Abstract – Due to the development and improvement in internet with high speed for the last few years and the availability of a large digital image collection, efficient image retrieval systems are required. Retrieval can be text-based and content-based. Content based image retrieval (CBIR) is a process of retrieve and display relevant images from large collection of image database on the basis of their visual content. CBIR is used for retrieval of images depending upon visual contents of images known as features. This paper focuses on color and texture based techniques for achieving efficient and effective retrieval of images. Color feature extraction is done by RGB to HSV conversion and color moments (mean and standard deviation). Texture feature extraction is acquired by Gabor wavelet filter and wavelet moments (mean and standard deviation). For classification of extracted features we have used support vector machine (SVM). Euclidian distances are calculated of every features for similarity measures.

Keywords – CBIR, Color feature, Euclidian distance, Gabor wavelet filter, RGB to HSV, SVM, Texture feature.

I. INTRODUCTION

The use of images in human communication is not a new concept, our cave-dwelling ancestors painted pictures on the walls of their caves, usage of maps and building plans for delivering information is almost certainly dates back to pre-Roman times. However the twentieth century has seen the growth and importance of images in all turns of life. Images play a vital part in the fields of medicine, journalism, education, advertising, design, and entertainment.

The need for efficient storage and retrieval of images recognized by managers of large image collections such as picture libraries and design archives for many years. After examining the issues involved in managing visual information, the participants concluded that images were indeed likely to play an increasingly important role in electronically-mediated communication. But, the significant research advances, relating collaboration between a numbers of disciplines, would be required before image providers could take full advantage of the opportunities offered. They recognized a number of critical areas where research was required, comprising, image query matching, indexing and feature extractions user interfacing and data representation.

Ongoing expansion of digital images requires improved methods for sorting, browsing, and searching through ever-growing image databases. Such databases are used by various professionals including doctors searching for similar clinical cases, editors looking for illustration images and almost everyone needs to organise their personal photos. Other applications comprise accessing video archives by means of similar key frames, detection of unauthorized image use, or cultural heritage applications. Former approaches to the image indexation were based on text descriptions and suffered not only from laborious and expensive creation but also imprecise description. Textual descriptions are influenced by personal background and expected utilisation, which is difficult or even impossible to predict.

Content-Based Image Retrieval (CBIR) systems are search engines for image databases, which index images according to their content. A typical task solved by CBIR systems is that a user submits a query image or series of images and the system is required to retrieve images from the database as similar as possible. Another task is a support for browsing through large image databases, where the images are supposed to be grouped or organised in accordance with similar properties. Although the image retrieval has been an active research area for many years (Smeulders et al. (2000) [1] and Datta et al. (2008) [2]) this difficult problem is still far from being solved. There are two main reasons, the first is so called semantic gap, which is the difference between information that can be extracted from the visual data and the interpretation that the same data have for a user in a given situation. The other reason is called sensory gap, which is the difference between a real object and its computational representation derived from sensors, which measurements are significantly influenced by the acquisition conditions.

Content-Based Image Retrieval using Color Moments, Wavelet Moments & SVM Classifier

Arpit Sameriya
M. Tech. Scholar
Amity University, Jaipur, Rajasthan (India)

Ms. Bhawana Sharma
Asst. Professor
Amity University, Jaipur, Rajasthan (India)
The semantic gap is usually approached by learning of concepts or ontologies and subsequent attempts to recognise them. A system can also learn from the interaction with a user or try to employ combination of multimedia information. However, these topics are beyond the scope of this work and we refer to reviews Smeulders et al. (2000) [1] and Lew et al. (2006) [3] for further information. This paper concern two major issues while designing the CBIR system:

- Every image in the image data base is to be represented efficiently by extracting significant features.
- Relevant images are to be retrieved using similarity measure between query and every image in the image data base.

II. CONTENT-BASED IMAGE RETRIEVAL

Content based image retrieval (CBIR) is the application of Computer vision techniques to the image retrieval problem, the problem of finding a digital images in large databases. The Content Based Image Retrieval tries to solve this problem as it provides the means to index, examine and retrieve those images. Content-based retrieval uses the contents of images to represent the images. A typical content-based retrieval system is divided into off-line feature extraction and online image retrieval. A basic framework for content-based image retrieval is illustrated in Figure 1 [4].

In off-line stage, the system automatically extracts visual attributes (color, shape, texture, and spatial information) of each image in the database based on its pixel values and stores them in a different database within the system called a feature database. The feature data (also known as image signature) for each of the visual attributes of each image is very much smaller in size compared to the image data, thus the feature database covers an abstraction (compact form) of the images in the image database. One advantage of a signature over the original pixel values is the significant compression of image representation. But, a more significant aim for using the signature is to gain an improved correlation between image representation and visual semantics [4].

In on-line image retrieval, the user can submit a query example to the retrieval system in search of desired images. The method represents this example with a feature vector. The distances (i.e., similarities) between the feature vectors of the query example and those of the media in the feature database are then computed and ranked. Retrieval is directed by applying an indexing scheme to provide an efficient way of searching the image database. Then system ranks the search results and then returns the results that are most similar to the query samples. If the user is not satisfied with the search results, he can provide relevance feedback to the retrieval model, which has a mechanism to learn the user’s information needs.

Content Based Image Retrieval is a task of searching images from a database and retrieval of an image, which are seems to be visually similar to a given example or query image. Content-based image retrieval uses the visual contents of an image such as shape, texture, color, and spatial layout to represent and index the image. In content-based image retrieval systems, the visual stuffing of the images in the database are extracted and described by multi-dimensional feature vectors. These feature vectors can be computed by different methods available to the users. The CBIR system consists of following components:

**Query Image**
It is the image to be search in the image database whether the same image is present or not or how many are similar kind images are exist or not.

**Image Database**
It consists of n number of images depends on the user choice.

**Feature Extraction**
It extracts visual information from the image and saves them as features vectors in a features database. The
feature extraction finds the image description in the form of feature value (or a set of value called a feature vector) for each pixel. These feature vectors are used to compare the query with the other images and retrieval.

**Image Matching**
The information about each image is stored its feature vectors for computation process and these feature vectors are matched with the feature vectors of query image which helps in measuring the similarity.

**Resultant Retrieved Images**
It searches the previously maintained information to find the matched images from database. The output will be the similar images having same or very closest features as that of the query image.

![Figure 2: CBIR System and its various components](image)

### III. METHODOLOGY

Figure 3 depicts the block diagram of the proposed retrieval approach.

![Figure 3: Block diagram of proposed approach](image)
A single feature may lack sufficient discriminatory information to permit the retrieval of relevant images [5], so here multiple features utilizing a combination of color and texture features that have been extracted separately are used. When a query image is entered into the retrieval system, it must first be pre-processed. Then each of the channels is then wavelet decomposed into a wavelet image, after which color features and texture features are extracted from the transformed image. Rest of the methodology is explained in the following subsections:

Wavelet Transform
To increase effectiveness in CBIR wavelet features computed from discrete wavelet coefficients are assigned weights.

Extraction of Color Features
After the wavelet transform, to extract the color features first we apply RGB to HSV conversion.

The HSV values of a pixel can be transformed from its RGB representation according to the following formula:

\[
H = \cos^{-1} \frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{\frac{1}{2}[(R-G)+(R-B)]^2 + \frac{1}{2}[(G-B)^2]}}
\]

\[
S = 1 - \frac{2}{(R+G+B)}
\]

\[
V = \frac{(R+G+B)}{3}
\]

Color Moments
Color moments are measures that can be used differentiatate images based on their features of color. Once calculated, these moments provide a measurement for color similarity between images. These values of similarity can then be compared to the values of images indexed in a database for tasks like image retrieval. In order to improve the discriminating power of color indexing techniques, we divide the image horizontally into three equal non-overlapping regions and from each of the three regions, we extract from each color channel the first two moments of the color distribution. If we interpret the color distribution of an image as a probability distribution, then the color distribution can be characterized by its moments. If the value of the \(i^{th}\) color channel at the \(j^{th}\) image pixel is \(I_{ij}\) and the number of pixels is \(N\), then the index entries related to this color channel and the region \(r\) are known as the color moments followed by the formula:

Mean:

\[
E_{r,i} = \frac{1}{N} \sum_{j=1}^{N} I_{ij}
\]

Standard Deviation:

\[
s_{r,i} = \sqrt{\frac{1}{N} \sum_{j=1}^{N} (I_{ij} - E_{r,i})^2}
\]

Where \(E_{r,i}\) (1 ≤ \(i\) ≤ 3) represents the average color (mean) of the region \(r\) and \(s_{r,i}\) represents the standard deviation of the region \(r\).

The feature vector from the extracted color features is given by the formula:

\[
f_c = \{E_{1,r}, \sigma_{1,r}, E_{2,r}, \sigma_{2,r}, E_{3,r}, \sigma_{3,r}, \ldots \ldots \ldots \ldots \ E_{r,i}, \sigma_{r,i}\}\]

Where, (1 ≤ \(r\), \(i\) ≤ 3). And \(r\) represents the region and \(i\) represents the color channel.

Extraction of Texture Features
For a given image \(I(x,y)\) with size \(P \times Q\), its discrete Gabor wavelet transform is given by a convolution:

\[
G_{mn}(x,y) = \sum_{s} \sum_{t} I(x-s, y-t) \psi_{mn}(s,t)
\]

Where, \(s\) and \(t\) are the filter mask size variables, and \(\psi_{mn}\) is the complex conjugate of \(\psi_{mn}\) which is a class of self-similar functions generated from dilation and rotation of the following mother wavelet:

\[
\psi(x,y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \exp(j2\pi Ww)
\]

Where \(W\) is called the modulation frequency. The self-similar Gabor wavelets are obtained through the generating function:

\[
\psi_{mn}(x,y) = a^{-m} \psi(\tilde{x}, \tilde{y})
\]

Where \(m\) and \(n\) specify the scale and orientation of the wavelet respectively.

And \(m = 0, 1, ..., M - 1, n = 0, 1, ..., N - 1, \) and

\[
\tilde{x} = a^{-m}(x \cos \theta + y \sin \theta)
\]

\[
\tilde{y} = a^{-m}(-x \sin \theta + y \cos \theta)
\]

Where \(a > 1\) and \(\theta = \frac{\pi n}{N}\).

The variables in the above equations are defined as follows:

\[
a = \left( \frac{U_l / U_j} {\frac{1}{2}} \right) \frac{1}{M-1}
\]

\[
W_{m,n} = a^m U_l
\]

\[
\sigma_{x,m,n} = \left( \frac{a+1}{2} \right) \frac{1}{n a^m - (a-1) U_l}
\]

\[
\sigma_{y,m,n} = \left( \frac{a+1}{2} \right) \frac{1}{n a^m - (a-1) U_l}
\]

Wavelet Moments

After applying Gabor filters on the image with different orientation at different scale, we obtain an array of magnitudes:

\[
E(m,n) = \sum_{x} \sum_{y} |G_{mn}(x,y)|
\]

Where, \(m = 0, 1, ..., M - 1; n = 0, 1, ..., N - 1\)
These magnitudes represent the energy content at different scale and orientation of the image. The main purpose of texture-based retrieval is to find images or regions with similar texture. It is assumed that we are interested in images or regions that have homogenous texture, therefore the following mean \( \mu_{mn} \) and standard deviation \( \sigma_{mn} \) of the magnitude of the transformed coefficients are used to represent the homogenous texture feature of the region:

\[
\mu_{mn} = \frac{\sum_x \sum_y |g_{mn}(x, y)|}{P \times Q}
\]  
\[
\sigma_{mn} = \sqrt{\frac{\sum_x \sum_y [(g_{mn}(x, y) - \mu_{mn})]^2}{P \times Q}}
\]

A feature vector \( f_g \) (texture representation) is created using \( \mu_{mn} \) and \( \sigma_{mn} \) as the feature components [6]. Five scales and 6 orientations are used in common implementation and the feature vector is given by:

\[
f_g = (\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \ldots, \mu_{45}, \sigma_{45})
\]  

Combining the Features

The retrieval result using only single feature may be inefficient. It may either retrieve images not similar to query image or may fail to retrieve images similar to query image. Hence, to produce efficient results, we use combination of color and texture features.

Classification by Support Vector Machine (SVM)

Support Vector Machines have shown their capacities in pattern recognition. Find the best hyper-plane separating relevant and irrelevant vectors maximizing the size of the margin (between both classes) is the main aim of SVM classification method. Relevant and irrelevant vectors are linearly separable assumes in initial method. The whole image database separated into two classes by SVM.

The two classes are including the unlabelled images with two types they are relevant and irrelevant unlabelled images. The relevant labelled image is related to the relevant unlabelled images in the image database. In similar way the irrelevant labelled image is related to the irrelevant unlabelled images in the database. SVM is also classifying the unlabelled images in accuracy manner.

Image Database

For evaluation of the proposed method, it has been implemented using MATLAB (2010a) and tested on a general purpose WANG database [7] containing 1000 Corel images in JPEG format of size 384x256 or 256x384. The image set comprises 100 images in each of 10 categories. This database was used extensively to test many CBIR systems because the size of the database and the availability of class information allows for performance evaluation as can be seen in the following sections. This database was created by the group of professor Wang from the Pennsylvania State University and is available for download. This database was also used for classification experiments.

Similarity Measure by Euclidean Distance

The texture similarity measurement of a query image \( Q \) and a target image \( T \) in the database is defined by:

\[
D(Q, T) = \sum_m \sum_n d_{mn}(Q, T)
\]

Where,

\[
d_{mn} = \sqrt{(\mu_{mn}^{Q} - \mu_{mn}^{T})^2 + (\sigma_{mn}^{Q} - \sigma_{mn}^{T})^2}
\]

If \( f_g^Q = (\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \ldots, \mu_{45}, \sigma_{45}) \) denote texture feature vector of query image and \( f_g^T = (\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \ldots, \mu_{45}, \sigma_{45}) \) denote texture feature vector of database image, then distance between them is given by the Euclidean distance formula:

\[
\Delta d = \sqrt{\sum_{i=1}^n [(f_g^Q - f_g^T)^2]}
\]

Where \( n \) is the number of features, \( i = 1, 2, \ldots, n \). Both images are the same for \( \Delta d = 0 \) and the small value of \( \Delta d \) shows the relevant image to the query image.

IV. SIMULATION AND RESULTS

The simulation is carried out by using image processing toolbox of MATLAB 2010a.
V. CONCLUSION

We have developed an improved content-based image retrieval system. The system uses extraction of color moments (Color Feature Extraction) using RGB to HSV transform and extraction of wavelet moments (Texture Feature Extraction) using Gabor Wavelet Filter. Similarity measurement is done using Euclidean Distance. The algorithm has been implemented and tested using 500 COREL color images and found out the retrieval performance on the basis of evaluation parameters; precision, recall and F-score.

REFERENCE


