Shot Boundary Detection for a Video under the Influence of Illumination: A Adaptive Thresholding Approach

Parul S. Arora Bhalotra
Research Scholar (JJT Univ.)
Assistant Professor, ETC
G.H.R.C.E.M. Pune (India)

Nilam N. Ghuge
Research Scholar (JJT Univ.)
Assistant Professor, Electrical
B.S.I.O.T.R.(W), Pune (India)

B. D. Shinde
Professor, ETC
G.N.V.O.I.T. Tala
Raigarh (India)

ABSTRACT

Video Shot boundary detection is the process of automatically detecting the boundaries between shots in video. Shot Boundary detection is an important fundamental process in video data access, indexing, search and retrieval. The increased availability and usage of on-line digital video has created a need for automated video content analysis techniques. Detection of gradual transition and elimination of disturbances caused by illumination change is a major challenge in shot boundary detection technique. These disturbances are often mistaken as shot boundaries. It is a crucial task to develop a method that is not only insensitive to illumination change but also sensitive to detect shot change. An algorithm is proposed for shot boundary detection in presence of illumination change. This is very important for accurate detection of shot boundaries and very useful in content based analysis of video. First the algorithm removes illumination change using discrete cosine transform and discrete wavelet transform. Then shot boundaries are detected using normal difference & wavelet difference. A shot boundary is detected when the feature difference shows sharp change greater than threshold. Experimental study is performed on number of videos that include significant illumination change. The performance of proposed algorithm is better as compared to existing techniques

General Terms

Video shot boundary detection, Image processing, Pattern recognition.

Keywords

Video shot detection, Discrete wavelet transform, Discrete cosine transform, Threshold.

1. INTRODUCTION

Video shot boundary detection has been deeply studied in recent years and has found applications in different domains like video indexing, video compression, video access. Advances in digital technology have made many video archives readily available. Therefore scalable, efficient and effective tools for indexing and retrieving video are needed. With a large amount of information encoded in one video, typically the first step of any video processing tools is to segment the input video into elementary shots in which each shot is defined as a continuous frame from a single camera at a given moment. The detection of shot boundaries provides a base for nearly all video abstraction and high-level video segmentation approaches. Therefore solving the problem of shot boundary detection is one of the major prerequisites for revealing higher level video content structure. Depending on transition between the shots, the shot boundaries can be categorized into two types: abrupt transition and gradual transition. The gradual transition can be further classified into dissolve, wipe, fade in and fade out, according to the characteristics of the different editing effects. The existing methods on shot boundary detection are discussed below.

Likelihood ratio, pair-wise comparison and histogram comparison have been used as a different metric for shot boundary detection by Zhang et al. [1].Object motion & camera motion have been observed as major source of false positives by Boreczky and Lawrence [2]. They presented a comparison of several shot boundary detection classification techniques & their variations including pixel difference, statistical difference, compression difference Histogram, Edge tracking, discrete cosine transform, motion vector and block matching methods. It was seen that algorithm features that seemed to produce good results were region based comparisons, running differences & motion vector analysis. According to Boreczky combination of these three features may perform well to produce better results than either the region histogram or running histogram algorithm. Lienhart [3] has used color histogram differences. Standard deviation of pixel intensities and edge based contrast as a metric to find shot boundaries and tested results on diverse set of video sequences. Henjalic [4] have identified and analyzed the major issues related to shot boundary detection in detail. Knowledge relevant to shot boundary detection, shot length distribution, visual discontinuity pattern at shot boundaries and characteristic temporal changes of visual features around a boundary are needed to be considered for the study.

Gargi et al. [5] have evaluated and characterized the performance of number of shot detection methods using color histogram, moving picture expert group compression parameter information and image block motion matching. Ford et al. [6] have reported results on various histogram test statistics, pixel difference. Yuan et al. [7] have presented a comprehensive review of existing approaches and identified major challenges to shot boundary detection, according to them elimination of disturbances due to motion of large object and camera is a challenge in shot boundary detection. Sethi and Patel [8] have tested statistical test for changes in scene. Jinhui yuan et al. [9] employed three critical techniques i.e representation of visual content, construction of continuity signal & classification of continuity values are identified and formulated in the perspective of pattern recognition. The methods of classification are rule based classifiers and statistical machine learning. Module evaluation and system evaluation was done. For module evaluation the specific module varies in different approaches while the other module of the system retains the same implementation. The three mappings identified by the formal framework are research problem for pattern recognition, which have undergone relatively mature evolution.

In the proposed algorithm illumination change is removed using discrete cosine transform and discrete wavelet transform followed by detection of shot boundaries by applying the techniques of finding difference by conventional method i.e. normal difference and by wavelet method.

The rest of the paper is organized as follows, Section 2 describes discrete cosine transform method for illumination removal, section 3 describes discrete wavelet transform method for illumination change removal. We discuss results & conclude the paper in section 4 and section 5.

2. ILLUMINATION REMOVAL USING DCT

The discrete cosine transform (DCT) helps to separate the image into parts of differing importance with respect to the image's visual quality as shown in Figure 1. The DCT is similar to the discrete fourier transform, it transforms a signal or image from the spatial domain to the frequency domain.
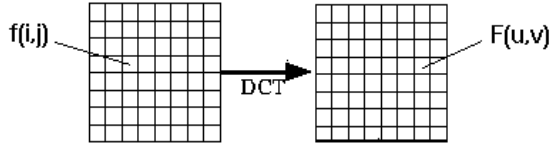
Figure 1:Spatial domain to frequency domain

The general equation for a 1D (N data items) DCT is defined by the following equation:

$$F(\mu) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} A(i). \cos\left[\frac{\mu.\pi(2i+1)}{2.N}\right] f(i) \qquad (1)$$

and the corresponding inverse 1D DCT transform is simple $F^{-1}(u)$, i.e.: where

$$A(i) = \frac{1}{\sqrt{2}} \qquad for\ \xi = 0 \qquad (2)$$
$$1 \qquad\qquad otherwise$$

General equation for a 2D (N X M image) DCT is defined by the following equation:

$$F(\mu,\vartheta) = \left(\frac{2}{N}\right)^{\frac{1}{2}}\left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1}\sum_{J=0}^{M-1} A(i).A(j). \cos\left[\frac{\mu.\pi(2i+1)}{2.N}\right]$$
$$\cos\left[\frac{\mu.\pi(2j+1)}{2.M}\right].f(i,j) \qquad (3)$$

and the corresponding inverse 2D DCT transform is simple $F^{-1}(u,v)$, i.e.: where

$$A(\xi) = \frac{1}{\sqrt{2}} \qquad for\ \xi = 0 \qquad (4)$$
$$1 \qquad\qquad otherwise$$

The input image is N X M. f(i,j) is the intensity of the pixel in row i and column j. $F(\mu,\vartheta)$ is the DCT coefficient in row k1 and column k2 of the DCT matrix. For most images, much of the signal energy lies at low frequencies, these appear in the upper left corner of the DCT. Compression is achieved since the lower right values represent higher frequencies, and are often small enough to be neglected with little visible distortion. The DCT input is an 8 X 8 array of integers. This array contains each pixel's gray scale level. 8 bit pixels have levels from 0 to 255. The output array of DCT coefficients contains integers. It is computationally easier to implement and more efficient to regard the DCT as a set of basic functions which gives a known input array size (8x8) can be precomputed and stored. This involves simply computing values for a convolution mask (8x8 window) that get applied.

The values as simply calculated from the DCT formula. We have set (0,0) component of 2D DCT to zero and then take inverse 2D DCT to remove illumination/change.

3. ILLUMINATION REMOVAL USING DWT

Wavelet is a suitable tool to decompose an image signal into subbands. Not only can it gives the desired low frequency and high frequency information but also it is fast and easy to compute and require only linear time in the size of the image. The discrete wavelet transform corresponds to the continuous wavelet transform of a sampled sequence x(n)= x(nT), where T is sampling period. The discrete wavelet transform (DWT) applies to discrete time signals, where both time and time-scale parameters are discrete. A DWT output on scale J consists of wavelet coefficients computed for j = 1,...,J

$$\{x(n); 2^j, k2^j\} = C_{j,k} = \sum n x_n h_j^*(n - 2^j k) \qquad (5)$$

The discrete wavelet transform can be used to analyze or decompose the signals and images. This process is called decomposition or analysis as shown in Figure 2. The other half of the story is how those components can be assembled back into the original signal without loss of information. This process is called reconstruction or synthesis. The mathematical manipulation that effects synthesis is called the inverse discrete wavelet transforms (IDWT). To synthesize a signal we reconstruct it from the wavelet coefficients.
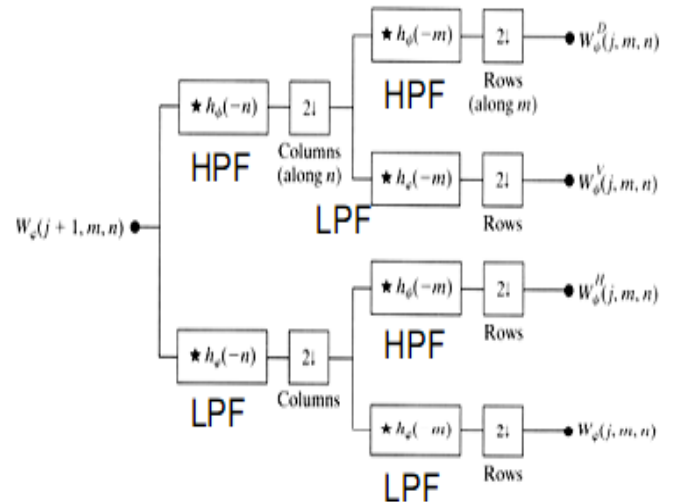

Figure 2: One stage 2D analysis bank

Analysis filter bank block decomposes a broadband signal into a collection of subbands with smaller bandwidths and slower sample rates. The block uses a series of highpass and lowpass filters to repeatedly divide the input frequency range as illustrated in the Figure 3. The filter bank's highpass and lowpass filters are provided by vectors of filter coefficients. In order to be efficient and useful. A method designed for analysis also has to be able to perform synthesis. The wavelet method achieves this. Synthesis is the reciprocal operation of analysis.
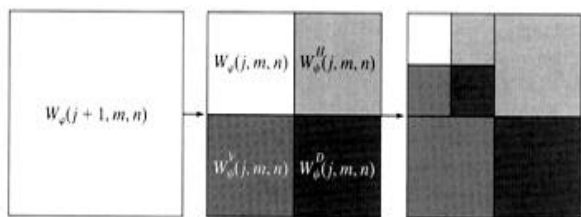
Figure 3: Two stage Decomposition

The block takes subbands of a signal and uses them to reconstruct the signal by using a series of highpass and lowpass FIR filters as illustrated in the Figure 4. The reconstructed signal has a wider bandwidth and faster sample rate than the input subbands.
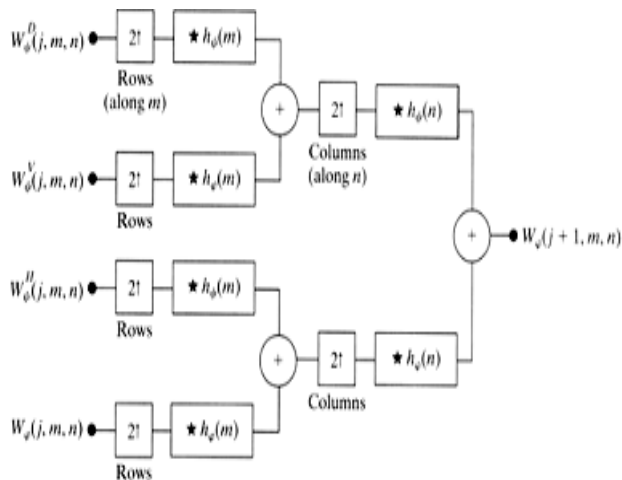

Figure 4: One stage 2D synthesis bank

We have set the lowest approximation cut off equal to zero and then take 2D Inverse discrete wavelet transform.

4. RESULTS & DISCUSSION

Figure 5 shows original image which is given as input to the DCT block. Figure 6 is the result after illumination removal by DCT method and Figure 7 shows graph between no. of frames versus frame difference using DCT.
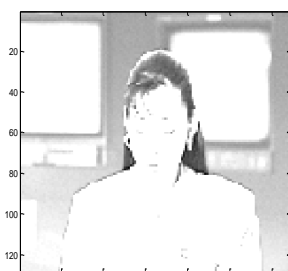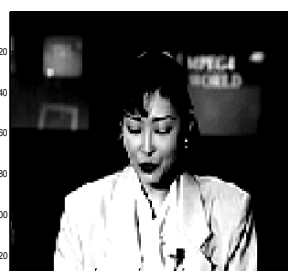

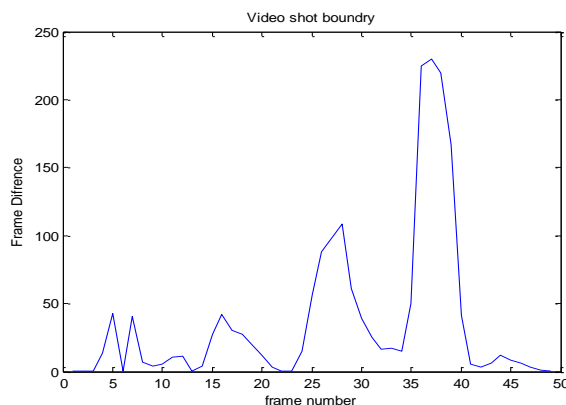Figure 5: Original Image

Figure 6:After illumination removal
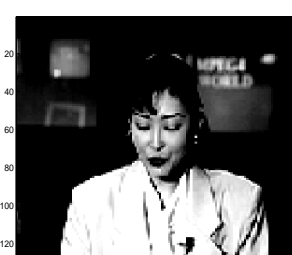

Figure 7: Graph1

Figure 8:Original Image

Figure 9:After illumination Removal

Figure 8 shows original image which is given as input to the DWT block. Figure 9 is the result after illumination removal by DWT method and Figure 10 shows graph between no. of frames versus frame difference using DWT.
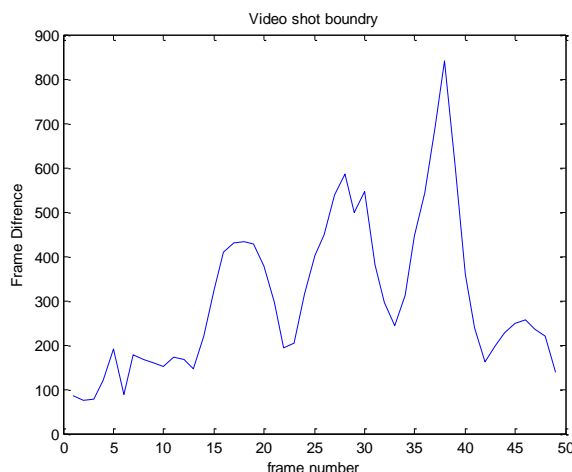

Figure 10: Graph 2

The methods used for shot boundary detection are normal difference method and wavelet difference method. In normal difference method difference between two consecutive frames is measured, if the difference measured is maximum then shot boundary is detected.

In wavelet difference method initially discrete wavelet transform of the two consecutive frames is calculated and then LL component is considered for calculating the difference between the two consecutive frames.

To calculate this difference Euclidean distance formula is used. The Euclidean distance or Euclidean metric is the "ordinary" distance between two points. The Euclidean distance between points p and q is the length of the line segment connecting them .The Euclidean distance between points p and q is the length of the line segment connecting them. Then the distance from p to q or from q to p is given by

$$d(p,q) = d(q,p) + \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \qquad (6$$

where in Cartesian coordinates
p = (p$_1$, p$_2$,..., p$_n$),
q = (q$_1$, q$_2$,..., q$_n$)

The Euclidean distance measure is used to measure the similarity between the frames. Based on the set threshold, a shot is said to be detected if the dissimilarity between the frames is higher than the threshold value.

A shot boundary is detected when the feature difference calculated shows sharp change greater than predefined threshold determined by an adaptive scheme. The continuity signal which is graph of frame difference *dm* against frame index on time axis is shown in Figure 11.
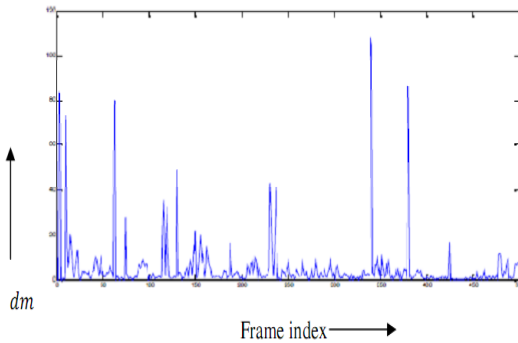

Figure 11: Graph 3

The problem of choosing the appropriate threshold is a key issue in the shot detection algorithms. Heuristically chosen global thresholds are inappropriate as experiments have shown that the threshold for determining a segment boundary varies from one shot to another which must be based on the distribution of the frame-to-frame differences of shots. So adaptive threshold is more reasonable than global threshold. In proposed algorithm, local window based threshold calculation method is verified. A window of m frames is built and then the mean value μ is calculated and the threshold T is defined as
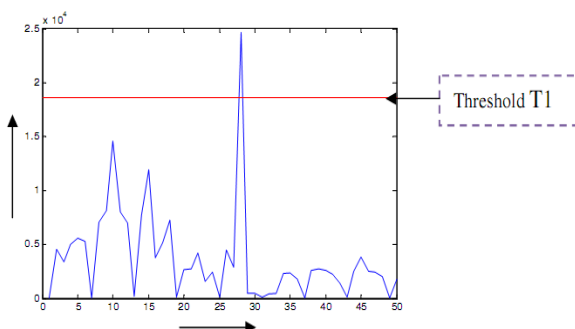
T= 5× (Mean μ of the widow size m) (6)

Figure 12: Shot boundary detected by Threshold

Where value of m is decided as 50 empirically. The dissimilarity value greater than threshold T1 indicates shot boundary. The Figure 12 shows shot boundary being detected between frame 28 and frame 29 in the experimental data set.

5. CONCLUSIONS

Disturbances caused by illumination change are often mistaken as shot boundaries, and its elimination is the major challenge to the shot boundary detection algorithms. In this paper, we address this issue by developing the structure features that are invariant to illumination using discrete wavelet transform & discrete cosine transform. Then to detect shot boundaries we used Normal difference method i.e. conventional method & wavelet method. A shot boundary detected if the difference in the frames is higher than the threshold value.

Our proposed algorithm has been tested on various videos, and is successful in avoiding disturbances due to illumination change and successfully detecting the shot boundaries.

6. REFERENCES

[1] H. J. Zhang, A. Kankanhalli, and S. W. Smoliar, 1993, ''Automatic partitioning of full-motion video'', Multimedia systems vol.1, pp. 10–28.

[2] J. S. Boreczky and L. Rowe, Feb. 1996, "Comparison of video shot boundary detection techniques", in Procedings IS&T/SPIE Storage and retrieval for still image and video databases IV, vol. 2670, pp. 170–179.

[3] R. Lienhart, Sept. 2001, "Reliable transition detection in videos: A survey and practitioners guide", International journal of image and graphics, vol.1, no. 3, pp. 469 – 486.

[4] A. Hanjalic, Feb. 2002, "Shot-boundary detection: Unraveled and resolved?", IEEE transaction circuits system video technology, vol.12, no. 2, pp. 90 – 105.

[5] U. Gargi, R. Kasturi, and S. H. Strayer, Feb.2000, "Performance characterization of video shot-change detection methods", IEEE transaction circuits systems video technology, vol.10, no.1, pp.1 – 13.

[6] R ford, C Robonson, D Temple & M gelach, 2000, " Metrics for short boundary detection in digital video system", vol. 8, pp. 37-46.

[7] J. yuan, H. wang, L. xiao, W. zheng, J Li, F Lin, et al. Feb. 2007, "A formal study of shot boundary detection on circuit & systems for video technology", vol. 17,no. 2, pp. 168-186.

[8] K. sethi & N. Patel, Feb 1995, "A statistical approach to scene change detection", SPIE, proceedings on storage and retrieval for image & video database III, vol. 2420, pp. 329-338.

[9] Jinhui Yuan, Huiti wang, lan Xiao, Wujie Zheng, Jianmin Li, Fuzong Lin, & Bo Zhang , 2007, "A formal study of shot boundary detection", IEEE Transaction on circuits & systems for video technology", vol. 17, no. 2, pp. 234-239.