Interactive Semantic Image Retrieval

Pushpa B. Patil* and Manesh B. Kokare**

Abstract—The big challenge in current content-based image retrieval systems is to reduce the semantic gap between the low level-features and high-level concepts. In this paper, we have proposed a novel framework for efficient image retrieval to improve the retrieval results significantly as a means to addressing this problem. In our proposed method, we first extracted a strong set of image features by using the dual-tree rotated complex wavelet filters (DT-RCWF) and dual tree-complex wavelet transform (DT-CWT) jointly, which obtains features in 12 different directions. Second, we presented a relevance feedback (RF) framework for efficient image retrieval by employing a support vector machine (SVM), which learns the semantic relationship among images using the knowledge, based on the user interaction. Extensive experiments show that there is a significant improvement in retrieval performance with the proposed method using SVMRF compared with the retrieval performance without RF. The proposed method improves retrieval performance from 78.5% to 92.29% on the texture database in terms of retrieval accuracy and from 57.20% to 94.2% on the Corel image database, in terms of precision in a much lower number of iterations.

Keywords—Content-based Image Retrieval (CBIR), Relevance Feedback (RF), Rotated Complex Wavelet Filters (RCWFs), Dual Tree Complex Wavelet, and Image retrieval

1. INTRODUCTION

1.1 Motivation

Recently, there is a rapid growth of digital image data on the Internet and in digital libraries. The advent of the Internet has made information sharing and access easier. Internet users are indulging in information exchange. Retrieving information from the World Wide Web has become a common practice. However, with the day-by-day increase in the size of the web and the increase in the heterogeneity of information, due to there being an abundant amount of information, these things have made classical information retrieval techniques ineffective. Searching for and retrieving information as desired has become a very important challenge. Hence, nowaday's image retrieval has became an active research field. As the databases grew larger, the traditional keywords based method to retrieve a particular image has become inefficient due to the following limitations:

Corresponding Author: Pushpa B. Patil
*   Dept. of Computer Science and Engineering, BLDEA's V. P. Dr. P. G. H. College of Engineering and Technology, Bijapur, Karnataka, India (pushpa_ms@rediffmail.com)
** Dept. of Electronics and Telecommunication Engineering, SGGS, Institute of Engineering and Technology, Nanded, Maharastra, India (mbkokare@yahoo.com)

- Image annotation is a tedious task, since it is practically impossible to annotate all of the images in a large scale database.
- For a large dataset, it requires more skilled labors to annotate the images in a database manually.

To overcome these limitations, researchers have turned their attention to CBIR. There are different ways to retrieve the images in CBIR. [1]-[4] presented a comprehensive and recent extensive literature survey on content based image retrieval. In CBIR systems, low level image features are extracted based on visual content, such as color, shape, and texture, which are represented by feature vectors instead of a set of keywords. However, user's are more interested in high level concepts than in retrieving similar images that are based on a simple low level feature. Hence, there is a big challenge in CBIR to reduce this semantic gap between the low level features and high level concepts. In order to reduce this gap, relevance feedback was introduced into CBIR [5]-[6]. Relevance feedback was initially developed for document retrieval. However, now it has become popular in CBIR within a short of period and it will remain an active research area, due to there being more ambiguities that arise in the interpretation of images than with words, which makes user interaction a neccessity. In addition, judging an image is faster than judging a document, since an image reveals its content almost instantly to a human obsever [12]. To overcome these problems we have proposed a novel method in this paper. For our proposed method, we first used our recently designed 2-D rotated complex wavelet filters [18] and dual-tree complex wavelet transform jointly in order to efficiently extract the textural features of textured and real world scenic images in 12 different orientations. Second, to reduce the significant gap between low level features and high level concepts, we are proposing a novel SVM based relevance feedback algorithm, which provides efficient retrieval performance that has very few feedback iterations. Third, extensive experiments on standard database show that there are significant improvements in terms of retrieval performance, as compared to earlier approaches based on retrieval without feedback [18], with relevance feedback based on AdaBoost[26], Single_RBF, and the RBF guassian Function[25]. The proposed method improves retrieval performance from 78.5% to 92.29% on the texture database in terms of retrieval accuracy and from 57.20% to 94.2% on the Corel Image Database, in terms of precision in a much lower number of iterations.

1.2 Related Work

Recently, many researchers began to consider the RF as a classification or semantic learning problem. In this approach a user provides positive and/or negative examples, and the systems learn from such examples to separate all data into relevant and irrelevant groups. Hence, many classical machine learning schemes may be applied to the RF, which include decision tree learning [7], Bayesian learning [8]-[9], Support Vector machines [10], boosting [11], and so on. There is a good review on RF in [12]. The process of learning is a very difficult task in RF [12]-[14], due to the following reasons:

- Training data is very small, which is less than the dimension of the feature space. This makes it difficult to apply most of the learning methods, such as the linear discriminate fisher classifier and relevance (RVM). Though, the RVMs are sparser than the SVMs and use less number of kernel functions.

- Training data is asymmetrical, which creates too much of an imbalance between the relevant and irrelevant images.
- In RF, for every iteration we have to perform both training and testing online, which requires more real time usage.

Recently, most of the work in RF is based on SVMs [14]-[17] because they minimize the measure of errors on the training set, while simultaneously maximizing the margin between relevant and irrelevant images. A SVM is a highly effective mechanism for avoiding over fitting, which leads to a good generalization. It is a sparse model, so the process of learning and evaluation is faster for the medium-sized training data.

For a visual representation of the images, we employed the global texture features presented in [18], which provide very efficient performance. Much of the work on RF uses the low-level representation using a discrete wavelet transform (DWT) [15], Gobor filters [16], and a Co-occurrence matrix [19][20] for extracting texture features. In order to retrieve general purpose images like artificial objects and natural scenes textural features are usually combined with colors and shapes to obtain a better retrieval performance. However, they still suffer from poor directional sensitivity, shift variants, and redundancy. From these combined features we may get better retrieval performance, but not an efficient one because as we increased the number of features, which increases the dimensionality of the feature space. With such a high dimensional feature space, RF may become impractical for even medium sized databases [14]. In order to store and process these high dimensional feature vectors it requires more memory space and time. So, to make a retrieval system efficient, we have to consider two factors—namely, time complexity and space complexity—together with the better retrieval performance. To overcome the above problem, we are proposing the use of new rotated complex wavelet filters for feature extraction.

1.3 Main Contribution

In this paper we have used our earlier recent work [18] to extract more compact effective low-level features, in order to improve the retrieval performance in terms of speed, storage, and accuracy by using the rotated complex wavelet filters and dual tree complex wavelet transform jointly. Furthermore, to reduce the significant gap between low-level features and high-level concepts, we have proposed a new RF approach that uses the left skewed relevant binary tree of the SVM while neglecting the right skewed irrelevant binary tree in every iteration of the feedback. This helps to exhibit the better results in a lower number of iterations. The proposed method has a threefold advantage over earlier approach. First, the proposed RF framework provides efficient retrieval performance in very few feedback iterations. Second, a proposed approach uses both the relevant and irrelevant examples for learning. Finally, the proposed approach uses the linear kernel function, which gives better performance even though the training samples are smaller than the dimensionality of feature space. Our extensive experiments, which used the proposed RF with SVM on a standard texture database and Corel database, show significant improvements with respect to retrieval performance in comparison with the earlier RF approach based on AdaBoostRF[26], Single_RBF, and the RBF Gaussian Function [25].

The rest of the paper is organized as follows: we briefly discuss the dual-tree complex wavelet, and dual tree rotated complex wavelet in Section 2. The proposed RF method using SVM is

discussed in Section 3. In Section 4, image retrieval and feature database creation are discussed. In Section 5, experimental results are discussed and finally, the conclusion is given in Section 6.

2. VISUAL REPRESENTATIONS OF THE IMAGES

2.1 DT-CWT

Real DWT has poor directional selectivity and it lacks shift invariance. Drawbacks of the DWT are overcome by the complex wavelet transform (CWT) by introducing limited redundancy into the transform. But still it suffers from problems where things like no perfect reconstruction is possible in the case of using CWT decomposition beyond Level 1, when the input to each level becomes complex. To overcome this, Kingsbury [21] proposed a dual tree complex wavelet transform (DT-CWT), which provides perfect reconstruction along with providing the other advantages of a complex wavelet, which is DT-CWT. This introduces a limited amount of redundancy and provides perfect reconstruction along with providing the other advantages of complex wavelets. The DT-CWT is implemented using separable transforms and by combining subband signals appropriately. Even though it is non-separable yet, it inherits the computational efficiency of separable transforms. Specifically, the 1-D DT-CWT is implemented by using two filter banks in parallel and operates on the same data. For d-dimensional input, a L scale DT-CWT outputs an array of real scaling coefficients corresponding to the low pass subbands in each dimension. The total redundancy of the transform is 2^d and independent of L. The mechanism of the DT-CWT is not covered here. Please refer to [22] and [23] for a comprehensive explanation of the transform and details of filter design for the trees. A complex valued $\psi(t)$ can be obtained as:

$$\psi(x) = \psi_h(x) + j\,\psi_g(x) \tag{1}$$

Where $\psi_h(x)$ and $\psi_g(x)$ are both real-valued wavelets. The impulse responses of 6 wavelets associated with 2-D complex wavelet transform are illustrated in Fig. 1.
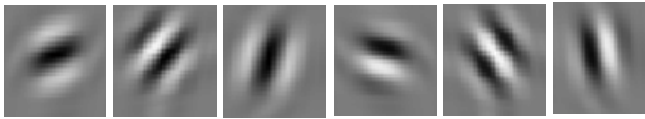

Fig. 1. Impulse response of 6 wavelet filters of a complex wavelet

2.2 DT-RCWF

Dual tree rotated complex wavelet filters were designed in 2005 [18]. Directional 2D RCWF are obtained by rotating the directional 2D DT-CWT filters by 45° so that decomposition is performed along new directions, which are 45° apart from the decomposition of CWT. The size of a newly obtained filter is $(2N-1)X(2N-1)$, where N is the length of the 1-D filter. The decomposition of an input image with 2-D RCWF followed by a 2-D downsampling operation is performed up to the desired level. The computational complexity associated with RCWF decompo-

sition is the same as that of a standard 2-D DT-CWT, if both are implemented in the frequency domain. The set of RCWFs retains the orthogonality property. The 6 subbands of 2D DT-RCWF gives information that is strongly oriented at ($30^{\circ}, 0^{\circ}, -30^{\circ}, 60^{\circ}, 90^{\circ}, 120^{\circ}$). The mechanism of the DT-RCWF is explained in our earlier work [18]. The 2D DT-CWT and RCWF jointly gives 12 different directional information on images in the directions $\left\{ \begin{array}{l} (+15^{\circ}, +45^{\circ}, +75^{\circ}, -15^{\circ}, -45^{\circ}, -75^{\circ}), \\ (0^{\circ}, +30^{\circ}, +60^{\circ}, +90^{\circ}, 120^{\circ}, -30^{\circ}) \end{array} \right\}$. Whereas, the standard DWT gives the directional information of images in only 4 directions $\left\{ 0^{\circ}, \pm 45^{\circ}, 90^{\circ} \right\}$. The impulse response of 6 wavelets associated with a rotated complex wavelet filter is shown in Fig. 2.
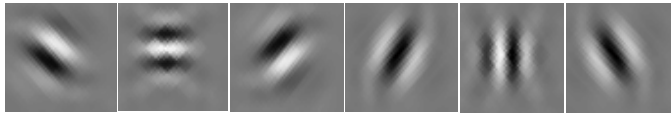

Fig. 2. Impulse response of six rotated complex wavelet filters

3. PROPOSED SVM BASED RELEVANCE FEEDBACK FRAMEWORK

The fundamental concept of RF is to learn the semantic gap between the low-level features and the high-level concepts by establishing interactions between the user and the retrieval system. This should be done so that the system refines the retrieval performance based on the relevance judgments provided by the user. Generally speaking, RF is designed to bridge the semantic gap between low level features and high level concepts (users feedback) for enhancing performance.

3.1 Overview of the Proposed Framework

Fig. 3 shows the block diagram of the proposed system. First, the user provides a query to a CBIR system for searching for desired images in the database. Then, the CBIR system computes the similarity between the user query and the images in the database by extracting the low-level features. Images with high similarities are returned to the users in the initial stage. Second, the user judges the relevance of the initially returned results and submits his/her judgments to the CBIR system. Where, U/R is the user's relevance feedback.

A RF algorithm reduces the significant gap between low−level features and high-level concepts by refining the initial retrieval results based on the user relevance judgments, and it returns an improved set of results to the user. Typically, the above procedure is repeated for a number of times to achieve satisfactory results. Finally, after some feedback iterations, the performance of the system remains the same. Unlike traditional relevance feedback, in our approach after every feedback iteration we get 2 sets of images—namely, a relevant image set and an irrelevant image set. For every iteration we have provided, only the images that are the most relevant to the learning system are needed to learn the semantic concept of the images based on user relevance judgments. This greatly improves retrieval performance within a few feedback iterations and
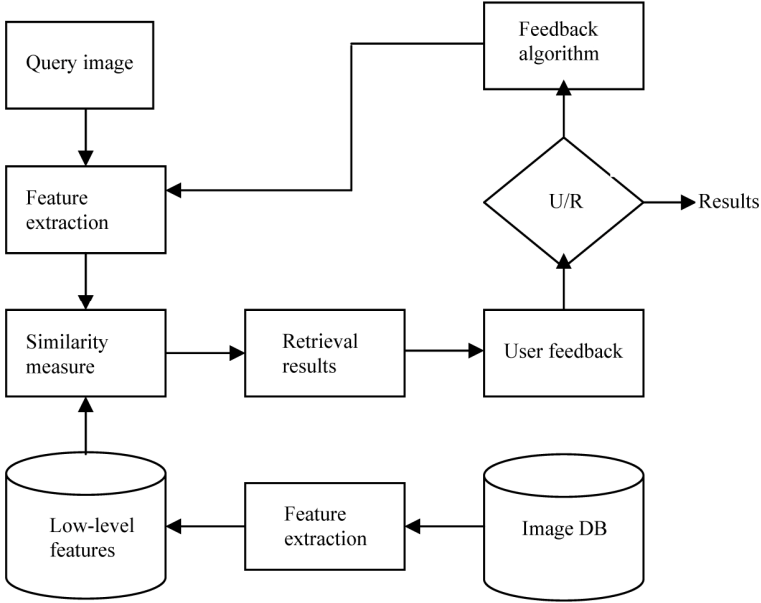
Fig. 3. System architecture

optimizes the testing process.

3.2 Support Vector Machine Framework

Here we briefly introduce the basic concepts of two classes of SVMs [24]. On pattern classification problems, SVMs provide a very good generalization performance in empirical applications. We begin our discussion of support vector machines for a two-class classification problem by using linear models of the form:

$$y(X) = W^T \phi(X) + b \tag{2}$$

Where $\phi(X)$ denotes a fixed feature-space transformation, and we have made the bias parameter b explicit. The training data set comprises N input vectors $X_1, \cdots, X_N$, with corresponding target values $t_1, \ldots, t_N$, and new data points are classified according to the sign of $y(X)$.

Given a training set of instances labeled pairs (X_i, t_i), $i = 1 \ldots N$, $X_i \in R^n$ and $t \in \{1, -1\}$. The SVM requires a solution for the following optimization problem:

$$\min_{w, b, \xi} \quad \frac{1}{2} W^T W + C \sum_{i=1}^{l} \xi_i \tag{3}$$

$$\text{Subject to:} \quad t_i (W^T \phi(X_i) + b) \geq 1 - \xi_i \tag{4}$$

$$\xi_i \geq 0$$

Where training vectors X_i are mapped into a higher dimensional space by the function ϕ. SVM finds a linear separating hyper-plane with a maximal margin in this higher dimensional space. Furthermore, $k(X, X') = \phi\left(X\right)^T \phi(X')$ is called the kernel function. The basic kernel functions of an SVM are:

1) Linear: $k(X_i, X_j) = X_i^T X_j$

2) Radial basis function (RBF): $k(X_i, X_j) = \exp(-\gamma \left\| X_i - X_j \right\|^2), \quad \gamma > 0$

3) Polynomial: $k(X_i, X_j) = \left(\gamma X_i^T X_j + r\right)^d, \quad \gamma > 0$

where γ, r, and d are the kernel parameters.

3.3 Kernel Selection for a Proposed RF

The RBF kernel nonlinearly maps samples into a higher dimensional space. Furthermore, the linear kernel is a special case of RBF, and they were shown that the linear kernel with a penalty parameter C has the same performance as a RBF kernel with some parameters (C, γ). In addition, the sigmoid kernel behaves like RBF for a certain parameter. The reason for this, is the number of hyper parameters, which influences the complexity of kernel selection. Finally, the RBF kernel has less numerical difficulties as compared to polynomial kernels. Since polynomial kernel values may go to infinity if $\gamma X_i^T X_j + r > 1$ or zero $\gamma X_i^T X_j + r < 1$ when the degree is large. However, if the dimension of the feature space is large, one may not need to map data to a higher dimensional space. That is, the nonlinear mapping does not improve the performance. So we have used the linear kernel with $C = 0$. It performs better when the number of training samples is smaller than the dimensionality of the feature space.

3.4 The Proposed Semantic Image Retrieval System

The following algorithm reduces the semantic gap between the low-level features and the high-level concepts by using SVM.

Algorithm 1: Proposed SVM-Based Relevance Feedback

Input: q: user query
DB: Image database
  P: Relevant images
  N: Irrelevant images
Output: Result
    Begin
          Result= CBIR (DB, q);
Repeat until user satisfaction or result remains the same
    (P, N)=Labeling (Result);
    T= (P U N);
    (PI, NI)=SVMLearner (T, DB);

```
            for each x Є PI do
                Dx =CanberraDist(x, query);
            end
        SortDist (D);
    Result=DisplayTop20 (D);
    DB=DB-NI;
    Repeat end
end
```

In summary, the pseudo code of the proposed method is presented in Algorithm 1, where the *CBIR* function follows the traditional content-based image retrieval mechanism by using low-level features. The *Labeling* function, which is an interactive mechanism between the user and the retrieval system in order to get refined results based on user judgments about the result of every iteration. To reduce the semantic gap, we have used SVM for the semantic learning of the retrieval system, which returns the results as relevant and irrelevant based on the relevant judgments of the user. From these results we have considered only relevant images for speeding up the retrieval system and for better semantic learning in aspect to the user perception. For optimizing the testing process, the irrelevant images that are obtained from the learner are removed from the image database. In each iteration, the consideration of relevant images helps to reduce the semantic gap between the low-level features and high-level user perception. The relevant images are ranked by using the Canberra distance measure eq. (11). Hence, this improves the performance of the semantic image retrieval system.

4. FEATURE DATABASE CREATION

To construct the feature vectors of each image in the database, we decomposed each image using DT-CWT and DT-RCWF up to third level. Features based on the Energy and Standard Deviation (STD) were computed separately on each subband and the feature vector was formed using these two parameter values. The retrieval performance, combined with these two feature parameters, always performs better than that using these features individually [18]. The Energy E_k and Standard Deviation σ_k of kth subband is computed as follows:

$$E_k \quad = \quad \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left| W_k(i,j) \right| \tag{5}$$

$$\sigma_k = \left[\frac{1}{M \times N} \sum_{i=1}^{N} \sum_{j=1}^{M} \left(W_k(i,\ j) - \mu_k \right)^2 \right]^{\frac{1}{2}} \tag{6}$$

where $W_k(i,\ j)$ is the kth wavelet-decomposed subband, $M \times N$ is the size of the wavelet decomposed subband, and μ_k is the mean of the kth subband. The resulting feature vector

using energy and the standard deviation are $\bar{f}_E = \begin{bmatrix} E_1 & E_2 & ... & E_n \end{bmatrix}$ and $\bar{f}_\sigma = \begin{bmatrix} \sigma_1 & \sigma_2 & ... & \sigma_n \end{bmatrix}$ respectively. So the combined feature vector is:

$$\bar{f}_{\sigma\mu} = \begin{bmatrix} \sigma_1 & \sigma_2 & ... & \sigma_n & E_1 & E_2 & ... & E_n \end{bmatrix} \tag{7}$$

4.1 Normalization

It is important to normalize the data before applying it to the proposed RF using SVM and AdaBoost learners. Normalization avoids features in greater numeric ranges and it dominates those in smaller numeric ranges. It also avoids the numerical difficulties during the calculation. Because kernel values usually depend on the inner products of feature vectors (e.g., the linear kernel and the polynomial kernel for large attribute values might cause numerical problems). So we normalized the feature vector $\bar{f}_E$ and $\bar{f}_\sigma$ by applying the following statistical normalization method as given in Eq. (8) and (9) respectively. We observed that with normalization the results are better in relevance feedback, as compared to the results without normalization.

$$\overline{f}_{VE} = \frac{\bar{f}_E - \mu_{\bar{f}_E}}{\sigma_{\bar{f}_E}} \tag{8}$$

$$\overline{f}_{V\sigma} = \frac{\bar{f}_\sigma - \mu_{\bar{f}_\sigma}}{\sigma_{\bar{f}_\sigma}} \tag{9}$$

Where $\mu_{\bar{f}_E}$, $\mu_{\bar{f}_\sigma}$ and $\sigma_{\bar{f}_E}$, $\sigma_{\bar{f}_\sigma}$ are the mean and the standard deviation of $\bar{f}_E$, $\bar{f}_\sigma$ respectively. Finally, the resultant feature vector will be the combined normalized vector of:

$$f_V = \begin{bmatrix} \overline{f}_{V\sigma}, \overline{f}_{VE} \end{bmatrix}. \tag{10}$$

The features vectors are constructed using these two parameters. The length of the feature vector will be equal to (no. of subbands × no. of feature measures used in combination). Let us assume that there are a total of n subbands and a combination of two feature measures, then the length of feature vector will be equal to ($n \times 2$). For the creation of a feature database, the above procedure is repeated for all the images in the database and these feature vectors are stored in the feature database.

4.2 Image Matching

The query image is one of the 1,856 images from the texture image database. The query image is further processed to compute the feature vector, as given in Section 4.1. The Canberra distance metric is used as a similarity measure. If x and y are the feature vectors of the database and query image, respectively, and have dimension d, then the Canberra distance is

given by:

$$Canb\ (x,\ y)\ =\ \sum_{i=1}^{d}\ \frac{|x_i - y_i|}{|x_i| + |y_i|} \tag{11}$$

5. EXPERIMENTS AND DISCUSSION

To evaluate the performance of a proposed system, we have used the Brodatz texture photographic album [18] and the Corel Image Database [25]. The experiments were conducted using a MATLAB 7.0 with an Intel core2Duo, which is a 1 GB RAM machine.

5.1 Texture Image Database

The texture database used in our experiment consists of 116 different textures [18]. We used 108 textures from the Brodatz texture photographic album, 7 textures from the USC database, and 1 artificial texture. The size of each texture image is 512×512. Each 512×512 image is divided into sixteen 128×128 non-overlapping subimages, thus creating a database of 1,856 texture images.

5.2 Corel Image Database

This database contains 1,000 color photographs that have a resolution of 384x256 pixels and that cover a wide range of semantic categories, from natural scenes to artificial objects [25]. The database is partitioned into 10 categories, each with 100 photographs.

5.3 Performance Measures

In order to obtain our experimental results, we conducted 2 different sets of experiments by using the proposed method, which we applied to 2 different standard databases. We did so in order to compare the retrieval performance in RF while still considering the top 20 image retrieval. For a retrieval task, it is significant to define a suitable metric for performance evaluation. We employed the following two performance measures:

$$recall(N) = \frac{R_n}{M} \tag{12}$$

$$precision(N) = \frac{R_n}{N} \tag{13}$$

Where M is the total number of relevant matches in the database, N is the number of retrievals and R_n is the number of relevant matches amongst the retrievals.

Our first experimental results were evaluated by randomly selecting one query image from each of the 116 classes from the texture database. For each experiment, one image was selected at random as being the query image from each category and thus the retrieved images were

obtained. Then, the users were asked to identify the images that are related to their expectations from the retrieved images. These selected images were used as the feedback images for the next iteration. Finally, we computed the average accuracy of all of the categories in the database. Each image category contains 16 images. The feedback processes were performed 5 times. However, in RF, we can perform the number of iterations repeatedly until the result remains the same or user satisfaction has been obtained. The reported results of the average accuracy are obtained by taking an average over the 116 texture database queries. Fig. 4 provides a detailed comparison of the average retrieval accuracy that was obtained by using SVMRF and AdaBoostRF on every feedback iteration of the randomly selected image from each category of the texture database.

The proposed RF using a SVM gives a better retrieval performance on the Brodatz texture database, which contains texture images. Second, from Fig. 4 we observed that the retrieval performance of SVMRF was better than AdaBoostRF [26]. However, there is a rapid increase in retrieval performance with each feedback iteration of the proposed RF when both learning algorithms are used. Retrieval performance is improved from 91.75% to 92.29% in comparison to AdaBoostRF. The results are tabulated in Table 1.

Our second experimental results were evaluated on a Corel image database. In this database, there are 10 categories of images and in each category 100 natural color images. For testing we have randomly selected 5 images from each category as a query image (altogether 50 images). The reported results of average precision are obtained by taking an average over the 50 queries.

Table 1. The average accuracy on each feedback iteration for the texture database

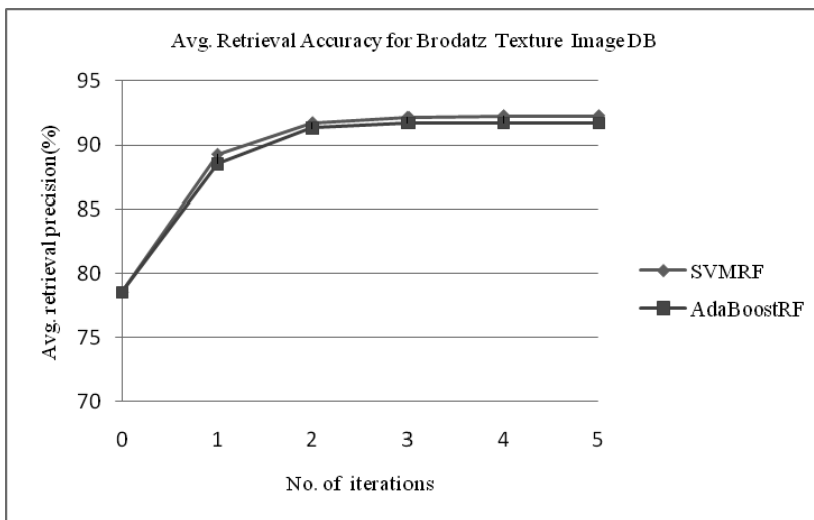| Approach	CBIR (without RF)	1st iteration	2nd iteration	3rd iteration	4th iteration	5th iteration
SVMRF	**78.50**	**89.27**	**91.75**	**92.18**	**92.29**	**92.29**
ADABoostRF [26]	78.50	88.52	91.32	91.70	91.70	91.70

Fig. 4. Average accuracy versus iteration curves for texture images

Fig. 5 describes the detailed comparison of the average retrieval performance obtained when using SVMRF, AdaBoostRF [26], Single_RBF, and the RBF Gaussian function [25] on every feedback iteration on Corel images. We observed in Table 2, that the proposed method yields better retrieval performance than the Single_RBF and RBF Gaussian function that was proposed by Rongtao et al. in 2007 [25] as the number of iterations increased. However, in the proposed system, the removal of the irrelevant group of images from the database in every feedback iteration cannot allow for the improvement of the retrieval performance after some iteration. Hence, the results become stationary after a few feedback iterations.

Table 2. The average precision on each feedback iteration for the Corel image database

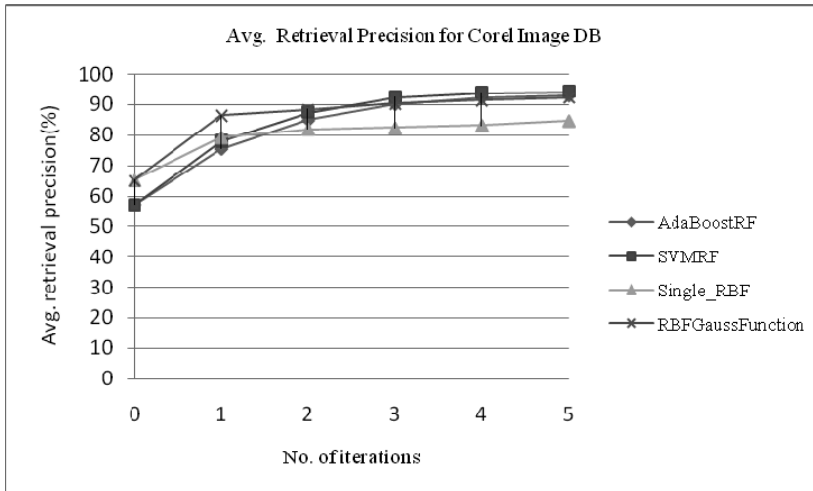| Approach	CBIR	1st iteration	2nd iteration	3rd iteration	4th iteration	5th iteration
RBFGaussFunction[25]	65.2	86.5	88.4	90.4	91.5	92.3
Single_RBF[25]	65.2	79.2	81.9	82.3	83.1	84.6
AdaBoostRF[26]	57.2	75.4	84.8	90.0	92.2	92.8
SVMRF	57.2	**78.0**	**87.5**	**92.3**	**94.0**	**94.2**

Fig. 5. Average accuracy versus iteration curves for Corel images

5.4 Image Retrieval Examples

Retrieval examples with the proposed method and earlier methods with and without the relevance feedback are shown in Figs. 6(a)-6(d). Fig. 6(a) is the result of CBIR without RF using combined features (RCWT+DT-CWT). It was observed that among the top 20 retrieved images, 8 images belong to the desired category (i.e., Images 1-6 and Images 16 and 20) and that the remaining 12 belong to the irrelevant category. Hence, there is a 50.0% retrieval precision of CBIR without RF. From Fig. 6(b) we can observe that there is a rapid increase in performance (i.e., from 50.0% to 100%) using SVMRF. Figs. 6(c)-6(d) shows the performance improvement of the approach using AdaBoost for a texture database. From Figs. 6(c) to 6(d), we can observe

that retrieval accuracy increased from 81.25% to 93.75% from the first iteration to the second iteration of relevance feedback and it remains the same in further iterations. It shows that AdaBoostRF is a bit slower to achieve better results than the SVMRF is.
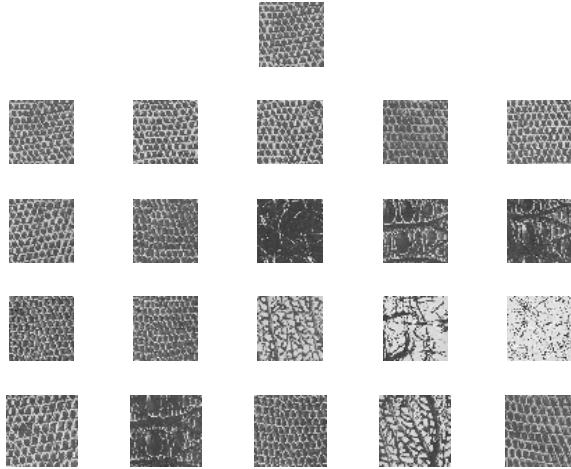

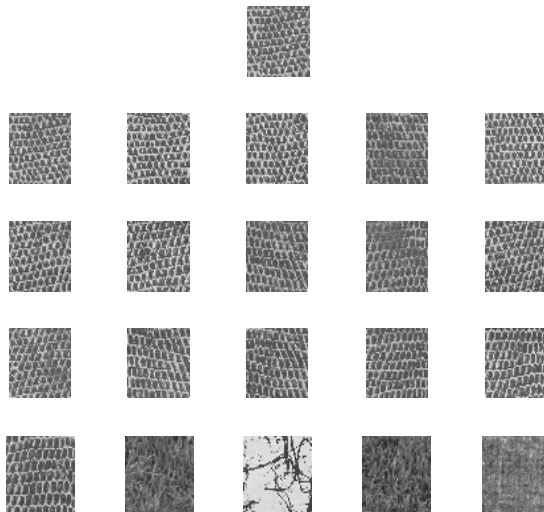Fig. 6(a). The result of CBIR using combined features (RCWF +DT_CWT) (8/16)

Fig. 6(b). The result after the first feedback iteration using SVMRF (16/16)
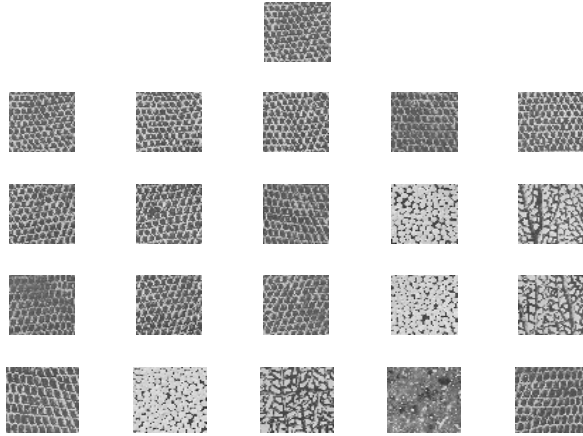
Fig. 6(c). The result after the first feedback iteration using AdaBoostRF (13/16)
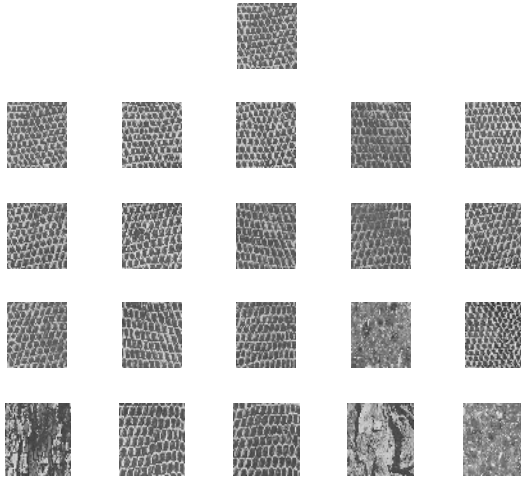

Fig. 6(d). The result after the second feedback iteration using AdaBoostRF (15/16)

6. CONCLUSION

In this paper, a novel relevance feedback framework has proposed, which can employ any machine learning algorithm that is applicable to RF. In this paper, we tested the proposed RF framework using SVM and AdaBoost, since SVM and AdaBoost both work well for small training data. The experimental results indicate that with the proposed method retrieval, accuracy is increased from 78.5% to 91.70 % and 92.29% when using AdaBoostRF and SVMRF, respectively on the texture database in only 5 iterations of relevance feedback. The retrieval precision for the Corel image database is increased from 57.20 to 92.3% and 94.2% in 5 iterations of RF by using AdaBoostRF and SVMRF, respectively. In the future, to improve the overall performance of the system, one can extend the proposed RF framework, which considers a linear combination of the most relevant images obtained from the both SVMRF and AdaBoostRF. This

is so that it can increase the training data size, which in turn will increase the retrieval performance.

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, No. 2 article 5,2008, 5:1-5:60.

[2] Y. Rui, T. S. Hung, and S. F.Chang, "*Image retrieval: Current Techniques, Promising Directions and Open Issues,*" J. Visual Comm. and Image Representation 10, (1999), pp. 39-62.

[3] A. W. M Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, "*Content –based Image Retrieval at the End of the Early Years,*" IEEE trans. Pattern Anal. Machine Intell. 22(12), 2000, pp.1349-1380.

[4] M. B. Kokare, P.K. Biswas, and B.N. Chatterji., *"A Survey on Current Content-based Image Retrieval Methods,"* IETE J. Res. 48(3&4), 2002, pp. 261-271.

[5] Y. Rui, T. S. Huang, and S. Mehrotra,. *"Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval,"* IEEE Transactions on Circuits and Systems for Video Technology, Vol. 8(5), 1998, pp. 644-655.

[6] Y. Rui, T. S. Huang, and S. Mehrotra, *"Content-based Image Retrieval with Relevance Feedback in MARS,"* in Proc. IEEE Int. Conf. on Image proc., 1997.

[7] S. D. MacArthur, C. E. Brodley, and C. R. Shyu, *"Relevance Feedback Decision Trees in Content-based Image Retrieval,"* in Proc. IEEE Work-shop Content-based Access of Image and Video Libraries, 2000, pp.68-72.

[8] I. J. Cox, M.L. Miller, T.P. Minka, T.V.Papathomas, and P.N.Yianilos, *"The Bayesian Image Retieval System, PicHunter: Theory, Implementation and Psychophysical Experiments,"* IEEE Tran. on Image Processing, vol. 9,Issue 1 , (2000), pp.20-37.

[9] Z. Su H. Zhang, S. Li, and S. Ma, *"Relevance Feedback in Content-based Image Retrieval: Bayesian framework, Feature Subspaces, and Progressive Learning,"* IEEE Trans. Image Process. vol. 12, no. 8, 2003, pp. 924-936.

[10] S. Tong and E. Chang, "*Support Vector Machines Active Learning for Image Retrieval,*" Proc. ACM Multimedia, 2001.

[11] K. Tieu and P. Viola, *"Boosting image retrieval,"* in Proc. IEEE Conf. Computer Vision Pattern Recognition, 2003, pp. 228-235,

[12] Zhou, X. S. and Huang, T. S., "*Relevance Feedback in image retrieval: A Comprehensive review*," Multimedia systems, 8, 6, 2003, pp. 536-544.

[13] Zhi-Hua Zhou, Ke-Jia Chen, and Hong-Bin Dai, "*Enhanced Relevance Feedback in Image Retrieval Using Unlabeled Data*," ACM trans. on informations systems, vol.24, issue 2, 2006, pp. 219-244.

[14] Marin Ferecatu, Nozha Boujemaa, Michel Crucianu, *"Semantic interactive image retrieval combining visual and conceptual content description,"* ACM multimedia systems Journal, vol. 13, No. 5-6, 2008, pp. 309-322.

[15] C. H. Steven Hoi, R. L. Michael, and J. Rong ,*"A Unified Log-Based Relevance Feedback Scheme for Image Retrieval,"* IEEE trans. on Knowledge and Data Engineering, vol. 18, no. 4, 2006.

[16] S. Tong and E. Chang, "*Support vector machine active learning for image retrieval,*" In Proceedings of the 9th ACM international conference on Multimedia, 2001, pp. 107-118.

[17] S. Tong and D. Koller, *"Support vector machine active learning with applications to text classification",* 17th International Conference on Machine Learning, Morgan Kaufmann 2000, pp. 999–1006.

[18] M. B. Kokare, P.K. Biswas, and B.N. Chatterji, *"Texture Image retrieval using New Rotated Complex Wavelet Filters,"* IEEE Trans. on systems, man, and Cybernetics-Part B: Cybernetics, vol. 35, no.6, 2005 , pp.1168-1178.

[19] Chiou-Ting Husu and Chuech-Yu Li, *"Relevance Feedback Using Generalized Bayesian Framework with Region Based optimization Learning",* IEEE trans. on Image Processing, vol. 14, No.10, Oct. 2005.

[20] Anca Loredana Ion, Liana Stanescu, and Dan Burdescu, *"Semantic Based Image Retrieval using Relevance Feedback"*, International Conference on Computer as a Tool, 2009, pp. 303-310.

[21] N.G. Kingsbury, *"Image processing with complex wavelet"*, Phil. Trans. Roy. Soc. London A, vol. 357, 1999, pp. 2543-2560.

[22] N. G. Kingsbury, *"Complex wavelets for Shift Invariant Analysis and Filtering of Signals"*, J.App. Comput. Harmon. Anal., vol. 10, no.3, (2001), pp.234-253.

[23] I. Selesnick, R. Baraniuk, and N. Kingsbury, "*The dual-tree complex wavelet transform,*" IEEE Signal Process. Mag., vol.22, no. o6, 2005, pp.123-151.

[24] V. N. vapnik. Statistical Learning Theory. 1998.

[25] Rongtao Ding, Xinhao Ji, and Linting Zhu, "*Research on the relevance feedback-based image retrieval in Digital library*", PWASET, vol. 25, ISSN 1307-6884, 2007, pp. 48-52.

[26] Pushpa B. Patil and M. B. Kokare, "*Semantic Image Retrieval using Relevance Feedback*", International Journal of Web & Semantic Technology (IJWesT), Vol.2, No.4, 2011, pp. 139-148.

Pushpa B. Patil
received B. E., and M. Tech. from the Department of Computer Science and Engineering, Karnataka University Darawad, Visveshwaraya Technological University Belgaum, India in the years 1996, 2006 respectively. From 1997-2000, she was worked as lecture in Computer Science Department at MBE's Engineering College, Ambajogai, Maharastra, India. In 2000, she joined as a lecturer in the Department of Computer Science at B. L. D. E' s. Institute of Engineering and Technology, Bijapur, Karnataka, India, where she is presently holding position of Associate Professor and doing PhD at S.R.T.M. University, Nanded, Maharastra, India. Her research interests include image processing, pattern recognition, and relevance feedback in Content Based Image Retrieval. She is a life member of Indian Society for Technical Education and Institute of Engineers.

Dr. Manesh Kokare
received the Diploma in Industrial Electronics Engineering from Board of Technical Examination, Maharashtra, India, in 1990, and B.E. and M. E. Degree in Electronics Engineering from Shri Guru Gobind Singhji Institute of Engineering and Technology (SGGSIE&T) Nanded, Maharashtra, India, in 1993 and 1999 respectively, and Ph.D. from the Department of ECE, Indian Institute of Technology, Kharagpur, India, in 2005. Since June1993 to Oct1995, he worked with Industry. From Oct 1995, he started his carrier in academics in the Department of Electronics and Telecommunication Engineering at SGGSIE&T, Nanded, where he is presently holding position of Associate Professor. He has published more than 50 papers in international and national journals and conferences. He received the prestigious Career Award for Young Teachers (CAYT) for the year 2005 from All India Council for Technical Education (AICTE), New Delhi, India. In December 2009, he honored with "Best Achiever of SGGSIE&T Alumni". Recently Dr. Kokare has been awarded "BOYSCAST" Fellowship for the year 2010-2011 by the Department of Science and Technology, Government of India to carry out his Post Doctoral research work at University of California Santa Barbara, USA. He is a life member of System Society of India, Indian Society for Technical Education, Institution of Electronics and Telecommunication Engineers, and Member of IEEE.