Color Difference Histogram for Feature Extraction in Video Retrieval

Anjali Sharma
Computer Science and Engineering
Meerut Institute of Engineering and Technology, Meerut, India
anjali.shr.16@gmail.com

Ajay Kumar Singh
Computer Science and Engineering
Meerut Institute of Engineering and Technology, Meerut, India
ajay41274@gmail.com

Abstract— Content Based Video retrieval is another facet of Content Based Retrieval systems. It is to be noted that videos are far more complex than images as they are a sequence of changing images and audio tracks. Here the content of the videos itself is matched with the query of the user. Video retrieval consists of a series of steps such as segmentation, frame selection, feature extraction and similarity measure. This paper is based on the fact that a video is a number of frames i.e. sequence of individual images which are moved at a certain speed. Therefore video is a sequence of moving images and same procedures used of Content based Image Retrieval for feature extraction can be applied to the video retrieval as well. The paper presents the approach called Color Difference Histogram which exploits the perceptually uniform color differences between two points under different background in accordance with the color and the edge orientations. As the system is based on the content of the videos, the query used here is a snap shot taken from the video recorder.

Keywords- Video Retrieval; Content based retrieval; feature extraction; color difference histogram

I. INTRODUCTION

Today the usage of multimedia data around the world has been increased to a different level. The development of all in one electronic gadget such as PDAs, smart phones laptops and digital cameras the generation of images, videos and audios is just a click away [1]. A lot of research is being done to find better options for searching the information. Out of different types of multimedia data available, such as images, videos and audios; videos are far more complex and hard to retrieve. The traditional methods of web search for video retrieval were limited and less efficient as there were lack of understanding and limited processing of video contents [2]. The methods used were based on annotations and metadata available. Therefore they suffer with all the drawbacks of traditional text-based search systems [3]. In order to remove the demerits and get a robust image retrieval system the concept of Content Based Retrieval came into existence. It is evident from the name that the retrieval is carried out on the basis of the content of the query itself. Just as images the Content Based Video retrieval is desirable because most of the video retrieval systems are based on the traditional annotation based search.

Just as images the video contents are of two types; low level like color [4], shape [5] and texture [6] and high level contents such as temporal activity, entropy and other statistical features. Most of the retrieval processes are based on the low level feature extraction. But feature extraction is not the first step for the video processing. A complete video retrieval process consists of a series of processes. Figure 1 shows the sequence of the processes. The steps are as follows:

- Query for video retrieval: the query for the retrieval process could be a short clip or a snap shot taken from the video recorder.
- Video segmentation: this process segments the video into shots; generally used when we take a video clip as a query.
- Key frame selection: this process selects a main or a key frame from which the feature vectors to be extracted.
- Feature extraction: it generates the feature vector from the key frame.
- Similarity measures: these are generally the distance metrics which are chosen in accordance with the method used to generate the feature vector.

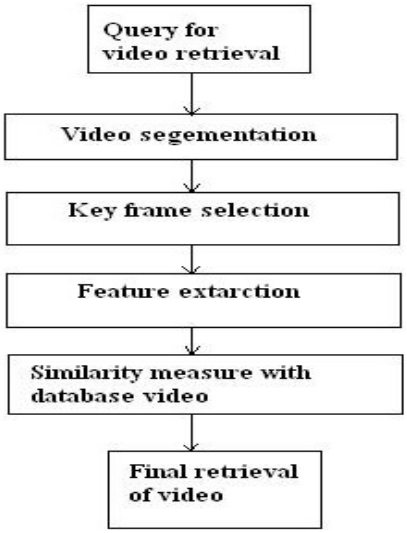

Figure 1 Basic video retrieval process

This paper focuses on a method for feature extraction which is proved to be a better option for generating feature vectors for Content Based Image Retrieval [7]. The method is generation of a histogram known as Color Difference Histogram which will be discussed in later sections.

II. RELATED WORK

A considerable work has been done in the field of content based information retrieval exploiting the low level contents of images and videos such as color, shape and texture [8]. Talking about the querying system in the image retrieval system; it is done keeping in mind the visual content of the images such as color, shape, sketch or simply a similar image. It is to be noted that the video retrieval system is just an extension of the process of image retrieval. It is done by adding two processes of segmentation and key frame selection. Later on the work is similar as that of image feature extraction. Key frame extraction can be done in several ways: Video Shot Method, Content Analysis Method, Cluster Based method and motion based analysis.

But in a general way video retrieval can be divided into two simple steps: (1) extraction of video segments to get the characteristics features and (2) find a similarity comparison to fetch similar videos from the database. In this regard Content Based Video Retrieval is considered one of the most difficult.

Just as the images have several ways of processing and retrieval videos also have different ways, for example video retrieval based on colors [9], video retrieval based on texture, video retrieval based on entropy. All these different manners of retrieving videos effectively and efficiently allowed Content Based Video Retrieval to be a promising way to develop future video search engines [10].

Key issue in content based video retrieval is effective motion feature extraction. Features can be both either low or high level features. Color histograms and Edge histograms are mostly used in order to detect color and edge features [11]. The physical properties of images are represented by texture. Various algorithms have been designed for texture analysis, such as gray level co-occurrence matrices. This approach explored the gray level spatial dependence of texture [12].

III. PROPOSED METHOD

This paper focuses on the combine the color and the texture feature extraction in a single algorithm to make programming effort easier and less time consuming. Some algorithms can ultimately combine color and texture together these include the color edge co-occurrence histogram integrative co-occurrence matrices, the Texton co-occurrences matrix, the micro-structure descriptor and the multi-Texton histogram [13].

The main concern was to combine the color and the texture feature extraction in a single algorithm to make it less time consuming.

The paper proposed Color Difference Histogram which is used to describe image features for image retrieval. This histogram is entirely different from existing histograms, and the proposed color difference histogram uses the perceptually uniform color difference as the histogram values. However, most of the existing histogram techniques merely focus on the frequency or number of pixels, which are used as the histogram values. In the proposed algorithm, orientation and perceptual color information have been combined in the unified framework, and both of their spatial layouts have been considered. The same algorithm can be extended to extract features of the images.

A. Color Difference Histogram (CDH)

When talking about color images the perceptually uniform color difference between colors and edge orientations cover on rich variety of visual information, it is very useful information and plays an important role in image content analysis and understanding. Figure 2 shows the process of plotting the CDH.

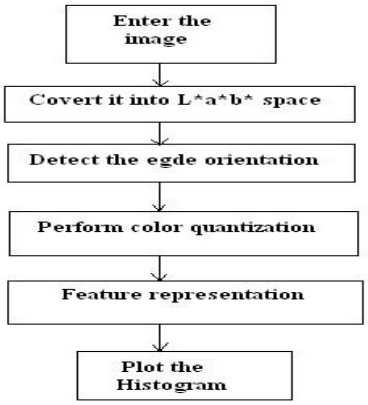

Figure 2 Process of Plotting CDH

The process involves converting a RGB color image or snap shot taken from the video to convert into L*a*b* color space [14]. This can be done using MATLAB image processing toolbox easily. It converts the image according to illuminanat D65 [15].Then we calculate the gradients of the image vectors. In order to extract chromatic information the process of quantization is carried out. The aim of color quantization is to pick and allot a limited set of colors for representing a given color image. The L*a*b* space used is quantized into 90 color which are perceptually uniform to human perceptual system.

The color features (H_{color}) and the other is for the edge orientations (H_{ori}).

$H_{color} = \sum\sum((\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2)^{½}$

H_{color} can represent the perceptually uniform color difference between neighboring edge orientations using color index information as a constraint, leading to a 90-dimensional vector.

$H_{ori} = \sum\sum((\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2)^{½}$

H_{ori} can represent the perceptually uniform color difference between neighboring color indexes with edge orientation information as a constraint, leading to an 18-dimensional vector; in total, a 90+18=108 dimensional vector is obtained for the final image features during image retrieval [16, 17].

The paper used improved version of Canberra distance as the similarity measure [18]. For each group representative image a 108-dimensional vector is calculated as $R = [R_1, R_2,\ldots R_{108}]$ and same 108-dimensional vector is calculated for the browsed query image as $Q = [Q_1, Q_2,\ldots Q_{108}]$. The distance is given as:

$D(R,Q) = \sum^{108}_{i=1}(|R_i - Q_i|)/(|R_i + f| + |Q_i + s|)$

Where, $f = \sum^{108}_{i=1}(R_i/108)$ and

$s = \sum^{108}_{i=1}(Q_i/108)$

IV. EXPERIMENTAL RESULTS

A. Datasets

Total 8 videos of less than 10 second duration are used for snap shot extraction and retrieval purposes.

B. Numerical Results

TABLE I 90-Dimensional color vector

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0.1033	0.1205	0.0432	0.1618	0.0861	0.6741	0.1874	0.2302
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0.0564	0.0042	0.1321	0.1838	0.3133	0.5434	1.3750	2.6590	1.5398	0.8500
0.0440	0.1367	0	0	0	0	0	0	0	0

TABLE II 18-Dimensional Orientation vector

0.5149	0.5562	0.6860	1.1685	1.8618	0.8200	0.6415	0.4146	0.4076
0.3583	0.4008	0.5173	0.8458	1.0740	0.4654	0.4013	0.3735	0.5747

TABLE III Canberra Distance

580.1613	579.9986	581.8208	580.0653	581.6373	580.9944	580.0153	580.1521

TABLE IV RESULTS

Result	
Minimum Value	579.9985
Index	2
Video Selected	2

C. Snap shots

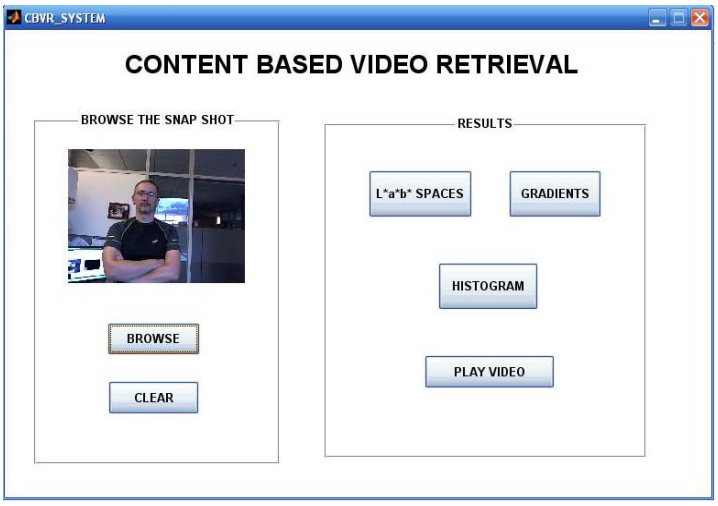

Figure 3 Browsing a query snap shot

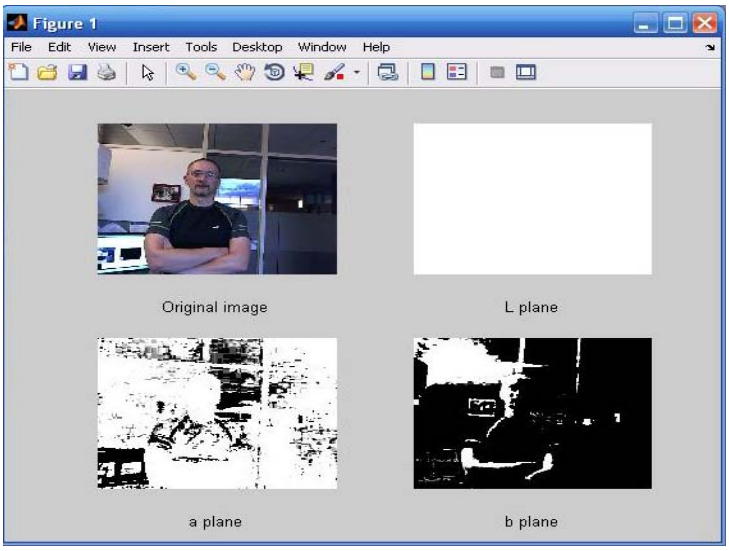

Figure 4 L*a*b* spaces

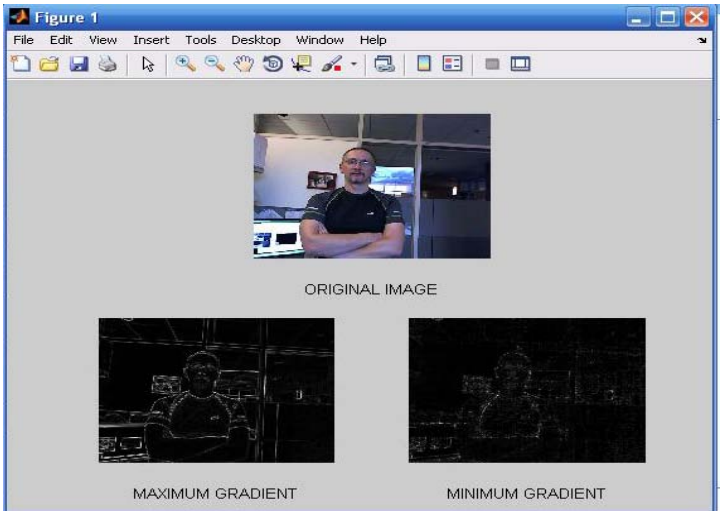

Figure 5 Minimum and Maximum Gradients

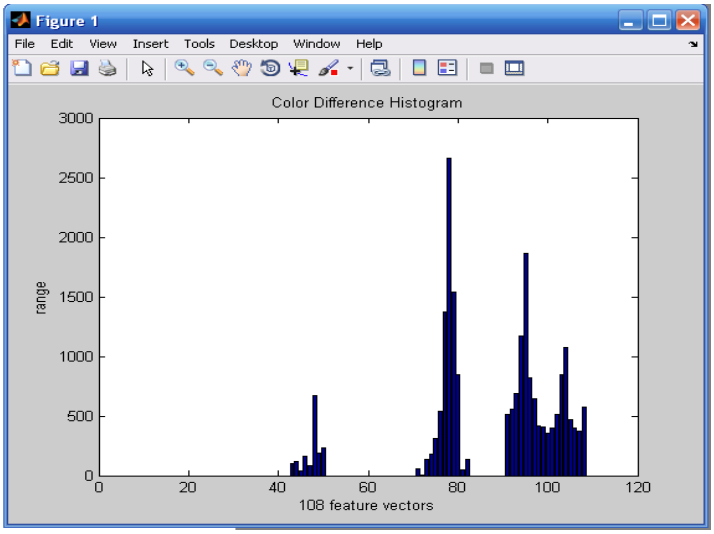

Figure 6 Color Difference Histogram

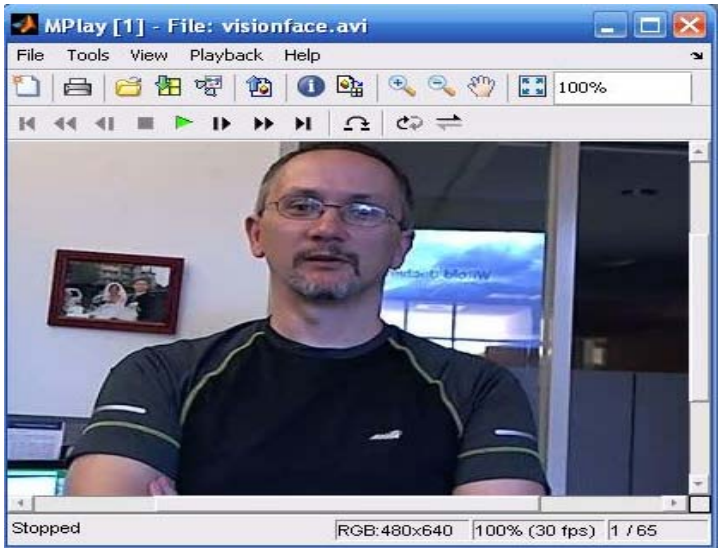

Figure 7 Searched Video clip

V. CONCLUSION

Apart from the issue of getting searched results another concern of large scale multimedia retrieval is development of an efficient algorithm for the retrieval system. Video retrieval problem is ultimately reduced to image retrieval after key frame selection. Most image retrieval techniques uses separate algorithms for different features. The proposed method reduces both the programming effort by reducing the number of algorithms for different discriminating features and similarity measures and the computation time by retrieving images from the group rather than comparing individual images of the database. The proposed method considers both the color and the spatial layout of the images to a significant extent by using as 108-dimesional vector.

REFERENCES

[1] Sharp. J-SH07. http://www.j-phone.com/f e/j/products/jsh07/ back.html and http://www.sharp.co.jp/products/jsh07/ index.html
[2] Farag, W. & Abdel-Wahab, H.Video Content-Based Retrieval Techniques,Idea Group Inc,2004,. pp.115-124.
[3] K. Chang and A. Hsu, “Image information systems: where do we go from here?,” IEEE transactions on Knowledge and Data Engineering, vol. 4, 1992, pp. 431–44
[4] Gao, X. and X. Tang, 2002. Unsupervised video shot segmentation and model-free anchorperson detection for news video story parsing. IEEE Trans. Circuits Syst. Video Technol., 12: 765-776.
[5] W. Burger, M.J. Burge, Principles of Digital image processing: Core Algorithms, Springer, 2009.
[6] B.S. Manjunathi, W.Y. Ma, Texture features for browsing and retrieval of image data, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (8) (1996) 837–842.
[7] Flickner, M., Sawhney, H. and et al. Query by Image and Video content: The QBIC system. In *IEEE Computer*, Vol.28, No.9, pp. 23-32, September 1995.
[8] P. Aigrain, H. Zhang, D. Petkovic, Contentbased Representation and Retrieval of Visual Media: A State-of-the-Art Review, *Multimedia Tools and Applications*, Kluwer Academic Publishers, 3(3),1996, 179-202.
[9] A. M. Ferman, A. M. Tekalp, and R. Mehrotra, “Robust color histogram descriptors for video segment retrieval and identification,” IEEE Transactions on Image Processing, Vol. 11, No. 5, pp 497-508, 2002.
[10] Steven C.H. Hoi and Michael R. Lyu, “A Multimodal And Multilevel Ranking Framework For Content-Based Video Retrieval”, 2007 International conference on Acoustics, speech, and Signal processing, Hawaii, USA, 15-20 April 2007.
[11] Paul Bourke, “Correlation”, from http://local.wasp.uwa.edu.au/~pbourke/miscellaneous/correlate/,1996.
[12] N. Chauhan and M. Goyani, “Enhanced Multistage Content Based Image Retrieval,” IJCSMC, vol. 2, Issue. 5, 2013, pp. 175–179.
[13] G.-H. Liu, L. Zhang, et al., Image retrieval based on multi-texton histogram, Pattern Recognition 43 (7) (2010) 2380–2389.
[14] W. Burger and M.J. Burge, Principles of Digital image processing: Core Algorithms, Springer, 2009.
[15] R.C. Gonzalez and R.E. Woods, Digital Image Processing, 3rd edition, Prentice- Hall, 2007.
[16] Y. Rubner, J. Puzicha, C. Tomasi, J.M. Buhmann, Empirical evaluation of dissimilarity measures for color and texture, Computer Vision and Image Understanding 84 (1) (2001) 25–43.
[17] Bakker, E., Lew, M.: Semantic video retrieval using audio analysis. In: International Conference on Image and Video Retrieval, Lecture Notes in Computer Science, vol. 2383, Springer (2002) 260–267
[18] Robles, Oscar D., et al. "Towards a content-based video retrieval system using wavelet-based signatures." 7th IASTED International Conference on Computer Graphics and Imaging-CGIM. 2004.