International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 4 660 - 664

Content-based Image Retrieval using Color and Geometry

Jeebachh Jha
PG Student, Dept. of I.T.,
VESIT College, Mumbai University, Maharashtra,
India- 400074
jhajeebachh@gmail.com

Prof. Manoj Sabnis
Associate Professor, Dept. of I.T.,
VESIT College, Mumbai University, Maharashtra,
India- 400074
manojsab67@yahoo.co.in

***Abstract*—**The increased need of content based image retrieval technique can be found in a number of different domains such as Data Mining, Education, Medical Imaging, Crime Prevention, Weather forecasting, Remote Sensing and Management of Earth Resources. With the development of Multimedia data types and heavy increase in available bandwidth, there's a huge demand of Image Retrieval system Content based image retrieval system uses color and geometry means to store, retrieve, sort and print any combinations of the images. The retrieval of images is, for the majority of search engines, available for collecting data from the image, this can be an image file name, html tags and surrounding text. This left the actual image more or less ignored. CBIR uses methods that analyze the actual bits and pieces i.e. color, shape, texture and spatial layout. There have been different approaches such as feature extraction, indexing and retrieval process. One approach is to make an attempt to classify the image into a more textual described context. With the image classified, it can be retrieved using more traditional and better retrieval methods. Our system Content Based Image Retrieval which is based on color and geometry, the system exactly does feature extraction in first step by using color, texture and shape (geometry) on images which gives there features which can be used to classify the image into different groups using distance formulas. Also the system gives relevant images as well as irrelevant images. The project thus going to work on relevance feedback of user which helps to improve the overall results.

Keywords*-Content Based Image Retrieval, Color, Shape, Lab space, HSV.*

*****

I. INTRODUCTION

With the development in multimedia, computational power of processors and reduction in the price of memory causes need to switch from the historical approach to the new approach. Maintenance of huge databases is very crucial and especially when large collection of digital images need to be maintained. In many areas where digitized images are required like Academia, Multimedia, Journalism, Hospitals, GIS, Crime Prevention, Pattern Recognition, Statistics and many more. Thus over the decade the volume of digital image is increasing very exponentially. Thus several researchers are working on it to maintain the large volume of databases and to indexing the features of data. [1].

Generally the increased need of content based image retrieval technique can be found in a number of different domains such as Data Mining, Education, Medical Imaging, Crime Prevention, Weather forecasting, Remote Sensing and Management of Earth Resources. With the development of Multimedia data types and heavy increase in available bandwidth, there's a huge demand of Image Retrieval system [2].

In previous system, feature was extracted only from text documents but due to the development in media like images and videos, the feature extraction has become more valuable. For this Selection of extracted features plays an important role in Content Based Image Retrieval. These features are intended for selecting, indexing and ranking images according to their potential interest to the user.

Good feature selection also allows the time and space cost of the retrieval process to be reduced. Content based image retrieval system uses color and geometry means to store, retrieve, sort and print any combinations of the images. The retrieval of images is, for the majority of search engines, available for collecting data from the image, this can be an image file name, html tags and surrounding text. This left the actual image more or less ignored. CBIR uses methods that analyze the actual bits and pieces i.e. color, shape, texture and various other feature. There have been different approaches such as feature extraction, indexing and retrieval process. One approach is to make an attempt to classify the image into a more textual described context. With the image classified, it can be retrieved using more traditional and better retrieval methods.

A. Motivation

Initially there were text documents, there were systems for their style, retrieval, sorting and presenting. These operations on the due work done by QBT i.e. query by text method. In this it has to be found out which are the keywords in the text and these keywords were suggested to the uses for the Text retrieval.

Then due to the development in internet and representation visualizations techniques were introduced where the images and video could convey more data information as compared to Text. Now a new technique called as QBE query by extraction was introduced where images were search by images and not by text.

For this expose of image features were required in a large ways. The image features like intensity, color, texture, brightness, illumination conditions and geometrical relocation where examined. This gives rise to different CBIR system which worked on different parameters presently exhibited by the images. Motivated by this idea and the having challenge in getting the accurate parameters for better results, our work is concentrated on two of the most prominent features available in image i.e. Color and Geometry.

B. Objective

Objective of this system i.e. Image Retrieval Using color, geometry are:

- To develop an application that makes the retrieval of images from the image database easy and efficient Query Based extraction method.
- To extract the image features and get similar images from the databases of images.
- Wherever color features are not efficient, geometry features to be used as replacement to detect right image and vice versa.
- Also using relevance feedback of user system to generate the module which is more user friendly.

Content-Based Image Retrieval (CBIR) uses the visual contents of an image such as color, shape, texture, and spatial layout to represent the image and also for index the image. The block diagram (Fig 1) of a CBIR system can be understood as a basic building blocks that communicate with each other to retrieve the database images according to a given query. In normal content-based image retrieval system, the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. The feature database made by these feature vectors of the images. To retrieve images, users provides query images to the retrieval system. Figure 1 shows the proposed CBIR what we done.

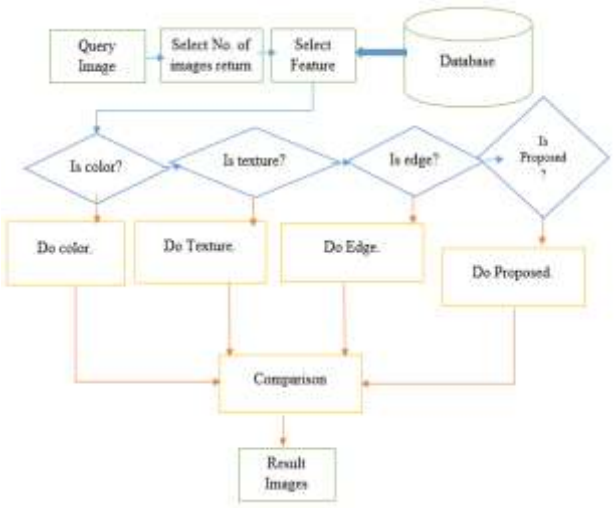

Figure 1. Block diagram of proposed content-based image retrieval system.

We noted some Characteristics of CBIR include like Image retrieval by image content, visually similar images to Query image, No keywords, Low level features like color, shape are used, etc.

The rest of the paper is as follows. Section 2 briefs the considered different schemes based on CBIR for the analysis. Section 3 represents the feature extraction techniques considering different features. Section 4 represents the proposed system. Section 5 gives all experiments and results about the proposed system and the paper concludes with section 5.

II. LITERATURE SURVEY

From the past decade, the search for highly effective and efficient techniques of CBIR is a dynamic focus of research. Various previous works related to Content Based Image Retrieval are as follows:

In literature [1], the author gives specified path to use these primitive features color and shape to retrieve the desired image. In this CBIR system uses color, texture and shape fused features to retrieve desired image from the large database and hence provides more efficiency or enhancement [1]. Some techniques have used global color and texture features whereas some have used local colors and texture features. In [2], the various useful feature extraction techniques are Principal Component Analysis (PCA), linear discriminant analysis (LDA), Independent Component analysis (ICA). In this work, the author analyses the performance of feature extraction techniques (PCA, LDA, and ICA) and classification techniques (SVM, NN) used in CBIR. In [3], author used generic Fourier shape descriptor technique. Contrasting to previous approaches, such as using a single color histogram for the whole image ,or local color histogram for a fixed number of image cells ,the one texture & shape features are analyzed. Gabor filter is used to extract texture features from images. Morphological c1ustring operation combined with Gabor filter gives better retrieval accuracy. After applying Gabor filter on the image, texture features such as mean and standard deviations are calculated. This forms the feature vector. Shape feature is extracted by using Fourier Descriptor and the centroid distance [3].

In [4], authorsanalyses large database, based on this a new content based image retrieval method using the feature analysis of edge extraction and median filtering is proposed. Aim at noise reduction while edge detail preservation. After feature extraction histogram comparison is applied for feature vector calculation and similarity measurement [4]. In [5], author gives the visual signature based on region was attracted more attention. To get the signature based on region, the crucial step is image segmentation, and reliable image segmentation is also critical to get the image shape description. In [6], presents a novel framework using color and shape features by extracting the different components of an image using the Lab and HSV color spaces to retrieve the edge features. The performances of the HSV and Lab color space approach have been compared with Gray and RGB approach. Accordingly the Lab color space approach gives better performance than RGB and HSV [6]. In [7], the paper state that Color is identified as low level feature. So, image retrieval using only color features often provide very unsatisfactory results because two retrieved image have similar color. As the solution of this problem this paper describes a novel algorithm for Content Based Image Retrieval (CBIR) based on Color Edge Detection and Discrete Wavelet Transform (DWT).

In [8], the author improve the retrieval performance, combined texture and shape features are utilized, because many features provide more information than the single feature. Then images are extracted based on their Euclidean distance [9]. The performance is evaluated using precision-recall graph. Gradient method for shape feature extraction and retrieval of similar image from image database is reported in literature [10]. The literature [11] proposed the gradient method for extraction of shape features from the image and to retrieve the similar image from the database. The comprehensive works of talwar [9], Bansal [15] provide some of the most influential surveys on the CBIR. The extensive work of Singh [13], Chaudhari[10], Daisy [3] also outstands to describe the functionality of CBIR systems. Pujari [7], Sarvanan [13] done a great work with color feature. Finally, the recent study of Xie [11], Gui [14] and Demir [16] gives an actual overview of the enhanced of CBIR

and tackle its major future challenges by implementing the relevance feedback using supervised learning algorithms.

III. FEATURE EXTRACTION WITH EXISTING FEATURES

As we mention in our survey paper [22], the basis of any content-based image retrieval technique is a visual feature extraction. In a bigger sense, features may include both visual features (color, shape, etc.) and text-based features (key-words, annotations). The visual feature scope can be further classified as high-level features and low-level features. One of the keys of a CBIR system is the selection of the features to represent an image. There does not exist a single best representation for any given visual feature because of the complex composition and perception subjectivityof visual data. Multiple approaches have been introduced for each of these visual features and each of them characterizes the feature from a different perspective [3].

3.1 Color

Color is a perception that depends on the response of the human visual system to light and the interaction of light with objects [1] [7]. It is the product of the surface spectral reflectance, illuminant and sensor sensitivity (i.e. of digital sensors or of cones in the human eye). Color is widely used visual features in content-based image retrieval among all the features. It is relatively robust to background complication.It is independent of image orientation and size. The key issues in color feature extraction include the color space, color quantization, and the choice of similarity function [3]. We have to first determine the color space to use, if we want to describe an image by its color features. There exist different space models such as RGB, HSV, CIE L*a*b*, etc. The best representation depends on the special needs of the application.

3.1.1 Methods of representation

Each feature may have several representations as color histograms [7], color moments [2], color coherence vectors [10] image color feature. Moreover, numerous variations of the color histogram itself have been proposed which differs in the selected color-quantization scheme as we studied. The mathematical operations of the pixel values represented in a certain color space are related to color descriptors. Some of the most popular descriptors are:

3.1.1.1 Color histogram

The main method of representing color information of images in CBIR systems is through color histograms [7]. A color histogram is a type of bar graph, where each bar represents a particular color of the color space being used. Statistically, a color histogram is a way to approximate the joint probability of the values of the three color channels. By splitting the range of the data into equally sized bins we get the most common form of the histogram. Then for each bin, the number the colors of the pixels in an image that fall into each bin are counted and normalized to total points, which gives us the probability of a pixel falling into that bin. It does not take into consideration the spatial information of pixels is one of the main drawbacks of the color histogram. If two images have similar color distributions, still very different images can be considered similar.

Steps for Color Histogram:
Step 1. Convert RGB image into HSV color space.
Step 2. Color quantization is carried out using color histogram.
Step 3.Divide the total number of pixels to get normalized histogram.
Step 4. Repeat step1 to step3 on database images.
Step 5. Calculate the similarity matrix of query image and the image present in the database.
Step 6. Repeat the steps from 4 to 5 for all the images.
Step 7. Retrieve the images.

3.2 Shape

Shape is a visual feature and it is one of the important primitive features for describe image content. It contains all the geometrical information of an object in the image. These information does not change generally change even when orientation or location of the object are changed.

3.2.1 Methods of representation

Two major steps in shape feature extraction: object segmentation and shape representation. Once objects are segmented, their shape features can be represented and indexed. In general, shape representations may divided into two types, boundary-based and region-based. The boundary-based uses only the outer boundary of the shape.The region-based uses the entire shape region.

The canny edge detection algorithm runs in 5 separate steps:
Step 1.Smoothing i.e. Blurring of the image to remove noise.
Step 2. Finding gradients i.e. the gradients of the image has large magnitudes, the edges should be marked.
Step 3. Non-maximum suppression: Only local maxima should be marked as edges and not global.
Step 4. Double thresholding i.e. Potential edges are determined by thresholding technique.
Step 5.Edge tracking by hysteresis i.e. Final edges aredetermined by suppressing all edges that are not connected to a very certain edge.

3.3 Texture

Texture is defined as structure of surfaces formed by repeating a particular element in relative spatial positions. Image textures are defined as images of natural textured surfaces.Also of artificially created visual patterns. It contains important information about the structural arrangement of the surface. It is a features that describes the relationship of the surface to the surrounding environment.as well as the distinctive physical composition of a surface.
Gabor wavelet is widely used to extract texture from the images for retrieval. It has been shown to be very efficient. Basically Gabor filters are a group of wavelets. Each wavelet capturing energy at a specific frequency and specific orientation.

Steps for formulation of Texture feature vector of an image:
Step 1.Select an image of size XYZ.

Step 2.Convert the query image from RGB to Gray Scale Image.
Step 3.Create 24 Gabor filter images.
Step 4.Convolute the image with 24 Gabor filters fmn (x,y)to form 24 Gabor filters images Gmn (x,y).
Step 5.Calculate Mean (μ) and Standard deviation (σ) for each Gabor filtered images.

IV. Feature Extraction with Proposed Method

The basis of any content-based image retrieval technique is a visual feature extraction. In a bigger sense, features may include both visual features (color, shape, etc.) and text-based features (key-words, annotations). But single feature is not enough to make accurate system with more feasibility. As some of the cases are fails with edge feature extraction as well as color is not effective for some cases. For that reason, we propose a system which gives feature extraction with respect to geometry. For this reason, we made use of more than one feature from we get a perfect combination of feature extraction technique. Below we mention the steps of our proposed system.

Steps for Proposed System:

Step 1. Convert RGB color space image into HSV color space and Gray Scale Image.

Step 2.Color quantization is carried out using color histogram by assigning 8 level each to hue, saturation and value to give a quantized HSV space with 8x8x8=512 histogram bins.

Step 3. The normalized histogram is obtained by dividing with the total number of pixels.

Step 4. Create 24 Gabor filter images (scale (m)-4 and orientation (n)-6)

Step 5. Convolute the image with 24 Gabor filters fmn (x,y)to form 24 Gabor filters images Gmn (x,y).

Step 6. Repeat step1 to step5 on an image in the database.

Step 7. Calculate the similarity matrix of query image and the image present in the database and Calculate Mean (μ) and Standard deviation (σ) for each Gabor filtered images.

Step 8. Repeat the steps from 6 to 7 for all the images in the database.

Step 9. Retrieve the images.

V.Experiment and Results

By looking at the results provided by proposed algorithm and the results obtained by study different feature extraction system compared against various factors and following points are worth noted. The retrieval performance of proposed technique is shown in terms of Precision and Recall. These two measures are given by equations (1) and (2) respectively,

$$\text{Precision} = \frac{\text{No.of relevant images retrived}}{\textit{Total no.of images retrived}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$\text{Recall} = \frac{\text{No.of relevant images retrived}}{\textit{Total no.of relevant images}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

TABLE I. Comparison between Edge, Color, Texture & Geometrybased on Number of images return correctly

No. of images returned	Edge	Color	Texture	Geometry
1-3	YES	YES	YES	YES
4-6	NO	YES	YES	YES
7-9	NO	NO	YES	YES
10-12	NO	NO	NO	YES
13-15	NO	NO	NO	YES

Table 1 shows the Comparison between Edge, Color, Texture & Geometry based on Number of images return correctly. Number of features increases the accuracy of the system. The requester has been provided requested images by our proposed system on the basis of different features.

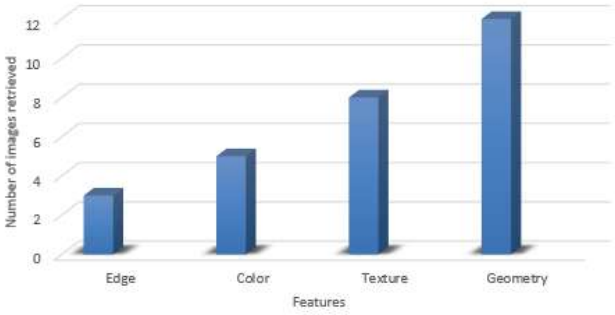

Figure 2. Edge vs. Color vs. Texture vs. Geometry

Figure 3 shows difference between Edge vs. Color vs. Texture vs. Geometry with respect to Number of images retrieved.

TABLE II. Comparison between Edge, Color, Texture and Geometry using Image Clusters

Image Cluster	Edge	Color	Texture	Geometry
Buses	20%	40%	40%	80%
Cats	20%	40%	40%	60%
Flowers	40%	60%	60%	80%
Dogs	20%	40%	40%	60%
Dinosaurs	40%	40%	60%	80%
Foxes	20%	40%	40%	60%
Tiger	20%	40%	60%	60%
Buildings	20%	40%	40%	60%

From the Table 2, we show the comparison between Edge, Color, Texture&Geometry using Image Clusters. First we consider the image clusters like Buses, Cats, Flowers, Dogs, Dinosaurs, Foxes and Tiger. There are many feature extraction type which is covered in our proposed system.

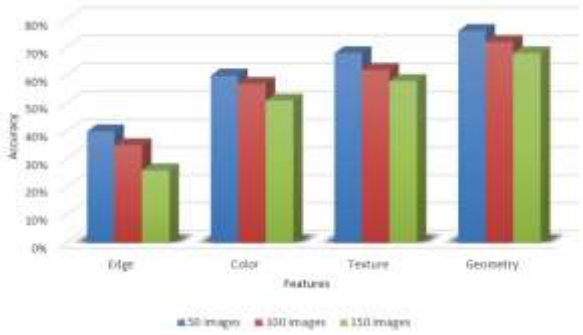

Figure 3. Comparison between Edge, Color, Texture, and Geometry based on number of images process

Figure 4 shows the comparison between Edge, Color, Texture, Geometry based on number of images process. This shows clearly that the proposed system make difference with efficiency compare to existing system.

VI. Conclusion and Future Work

CBIR Using color and geometry is to extract the image features and get similar images from the databases of images. Our system Content Based Image Retrieval using color and geometry, the system exactly do feature extraction in first step by using texture, color and shape(geometry) on images which gives features which can be used to classify the image into different groups using distance formulas. But our proposed algorithm helps to improve the results of the existing system with huge margin.We can make a better system by combining different features to retrieve accurate image as an output. In our work, it is clearly shown that when we use combined two feature then retrieval performance is improved. In future relevance feedback also be our aim of research.

Acknowledgment

JeebachhJha is pursuing M.E (I.T) from V.E.S. Institute of Technology, Mumbai University. He did his graduation B.E (IT) from Mumbai University, Maharashtra. He iscurrently working on Content based image retrieval with relevance feedback.

References

[1] Ms. Apurva N. Ganar, Prof. Sachin M. Jambhulkar , "Enhancement of image retrieval by using colour,texture and shape *features'' International Conference on Electronic Systems, Signal Processing and Computing Technologies*, 2014.

[2] D.Jeyabharathi, Dr.A. Suruliandi, "Performance Analysis of Feature Extraction and Classification Techniques in CBIR" *International Conference on Circuits, Power and Computing Technologies* ,2013.

[3] M Mary Helta Daisy, Dr.S. TamilSelvi " Combined texture and Shape Features for Content Based Image Retrieval" *International Conference on Circuits, Power and Computing Technologies*,2013.

[4] Hui Zhao,Pankoo Kim and Jongan Park, "Feature analysis based on Edge Extraction and Median Filtering for CBIR" *11th International Conference on Computer Modeling and Simulation,UKSim* 2009.

[5] Dewen Zhuang, Shoujue Wang, "Content-based image retrieval based on integrating region segmentation and relevance feedback" IEEE 2010.

[6] JagadeeshPujari ,Pushpalatha S.N, " Content-Based Image Retrieval using Color and Shape Descriptors" IEEE 2010.

[7] Swati Agarwal,A. K. Verma, "Content Based Image Retrieval using Color Edge Detection and Discrete Wavelet Transform", IEEE 2014.

[8] ManimalaSingha* and K.Hemachandran "Content Based Image Retrieval usingColor and Texture", *Signal& Image Processing: An International Journal (SIPIJ)* Vol.3, No.1, February 2012

[9] Hui Xie, Ying Ji, Yueming Lu,"An Analogy-Relevance Feedback CBIRMethod Using Multiple Features" IEEE 2010.

[10] Ajay Kumar Bansal, Swati Mathur, "Feature Extraction in Content Based Image Retrieval: A Review" *International Conference on Advancement in Information Technology* (ICAIT- 23 February 2013).

[11] S. Mangijao Singh, K. Hemachandran, "Content-Based Image Retrieval using Color Moment and Gabor Texture Feature", *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 5, No 1, September 2012.

[12] Amandeep Khokher, Rajneesh Talwar, "Content-based Image Retrieval: Feature Extraction Techniques and Applications" *International Conference on Recent Advances and Future Trends in Information Technology* (iRAFIT2012).

[13] Reshma Chaudhari, A. M. Patil, "Content Based Image Retrieval Using Color and Shape Features" *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* Vol. 1, Issue 5, November 2012.

[14] A.Ramesh Kumar, D.Saravanan, "Content Based Image Retrieval Using Color Histogram", *International Journal of Computer Science and Information Technologies*, Vol. 4 (2), 242 – 245, 2013.

[15] Gui-Zhi Li, Ya-Hui Liu, Chang-Sheng Zhou, "Relevance Feedback Using Semi-Supervised Learning Algorithm For Image Retrieval", *International Conference on Machine Learning and Cybernetics*, Tianjin, 14-17 July, 2013.

[16] Begüm Demir, Lorenzo Bruzzone, "A Novel Active Learning Method in Relevance Feedback for Content-Based Remote Sensing Image Retrieval", *IEEE Transactions On Geoscience And Remote Sensing*, Vol. 53, No. 5, May 2015.

[17] Jeebachh Jha, Manoj Sabnis, "A Survey on Feature Extraction Techniques in Content-based Image Retrieval", *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 3, Issue 12, December 2015.