Scene Change Detection Based on Twice Difference of Luminance Histograms

Xinying Wang¹, K.N.Plataniotis², A. N. Venetsanopoulos¹ ¹ Department of Electrical & Computer Engineering University of Toronto (Wang, Visiting Scholar ,Xinjiang University) 10 King's College Road, Toronto, Ontario, M5S 3G4 CANADA

u

 ² Department of Math, Physics & Computer Science Ryerso n Polytechic University
 350 Victoria Street, Toronto, Ontario,M5B 2K3 CANADA

Abstract: - This paper introduces a new method for video scene change detection. The proposed method utilizes the luminance histogram twice difference in order to determine the dynami threshold needed to evaluate the break. The adaptive determination of the threshold, minimizes the nume ber of incorrect decisions leading to a robust and accurate determination of the actual scene break. The method is simple, computationally attractive and capable to detect changes in a variety of visual inputs. Experimental results indicate that the new method constantly outperforms existing techniques that are based on static thresholds.

Key-Words :- Scene Change Detection, Luminance Histograms, Twice difference, Dynami Threshold, RGB Color Space , YCbCr Color space .CSCC'99 Proceedings, Pages:4151-4155

1. Introduction

Video is the most popular source of multimedia information. It combines all other medi information, such as, text, image, graphic, audio and etc into a single data stream. Scene change detection, an effective method for segmenting a video sequence into significant components, generally called shots, has been recognized as one of the important research areas in recent years [1], [2]. It is an important technology for video editing, video indexing, and motion compensation, etc.

Scene change, is defined as an unbroken sequence of frames recorded by a single

camera, which forms the building block of video stream. After shots are segmented, key frames can be extracted from each shot. Key frame is the frame which can be used to represent the salient content of the shot. Depending on the content complexity of the shot, one or more key frames can be extracted from a single shot. Therefore, The first problem is to develop an effective algorithm of detecting scene changes [2].

Many scene change detection methods can be found in the literature, but some of them are either computationally expensive or ineffective [3], [4]. Express the qualitative difference in content of frames through the video sequence by a suitable metric to find out a frame to frame comparison value. Shots change could be detected by using an appropriate threshold. Hard thresholds cannot perform equally well for all videos, which must be assigned to tolerate variations in individual frames, while still ensuring a desired level of performance. A high threshold value can prevent false positive cuts to be accepted, but increases the number of missed shots. Conversely, a low threshold value enable consistent cuts to be accepted, but increases the number of false positive shots [5]. Therefore, it is a key problem to design an algorithm to obtain an adaptive threshold. Because the high accuracy is a more important requirement in automatic video segmentation than other applications. Only with a correct threshold, we can get high accuracy results. Thus, the threshold value selected is an appropriate but fixed value, which depends on the chosen metric and experiments show that false and missed cuts persist. This problem can be resolved using a dynamic threshold which could vary through the video sequence. This paper presents a method to calculate an adaptive threshold to improve the accuracy of scene change detection.

2. Approach

A scene change signifies either a significant change in the overall color composition or a significant change in the object location or both. To detect this content change automatically, we propose to use luminance (Y) histogram difference metric hereby term as LHD. The luminance(Y) histogram based metric is robust to camera as well as object motion. The luminance(Y) image is gray scale one. It is with higher clarity than common gray scale ones, and luminance(Y) histogram information is more concentrate, they are extremely sensitive to object and camera motion. Another advantage of the use of luminance (Y) histograms is that they are rotation and translation invariant for a constant background. Therefore, we consider the YCbCr color space motivated by its adoption in MPEG applications. It has also been shown that the luminance and chrominance information in YCbCr color space can be interpreted independently for scene change detection.

2.1 Color Space Convert

An important parameter in the performance of the scene change detection algorithm is the color space used. The RGB color space is the most widely used color format, yet it does not properly reflect the properties of the human visual system. The YCbCr color space provides a better alternative to the RGB color system, especially when the luminance(Y) histogram is more coarsely quantized than the chrominance channels. It is also the color format adopted in the compression standards M-JPEG and MPEG, and hence is computationally appropriate as well [2].

Most scene detection techniques utilizing the RGB color space [5], [6]. Although RGB color space is the most widely used color format in everyday applications, yet it does not properly reflect the properties of human visual system. Otherwise, in RGB color space, one has to calculate three metrics. It is very time and space consuming. The YCbCr color space provide a better alternative to the RGB color system, especially, YCbCr luminance(Y) is a gray scale one, but it is different from the common gray scale space. It includes much more information to perform scene change detection. While these metrics take into account the spatial changes, they are extremely sensitive to object and camera motion.

In order to build luminance(Y) histogram form the RGB image images, we must convert sequence from RGB values to YCbCr color space. YCbCr MAP is a M-by-3 metric that contains the YCbCr luminance(Y) and Chrominance (Cb and Cr)color values as corresponding row in the RGB color map. In RGB color space, in order to delete scene change, we have to calculate H_R, H_G, H_B three metrics. It is very time and space consuming. But in YCbCr color space, because of luminance(Y) characteristics, we only need to calculate the luminance(Y) histogram H_Y metric. Thus, we both simplify the calculation process and save time and space.We take Y metric from YCbCr M-by-3 metric, then build luminance(Y) histogram.



Fig. 1

2.2 Twice Difference Algorithm

Based on the luminance(Y) histograms, we use the histogram comparison method, which is a widely know method for detecting scene change.

$$D_{Y}(i) = \sum_{j=1}^{256} |H_{i}(j) - H_{i+1}(j)| ,$$

$$i = 1, 2, ..., n \text{ #frames} \qquad (1)$$

In Eqn. (1), $H_i(j)$ denote the luminance(Y) histogram of Y component, respectively, of frame i, and j denotes one of the G bins. The component $D_Y(i)$ simply provides quantitative measure of gray scale between successive frames.

Based on luminance (Y) histograms difference result, performing twice difference and decision condition to get adaptive Threshold. The algorithm is following:

In this algorithm, we first based on the Eqn.(1) results, i.e difference luminance histograms

results D_Y (i), i = 1, 2, ..., n #frames, performing following decision condition :

$$|D_{\mathbf{Y}}(\mathbf{i}) - D_{\mathbf{Y}}(\mathbf{i}+1)| > D_{\mathbf{Y}}(\mathbf{i}+1))\mathbf{x} l$$
,
 $0 < l < 1$

if $|D_Y(i) - D_Y(i+1)|$ satisfied this decision condition, save $max(D_Y(i), D_Y(i+1))$ in $D_d(k)$, this value should be scene change ones. like this, we get all of $D_Y(i)$ which satisfied the decision condition, and save them in $D_d(k)$, then calculate minimum $T = \{ D_d(k) \}$ as the adaptive threshold. Finally, use this threshold in the Eqn.(1) result, to find all of $D_Y(i)$ satisfied decision condition $D_Y(i)T$, and marked their positions. So far, we get all of the shots in the correspond sequences.

In this algorithm, l is an important value. l is not a prespecified constant, and varies adaptively depending on the content. For instance, l is large in a sports or dance clip, while is small for a news clip. We summarize the steps of the unsupervised change detection algorithm in following:

Step 1: Convert RGB color space value to YCbCr color space .

Step 2: Extract luminance(Y) metric Y from YCbCr M-by-3 metric.

Step 3: Calculate luminance histograms H_1 , H_2 ,..., H_n .

Step 4: Calculate the luminance histograms difference by Enq.(1), get $D_Y(i)$, I=1,2,...,n-1.

Step 5: Performance twice difference:

Step 5a: if $|D_Y(i) - D_Y(i+1)| > D_Y(i+1) xl$,

 $D_{Y}(1+1)$

Step 5b: Save all of the $max(D_Y(i),D_Y(i+1))$ satisfied

the above decision condition in the $D_d(k)$.

Step 5c: Calculate the adaptive threshold: $T = \min \{D_{i}(k)\}$

$$I = \min\{ D_d(k), \}$$

Step 6: Use T in $D_Y(i)$, I=1,2,...,n-1., label Scene change frames.

3. Results and Discussion

In this section, we present the results of the adaptive threshold obtained from 100 frames of video sequence named Foreman. The

frames are of resolution 144x176 with pixel represented by 24 bit of data, 4x8x4 color map. First, we convert the sequence form RGB value to YCbCr color space, then take Y metric and build luminance(Y) histograms $H_1, H_2, \ldots H_{100}$, second, we used the histogram comparison method to get :

$$\begin{split} D_{Y} \ (i) &= \ \sum_{j=1}^{256} \ \mid H_{i}(j) - H_{i+1}(j) \quad \mid \ , \\ &i = 1, \ 2 \ , \dots, 100 \ \# frames \ \ (1) \end{split}$$

then, we use above algorithm to perform twice difference to find all of D_Y (i) satisfied the decision condition and calculate the adaptive threshold. In this sequence, we chose l=0.25, computed T = 3312, finally, we use T in the Eqn.(1), to find all of D_Y (i) T, we get 11 shots in this sequence. The results show no one is missed and no false positions. Almost 100% correct. (Fig.2)

4. Conclusion

We have presented in this paper a twice difference of luminance histograms based framework for temporal segmentation. The method has been introduced to perform unsupervised. The detection performance of the algorithm was further improved by twice difference of luminance histograms method,

Which helped significantly reduce the number of missed and false shots. This approach gives the users a useful method to select adaptive threshold to more effective detect the scene change. This method is easy to understand and achieve for users. The results show, as long as correctly chose l, following the algorithm steps, we can get high accuracy results. However, this method need to be further transitions expended to detect complex between scene change as fades and dissolve. This brings us the question of whether the luminance (Y) histograms difference and condition used is able to express these types

of transitions on the suitable decision conditions.

Reference:

[1] X1, Yueting Zhuang and etc, " Adaptive key frame extraction using unsupervised clustering,"

IEEE International Conference on Image Processing Vol. 1 of 3,1998, pp866-870. [2] X1,A. Müfit Ferman and A.Murat Tekalp, "Efficient filtering and clustering methods for tempo-

ral video segmentation and visual summarization," *J.Vis Comm. and Image Rep*, Vol.9,no 4

(special issue),1998, pp336-351.

[3] X1, Bilge Günsel and A.Murat Tekalp, "Content-based video abstraction," *IEEE International*

Conference on Image Processing Vol. 3 of 3,1998, pp128-132.

[4] X1, Cuneyt Taskian and etc, "A Compressed video database structured for active browsing and

search," *IEEE International Conference on Image Processing* Vol. 3 of 3,1998, pp133-137.

[5] X1, Mohsen Ardebilian, "Improvement of shot detection methods based on dynamic threshold

selection," SPIE vol.3229.0277, 1997, pp14-22.

[6] X1, M.R. Naphade and etc, "A High performance algorithm for shot boundary detection using

multiple cues,"*IEEE International Conference on Image Processing* Vol.1 of 3,1998, pp884-887.



















