Comparative Performance of Shot Change Detection Techniques in MPEG Stream

Ruaa Hussam Ali Al-Mallah
Iraq, Mosul, Technical College
ruaa.almallah@gmail.com

Abstract

This paper deals with the detection of shot changes in MPEG video stream. There are two main types of shot changes in a video. The first occurred abruptly, while the second type occurred gradually. This research is focused on the abrupt shot change detection. There are many techniques to detect a sudden shot change in video. In this research, the famous three techniques are used which are pixel difference, XOR operator, and Edge Change Ratio (ECR).

Two new algorithms proposed with these techniques to improve the behaviors and to find the best results of these techniques; the first algorithm depends on grayscale levels and the second algorithm depends on edge detection operation by Soble operator. Later on, these two algorithms are applied with the three techniques in two videos, which have abrupt shot change to measure maximum difference values between each two frames. The first algorithm with pixel difference technique gives an accurate result 100%, while the perfect results are given with second algorithm applied with ECR and XOR techniques.

Keywords: Abrupt shot detection, Edge Detection, Cut/Gradual Transition, Shot Boundary Detection, Edge Change ratio.

الخلاصة

يعنى هذا البحث بعملية تحديد مقاطع التغيير في المشهد الفديوي . هنالك طريقتان رئيسيتان للتغيير الفيديوي. الاولى تحدث فجائية بينما تحدث الثانية تدريجية. اعتمد هذا العمل على ثلاث تقنيات تستخدم مع كشف التغيير الفجائي. وهذه التقنيات هي تحديد تغيير الحوافECR , الاختلاف النقطي Pixel Difference و XOR . تم تصميم خوارزميتان جديدتان للعمل مع هذه التقنيات لتحسين من عمل هذه التقنيات وايجاد افضل النتائج، حيث اعتمدت الخوارزمية الاولى على قياس المستويات الرمادية للصورة والخوارزمية الثانية اعتمدت على تحديد الحواف الخارجية للصورة باستخدام طريقة الفلتر صوبل. ثم تم تطبيق هاتين الخوارزميتين مع الطرق الثلاث على فيديوين ذو مشاهد فجائية التغيير من اجل قياس فترة الاختلاف العظمى بين صورتين فجائيتين . واشارت النتائج على الحصول على دقة 100% عند استخدام الخوارزمية الاولى مع تقنية الاختلاف النقطي Pixel Difference. بينما تم الحصول على النتائج الدقيقة عند تطبيق الخوارزمية الثانية مع تقنية تحديد تغيير الحواف ECR و تقنية XOR.

الكلمات الدالة: التغيير المشهد الفجائي, تحديد الحواف , الانتقال التدريجي– الفجائي, تحديد حدود المشاهد, تغيير نسبة الحواف.

1.Introduction

The detection of shot boundaries in digital videos and compressed video domain are presents an essential step in video-analysis system. Any video stream consists of a number of shots, each shot is sequence of frames pictured using a single camera. Switching from one camera to another indicates a transition from one shot to another. Shot Boundary Detection has been an area of active research. Many automatic techniques have been developed to detect frame transition in video sequences. Shot boundaries can be classified into two main categories: abrupt and gradual [Waleed & Hussein, 2005]. A cut is an abrupt shot change that occurs over a single frame while a gradual is a slow change that occurs in a number of consecutive frames as fade, dissolve and wipe. A fade out is a slow decrease in brightness resulting in a black frame; a fade in is a gradual increase in intensity starting from a black image. Dissolves show one image superimposed on the other as the frames of the first shot get dimmer and those of the second one get brighter.

A video scene change detection method is necessary and had been used in many advanced video applications such as video on demand (VOD) surveillance, remote sensing, video Indexing, civil infrastructure, and underwater sensing, [Dolley & Manisha, 2012], was proposed The TAMI algorithm by adjusting the group of picture (GOP) structure. It loads frames from one GOP into frame memory. This algorithm detects the position of abrupt scene changes and scene segmentation points and determines the GOP structure.

The major techniques that have been used for shot boundary detection are pixel differences, statistical differences, histogram comparisons, edge differences, compression differences, and motion vectors. The only reported comparisons of this technique applied the tested methods to a small number of short test sequences [Purnima & Srinivasan, 2011].

A new algorithm for shot boundary detection was presented in the work of [Yousri *et al.*, 2000]. It combines both photometric and motion information. The evaluation of the feature points based a histogram method, especially in a scene characterized by large motion originating from objects as well as from the camera. This algorithm depends on the stability of the feature point's extraction and tracking steps under different transformations.

Colin O'Toole *et al* [Colin *et al.*, 1999], is presented a detailed evaluation of a histogram-based shot cut detector. In their work shot boundaries for in such broadcast video are determined by selecting the similarity of thresholds.

Shu- Ching and *et al.*, [Shu- Ching *et al.*, 2011], proposed an innovative scene change detection method using the unsupervised segmentation algorithm and object tracking technique, and showed the performance using the sample MPEG-1 news video data. The key idea of the matching process in scene change detection is to compare the segmentation mask maps between two successive video frames. In addition, the object tracking technique is employed as a complement to handle the situations of scene rotation without any extra overhead. Unlike many methods using the low-level features of the video frames, the proposed method is not sensitive to the small change in luminance or color and it has high precision. [Truong *et al.*, 2000], proposed techniques for cut, fade, and dissolve detections. An adaptive threshold technique to detect peaks in the color histogram difference curve is presented for detecting hard cuts. Huang *et al.*, [Huang, *et al.*, 2008], proposed an approach based on local key point matching of video frames to detect abrupt and gradual transitions. By matching the same objects and scenes using contrast context histogram (CCH) in two adjacent frames, the method decided that there is no shot change.

Zabik proposed Edge Change Ratio (ECR) technique to deal with MPEG specific effects in pixel and based difference measurements. Edge Change Ratio technique relies on the fact that the edges of the objects within the frames would definitely change across a boundary [Zabik &Miler, 1995].

The remainder of the paper is organized as follows. Section 2 explains the theory of video shot boundaries detection and describes the methodology (Edge Change Ratio (ECR), Pixel difference and XOR operator) techniques. Section 3 describes the experimental results and section 4 comments some conclusions about the results.

1. Theory

Shot Boundary Detection has been an area of active research. Many automatic techniques have been developed to detect frame transition in video sequences. Shot boundaries can be classified into two main categories: abrupt and gradual. A cut is an abrupt shot change that occurs over a single frame while a gradual is a slow change that occurs in a number of consecutive frames as fade, dissolve and wipe. A fade out is a slow decrease in brightness resulting in a black frame; a fade in is a gradual increase in intensity starting from a black image. Dissolves show one image superimposed on the other as the frames of the first shot get dimmer and those of the second one get brighter. As illustrated in Fig (1).

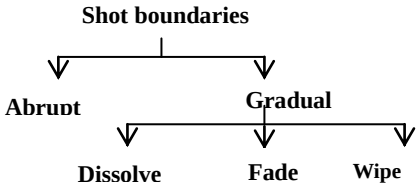


Fig (1) Types of shot boundaries.

Edge frame in the digital video is turned from RGB image to gray scale image and Prewitt filtering applied to detect edges.

The method looks for similar edges in adjacent frames to detect a shot boundary [Guozhu and Junming, 2009]. The simplest technique is by comparing the pixel values of two corresponding frames by using XOR operation and when used absolute of subtraction between two frames. The way alternate to pixel matching is using a gray-scale or color histogram. Another way of detecting shot boundary is by edge change ratio (ECR). This paper presents simplest and fastest methods for detecting abrupt scene detection change in video. The techniques which used in this research are:

1.1 Edge Change Ratio (ECR) Technique:

Edge Change Ratio technique relies on the fact that the edges of the objects within the frames would definitely change across a boundary. The percentage of edges that enter and exit between the two frames is computed. The basic idea of edge change fraction (ECF) comparison method is summarized as following:

1. Detect edges in two consecutive frames f_n and f_{n-1} respectively.
2. Count the number of edge pixels σ_n and σ_{n-1} in frame fn and fn-1.
3. Define the entering and exiting edge pixel X_n^{in} and X_{n-1}^{out}.

Suppose we have two images f_n and f_{n-1} The entering edge pixels X_n^{in} are the fraction of edge pixels in f_{n-1} which are farther than a fixed distance r away from the closest edge

pixel in. Similarly exiting edge pixels X_{n-1}^{out} are the fraction of edge pixels in $\boldsymbol{f_n}$ which are farther than a fixed distance r away from the closest edge pixel in $\boldsymbol{f_{n-1}}$. The Edge Change Ratio $ECR(n,1)$, between the frames f_n and f_{n-1} as define in [Zabik &Miler, 1995] is shown below:

$$ECR(n,1) = max(\frac{X_n^{in}}{\sigma_n}, \frac{X_{n-1}^{out}}{\sigma_{n-1}}) \quad \text{..} \quad (1)$$

Where, σ_n is the number of edge pixels in the frame n, and X_n^{in} and X_{n-1}^{out} are the entering and exiting edge pixels in frames f_n and f_{n-1} respectively. Sobel Edge Detector operator is applied to the incoming frames to implement the above mentioned formula. Obtained ECR value varies between 0 and 1. Maximum ECR value above some threshold θ represents shot change between the frames.

2.2 Pixel Difference Technique:

The idea behind the algorithm is that the intensities of the pixels changes rapidly at the shots. The pixel subtraction operator takes two consecutive images as input and produces the value which is the sum of the difference between corresponding pixel values of first image with that of the second image. As we are interested only in the change in intensity value from one frame to another the difference is taken to be an absolute value. Mathematically it is represented as

$$D(i,i-1)=\sum | f_i - f_{i-1} | \quad \text{..} (2)$$

If D(i,i-1)>T, where T is some threshold then shot change is assumed .
The difference between the two frames is relative large and it is possible to be abrupt shot boundary (ASB). Thus hard cuts and shot transitions can be detected as single peaks in the time series of the sum of pixel differences of contiguous frames from a given source [Purnima & Srinivasan, 2011].

2.3 XOR Operator Technique:

The idea behind the algorithm is that the intensities of the pixels changes rapidly at the shots. XOR operation is applied to compare the differences between these two frames (f_i,f_{i-1}). Between two pixels p1 and p2, at the same (x, y) coordinates of the two binary images f_i and f_{i-1} the existence of a change can only mean that either "p1 is white and p2 is not" or "p1 is not white and p2 is." This directly implies the XOR operation in binary logic. Hence the obvious solution to the change detection problem is XOR-ing the two binary images pixel wise as:

$$C(x,y) = f_i(x,y) \oplus f_{i-1}(x,y) \quad \text{..} \quad (3)$$

This operation gives '0' for pixels having the same value in both images, and gives '1' for pixels having different values. Therefore, white pixels in the resulting binary image *C(x, y)* represent the changed regions [Purnima.S. 2011].

Varying threshold value in binary image for detecting maximum change with two frames comparison .It calculated by the following equation:

$$C(x,y)=max(xor(f_i(x,y) - f_{i-1}(x,y))) \quad \text{...} (4)$$

2. Practical Implementation:

One possible approach for detection and classification of scene breaks in video sequences is by detecting the appearance of intensity edges in a frame that are a fixed distance away from the intensity edges in the previous frame. This paper examined simplest and faster methods for detecting abrupt shot change detection in MPEG stream. The used

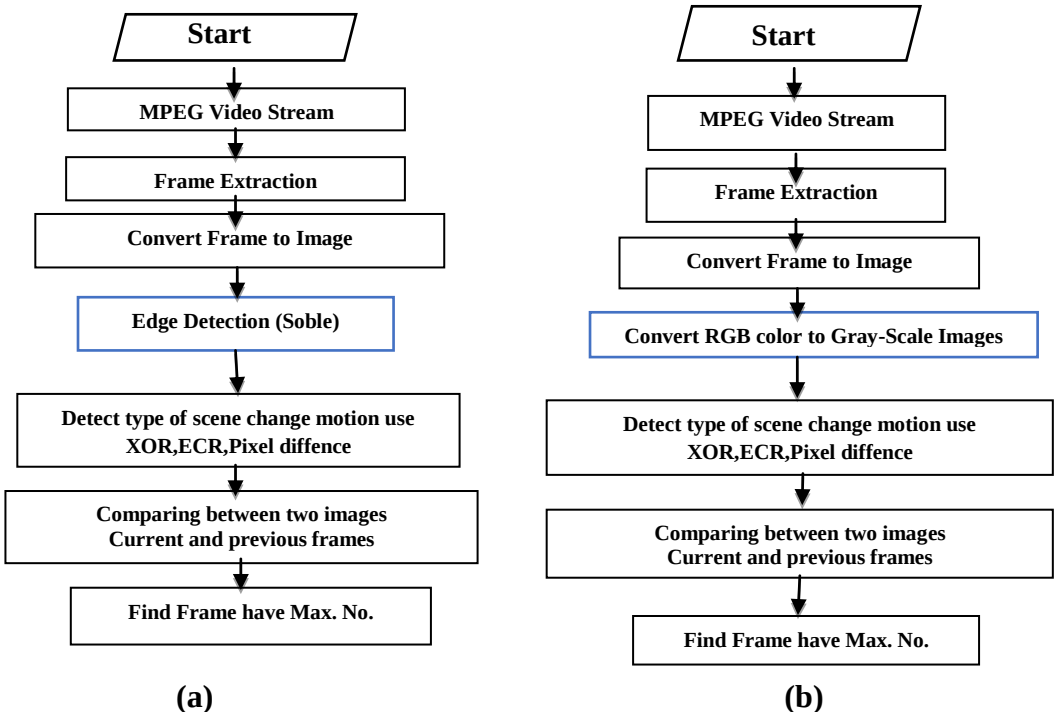
methods are Edge Change Ratio (ECR) technique, Pixel difference technique and XOR operation technique and by varying threshold value in edge detection image for detecting maximum change with two frames comparison and has better performance than the traditional approaches.The steps of this reasearch were broken down into the following steps and illustrated in fig.(2a,2b).

Step 1. Frame Extraction: The imaging process takes in an MPEG-1 video (RGB 24) bit stream and converted to multiple frames. Selected two frames for processing (current frame f_i and previous frame f_{i-1}) depending in algorithm.

Step 2. Gray Scale Image: Convert RGB image values to Gray image values by forming a weighted sum of the R,G and B components as the following equation (5):

Gray values =0.299 R+0.587 G + 0.114 B ..... (5)

Step 3. Binary Image: Convert gray image to binary image as white pixel '1' value and black pixel '0' value.



(a) **(b)**

Fig.2 (a) Abrupt Shot change detection algorithm with Edge Detection image
(b) Abrupt Shot change detection algorithm with Gray Scale image

Step 4. Soble Filter: Convert binary image to edge detection to increase speed operation. The edges are calculated by the Sobel edge detector. Technically, it is a discrete differentiation operator computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the Sobel operator is either the corresponding gradient vector or the norm of this vector. The components of the gradient may be found using the following approximation [Vincent, 2009]:

$$\frac{\partial f(x,y)}{\partial x} = \Delta x = \frac{f(x+dx,y)-f(x,y)}{dx} \quad \ldots (6) \qquad \frac{\partial f(x,y)}{\partial y} = \Delta y = \frac{f(x,y+dy)-f(x,y)}{dy} \quad \ldots (7)$$

Where ∂x and ∂y measure distance along the *x* and *y* directions respectively. Intensity edges are formed from high luminance pixels that reside around the edges of main objects in an image.

Step 5. Difference between two frames: The difference between the binary values for each frame is calculated by XOR operator, ECR technique and Pixel difference techniques with current frame f_i and previous frame f_{i-1}. After varying threshold value to get the best result, the result is a binary image. This step has been used with both in gray scale image by depending on flow chart that illustrated in fig. (2a) and in edge detection operator as illustrated in fig. (2b).The value of the numbers of white pixels are calculated in all frames of video, the high peak shows where the abrupt scene changes occurred.

3. Expermental Results And Discussion

Two samples of videos clip were used for testing accuracy of the methods (ECR, XOR and pixel difference) to detect abrupt shot change. The parameters of these two samples of videos are illustrated in Table (1).

Table (1) parameters of two videos

Parameters	**Video1**	**Video2**
Type	Video Clip (avi file)	Video Clip (avi file)
Bit stream	RGB 24 bit	RGB 24 bit
Frame height	288	120
Frame width	352	160
Frame rate	25 frames/second	15 frames/second
Data rate	608 kbps	520kbps
Total bit rate	736 kbps	736kbps
No. of frames	24 frames	92 frames
No. of Abrupt shot	1	2

4.1 Gray Image Algorithm:

The first algorithm is used a sample video1, the gradual change of scene is appeared from Frame #1 to Frame #22 and then abrupt change of shot is occurred in Frame #23. In The second sample video 2, the gradual change of scene is appeared from Frame #1 to Frame #39 and then abrupt shot change is occurred in Frame #40 then gradual scene change appeared until frame #71 then another cut is appeared in frame # 72 after that appeared gradual scene change until frame No.# 91. In this stage all frames were converted from RGB color to Gray scale images. Then the techniques were applied for detecting abrupt shot in video stream:

3.1.1 XOR operator technique: The first attempt was used (XOR operator) with test first sample of MPEG stream. The result was appeared as illustrated in Fig.(3) and the achieved real time to detect a sudden scene change was 1 second. It was find that 5 frames in the sequences (#1,#2,#12,#19,#22) have large differences values between two frames (current frame and previous frame) by using this method . All these 5 frames are false frames except frame #22 is true frame (abrupt shot) so the accuracy was (82.6 %) depending to the equation below [Guozhu Liu, 2009]:

$$\% \text{ Accuracy} = \frac{\#\,Total\ Frames\ -\ (\#\,Detect\ Frames - \#\ true\ Frames)}{\#\,Total\ Frames} * 100\% \quad \text{..... (8)}$$

When second video sample was used XOR operator, the result was illustrated in fig. 4. It was find that 24 frames have large differences values. All these 24 frames are false frames except two frame in sequence #40 and #72 is true frame (abrupt shot) so the accuracy was (76.6 %) and the execution time was 7.2 second. The (Tic-Tac) command function in MATLAB program have been used to determine the real time value.

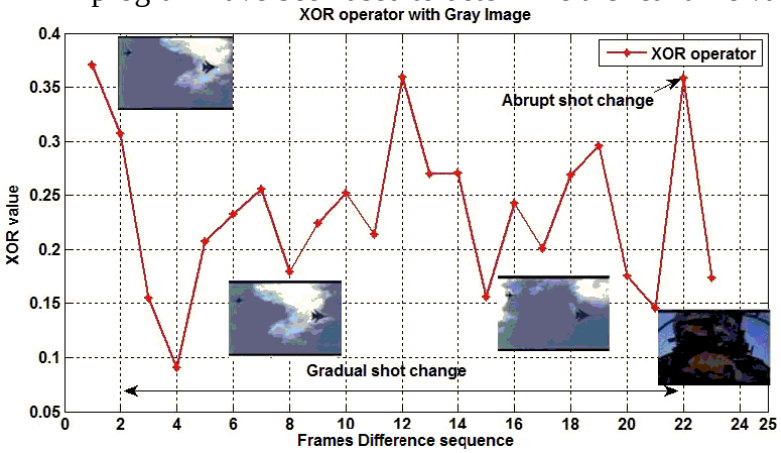


Fig 3.XOR with Gray in video1

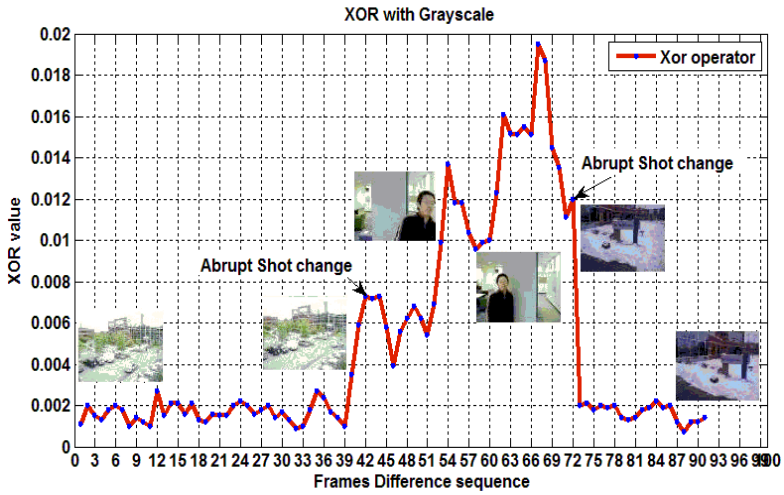


Fig.4 XOR with gray in video 2

4.1.2 ECR Technique:

ECR technique was applied with first test template and the time execution was 1.36 second. The result was appeared as illustrated in Fig.(5), It was find that nine frames in the sequences (#1,#3,#4,#5,#6,#7,#8,#18,#22) have large differences values. All these nine frames are false frames except frame #22 is true frame (abrupt shot) so the accuracy was (65.2 %). When second video sample was with ECR technique, the result was illustrated in fig.(6). It was find that 5 frames have large differences values. All these 5 frames are false frames except two frames in sequence #40 and #72 is true frame (abrupt shot) so the accuracy was (96.6 %) and the execution time was 7.4 second.

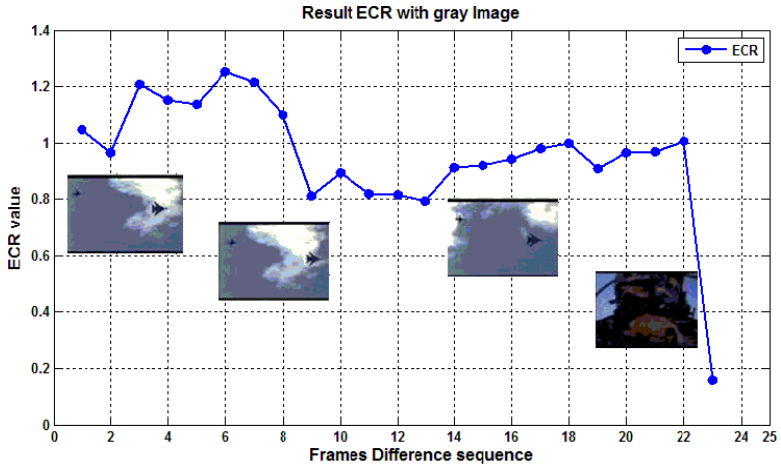


Fig 5. ECR with Gray in video1

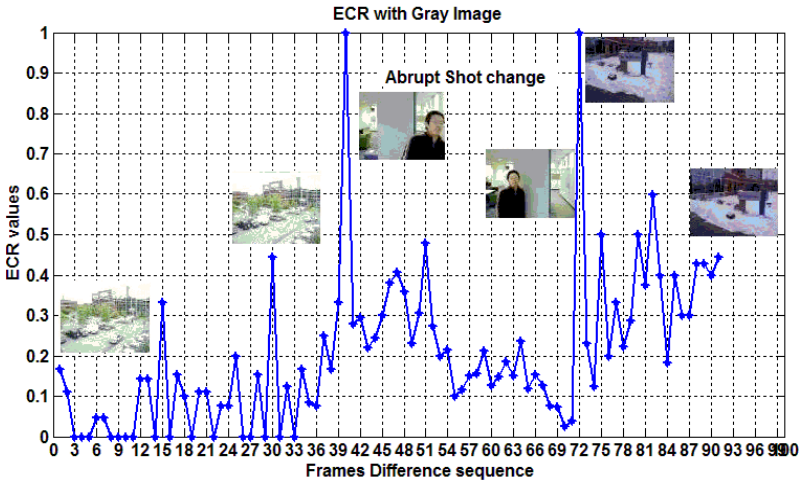


Fig. 6 ECR with Gray in video 2

4.1.3 Pixel differences

Pixel difference technique was applied with first test template and the time execution was 1.06 second. The result was appeared as illustrated in Fig.(7), It was find that one frame in the sequences (#22) have large differences values that a true frame (abrupt shot), so the accuracy was (100%).Time execution was (1.06) sec respectively. When second video sample was used with Pixel difference technique, the result was illustrated in fig.(8). It was found that two frames have large difference values and those two frames belong to two different shot in sequence #40 and #72,so the accuracy was (100 %) and the execution time was 6.5 second.

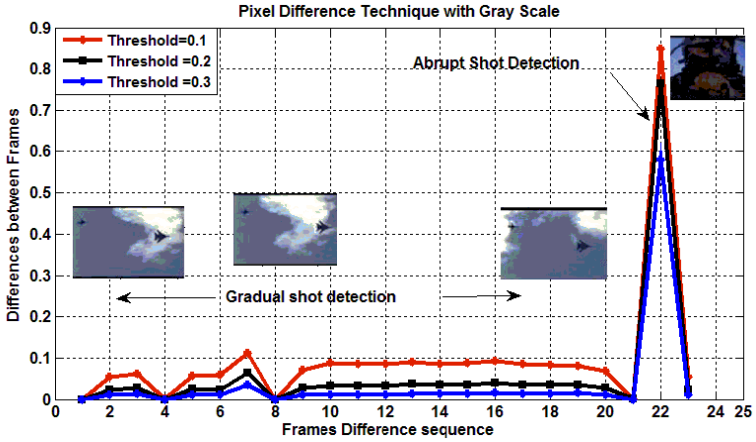


Fig. 7 Pixel operator with Gray in video1.

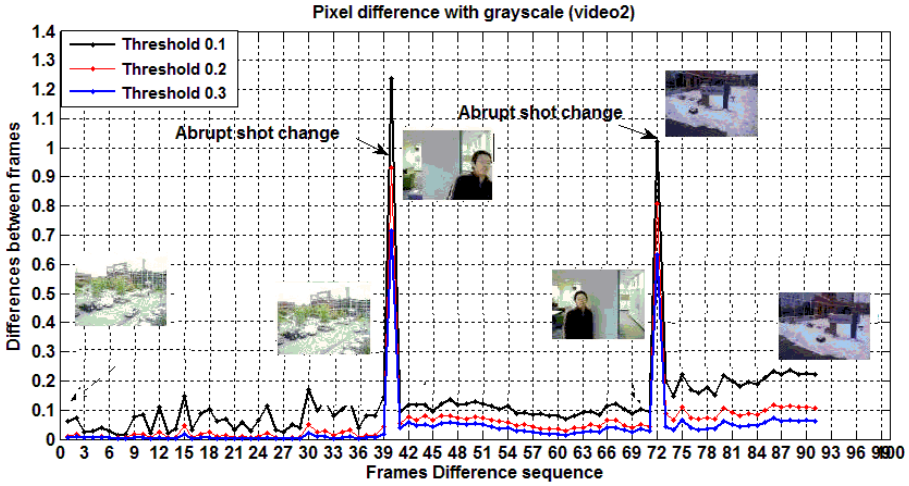


Fig. 8 Pixel difference with Gray in video 2

4.2 Edge Detection:

The second algorithm to detect abrupt shot change in video stream was used depending on edge detection by Soble filter. Then the techniques of (ECR, XOR and Pixel Difference) were applied. The results are compared with that of the first algorithm that appeared from gray image.

In Pixel difference method, the number of error frames was 12 frames with changing of threshold values. So that the percentage of accuracy was 80% in video 1 and 86.6 % in video 2 as illustrated in fig. (12, 13) respectively.

4.2.1 XOR operator technique:

In the first sample video, The result was appeared as illustrated in Fig.(9) and the execution time for detecting sudden scene change was 1.09 second. It was found that two frames have large differences values. One of them was shot change detected. So the accuracy was (95.6 %). In the second video sample, the time execution was 6.6 sec when detect abrupt shot change as illustrated in Fig (10). It was found that 16 have large differences values. Only two abrupt shot change frames occurred in frame sequence in #40 and # 72, so the accuracy was (84.7 %).

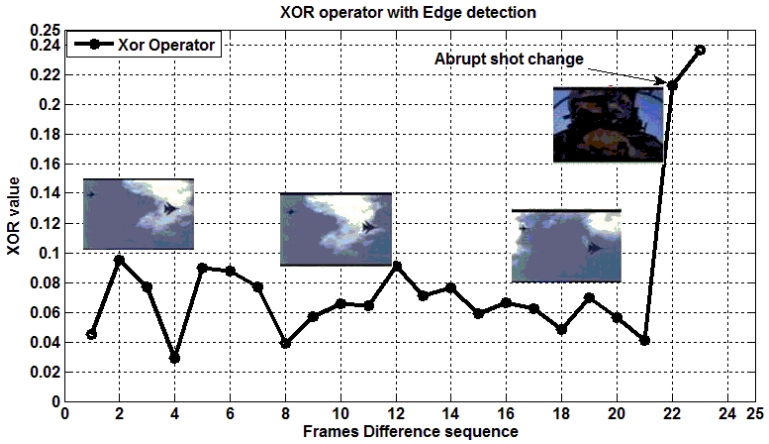


Fig. 9 XOR technique with edge detection image in video1.

4.2.2 ECR Technique:

ECR technique was applied with first test template and the time execution was 1.4 second. The result was appeared as illustrated in Fig. (11), the accuracy was (95.2 %). When second video sample was applied with ECR technique, the result was illustrated in Fig. (12).

It was found that two frames(#40 and #72) have large differences values, so the accuracy was (100 %) and the execution time was 6.64 second.

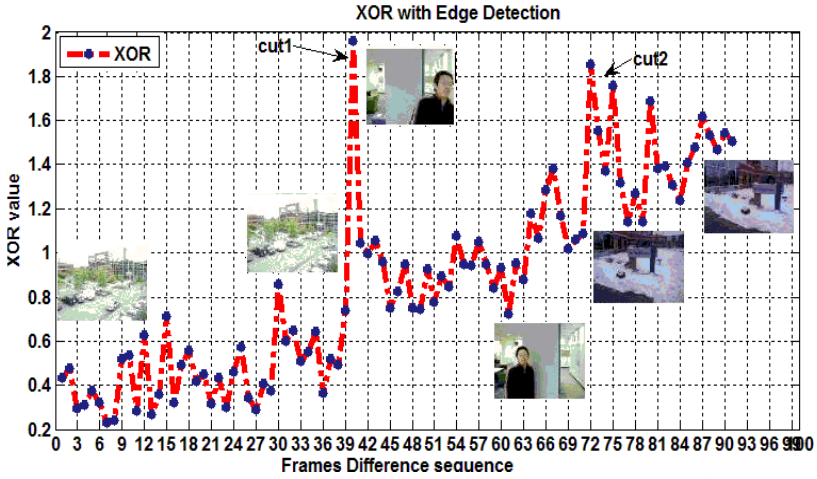


Fig 10 XOR with Edge detection in video 2.

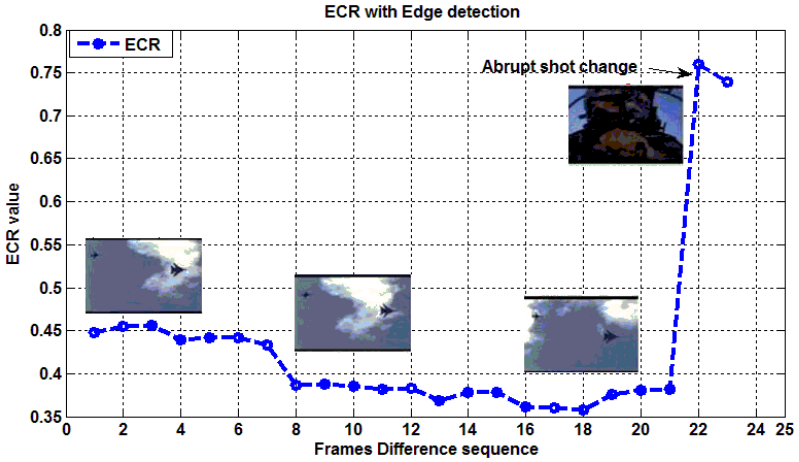


Fig. 11 ECR with Edge detection in video1

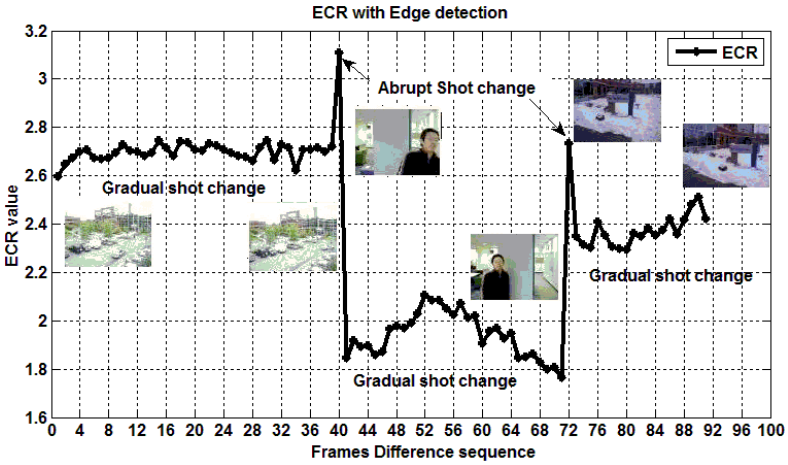


Fig. 12 ECR with Edge detection in video2

4.2.3 Pixel differences:

Pixel difference technique was applied with first test template. The result was appeared as illustrated in Fig.(13) and the execution time for detecting sudden scene change was 1.09 second. It was found that two frames have large differences values. One of them was shot change detected. So the accuracy was (95.6 %). In the second video sample, the time execution was 6.6 sec when detect abrupt shot change as illustrated in Fig (14). It was found that 16 sequence frames have large differences values. Only two abrupt shot change frames occurred in frame sequence in #40 and # 72, so the accuracy was (84.7 %).

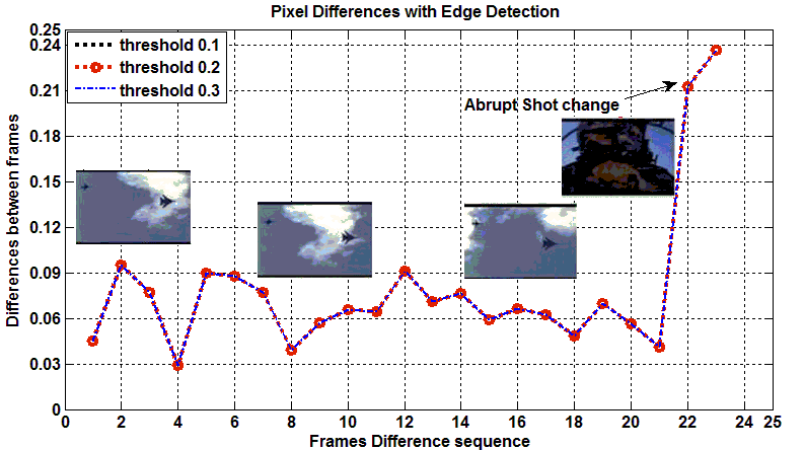


Fig.13 Pixel Difference with edge detection in video1

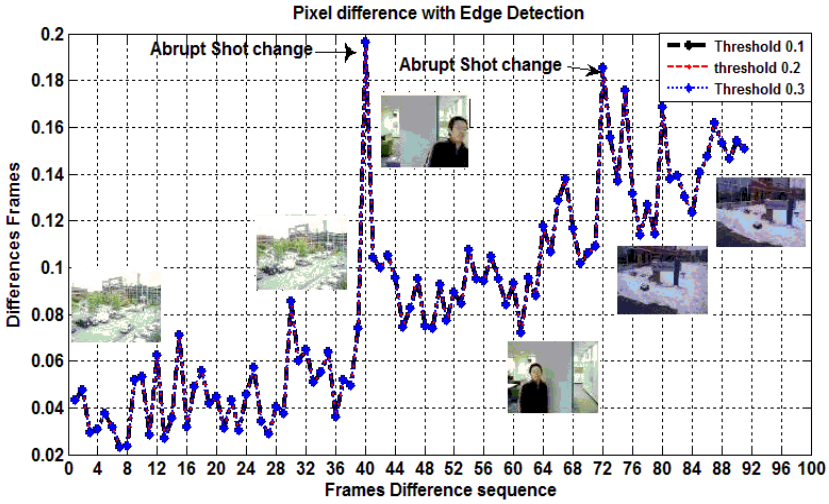


Fig.14 Pixel Difference with edge detection in video2

4.Conclusions

There are many techniques to detect a sudden shot change in video. In this research, the famous three techniques are used which are: pixel difference, XOR operator and Edge Change Ratio (ECR). Two new algorithms (Gray scale level and Edge detection) have been proposed to improve the behaviors and to find the best results of these techniques. We concluded that ECR and XOR operator techniques give more accurate results with edge detection algorithm comparing with gray scale algorithm at minimum achieved real time as can be noticed in table 2 and 3. And pixel difference technique give accurate results as well with gray scale algorithm as illustrated in table 2 and 3 for first and second video respectively.

Table (2) Accuracy and real time of three techniques with gray scale and edge detection algorithms in video 1

Video1	***Accuracy***		***Real Time***	
Methods	*Gray Scale*	*Edge detection*	*Gray Scale*	*Edge detection*
XOR	82.6%	**95.6%**	1.06 sec	1.09 sec
ECR	65.2%	**95.2%**	1.4 sec	**1.5 sec**
Pixel Difference	**100 %**	95.6%	**1.06 sec**	1.08 sec

Table (3) Accuracy and real time of three techniques with gray scale and edge detection algorithms in video 2

Video2	***Accuracy***		***Real Time***	
Methods	*Gray Scale*	*Edge detection*	*Gray Scale*	*Edge detection*
XOR	76.6%	84.7 %	7.2 sec	6.64 sec
ECR	96.6 %	**100 %**	7.6 sec	**7.4 sec**
Pixel Difference	**100%**	84.7 %	**6.5 sec**	6.8 sec

6. REFERENCES:

Colin O'Toole, Alan Smeaton, Noel Murphy and Sean Marlow, 1999,"Evaluation of Automatic Shot Boundary Detection on a Large Video Test Suite", Challenge of Image Retrieval, Newcastle,

Danisman, T. Alpkocak, A. 2006, "Video Shot Boundary Detection",Proceedings of the TREC Video Retrieval Evaluation (TRECVID), Dokuz Eylul University.

Dolley Shukla and Manisha Sharma, 2012,'Overview of Scene Change Detection Application to Watermarking', International Journal of Computer Applications, Volume 47– No.19.

Dong Zhang, Wei Qi and Hong Jiang Zhang,2001,"A New Shot Boundary Detection Algorithm",Advance in Multimedia Information Processing, PCM, lecture Notes in Computer Science Vol.(2195),pp.63-70

Guozhu Liu, and Junming Zhao,2009,'Key Frame Extraction from MPEG Video Stream', Huangshan, P. R. China, 26-28,Dec., pp. 007-011.

Huang, C. ; Lee, H.; Chen, C. 2008," Shot change detection via local keypoint matching", IEEE Transactions on Multimedia 10 (6) pp.1097–1108.

Jung woo Lee, 1994, 'Temporally adaptive motion interpolation exploiting temporal masking in visual perception' , Image Processing, IEEE Transactions on, Vol.3 ,Issue 5, pp. 513-526.

Patrick Seeling, 2010,"Scene Change Detection for Uncompressed Video", Department of Computing and New Media Technologies, Technological developments in Education and Automation, pp: 11-14, publisher :Springer Netherlands,ISBN:978-90-481-3656-8.

Purnima.S.Mittalkod, G.N.Srinivasan, 2011,"Shot Boundary Detection Algorithms and Techniques', Research Journal of Computer Systems Engineering- An International Journal Vol 02, Issue 02.

Shu- Ching Chen, Mei-Ling Shyu and etal, 2011,"video scene change detection method using unsupervised segmentation and object tracking". Multimedia and Expo,ICME, IEEE international conference, Computer Society, ISBN : 0-7695-1198-8/01.

Truong, B. ; Dorai C. ; Venkatesh, S. 2000," New enhancements to cut, fade, and dissolve detection processes in video segmentation", Proceedings of the Eighth ACM International Conference on Multimedia, pp. 219–227.

Vincent O. R. and Folorunso, O. 2009,'A Descriptive Algorithm for Sobel Image Edge Detection', Proceedings of Informing Science & IT Education Conference (InSITE).

Waleed E. Farag, and Hussein Abdel-Wahab, 2005,"Video Data Management and Information" ,IRM press, pp.:(194-195), USA,copyright.

Young-Suk Yoon, Won-Young Yoo and Young-Ho Suh, 2012,"A Robust Scene Change Detection Using Mode Distribution in H.264/AVC", The Sixth International Conference on Advances in Semantic Processing, SEMAPRO.

Yousri Abdeljaoued, Touradj Ebrahimi and *et al.,* 2000 ,' A New Algorithm for Shot Boundary Detection", In proceedings of the 10th European Signal Processing Conference pp. :151-154.

Zabik R.,J. Miler, 1995,"A feature based algorithm for detecting classifying scene breaks", proceeding of ACM Multimedia, San Francisco,CA.,pp:189-200