Compact Image Signature Generation: An Application in Image Retrieval

Manish Chowdhury[1], Sudeb Das[2] and Malay Kumar Kundu[3] *Senior Member, IEEE*

Machine Intelligence Unit, Indian Statistical Institute, 203 B.T.Road, Kolkata-108, India

Email: [1]st.manishc@gmail.com, [2]to.sudeb@gmail.com, [3]malay@isical.ac.in

Abstract—In this article, we have proposed a novel Content Based Image Retrieval (CBIR) system, where each image in the database is represented by a compact image signature which is computed using the Nonsubsampled Contourlet Transform (NSCT) and Fuzzy-C-means (FCM) technique. To improve the retrieval accuracy, the proposed system incorporates Least Square Support Vector Machine (LS-SVM) based classifier, Earth Mover's Distance (EMD) and Relevance Feedback Mechanism (RFM). Extensive experiments were carried out to evaluate the effectiveness of the proposed system on SIMPLIcity image database consisting of 1000 images. Experimental results and comparisons show that the proposed CBIR system performs efficiently in image retrieval domain.

I. INTRODUCTION

Over the last few decades, the tremendous growth in the quantity of images stored in digital forms, have raised many demanding issues concerning effective digital image management. The task of retrieving relevant images from a large image database (DB), by measuring similarities between the query image and the database images, using automatically derived features like color, texture, shape etc., has became a potential area of research. High retrieval efficiency and less computational complexity, are the desired characteristics of an efficient CBIR system [1].

Wavelet transform (WT) based low level features, provide a unique representation of the image, and are highly suitable for characterizing textures of the image [2]. But the problem with WT is that it is inherently non-supportive to directionality and anisotropy. To overcome these limitations of WT, recently a theory called Multi-scale Geometric Analysis (MGA) for high-dimensional signals, has been proposed and several MGA tools were developed like Ridgelet, Curvelet and Contourlet etc., [3]. CBIR systems based on these MGA tools have been found to perform better than WT based traditional CBIR schemes [4]. However, contourlet transform is shift sensitive and aliasing in both space and frequency domains. To alleviate these defects of contourlet, a new modified version of contourlet transform is proposed by combining both non-subsampled pyramid (NSP) and non-subsampled directional filter bank (NSDFB), known as non-subsampled contourlet transform (NSCT) [5].

To improve the retrieval accuracy and to decrease the computational complexity, many machine learning techniques have been used in the CBIR paradigm. Researchers have applied Multilayer Perceptron (MLP) for hierarchical medical image classification [6], [7]. However, the training process of MLP is difficult due to the involvement of redundant and irrelevant input variables, which affects the final results of the MLP model severely. Similarly, the training with Support Vector Machine (SVM) is computationally expensive for high dimensional data sets. Therefore, to reduce the time complexity, recently researchers have focused on least square version of SVM, i.e., LS-SVM [8]. LS-SVM is able to deliver higher classification accuracy than the other existing data classification algorithms as reported in the [9].

The performance of a CBIR system, strongly depends both on the availability of the suitable features (for proper representation of the semantic aspects of the images automatically), and also the effectiveness of the used similarity measure [10]. Traditionally, the most commonly used similarity measures in CBIR are City-block / Manhattan distance (MD) and Euclidean distance (ED). MD is easy to compute but it is very sensitive to the orientation of an object, whereas, ED suffers from the scaling problem [11]. In contrast to the distance measures mentioned above, Earth Mover Distance (EMD) measure is a variable size descriptions of distributions [12]. In EMD, ground distance is calculated between two signatures of the same feature space. EMD based CBIR systems have been found to perform better than CBIR systems based on MD or ED [13].

Relevance Feedback Mechanisms (RFMs) have been used as an effective tool to provide significant performance boost in CBIR systems through continuous learning and interaction with the end-users. Most of the RFMs employ two approaches, namely query vector moving technique and feature re-weighting technique [14]. Feature re-weighting technique works by using both the relevant and the irrelevant retrieved image's information [15]. But in most of these cases, time complexity per iteration are high and accuracy of the relevant images are low.

In this article, a novel CBIR system is proposed, where each image in the database is represented by an image signature which is computed using NSCT and FCM clustering. After feature extraction, class of the query image is identified using LS-SVM classifier. EMD is used to find the similarity between the query image's signature and other image signatures of the DB. The proposed algorithm stops, if the user is satisfied with the retrieval result obtained in the first pass of the algorithm, otherwise Fuzzy Relevance Feedback (FRF) using a weighted EMD and fuzzy entropy based feature evaluation technique are used to improve the retrieval result. To evaluate the effectiveness of the proposed CBIR system, its performance is compared against two other well-known existing CBIR systems.

The paper is organized as follows: Section II describes the detail description of the compact signature generation proce-

dure. Multi-class prediction mechanism using LS-SVM based on the image signatures is discussed in Section III. Section IV explains the FRF mechanism with automatic weight updation based on the EMD similarity measure. Experimental results and conclusion is discussed in Section V and Section VI, respectively.

II. Image Signature Computation

The motivation behind using NSCT to extract the image features, is its multi-resolution, localization, multi-directionality and anisotropy properties. An image I of size $M \times N$ of the database is converted from RGB color space to Y-Cb-Cr color space, prior to NSCT decomposition. This ensures that the textural characterization of the image is independent of the color characterization. NSCT decomposition over the intensity plane characterizes the texture information, while the NSCT decomposition over chromaticity planes characterize color. NSCT is applied to Y, Cb and Cr planes of I separately. An l-level NSCT decomposition of an image plane Ip (Y, Cb, Cr), results in 2^l sub-bands, each having the same size $M \times N$ that as of the image I. Let the sub-bands of each image plane (Ip) be denoted by S_b^{Ip}, where $b = 1, 2, 3, \cdots, 2^l$.

Natural images exhibit spatial variations of the textures. So, texture based retrieval of images assume that textures region may not be homogeneous over very large areas. A localized characterization of textures thus becomes necessary. Hence, for each sub-band S_b^{Ip}, a feature map FM_b^{Ip} is computed by using Eq. (1) utilizing the concept of 'local energy' over a neighborhood W_{ij} of size $p \times q$, centered around a coefficient with coordinates (i, j). The size of the window W_{ij}, is determined using the Spectral Flatness Measure (SFM).

$$FM_b^{Ip}(i,j) = \sum_{(p,q)\epsilon W_{ij}} |S_b^{Ip}(p,q)|G(i-p,j-q) \quad (1)$$

where, $G(i, j)$ is a Gaussian low-pass (smoothing) filter of the form

$$G(i,j) = \frac{1}{\sqrt{2\Pi}\sigma}e^{-(\frac{1}{2\sigma^2})(i^2+j^2)} \quad (2)$$

Use of a Gaussian (weighting) window results in less sparse points (i.e., denser feature distributions) as compared to when uniform weighting window is used. After computing the feature maps for an image I of the database, each feature map is clustered using FCM clustering algorithm, into m different clusters.

In the proposed CBIR system, an image I is represented in terms of a signature

$$Sig_j = \{(C_1, w_{C_1}), (C_2, w_{C_2}), ..., (C_m, w_{C_m})\} \quad (3)$$

where, Sig_j is the signature of the j-th image in the database, C_i represents i-th cluster centroid ($i = 1, 2, ..., m$) and w_{C_i} indicates the fraction of pixels belonging to the cluster C_i. C_i constitutes the NSCT features over n feature maps of each Y, Cb and Cr plane and is represented by

$$C_i = [f_{1Y}^i, ..., f_{nY}^i, f_{1Cb}^i, ..., f_{nCb}^i, f_{1Cr}^i, ..., f_{nCb}^i] \quad (4)$$

In the proposed method, the number of decomposition level is set at $l = 3$, which results in $8(= 2^l)$ number of sub-bands for each image plane I_p. The number of feature maps for each image I of the database is 24 ($= 8 \times 3$; 3 color planes/image and 8 feature maps/color plane), and the number of clusters for each feature map is kept at $3(= m)$, as it gives results up to the expectation at minimum cost of computation. Increasing the number of clusters may include finer details with increase in complexity, as well as the uncertainties of characterizing the perceptual content may increase. So in the proposed system, the dimension of an image signature become $75(= 72(= 3 \times 24) + 3)$.

III. Multi-Image-class prediction using LS-SVM Classifier

The most critical drawback of Support Vector Machine (SVM) is its high computational complexity for high dimensional data sets. To reduce the computational demand, the least square version of SVM (LS-SVM) is adopted as classifier in this paper. LS-SVM avoids solving quadratic programming problem and simplifies the training procedure [8]. Considering a linearly separable binary classification problem:

$$(x_i, y_i)_{i=1}^n \quad and \quad y_i = \{+1, -1\} \quad (5)$$

where x_i is an n-dimensional vector and y_i is the label of this vector. LS-SVM can be formulated as the optimization problem:

$$\min_{w,b,e} \mathcal{J}(w,b,e) = \frac{1}{2}w'w + \frac{1}{2}C\sum_{i=1}^n e_i^2 \quad (6)$$

subject to the equality constraints

$$y_i[w'\varphi(x_i) + b] = 1 - e_i \quad (7)$$

where $C > 0$ is a regularization factor, b is a bias term, w is the weights vector, e_i is the difference between the desired output and the actual output and $\varphi(x_i)$ is a mapping function.

The lagrangian for problem of Eq.6 is defined as follows:

$$\mathcal{L}(w, e_i, b, \alpha_i) = \min_{w,b,e} \mathcal{J}(w,b,e) -$$
$$\sum_{i=1}^n \alpha_i\{y_i[w'\varphi(x_i) + b] - 1 + e_i\} \quad (8)$$

where α_i are Lagrange multipliers. The Karush-Kuhn-Tucker (KKT) conditions for optimality $\frac{\partial \mathcal{L}}{\partial w} = 0 \rightarrow w = \sum_{i=1}^n \alpha_i y_i \varphi(x_i)$; $\frac{\partial \mathcal{L}}{\partial e_i} = 0 \rightarrow \alpha_i = Ce_i$; $\frac{\partial \mathcal{L}}{\partial b} = 0 \rightarrow \sum_{i=1}^n \alpha_i y_i = 0$; $\frac{\partial \mathcal{L}}{\partial \alpha_i} = 0 \rightarrow y_i[w'\varphi(x_i) + b] - 1 + e_i = 0$,is the solution to the following linear system

$$\begin{bmatrix} 0 & -Y \\ Y & \varphi\varphi' + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \overline{1} \end{bmatrix} \quad (9)$$

where $\varphi = [\varphi(x_1)'y_1, ..., \varphi(x_n)'y_n], Y = [y_1, ..., y_n]$, $\overline{1} = [1, ..., 1]$, and $\alpha = [\alpha_1, ..., \alpha_n]$.

For a given kernel function K(,) and a new test sample point x, the LS-SVM classifier is given by

$$f(x) = sgn[\sum_{i=1}^n \alpha_i y_i K(x, x_i) + b] \quad (10)$$

LS-SVM was originally develop for binary classification problems. But, a number of methods have been proposed by various researchers for extension of binary classification to

multi-classification problem. However, one need an appropriate method for solving this multi-class problem. It's essentially separate M mutually exclusive classes by solving many two-class problems and combining their predictions in various ways. One such technique which is commonly used is Pairwise Coupling (PWC) or one-vs.-one is to construct binary SVMs between all possible pairs of classes. PWC uses $M*(M-1)/2$ binary classifiers for M number of classes, each of which provides a partial decision for classifying a data point. During the testing of a feature, each of the $M*(M-1)/2$ classifiers votes for one class. The winning class is the one with the largest number of accumulated votes. Hsu.et.al shows that the PWC method is more suitable for practical use than the other methods discussed in [16]. Hence, we use the one against one multi-class image classification method based on the LS-SVM tool by combining all pairwise comparison of binary classifiers [8].

IV. EARTH MOVER'S DISTANCE

Earth Mover's Distance (EMD) by its definition extends to distance between sets or distributions of elements, thereby facilitating partial matches [12]. If p_i be the centroid of each cluster of a signature $P_i = \{(p_1, w_{p_1}), ..., (p_m, w_{p_m})\}$ with m clusters and w_{p_i} the weight of each cluster, similarly q_i be the centroid of each cluster of a signature $Q = \{(q_1, w_{q_1}), ..., (q_n, w_{q_n})\}$ with n clusters and $D = [d_{ij}]$ is the ground distance matrix, where $d_{ij} = d(p_i, q_j)$. Computing EMD thus becomes finding a flow between p_i and q_j which minimizes the overall cost. A flow between P and Q is any matrix $F = (f_{i\ j}) \epsilon R^{m \times n}$. Thus, EMD is defined as the work normalized by the total flow i.e. $EMD(P, Q) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} d(p_i, q_j) f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}}$ subject to the constraints (a) $f_{ij} \geq 0, 1 \leq i \leq m, 1 \leq j \leq n$, (b) $\sum_{j=1}^{n} f_{ij} \leq w_{p_i}, 1 \leq i \leq m$, (c) $\sum_{i=1}^{m} f_{ij} \leq w_{q_j}, 1 \leq j \leq n$, and (d) $\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij} = \min(\sum_{i=1}^{m} w_{p_i}, \sum_{j=1}^{n} w_{q_j})$.

A. Fuzzy Relevance Feedback (FRF) using EMD

The proposed algorithm stops at first pass if the user is satisfied with the retrieval result of the first pass. Otherwise, the information obtained from the set of relevant and irrelevant images as marked by the users are used to automatically specify the weights of the component features. Weights computations are based on a measure defined as Feature Evaluation Index (FEI). The FEI which automatically estimates the importance of an individual feature can be obtained by considering a pattern classification problem. Let $C_1, C_2, ..., C_m$ are m pattern classes in N dimensional features space where class C_j contains, n_j number of samples.

Let the features values along the qth coordinate along class C_j is assigned a fuzzy membership function between 0 and 1, using a standard S-type membership function [17]. Entropy (H) of C_j which gives the measure of intraset ambiguity is then given by

$$H = (\frac{1}{n_j \ln 2}) \sum_{i} S_n(\mu(f_{iqj})); i = 1, 2...n_j \quad (11)$$

where

$$S_n(\mu(f_{iqj})) = -\mu(f_{iqj}) ln\mu(f_{iqj}) - \{1 - \mu(f_{iqj})\} ln\{1 - \mu(f_{iqj})\} \quad (12)$$

is the Shannon's function. Entropy is dependent on the absolute values of membership (μ). $H_{min} = 0$ for μ=0 or 1, $H_{max} = 1$ for μ=0.5 [17].

H_{qj} is the entropy of class C_j along q^{th} dimension over n_j number of sample and denotes intraset ambiguity. H_{qk} is the entropy of class C_k along q^{th} dimension over n_k number of samples. H_{qjk} then denotes the interset ambiguity along q^{th} dimension between class C_j and C_k with $(n_j + n_k)$ number of samples. The FEI_q for the qth component is then defined as

$$FEI_q = \frac{H_{qjk}}{H_{qj} + H_{qk}} \quad (13)$$

The criteria of a good feature is that FEI_q should be decreasing after combining C_j and C_k as the goodness of the qth features in discriminating pattern classes C_j and C_k increases. Lower value of FEI_q, indicates better quality of importance of the qth feature in recognizing and discriminating different classes. This approach is used in estimating the importance of each feature component. The weight w_q of the qth component feature is a function of the evaluated FEI_q:

$$w_q = F_q(FEI_q) \quad (14)$$

In conventional CBIR approaches an image I is usually represented by a set of features, $F = \{f_q\}_{q=1}^{N}$, where f_q is the qth feature component in the N dimensional feature space. In the proposed system, the number of classes are two of which one class constitute the relevant images $I_r = \{I_{r1}, I_{r2}, ..., I_{rm}\}$ and the other consists of the irrelevant images $I_{ir} = \{I_{ir1}, I_{ir2}, ..., I_{irm}\}$. H_{qj} is computed from $I_r^{(q)} = \{I_{r1}^{(q)}, I_{r2}^{(q)}, I_{r3}^{(q)}, ..., I_{rk}^{(q)}\}$. Similarly, H_{qk} is computed from the set of images where, $I_{ir}^{(q)} = \{I_{ir1}^{(q)}, I_{ir2}^{(q)}, I_{ir3}^{(q)}, ..., I_{irk}^{(q)}\}$. H_{qkj} is computed combining both the sets. Images are ranked according to the similarity measures. The user marks the relevant and irrelevant set from 20 returned images, for automatic evaluation of the FEI [17].

The results retrieved from the 1st pass are obtained by measuring EMD between the signature of the query image and the stored images in the database. A better retrieval is then obtained if the user is not satisfied with the result, by using a weighted distance from user RF at successive iterations. In this cases, weights of each component features of different planes is determined from the feature evaluation mechanism described above. Perceptual importance as used in the JPEG 2000 is Y : Cb : Cr = 4 : 2 : 1. Here, the weights are chosen heuristically which is based on the convention "Human visual system is less sensitive to chrominance than luminance". However, an automatic scheme which chooses the weights depending on the color-texture complexity of the image will certainly boost the performance of the CBIR system.

From the set of marked images (I_r and I_{ir}), the weights of the features computed over each Y, Cb and Cr are estimated as follows: For each cluster p_i the NSCT based sub-band features of each of Y, Cb and Cr planes are considered. To compute the FEI along each component plane e.g. plane Y are $F_Y = \{f_{i1Y}, ..., f_{i_qY}\}$, where $i = 1, 2, ..., m$ (m clusters) and $q = 1, 2, ..., n$ (n sub-band features). Similarly features are considered for Cb and Cr plane. The $(FEI)_{qY}$ for the

3

component feature q of Y plane is

$$(FEI)_{qY} = \frac{H_{qTotal}}{H_{qRel} + H_{qIrrel}} \qquad (15)$$

H_{qRel}, H_{qIrrel} and H_{qtotal} are the entropies along the qth dimension of relevant, irrelevant and total returned images respectively. The $(FEI)_{qCb}$ and $(FEI)_{qCr}$ of the other two planes are computed similarly [17].

The overall weight factor for the Y plane is given by

$$W_Y' = \sum_{i=1}^{m} \sum_{q=1}^{n} (FEI)_{qY} \qquad (16)$$

Similarly, the overall weight factors W_{Cb}' and W_{Cr}' for Cb and Cr planes are computed respectively.

A normalization process is used to ensure proper emphasize in each plane even if their features values are of different dynamic ranges. The relative weight factor for the Y plane is as follows:

$$W_Y = \frac{W_Y'}{W_Y' + W_{Cb}' + W_{Cr}'} \qquad (17)$$

Using similar formula, W_{Cb} and W_{Cr} are computed. The weights W_Y, W_{Cb} and W_{Cr} reflects user's different importance on the representative feature map for computing overall similarity between images. Multiplying with the weights, actually modify the ground distance $d(p_i, q_j)$ as describe in Section IV, but keep the distribution, $i.e.$, the number of pixels in each cluster remains unchanged and the total weight remains the same. So, the weighted EMD, $i.e.$, $EMD_g(P, Q)$, and $g\epsilon(W_Y, W_{Cb}, W_{Cr})$ is computed from the work flow:

$$WORK(F, P', Q') = \sum_{i=1}^{m} \sum f_{i,j} d(g(p_i), g(q_j)) \qquad (18)$$

where, the centroids p_i and q_j are transformed to $p_i' = [W_Y p_{i_Y}, W_{Cb} p_{i_{Cb}}, W_{Cr} p_{i_{Cr}}]$ and $q_j' = [W_Y q_{j_Y}, W_{Cb} q_{j_{Cb}}, W_{Cr} q_{j_{Cr}}]$, with the weight updation factor g respectively. The EMD is computed upto k^{th} iteration till it converges, $i.e.$, $W(F^{(K+1)}, P_K'^{(K+1)}, Q_K'^{(K+1)}) \leq W(F^{(K)}, P_K'^{(K)}, Q_K'^{(K)})$. As it is assumed that similar images will have nearly same signature. After multiplying with weights it generate a sequence of work flows $W(F^{(k)}, P_k'^{(k)}, Q_k'^{(k)})$, which is expected to vary in a similar fashion for similar images. As a result, the ranks of the relevant images are not affected much.

In case of irrelevant images, although the signature map may be different, the EMD distance as obtained from the work flows is nearly equal. After FRF, the ground distance (d_{ij}) is modified by the weighting factor W where $W\epsilon[W_Y, W_{Cb}, W_{Cr}]$. The search space that varies over each plane becomes elliptic for unequal weights. As a result, irrelevant images are discarded to a large extent and more relevant images are included due to RF mechanism because the ground distance varies in accordance to the importance of the component planes [17].

V. EXPERIMENT RESULTS

We evaluated the effectiveness of the proposed CBIR algorithm on SIMPLIcity [18] image dataset. SIMPLIcity database consists of 1000 images with 10 different categories and each category contains 100 images. This semantic categorization of the dataset is determined by the creators which reflect the human perception of image similarity. In this proposed approach, we have used 3 levels NSCT decomposition with "9-7" pyramidal and "pkva" directional filters. The experiments were carried out on a Dell Precision T7400 PC with 4GB RAM and was implemented using MATLAB R2008a. To make a fair comparison, the retrieval results obtained using the proposed CBIR system, are compared with some of the state of the art CBIR systems, such as Wang [18] and Lai [19].

LS-SVM classifier is trained with 70% and tested with 30% of the leveled data using stratified random sampling method from the image database. For the training, we use the radial basis function (RBF),$K(x_i, x_j) = exp(-\gamma \parallel x_i - x_j \parallel^2), \gamma > 0$, as the kernel. There are two tunable parameters while using the RBF kernels in LSSVM classifier: C and γ. In our experiment, C = 100 and $\gamma = 0.02$ with accuracy 75% is selected for SIMPLIcity image database.

In image retrieval systems, user consists of a query formulation part. Commonly used query formations are: category browsing, query by concept, query by sketch, and query by example. Our system follows the query by example procedure. In query by example procedure, the user formulates a query by providing an example image to the system. The system initially constructs an image representative feature vector based on various high level and low level features of the image. This image representative feature vector is used in the subsequent steps of the system to classify the image. After the class identification of the query image, similar images are retrieved from the image database. If the user is satisfied with the retrieved results without using RFM then the retrieval process stops at that point. If, however the user is not satisfied, he/she can select top query based relevant images as positive feedbacks and the remaining as negative examples for updating the weight parameters and revising the features. Using these feedbacks, FEIs and weights of the features are recomputed as described in Section IV-A. Then, all the images are re-sorted based on the weighted similarity metric. If the user is still not content with the result, s/he can repeat the process.

Fig. 1 and Fig. 2, shows the retrieval results from SIMPLIcity database without using the LS-SVM classifier, where the results are based on the combination of "*NSCT+EMD+FRF*". In the example of Fig. 1, an image from the "Beach" class was used as the query image, similarly an image from "Elephant" class was input to the system "*NSCT+EMD+FRF*" as the query image in the example of Fig. 2. It is clear from the results given in Fig. 1 and Fig. 2 that the performance of the CBIR system based on the combination "*NSCT+EMD+FRF*" was improved with the iteration of used FRF mechanism. But the overall system performance of the combination "*NSCT+EMD+FRF*" is not satisfactory as is evident from the high number of misclassified images (6/20 in Fig. 1 and 5/20 in Fig. 2 after second iteration of FRF).

Therefore, to improve the performance of the retrieval system we used the "*NSCT+LSSVM+EMD+FRF*" combi-
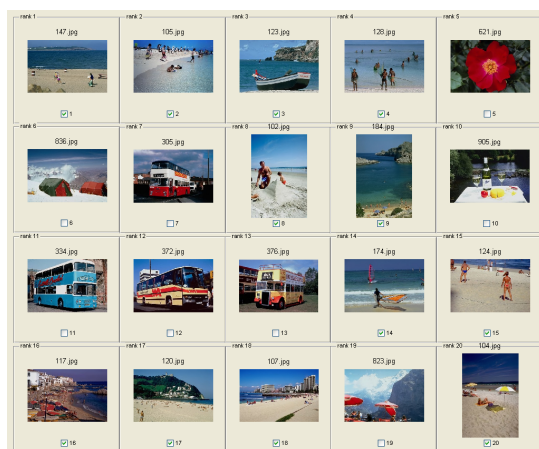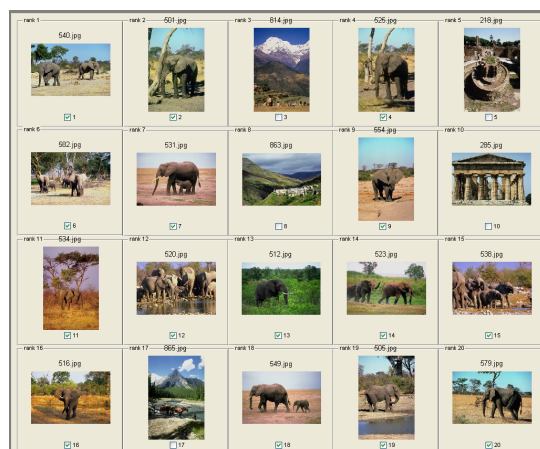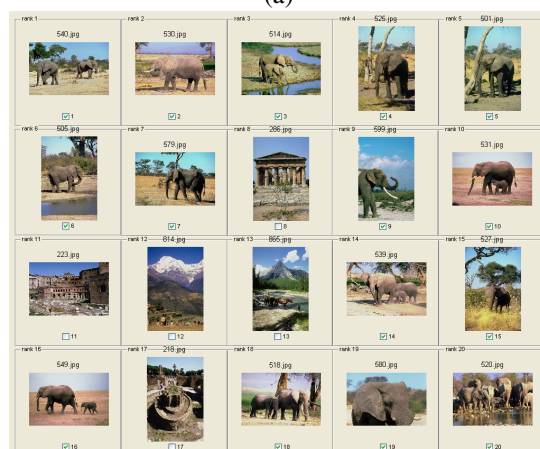
(a)
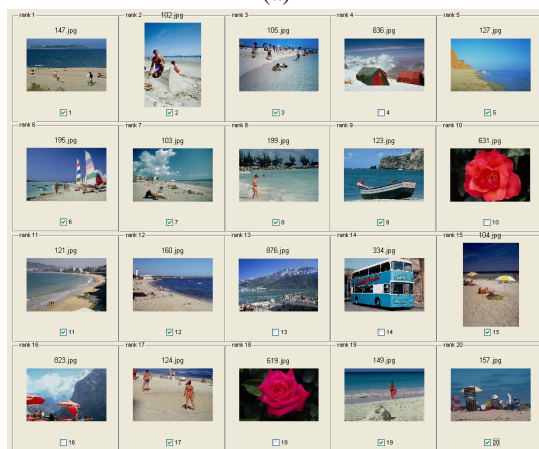
(b)

Fig. 1. Results on SIMPLIcity using NSCT+EMD+FRF (Top left side image is the query image) (a) First pass of the retrieval set Using NSCT features and EMD (12/20) (b) Second Iteration with weighted EMD (14/20, improved ranking).

(a)

(b)

Fig. 2. Results on SIMPLIcity using NSCT+EMD+FRF (Top left side image is the query image) (a) First pass of the retrieval set Using NSCT features and EMD (15/20) (b) First Iteration with weighted EMD (15/20, improved ranking).

nation incorporating the advantages of the LS-SVM classifier. Fig. 3 and Fig. 4, show the examples of the retrieval results obtained by our proposed system based on the combination "*NSCT+LSSVM+EMD+FRF*", using the same query images (query images from "Beach" and "Elephant" classes). It is evident from the results given in Fig. 3 and Fig. 4 that our proposed system based on the combination "*NSCT+LSSVM+EMD+FRF*" performs better than the combination "*NSCT+EMD+FRF*". From the given results shown in Fig. 3 and Fig. 4, it is clear that in these cases not only all the retrieved images are from the same class as of the query image, but after iteration of the FRF mechanism, the ranking of the retrieval images gets improved.

To make a fair comparison with other existing CBIR systems, we computed two statistical measures: average precision and average recall as the evaluation metrics. It computes the precision and recall for all the query image, and then takes the average of precisions and recalls over all the images as the final evaluation result.The precision and recall is defined as follows;

$$Precision\ (P) = \frac{N_{RIR}}{N_{RIR} + N_{IRIR}} \qquad (19)$$

and

$$Recall\ (R) = \frac{N_{RIR}}{T_{RID}} \qquad (20)$$

where, N_{RIR} is the Number of relevant images retrieved, N_{IRIR} is the Number of irrelevant images retrieved and T_{RID} is the total number of relevant images in the database.

We compared our method with two other existing methods such as [18] and [19]. Only the retrieval results obtained from low level visual features (without using FRF) were used to evaluate the superiority of our approach than the two aforementioned methods. Table. I demonstrate that our method outperformed for all the classes in comparison with the method proposed by Wang et al. [18] and for some classes the method proposed by Lai et al. [19] gives better result. But in overall, our method yields the best average precision value. The "Bold" values in the Table. I indicate the highest retrieval performance in terms of average precision.
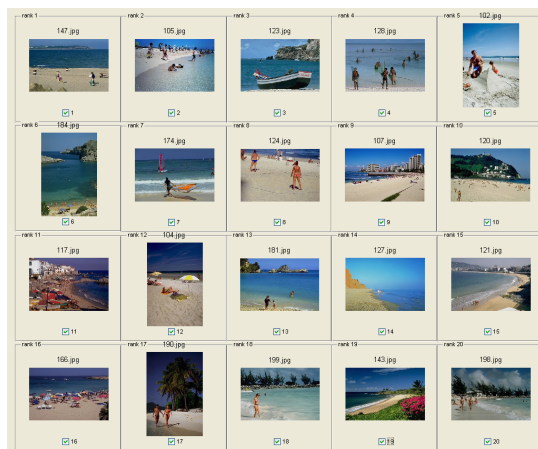
(a)

(b)

Fig. 3. Results on SIMPLIcity using NSCT+LSSVM+EMD+FRF (Top left side image is the query image) (a) First pass of the retrieval set Using NSCT features and EMD (20/20) (b) Second Iteration with weighted EMD (20/20, improved ranking).

(a)

(b)

Fig. 4. Results on SIMPLIcity using NSCT+LSSVM+EMD+FRF (Top left side image is the query image) (a) First pass of the retrieval set Using NSCT features and EMD (20/20) (b) First Iteration with weighted EMD (20/20, improved ranking).
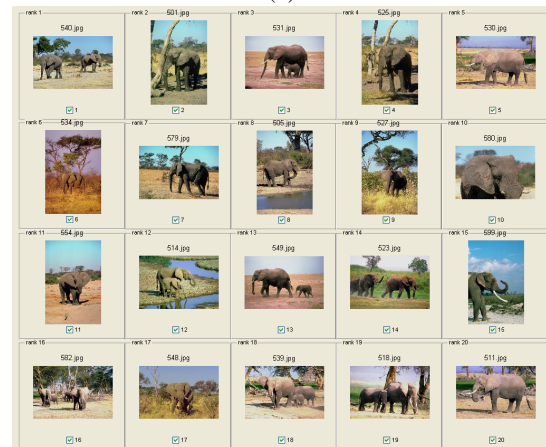
Class	Wang [18]	IGA-CBIR [19]	Proposed Method
Africans	48	**99**	84.83
Sea	32	**85**	71.78
Building	35	**93**	83.95
Buses	36	80	**88.41**
Dinosaurs	95	71	**100**
Elephant	38	**100**	96.42
Flower	42	82	**99.74**
Horse	72	79	**98.11**
Mountain	35	56	**69.45**
Food	38	61	**92.79**
Average	47	80.6	**88.55**

TABLE I. COMPARISONS WITH OTHER EXISTING CBIR SYSTEMS WITHOUT RELEVANCE FEEDBACK IN TERMS OF AVERAGE PRECISION

We also compared our FRF model with two existing methods of RF mechanism. In [20], Rui et al. employed the weight assignment technique and obtained 52% average precision in three to four iterations from SIMPLIcity database. The other method proposed by Min et al., [21] used fuzzy based SVM in their RF model and achieved up to 80% average precision in three to four iterations whereas, our FRF model have achieved nearly 98% average precision in the same number of iterations as shown in Fig. 5.

VI. CONCLUSION

In this paper, we have presented an efficient interactive CBIR system. NSCT and FCM is used to extract a compact image signature based on the geometrical information. To enhance the retrieval result accuracy, pre-classification using LS-SVM classifier is used to classify the query image and retrieved the similar types of images from that class only. If the pre-classification is wrong then our whole retrieval result will be wrong. To overcome this problem we are trying to implement fuzzy ranking membership function. In addition, FRF mechanism with fuzzy entropy based feature evaluation technique is used with EMD to improve the retrieval results further. Experimental results of the proposed approach have shown the significant improvement in retrieval performance. As the future scope of research, the proposed technique could be tested for retrieval of videos in conjunction with motion information.

REFERENCES

[1] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40, pp. 1–60, 2008.
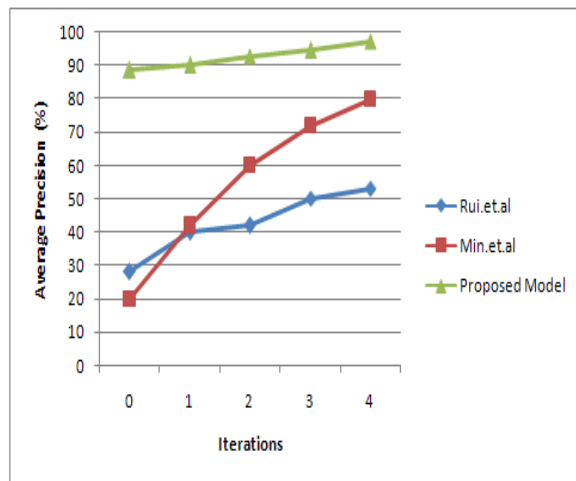
Fig. 5.   Comparison with others Relevance feedback Mechanism

[2]   M. Acharyya and M. K. Kundu, "Extraction of features using M-band wavelet packet frames and their neuro-fuzzy evaluation for multi-texture segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1639–1644, 2003.

[3]   M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, pp. 2091–2106, 2005.

[4]   J. Bai, X. Wang, and L. Jiao, "Image retrieval based on color features integrated with anisotropic directionality," *Journal of Systems Engineering and Electronics*, vol. 21, pp. 127–133, 2010.

[5]   A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, pp. 3089–3101, 2006.

[6]   H. Pourghassem and H. Ghassemian, "Content-based medical image classification using a new hierarchical merging scheme," *Computerized Medical Imaging and Graphics*, vol. 32, no. 8, pp. 651–661, 2008.

[7]   D. Mohandass and J. Janet, "A novel image retrieval technique for enhanced telemedical applications," in *Proc. of 3rd Int. Conf. on Trendz in Information Sciences and Computing (TISC)*, vol. 1, 2011, pp. 46–50.

[8]   J. A. K. Suykens, T. V. Gestel, J. D. Brabanter, B. D. Moor, and J. Vandewalle, *Least Squares Support Vector Machines*.   World Scientific, Singapore, 2002.

[9]   W. Chu, C. J. Ong, and S. S. Keerthi, "An improved conjugate gradient scheme to the solution of least squares svm," *IEEE Transactions on Neural Networks*, vol. 16, no. 2, pp. 498–501, 2005.

[10]   D. Heesch, "A survey of browsing models for content based image retrieval," *Multimedia Tools Application*, vol. 40, no. 2, pp. 261–284, 2008.

[11]   S. Santini and R. Jain, "Similarity measures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 871–883, 1999.

[12]   Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Interational Journal of Computer Vision*, vol. 40, no. 2, pp. 99–121, 2000.

[13]   W. Horsthemke, D. Raicu, and J. Furst, "Task-oriented medical image retrieval," in *Proc. of MICCAI 2007 Workshop on Content-based Image Retrieval for Biomedical Image Archives: Achievements, Problems, and Prospects*, 2007, pp. 31–44.

[14]   Y. Rui, T.S.Huang, and S.Mehrotra, "Relevance feedback: a power tool for interactive content-based image retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, pp. 644–655, 1998.

[15]   Z. Jin, I. King, and X. Li, "Content-based image retrieval by relevance feedback," in *Proc. of VISUAL 2000*, 2000, pp. 521–529.

[16]   C. W. Hsu and C. J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transaction on Neural Networks*, vol. 13, no. 2, pp. 425–425, 2002.

[17]   S. K. Pal and D. D. Majumder, *Fuzzy Mathematical Approach To Pattern Recognition*.   Willey Eastern Limited, 1985.

[18]   J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics-sensitive integrated matching for picture libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 947–963, 2001.

[19]   C. C. Lai and Y. C. Chen, "A user-oriented image retrieval system based on interactive genetic algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 10, pp. 3318–3325, Oct 2011.

[20]   Y. Rui and T. S. Huang, "Optimizing learning in image retrieval," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2000, pp. 236–243.

[21]   R. Min and H. D. Cheng, "Effective image retrieval using dominant color descriptor and fuzzy support vector machine," *Pattern Recognition*, vol. 42, pp. 147–157, 2009.