Image Retrival using Feature of Color and Shape

Chaithra C.C
B.E., (M.Tech), Department of CS &E
Adichunchanagiri Institute of Technology
Karnataka, India

Sunitha M.R
Associate Professor
Department of CS &E
Adichunchanagiri Institute of Technology
Karnataka, India

***Abstract*-Image retrieval system is an efficient and effective tool for handling the image database. Content based image recovery system allow the user to enter a query then CBIR retrieve the image which are stored in the large image database, which are comparable to the query image. CBIR treated the entered image as query image and searches in database as visual features like color, shape, texture, etc. CBIR main focused on the visual feature of the image. Content based image recovery is a challenging problem because it searches the query image in large image database. So, in content based retrieval it used feature extraction. Feature extraction means extracting unique and valuable information from the image. Feature withdrawal is initial and important step in the image retrieval in CBIR. Here CBIR uses color feature extraction for image retrieval. Wavelet transform is used to retrieve image based on texture and shape based retrieval. Main aim of this proposed work is organization of the image using K nearest neighbor Algorithm.**

Keywords—CBIR, color moment, Gabour wavelet,KNN classification.

I INTRODUCTION

Due to advance in digital photography, network speed and storage capacity, storage of large amount of high quality images is possible. Today's Digital image is very useful in wide range of application like medical, military, crime prevention, architecture and many more applications. Image recovery system is a computer system very useful for the image retrieval in the large image database. Most traditional image retrieval are using when it stores the image in database it adding metadata such as captioning, keyword, or notes about the how image is seen so that is useful for retrieving a image in large image database. Initially image retrieval is used in web application and semantic web application during this application user enters the information he wants then the internet searches in very big data and retrieve the information what actually user required. Image search is focused data search useful for finding an image. To search for image user may give a query term like keyword, image file or link, or user can give a click on to the image. System checks that query image and retrieves the image which is similar to the query images. Image meta search searching of image in database on based on keyword, text, etc. system takes that search box and retrieve that related information from huge image database.

Image recovery system is used for retrieved the image which are similar to the query image. Recovered the images based on color, shape, texture, and description of the image. Images are kept in database like labeled and without labeled image or unlabeled image. Based on how images are retrieved is classified as two types they are: 1)Text Based Image Retrieved (TBIR) 2)Content Based Image Retrieved (CBIR).

A. Text Based Image Retrieved (TBIR).

In text based image recovery user need to give the description of the query image. In the database each images are stored in the form of descriptive form. Database manager need to give importance to every images and write a description of the images and store. When user give a query image in the form of description form the retrieval system searches in database then give a resulted image which are similar to the request image.

Text based image recovery is also called information retrieval because it uses information of image during storing the image in database and also useful in the searching images also. Search can be based on occupied text or other content based indexing. An information retrieval process begins when user give a query. Queries are proper statement of information needs, for example in web application user give the information in the search box and system considered as a single object and searches in the database and retrieve the information which are related to the query information.

A thing is an entity represented by information in a content group or database. User queries are matches based on that information database. Since the information is not available in database then system retrieve slightly similar information if not have information than it may not retrieve the information. Text based image retrieval is also called as information retrieval.

The performance of an info retrieval system is measure by the how it retrieve the information which are related to the query information. Retrieval measure is calculated as precision and recall. If this measure are low than it shown that information are retrieved so user satisfy the retrieval system. Many more measure are proposed to calculated how fast the image is retrieved from the large image database.

Disadvantage of this text based retrieved is it take a huge time, loss of information, more expensive task, time consuming and also complex one. Text based image uses description information for the each image when images are added to the database. So retrieving image in database using keyword is complex. In text based image retrieval user or admin need to enter the information of each and every image during the storage process, user give a query in the form of information of the image as query.

B. Content Based Image Retrieved (CBIR).

The content based image recovery is invented in 1992 when it was used by T. Kato to describe experimentations into automatic retrieval of image from the database. Images are retrieved based on color and shape in the CBIR. Compared to text based image retrieval content based image retrieval system is more effective. Because it uses feature of the images like color, shape, and texture information. When user given a query image then CBIR produce a feature withdrawal from that image and searches in the large image database. Content based image retrieval system is working on various type of format of the images like TIFF, Photo CD, GIF, JPEG, and PNG 1.2.

Content based image recovery system is also called as query by image content. Content based image recovery is opposed the traditional concept based approach. Content based visual information retrieval is application of computer idea techniques to the image retrieval problem that is the problem of penetrating of images in large image database. Content based retrieval means searching analysis is based on the content of the image rather than the metadata such as keyword, explosion of the image. The content of the images is referred as shape, text, and color. If user wants to store image in the database if manually than user give keyword for each image and stored in database it is time consuming. So CBIR is very useful for only writing one code than it apply to all images to store image in database.

Many content based image recovery systems were developed but the problematic of retrieving image based on pixel is unsolved.

1. Request technique

Dissimilar application of CBIR makes use of different type of query. User given a image than CBIR system consider that image as query image and procedure continue that query image. User can give a predefined image which are shown in the directory or it select randomly. User draws a rough calculation of images they are looking and specify the proportion of each color. CBIR system sees that shape it calculate the edge detection and make histogram based on the proportion of the color. It is helpful for the image retrieval in the large database.

2. Semantic retrieval

In semantic retrieval user, may give a request like “find pictures of AIT”. This type of task is very difficult to perform in the computer because AIT image is not taken in same directions. Therefore, many CBIR systems uses low feature like color, shape and texture information of the images. In general image retrieval system needs a human feedback in order to identify high level concept.

3.Human Interaction

CBIR system allows the user interaction during the query entry. User is the main entry point to the system. User need to give what type of images stored in the database and user give the they want this image is known as query image.

4.Other query method

Query image can be taken as browsing from the internet, or take a picture from a camera. User can give a query image by sketch how the image is seen and user may give a proportion of the color used in the image.

II. LITERATURE REVIEW

A characteristic image retrieval system includes three major components: i) feature extraction ii) high dimensional indexing and iii) system design . An image is represented as a color, shape and texture. While several image retrieval systems rely on only one feature for the extraction of relevant images.

The process of determining the combination of features that is most representative of a specific query image is called feature selection. Works has been done on color and texture feature extraction algorithms

A. color attribute extraction

Color description include the conservative color histogram, the unclear color histogram, the color correlogram and a more fresh color-shape based feature. The extraction of the color-based description follows a similar development in each of the four methods: Color space, Color feature extraction, Distance function and quantization of color space.

i. conservative color histograms

The conservative color histogram of an image indicate the frequency of happening of every color in an image. It refers to the likelihood mass function of the image strength which captures the intensities of the color channels RGB color space in R,G and B, or the HSV color-space in H,S and V, and likewise for other color spaces . The conservative color histogram can be represented as X Y and Z (x,y,z) = N . Prob(X=x, Y=y, Z=z), where X,Y and Z are the three color channels, and number of pixels is signify by N in the image [1].

ii. color correlogram

The color correlogram courier how the three-dimensional of pairs of colors changes with distance. A color correlogram in an image is defined bycolor pairs present in a table indexed. Global correlation is not efficient than The local associations between different colors in an image, special correlations can be caught with a small value of distance. An well-organized algorithm for computing the color correlogram exists and is described in [2]. The computation is based on the image size.Color correlogram method encrypts local and also global spatial information, and it work well for coarse color images [3].

B.Shape feature extraction

In image retrieval, as per applications, shape representation are required to be either invariant to translation, rotation, and scaling or not. Hence, two groups of shape symbols can be illustrious, boundary and region based. The first utilizes only the outer boundary of the shape while the other access the whole shape region . The most successful legislatures

for these two categories are Fourier descriptor and moment invariants. The main idea of a Fourier descriptor is to use the Fourier misshapen boundary as the shape feature. Some initial work can be found in [4]. Rui et al. proposed a altered Fourier descriptor which is both robust to noise and invariant to geometric alterations. In Hu identified seven such moments. Many improved Records of 3rd International Conference on Excellence Up gradation in Engineering.

Gross and Latecki developed an approach which conserved the qualitative differential geometry of the object boundary, even after an image was digitized . In [5], invariants is proposed to represent complex objects in a cluttered scene a outline of algebraic curves by parts or patches.

III. METHODOLOGY

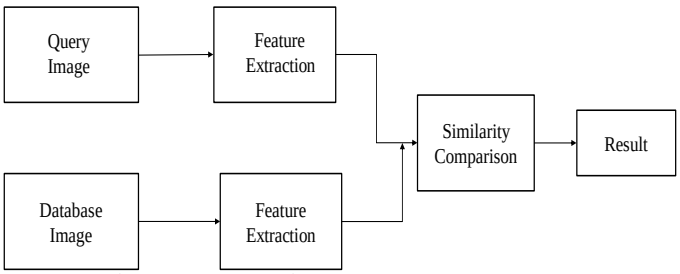

fig 1: Architecture of Content based image retrieval

The system architecture of proposed system is illustrated in the fig 1. The different stages of the process of feature extraction of the image are mentioned in two blocks one for query feature extraction and another for images in the database. User give the image as a query image the CBIR system process the feature extraction to that query image. And also it make feature extraction to the images which are in the database. If both image characteristic are equal than it check in the similarity comparison block and provide the resulted images. From a collection of significant size. Image processing covers a much broader field, including image improvement, compression, transmission, and clarification. While there are grey areas the distinction between normal image analysis and CBIR is usually fairly clear-cut. An example may make this clear. Many police forces now use programmed face recognition systems

- ***Content Based Image Retrieval System***

The initial use of the period content-based image retrieval in the works appears to have been by, to describe his experiments into automatic recovery of images from a database by color and shape feature. The term has since been extensively used to describe the process of recovering desired images from a large assembly on the basis of features (such as color, texture and shape) that can be mechanically extracted from the images themselves. The features used for retrieval can be either primitive or semantic, but the extraction process must be typically automatic. Recovery of images by manually-assigned keywords is positively not CBIR as the term is generally understood even if the keywords label image content.

CBIR varies from classical information retrieval in that image databases are basically formless, since digitized images consist decently of arrays of pixel strengths, with no inherent meaning. One of the key matters with any kind of image processing is the need to abstract useful information from the fresh data before any kind of mental about the image's contents is possible. Image databases thus differ fundamentally from text databases, where the raw material has already been logically structured by the author . There is no equal of level 1 recovery in a text database.

Wavelets in edge detection

Examination of image contents has been a considerable target for computer vision as well as image processing investigators since last few decades. An image carries change of information concerning contour, color, as well as orientation. The first step for contour extraction begins with the detection of edges. This practicality exposes the real significance of edge detection techniques in image processing field. Edge detection has a extensive range of applications in image compression, enhancement of images, watermarking, morphological operations, and refurbishment process and so on. The most important benefit of edge detection is that it reduces the huge data in an image, maintenance the structural attributes for further processing. The introduction of multi-sensory image fusion techniques have pointed towards a new measurement of research work in edge discovery process. This edge detection is useful for shape based retrieval.

The image in the large database is retrieved by a feature withdrawal. Feature means typical of the object. Feature extraction plays an important role during image retrieval. Feature is classified into 3 types during image processing they are low, middle and high. low level feature are called color and texture. Middle level feature is shape and high level feature is semantic gap between the object.

Most common features used in content based image retrieval are:

1.Color based retrieval:

Initially which images is need to store in a large image database first finding a histogram for each images and stored histogram along with the image in database. Histogram is a proportion of each color means how much percentage of each color are contributed to form that image. User give a query image then CBIR system generate histogram for that image and searches in the image database. If similar image having a same histogram, then retrieved system retrieve that image. Here user may give proportion of each color to finding image.

2. Shaped based retrieval:

Every object is identified by shape so it is one of the important technique. consider if the image is grayscale then CBIR not used a color feature extraction that time shaped based retrieval is a primitive level to recognized the object.

3.Texture based retrieval:

It is not very useful but useful in the distinguishable areas of images with similar colors lie sky and sea, leaves and grass etc. Texture based image retrieval identify degree of contrast,

regularity and randomness. This set not only defines the texture but also show where the texture present in the image.Texture is a very to represented in the image. It is very difficult to identify the texture location in the image so it is represented as two-dimensional gray level variation.

4.Content comparison using image distance measure

The most common method for comparing two images in CBIR is using distance measure. It calculating distance between the query image what user give and image in the database. The image distance measure compares the resemblance of two images in various dimensions such as color texture, shape and other. If the similarity value greater than 0 than it shows that there is similarity in the query image and image in the database. Search result is sorted based on the distance to the query image and image in the large image database. Many measure of image distance is developed.

5.Performance Measure

CBIR performance is analyzed by computing the values of precision and recall. Content based image retrieval includes the standard measures. The measure is recall and precision. These measures are used to improve the image retrieval process with various measures.

Precision:

Precision is the part of retrieved images that are relevant to the input image.

Precision= Total number of retrieval related image / Total number of retrieved image.

Recall:

Recall is the part of the images that are related to the query that are successfully retrieved.

Recall= Total number of retrieval image / Total number of related image.

- ***Current level 1 content based techniques***

In dissimilarity to the text-based method of the system content based technique works on a totally dissimilar information recovering stored images from a collection by associating features mechanically extracted from the images in a database. The commonest description used are exact events of colour, texture or shape; hence nearly all present CBIR systems, whether saleable or experimental. A typical system allows users to express queries by submit an ex of the type of image being required though some offer replacements such as selection from a palette or sketch input. The system then classifies those stored images whose feature values match those of the query most closely, and shows thumbnails of these images on the screen Mistake Orientation source not found.

- ***Colour retrieval***

Numerous methods for recovering images on the basis of colour resemblance is been labelled in this paper , but most are variations on the same basic idea. Each image having a color and calculating a histogram of each image and stored in a database. At searching time the user can either give a amount of color the image include or the image. (ex 75% of red and 25% of green) or submit an images. It searches in the database and matches based on histogram, then send a requested image which are similar to query image.

- ***Texture retrieval***

The capability to recover images on the basis of texture comparison may not look very useful. But the capability to match on texture similarity can often be useful in unique between parts of images with similar color (such as sky and sea, or leaves and grass). Fundamentally, it calculates brightness of selective pairs in an image. From these it is possible to calculate measures of image texture such as the degree of difference, coarseness, directionality and uniformity, or periodicity, directionality and chance. Other methods of information analysis for recover from the database include the use of Gabor filters and fractals. Texture queries can be stated in a similar manner to color requests, by selecting examples of desired surfaces from a palette, or by providing an example query image. The system then retrieves images with texture measures most similar in value to the query. A recent postponement of the technique is the texture dictionary, which retrieves textured regions in images on the basis of comparison to automatically-derived code words on behalf of important programs of texture within the gathering.

- ***Shape retrieval***

The ability to retrieve by shape is possibly the most understandable necessity at the original level. different texture, shape is a equally well-defined concept and there is considerable suggestion that natural objects are mostly familiar by their shape. A number of features representative of object shape are computed for every object known within each stored image. Requests are then answered by computing the same set of features for the query image, and retrieving those stored images whose constructions most closely match those of the query. Two key types of shape feature are usually used global features such as aspect ratio, circularity and moment invariants and local features such as sets of successive boundary segments. Another methods proposed for shape matching have included stretch bend of templates, contrast of directional histograms of edges extracted from the image, and shocks, misused representation of object shape that can be compared using graph matching techniques. Requests to shape recovery systems are framed either by identifying an example image to act as the query or as a user-drawn sketch.

Algorithm

1. Image database having a group of images

- User stores more than 100 images in the database the images are in any one of the format .bmp, .jpg or .tiff.
- First of all images in database converted into RGB model to HSV model.

2. Feature Extractions

- In feature extractions is taken out by using color and texture. For color feature extraction color histogram is used and also color moment used. For texture feature extraction wavelet transform and Gabor transform are used.
- Images are registered with their respective color and texture format.
- This feature is forwarded to Feature Vector Module.

3. Similarity Measure

- To find similarities between requested image and image in the database uses distance between two pixels in an images. To find a distance uses the proposed work uses Euclidean distance.

4. Retrieving the image

- When user give a requested image to content based retrieval system calculating a feature vector and it calculate feature vector in the database images both are compared in the similarity comparison.
- Finally image which is similar to the requested image is retrieved.

IV. RESULT AND DISCUSSION

This content based image recovery system is implemented using Matlab and in a backend it uses SQL for storage purpose.. When user starts using this system first, they are directed to GUI window where they get an option to select whether they want to search or they want to insert into database which can show here.

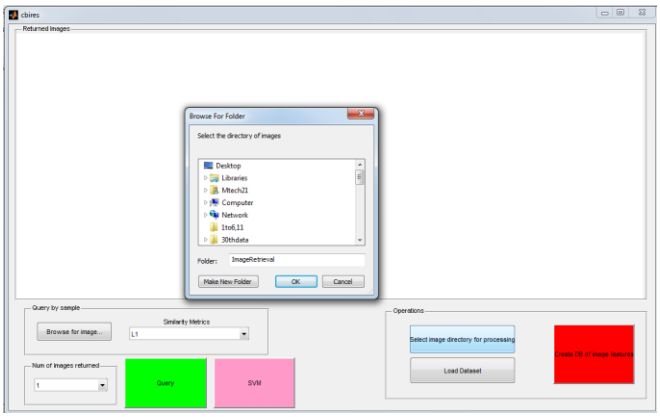

Fig 2. Display Browse for folder.

In fig 2 Graphical user interface allow the user to select image directory for processing. This window display many directory user can specifiy folder and OK button for processing

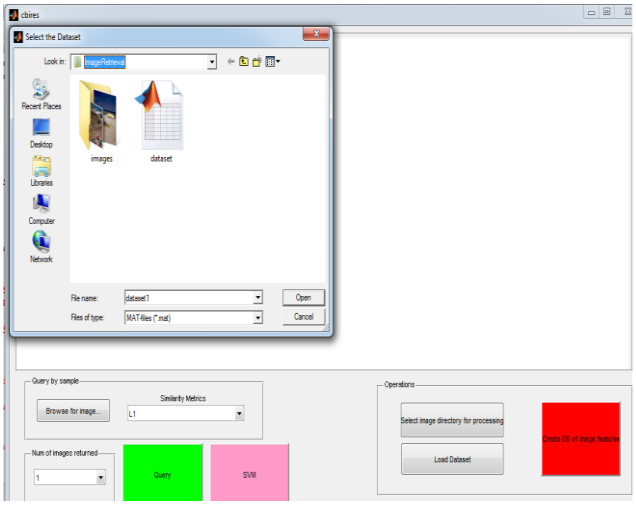

Fig 3: Select the database

In above fig 3 it allow the user to select the dataset to load the database. If database is not created then user need to create a database. In dataset it containing many image which are compare with the query image.

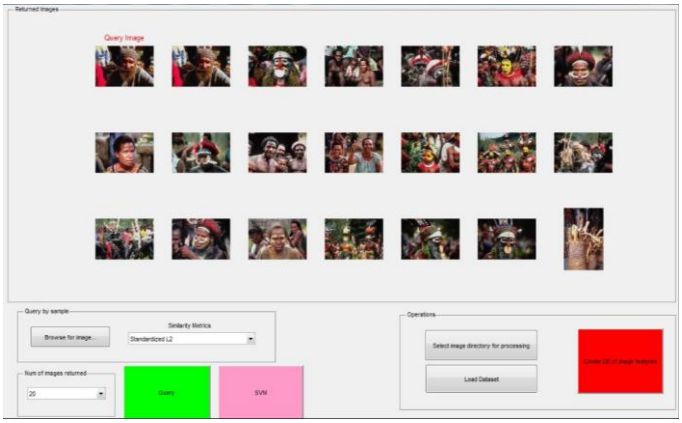

Fig 3: Relavant images are retrieve.

In above fig 3 it shows the query image and also related image which are similar to the query image. Precision and recall is calculated is based on number of images are retrieved.

V. CONCLUSION AND FUTURE WORK

Content based image recovery system is very useful for recover the image from big image file. Every image have some characteristic is known as feature. Feature extraction is very useful for retrieve the images. There are some feature is used in CBIR technique like color feature, shape and texture. By using color feature extraction CBIR technique retrieve the images with similar proportion of the color images. Shape retrieval is using for gray scale images, if the image does not have the proportion of color then CBIR technique uses shape feature to retrieve the image from large database. Texture based retrieval is used only when distinguishable object are there in the query image. KNN algorithm is used for

comparing query image to training sample and retrieve less relevant images.

Future enhancement:

Advances and studies are going on for further improvements in design and performance of content based image retrieval systems.

In future below method can be implemented:

- Shape and Texture investigation.
- Color Image histogram.
- Image ranking in Euclidean Distance method

REFERENCES

[1] G. Jaiswal, A. kaul, National Conference on computing, communications and control, (ccc-09).

[2] M. Saad, Content Based Image recovery- Literature Survey, EE 381k , Multi-Dimensional Digital Signal Processing,2008.

[3] S. Huang and S. Chang, "Images Retrieval: Current Technique, Promising Directions and Open Issues ", Journal of Visual Communication and Image Representation, vol. 10, pp. 39-62, March 1999.

[4] Jia, Z. Wang and Z. Lv, "A New and Effective Image Retrieval Method Based on Combined Features", Proc. IEEE Int. Conf. on Image and Graphics, vol. I, pp. 786-790, August2007

[5] S.F. Chang, " Automated images retrieval using color and texture", Columbia University, Tech. Rep CU/CTR 408-95-14, July 1995.