

Correlated Networks for Content Based Image Retrieval

Aun Irtaza

*Department of Computer Science, National University of Computer & Emerging Sciences
Islamabad, Pakistan*

Email: aun.irtaza@gmail.com

M. Arfan Jaffar

*College of Computer and Information Sciences,
Al-Imam Muhammad ibn Saud Islamic University, Riyadh, Saudi Arabia*

Email: arfan.jaffar@ccis.imamu.edu

Eisa Aleisa

*College of Computer and Information Sciences,
Al-Imam Muhammad ibn Saud Islamic University, Riyadh, Saudi Arabia*

Email: aleisa@ccis.imamu.edu

Received 4 April 2012

Accepted 4 February 2013

Abstract

Efficient CBIR systems are based on three things, (1) how they represent the repository images in the form of signature; (2) how they measure the similarity of the database images with query image, (3) how they retrieve the semantically similar images in response of a query image. The paper is focusing on these three things. For signature development, curvelet transform, wavelet packets, and Gabor filters based signature development is introduced. For measuring the similarity, Pearson correlation is used as a distance measure; and for retrieving the semantically similar images in response of query images, Neural Network based architecture for content based image retrieval is presented. These things ensure the retrieval of images in a robust way. To elaborate the effectiveness of the presented work, the proposed method is compared with several existing CBIR systems, and it is proved that the proposed method has performed better than all comparative systems.

Keywords: CBIR, Pearson Correlation, Neural Network based Semantic Association.

1. Introduction

Modern image search engines which retrieve images on the base of actual visual contents are referred to as Content Based Image Retrieval (CBIR) systems. CBIR systems have found vast applications in many fields like architectural design, surveillance systems, geographical information systems, data mining, remote sensing, fabric design, medical image retrieval, Internet image search, video search, and communication systems [26]. Efforts are carried out for determining the ways through

which retrieval of semantically correct images in response of any query image can be assured. If we become able to develop an efficient and robust CBIR system we can overcome the linguistic problems as well which appears while describing and sharing images, as suffered by all keyword based image retrieval systems.

Main working theme of any CBIR system is to calculate the feature vector of repository images and characterize some visual properties against them. These feature vectors (also known as image signatures) serves

as feature database and are used for measuring similarity amongst the images in the image repository. User provides a query image, and the CBIR system computes feature vector for it, and then compare it with feature database using any similarity or dissimilarity metric like Manhattan distance. Most similar images are returned to the user as system response. But this approach has two main flaws due to which CBIR systems are unable to produce good results: (1) they rank the retrieved images on the base of distance or similarity with query image and generate the results, in this regard they do not verify their output. The problem with this approach is that many images may appear as the response images while they are not relevant at all, e.g. as shown in figure 1: due to the lack of verification, CBIR systems may bring dog's image in response of girl's image as both images are similar in features. (2) Secondly they do not consider similarity amongst neighbors of the query image for output finalization; therefore the generated output is very inconsistent, as probability of right semantic association on the base of single image (query image) is far less than the right semantic association probability of multiple images (top neighbors) [6]. Keeping the above mentioned points in mind; in this paper we focused on finding the ways through which such kind of shortcomings can be avoided and performance of CBIRs can be enhanced. A backpropagation neural network structure is proposed which is trained on sub repository of images generated from the main image repository and utilizes the right neighborhood of the query image. The aim of this training is to ensure the retrieval of semantically correct images in response of query images. To further enhance the performance of proposed system, a powerful feature extraction and neighbor selection way is introduced which is based on in-depth texture analysis of images. For this hybrid texture features are used which are the fused version of features obtained from curvelet transform of the images wavelet packets and Eigen values of Gabor filter. For neighbor selection Pearson correlation is used.

Rest of paper is organized as follow. Section 2 provides the details of related work done in the area of content based image retrieval. Section 3 provides the technical background required for signature development and similarity calculation; Wavelet Packets, Gabor Features, and Pearson Correlation are



Fig. 1. Visual features of girl's image and the dog image are very similar, but both images belongs from different semantic class. (Image is taken from <http://hair-craft.co.uk>).

discussed in detail. Section 4 is focusing on the proposed method introduced in this paper. Experimentation and results are covered in section 5. Finally we concluded in section 6.

2. Related Work

In the commercial domain, QBIC system by IBM [1] is considered as the first CBIR system. After QBIC system many additional systems were developed at IBM, NEC AMORA, VIRAGE, Interpix and Bell Laboratory [2]. Common ground of all proposed CBIR systems is their intention to address image searching problem in a more effective way by addressing new feature types and image similarity detection measures. For this the focus of research is on texture [11, 12, 13, 14], color [14, 15], or shape [16] features or any of these combinations [13]. As far as the scope of features is concerned, features can be divided into two main categories: global features and local features. Global feature category includes color histogram [3, 4], texture histogram [7], color layout [5] of the whole image, and features selected from multidimensional discriminant analysis of a collection of images [23]. While local feature category includes color, texture [9], and shape features for sub images, segmented regions [2,6], or interest points. Color histogram tells the global distribution of colors in the images. It is insensitive to small variations in the image structure. But, it has two major drawbacks. Firstly, they are unable to fully accommodate the spatial information, and also they are not unique. Images with similar color distribution and different object composition produce very similar

histograms. Moreover, similar images of same point of view carrying different lighting conditions create dissimilar histograms. Many researchers suggested the use of color correlogram for avoiding inconsistencies involving the spatial information [26]. Some other visual features are also proposed for content based image retrieval, such as salient points and spatial features. SIFT [24] and SURF [25] are the well known visual features based on salient points. These visual features have shown promising results for image retrieval, but due to the high dimensionality of SIFT and SURF based feature vectors, these features can cause over-fitting problem when employed with support vector machines for image retrieval purposes. In [24] a bag of features based model is presented by integrating SIFT and LBP features, and are combined with weighted k means clustering algorithm for image retrieval. In Rao et al. [29], as a first step image is divided into eight coarse partitions. After the coarse partitioning, the centroid of each partition is selected as its dominant color. Texture of an image is obtained by using GLCM (Gray Level Co-occurrence Matrix). Color and texture features of an image are normalized, and after that shape information is captured through edge images obtained by Gradient Vector Flow fields. The combination of the shape color and texture features of an image is provided as a feature set for image retrieval. Weighted Euclidean distance of color, texture and shape features of gray level edge images of RGB space is used in retrieving the similar images. But this method suffers from boundary delocalization and is not able to capture the contours effectively.

In [2] James Z. Wang et.al, introduced a semantics classification method which uses a wavelet based approach for feature extraction and then for the image comparison they used image segmentation based region matching approach. IRM proposed in this work is not efficient for texture classification due to uncertain modeling. So to address this issue, their idea was further processed by Yixin Chen et.al [6]. They introduced an un-supervised clustering based technique, which generate multiple clusters of retrieved results and give more accurate results as compare to the previous work, but their method suffers from issues like numbers of clusters identification, and segmentation uncertainty,

due to which the results of this technique are not reliable.

In [7] Michael Lama et.al, proposed a content based image retrieval system for computed tomography nodule images. Their system generates feature vector through GLCM, Gabor filters, and Markov random fields. On the extracted features they perform Euclidean, Manhattan and Chebychev distances to find the relevant images from image database. In the same way Kerstin Bunte et.al [8] has used CBIR in dermatology. They have used two different methods to learn favorable feature representations Limited Rank Matrix Learning Vector Quantization (LiRaMLVQ) and Large Margin Nearest Neighbor (LMNN). Both methods use labeled training data and provide discriminant linear transformation of the original features. But these techniques do not report the comparison with other techniques. Therefore it is a difficult thing to judge their effectiveness.

3. Signature Development and Similarity Calculation

The efficient retrieval in any CBIR system, is strongly dependent on the fact, that how efficiently the visual contents present in the image are represented in the form of image signature. The method we used for signature development works in three dimensions. In the first phase, it analysis the input image on the base of best nodes of wavelet packets tree; and as a second step it uses one approximation image to perform the detail Gabor analysis and selects a set of Eigen values to completely represent the Gabor features. As a third step it analysis the image through curvelet transform. Then by fusing these three types of feature vectors it generates the corresponding feature vector for image.

Another important thing for any CBIR system is, its ability to compute the similarity between images. For similarity calculation we used Pearson correlation; the benefit of this scheme is that it ranks the images in the range of most similar target images to most dissimilar target images. In this section we first present the details of wavelet packets decomposition, Gabor features, and curvelet transform of the images. After this, the details

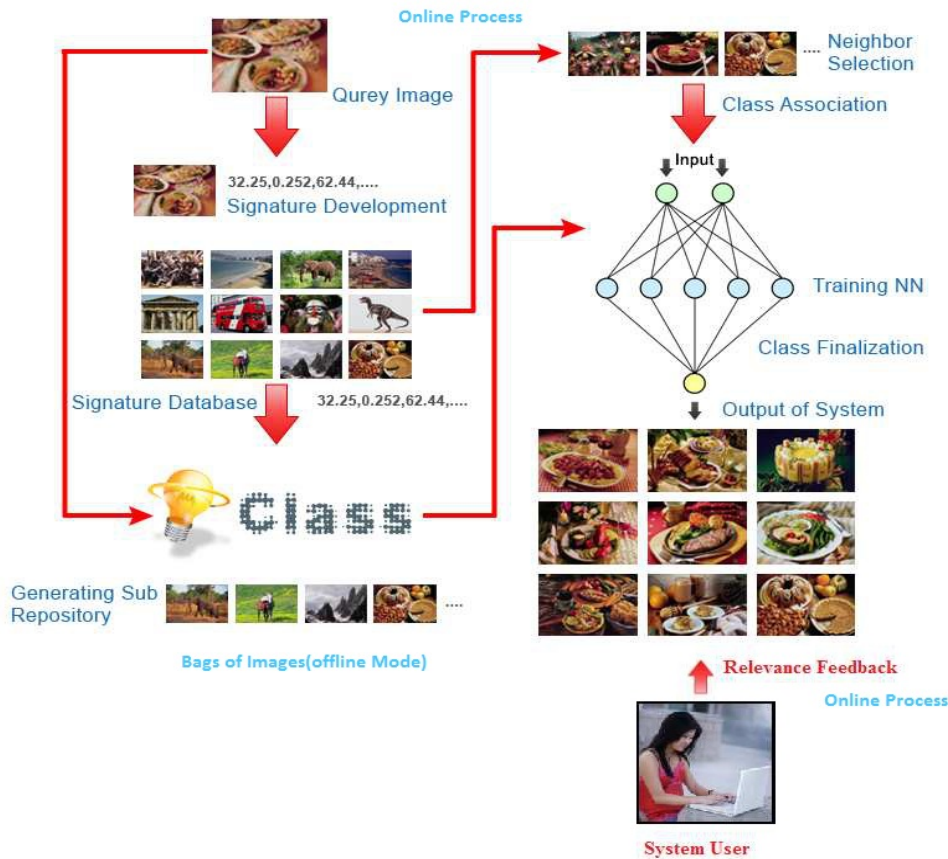


Fig. 2. Architecture of the proposed method

about the signature development come. Then details regarding similarity calculation are provided.

3.1 Wavelet packets

To analyze the signals in a rich way the wavelet packet method is used which is a generalization of wavelet decomposition and can be described by the collection of functions $\{WJ(x) | J \in \mathbb{Z}^+\}$ obtained from [11]: Where 'p' is a scale index and 'l' is the translation index. $W_0(x) = \Phi(x)$ is the scaling function. $W_1(x) = \psi(x)$ is the basic wavelet function [20]. h_k and g_k are the

$$2^{\frac{p-1}{2}} W_{2n}(2^{p-1}x - l) = \sum_m h_m - 2l^2 \frac{p}{2} W_n / (2^{p-1} - m) \quad (1)$$

$$2^{\frac{p-1}{2}} W_{2n+1}(2^{p-1}x - l) = \sum_m g_m - 2l^2 \frac{p}{2} W_n / (2^{p-1} - m) \quad (2)$$

quadratic mirror filters. Wavelet packets are well

localized in both time and frequency and thus provide an attractive alternative to pure frequency (Fourier) analysis. For a given orthogonal wavelet function, we obtain a library of bases called wavelet packet bases. Each of these bases offers a particular way of coding signals, reconstructing exact features and preserving global energy. The inverse relationship between wavelet packets of different scales can be shown through [11]:

$$2^{\frac{p}{2}} W_{2n+1}(2^{p-1}x - l) = \sum_l h_k - 2l^2 \frac{p-1}{2} W_{2n}(2^{p-1}x - l) + \sum_l g_k - 2l^2 \frac{p-1}{2} W_{2n+1}(2^{p-1}x - l) \quad (3)$$

Equation (3) can be used to calculate the wavelet packets. Coefficients of coarser scale can be calculated using eq. 1 and eq. 2 as

$$S_{2n,l}^{p-1} = \sum_m hm - 2lS_{n,m}^p \tag{4}$$

$$S_{2n+1,l}^{p-1} = \sum_m gm - 2lS_{n,m}^p \tag{5}$$

The main difference between normal wavelet decomposition and wavelet packets decomposition is that despite of just splitting the approximation components wavelet packets decomposes the detail components as well. So by this richest analysis becomes possible. Wavelet packets procedure results in a large number of decompositions and its explicit enumerations are unmanageable. So it is necessary to find the optimal decomposition with respect to some reasonable criterion. One convenient criterion can be the selection of tree nodes on the base of best entropy values. In this paper we used Shannon entropy measure to calculate the entropy. This can be calculated as:

$$E(S_i) = - \sum s_i^2 \log(s_i^2) \tag{6}$$

Using the Shannon entropy the optimal or the best tree can be calculated using the following scheme. A node N will be split into two nodes N1 and N2 if and only if the sum of the entropy of N1 and N2 is less than the entropy of N. This is a local criterion based only on the information available at the node N. It will lead in the form of a tree which is of much smaller size than the actual tree.

3.2 Gabor features

Gabor filter is a sinusoid function modulated by a Gaussian and extracts feature information from an image in the form of a response images by applying varying parameters [7]. Such filters after convolution generate the set of response images which are based on different frequencies and orientations; and from these response images we generate the representative feature vectors based on salient points. In our implementation the filters we used are defined by the following equation [7]:

$$G(x, y) = e^{(\frac{x_\theta^2 - \gamma^2 y_\theta^2}{\alpha^2} + \frac{2\pi x_\theta i}{\lambda})} \tag{7}$$

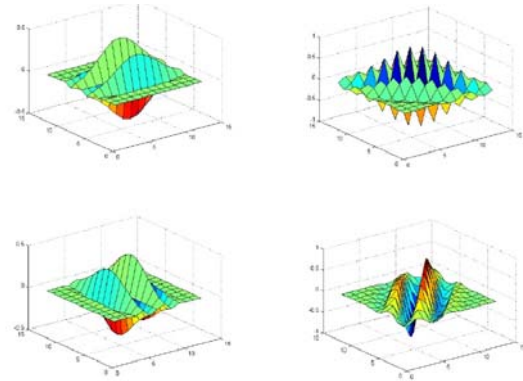


Fig. 3. Gabor filters visualization with different Parameters

Where

$$x_\theta = x \cos\theta + y \sin\theta \tag{8}$$

$$y_\theta = -x \sin\theta + y \cos\theta \tag{9}$$

And σ is the standard deviation of the Gaussian function, λ is the wavelength of the harmonic function, θ is the orientation, and γ is the spatial aspect ratio which is left constant at 0.5. The spatial frequency bandwidth is the ratio σ/λ and is held constant and equal to .56. Thus there are two parameters which changes when forming a Gabor filter θ and λ . The input image is divided into 9x9 non-overlapping regions. The Gabor filter is then convolved with different parameters; and it will generate the response images. As per the work done by Michael Lama et.al [7] we are using only the odd component of the Gabor filter which does not produce imaginary output:

$$G_o(x, y) = \exp(\frac{x_\theta^2 - \gamma^2 y_\theta^2}{\alpha^2}) \sin(\frac{2\pi x_\theta i}{\lambda}) \tag{10}$$

We convolve the image with 12 Gabor filters tuned to four orientations (θ) and three frequencies ($1/\lambda$). Orientation varied from 0 to $3\pi/4$ (stepping by $\pi/4$) and frequency varied from .3 to .5 (stepping by .1, figure 3).

3.3 Curvelet transform

Curvelet transforms are the extension of the ridgelet transform to multiple scale analysis, and covers the complete spectrum of frequency. By tuning ridgelet into different orientations and scale we can create

curvelets i.e. given an image function $f(x,y)$, the continuous ridgelet transform is given as [26]:

$$\mathfrak{R}_f(a, b, \theta) = \int \int \psi_{a,b,\theta}(x, y) f(x, y) dx dy \quad (11)$$

where $a > 0$ is scale, $b \in \mathbb{R}$ is the translation and $\theta \in [0, 2\pi]$ is the orientation. The ridgelet is defined as [23]:

$$\psi_{a,b,\theta}(x, y) = a^{-\frac{1}{2}} \psi\left(\frac{x \cos \theta + y \sin \theta - b}{a}\right) \quad (12)$$

Ridgelets are oriented along angle θ and is constant along lines $x \cos \theta + y \sin \theta = \text{const}$. A ridgelet is linear in edge direction and is much sharper than a conventional sinusoid wavelet [23]. Now if we compare it with 2-D wavelets transform then according to the wavelet domain:

$$\psi_{a1,a2,b1,b2}(x, y) = a1^{-\frac{1}{2}} a2^{-\frac{1}{2}} \psi\left(\frac{x - b1}{a1}\right) \psi\left(\frac{y - b2}{a2}\right) \quad (13)$$

By comparison we can see that, the ridgelet is similar to the 2-D wavelet except that the point parameters ($b1, b2$) are replaced by the line parameters (b, θ). In other words, the two transforms are related by:

Wavelet: Ψ scale, point-position

Ridgelet: Ψ scale, line-position

So this means that ridgelet can be tuned to different orientations and different scales to create the curvelets, the benefit of this scheme is to have a complete cover of the spectrum in frequency domain. The digital curvelet transform is taken on a 2-D Cartesian grid $f[m,n]$, $0 \leq m < M, 0 \leq n < N$ using following equation:

$$C^D(a, b, \theta) = \sum_{0 \leq m < M, 0 \leq n < N} f[m, n] \psi_{a,b,\theta}^D[m, n] \quad (14)$$

$< M, 0 \leq n < N$ using following equation:

3.4 Signature development

For the signature development, we computed the complete Shannon entropy based wavelet packets tree of the repository images up to the 3rd level. This results in the form of 64 nodes of wavelet packets tree. But we are concerned with only those nodes which have the best entropy values. So for this purpose we generated the best entropy valued tree as corresponding best tree as shown in figure 4. Nodes of the best tree are used for the wavelet packets feature generation.

Smallest approximation image of best tree is also used for Gabor analysis after computing the response images,

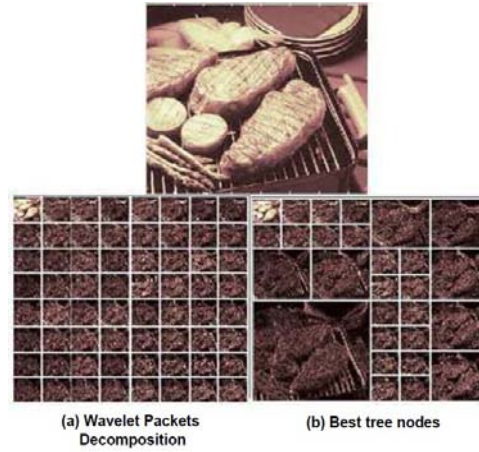


Fig. 4. Top image is the input Image and bottom images are the representation of corresponding full wavelet packets tree decomposition and nodes obtained from best tree of the wavelet packets tree.

according to the method already described. For signature development, the scheme we applied works as follows:

- We obtain the best entropy based wavelet packets tree, and twelve Gabor based response images after parameters application.
- On these tree nodes and response images, we obtain the eigenvector [30] corresponding to every best tree node, and Gabor response image. Eigenvectors are the linear transformations of the 2D squared matrices [30] and are widely used in computer vision applications like face and voice recognition and feature vector dimensionality reduction [31]. An eigenvector of a square matrix is a non-zero vector that is when multiplied by the matrix will yield a vector that will differ from the original at most by a multiplicative scalar [30]. The benefit of the scheme is that, the multiple 2D best tree nodes, and Gabor response images will be represented by 1D vector.
- The size of the wavelet packets feature, and Gabor response features in the form of eigenvectors are still very large; therefore we take the mean of every eigenvector and merge it in one vector. This vector serves for us the corresponding Gabor feature vector.

So after the application of these steps, we will have feature representation for wavelet packets transform and Gabor filters.

Now we use following method to take the representation of images through curvelet transform via wrapping.

- Curvelet transform represent images in the form of multiple bands and sub-bands. For every sub-band we calculate variance.
- For every band we take the mean of the variances of sub-bands of that band.
- We collect the mean values in one vector which represent the curvelet transform of image.

Aggregation of all these features in a single vector represents signature for any image.

3.5 Similarity calculation

For detecting the similar images from image repository we used Pearson correlation method. The reason for our choice is that Pearson correlation has the ability to describe linear relationship between two entities. Through Pearson correlation, it is possible to measure that, two entities are related with each other to which extent. Pearson correlation is widely used in clustering to measure the degree of association between records, but it is not common to use it in image classification and retrieval purposes. The similarity results are in the range of -1 to +1, correlation of -1 suggests that two records (images in our case) are entirely opposite. While 0 as the correlation value suggests that there is no correlation between these two records. While +1 suggests that this is the similar record as that of the query. Pearson correlation can be computed using following formula:

$$r = \sqrt{\frac{\sum xy - \frac{\sum x \sum y}{n}}{(\sum x^2 - \frac{(\sum x)^2}{n})(\sum y^2 - \frac{(\sum y)^2}{n})}}$$

Where 'x' is signature of query image and 'y' is the signature of repository element. Both 'x' and 'y' are the vectors. The operation will be performed on the base of respective elements in both vectors.

4. Proposed Method

This section starts with the discussion of the general framework which we introduced in this paper. Then neural network based semantic association of the images is introduced. This is followed by a discussion of the parameter selection and retrieval method.

4.1 Overview of technique

We perform two types of retrievals from signature repository for any query image. First type of retrieval is based on its K nearest neighbors; and second type of retrieval is based on its associated semantic class. The reason we used these two types of retrievals is that in case if our semantic class is matched wrongly; the retrieval system shouldn't put us in a complete blind state where every associated result is a wrong result. So in such kind of situation semantic type can be compromised with object composition which is present in the query image. The input of the algorithm consists of $R \geq 2$ example images in every semantic class. On these example images we define a backpropagation neural network structure for every class which is trained on these example images. Then with the help of neural network association for query image and its 'K' neighbors, we finalize the semantic class for query image. After semantic association, images from retrieved semantic class are ranked on the base of similarity with respect to query image. The system also ranked the 'M' closest images on the base of similarity with query image; and as the output system returns these two types of the clusters. In which one will be the most representative one.

4.2 Semantic association using Neural Networks

For associating repository images with their actual semantic class; we generated the sub repository from main image repository having images from 'M' known classes, and every class contains $R \geq 2$ images. The value of 'R' can vary for every class but ideally it should be same for all classes. On this sub repository we define class specific neural networks; which are trained with the concept of one against all classes (OAA) to classify the input images according to their semantic class. For training purposes for a specific class we define training set $\Omega_{tr} = \Omega_{pos} \cup \Omega_{neg}$ where Ω_{pos} means those images which belong from that class for which network is defined and are labeled with '1'; and Ω_{neg} means all other images in the sub repository which do not belong from that specific class and are labeled with '0'. In this way we define training sets for all classes and train neural networks upon them. All repository images are tested against all trained neural networks and on the base of decision function and association rules we

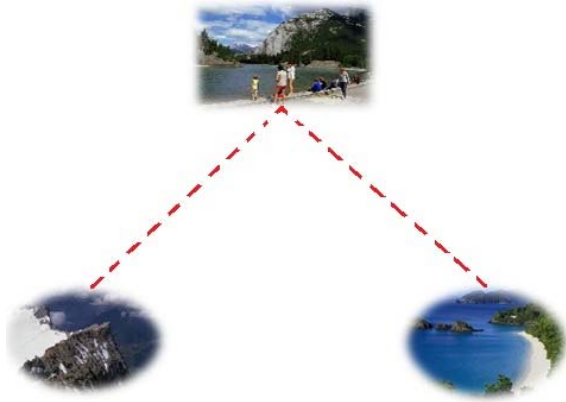


Fig. 5. Composition of objects suggest that input image can be associated with mountains and beach classes associate them with their specific semantic class. Due to the object composition present in the images many images may tend to belong from more than one semantic class as shown in figure 5. So in this case it is possible that decision function may associate them with undesired semantic class; so it is required that the semantic association process should be further enhanced. This is the reason that, for the finalization of semantic class for any image we do consider its top 'K' neighbors as well, and according to the majority voting rule (MVR), semantic class of the input image is finalized. MVR is represented as [22]:

$$C^*(x) = \text{sgn} \left\{ \sum_i C_i(X) - \frac{K-1}{2} \right\} \quad (15)$$

Where $C_i(X)$ is the class wise association factor of input image and its top neighbors. MVR counts the largest number of classifiers that agree with each other [22]. So according to equation (15) the class of input image is one which has a maximum combined association value for itself and its nearest neighbors, as represented in

$$C^*(x) = \text{argmax} \left(\sum_i m^* \right) \quad (16)$$

equation (16). Process of class finalization is graphically represented in figure 6. The process of semantic association is applied to all repository images and the results of association are stored in a file on disk. So by this way we prepare a database of semantic associations, which will be further utilized to generate the system output against query images.

4.3 Parameter selection

To implement the technique presented in this paper, right parameter selection is very important. As we are using neural networks for classifying images into multiple semantic classes, so performance of network is strongly dependent on the parameters which we choose for training purposes. While selecting the parameters we have to keep this thing before us that due to the parameters, computational time shouldn't increase in a way that it cannot be used in real time applications e.g. in case of deep neural networks or more neurons in hidden layer the time required to train the network is far more than the networks having less neurons and same capability. For the finalization of semantic class against any query image we are using top 5 images as the closest neighbors to make the decision of semantic association. These images are tested against class specific neural networks. The structure of these neural networks can vary for every class. But the common thing between all of these neural network structures is that, all of them have single hidden layer of neurons. In our implementation, number of neurons in hidden layer are 20, and μ from 0.01 to 0.0001 for all classes. Performance of a specific network can be increased by changing the parameters but this change may affect other classes. Therefore we have to choose the parameters which are most optimal. The performance goal is met with the help of neighbors and class finalization process which is already described. Class specific network performances measured on mean squared error with number of neurons in hidden layer and MSE goal are plotted in figure 7.

4.4 Content Based Image Retrieval

It is possible that for some complex images, our semantic association system may associate them with wrong semantic class. So there should be a mechanism through which such kind of situation can be avoided. To achieve this goal, the system returns us two image sets in the form of representative images. In which one image represents the retrieved semantic class; and other image represents the set of images which are supposed to have the same object composition as that of the query image. The benefit of this approach is that our system will never stuck in a blind state where every retrieved result is a wrong result (figure 9).

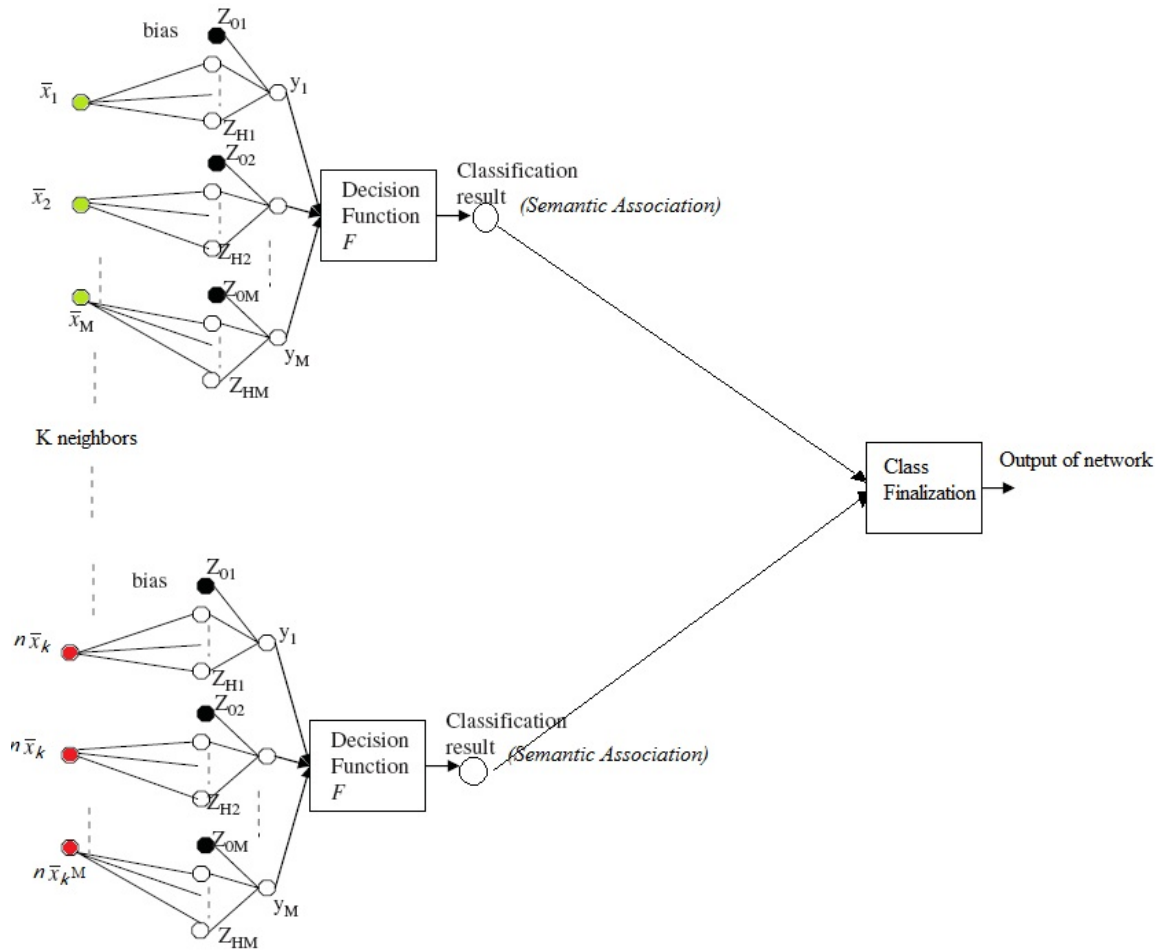


Fig. 6. Process of semantic association and class finalization.

The top matched image from the set of 'K' nearest neighbors represents the set of images suppose to have the same object composition. While most similar image from the sub repository of the obtained semantic class is returned to represent the semantic class. User has the option to see both sets of images having equal images. System also enables relevance feedback upon the images which appear similar to the query image in terms of distance. Relevance feedback gives freedom to the user to guide image retrieval system in case of complex queries to achieve desired output. Therefore in the situations where user is not looking for a particular class of images, and wants to seek the answer of particular problem in mind; can also utilize image retrieval system according to her search intentions.

5. Experiment and Results

This section provides the details of experiments performed to test our technique. The section is composed of following subsections. Section 5.1 gives the details of the image database we used for experimentation. Section 5.2 aims to illustrate the performance of the system through query examples. Section 5.3 elaborate the effectiveness of proposed system through precision and recall. Section 5.4 is about relevance feedback.

5.1 Database description

For designing an image retrieval system, selection of a suitable image database is a critical and important step. At the present time, there is not a standard image

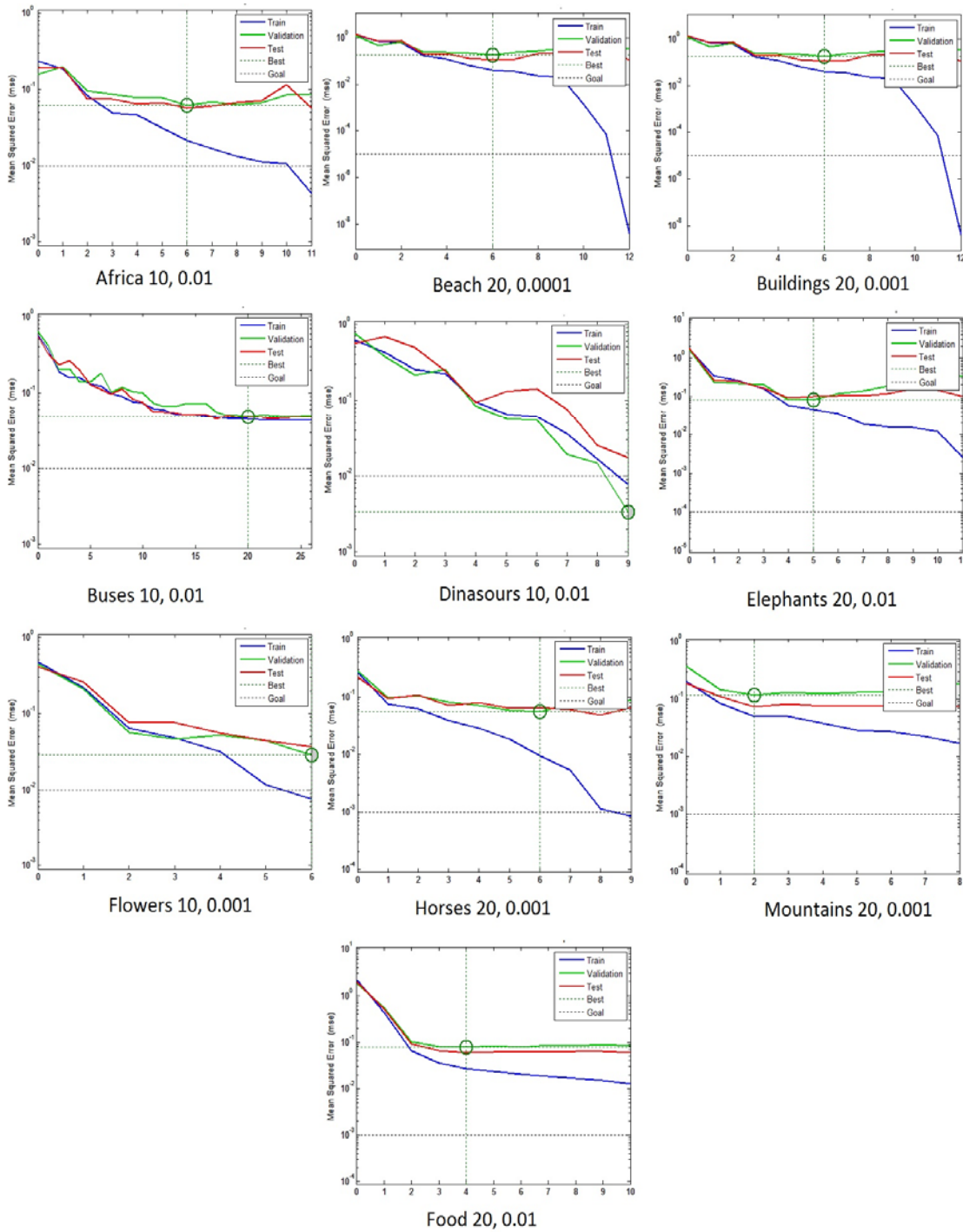


Fig. 7. Class specific Neural Networks training performance graph. Plotted mean squared error(MSE).

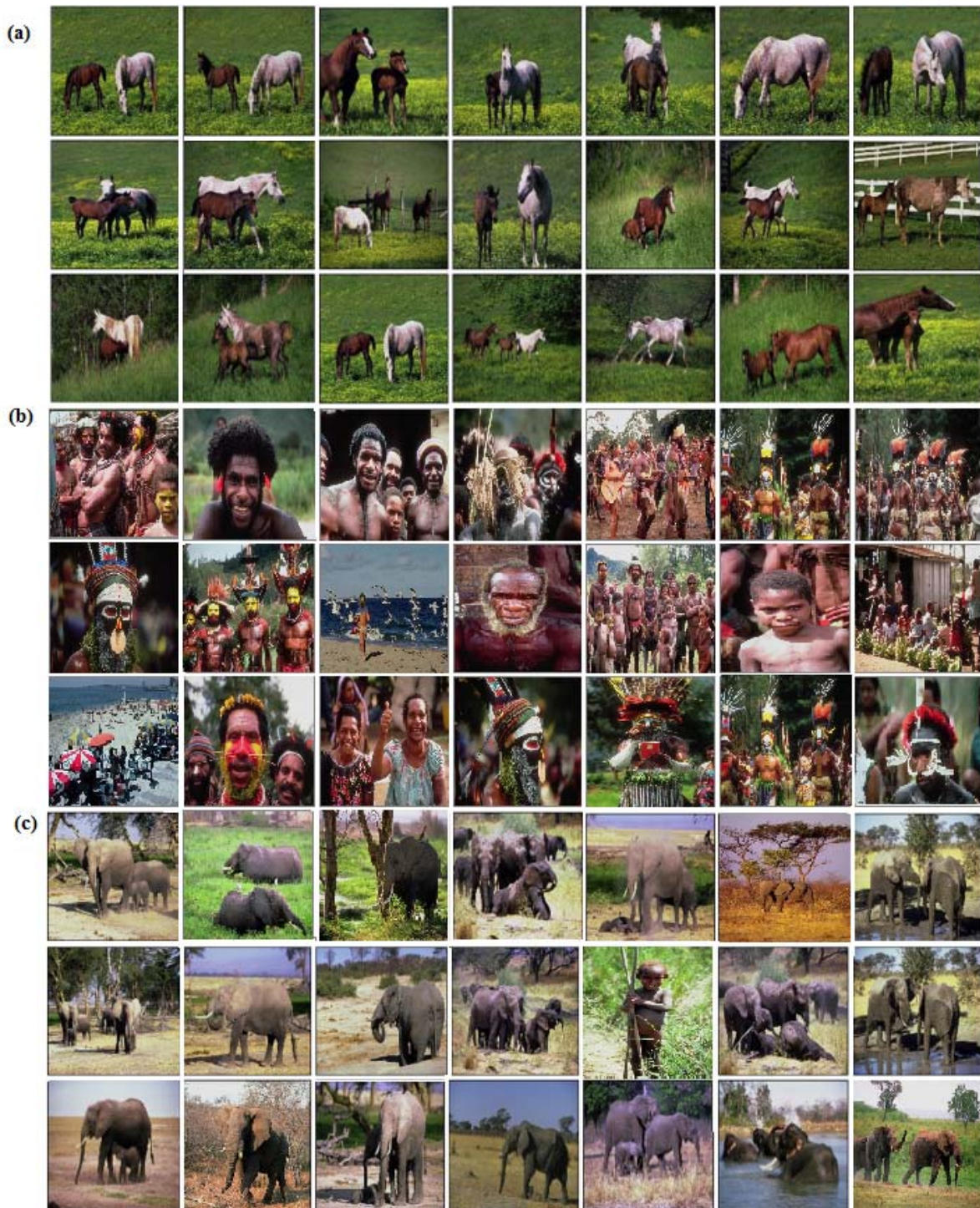


Fig. 8. Image retrieval results for class Horses, Africa, and Beach. 1st image in every group is the query image and all other images in that group are retrieved results.



Fig. 9. Some snapshots of the results, Left column has semantic class association results, while Right column has K nearest neighbors result. Top image is the query image and bottom images are the retrieved images.



Fig. 10. Sample images of each category of Corel dataset.

database for this purpose [14]; and also there is no agreement on number of images and type of images in the image database [14]. As CBIR systems are if not developed for domain specific applications, then it must include various semantic groups in the image database. This is the reason that we used COREL image

dataset for this purpose. Corel dataset is widely used in literature for image retrieval purposes [6]. Corel dataset covers a wide range of semantic groups from natural scenes to artificial objects for CBIR experiments. Dataset has 1000 images, and is partitioned into ten semantic categories namely African people and village, beach, buildings, buses, dinosaurs, elephants, flowers, horses, mountains, and food (Figure 10). Each category has 100 images, and size of each image is either 384X256 or 256X384. Partitioning of the dataset into semantic categories is determined by the creators and reflects the human perception of image similarity [36]. To further elaborate the performance of the proposed method, experiments are also carried out on Columbia object image library (COIL) [26]. Dataset has 7200 images from 100 different categories. A sample of the COIL dataset is presented in figure 14.

5.2 Query examples

As a first step for the implementation of proposed method, gray-scale versions of repository images are generated. This pre-processing is required to perform the feature extraction in a cost effective way. Therefore as a backend process feature extraction is performed on gray-scale versions of images and to display the results we display the color versions of the retrieved images. Each image in the image dataset is tested against the trained neural networks and its semantic class is determined. Results of semantic association are stored in a file which serves as the association database. For performance evaluation of proposed system, one image from horses, Africa, and elephant categories is randomly selected from Corel dataset; and then retrieved results are displayed against them. Query image results are displayed in figure 8.

The response of the system can be observed from the number of correctly retrieved images in response of query images. Therefore from query results it can be

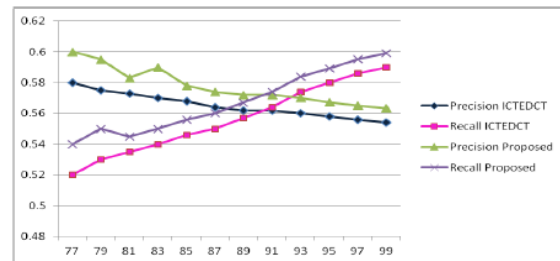


Fig. 11. Precision and Recall results on Coil dataset..

Table 1. Comparison of average precision obtained by proposed method with other standard retrieval systems

[28, 6, 18, 29, 27].

Class	Proposed Method	CTDCIRS	CLUE	CTCHIRS	Babu Rao	EMSVM
Africa	0.66	0.56	0.49	0.68	0.42	0.50
Beach	0.59	0.54	0.37	0.54	0.39	0.70
Buildings	0.64	0.61	0.37	0.56	0.43	0.20
Buses	0.84	0.89	0.64	0.88	0.65	0.80
Dinosaurs	0.93	0.98	0.95	0.99	0.97	0.90

Table 2. Comparison of average recall obtained by proposed method with other standard retrieval systems

[28, 6, 18, 29, 27].

Class	Proposed Method	CTDCIRS	CLUE	CTCHIRS	Babu Rao	EMSVM
Africa	0.13	0.11	0.10	0.14	0.08	0.10
Beach	0.12	0.11	0.07	0.11	0.08	0.14
Buildings	0.12	0.12	0.07	0.11	0.08	0.04
Buses	0.17	0.17	0.13	0.13	0.13	0.16
Dinosaurs	0.19	0.20	0.19	0.20	0.19	0.18

easily observed that performance of system is very good.

5.3 Retrieval precision and recall evaluation

To evaluate the effectiveness of the proposed method, we determined how many relevant images are retrieved in response of a query image. For this retrieval effectiveness is defined in terms of precision and recall rates. Precision is used to measure specificity of image retrieval, and recall is used for measuring sensitivity or true positive rate. Experiments results are reported after running five times on ten query images randomly selected from each image category. For each query image, relevant images are considered to be those images only, which belong to the same category as that of the query image. Based on this concept, the retrieval precision and recall are defined as:

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total number of images retrieved}}$$

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{Total number of relevant images}} \quad \text{Top 20}$$

retrieved images are used to compute the precision and recall rates.

5.3.1 Comparison results on Corel dataset

In order to show the superiority of proposed technique, it is compared with CTDCIRS [28], CLUE [6], CTCHIRS [18], Babu Rao [29], and EMSVM [27]. The results of comparative techniques are obtained from the original research work reported by the corresponding authors. Table 1 describes the class wise comparison of the proposed system with comparative systems in terms of average precision values. Same results are

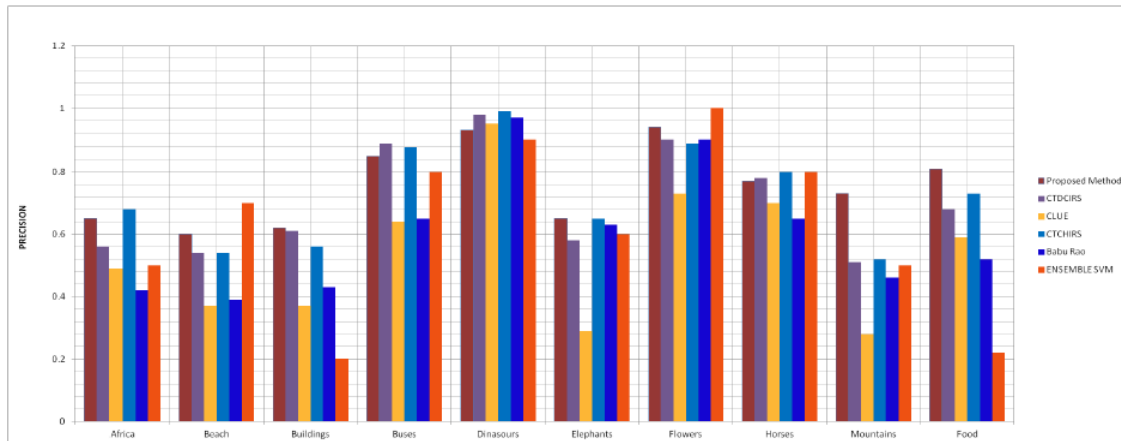


Fig.12. Comparison of average Precision obtained by proposed method with other standard retrieval systems [28, 6, 18, 29, 27].

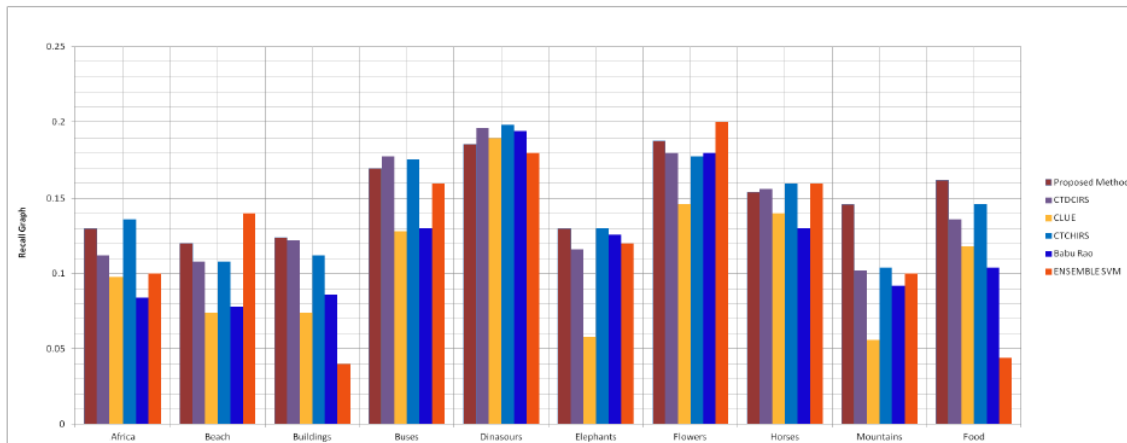


Fig.13. Comparison of average recall obtained by proposed method with other standard retrieval systems [28, 6, 18, 29, 27].

graphically illustrated in fig 12. Similarly Table 2 presents comparison of the proposed system with comparative systems in terms of average recall values. Recall results are graphically illustrated in fig 13. From results it can be observed that, there is no system which is giving highest results in all categories. But in terms of overall accuracy our proposed system has outperformed all other systems.

5.3.2 Comparison results on COIL dataset

We compared the proposed method with ICTEDCT [26] on Coil dataset. Five images are selected from each image category and then performance of both systems is compared against each category. From the results elaborated in figure 11, it can be clearly observed that

proposed method is giving higher recall and precision rates as compare to ICTEDCT [26]. Hence from the results of proposed method on Coil and Corel datasets, we can say that proposed method is much more precise and effective as compare to other CBIR systems.

5.4 Relevance Feedback

Relevance feedback (RF) system we proposed generates the initial output against a query image on the base of Pearson correlation. For this it returns a set of images to the user which appears close to the query image in terms of distance as described in section 4. User gives initial feedback by selecting only positive images in returned output, rest of the images in the returned output are considered as negative feedbacks. We train



Fig.14. A sample of coil database[26].

backpropagation neural networks on these two classes of inputs and output is generated. User gives feedback and process continues until user become satisfied from the output. In our experiments, we randomly selected 100 images from Corel dataset and then relevance feedback is automatically performed by the computer simulation. All query relevant images (i.e., images with the same concept as the query) are marked as positive feedback samples and all the other images are marked as negative feedback samples. Method is tested on top 20 returned images. We used 9 iterations of RF for our experiments; in which the 0th iteration returns the results obtained through Pearson correlation. The performance of the system has been measured as averaged precision, and recall values (Fig. 15). Precision and recall curves are the average values of 100 queries. The precision curve evaluates the

effectiveness of a given algorithm and recall curve evaluates the robustness of the algorithm.

6. Conclusion

The paper has introduced an image retrieval system, which is based on the concept of semantic class association through trained neural networks. For an efficient image retrieval system, it is necessary that it must focus on three main things. One is to uniquely represent the thumb impact of repository images which is the image signature; secondly to measure the similarity with other repository images it must have an efficient similarity calculation way; and lastly it must be able to retrieve the image results which are semantically similar to the query image. So in this paper we focused on these three key issues. Wavelet packets, Gabor filters, and curvelet transform based signature

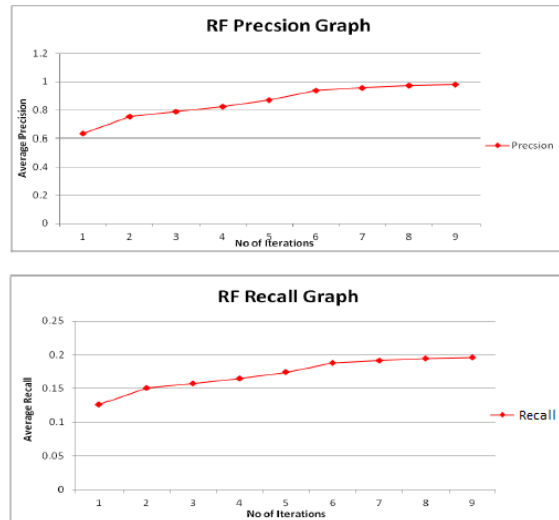


Fig. 15. Precision and Recall graphs for Relevance feedback. development is introduced. For similarity detection, Pearson correlation based similarity calculation is introduced; and for efficient image retrieval, neural networks based technique is introduced which also utilizes the right neighborhood of the query images. According to which semantic association occurs in a more systematic way. To avoid the risk of mis-association relevance feedback is also incorporated. The results of the proposed method are compared with several standard retrieval systems, and through experiments it is proved that, proposed system has outperformed all other systems in terms of average precision and recall values.

Acknowledgements

This research work is supported by Higher Education Commission, Pakistan.

References

1. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Domet al., Query by Image and Video Content: The QBIC System. *IEEE Computer* **28**, (1995).
2. James Z. Wang, Jia Li, Gio Wiederhold, SIMPLicity: Semantics-Sensitive Integrated Matching for Picture Libraries. *IEEE Transactions on pattern analysis and machine intelligence* **09** (2001)

3. C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz, Efficient and effective querying by image content. *J. Intell. Inf. Syst* **3** (1994) 231-262.
4. A. Gupta and R. Jain, Visual information retrieval. *Commun.. ACM* **40** (1997) 70-79.
5. J. R. Smith and S.-F. Chang, VisualSEEK: a fully automated content based query system. *In: Proc. 4th ACM Int. Conf. Multimedia* (1996), pp. 87-98.
6. Yixin Chen, James Z. Wang,, Robert Krovetz, CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning. *IEEE Transactions on Image Processing* **14** (2005).
7. Michael Lama, Tim Disney et al., Content based image retrieval for pulmonary computed tomography nodule images. *In: SPIE Medical Imaging Conference*, (San Diego, 2007)
8. Kerstin Bunte, Michael Biehl, Marcel F. Jonkman, Nicolai Petkov, Learning effective color features for content based image retrieval in dermatology. *Pattern Recognition* **44** (2011) 1892-1902.
9. G. Quellec, M. Lamar et al., Wavelet optimization for content-based image retrieval in medical databases. *Medical Image Analysis* **14** (2010) 227-241.
10. J.Z. Wang, G. Wiederhold, O. Firschein, and X.W. Sha., Content Based Image Indexing and Searching Using Daubechies' Wavelets. *Int'l J. Digital Libraries* **4** (1998) 311-328.
11. Andrew Liane, Jian Fan, Texture classification by wavelet packets signature. *IEEE Transactions on pattern analysis and machine intelligence* **15** (1993).
12. Tomasz Andrysiak, Michal Chora's., Texture Image retrieval based on hierarchal Gabor filters. *Int. J. Appl. Math. Comput. Sci.* **15** (2005) 471-480.
13. N Gnaneswera Rao, Dr. V. Vijaya Kumar, V. Vinkata Karishna., Texture based image indexing and retrieval. *IJCSNS International Journal of Computer Science and Network Security* **09** (2009)
14. Dengsheng Zhang, Improving image retrieval performance by using both color and texture features. *IEEE conference on Image and Graphics* (2004)
15. Zhang Lei, Lin Fuzong, Zhang Bo, A CBIR method based on color-spatial feature. *In: TENCON*

99. *Proceedings of the IEEE Region 10 Conference* (2002)
16. Xiang Sean Zhou, Thomas S. Huang, Edge based structural features for content based image retrieval. *Pattern Recognition Letters* **22** (2001) 457-468
 17. Wei Bian, Dacheng Tao, Biased Discriminant Euclidean Embedding for Content-Based Image Retrieval. *IEEE Transactions on Image Processing* **19** (2010)
 18. Chuen-Horng Lin, Rong-Tai Chen, Yung-Kuan Chan, A smart content-based image retrieval system based on color and texture feature. *Image and Vision Computing* **27** (2009) 658-665
 19. N. Jhanwar, S. Chaudhurib, G. Seetharamanc, B. Zavidovique., Content based image retrieval using motif co-occurrence matrix. *Image and Vision Computing* **22** (2004) 1211-1220
 20. André Tavares da Silva, Alexandre Xavier Falcão, Léo Pini Magalhães, Active learning paradigms for CBIR systems based on optimum-path forest classification. *Pattern Recognition* **44** (2011) 2971–2978
 21. Ishrat Jahan Sumana, Md. Monirul Islam, Dengsheng Zhang and Guojun Lu, Content based image retrieval using curvelet transform. In: *IEEE 10th workshop on multimedia and signal processing*, (Gippsland Sch. of Inf. Technology Monash University, Churchill, VIC 2008)
 22. Dacheng Tao, Xiaoou Tang, Xuelong Li, Xuelong Li, Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. *IEEE Transactions on pattern analysis and machine intelligence* **28** (2006) 1088–1099
 23. Wei Bian, Dacheng Tao, Biased Discriminant Euclidean Embedding for Content-Based Image Retrieval. *IEEE Transactions on Image Processing* **19** (2010)
 24. Xiali Yuan, Jing Yu, Zengchang Qin, and Tao Wan: A SIFT-LBP image retrieval model based on bag of features. *IEEE International Conference on on image processing* (2011).
 25. K. Valmunurang et. al., Content based image retrieval using SURF and color moment. *Global journal of computer science and technology* **11** (2011)
 26. Youssef SM., ICTEDCT-CBIR: Integrating curvelet transform with enhanced dominant colors extraction and texture analysis for efficient content-based image retrieval. *Comput Electr Eng* (2012), <http://dx.doi.org/10.1016/j.compeleceng.2012.05.010>
 27. Ela Yildizer, Ali Metin Balci, Mohammad Hassan, Reda Alhajj, Efficient content-based image retrieval using Multiple Support Vector Machines Ensemble. *Expert Systems with Applications* **39** (2012) 2385–2396
 28. BabuRao M, Rao BP, Govardhan A. CTDCIRS: content based image retrieval system based on dominant color and texture features. *Int J Comput Appl.*, **18** (2011) 0975–8887
 29. Babu Rao M, Prabhakara Rao B, Govardhan A., Content based image retrieval system based on dominant color, texture and shape. *Int J Eng Sci Technol (IJEST)*. **4** (2011) 2887–2896
 30. Lucas Parra, Paul Sajda: Blind separation via generalized eigenvalue decomposition, *Intl. Journal of machine learning research* **4** (2003) 1261--1269
 31. Dr. H.B.Kekre, Sudeep D. Thepade, Akshay Maloo, CBIR Feature vector dimension reduction with eigenvectors of Covariance Matrix using Row, Column and Diagonal Mean Sequences, *International Journal of Computer Applications*. **3** (2010)
 32. Chih-Chin Lai, Ying-Chuan chen: A user-oriented image retrieval system based on interactive genetic algorithm, *IEEE Trans. Instrumentation and Measurement.*, **60** (2011) 3318–3325.