ABSTRACT: With the development of the Internet at large scale and the availability of various image capturing devices such as digital cameras, smart mobile phones, image scanners, digital image collection is increasing rapidly. With the result of that, content based image retrieval has grown in different areas to research. Efficient image retrieval tools are needed for users in various domains including remote sensing, crime prevention, fashion, medicine, publishing, architecture, etc. To achieve content based image retrieval, many retrieval systems have been developed. Two frameworks are accessible to perform the information retrieval which are text based retrieval and content based retrieval. But text based approach has disadvantages, firstly, more manual work is required to annotate the images and second is annotation inaccuracy. Therefore, to overcome these drawbacks, content based retrieval becomes a major topic in research. In this work, we have chosen this topic for our thesis work and explored autonomous sub-class based image retrieval depending upon color and texture of the query image dominant object by using uniform quantization and fuzzy methods of clustering to reduce the colors in the database images. Texture colors of the object have been used as a variable parameter to retrieve images according to human perception vision. The experimental results show efficient outputs in terms of recall and precision ratio. Keywords: content based image retrieval; low-level descriptors, particle swarm optimization.

I. INTRODUCTION

Due to the existence of a number of image retrieval systems, Content Based Image Retrieval system has become a challenging task. It is a process to find images similar in visual content to a given query from an image database. It is usually performed based on a comparison of low-level descriptors, such as color, texture, and shape features, extracted from the images themselves. It provides accurate image retrieval results which are used in fashion designing, remote sensing, medical areas. Early techniques of image retrieval were based on the manual textual annotation of images, a cumbersome and also often a subjective task. Texts alone are not sufficient because of the fact that interpretation of what is seen is hard to characterize by them. Hence, contents in an image, color, shape, and texture, started gaining prominence. Content-based image retrieval (CBIR) has become an important research area in computer vision due to huge digital image collections on World Wide Web. There are collections of images from art museums, medical institutes, and environmental agencies, to name a few. In the commercial sector, companies have been formed that are making large collections of photographic images of real-world scenes available to users who want them for illustrations in books, articles, advertisements, and other media meant for the public at large. The largest of these companies have collections of over a million digital images that are constantly growing bigger. Incredibly, the indexing of these images is all being done manually—a human indexer selects and inputs a set of keywords for each image. Each keyword can be augmented by terms from a thesaurus that supplies synonyms and other terms that previous users have tried in searches that led to related images. Keywords can also be obtained from captions, but these are less reliable. Content-based image retrieval research has produced a number of search engines. The commercial image providers, for the most part, are not using these techniques. The main reason is that most CBIR systems require an example image and then retrieve similar images from their databases. Real users do not have example images; they start with an idea, not an image. Some CBIR systems allow users to draw the sketch of the images wanted. Such systems require the users to have their objectives in mind first and therefore can only be applied in some specific domains, like trademark matching, and painting purchasing. Thus, the recognition of generic classes of objects and concepts is needed to provide automated indexing of images for CBIR. However, the task is not easy. Computer programs can extract features from an image, but there is no simple one-to-one mapping between features and objects. While eliminating this gap completely may require a very long time. Image features based on color and texture can be built and utilized to retrieve the required images.

Image Features

One of the key issues in querying an image out of databases by similarity is the choice of appropriate image descriptors and corresponding similarity measures. In this section, we present a brief review of considered low-level visual features in our approach.

1. Texture Descriptor:

Texture is an important aspect that refers to intrinsic surface properties of an object and their relationship to the surrounding environment. If appropriate texture descriptors are chosen, the performance of the CBIR can be improved. We use a gray level co-occurrence matrix (GLCM), which is a simple and effective method for representing texture. The entropy (E) is used to confine the textural information in an image and is defined as follows:
Usage of pattern recognition in image retrieval system: A basic prerequisite for most object recognition and classification systems is the task of segmenting and labeling objects within a scene [8]. After that, their relevant features can be extracted and used to classify, compare, cluster, or recognize the objects in question. Image retrieval uses basic principles of pattern recognition from concepts in databases, machine learning, statistics, patterns and 'soft' computing. Segmentation algorithms, unsupervised and supervised classification methods, descriptive and predictive spatial models and algorithms for large time series analysis will be presented to assist experts in their knowledge discovery for image retrieval. As images are mixes of numerical values from green, red and blue components taken with the help of cameras, we can process the content in the images by help of numbers of tools/software’s. Matlab is strong tool in image processing. It has been used on large scale by scholars, researchers, scientists, mathematicians all over the world.

Computer Vision: The front end of most vision systems consists of edge detection as preprocessing. The vision of objects is easy for the human because of the natural intelligence of segmenting, pattern matching and recognizing very complex objects. But for the machine, everything needs to be artificially induced and it is not so easy to recognize and identify objects. Towards Computer vision, the Machine needs pattern recognition; extracting the important features so as to recognize the objects, where the boundary detection or the edge detection is very crucial. Edge detection is finding the points where there are sudden changes in the intensity values and linking them suitably. All standard Computer vision aims to duplicate the effect of human vision by electronically perceiving and understanding an image. Giving computers the ability to see is not an easy task. Over the recent years, analysis of images such as segmentation, Edge Detection, Boundary detection, classification, clustering and texture property extraction were attracted the attention of many Researchers in the image processing and pattern recognition area.

Fig 1: Pattern-Classifier

II. RELATED WORK

M. Venkat Dass, Mohammed Rahmuth Ali, Mohammed Mahmood Ali[1] explained that in recent years, with the development of digital image techniques and digital albums in the Internet, the use of digital image retrieval process has increased dramatically. An image retrieval system is a computer system for browsing, searching and retrieving images from large databases of digital images. In order to increase the accuracy of image retrieval, a content-based image retrieval system (CBIR) based on interactive genetic algorithm (IGA) is proposed. Color, texture and edge have been the primitive low level image descriptors in content based image retrieval systems. In this paper we proposed a system that splits the retrieval process into two stages. In the query stage, the feature descriptors of a query image were extracted and then used to evaluate the similarity between the query image and those images in the database. In the evolution stage, the most relevant images were retrieved by using the IGA. IGA is employed to help the users identify the images that are most satisfied to the users’ need. The experimental evaluation of the system is based on a 10000 WANG color image database.

Jian-Ping Li et al. [2] explained, studied and compared the image retrieval techniques on the basis of color, texture and shape. To address this problem, content based image retrieval techniques are applied for information retrieval.

Muhammad Hammad Memon et al. [3] proposed a process of image refining retrieval result by exploiting and fusing.
unsupervised feature technique Principal component analysis (PCA) and spectral clustering. Gwenole Quellec et al. [4] a simple image characterization, which is based on the standardized moments of the wavelet coefficient distributions, is presented. An algorithm is proposed to compute this image characterization almost instantly for every possible separable or non separable wavelet filter.

Hatice Cinar Akakin et al. [15] described the design and development of a multilited content-based image retrieval (CBIR) system for microscopic images utilizing a reference database that contains images of more than one disease. Divya Ragatha Venkata et al. [16] put forward the model for a search engine where an image could be uploaded from the local database of the user to retrieve the information about it from the internet. It was similar to traditional keyword search used the most search engines with only one difference being that here the image is uploaded as a query rather than textual keywords

Bo Yang et al. [7] presented the study which focus on dimensionality reduction and image index of Web image. Y. Zhang, Z. Jia, and T. Chen et al. [8] presented an approach which encodes more spatial information through the geometry-preserving visual phrases (GVP). In addition to co-occurrences, the GVP method also captures the local and long-range spatial layouts of the words. This GVP based searching algorithm increases little memory usage or computational time compared to the BoV method. Moreover, it was shown that this approach can also be integrated to the min-hash method to improve its retrieval accuracy.

H.-W. Yoo, H.-S. Park et al. [9] suggested two kinds of indexing keys to prune away irrelevant images to a given query image: major colors’ set (MCS) signature related with color information and distribution block signature (DBS) related with spatial information. After successively applying these filters to a large database, only small amount of high potential candidates were gotten that are somewhat similar to a query image and made use of the quad modeling (QM) method to set the initial weights of two-dimensional cell in a query image according to each major color. Finally, more similar images were retrieved from the database by comparing a query image with candidate images through a similarity measuring function associated with the weights. In that procedure, a new relevance feedback mechanism was used. This feedback enhances the retrieval effectiveness by dynamically modulating the weights of color-spatial information. Experiments show that the proposed system is not only efficient but also effective.

III. GENERAL APPROACH USED

As has been already explained in literature survey, considered it as base for Content Based Image Retrieval System using Color Quantization using Particle Swarm Optimization technique and applied on different datasets collected from World Wide Web. In order to verify the algorithm it has implemented on variety of Images. A few out of total images used for the experimental results have been shown below. Images have been chosen with different lightning conditions, objects, shapes and using different locations of the camera from the objects.

![Figure 2: A sample of images from the database](image-url)

Flowchart of Presented work

Start

Read database images in MATLAB workspace

Generate color matrix using RGB color space for total color

Approximate colors in database images using uniform quantization and using reduced color

Apply FCM clustering to segment images into fewer clusters

Segmenting the database image into two classes containing dominant colors and other

Input query image and segment the foreground object:

Generate color matrix for each image and matches it with query image. Choose a threshold to include the image as retrieved

Retrieve images by inputting different energy via as well as nature colors found

Flowchart of Presented work

End

Flowchart of Presented work
2. Contrast Adjustment: It has been found that contrast adjusted image has high contrast values which reduces shading effect in the images and hence enhance the color display sensitivity of the images.

3. Uniform Quantization Based on Dominant Color Matrix
As total number of colors are reduced in quantized image as comparison to original image. Colors are approximated according to dominant color matrix find in step one.

4. Segmentation of Image Using Fuzzy C-Means Clustering Method:
In the color segmentation and pattern recognition, cluster is one the most important processes.
FCM algorithm is an unsupervised clustering method. From results shown above, it has been found that the FCM clustering method attempts to minimize the objective function by organizing the data into clusters which can be later fed for image retrieval process.

5. Object Detection in Query Image:
By using the segmented image high intensity pixels, corresponding values from FCM clustering and put into an array, which stores the texture pattern of the object and by using this texture pattern pixels from database images are kept as it is along with discarding other pixel area.
After that images will be retrieved based on color quantization.

IV. RESULTS
For image retrieving, a colored query image is selected from database. After that only the object is localized to make it less complex.

The algorithm color parameter is varied to retrieve different type of output images for matching human perceptive view of image retrieval. It has been found that increasing the colors to be matched with query image and database image, there is high precision rate.

<table>
<thead>
<tr>
<th>Colors in object texture</th>
<th>True positive</th>
<th>False positive</th>
<th>False negative</th>
<th>Similar object in database</th>
<th>Recall Ratio</th>
<th>Precision ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>14</td>
<td>3</td>
<td>4</td>
<td>18</td>
<td>.78</td>
<td>.82</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>1</td>
<td>8</td>
<td>18</td>
<td>.56</td>
<td>.91</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>0</td>
<td>14</td>
<td>18</td>
<td>.22</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>0</td>
<td>15</td>
<td>18</td>
<td>.17</td>
<td>1</td>
</tr>
</tbody>
</table>

Table: Comparison of retrieval precision

V. CONCLUSIONS
In this work, a method is proposed for image retrieval based on colors in an image. There is lot of work done found in the literature in which different classes of objects has been retrieved but there is less work found on sub-classes. In this work, database of flowers is chosen for retrieving sub-classes of flowers. The algorithm works well in retrieving similar type of flowers depending upon query image. PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles. The detailed information will be given in following sections.
Compared to GA, the advantages of PSO are that PSO is easy to implement and there are few parameters to adjust. PSO has been successfully applied in many areas: function optimization, artificial neural network training, fuzzy system control, and other areas where GA can be applied.

REFERENCES


