Mathematical Problems of Computer Science 44, 77--84, 2015.

Video Shot Detection Method Based on Histogram Comparison Procedure

Manuk K. Zakaryan

Russian - Armenian (Slavonic) University
e-mail: zakaryanmanuk@yahoo.com

Abstract

Video segmentation and shot detection are the main components in automatic video indexing, archiving, editing and retrieval. For the last fifteen years a lot of techniques have been proposed some of which have reliable performances in the video processing, especially for the video cut detection, key frame extraction and browsing technique. They have a big range of usage mainly in medical diagnosis, military, retail catalog, traffic control, security, etc. In this paper we provide a comparative analysis of histogram and certain structural properties based methods for video cut detection problems. Advantages and disadvantages of the mentioned methods are shown.

Keywords: Video segmentation, Shot detection, Image similarity measure, Histogram.

1. Introduction

Saying content-based image retrieval we understand an extraction of a range of images which is relevant with the given image from a large database of images. Here we first need to get the feature of each image data and store it in a table. The next step is to perform some similarity measurement algorithm to find similar image data. The main question firstly comes to mind "what is the best feature to choose?"

There are many image features like color, texture, shape. The first step is to extract these features from image.

There are several video segmentation algorithms currently proposed in the literature which are based on various statistical properties of the shots. We briefly consider some of them.

- Pixel values direct comparison method - one of the well-known and widely used methods based on Mean Square Error method (PSNR).
- Color distribution method - mostly used method is color histogram comparison.
- Optical flow method - mainly used when there is a panning or zooming of video camera.
- Object edges method
- Macroblock type method

Nevertheless, as it is mentioned in [1] they are developed for the specific range of problems and may sometimes be not useful for a wide variety of video sequences. More precise analysis of digital video segmentation is given in [2]. In this paper we will mainly discuss the image histogram comparison and the proposed method named W^2.

It is known that human is more sensitive to the color features of an image than to shape or texture. Existing retrieval systems based on color features are more popular. Here are some methods which extract color features - histogram, color moments, color correlogram, color coherence vector [3], etc. Sometimes it is not effective to use only color-based features for similarity measure, because it may produce a problem in similar color distribution in different images. So this kind of methods are mainly combined with some other similarity measures to exclude this disadvantage.

2. Color Histogram Analysis Based Method

Histograms are collected counts of data organized into a set of predefined bins. Saying data we are not restricting it to be a color. The data collected can be whatever feature we find useful to describe our image. In our case it will be intensity values. Let's imagine a matrix which contains information about an image pixels intensities in the range from 0 to 255.

To compare two histograms to express how well both histograms match we choose a metrics X^2 mostly known as *Chi-Square* metrics [4].

Let us have two histograms H_1 and H_2 that we want to compare. N is the number of pixels in the first image, and M is the number of pixels in the second image. So the difference $d(H_1, H_2)$ between these two histograms will be expressed by this formula:

$$d(H_1,H_2)=\sum_{i=1}^{K}\frac{\left(\frac{n_i}{N}-\frac{m_i}{M}\right)^2}{\frac{n_i}{N}+\frac{m_i}{M}}, \qquad (1)$$

where K is the number of bin segments, n_i and m_i are the numbers of pixels in i-th bin for H_1 and H_2 histograms, correspondingly. For segmentation problems we are splitting the video to the same size of frames, as in the experiments $M = N$.

For shot boundary detection we use this method and compare consecutive frames histograms with (1) equation and where it gets its maximum value there we consider that cut has taken place.

Another metrics that is used for comparison of histograms is *Correlation* method [5]. To compare H_1 and H_2 histograms we are using the following equation:

$$d(H_1,H_2)=\frac{\sum_I\left(H_1(I)-\overline{H}_1\right)\left(H_2(I)-\overline{H}_2\right)}{\sqrt{\sum_I\left(H_1(I)-\overline{H}_1\right)^2\sum_I\left(H_2(I)-\overline{H}_2\right)^2}}, \qquad (2)$$

where

$$\overline{H}_k=\frac{1}{N}\sum_J H_k(J),$$

and N is the total number of bins.

Histogram intersection (HI), proposed by Swain and Ballard, is a straightforward method to calculate the matching rate between two histograms for this purpose [6].

$$d(H_1,H_2)=\sum_I \min\left(H_1(I),H_2(I)\right). \qquad (3)$$

The resultant fractional matching value between 0 and 1 is actually the proportion of pixels from the model image that has corresponding pixels of the same color in the target image. A higher histogram matching rate indicates higher similarity between the model image and the target image.

Bhattacharyya distance. Bhattacharyya coefficient is one of the criteria, which gives a measure of similarity between the probability density functions (spectral information) of two images [7, 8]. It is a divergence-type measure which has a straightforward geometric interpretation:

$$d(H_1, H_2) = \sqrt{1 - \frac{1}{\sqrt{\bar{H}_1 \bar{H}_2 N^2}} \sum_I \sqrt{H_1(I), H_2(I)}}. \quad (4)$$

3. Experimental Results

We will apply the usage of the above mentioned methods for video anni008.wmv taken from famous TRECVID library of video sequences. Segmentation results are shown in figures below. The video contains 15 shots.

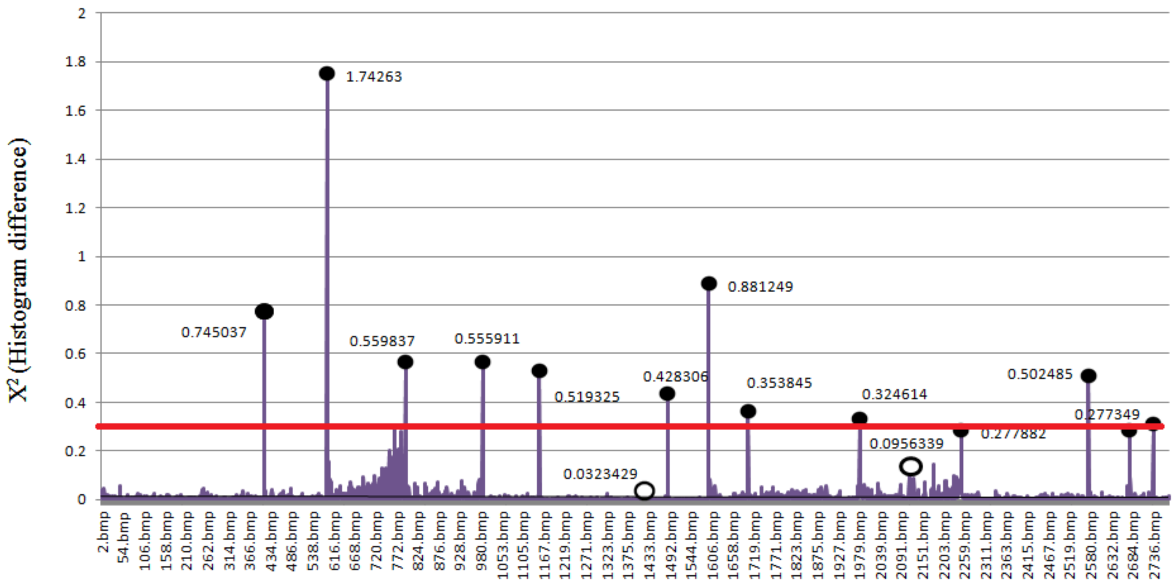

Fig. 1. Comparison of Histogram values with Chi-Square method.

In Fig.1 we provide experimental results comparing frames with Chi-Square method. We got 13 cuts that have been accurately detected and 2 cuts that have not been detected. The threshold value was 0.3.

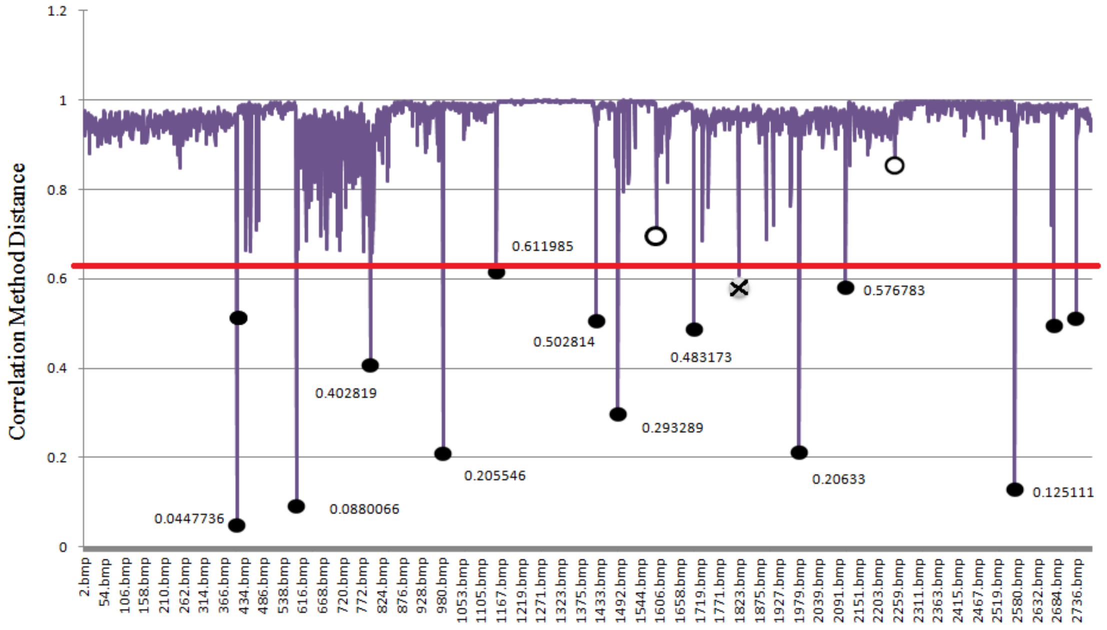

Fig. 2. Comparison of Histogram values with Correlation method.

In Fig.2 results with Correlation method are provided. It accurately detected 13 cuts, but 2 cuts were not detected and there was also one false cut. Threshold value was 0.65.

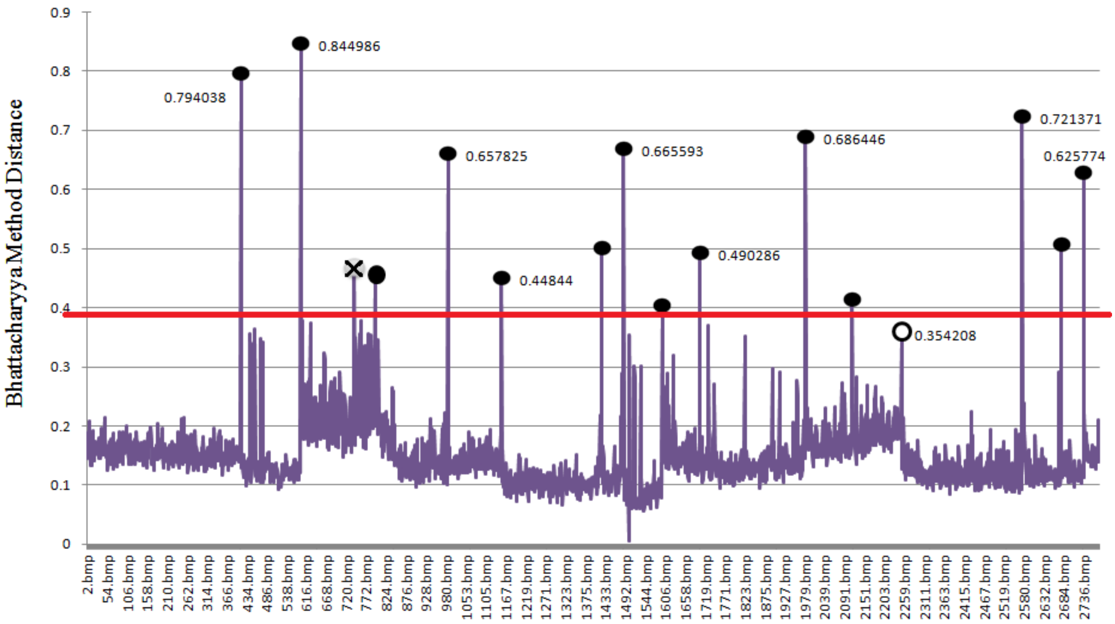

Fig.3. Comparison of Histogram values with Bhattacharyya method.

Fig.3 illustrates comparison results with Bhattacharyya method. 14 cuts have been detected, and there is one missed cut and one false cut. Threshold was selected equal to 0.39.

● - Real cuts that have been detected
○ - Real cuts that have **NOT** been detected
✕ - False cuts

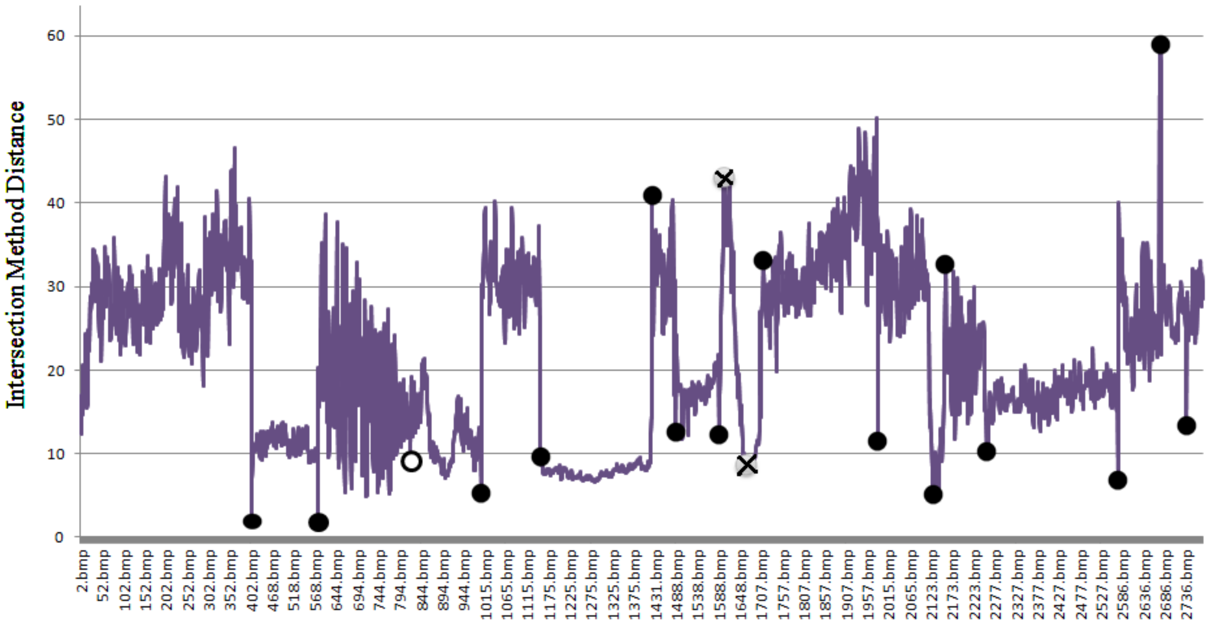

Fig. 4. Comparison of Histogram values with Intersection method.

The results are given in Figures from 1 to 4. As seen from the data, all 4 methods did not give absolute results. As it was mentioned previously the main two disadvantage of this method is that different images may have the same color distribution (cuts that have not been detected) and the change of brightness may totally affect the histogram values (false cuts). In Figure 5, 2 structurally different images are illustrated that have almost the same histograms.

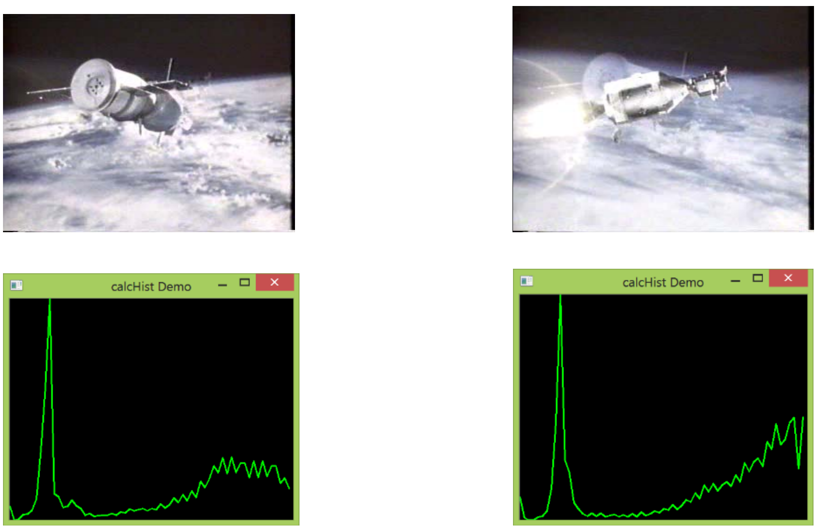

Fig. 5. Consecutive frame of the cut and corresponding histograms.

In experimental results we also provide a comparison of these methods with the proposed method which is based on structural properties of the image and called W^2

$$W^2 = \frac{\min(b_1, b_2)\min(c_1, c_2)}{\max(b_1, b_2)\max(c_1, c_2)}, 0 < W^2 \le 1, \quad (5)$$

where the corresponding parameters b and c are represented as statistical estimations got from the corresponding samples of gradient magnitude [9, 10].

Frame 1 Frame 2 Frame 3

Fig. 6. Frames to be compared.

In Table 1 we provide comparison results of four histogram methods and also provide the results got with W^2 method. It was done for consecutive frames given in Fig.6. As we can see Correlation and W^2 methods give results in range [0, 1] and if the consecutive frames are similar they give values near 1, for Chi-square and Bhattacharyya methods, the more similar are images the less are results near 0. Intersection method gives higher values if images are similar and decreased when the cut appears.

Table 1. Metric results of comparison of 3 consecutive frames with 4 histogram-based and W^2 methods.

Method	Fr. 1 - Fr. 1	Fr. 1 - Fr. 2	Fr.2 - Fr. 3
Chi-square	0.0000	0.0104	0.0923
Correlation	1.0000	0.9466	0.5676
Bhattacharyya	0.0000	0.2213	0.5106
Intersection	24.4915	24.3512	13.0642
W^2	1.0000	0.9513	0.6087

4. Conclusion

In this paper we have performed comparative analysis of histogram-based methods. The experimental results have shown that though they work fine in most cases of video segmentation problems, but they have two main disadvantages: the histogram comparison may give a wrong result when the brightness of similar frames has been changed, it explains false cuts that we saw in test results. It very often is happening when there are light effects in the scene or it has been taken outside. Second problem is that two different images may have the same intensity and that is why we got cuts that had not been detected. In some cases it can occur that one of the considered methods provides himself better than the others. So to achieve more accurate results it is suggested to combine the mentioned methods.

References

[1] R. Lienhart, “Reliable transition detection in videos: A survey and practitioners guide”, *International Journal of Image and Graphics*, vol. 1, no. 3, pp. 469--486, 2001.

[2] I. Koprinska and S. Carrato , “Temporal video segmentation: a survey”, *Signal Process.: Image Communication*, vol. 16, no. 5, pp. 477 – 500, 2001.

[3] R. Brunelli and O. Mich, “Histograms analysis for image retrieval”, *Pattern Recognition*, vol. 34, no.8, pp. 1625-1637, 2001

[4] X. Jiang, T. Sun, J. Li and Xi Chen, “A novel shot edge detection algorithm based on chi-square histogram and macro-block statistics”, *International Symposium on Information Science and Engieering*, vol. 2, pp. 604 - 607, 2008.

[5] M. Sizintsev, O. N Toronto, K. G. Derpanis and A. Hogue, “Histogram-based search: A comparative study”, *IEEE Conference on Computer Vision and Pattern Recognition, CVPR* 2008, pp. 1 - 8, June 2008.

[6] W. Jia, H. Zhang, X. He and Q. Wu, “A comparison on histogram based image matching methods”, *IEEE International Conference on Video and Signal Based Surveillance*, 2006. AVSS '06, p.97, Nov 2006.

[7] M. S. Khalid and M. Bilal Malik, “Biased nature of Bhattacharyya coefficient in correlation of gray-scale objects”, *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, pp. 209 - 214, Sept. 2005.

[8] S. Dubuisson, “The computation of the Bhattacharyya distance between histograms without histograms”, *2nd International Conference on Image Processing Theory Tools and Applications (IPTA)*, pp. 373 - 378, July 2010.

[9] D. Asatryan and K. Egiazarian, “Quality assessment measure based on imagestructural properties”, *Proc. of the International Workshop on Local and Non-Local Approximation in Image Processing*, Finland, Helsinki, pp. 70-73, 2009.

[10] D. Asatryan and M. Zakaryan, “Method for video shot detection and separation”, *International Journal "Information Models and Analysis"*, vol. 3, no. 3, p. 247-251, 2014.

Submitted 04.08.2015, accepted 20.11.2015

Սյունապատկերների համեմատման եղանակով տեսահաջոր-դականության հատվածավորման մեթոդ

Մ. Զաքարյան

Ամփոփում

Տեսահաջորդականության հատվածավորումը և անցումների հայտնաբերումը տեսանյութի ինդեքսավորման, արխիվացման, խմբագրման և որոնման մեթոդների հիմնական բաղադրիչներն են: Վերջին 15 տարիներին առաջարկվել են բազմաթիվ եղանակներ, որոնց նպատակն է հատվածավորումը հուսալիորեն իրականացնել հատկապես տեսահատվածների կտրուկ անցումների համար: Սույն հոդվածում ուսումնասիրվել են տեսահաջորդականության սյունապատկերների, ինչպես նաև պատկերի կառուցվածքային որոշակի հատկությունների վրա հիմնված եղանակներ: Բերվել են նշված եղանակների համեմատական վերլուծության արդյունքներ:

Метод сегментации видеопоследовательности, основанный на сравнении гистограмм

М. Закарян

Аннотация

Сегментация видеопоследовательности является основным компонентом при индексации, архивации, редактировании и поиске видеоматериалов. За последние 15 лет были предложены многочисленные методы, предназначенные для надёжной сегментации, особенно при резких переходах видеосегментов. В настоящей работе исследованы методы, основанные на гистограммы видеопоследовательности, а также на определённые структурные свойства изображения. Приведены результаты сравнительного анализа указанных методов.