RESEARCH ARTICLE OPEN ACCESS

Content based Image Retrieval from Forensic Image Databases

Swati A. Gulhane*, Dr. Ajay. A. Gurjar**
*(Department of Electronics and Telecommunication, Sant Gadge Baba Amravati University, Amravati-444606)
**(Professor , Department of Electronics and Telecommunication, Sant Gadge Baba Amravati University, Amravati-444606)

ABSTRACT
Due to the proliferation of video and image data in digital form, Content based Image Retrieval has become a prominent research topic. In forensic sciences, digital data have been widely used such as criminal images, fingerprints, scene images and so on. Therefore, the arrangement of such large image data becomes a big issue such as how to get an interested image fast. There is a great need for developing an efficient technique for finding the images. In order to find an image, image has to be represented with certain features. Color, texture and shape are three important visual features of an image. Searching for images using color, texture and shape features has attracted much attention. There are many content based image retrieval techniques in the literature. This paper gives the overview of different existing methods used for content based image retrieval and also suggests an efficient image retrieval method for digital image database of criminal photos, using dynamic dominant color, texture and shape features of an image which will give an effective retrieval result.

Keywords **–** Dominant Color, Forensic science , Image retrieval, Texture

I. INTRODUCTION

The necessity of efficiently querying generally available image data has improved with the increase in the availability of huge quantities of such data. Hence, content-based video data retrieval proves to be a challenging and crucial problem. Due to the proliferation of video and image data in digital form, Content-based image retrieval (CBIR) has become a prominent research topic. Therefore an important problem that needs to be addressed is fast retrieval of images from large databases. To find images that are perceptually similar to a query image, image retrieval systems attempt to search through a database.

Content-based Image Retrieval (CBIR), also known as query by image content (QBIC), is all about the searching of digital images in large databases. "Content based" means that the search will analyze the actual contents of the images. The term "content" in this context might refer to colors, shapes, textures etc. CBIR can greatly enhance the accuracy of the information being returned and is an important alternative and complement to traditional text-based image searching.

In this paper, we proposed the method of Content based Image Retrieval from large collection of forensic images.

II. OVERVIEW OF EXISTING METHODS

In [1], an efficient image retrieval technique which uses dominant color and texture features of an image is proposed. The proposed method yielded higher average precision and average recall with reduced feature vector dimension. In [2], a new and effective color image retrieval scheme for combining all the three i.e. color, texture and shape information, which achieved higher retrieval efficiency. The color and texture features are extracted using color quantization and steerable filter decomposition whereas shape feature is extracted using pseudo Zernike moment. Trademark image retrieval (TIR) system is proposed in [3] to deal with the vast number of trademark images in the trademark registration system. The proposed approach commences with the extraction of edges using the Canny edge detector, performs a shape normalization procedure, and then extracts the global and local features. In [4], a further exploration and study of visual feature extraction is done. A new image retrieval method based on HSV color and GLCM texture features of image sub-blocks with one to one matching is proposed. The method based on color and texture features of image sub-blocks has better retrieval performance compared with the Image retrieval system using only HSV color, only GLCM texture and combined HSV color and GLCM texture.

An image retrieval system is presented in [5], which used HSV color space and wavelet transform approach for feature extraction. A comprehensive survey, highlighting current progress, emerging directions, the spawning of new fields, and methods for evaluation relevant to the field of image retrieval is presented in [6]. It consider that the field will experience a paradigm shift in the foreseeable future, with the focus being more on application-oriented, domain-specific work, generating considerable

impact in day-to-day life. Dominant color descriptor (DCD) is one of the color descriptors proposed by MPEG-7 in [7] that has been extensively used for image retrieval. A content-based image retrieval method based on an efficient combination of multi resolution color and texture features is proposed in [8]. In [9], a detailed evaluation of the use of texture features in a query-by-example approach to image retrieval is presented. Image retrieval mechanism is explored in [10], based on combination of color and texture features.

The fundamental techniques for content based image retrieval are:

- Spatial Analysis
- Temporal Analysis

Spatial Analysis:
Spatial Analysis uses parameters like color, shape and texture.
Color:

Color is the most extensively used visual content for video retrieval. Its three dimensional values makes its discrimination potentiality superior to the single dimensional gray values of videos. Retrieving the videos using color is done by computing a color histogram which identifies the proportion of pixels within an image holding specific values.

Shape:

The usage of shape features for video retrieval is restricted because robust and accurate video segmentation is difficult to achieve. Shape, moreover, does not refer to the shape of a video but to the shape of a particular region that is being sought out. Shapes can be determined by applying segmentation or edge detection techniques.

Texture:

Texture representations can be classified as structural and statistical. There are the visual patterns in the videos and how they are spatially defined. Texture represented by texels gives the relative brightness of consecutive pixels and finds the degree of contrast, regularity, coarseness and directionality which classifies textures as 'smooth', 'rough' etc.

Temporal Analysis:

Temporal Analysis is done based on parameters like motion and audio.

Motion:

Motion is an intangible factor which is used for querying. Though tapping the motion vectors for the videos is a task which is why still motion based querying is in its infancy, once this concept is perfected it will help in video retrieval to the best accuracy level and is more dependable than visual and textual query.

Audio:

Unlike popular belief, from content-based video retrieval point of view, the audio information can be even more important than the visual part since it is mostly unique and significantly stable within the entire duration of the content. However, audio-based studies lag far behind the visual counterpart and the development of robust and generic systems for audio content management is still in its infancy.

Color, texture and shape features have been used for describing image content. Color is one of the most widely used low-level visual features and is invariant to image size and orientation. As conventional color features used in CBIR, there are color histogram, color correlogram, and dominant color descriptor (DCD). Color histogram is the most commonly used color presentation, but it does not include any spatial information. Color correlogram describes the probability of finding color pairs at a fixed pixel distance and provides spatial information. Therefore color correlogram yields better retrieval accuracy in comparison to color histogram. Color autocorrelogram is a subset of color correlogram, which captures the spatial correlation between identical colors only. Since it provides significant computational benefits over color correlogram, it is more suitable for image retrieval. Texture is also an important visual feature that refers to innate surface properties of an object and their relationship to the surrounding environment. Many objects in an image can be distinguished solely by their textures without any other information. Texture may consist of some basic primitives, and may also describe the structural arrangement of a region and the relationship of the surrounding regions. Although shape, color and texture are undoubtedly important visual features for image representation, there is still little understanding of how best to implement these attributes for image retrieval. In our approach we have used the texture features using gray-level co-occurrence matrix (GLCM). Shape feature has been extensively used for retrieval systems. Shape signatures are computed from blurred images and global invariant moments are computed as shape features.

III. PROPOSED METHOD

To describe image from the different aspects for more detailed information in order to obtain better search results and to express more image information, we consider the dominant color, texture and shape features combined. The proposed method is based on dominant color, texture and shape features of image.

A. Retrieval Algorithm

Step 1: Uniformly divide each image in the database and the target image into 8-coarse partitions

Step 2: For each partition, the centroid of each partition is selected as its dominant color.
Step 3: Obtain texture features (Energy, Contrast, Entropy and inverse difference) from GLCM.
Step 4: Obtain invariant moments of Gradient Vector Flow Fields as shape features
Step 5: Construct a combined feature vector for color, texture and shape.
Step 6: Find the distances between feature vector of query image and the feature vectors of target images using weighted and normalized Euclidean distance.
Step 7: Sort the Euclidean distances.
Step 8: Retrieve most similar images with minimum distance.

1. Color Feature Representation

In general, color is one of the most dominant and distinguishable low-level visual features in describing image. To retrieve images, such as QBIC system and Visual SEEK many CBIR systems employ color. DCD contains two main components: representative colors and the percentage of each color. DCD describe the color distribution in an image or a region of interesting and can provide an effective, compact, and intuitive salient color representation.

2. Extraction of dominant color of an image

The procedure to extract dominant color of an image is as follows:

The selection of color space is not a critical issue for DCD extraction. Therefore, the RGB color space is used for simplicity and without loss of generality. The RGB color space is uniformly divided into 8 coarse partitions. If there are several colors located on the same partitioned block, they are assumed to be similar. After the above coarse partition, the centroid of each partition is selected as its quantized color.

3. Extraction of texture of an image

A texture representation for image retrieval based on GLCM is used. Texture features are extracted from the statistics of this matrix.

4. Extraction of shape of an image

Shape information is captured in terms of the edge image of the gray scale equivalent of every image in the database. We have used gradient vector flow (GVF) fields to obtain the edge image. Gradient vector flow (GVF) is a static external force used in active contour method.

IV. OBJECTIVES OF PROPOSED METHOD

Research and development issues in CBIR cover a range of topics, many shared with mainstream image processing and information retrieval. Some of the most important objective can be:

- understanding image user's needs and information-seeking behaviour.
- to develop a technique which captures color and texture descriptors of an image, and has a shape descriptor in terms of invariant moments computed on the edge image.
- identification of suitable ways of describing image content.
- extracting such features from raw images.
- providing compact storage for large image databases.
- matching query and stored images in a way that reflects human similarity judgments.
- efficiently accessing stored images by content.
- providing usable human interfaces to CBIR systems.

Tool used: MATLAB, Data Source: Bunch of criminal Images.

V. METHODOLOGY

We will implement an efficient image retrieval technique which uses dynamic dominant color, texture and shape features of an image. As a first step, an image is uniformly divided into 8 coarse partitions as shown in Fig. 1

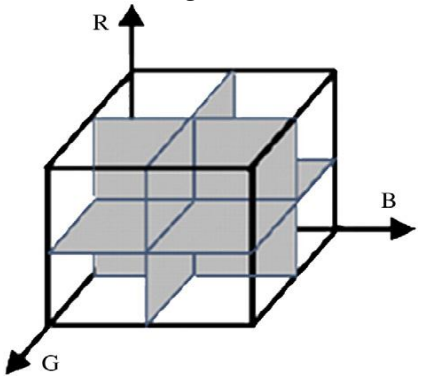

Figure. 1. The coarse division of RGB color space.

The centroid of each partition is selected as its dominant color after the above coarse partition. By using Gray Level Co-occurrence Matrix (GLCM), texture of an image is obtained. Color and texture features are normalized. Using Gradient Vector Flow fields, shape information is captured in terms of edge images computed. To record the shape features, invariant moments are then used A robust feature set for image retrieval is provided by using the combination of the color and texture features of an image in conjunction with the shape features. In retrieving the similar images of video, weighted Euclidean distance of color, texture and shape features is used.

The features including texture, color and shape are extracted from the database images and it is stored in the feature library. For retrieving, the query image is given to the proposed system that extracts the features and the features are matched with the

feature library using the designed image matching. The matching score computed is used to retrieve the images from the dataset and the retrieved images for the corresponding input query image is given in the following Fig 2.

(a)

(b)

Figure 2: a) Query image b) Retrieved images from the large database of images which are similar to the query image.

VI. BLOCK DIAGRAM OF PROPOSED METHOD

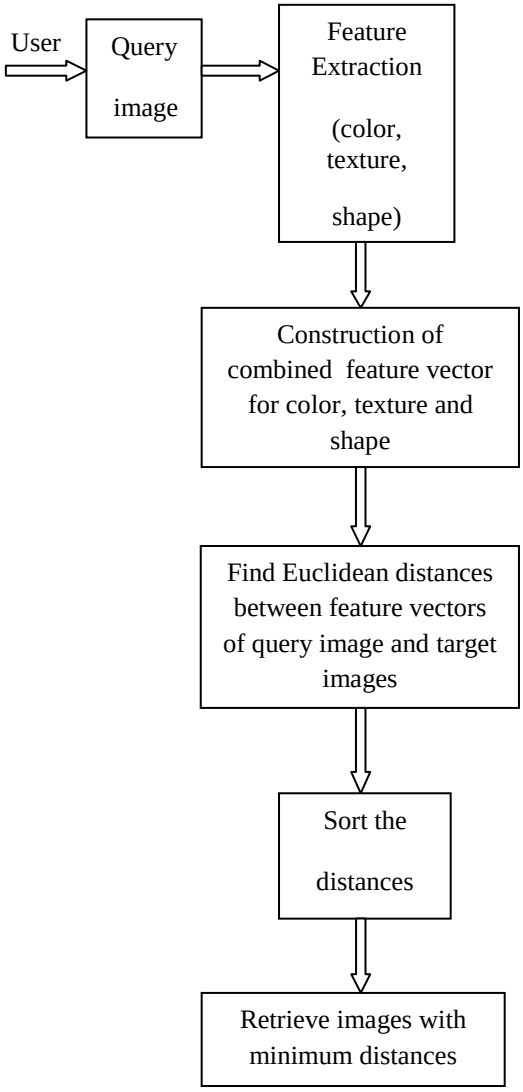

Figure 3: Block Diagram of Proposed Method

VII. CONCLUSION

In the beginning, we explained that image retrieval has become the need of the today's world. Next this paper has given the overview of different existing methods used for content based image retrieval. Further we discussed the proposed method for content based image retrieval for forensic application using dominant color texture and shape which gives the effective retrieval performance.

REFERENCES

[1]. M.Babu Rao, Dr. B.Prabhakara Rao, Dr. A.Govardhan, "*Content based image retrieval using Dominant color and Texture features*", *International Journal of Computer science and information security, Vol.9 issue No: 2, February 2011.pp:4146*

[2]. X-Y Wang et al., "*An effective image retrieval scheme using color, texture and shape features*", Comput. Stand. Interfaces (2010), doi:10.1016/j.csi.2010.03.004

[3]. Chia-Hung Wei, Yue Li, Wing-Yin Chau, Chang-Tsun Li, "*Trademark image retrieval using synthetic features for describing global shape and interior structure*", Pattern Recognition, 42 (3) (2009) 386–394.

[4]. FAN-HUI KONG, "*Image Retrieval using both color and texture features*", *proceedings of the 8th international conference on Machine learning and Cybernetics, Baoding, 12-15 July 2009.*

[5]. JI-QUAN MA, "*Content-Based Image Retrieval with HSV Color Space and Texture Features*", *proceedings of the 2009 International Conference on Web Information Systems and Mining.*

[6]. Ritendra Datta, Dhiraj Joshi, Jia Li, James Z. Wang, "*Image retrieval: ideas, influences, and trends of the new age*", ACM Computing Surveys, 40 (2) (2008) 160.

[7]. Nai-Chung Yang, Wei-Han Chang, Chung-Ming Kuo, Tsia-Hsing Li, "*A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval*", *Journal of Visual Communication and Image Representation, 19 (2) (2008) 92–105.*

[8]. Young Deok Chun, Nam Chul Kim, Ick Hoon Jang, "*Content-based image retrieval using multiresolution color and texture features*", IEEE Transactions on Multimedia, 10 (6) (2008) 1073–1084.

[9]. P. Howarth and S. Ruger, "*Robust texture features for still-image retrieval*", *IEEE Proceedings of Visual Image Signal Processing, Vol.152, No. 6, December 2005.*

[10]. S. Liapis, G. Tziritas, "*Color and texture image retrieval using chromaticity histograms and wavelet frames*", IEEE Transactions on Multimedia 6 (5) (2004) 676