Feature Extraction in Content based Video Retrieval

Prof. Priyadarshani A. Bandgar
Head, Department of Information Technology Engineering
JSPM's BSIOTR(W), Pune, Maharashtra, India

Pooja Naik
Department of Information Technology Engineering
JSPM's BSIOTR(W), Pune, Maharashtra, India

Reshma Khaire
Department of Information Technology Engineering
JSPM's BSIOTR(W), Pune, Maharashtra, India

Abstract- Video retrieval is a wide spectrum of promising applications, motivating the researchers from worldwide. This paper provides an overview of the content-based video retrieval and focus on methods used for the key frame extraction in content based video retrieval. It contains video segmentation, feature extraction and similar video retrieval from the video database. Here, we are using three methods for features extraction. Finally, we analyze which is the best method for the feature extraction of videos from video database or from the internet.

Keywords- CBVR (content based video retrieval), Average RGB, Co-occurrence, Geometric Moment.

I. Introduction

A number of video features are based on color, texture and shape attributes. Key frame extraction is one of the important method which focused on the video structure analysis. In this method static key frame features, object feature and motion framework included. For the video data mining we used the feature extraction. For the classification purpose we extract the features among many possible features of video. In this paper color video classification is done on feature extracted from average RGB values of color component and it having better efficiency and insensitivity. Then on the basis of shape features are extracted with the help of geometric moment method. Then texture features are extracted based on co-occurrence method. The features are extracted on the basis of above three methods.

Video feature extraction and retrieval has attracted much attention in recent years. Feature extraction is fundamental in Video retrieval systems. A new feature extraction method based on co-occurrence matrix for texture features, average RGB for color features and geometric moment for shape features. Retrieval of Multimedia data is an important and promising research field. Content based video retrieval is becoming a highly recommended trend in many video retrieval systems. Key frame extraction is the first and foremost step in the process of video retrieval. Then the second step is feature extraction. It is fundamental to any kind of video analysis and video application since it enables segmentation of a video into its basic components. In order to extract the important key frames from a video, we need to segment it first, usually into shots, and then analyze which will be the most representative frame in the set of frames that compose the detected shot.

Histograms have been widely used by many video summarization techniques, several models such as have used histograms as a visual descriptor. The main reason for using histograms in video summarization is that it provides significant information about a frame and it is not computationally expensive. That is reason we use histograms in our model. Applying histograms to our problem will provide us the visual information describing the color distribution in the frames. We can use this information to group similar frames and extract possible key frames. In the concept of histogram evolution is used for summarizing the video.

This paper is organized as follows. Section2 contains Related work. Section3 Key frame extraction. Section4 Methodology, Section 5 summarizes and concludes the paper.

II. RELATED WORK

Color video classification is done on features extracted from histograms of color components. The benefit of using color video histograms are efficiency and insensitivity to small changes in camera view-point i.e. translation and rotation. As a case study for validation purpose, experimental trials were done on a database of about 100 videos divided into four different classes has been reported and compared on histogram features for RGB.

Several methods for retrieving images on the basis of color similarity have been described in the literature, but most are variations on the same basic idea. Each image added to the collection is analyzed to compute a color histogram which shows the proportion of pixels of each color within the image. The color histogram for each image is then stored in the database. At search time, the user can either specify the desired proportion of each color (75% olive green and 25% red, for example), or submit an example image from which a color histogram is calculated.

III. FEATURE EXTRACTION

To extract features according to video structural analysis results is the based on static features on key frames. These key frames are used for feature extraction. We focus on the visual features suitable for video retrieval. These mainly include features of key frames, objects, and motions. Key frames include retrieval of static features like color, texture and shape.

- *Static Features of Feature Extraction:*

In this we extract feature of videos based on Color, Texture, Shape .The key frame is used for feature extraction. The key frames of a video reflect the characteristics of the video to some extent. Traditional image retrieval techniques can be applied to key frames to achieve video retrieval. The static key frame features useful for video indexing and retrieval are mainly classified as color-based, texture-based, and shape-based.

1. Color-Based Features:

Histograms are widely used in computer vision. They can describe the color features when applied to a video frame. In this approach, we use a RGB (Red, Green and Blue) histogram of a frame, taking into consideration that the videos used in our experiments are colored. The original model uses grayscale histograms when RGB values are not present.

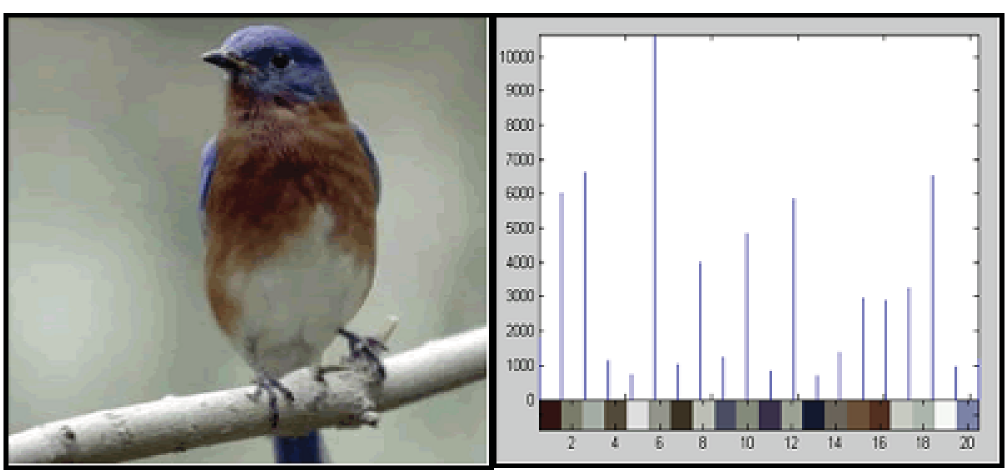

Fig: Color Histogram

We use average RGB to calculate color similarity. Average RGB is to compute the average value in R, G, and B channel of each pixel in an image, and use this as a descriptor of an image for comparison purpose. Gridded color distance is the sum of the color distances in each of the corresponding grid squares.

The Average RGB color histogram provides distribution information of colors for a given video frame. Let us consider that the values of each color channel goes from 0 to 255, making a total of 256 values and that for a given color frame each pixel would contain a combination of these three color channels. Therefore a RGB histogram for a frame should be represented by a structure of size 256 * 256 * 256, i.e., the RGB cube. A minute of video usually has more than 400 frames, then it would be computationally expensive to operate such a structure for each frame. In order to reduce this complexity, we are removing the redundant frames and uses only key frames.

For each of the images classified as White a pixel average was generated for each of the RGB coordinates. With the average values in hand, the minimum, average and maximum values were calculated. The same procedure was carried out for the brown and black tones. These values were then used to put together the fuzzy sets from each of the RGB coordinates for each of the classified tones.[2] One can consider that the universes of discourse for this system are the values that belong to the interval [0, 255], and which represent the pixel values in RGB, and the quantity of points generated is 255.

2. Texture-Based Features:

Texture is that property of surfaces that describes visual patterns. Co occurrence Matrix method is used for Texture-Based Features extraction. Texture represented by pixels gives relative brightness of consecutive pixels and finds the degree of contrast, regularity, coarseness and directionality which classifies textures as 'smooth', 'rough' etc. Texture is a visual pattern where there are a large number of visible elements densely and evenly arranged.

A texture element is a uniform intensity region of simple shape which is repeated. We divide the image into 55 blocks and compute texture features using Gabor-wavelet filters in each block. The merit of texture-based features is that they can be effectively applied to applications in which texture information is salient in videos. However, these features are unavailable in non texture video images. Depending on the texture on the foreground and background regions in a given video, we will get the trajectories from both parts from input video and from database videos.

Many statistical texture features are based on co-occurrence matrices representing second-order statistics of grey levels in pixel pairs in an image. The matrices are sufficient statistics of a Markov/Gibbs random field with multiple pair wise pixel interactions. A co-occurrence matrix shows how frequent is every particular pair of grey levels in the pixel pairs, separated by a certain distance *d* along a certain direction *a*.

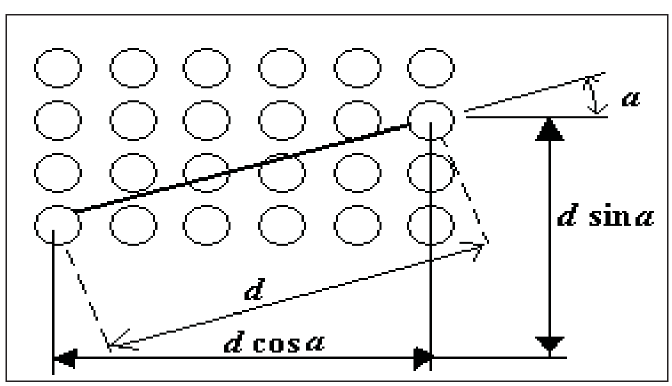

Fig: Co-occurrence method

The co-occurrence matrix to make it possible to determine the frequency of appearance of a formed "reason" for two pixels separated by a certain distance D in a direction (theta) particular was compared to the horizontal one. This definition enables us to choose the parameters of matrices of co-occurrence which give a better percentage of classifications. In this study, the percentage of classification for each texture for various values of d =1 and (theta) of (0°, 45° 90°, 135°) was determined[4][5].

3. *Shape-Based Features:*

Shape goes one step further than color and texture. It requires identification of regions to compare. There have been many shape similarity measures suggested for pattern recognition that can be used to construct shape distance measures. Shape is the characteristic surface configuration that outlines an object giving it a definite distinctive form.

Shape descriptions are an important task in content-based video retrieval. It is a mapping that converts the shape space into a vector space and satisfies the requirement that two similar shapes will also have close-to-identical shape descriptors. The two-dimensional moment (2-D frame) of 2-D object R is defined as:

$$M_{pq} = \sum \sum_{R} x^p y^q f(x, y)$$

where f (x,y) is the characteristic function describing the intensity of R, and p+q is the order of the moment.

The problem of moment-based shape orientation and symmetry classification is jointly considered. A generalization and modification of current state-of-the-art geometric moment-based functions is introduced. The proposed approach removes the requirement for accurate shape centroid estimation, which is the main limitation of moment-based methods, operating in the image spatial domain. The proposed framework demonstrated improved performance, compared to state-of-the-art methods [1]

Moment functions computed from an image have been extensively used as invariant feature descriptors in applications involving pattern recognition, image classification and template matching. They have also been used to extract the primary geometrical attributes of an image shape for 2D pose estimation. Geometric moments evaluated at high orders tend to become numerically unstable due to the monomial structure of the kernel functions. Image coordinate values are usually normalized to a value less than 1 by dividing by the image size, in order to eliminate this problem.[3]

IV. ALGORITHMS USED

1. Euclidean Distance Algorithm:

Euclidean Distance Algorithm is used to measure the distance between two pixels in multi-dimensional space between two frames.

$$D_i^{\,2}(x, y) = \sum (x_i - y_i)^2$$

The Advantage of Pixel Difference is Simple, Easy to Implement. The Euclidean distance mapping algorithm uses mapping of distances between two frames the map labels each pixel of the image with the distance to the nearest obstacle pixel. A most common type of obstacle pixel is a boundary pixel in a binary image. Here using this algorithm we are going to calculate the distance between two consecutive frames. The frames from the input video and the videos from the databases are extracted and distances between the frames are mapped to find out the videos for required input.

IV.CONCLUSION

We have presented a review on recent developments in visual content-based video retrieval. This paper focused on the static features color, texture and shape using methods: average RGB, co-occurrence and geometric moment respectively. Using these three methods we extract the static features from query video of frames and compared it with database video of frames. At the end of this survey, we have discussed future directions such as affective computing-based video retrieval and distributed network video retrieval.

REFERENCES

[1] A unifying approach to moment-based shape orientation and symmetry classification Georgios Tzimiropoulos, Nikolaos Mitianoudis and Tania Stathaki. IEEE TRANSACTIONS ON IMAGE PROCESSING 1

[2] Fuzzy Classification of Human Skin Color in Color Images A. G. Boaventura, V. M. Volpe, I. N. da Silva, A. Gonzaga

[3] FAST COMPUTATION OF GEOMETRIC MOMENTS AND INVARIANTS USING SCHLICK'S APPROXIMATION R. MUKUNDAN Department of Computer Science and Software Engineering University of Canterbury, Christchurch, New Zealand mukundan@canterbury.ac.nz International Journal of Pattern Recognition and Artificial Intelligence Vol. 22, No. 7 (2008) 1363–1377 c_ World Scientific Publishing Company

[4] IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 11, NO. 1, JANUARY 2014 Retrieval of Forest Stand Age From SAR Image Texture for Varying Distance and Orientation Values of the Gray Level Co-Occurrence Matrix Isabelle Champion, Christian Germain, Jean Pierre Da Costa, Arnaud Alborini, and Pascale Dubois-

[5] A Multiresolution Approach for Noised Texture Classification based on the Co-occurrence Matrix and First Order Statistics M. Ben Othmen, M. Sayadi, and F. Fnaiech World Academy of Science, Engineering and Technology 39 2008.

[6] IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 41, NO. 6, NOVEMBER 2011 797 "A Survey on Visual Content-Based Video Indexing and Retrieval" Weiming Hu*, Senior Member, IEEE*, Nianhua Xie, Li Li, Xianglin Zeng, and Stephen Maybank.