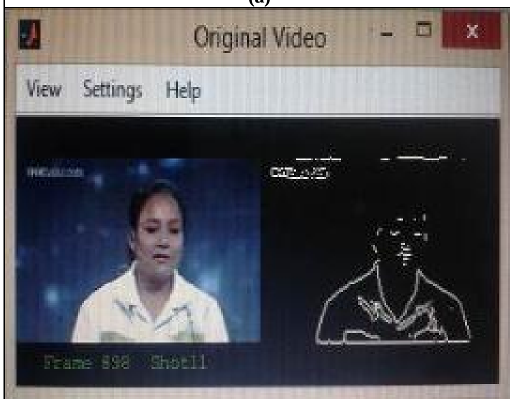
S N	Frame Rate Display/ Time Offset	The Ride	Wild Ocean Row	Team No Limit Motorcycle Stunt Show
1	0-2	2.2	3.0	2.9
2	2-4	2.2	2.5	2.9
3	4-6	2.1	1.9	2.8
4	6-8	2.1	1.8	2.1
5	8-10	2.4	1.8	2.1
Average Frame Rate Display		2.2	2.2	2.56

Table 3: Performance accuracy of the Scene Change Detection method

S. N.	Video Sequence	True	False	Missed	Recall (%)	Precision (%)	F-measure (%)
1	Shiv Khera Motivational Video	8	5	3	61.53	72.72	66.04
2	Arunima Sinha Motivational Video	15	0	4	100	78.94	88.20

Simulation of MATLAB results of two above video sequences in Table 3 are as shown in figure 8, figure 9 and figure 10.

(a)



(b)

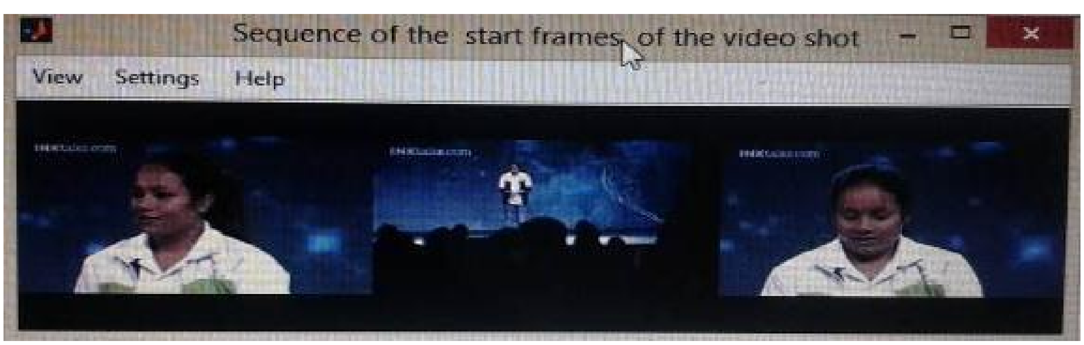
figure 8: Original video sequences and Edge detected images

(a)

(b)

Figure 9: Image matrix viewer

(a)



(b)

Figure 10: Sequences of video frames

CONCLUSION

The performance of SCD model in the video is giving the expected results better than other techniques in its category.

The threshold is dynamically allotted to each video sequences. Thus, proposed method is stable in scene change detection ability on various video sequences. This paper concentrate on developing a fast, as well as a well-performed video SCD method. The experimental results on several database showed that the proposed method can achieve a precision rate of greater than 70% and recall rate is achieve up 100% result, But it is not for true for all video sequences. Some basic input is needed to improve the performance of the model.

Some issues are considered into account for future research directions. Different video features can be used in order to improve the results.

REFERENCES

Ankita, P., Chauhan, Rohit R. Parmar, Shankar K .Parmar & Shahida G. Chauhan. (2013).Hybrid approach for video compression based on scene change detection. *Signal Processing, Computing and Control (ISPCC), IEEE*, 1-5.

Brandt, J., Trotzky, J. & Wolf, L. (2008). Fast frame-based scene change detection in the compressed domain for MPEG-4 video. *The Second International Conference on Next Generation Mobile Applications, Services, and Technologies, IEEE*, 514 -520.

Chung-Lin Huang & Bing-Yao Liao. (2001). A Robust scene-change detection method for video segmentation. *IEEE Transactions on Circuits and Systems for Video Technology*, 11:12.

Ding J. R. & Yang, J.F. (2008).Adaptive group-of-pictures and scene change detection methods based on existing H.264 advanced video coding information. *Image Processing, IET*, 2(2), 85 - 94

Ding J.-R. & J.-F. Yang. (2008). An effective error concealment framework for h.264 decoder based onvideo scene change detection. *IET Image Process,* 2(2):85–94.

Fernando W.A.C., Canagarajah, C.N., & Bull, D. R. (1999).Wipe scene change detection in video sequences. *Image Processing IEEE*, 3, 294-298.

Fernando, W. A. C., Canagarajah, C. N., & Bull, D. R. (2001). Scene change detection algorithms for content based video indexing and retrieval. *Electronics & Communication Engineering Journal*, 117-126.

Man-Hee Lee, Hun-Woo Yoo, & Dong-Sik Jang. (2006). Video scene change detection using neural network: improved art2. *Elsevier Ltd*., 31(1), 13–25.

Neetirajsinh, J., Chhasatia, C., Trivedi, U., & Komal A. Shah. (2013). Performance evaluation of localized person based scene detection and retrieval in video. *Image Information Processing (ICIIP), IEEE Second International Conference*, 78 – 80.

Priyadarshinee, Adhikari, Neeta Gargote, Jyothi Digge & Hogade, B.G. (2008).Abrupt scene change detection. *World Academy of Science, Engineering and Technology,* 2, 557-562.

Ralph M. Ford, Craig Robson, Daniel Temple, &Michael Gerlach (1997).Metrics for scene change detection in digital video sequences. *International Conference on Multimedia Computing and Systems IEEE*, 610 - 611

Shukla Dolley & Manisha Sharma. (2012). Overview of Scene Change Detection– Application Watermarking. *International Journal of Computer Applications,* 47, 975-888.

Taehwan Shin, Jae-Gon Kim, Hankyu Lee & Jinwoong Kim. (1998). Hierarchical scene change detection in an MPEG-2 compressed video sequence.4 *Circuts and Systems IEEE,* 4, 253 - 256.

Ufuk Sakarya & Ziya Telatar. (2010).Video scene detection using graph-based representations.*Signal Processing: Image Communication 25 Elsevier*, :774–783.

Wensheng Zhou1 & Asha Vellaikal. (2001).On-line scene change detection of multicast video. *Journal of Visual Communication and Image Representation,* 35(27), 1-15.

Wing-San Chau, Oscar C. Au, Tak-Song Chong, Tai-Wai Chan, &Chi-Shun Cheung. (2005). Efficient scene change detection in mpeg compressed video using composite block averaged luminance image sequence. *Fifth International Conference on Information, Communications and Signal Processing IEEE*, 688 - 691.

Xiaoquan Yi & Nam Ling. (2005).Fast pixel-based video scene change detection. *IEEE*, 4, 3443 - 3446.