Image Retrieval using Fusion of Color – Size and Texture Features

N.Puviarasan
Associate Professor
Dept. of CSE
Annamalai University
Tamilnadu, India

R.Bhavani
Professor
Dept. of CSE
Annamalai University
Tamilnadu, India

T.S.Arthi
PG Student
Dept. of CSE
Annamalai University
Tamilnadu, India

ABSTRACT
Due to the revolutionary explosion of internet and digital technologies, the requisite to have a system that organizes the copiously available digital images for easy categorization and retrieval has been imposed. Nowadays, Content Based Image Retrieval (CBIR) has become a solution and source of accurate and fast retrieval. CBIR uses the visual contents to retrieve relevant images from large databases according to user's interests. The visual contents (color, texture, shape etc) serve as the features for the images. Features are measurements of ultimate interest analyzed from an image. In this paper, a new type of visual feature named Color-Size feature which integrates the information of both color and size of the image in terms of number of segments is proposed. Initially the images are segmented using Watershed segmentation approach. Different images would yield different number of segments that has to be taken into account for the extraction of features. From the segmented image the Color-Size features are extracted using Color-Size Histogram. Gabor texture and GLCM (Gray Level Co-occurrence Matrix) are employed to extract texture features. The feature extraction process is exercised for both the query image and images stored in database. After the extraction of mentioned features in the proposed system, the relevant images are retrieved for the given user's query image with respect to closest distance among the feature vectors. In this paper, the fusion of Color-Size with Gabor and Color-size with GLCM texture are proposed and it is deduced that the compounding of Color-Size with Gabor yields better results.

General Terms
Image Processing

Keywords
Content based Image Retrieval; Color-Size; Feature vector; Visual features; Watershed approach;

1. INTRODUCTION
Images have always been an inevitable part of human communication and its root millennia ago. To express ideas and convey information, humans always preferred concrete visual means (images, painting) to more extent. The importance of Content-Based Image Retrieval is motivated by the increasing hope for retrieving images from growing digital image databases over the Internet. As the size of image databases grow exponentially, the running of large image databases became difficult which leads to the motivation of research communities to explore new algorithms for feature extraction. The most common two grouping of image retrieval methods are text based and content based. Textual notation of images becomes impractical and time consuming. It cannot capture the visual contents of images. The amount of labor required to annotate every single image, as well as the difference in human perception when describing the images which lead to inaccuracies during the retrieval process. It is often introspective, context-sensitive [1]. Hence there is a need for better system. Problems with text-based access to images have shifted the focus of the researchers to content based image retrieval.

Content based image retrieval, a technique which uses the visual contents of an image to search and retrieve relevant and similar images relating to the query image from huge image databases in accordance with user's interests. Active research in CBIR is geared towards the development of methodologies for analyzing and interpreting image databases. In CBIR system, it is common to group the image features in three main classes: color, texture and shape [2].

IBM [3] has introduced several versions of the QBIC system which is widely used for image retrieval purposes. Using the system, users are able to search the image by identifying certain characteristics of the image. Latest versions of QBIC use regional segmentation function to segment the image into different regions. The main disadvantage of this system is that it is sensitive to the variation of illumination.

A CBIR scheme [4] has shown the comparative analysis of various feature extraction techniques such as Average RGB, Co-occurrence, Local color histogram, Global color histogram and Geometric moments. An improvement has been stated in image retrieval by introducing the idea of Query modification through image cropping.

A method of image retrieval using the histogram, color and edge detection features. Image segmentation is used with the intention of getting better accuracy percentage and it has proved to be successful approach [5].

The image properties are analyzed in this paper [6] by using image processing algorithms. For color the histogram of images is computed, for texture histograms of image are computed and entropy, smoothness and uniformity are calculated. For retrieval of images Euclidean distance measure is used.

The paper [7] proposed the extraction of low – level image feature like color histogram, color coherence vector. Then edge detection techniques are applied to get better output. Finally Manhattan distance is used to find the similarity among the images.

The texture feature is analysed in the paper [8] which proceeds by evaluating the structural and statistical approaches of texture analysis. It has compared the performance of both approaches and given the average precision values.

An image retrieval application [9] has performed a simple color-based search using color, texture and shape features. This paper has incorporated fuzzy histogram for color, Tamura features for texture and Moment invariants for shape features.

A new CBIR scheme incorporating color, texture as shape features is proposed in [10]. Histogram, Gabor filter and Moment invariants are used for extraction of mentioned features respectively.

2. ARCHITECTURAL DESIGN OF PROPOSED WORK

It is proposed to extract Color-Size and Texture features from the images which comprises of five steps. The architecture for the proposed work is shown in Figure 1. Following methodologies are used in the proposed system.

- Image Pre-processing
- Segmentation
- Feature Extraction
- Fusion of features
- Similarity analysis

The rest of the paper is organised as follows. Section 3 presents the proposed system framework. Section 4 explains the feature extraction. Fusion of features and Similarity analysis are brought out in Section 5 & 6 respectively. Section 7 displays the experimental results and Section 8 provides the conclusion of proposed methodology.

3. SYSTEM FRAMEWORK

3.1 Image Pre - Processing

Initially, the query image undergoes pre-processing. It is performed to remove the distortions and to enhance the image for further processing. Image filtering [11] is not only used to improve image quality but also is used as a pre-processing stage in many applications. Image noise is the random variation of brightness or color information in images produced by the sensor and circuitry of a scanner or digital camera.The noise is filtered using Median filter which serves the best for filtering the noise in the proposed system by after having a comparative analysis among the Averaging, Median and Wiener filters based on the PSNR (Peak Signal to Noise Ratio) values.

Median filter is a non linear digital filtering technique, often used to remove salt and pepper noise [12, 13]. It is widely used in digital image processing because under certain conditions, it preserves edges whilst removing noise. Here, the neighboring pixels are ranked according to brightness (intensity) and the median value becomes the new value for the central noisy pixel. It does not shift boundaries, as it can happen with conventional smoothing filters (a contrast dependent problem). The PSNR metric computes the peak signal-to-noise ratio, in decibels, between two images [11]. This ratio is often used as a quality measurement between the original and a compressed image. The higher the PSNR, better the quality of the compressed or reconstructed image. PSNR is most easily defined via the Mean Squared Error (*MSE*).

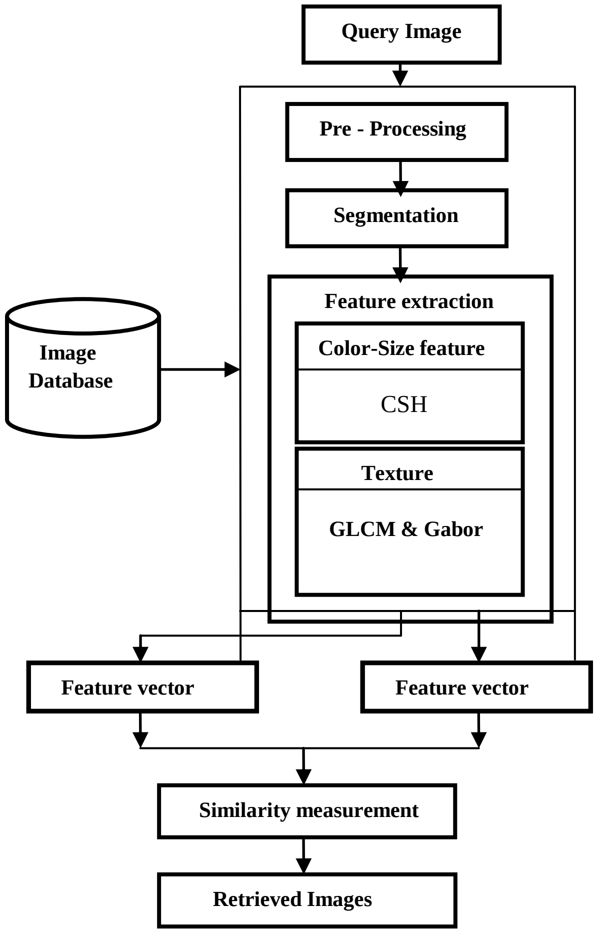


Fig 1: Architectural design of proposed work

Given a noise-free $m \times n$ monochrome image I and its noisy approximation K, MSE is defined as:

$$MSE = \frac{1}{m\,n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \tag{1}$$

To assess the comparative performance of various filters for removing salt-and pepper noise, PSNR is computed as follows,

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{2}$$

Here, MAX_I is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. For color images with three RGB (red, green, blue) values per pixel, the definition of PSNR is the same except the MSE is the sum over all squared value differences divided by image size and by three. Figure 2 shows the PSNR graph among Median, Wiener and Averaging filters.

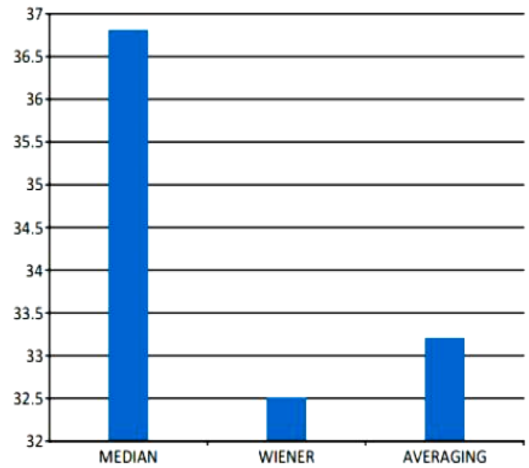


Fig 2: PSNR graph for Median, Wiener, Averaging filters

3.2 Segmentation

Image segmentation is the most important field of image analysis which refers to the process of partitioning a digital image into multiple segments [5]. The segments possess sets of pixels also known as super pixels. Each of the pixels in the segments is similar with respect to some characteristic or computed property, such as color, intensity, or texture.

Watershed Segmentation is used in the proposed work which is a widely used method for image segmentation in the field of mathematical morphology. It is a very predominant segmentation scheme ensures the closed region boundaries and gives solid results which automatically separates or cuts apart particles that touch. It uses concepts from mathematical morphology to partition images into homogeneous regions. Watersheds often produce more stable results than other segmentation approaches based on detection of discontinuities, thresholding, and region processing.

In this representation, the image is considered as a topographic relief, the numerical value of each pixel determining the corresponding point elevation. A rain drop hitting any point will travel along the greatest gradient towards the nearest local minima, or in some points two or more local minima. The set of all pixels leading to the same local *minimum* is called a *catchment basin*. The set of all the different catchment basins constitutes a non intersecting partition of the image, i.e. a *segmentation* of the image.

3.2.1 Remedy for over-segmentation

This method can suffer from over-segmentation, which occurs when the image is segmented into an unnecessarily large number of regions [13]. To overcome, small local minima in the gradient image which consists of a *small number of pixels* or *pixels having low contrast with their neighbors*,that are eliminated by assigning two scaling parameters: r and h in the proposed system. Parameter r is the size of the structuring element of dilation operators, whose application eliminates local minima of size less than r pixels, and parameter h is the height of elevation used to remove the local minima with low contrast. These two parameters can be used to control the coarseness of the segmentation results: as r or h increases, the number of regions generated decreases. The images are segmented by assigning r=1 and h=3 in the proposed work is exhibited in Figure 3.

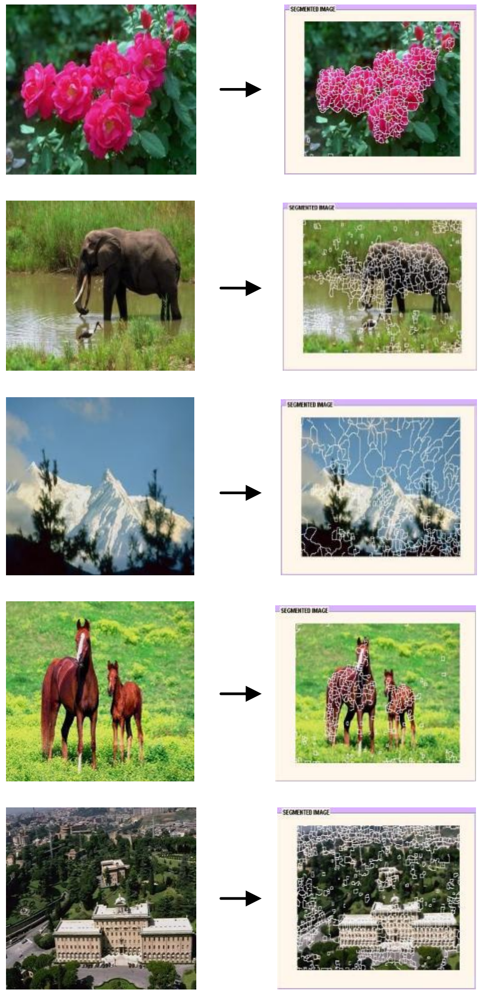


Fig 3: Watershed segmented images with r=1 and h=3

4. FEATURE EXTRACTION

Features are measurements of ultimate interest analyzed from an image. A feature is a characteristic which can capture a certain visual property of an image either globally for the entire image or locally for regions or objects. There exist many features like color, shape, texture etc. It is necessary to extract these features from image for analyzing the similarity among different images which is mandatory for retrieving relevant images. Feature extraction is the process of mapping the image pixels into the feature space where it is stored as feature vectors.

4.1 Color – Size

One of the important features that make possible the recognition of images by human is color. Color is a property which is used for providing discrimination among various objects and hence adopted as the most dominant feature in CBIR (ex: QBIC). Usually colors are defined in three dimensional color spaces [9]. These could be **RGB** (Red, Green, and Blue), **HSV** (Hue, Saturation, and Value) or **HSB** (Hue, Saturation, and Brightness). The ***Size*** feature denotes the size of the image in terms of number of segments. Different images would yield different number of segments that has to be taken into account for the extraction of features.

4.1.1 Color Histogram

It is the method of representing color information. It is a type of bar graph, where each bar denotes a particular color of the color space being used. A histogram of an image is produced first by discretization of the colors in the image into a number of bins, and counting the number of image pixels in each bin [9]. The histogram provides a compact summarization of the distribution of data in an image. Thus it gives the count of pixels in an image representing a particular color. The Figure 4 presents the sample extracted Color-Size feature values. Hence in the proposed system, the mean value of each channel is computed and stored as feature vector.

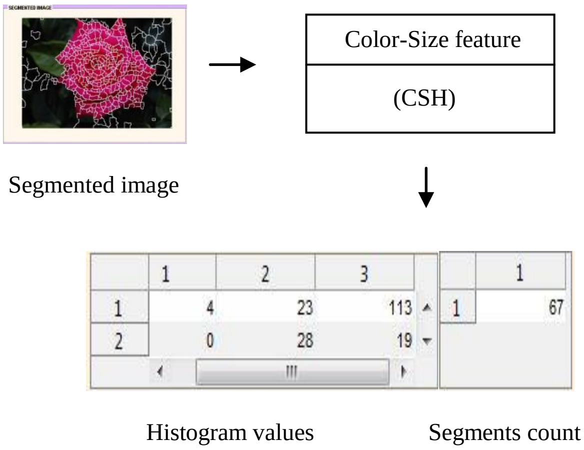


Fig 4: Extracted sample Color-Size values

4.1.2 The Proposed Size Feature

The **Size** feature denotes the count of segments in the watershed segmented image. In this paper, the size component is embedded over the histogram with the effort of increasing the accuracy rate in retrieval process. Therefore Color-Size component encompasses four feature values with three components representing the RGB value and single component denoting the number of segments. Thus, both color and size factors are incorporated together for the effectiveness of presenting more similar and relevant images during the retrieval process.

4.2 Texture

It is one of the main features utilized in image processing and vision which denotes the visual patterns. Texture describes the distinctive physical composition of a surface and contains information about structural arrangement of surface (clouds, leaves, bricks etc). Since an image is composed of pixels, texture can be defined as the entity consisting of mutually related pixels and group of pixels. This group of pixels is called as texture primitives or texture elements referred as *texels*. The methods of characterizing texture fall into major categories: *Statistical* and *Structural* [8]. The former qualifies texture by the statistical distribution of the image density and latter describes texture by identifying structural primitives and their placement rules. GLCM and Gabor filter fall under Statistical and Structural categories respectively.

4.2.1 Gray – Level Co occurrence matrix

This gray-level spatial dependence matrix based calculations fall under second-order statistics. Haralick et. al. suggested a set of fourteen features which can be extracted from the co-occurrence matrix [16]. The elements in this matrix are the relative frequencies of occurrence of grey level combinations among pairs of image pixels. This matrix considers the relationship of image pixels in different directions, such as horizontal, vertical, diagonal and anti diagonal. It is two-dimensional array, P, in which both the rows and columns represent a set of possible image values. A GLCM P_d [i, j] is defined by first specifying a displacement vector d = (dx, dy) and counting all pairs of pixels separated by d having gray levels i and j. $P_d[i,j] = n_{ij}$, where n_{ij} is the number of occurrences of the pixel values (i,j) lying at distance d in the image. The following important five features are extracted from the co-occurrence matrix in the proposed system.

a) Contrast: It is a measure of the local variations present in an image. Computed by,

$$C(k,n) = \sum_i \sum_j (i-j)^k P_d[i,j]^n \tag{3}$$

b) Homogeneity: Measures the uniformity of the non zero entries in the GLCM given as

$$C_h = \sum_i \sum_j \frac{P_d[i,j]}{1+|i-j|} \tag{4}$$

c) *Entropy:* A measure of the image's spatial disorder which is calculated as,

$$C_e = -\sum_i \sum_j P_d[i,j] \ln P_d[i,j] \tag{5}$$

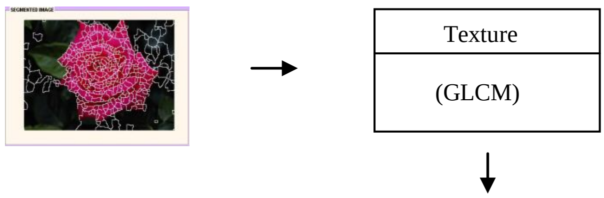
d) *Energy:* A feature which measures the local homogeneity and therefore represents the opposite of entropy. Basically it renders about the uniformity of the texture and computed by,

$$L_e = \sum_{i=1}^{m} \sum_{j=1}^{n} |C(i,j)| \tag{6}$$

e) *Correlation:* It is a measure of image linearity. It will be high if an image contains a considerable amount of linear structure.

$$C_c = \frac{\sum_i \sum_j [i.jP_d[i,j]] - \mu_i \mu_j}{\sigma_i \sigma_j} \tag{7}$$

Thus five feature values constitute the GLCM vector and the Figure 5 depicts the sample extracted values.



Segmented image

		3	4	5
1	170	0.1790	2.5043	0.828

Fig 5: Extracted sample GLCM values

4.2.2 Gabor Filter

It provides a useful way to analyze the texture information of an image. In our work, Gabor texture feature is adopted to represent the texture features. To extract the Gabor texture feature of an image *I*, *I* is first filtered with a bank of scale and orientation Gabor filters, and then computes the mean and standard deviation of the output of filters[10]. When an image is processed by Gabor filter, the output is the convolution of the image *I(x, y)* with the Gabor function *g(x, y)* which is

$$r(x, y) = I(x, y) * g(x, y) \tag{8}$$

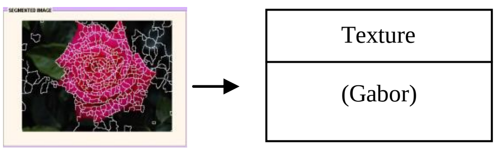
where * represents the 2D convolution. After applying Gabor filters on the image by orientation and scale, an array of magnitudes can be obtained as follows,

$$E(m,n) = \sum_x \sum_y |G_{mn}(x,y)|$$
$$m = 0, 1, \ldots, M-1;\ n = 0, 1, \ldots, N-1 \tag{9}$$

The magnitudes represent the energy content at different orientation and scale of image. The filter is applied with 4 scales and 3 orientations. The main purpose of texture based retrieval is to find images or regions with similar texture. The following mean μ_{mn} and standard deviation σ_{mn} of the magnitude of the transformed coefficients are used to represent the texture feature of the region,

$$\mu_{mn} = E(m,n)/P \times Q$$
$$\sigma_{mn} = \sqrt{\sum_x \sum_y \left(|G_{mn}(x,y)| - \mu_{MN}\right)^2 / P \times Q} \tag{10}$$

where *m* represents the scale and *n* represents the orientation. The feature vector that represents the texture features is created using mean μ_{mn} and standard deviation σ_{mn} as feature components. And these components are saved into two feature vectors having 12 values each and then these two vectors are combined in order to make the single feature vector that will be treated as an image texture descriptor. In this paper, twenty four feature values comprising mean and standard deviation of twelve values each are extracted and sample extracted values are presented in the Figure 6.



Segmented image

	9	10	11
1	0.4624	0.8289	0.4443

Fig 6: Extracted sample Gabor texture values

5. FUSION OF FEATURES

The Color-Size and Gabor features are fused together which accounts a vector of twenty eight values with four and twenty four values respectively. Then Color-size is combined with GLCM feature which encounters nine values with four and five values respectively.

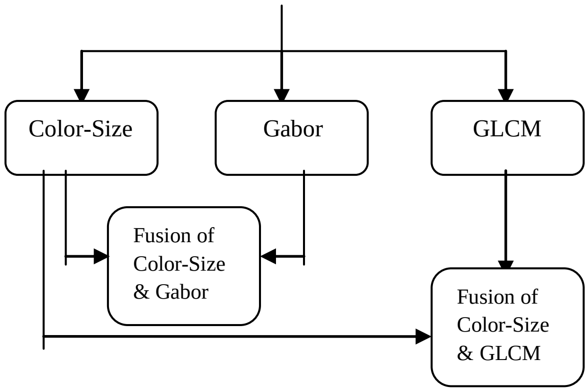


Fig 7: Fusion of proposed features

6. SIMILARITY MEASUREMENT

Similarity between two images is measured numerically that reflects the strength of connections between them. Similarity is crucial in obtaining relevant results. To obtain relevant results various researchers use different methods to measure similarity. For example, some researchers used fuzzy measures, histogram intersection, Euclidean distance and Manhattan distance. Euclidean distance is much suited when compared with other measures. It is employed to calculate the similarity between two feature vectors as follows,

$$E(I_q, I_d) = \sqrt{\sum_{i=1}^{n} [I_i^q - I_i^d]^2} \tag{11}$$

where I_q and I_d are image query and image database respectively, *i* is a feature range. Closer distance represents the higher similarity between images.

7. EXPERIMENTAL RESULTS

The simulations were taken place in MATLAB. The number of search results may vary depending on the number of similar images in the database. The development of powerful processing, faster and cheaper memories contribute much to CBIR growth. In order to assess the performance of the proposed method, Wang 1000 datasets [4] are applied. In our experiment, seven types of query images are selected, and 100 images are available in each category. Hence 700 images are taken from Wang's database. The test results of a sample query image are displayed in the Figures 8, 9, 10, 11 and 12.

7.1 Main GUI (Graphic User Interface)

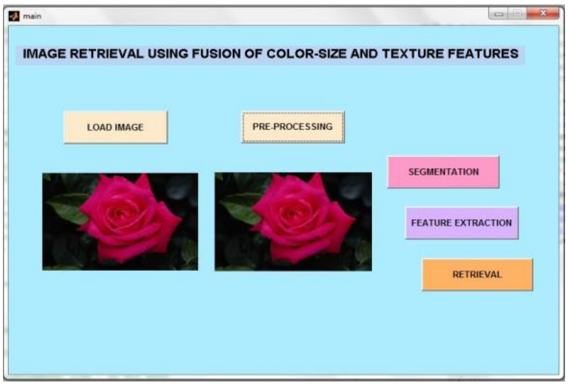
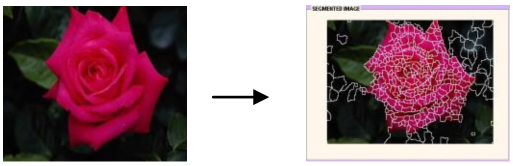


Fig 8: Main GUI for the proposed work

7.2 Watershed segmentation results



Input image Segmented image

Fig 9: Segmented result for given query image

7.2 Feature Extraction results

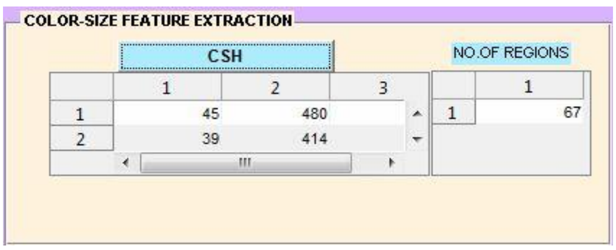


Fig 10: Extracted sample Color-size feature values

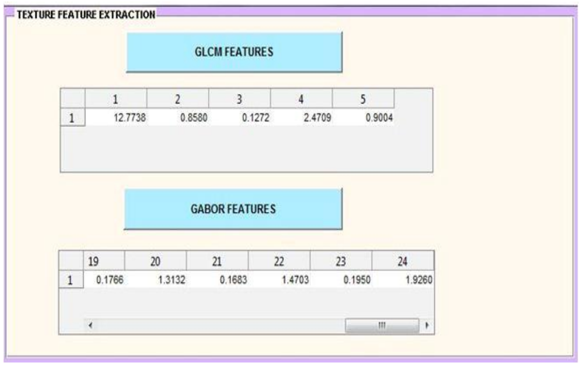
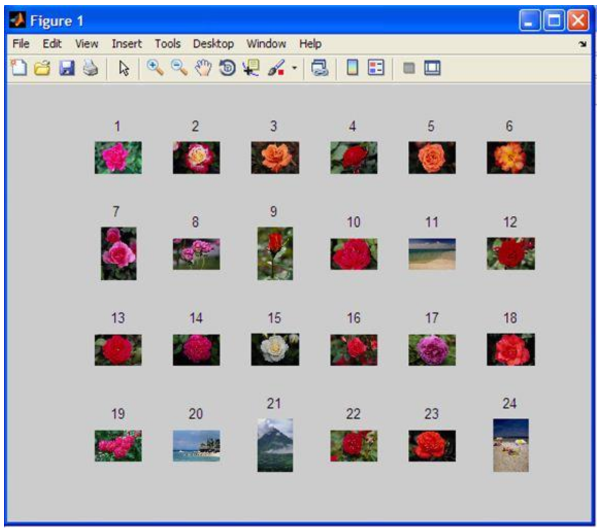


Fig 11: Extracted Sample Texture feature values

7.4 Image retrieval result for the query image

Input query image



Retrieved images for given query image

Fig 12: Image retrieval

7.5 Performance measures

The performance of a CBIR system can be measured in terms of its precision and recall. Precision measures the retrieval accuracy [16]; it is the ratio between the number of relevant images retrieved and the total number of images retrieved. Recall measures the ability to retrieve all relevant images in the database. It is the ratio between the number of relevant images retrieved and all of the relevant images in the database.

They are defined as follows:

$$\text{Precision (P)} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}$$

$$\text{Recall (R)} = \frac{\text{Number of relevant images retrieved}}{\text{Number of relevant images in database}}$$

The Table 1 shows the results of average precision and average recall for Color-Size features. The Table 2 shows the results of average precision and average recall for Texture features. The Table 3 shows the results of average precision and average recall for combined features.

Table 1.Average precision and recall for Color-Size features

Color - Size	CSH	
Category	Precision	Recall
Roses	0.80	0.69
Horses	0.77	0.71
Elephants	0.65	0.74
Dinosaurs	0.50	0.78
Buses	0.72	0.67
Mountains	0.71	0.69
Buildings	0.63	0.78

Table 2. Average precision and recall for Texture feature

Texture	Gabor		GLCM	
Category	P	R	P	R
Roses	0.68	0.60	0.53	0.67
Horses	0.62	0.65	0.48	0.71
Elephants	0.56	0.70	0.44	0.73
Dinosaurs	0.44	0.73	0.39	0.75
Buses	0.69	0.65	0.58	0.61
Mountains	0.70	0.63	0.55	0.64
Buildings	0.53	0.74	0.38	0.77

Table 3.Average precision and recall for combined features

Combined techniques	Average Precision (%)	Average Recall (%)
CSH + Gabor	80	63
CSH + GLCM	71	68

To evaluate the most efficient image retrieval measures, precision and recall scores are combined into a single measure of performance, known as the F-score. The formula for calculating the F-score is:

$$\text{F-score} = 2 * P * R / P + R \quad (12)$$

Accuracy rate for proposed system is calculated which is defined as follows:

$$\text{Accuracy} = (P + R) / 2 \quad (13)$$

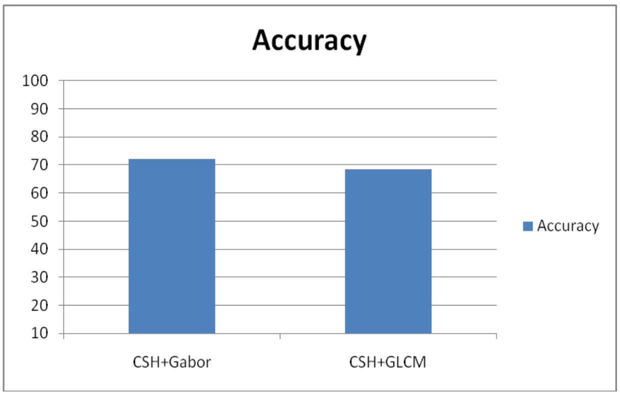


Fig 13: Chart showing Accuracy for fusion of CSH & Gabor and CSH & GLCM

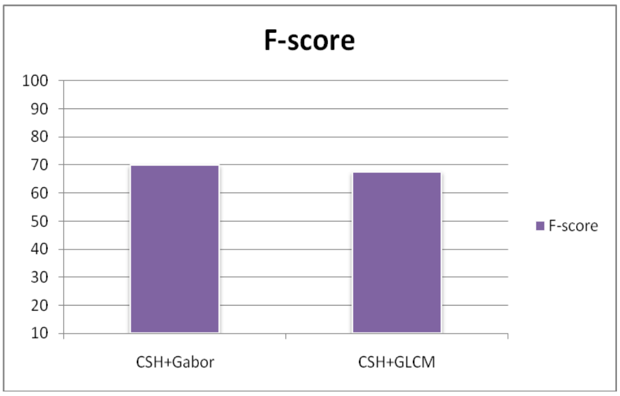


Fig 14: Chart showing F-score for fusion of CSH & Gabor and CSH & GLCM

8. CONCLUSION

The fusion of Color-size and Texture features has been proposed for developing CBIR. Initially the query image is pre-processed and segmented. From the segmented image the Color-Size Histogram (CSH) and Texture (Gabor, GLCM) features are extracted and stored as feature vector. The feature vector of query image is compared with the images stored in the database. The images with closest vector distance are retrieved as the similar results. The results are displayed on the basis of sorting of images with respect to minimum feature distance. The average Precision and Recall values are analyzed for individual as well as combined techniques. The test results show that the combination of feature extraction techniques yields better performance than individual techniques. Moreover, it is found that the fusion of Color-Size Histogram and Gabor wavelet performs well with in accordance with Accuracy rate and F-score.

9. ACKNOWLEDGMENTS

Our thanks to the expert named Aruna who have contributed towards development of the work.

10. REFERENCES

[1] Jagpal Singh, Jashanbir Singh Kaleka and Reecha Sharma , " Different Approaches of CBIR Techniques", International Journal of Computers and distributed Systems, Vol.1, Issue 2, 76-78, August 2012.

[2] A. W. M. Smeulders, M. Worring, S.Santini, A.Gupta and R.Jain, "Content-Based Image Retrieval at the End of the Early Years", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, Issue 12 , pp. 1349-1380, December 2000.

[3] Christos Faloutsos, Ron Barber, Myron Flickner, Jim Hafner, Wayne Niblack, Dragutin Petkovic, William Equitz, " Efficient and Effective Querying by Image Content ", Journal of Intelligent Information Systems, Vol. 3, No. ¾, pp. 231-262, 1994.

[4] Aman Chadha , Sushmit Mallik , Ravdeep Johar, "Comparative Study and Optimization of Feature-Extraction Techniques for Content base Image Retrieval", International Journal of Computer Applications Vol. 52 No.20, pp. 35-42,2012.

[5] Poulami Halder, Joydeep Mukherjee, "International Journal of Computer Applications", Vol.48, No.11, pp.25-31, June 2012.

[6] Mirnall Bansal, Gurpreet Kaur, Maninder Kaur, "Content Based Image Retrieval basd on Color", International Journal od Computer science and Technology", Vol.3, Issue 1, pp. 295 – 297, Jan. –March 2012.

[7] Kalyan Roy, Joydeep Mukherjee, "Image Similarity Measure using Color Histogram, Color Coherence vector and Sobel method", International journal of Science and Research", Vol.2, Issue 1, pp.538-543, January 2013.

[8] S.Selvarajah, S.R.Kodituwakku" Analysis and Comparison of Texture features for Content Based Image Retrieval", International Journal of Latest Trends in Computing", Vol.2, Issue 1, pp.108-113,March 2011.

[9] Swati.V.Sahhare, Vrushali G. Narse, "Design of feature extraction in Content Based Image Retreival(CBIR) using Color and Texture",International Journal of Computer Science & Informatics Vol. I, No.II, 2011.

[10] Kashif Iqbal, Michael O. Odetayo, Anne James,"Content Based Image Retrieval for biometric security using color ,texture and shape controlled by fuzzy heuristics", Journal of computer and system sciences", pp.1258-1277, October 2011.

[11] Zinat Afrose, "A Comparative Study on Noise Removal of Compound Images using Different Types of Filters", International Journal of Computer Applications Vol. 47, No.14, pp. 45-48, June 2012.

[12] Pawan Patidar, Manoj Gupta, Sumit Srivastava, Ashok Kumar Nagawat, "Image De- noising by various filters for different noise", International Journal of Computer Applications Vol. 9, No.4, pp. 45-50, 2010.

[13] Jayant S.Rohankar, Sachin U.Balvir , " A study of Content Based Image Retrieval (CBIR) System based on Correlation , Median Filtering and Edge Extraction" , International Journal of Electronics Communication and Engineering", Vol .4, No.1, pp. 251-256, 2011.

[14] Mohamed A.Helala, Mazen M.Selim, and Hala H.Zayed, "A Content Based Image Retrieval Approach based on Principal Regions Detection", International Journal of Computer Science Issues", Vol.9 ,Issue 4, No.1 , 204-213, 2012.

[15] Haralick R.M.Shanmugum K. and Dinstein I. (1973), Texture features for image classification, IEEE Transactions on System, Man and Cybernetics, SMC-3(6), broadatz Textures: A photographic album for Artists & designers, New York: Dover.

[16] Shankar M.Patil, "A Content Based Image Retrieval using color, texture and shape", International Journal of Computer Science and Engineering Technology", Vol.3 , No.9 ,pp.404-410, 2012.