DCT Based Scene Break Detection for Color Video using Variance of Histogram Difference of DC Image

Rakesh Trivedi¹, Mehul Shah² and Laukik Patel³

¹² Electronics and Communication Department, Gujarat Technological University/G. H. Patel College of Engineering & Technology Vallabh Vidhyanagar-388120, Gujarat, India
³ Electronics and Communication Department, Gujarat Technological University/Government Polytechnic, Gandhinagar-382026, Gujarat, India

Abstract
In current digital environment video is important parameter for interactive multimedia. In this article we propose new method for detection of scene breaks in color video using variance of histogram difference of consecutive DC images. DC image concept is used for analysis to reduce processing time. DCT is used for compression of color video. Simulation results shows how quickly and efficiently this algorithm detects cut regions in video frame sequences.

Keywords: Shot Boundary Detection, DCT, DC image, Variance, Macro block, histogram difference, cut.

1. Introduction
In digital environment video is used for communication and broadcasting. In this field there is need for extracting information from objects or images which needs efficient method to handle video. Along with method we need to represent more information with only fewer co-efficient to reduce traffic on data network. All this requires video to be segment in its lower levels called shot boundary detection in the field of shot boundary detection various methods are presented. Simple method is pair wise pixel comparison [1]. Total number of pixels that differ in value more than some threshold value is compared with some other threshold to determine if a shot boundary has been found. Its major disadvantage is that it is sensitive to camera and object motion. In [2] authors have presented pair wise pixel method that uses 3*3 averaging filter to reduce problem of camera motion and noise. Another method includes comparing intensities of consecutive frames [3] but, it is time consuming. In [4] author used block based method in which frame is divided in block and Cut is detected based on some threshold value. After block based method to further reduce problem of motion histogram technique is used [5]. In [6] author has presented block matching algorithm to estimate the motions between an image pair. The extracted motion vectors were used to reconstruct the next frame in the sequence and if the displaced frame difference exceeded some threshold a shot boundary was detected. However, [6] showed that the process of computing the pixel difference can still lead to false detections in the presence of sudden intensity changes or miss detections if two shots have similar intensities.

In our approach input video file is RGB file which needs to be converted into Gray file for analysis of shot boundary detection. In this method work is done in compressed domain means video is divided into smaller parts and each part is compressed by DCT which results into DCT co-efficient.

2. Shot Cut Detection
Shot cut detection means finding boundary of shot means finding starting and ending of shot. Here shot is defined as sequence of unbroken frames taken by single camera continuous in time and space. Transition between frames is of two types. First is gradual transition in which transition takes place over multiple frames. Typical examples are fade, dissolve, wipe etc. In abrupt transition completes within two consecutive frames only. Cut boundary shows change in pixel intensities or color. Example is cut effect in which one shot is completely replaced by another shot. Cut effect is best described by following Fig.1 here 2nd and 3rd frame shows cut effect. After 2nd frame there is sudden change from one picture to another picture. Here only two frames 2nd and 3rd frame are involved in transition which produce cut effect.
In compressed domain one of the simplest ways to find scene breaks is by using DC image concept. In this paper we have propose macro block based method [7]. Here entire frame is divided into number of block of size 16*16. this concept is used as macro block concept [7]. after dividing into block variance of each block is counted and all the values are combine for desired work. Macro block is an image compression component and techniques based on discrete cosine transform used on still images and video frames. Macro blocks are usually composed of two or more blocks of pixels. In the JPEG standard macro blocks are called MCU blocks. The size of a block depends on the codec and is usually a multiple of 4. In MPEG2 and other early codec the size is fixed at blocks of 8x8 pixels. In more modern codec such as h.263 and h.264 the overarching macro block size is fixed at 16x16 pixels, but this is broken down into smaller blocks or partitions which are either 4, 8, 12 or 16 pixels by 4, 8, 12 or 16 pixels. Color information of an 8x8 macro block in a 4:1:1 color space will be encoded into a Y Cb Cr format. The Luminance will be encoded at an 8x8 pixel size and the difference-red and difference-blue information each at a size of 2x2. In the decode process these will be stretched out to cover the 8x8 space. In this paper we have used DC image concept. If operation is done directly on real data then high amount of data must be tested and processing time required is more. For removing more time consumption problem DC image concept is used. In this original image is represented as spatially reduced form. And formulas are applied on data of DC image.

3. Algorithm for Cut Detection

Cut detection between frames is determined by DCT of frame first in which entire frame is divided into block of 16*16. After that DCT of each block is calculated. Formula for finding DCT of 16*16 blocks is given by Eq. (1)

\[ F(u,v) = \frac{C_u}{2} \sum_{x=0}^{15} \sum_{y=0}^{15} f(x,y) \cos \left( \frac{2x+1}{16} u \pi \right) \cos \left( \frac{2y+1}{16} v \pi \right) \]

(1)

DCT block contains one DC component and all other AC components. DC component carry most of the signal energy while AC components carry least energy. DC image of current frame is calculated and compared with DC image of next frame or frames. By using following Eq. (2) DC image of frame can be calculated.

\[ DC(i,j) = \frac{1}{256} \sum_{x=0}^{15} \sum_{y=0}^{15} DCT(x,y) \]

(2)

DCT(x, y) shows the 256 DCT values of one 16*16 block. And DC (i, j) shows the each component of DC image. After completing the calculation for all blocks one DC image corresponds to one image(frame). similarly find DC image and its components for second frame. Now we have two DC image for two consecutive frames. Now find histogram for both DC images. Calculation of histogram of DC image is taken which is calculated by Eq. (3)

\[ H(r_k) = n_k \]

(3)

In the above equation \( r_k \) is the Gray level, but for DC image concept it is mean of co-efficient of each block in image. And \( n_k \) is number of times the particular value \( r_k \) is occurring and \( H(r_k) \) is called histogram of DC image. After calculation of histogram, absolute value of difference of two successive DC image is taken which is calculated by Eq. (4)

\[ H_{diff}(k) = abs(h_1 - h_2) \]

(4)

Here \( h_1 \) and \( h_2 \) are DC image for frame 1 and frame 2. And abs function finds the absolute of each difference value. After finding difference in \( H_{diff}(k) \) variance of \( H_{diff}(k) \) is taken because \( H_{diff}(k) \) is itself having same size as DC images whose histogram difference is taken. So there will be one variance value for each successive frame difference of DC image. This calculation is of variance of DC image, and here in this algorithm variation in variance of frames is tracked.

\[ \text{Variance (D.C image)} = \frac{m^2 - \overline{m^2}}{m^2} \]

(5)

In the above equation \( m^2 \) represents square of mean of D.C image and \( \overline{m^2} \) represents the mean of squared D.C

**Fig. 1 Video cut effect.**
image means first D.C image is multiplied with itself and then mean of that squared D.C image is taken.

\[ m^2 = m \times m \]  \hspace{1cm} (6)

\[ \overline{m^2} = \frac{m^2}{m_1} \]  \hspace{1cm} (7)

m1=total pixels in DC image. Finally graph is plotted for number of frames versus Variance of histogram difference of consecutive DC images. High spikes and peaks represent cut effect. Overall algorithm can be summarized in following steps.

Step: 1 Read input video frame by frame.
Step: 2 Divide frame into blocks of size 16*16.
Step: 3 Apply DCT to all block.
Step: 4 Find variance of each block. Variances of each block are elements of DC image. We have one DC image with 16 variances as elements for one frame.
Step: 5 Apply same procedure to obtain DC image for next frame.
Step: 6 we have two DC images of two consecutive frames. Find histogram of both DC images.
Step: 7 Find absolute value of histogram difference of two DC images.
Step 8: Find variance of absolute value of histogram difference of two DC images. We have one variance for histogram difference of two consecutive DC images.
Step 9: Plot graph of Number of frames versus variance of absolute histogram difference of DC images.

4. DC Image

![Fig. 2 Generation of DC image for 16*16 Frame.](image)

Above Fig.2 shows how DC image is generated for frame. First frame of size say 256*256 is divided into block of16*16.DCT is applied on each of the block. Calculate variance of each block. Each block’s variance is elements of DC image. We have total 16 variances as elements of DC image.

5. Simulation and Results

The algorithm for detecting cut effect are applied on various movies in MATLAB 7.6.0 and obtained various useful results which shows the detection of cut effect. These results are as follows.

![Fig. 3 cut effect detection for fight movie for frame range 4000 to 5000 using variance of histogram difference of DC images](image)

Fig.3 shows cut detection for fight movie for frame range from 4000 to 5000. There are 6 high spikes and peaks represent abrupt change in histogram of both DC images. There are total 6 cut effect is detected. Location pairs are frame number 4110-4111, 4520-4521, 4524-4525, 4697-4698, 4869-4870, and 4911-4912.

Fig.4 shows cut detection for animation movie for frame range from 1 to 700. There are 3 high spikes and peaks represent abrupt change in histogram of both DC images. There are total 3 cut effect is detected. Location pairs are frame number 2-3, 7-8, 152-153.

Fig.5 shows cut detection for comedy movie for frame range from 2000 to 4000. There are 5 high spikes and peaks represent abrupt change in histogram of both DC images. There are total 5 cut effect is detected. Location pairs are frame number 3099-3100, 3436-3437, 3623-3624, 3724-3725, and 3762-3763. There are total 5 instants at which one picture is changing to another picture.
Fig. 4 cut effect detection for animation movie for frame range 1 to 700 using variance of histogram difference of DC images.

Fig. 5 cut effect detection for comedy movie for frame range 2000 to 4000 using variance of histogram difference of DC images.

Fig. 6 cut effect detection for funny video for frame range 1 to 2400 using variance of histogram difference of DC images.

Fig. 7 cut effect detection for election video for frame range 4000 to 5000 using variance of histogram difference of DC images.

Fig. 8 cut effect detection for social video for frame range 3000 to 5000 using variance of histogram difference of DC images.

Fig. 6 shows cut detection for funny video for frame range from 1 to 2400. There are 3 high spikes and peaks represent abrupt change in histogram of both DC images. There are total 3 cut effect is detected. Location pairs are frame number 1036-1037, 1060-1061; 1632-1633. There is total 3 instants at which one picture is changing to another picture. In this video between frame numbers 1 to 2400 only 3 times pictures are changing and abrupt transition occurs only 3 times.

Fig. 7 shows cut detection for election video for frame range from 4000 to 5000. There are total two cut effect are there. Location pairs are frame number 4470-4471 and 4476-4477.

Fig. 8 shows cut detection for social video for frame range from 3000 to 5000. Total two cuts are detected and its
location pairs are 3850-3851 and 4007-4008. Between this ranges two times pictures are changing and produce cuts.

6. Conclusions

In this paper we propose an algorithm for finding scene breaks within video frame ranges. Simulation results shows how efficiently and effectively this algorithm can detects cut effect. Concept of DC image is introduced for reduction of processing time and to improve processing speed. DCT is used for compression to represent frames by its meaningful information only.

References

Rakesh Trivedi has completed has B.E. (E. C.) in 2009 from A. D. Patel Institute of Technology, Sardar Patel University, Vallabh Vidhyanagar. He is pursuing his M. E. (CSE) in Communication System Eng. From G.H.Patel college of engineering and technology, vallabh vidhyanagar Gujarat, Technological University. His area of interest is video processing.

Mehul Shah is Associate Prof G.H. Patel college of engineering and technology, Vallabh Vidhyanagar Gujarat. He has more than 12 years of teaching experience. He has completed his M.tech (Electronics & Communication) form DDIT, Nadiad and pursuing PhD from IIT-Bombay. His areas of interest are wireless and mobile communication and image processing.