

CBIR System Based on Prediction Errors

A. MOHAMED UVAZE AHAMED, C. ESWARAN AND R. KANNAN

*Centre for Visual Computing, Faculty of Computing and Informatics
Multimedia University*

Cyberjaya, 63100 Malaysia

E-mail: mohamed.sha33@gmail.com; {eswaran; kannan.ramakrishnan}@mmu.edu.my

Content-Based Image Retrieval (CBIR) systems are widely used for local as well as for remote applications such as telemedicine, satellite image transmission and image search engines. The existing CBIR systems suffer from the limitations of storage space, data security and bandwidth requirement. To overcome these problems, a new method termed as CBIR-PE which makes use of prediction errors instead of actual images for storage, transmission and retrieval is presented. Identical artificial neural networks (ANNs) are employed both at the server and client sides to carry out the prediction. At the server side, only the error database comprising the difference between the original and the predicted pixel values is used instead of the actual image database. The prediction errors of the query image are matched with those in the server database to retrieve similar prediction error patterns. These errors are then combined with the predicted values available at the client ANN to reconstruct the actual images. Since only the prediction errors are employed, the proposed method is able to solve the problems of storage space, data security and bandwidth requirement. The proposed method is implemented in combination with a clustering technique called WBCT-FCM which makes use of wavelet based contourlet transform (WBCT) and fuzzy c-means (FCM) clustering algorithm. The performances of the proposed WBCT-FCM and CBIR-PE are evaluated using COREL-1k database. The experimental results show that the proposed methods achieve better results with respect to clustering and retrieval accuracies compared to the existing methods.

Keywords: content-based image retrieval, wavelet-based contourlet transform, fuzzy c-means clustering, artificial neural network, prediction error

1. INTRODUCTION

Nowadays digital images are widely used in various domains for storing the information. It is usually very difficult for a user to retrieve an image stored in a large database. Therefore, an efficient and reliable automated procedure is needed for indexing and retrieving the images from large databases [1]. Content-based image retrieval (CBIR) systems have become very popular in recent years [2-7]. The emergence of internet has paved the way for digital acquisition of information especially from different databases. Most of the existing approaches focus on the low level and semantic features for developing CBIR systems [8-10]. The image semantics implying the importance of an image has several classification levels [11, 12]. CBIR helps in retrieving images from an image database, which are semantically relevant to the query image based on the image features [13, 14].

Conventional CBIR systems are normally categorized into three types based on extracting models such as region based search, histogram and color layout. As the similarity search is based on the overall properties of images, most of the CBIR systems make

Received January 22, 2016; revised April 24 & June 23, 2016; accepted June 24, 2016.
Communicated by Chu-Song Chen.

use of a global search for performing matching between the query and database images. Implementing CBIR systems for remote applications involving large databases with limited storage space and bandwidth still remains a challenging problem. It is possible to overcome this problem by using image compression techniques. In this paper, an integrated prediction error based image retrieval system called CBIR-PE which combines both compression and retrieval operations are proposed. The proposed CBIR-PE makes use of prediction errors as features for retrieval instead of the actual image pixel values, thus it is able to overcome the limitations with respect to storage space and bandwidth. Before storing the images, they are grouped into a set of classes using a clustering system which is implemented in two steps. In the first step, the image texture features is extracted using wavelet-based contourlet transform (WBCT) [15]. In the second step, the extracted texture features are clustered into a set of classes using fuzzy c-means (FCM) clustering method [16]. The prediction is performed using a feed-forward artificial neural network (ANN) and the prediction errors (PE) are obtained as the difference between the original and the predicted image pixel values. The similarity matching between the prediction errors of the query image and those of the database images is carried out using the Euclidean distance similarity measure [17]. The performance measure of the proposed image retrieval system CBIR-PE is carried out using parameters such as precision, recall and accuracy with the COREL 1k database.

2. BACKGROUND STUDY

CBIR systems are being used for indexing and retrieval of visual data since two decades [18, 19]. CBIR helps the user to find images which are similar in sketch, color texture characteristics, compared to text based systems which make use of only textual keywords. Different types of CBIR systems have been proposed recently by many authors [20-22]. Image retrieval for medical applications (IRMA) has been developed to support semantic and formalized queries [23]. A semantic region based image retrieval (SRBIR) system has been proposed by Rajam & Valli [24]. In this method, an image is segmented into various regions to find the most dominant foreground region. Then the low-level features are extracted from the identified dominant foreground region. For the learning process, support vector machine-binary decision tree (SVM-BDT) has been used. Mohammed *et al.* [7] introduced a technique for content-based image classification and retrieval using pulse-coupled neural network (PCNN). This technique uses PCNN to improve the signature quality with k-nearest neighbor (K-NN) algorithm for classification and matching of the image.

Yonekawa & Kurokawa [25] proposed a method of learning to define the parameters for image matching in the pulse coupled neural network (PCNN). The performance of image matching using the PCNN is improved by the learning method. With increased size of image database, identification of similar images has become a challenging task for the researchers [26]. A CBIR system based on grouped bag-of-visual-words (GBoVW) model has been proposed by Alexandra Alfanindya *et al.* [6]. This method makes use of interest point's detector and descriptor called speeded-up robust features (SURF) for texture feature extraction.

A CBIR system called 'weighted histograms as input mean-shift and Gaussian mix-

tures' (WHMSGM) system have been proposed by Mohamed Bouker & Hervet [27]. In this method, a tool has been developed based on weighted color histograms to model the colors of an image as a set of 2D Gaussian distributions with mean shift analysis clustering schemes. A hierarchical multilevel image thresholding for image retrieval system has been proposed by Guan & Yan [28]. This system is based on maximum fuzzy entropy principle with Euclidean distance measure to retrieve similar images for the given query image, in this research work consider only Euclidean distance measure to measure the similarity between the query and input samples. Subrahmanyam *et al.* [29] proposed an image indexing and retrieval system, which uses the color histogram and spatial orientation tree as the vector points to match and retrieve similar images from the database. Pandey & Kumar [30] have proposed a method using FCM and local binary pattern (LBP) for image retrieval, which uses the feature vector formed by concatenating the cluster LBP histograms to find similar images.

Yildizer *et al.* [31] have proposed a CBIR system in which multiple support vector machines ensemble concept was made use of. In [31], image features were extracted using Daubechies wavelet transformation, and the performance results with respect to classification as well as retrieval were reported. These authors have defined the precision value assuming that the number of retrieved images varies from 1 to 100. A CBIR system based on statistical modeling approach has been proposed by Li *et al.* [32]. In this approach, wavelet features of images extracted using Daubechies-4 and Haar transform have been applied.

Lin *et al.* [33] have proposed an image retrieval system using a combination of image features such as color co-occurrence matrix, difference between pixels of scan pattern and color histogram. ElAlami *et al.* [34] have proposed an improved CBIR system using color and texture features and Artificial Neural Network model for classification. In [35], Irtaza *et al.* have proposed a CBIR system which employs wavelet packets and Eigen values of Gabor filters for image representation. To make the retrieval more accurate, a partial supervised learning scheme was applied in [35]. Huang *et al.* [36] have proposed a CBIR system based on texture similarity. In this paper, two image features namely, the composite sub-band gradient vector and the energy distribution pattern string from the sub-images of a wavelet decomposition have been employed. A fuzzy based matching algorithm has been used for checking the image similarity. Jhanwar *et al.* [37] have proposed a CBIR system using motif co-occurrence matrix (MCM) which is a combination of both color and texture features. The distance between the MCMs of two different images is used as the similarity measure for the images retrieval process.

3. RESEARCH METHODOLOGY

The block diagram representation of the proposed CBIR-PE system is shown in Fig. 1. The image dataset is first converted into a set of greyscale images. It is then grouped into R classes (1, 2... R) using fuzzy c-means (FCM) [16] clustering method based on features extracted using wavelet based contourlet transformation [15, 27, 38]. For each class of images, an optimal artificial neural network (ANN) configuration is selected to achieve the best prediction value. The server side database comprises two types, namely dataset of errors (DE) and dataset of class parameters (DCP). Assume that A and A' represent respectively the matrices comprising the original and the predicted pixel values of

an image. Also assume that the first row and the first column of the A' matrix contain 0 values and the remaining rows and columns contain the predicted values. In other words, we assume that the prediction process starts only for row and column greater than 1. In DE, the error matrix of each image denoted as PE matrix which is the difference between A and A' matrices is stored. It may be noted that the first row and the first column of a PE matrix will contain the original pixel values of the image which are used as initial values for ANN for prediction. The second dataset, namely, DCP comprises the FCM center values and the ANN configuration parameters (such as weights and biases) for each class of images (1, 2, ..., R). The DCP dataset is initially transmitted to the client side for synchronization purpose. In other words, the DCP data sets are assumed to be available both at the server and client sides. For retrieval operation, the query image is first clustered into one of (1, 2, ..., R) classes (say class x) based on WBCT and FCM operations. Using ANN_x , the PE^q matrix representing the prediction error values of the query image is obtained as shown in Fig. 1.

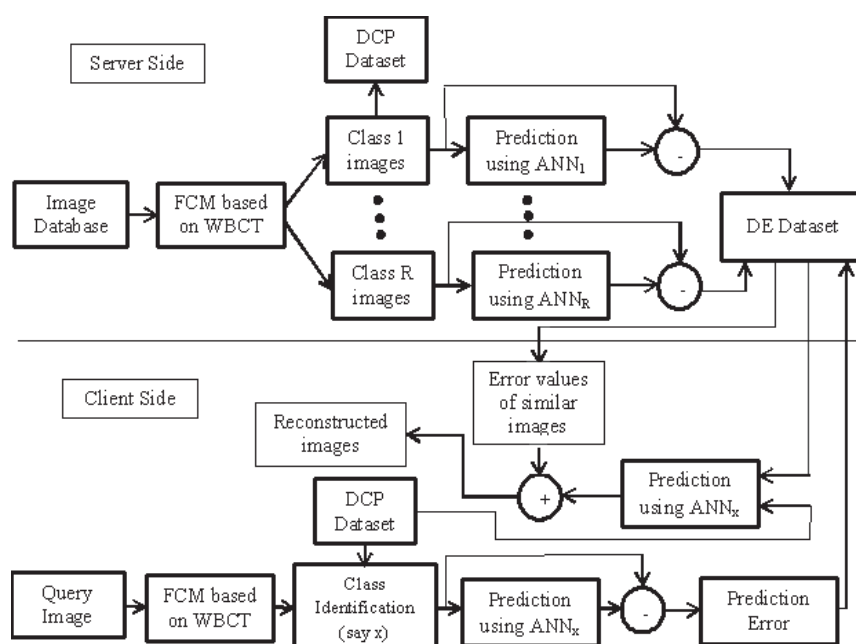


Fig. 1. Block diagram of the proposed CBIR-PE system.

Error matching is done between the PE^q matrix and the PE matrices of class x stored in the DE dataset at the server using the Euclidian distance (ED) measure. The PE matrices of the images with minimum ED are retrieved from the DE dataset. The retrieved images are reconstructed at the client side by combining the corresponding PE and A' matrices.

3.1 Wavelet Based Contourlet Transform

Wavelet based contourlet transform (WBCT) [15] has a similar structure as that of

contourlet transform. WBCT (Fig. 2) consists of two filter bank stages for smoothening the images and the contours. The first stage provides sub-band decomposition, where multi-scale decomposition using Laplacian pyramid is implemented in wavelet. The second stage of the WBCT is a directional filter bank (DFB) which is applied to all three high-pass sub-band for the angular decomposition. The WBCT provides a sparse representation, which resembles as the segmented contours of the images.

Fig. 3 illustrates the schematic representation of the dyadic scale level for the wavelet decomposition. At each wavelet transformation level (l), LH, HL, and HH high pass bands are obtained. Then DFB is applied to these high pass bands with the number of directions N_D at each level (l) fixed as shown in Eq. (1), where l represents the level of decomposition.

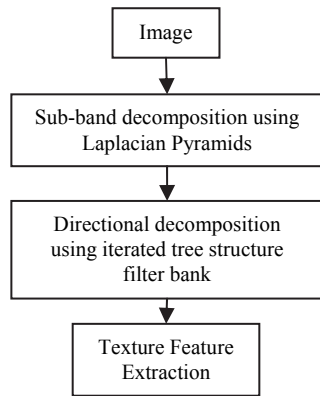


Fig. 2. WBCT block diagram.

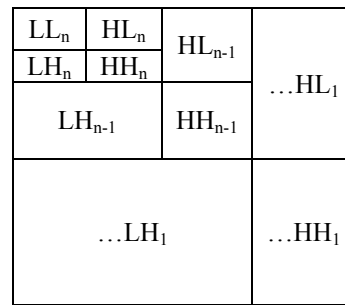


Fig. 3. n dyadic decomposition levels.

$$N_D = 2^l(1)$$

An l -level DFB has 2^l directional sub-bands with $G_k^l, 0 \leq k < 2^l$ equivalent synthesis filters and down sampling matrices $S_k^l, 0 \leq k < 2^l$ which are defined as in Eq. (2) [38].

$$S_k^l = \begin{cases} \begin{bmatrix} 2^{l-1} & 0 \\ 0 & 2 \end{bmatrix}, & \text{if } 0 \leq k < 2^{l-1} \\ \begin{bmatrix} 2 & 0 \\ 0 & 2^{l-1} \end{bmatrix}, & \text{if } 2^{l-1} \leq k < 2^l \end{cases} \quad (2)$$

If g_k^l represents the impulse response of the synthesis filter, G_k^l then $\{g_k^l[n - s_k^l m]\}, 0 \leq k < 2^l, m \in \mathbb{Z}^2$ represents the directional basis for $l^2(\mathbb{Z}^2)$.

Contourlet (ρ) is formed by the combination of the DFB basis along with the multiscale subspaces of the Laplacian pyramid (Ψ) as given in Eq. (3) [38].

$$\rho = \sum_{m, n \in \mathbb{Z}^2} g_k^l[n - S_k^l m] \Psi, 0 \leq k < 2^l \quad (3)$$

A rigorous mathematical analysis of WBCT can be found in [38].

Texture is one of the important properties or characteristics of the images, as the texture feature provides a regional geometric descriptor that helps in the image retrieval process. Though the texture feature may be adequate for image clustering, this feature alone may not give accurate results for image retrieval applications. To extract the texture feature, wavelet decomposition is applied to the image. After the wavelet decomposition, the three high pass bands LH, HL, and HH are obtained. These corresponding wavelet coefficients are used for texture discrimination [27].

3.2 Fuzzy C-Means Clustering

In hard clustering, texture feature data is extracted from the WBCT and partitioned into separate clusters. If two or more extracted texture features have similar characteristics, then they are grouped under the same cluster or class. Fuzzy c-means (FCM) [16, 30] is one of the widely used clustering algorithms for various applications. FCM algorithm allows assigning the data points to one or more clusters. Each data point has its own degree of membership (or probability), with respect to a given cluster. FCM algorithm separates each input image $Y = \{y_1, \dots, y_n\}$ into a collection of cluster centers $C = \{c_1, \dots, c_c\}$ using the image texture features according to the objective function J_m given in Eq. (4) [39].

$$J_m = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m |y_i - c_j|^2 \quad (4)$$

where, u_{ij} is y_i 's degree of membership in the j th cluster and m is any real value greater than 1. Fuzzy separation is an iterative optimization of the objective function with regard to the change or modification in the membership function u_{ij} and the cluster centers c_j . Eqs. (5) and (6) show the formulae to compute the degree of membership u_{ij} and cluster center c_j respectively [39]:

$$u_{ij} = \frac{|y_i - c_j|^{-\frac{1}{m-1}}}{\sum_{j=1}^c |y_i - c_j|^{-\frac{1}{m-1}}} \quad (5)$$

$$c_j = \frac{\sum_{i=1}^n u_{ij}^m y_i}{\sum_{i=1}^n u_{ij}^m} \quad (6)$$

Fuzzy c-means Clustering algorithm uses the following steps to cluster the available data points [16]:

- Step 1:** Initialize membership $U = [u_{ij}]$, $U^{(0)}$.
- Step 2:** Evaluate the fuzzy centroid $C^{(k)} = [c_j]$, where $j = 1, \dots, c$, using Eq. (6).
- Step 3:** Update the fuzzy membership $U^{(k)} = [u_{ij}]$, $U^{(k+1)}$ using Eq. (5).
- Step 4:** If $\|U^{(k)} - U^{(k-1)}\| < \varepsilon$, then STOP, else again compute centroid using step 2.

3.3 Artificial Neural Network

The Artificial neural network (ANN) comprises an input layer, a hidden layer and

an output layer [40]. In this work, ANN is used as a predictor for predicting image pixel values using neighboring pixels. The design of an optimal ANN predictor involves selection of appropriate number of input and hidden layer neurons as well as selecting the best activation functions for the hidden and output layers. The ANN is initially trained using a learning algorithm for updating the edge weights. Fig. 4 shows the proposed ANN model with three layers.

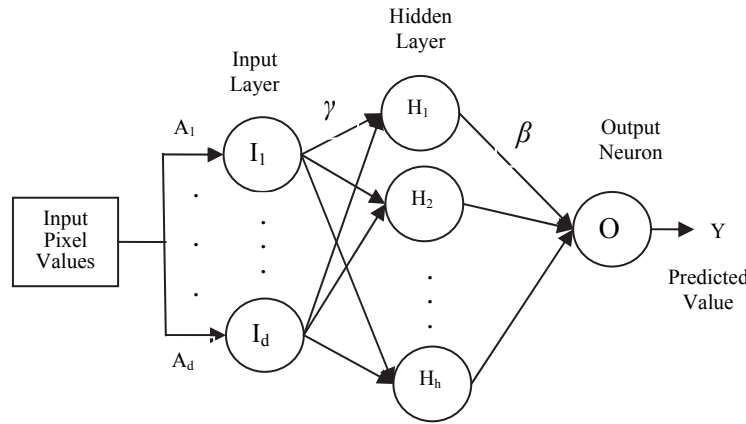


Fig. 4. Neural network structure.

The training phase of ANN involves assigning the weight values for the edges between the input and hidden layers (γ_{ji}) and also for the edges between the hidden and output layers (β_j). The weighted sum of the values at the input layer neurons is applied to each of the hidden layer neurons. A similar process is repeated between the hidden and output layers. The final output Y obtained at the output layer neuron is given in Eq. (7).

$$Y = \Omega \left\{ \beta_0 + \sum_{j=1}^h \beta_j \Phi \left(\gamma_{j0} + \sum_{i=1}^d \gamma_{ji} A_i \right) \right\} \quad (7)$$

where $(\beta_0, \beta_1, \dots, \beta_h, \gamma_{10}, \gamma_{11}, \dots, \gamma_{hd})$ represent the bias and weight parameters. The functions Φ and Ω in Eq. (7) represent the activation functions applied at the hidden and output layer neurons respectively.

Three types of activation functions, namely, linear, sigmoid and hyperbolic tangent which are defined as in Eqs. (8)-(10) respectively are considered for the hidden and output layer neurons in this work.

$$f(\lambda) = \lambda \quad \text{Linear} \quad (8)$$

$$f(\lambda) = \frac{1}{1 + e^{-\lambda}} \quad \text{Sigmoid (Logistic)} \quad (9)$$

$$f(\lambda) = \frac{1 - e^{-2\lambda}}{1 + e^{-2\lambda}} \quad \text{Hyperbolic tangent} \quad (10)$$

3.3.1 Optimal ANN configuration

Prediction Error (PE) matrix of an image is determined as the difference between

the actual and predicted pixel values. Selection of the optimal ANN configuration for each class of images involves finding the best set, namely, the number of input and hidden layer neurons, hidden and output layer activation functions and training algorithms which yields the minimum prediction error for that class. For this purpose, a performance measure termed as average root mean square of the PE matrix (ARMSPE) is defined for a class of images comprising say K images of $N \times N$ dimension as in Eq. (11).

$$ARMSPE = \frac{1}{K} \sum_{k=1}^K \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^N PE_{ij}^k}{N^2}} \tag{11}$$

where K represents the total number of images in the class.

3.3.2 ANN prediction

The ANN is trained with the training dataset images selected from the COREL 1k database. For each cluster, a separate optimal ANN model is obtained using the procedure described in the previous section. Fig. 5 (a) shows a sample A matrix comprising the pixel values of the original image. Fig. 5 (b) shows the corresponding $N \times N$ predicted matrix denoted as A' where the first row and column values are zeros and the remaining rows and columns contain the predicted values. For example, A'_{22} in Fig. 5b, which is the predicted value of A_{22} is obtained by using the values of A_{11} , A_{12} and A_{21} as inputs to the predictor. Similarly the predicted value of A_{23} represented as A'_{23} is obtained by using the values of A_{12} , A_{13} and A'_{22} and this procedure is repeated for the remaining pixel values.

The prediction error matrix (PE) of an image is given in the Eq. (12):

$$PE = A - A' \tag{12}$$

where A and A' represent the original and the predicted matrices of the image.

A_{11}	A_{12}	A_{13}	...	A_{1N}
A_{21}	A_{22}	A_{23}	...	A_{2N}
A_{31}
...
A_{N1}	A_{N2}	A_{N3}	...	A_{NN}

Fig. 5. (a) Original image matrix (A).

0	0	0	0	0
0	A'_{22}	A'_{23}	...	A'_{2N}
0	A'_{32}
0
0	A'_{N2}	A'_{N3}	A'_{NN}

Fig. 5. (b) Predicted matrix (A').

Fig. 6 shows the PE matrix, in which the first row and first column values will be the same as the original pixel values of the image.

$PE_{11} = A_{11}$	$PE_{12} = A_{12}$	$PE_{13} = A_{13}$...	$PE_{1N} = A_{1N}$
$PE_{21} = A_{21}$	PE_{22}	PE_{23}	...	PE_{2N}
$PE_{31} = A_{31}$
....
$PE_{N1} = A_{N1}$	PE_{N2}	PE_{N3}	...	PE_{NN}

Fig. 6. PE matrix of an image.

The PE matrices of all images are stored in the DE dataset at the server side. The ANN parameters, such as the number of hidden layer neurons, type of activation functions are stored in the DCP datasets both at the server and the client sides.

3.3.3 Image retrieval

Fig. 7 shows the overall image retrieval mechanism adopted in the proposed CBIR-PE system.

The mean value of a column in the PE^q matrix of the query image denoted as is \overline{PE}_j^q calculated as shown in Eq. (13):

$$\overline{PE}_j^q = \frac{\sum_{i=1}^N PE_{ij}^q}{N}, \text{ for } j=1, 2, \dots, N. \quad (13)$$

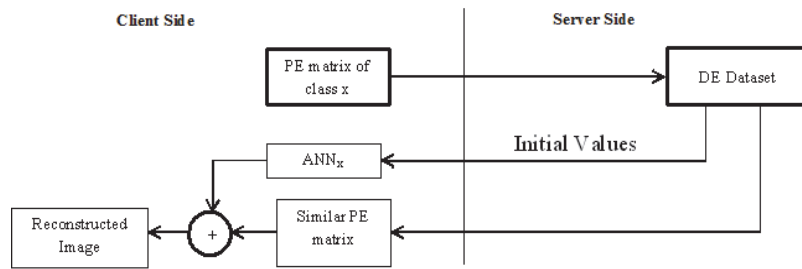


Fig. 7. Image retrieval mechanism.

Assuming that the query image belongs to class x , the Euclidean distance measure between the column mean values PE_j of the query image and the mean values of an image, k , in class x denoted as ED_k^x is calculated as shown in Eq. (14):

$$\overline{PE}_j^q = \frac{\sum_{i=1}^N PE_{ij}^q}{N}, \text{ for } j=1, 2, \dots, N. \quad (13)$$

$$ED = \sqrt{\left(\overline{PE}_1^q - \overline{PE}_1^k\right)^2 + \left(\overline{PE}_2^q - \overline{PE}_2^k\right)^2 + \dots + \left(\overline{PE}_N^q - \overline{PE}_N^k\right)^2} \quad (14)$$

Eq. (14) is calculated for $k = 1, 2, \dots, K$ where K represents the total number of images stored in class x .

Assume that a threshold value or number of top matches of say M is fixed for the number of images to be retrieved. Based on Eq. (14), M number of images will be retrieved in the ascending order of the ED_k^x values. In order to assign a more accurate ranking order for the retrieved M images, Euclidean distance measure ED' is applied at the pixel level using the PE matrices of the query image and the selected M images as shown Eq. (15).

$$ED' = \sqrt{\sum_{j=1}^N \sum_{i=0}^N \left(PE_{ij}^q - PE_{ij}^m\right)^2}, \text{ for } m = 1, 2, \dots, M \quad (15)$$

The selection of threshold value M will affect the retrieval performance as discussed in section 4.3.

In effect, image retrieval in the proposed system involves the following steps:

- Step 1:** WBCT is applied to the query image to extract the texture features.
- Step 2:** Using FCM clustering technique, the class x to which the query image belongs is identified.
- Step 3:** ANN_K predictor is applied to the query image to get the PE^q matrix of the image. The parameters of ANN_x are assumed to be stored in the DCP dataset (see Fig. 1).
- Step 4:** PE matrices which are close to the PE matrix of the query image are retrieved from the DE dataset by using Euclidean distance measure as shown in Eqs. (13)-(15).
- Step 5:** The first row and the first column of the PE matrix of a retrieved image contain the original pixel values of the image. By using these original pixel values as inputs to ANN_x , the predicted matrix (A') of the image is obtained.
- Step 6:** By combining the PE and A' matrices, the retrieved image is reconstructed.

4. EXPERIMENTAL RESULTS

The proposed system is tested using the general-purpose COREL 1k database images (URL <http://wang.ist.psu.edu/docs/related/>). Fig. 8 shows the sample images from Corel dataset, which contains 1000 images of 10 classes with 100 images in each class. The 10 classes of images are: African People, Beach, Building, Buses, Dinosaur, Elephants, Flower, Horse, Mountain and Food. Among the total 100 images in each class, 10 images are used as query images and the remaining 90 images are stored in the image database. After applying the clustering technique (*i.e.* FCM based on WBCT), the 900 images stored in the database are grouped into 10 classes. The ANN is trained for each class separately. For this purpose, 60% of the images in each class are used for training and the remaining images are used for testing based on the empirical results.

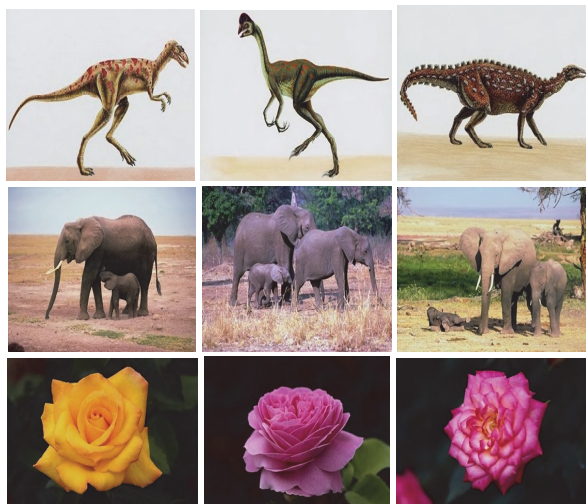


Fig. 8. Sample images of the COREL dataset.

4.1 Performance of the Proposed WBCT-FCM

The clustering performance is determined based on the confusion matrix shown in Fig. 9. The confusion matrix values obtained for the 10 classes using the proposed method are shown in Table 1. In the confusion matrix, the total number of images in each class before clustering (denoted as actual class) is 90 which is given by the sum of values in each row. The total number of images in each class after clustering (denoted as predicted class) is given as the sum of values in each column. For example, for class 1, the number of images in the actual class is 90 whereas the number of images in the predicted class is 76. The experimental results obtained with respect to clustering and optimal ANN configuration are given in this section.

Actual Class	Predicted Class	
	true positives (actual relevant images that were correctly classified)	false negatives (relevant images that were incorrectly marked as non-relevant images)
	false positives (images that were incorrectly labelled as relevant)	true negatives (all the remaining images, correctly discarded as non-relevant images)

Fig. 9. Confusion matrix.

Table 1. Confusion matrix obtained using the proposed WBCT-FCM method.

Actual Class \ Predicted Class	Predicted Class									
	1	2	3	4	5	6	7	8	9	10
1	54	10	0	0	0	0	10	0	8	8
2	8	65	1	1	0	1	4	3	2	5
3	2	3	76	0	0	0	5	0	1	3
4	2	0	0	85	0	2	0	1	0	0
5	0	0	0	0	90	0	0	0	0	0
6	3	3	0	0	0	71	7	1	2	3
7	0	0	2	1	0	2	85	0	0	0
8	0	0	0	0	0	0	0	82	4	4
9	4	3	0	2	0	0	9	1	70	1
10	3	4	3	1	0	8	0	2	5	64

The clustering or the classification accuracy of the proposed WBCT-FCM method for each class is measured based on the diagonal entries in the confusion matrix [31, 32]. The maximum accuracy value of 100 % is achieved for the class 5 (dinosaur’s images). The average clustering accuracy value achieved by the proposed method is 82 % which is significantly higher compared to the values of 62% and 64% reported in [31, 32] respectively.

Fig. 10 shows a comparison graph for the classification accuracy values achieved by the proposed method and two other known methods reported in [31, 32]. From Fig. 10, it is clear that the proposed method in general performs better than the other known methods [31, 32] except for class 7 and 9.

4.2 Determination of Optimal ANN Configuration

The optimal ANN configuration is determined as described in section 3.3.1 using different sets of activation functions, training algorithms, input and hidden layer neurons. For this purpose, the following training algorithms and activation function are considered.

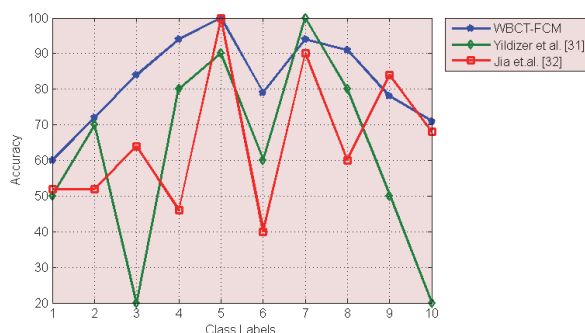


Fig. 10. Comparison of clustering accuracy with existing systems.

Training algorithms: Levenberg-Marquardt (LM), Resilient Backpropagation (RP), BFGS quasi-Newton backpropagation (BFG), Batch training with weight and bias learning rules and Bayesian regulation backpropagation (BR).

Activation functions: Hyperbolic tangent sigmoid (tansig), Log sigmoid (logsig), Triangular basis transfer function (tribas), Radial basis and Linear transfer function (purlin).

The number of input and hidden layer neurons which are used for testing are as follows:

Input layer neurons: 3, 4 and 5.

Hidden layer neurons: 5, 10 and 20.

Based on experimental results, it is found that for all the classes of images, input layer neuron = 3, hidden layer neuron = 10, yield the minimum ARMSPE values for the considered training algorithms and activation functions. For illustration, the ARMSPE values obtained for the first class of images (*i.e.* African people) with 10 hidden layer neurons and with input neurons = 3, 4 are given in Table 2. This table shows the results obtained using training algorithms LM and RP with three different types of activation functions (tansig, logsig and purelin).

Table 2. Optimal ANN prediction parameter selection.

Hidden, Output layer Activation functions	Number of Input Neurons = 3		Number of Input Neurons = 4	
	LM	RP	LM	RP
logsig, logsig	ARMSPE=23.552	ARMSPE=34.533	ARMSPE=19.459	ARMSPE=23.183
logsig, purlin	ARMSPE=15.647	ARMSPE=13.680	ARMSPE=30.338	ARMSPE=27.063
tansig, purlin	ARMSPE=12.868	ARMSPE=14.066	ARMSPE=35.822	ARMSPE=36.519

It can be noted from Table 2 that the following combination yields the minimum ARMSPE of 12.868 for class 1 images:

Number of input neurons = 3, LM training algorithm, tansig at the hidden layer with 10 hidden neurons and purelin at the output layer.

The optimal ANN configuration is determined for each class as described in Section 3.3.1 and the results obtained are given in Table 3.

Table 3. Optimal ANN configuration parameters.

Class No.	Class Name	Training Algorithm	Activation Function		ARMSPE
			Hidden Layer	Output Layer	
1	African People	Levenberg-Marquardt	Hyperbolic tangent sigmoid	Linear transfer function	12.868
2	Beach	Resilient Backpropagation	Log sigmoid	Log sigmoid	23.453
3	Building	Levenberg-Marquardt	Hyperbolic tangent sigmoid	Log sigmoid	19.362
4	Bus	Resilient Backpropagation	Triangular basis transfer function	Linear transfer function	20.904
5	Dinosaur	BFGS quasi-Newton backpropagation	Log sigmoid	Linear transfer function	26.399
6	Elephants	Levenberg-Marquardt	Hyperbolic tangent sigmoid	Log sigmoid	16.389
7	Flower	Resilient Backpropagation	Radial basis	Log sigmoid	19.952
8	Horse	BFGS quasi-Newton backpropagation	Triangular basis transfer function	Linear transfer function	14.793
9	Mountain	Resilient Backpropagation	Radial basis	Log sigmoid	17.847
10	Food	Levenberg-Marquardt	Hyperbolic tangent sigmoid	Linear transfer function	14.655

For all the classes, the number of input layer and hidden layer neurons are fixed as 3 and 10 respectively. In Table 3, ARMSPE denotes the average of the root mean square prediction error values of the images in each class.

4.3 Performance of the CBIR-PE System

The image retrieval system performance is evaluated using the precision and recall measures, which are defined as in Eqs. (16) and (17) [41].

$$Precision = \frac{\text{relevant images in retrieved images}}{\text{retrieved images}} \% \quad (16)$$

$$\text{Recall} = \frac{\text{relevant images in retrieved images}}{\text{relevant images in the database}} \% \quad (17)$$

Table 4 shows the average precision and recall values obtained for each class using the proposed CBIR-PE method, assuming that the threshold value fixed for the top matches is 20, 30, 40 and 50.

Table 4. Image retrieval performance of CBIR-PE method.

Class No.	Class Name	Precision value (%)				Recall value (%)			
		Threshold Value For Top Matches (M)				Threshold Value For Top Matches (M)			
		20	30	40	50	20	30	40	50
1	African People	70	67	65	62	16	22	29	34
2	Beach	75	73	70	60	17	24	31	33
3	Building	90	90	88	88	20	30	39	49
4	Bus	95	93	90	90	21	31	40	50
5	Dinosaur	100	100	100	100	22	33	44	56
6	Elephants	85	75	73	74	19	26	32	41
7	Flower	70	67	63	50	16	22	28	28
8	Horse	90	87	85	84	20	29	38	47
9	Mountain	75	73	68	62	17	24	30	34
10	Food	70	63	60	60	20	21	27	33
	Average	82	79	76	73	19	26	34	41

As the threshold value M increases, the increase in the number of relevant images will be normally less compared to the increase in the number of retrieved images and as a result, the precision value as given in Eq. (16) decreases. This fact is observed in the experimental results shown in Table 4. However with respect to the recall value, we note from Eq. (17) that as the value of M increases, the number of relevant images in the database remaining constant, the number of relevant images among the retrieved images normally increases and as a result, the recall value increases. This fact is reflected in the experimental results shown in Table 4.

In Fig. 11, the precision and recall values obtained by the proposed CBIR-PE method are compared with those obtained by using the existing methods.

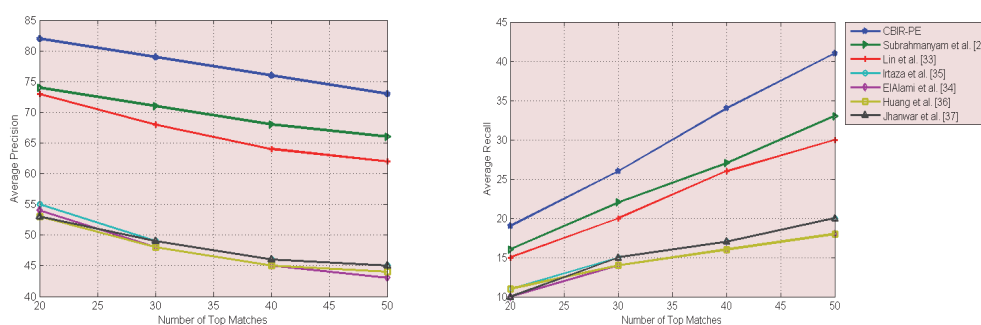


Fig. 11. Comparison of average precision and recall value with the existing systems.

From Fig. 11, it is noted that the proposed method yields higher precision and recall values compared to the values achieved by the known methods.

5. CONCLUSIONS

A new content-based image retrieval technique using prediction error concept called CBIR-PE system has been presented. The proposed method makes use of a new clustering scheme employing wavelet based contourlet transform and fuzzy c-means clustering for improving the performance of the system. A multilayer neural network predictor has been employed for predicting the pixel values using the neighboring pixels. The proposed method has the advantage of using a lower bandwidth since for retrieval only the error values are transmitted instead of the actual image pixel values. In addition, it also has the advantage of data security since the original images cannot be reconstructed without the optimal predictors. Since only the error values are stored in the database instead of the actual image pixel values, the proposed system requires significantly reduced storage space. From the experimental results, it is found that the proposed method achieves better results compared to the existing methods with respect to the performance parameters such as clustering accuracy, precision and recall.

ACKNOWLEDGEMENT

The authors thank Multimedia University, Malaysia for supporting this work through Graduate Research Assistantship Scheme. They also wish to thank the anonymous reviewers for their comments and suggestions which helped in enhancing the quality of the paper.

REFERENCES

1. C. B. Akgül, D. L. Rubin, S. Napel, C. F. Beaulieu, H. Greenspan, and B. Acar, "Content-based image retrieval in radiology: current status and future directions," *Journal of Digital Imaging*, Vol. 24, 2011, pp. 208-222.
2. H. L. Tang, R. Hanka, and H. H. Ip, "Histological image retrieval based on semantic content analysis," *IEEE Transactions on Information Technology in Biomedicine*, Vol. 7, 2003, pp. 26-36.
3. L. Zheng, A. W. Wetzel, J. Gilbertson, and M. J. Becich, "Design and analysis of a content-based pathology image retrieval system," *IEEE Transactions on Information Technology in Biomedicine*, Vol. 7, 2003, pp. 249-255.
4. H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applications – clinical benefits and future directions," *International Journal of Medical Informatics*, Vol. 73, 2004, pp. 1-23.
5. W. Hsu, L. R. Long, and S. Antani, "SPIRS: a framework for content-based image retrieval from large biomedical databases," *MedInfo*, Vol. 12, 2007, pp. 188-192.
6. A. Alfanindya, N. Hashim, and C. Eswaran, "Content based image retrieval and classification using speeded-up robust features (SURF) and grouped bag-of-visual-words

- (GBoVW),” in *Proceedings of IEEE International Conference on Technology, Informatics, Management, Engineering, and Environment*, 2013, pp. 77-82.
7. M. M. Mohammed, A. Badr, and M. B. Abdelhalim, “Image classification and retrieval using optimized pulse-coupled neural network,” *Expert Systems with Applications*, Vol. 42, 2015, pp. 4927-4936.
 8. L. Yang, O. Tuzel, W. Chen, P. Meer, G. Salaru, L. A. Goodell, and D. J. Foran, “Pathminer: a web-based tool for computer-assisted diagnostics in pathology,” *IEEE Transactions on Information Technology in Biomedicine*, Vol. 13, 2009, pp. 291-299.
 9. D. S. Channin, P. Mongkolwat, V. Kleper, K. Sepukar, and D. L. Rubin, “The caB-IG™ annotation and image markup project,” *Journal of Digital Imaging*, Vol. 23, 2010, pp. 217-225.
 10. H. Guan, S. Antani, L. R. Long, and G. R. Thoma, “Minimizing the semantic gap in biomedical content-based image retrieval,” *SPIE Medical Imaging, International Society for Optics and Photonics*, 2010, pp. 762807-762807.
 11. L. J. Li, H. Su, L. Fei-Fei, and E. P. Xing, “Object bank: A high-level image representation for scene classification and semantic feature sparsification,” *Advances in Neural Information Processing Systems*, 2010, pp. 1378-1386.
 12. C. Liu and G. Song, “A method of measuring the semantic gap in image retrieval: Using the information theory,” in *Proceedings of IEEE International Conference on Image Analysis and Signal Processing*, 2011, pp. 287-291.
 13. A. Kumar, J. Kim, W. Cai, M. Fulham, and D. Feng, “Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data,” *Journal of Digital Imaging*, Vol. 26, 2013, pp. 1025-1039.
 14. Y. Liu, D. Zhang, G. Lu, and W. Y. Ma, “A survey of content-based image retrieval with high-level semantics,” *Pattern Recognition*, Vol. 40, 2007, pp. 262-282.
 15. R. Eslami and H. Radha, “Wavelet-based contourlet coding using an SPIHT-like algorithm,” in *Proceedings of IEEE International Conference on Image Processing*, 2004, pp. 784-788.
 16. A. Fabijańska, “A fuzzy segmentation method for images of heat-emitting objects,” in *Proceedings of Iberoamerican Congress on Pattern Recognition*, 2009, pp. 217-224.
 17. A. Lakdashti, M. S. Moin, and K. Badie, “Reducing the semantic gap of the MRI image retrieval systems using a fuzzy rule based technique,” *International Journal of Fuzzy Systems*, Vol. 11, 2009, pp. 232-249.
 18. A. W. 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, 2000, pp. 1349-1380.
 19. Y. Chen, J. Z. Wang, and R. Krovetz, “Content-based image retrieval by clustering,” in *Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval*, 2003, pp. 193-200.
 20. S. B. Park, J. W. Lee and S. K. Kim, “Content-based image classification using a neural network,” *Pattern Recognition Letters*, Vol. 25, 2004, pp. 287-300.
 21. E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, “A simultaneous feature adaptation and feature selection method for content-based image retrieval systems,” *Knowledge-Based Systems*, Vol. 39, 2013, pp. 85-94.

22. C. Beecks, M. S. Uysal, and T. Seidl, "Content-based image retrieval with Gaussian mixture models," in *Proceedings of International Conference on Multimedia Modeling, Springer International Publishing*, 2015, pp. 294-305.
23. M. O. Güld, C. Thies, B. Fischer, and T. M. Lehmann, "A generic concept for the implementation of medical image retrieval systems," *International Journal of Medical Informatics*, Vol. 76, 2007, pp. 252-259.
24. I. F. Rajam and S. Valli, "SRBIR: semantic region based image retrieval by extracting the dominant region and semantic learning," *Journal of Computer Science*, Vol. 7, 2011, pp. 400-408.
25. M. Yonekawa and H. Kurokawa, "The content-based image retrieval using the pulse coupled neural network," in *Proceedings of IEEE International Joint Conference on Neural Networks*, 2012, pp. 1-8.
26. A. Torralba, R. Fergus, and W. T. Freeman, "80 million tiny images: A large data set for nonparametric object and scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, 2008, pp. 1958-1970.
27. M. A. Bouker and E. Hervet, "Retrieval of images using mean-shift and Gaussian mixtures based on weighted color histograms," in *Proceedings of the 7th IEEE International Conference on Signal-Image Technology and Internet-Based Systems*, 2011, pp. 218-222.
28. P. P. Guan and H. Yan, "A hierarchical multilevel image thresholding method based on the maximum fuzzy entropy principle," *Image Processing: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*, 2013, pp. 274.
29. M. Subrahmanyam, R. P. Maheshwari, and R. Balasubramanian, "Expert system design using wavelet and color vocabulary trees for image retrieval," *Expert Systems with Applications*, Vol. 39, 2012, pp. 5104-5114.
30. D. Pandey and R. Kumar, "Hybrid algorithm using fuzzy C-means and local binary patterns for image indexing and retrieval," *Soft Computing Techniques in Vision Science*, 2012, pp. 115-125.
31. E. Yildizer, A. M. Balci, M. Hassan, and R. Alhaji, "Efficient content-based image retrieval using multiple support vector machines ensemble," *Expert Systems with Applications*, Vol. 39, 2012, pp. 2385-2396.
32. J. Li and J. Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, 2003, pp. 1075-1088.
33. C. H. Lin, R. T. Chen, and Y. K. Chan, "A smart content-based image retrieval system based on color and texture feature," *Image and Vision Computing*, Vol. 27, 2009, pp. 658-665.
34. M. E. ElAlami, "A new matching strategy for content based image retrieval system," *Applied Soft Computing*, Vol. 14, 2014, pp. 407-418.
35. A. Irtaza, M. A. Jaffar, E. Aleisa, and T. S. Choi, "Embedding neural networks for semantic association in content based image retrieval," *Multimedia Tools and Application*, Vol. 72, 2014, pp. 1911-1931.
36. P. W. Huang and S. K. Dai, "Image retrieval by texture similarity," *Pattern Recognition*, Vol. 36, 2003, pp. 665-679.

37. N. Jhanwar, S. Chaudhuri, G. Seetharaman, and B. Zavidovique, "Content based image retrieval using motif cooccurrence matrix," *Image and Vision Computing*, Vol. 22, 2004, pp. 1211-1220.
38. M. N. Do and M. Vetterli, "Contourlets," *Studies in Computational Mathematics*, Vol. 10, 2003, pp. 83-105.
39. C. C. Hung, S. Kulkarni, and B. C. Kuo, "A new weighted fuzzy c-means clustering algorithm for remotely sensed image classification," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, 2011, pp. 543-553.
40. D. Patidar, N. Jain, and A. Parikh, "Performance analysis of artificial neural network and K Nearest neighbors image classification techniques with wavelet features," in *Proceedings of IEEE International Conference on Computer Communication and Systems*, 2014, pp. 191-194.
41. D. L. Olson and D. Delen, "Advanced data mining techniques," *Springer Science & Business Media*, 2008, pp. 138-145.



A. Mohamed Uvaze Ahamed received his B.E. and M.E. degrees in Computer Science and Engineering from Anna University, Chennai, India. He worked as an Assistant Professor in the Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India for the period of June 2012 to April 2014. He received a scholarship under Graduate Research Assistantship scheme from Multimedia University, Malaysia. Currently he is pursuing his Ph.D. research in the area of Image Retrieval and Compression.



C. Eswaran received his B.Tech, M.Tech and Ph.D. degrees from the Indian Institute of Technology Madras, India where he worked as a Professor in the Department of Electrical Engineering until January 2002. Currently he is working as a Professor in the Faculty of Computing and Informatics, Multimedia University, Malaysia. He is also serving as the Chairman of the university center of excellence, "Visual Computing". Dr. C. Eswaran served as a visiting faculty and research fellow in many international universities such as Ruhr University Bochum, Germany, Concordia University, Canada, University of Victoria, Canada, and Nanyang Technological University, Singapore. He has also received prestigious international fellowships such as Humboldt-Fellowship (Germany) and Advanced Systems Institute Fellowship (British Columbia, Canada). He has supervised successfully more than 50 Ph.D. / M.S students in the areas of Multimedia Systems, Communications, Digital signal Processing, Artificial Intelligence, and Health Informatics. He has published more than 240 research papers in these areas in reputed international journals and conferences. He has also carried out several sponsored research projects and served as an industrial consultant in these areas. He is serving as a reviewer for many international journals and conferences.



Kannan Ramakrishnan received his Bachelor of Engineering degree in Electronics and Communication and Postgraduate Diploma in Medical Instrumentation Technology, both from Coimbatore Institute of Technology, India. He also received his Master of Engineering degree in Computer Science from Regional Engineering College (now known as National Institute of Technology), Trichy, India. He has completed his Ph.D. in Information Technology from Multimedia University, Malaysia, where he is currently working as a Senior Lecturer in the Faculty of Computing and Informatics. He has teaching and research experience of over 25 years in the field of Information Technology. His current research interests include biomedical signal processing, data mining, time-series forecasting, data compression, and soft computing models. He is a Member of Indian Society for Technical Education, IEEE Signal Processing Society, and Technical Committee on Health Informatics Standards, Malaysia.