Semantic Learning in Interactive Image Retrieval

Pushpa B. Patil[1] and Manesh Kokare[2]

[1] Dept.of Computer Science and Engg.
BLDEA's, V.P.Dr. P. G. Halakatti CET. Bijapur, India
[2] Dept. of Electronic and Telecommunication
S.G.G.S, Institute of Engineering and Technology, Nanded, India
pushpa_ms@rediffmail.com,
mbkokare@sggs.ac.in

Abstract. This paper presents content-based image retrieval frameworks with relevance feedback based on AdaBoost learning method. As we know relevance feedback (RF) is online process, so we have optimized the learning process by considering the most positive image selection on each feedback iteration. To learn the system we have used AdaBoost. The main significances of our system are to address the small training sample and to reduce retrieval time. Experiments are conducted on 1856 texture images to demonstrate the effectiveness of the proposed framework. These experiments employed large image databases and combined RCWFs and DT-CWT texture descriptors to represent content of the images.

Keywords: Content-Based Image Retrieval (CBIR), Relevance Feedback (RF), Rotated Complex Wavelet Filters (RCWFs), Dual Tree Complex Wavelet (DT-CWT), and Image retrieval.

1 Introduction

Recently, there is a huge collection of digital images are available. Since there is a development of the internet and availability of digital devices such as scanners, digital cameras etc. With these huge collections, it is very essential to have efficient and effective tools for searching, browsing and retrieval of images. To do these tasks CBIR was introduced in the early 1980. CBIR uses the low level features like color, texture and shape to retrieve the most similar images stored in the database. With these low level features, user's perception on the images can not be achieved. Since different users perception is different on same images. It is the big disadvantage of the CBIR. To overcome this, relevance feedback (RF) was introduced into CBIR in 1998[5]. There is good review on CBIR [1-4].

RF is an online process, which tries to learn the user perception interactively; initially RF was designed for text-based information retrieval systems. Later it was introduced into CBIR during mid 1990's, with the involvement of the user in the retrieval loop to reduce the "semantic gap" between query representation (low level features) and user perception (high level concepts). RF has been proved to provide effective and efficient retrieval performance improvement in CBIR systems through interactive learning based on the user feedback [5-6].

D. Nagamalai, E. Renault, and M. Dhanushkodi (Eds.): DPPR 2011, CCIS 205, pp. 118–127, 2011.

1.1 Related Work

Recently, many researchers began to consider the RF as a classification or semantic learning problem. That is 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 good review on RF in [12]. The process of learning is very difficult task in RF [12]-[14], due to the following reasons

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

2) Training data is asymmetrical, which creates a too much imbalance between the relevant and irrelevant images.

3) In RF, for every iteration we have to perform both training and testing online, which takes more real time.

For visual representation of the images, we employed the global texture features presented in [18], which provides very efficient performance. Much of the work on RF uses the low-level representation using discrete wavelet transform (DWT) [15], Gobor filters [16] and co-occurrence matrix [19][20] for textures. In order to retrieve the general purpose images like artificial objects and natural scenes most of the time textural features are combined with color and shape to get better retrieval performance. However, they still suffers from the poor directional sensitivity, shift variant and, redundancy. From these combined features we may get better retrieval performance but not efficient one because as we increase number of features it increases the dimensionality of feature space. With such 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 our system efficient, we have to consider two factors namely time complexity and space complexity together with better retrieval performance. To overcome above problem, we propose to use the new rotated complex wavelet filters which gives both better and efficient retrieval performance.

1.2 Our Approach

In this paper we have used our earlier recent work [18] to extract more compact effective low level features, 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. To reduce the significant gap between low level feature and high level concepts, we have proposed a new RF approach and it is tested using AdaBoost. We found that proposed RF framework provides efficient retrieval performance in very few feedback iterations. A new relevance feedback approach, which is based on a ADABoost uses the relevant and irrelevant examples. Our extensive experiments using proposed RF with AdaBoost on standard texture database show significant improvements with respect to retrieval performance.

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. In section 3, we have, explained the concept of AdaBoost. In section 4, experimental results are discussed. Finally, the conclusion is given in section 5.

2 Image Descriptors

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 suffer from problem like no perfect reconstruction is possible using CWT decomposition beyond level 1, when input to each level becomes complex. To overcome this, Kingsbury [21] proposed a new transform, which provides perfect reconstruction along with providing the other advantages of complex wavelet, which is DT-CWT. The DT-CWT uses a dual tree of real part of wavelet transform instead using complex coefficients. 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 using two filter banks in parallel, operating on the same data. For d-dimensional input, a L scale DT-CWT outputs an array of real scaling coefficients corresponding to the lowpass 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. See [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 response of six wavelets associated with 2-D complex wavelet transform are illustrated in Fig. 1.

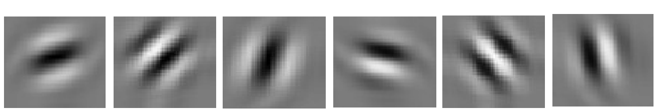

Fig. 1. Impulse response of six wavelet filters of complex wavelet.

2.1 DT-RCWF

Recently we have designed 2D- rotated complex wavelet transform [18]. Directional 2D RCWF are obtained by rotating the directional 2D DT-CWT filters by 45^0 so that decomposition is performed along new direction, which are 45^0apart from 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 input image with 2-D RCWF followed by 2-D downsampling operation is performed up to the desired level. The computational complexity associated with RCWF decomposition is the same as that of standard 2-D DT-CWT, if both are implemented in the frequency domain. The set of RCWFs retains the orthogonality property. The six subbands of 2D DT-RCWF gives information 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 provide us with more directional selectivity in the direction $\left\{\begin{matrix} \left(+15^{\circ},\ +45^{\circ},\ +75^{\circ},\ -15^{\circ},\ -45^{\circ}, -75^{\circ}\right), \\ \left(0^{\circ}, +30^{\circ},\ +60^{\circ},\ +90^{\circ},\ 120^{\circ},\ -30^{\circ}\right) \end{matrix}\right\}$ than the DWT whose directional sensitivity is in only four directions $\left\{0^{\circ}, \pm 45^{\circ}, 90^{\circ}\right\}$. The impulse response of six wavelets associated with rotated complex wavelet transform is illustrated in Fig. 2.

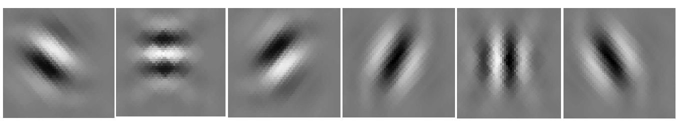

Fig. 2. Impulse response of six rotated complex wavelet filter

3 AdaBoost

AdaBoost was introduced in 1995 by Freund and Schapire [24] as an efficient algorithm of the ensemble learning field. It is used to boost the classification performances of a weak learner. It does this by combining a collection of weak classification functions to form a stronger classifier. AdaBoost combines iteratively the weak classifiers by taking into account a weight distribution on the training samples such that more weight is attributed to samples misclassified by the previous iterations.

Consider a two classification problem, in which the training data comprises input vectors $X_1, \ldots, X_N$ along with corresponding binary target variables $t_1, \ldots, t_n$, where $t_n \in \{-1,1\}$. each data point is given an associated weighting parameter w_n, which is initially set $1/N$ for all data points. We shall suppose that we have a procedure available for training a base classifier using weighted data to give a function $y(x) \in \{-1,1\}$. At each stage of the algorithm, AdaBoost trains a new classifier using a data set in which the weighting coefficients are adjusted according to the misclassified data points. Finally, when the desired number of base classifiers has been trained, they are combined to form a committee using coefficients that give different weight to different base classifiers.

3.1 Image Retrieval

3.1.1 Feature Database Creation

To conduct the experiments, each image from database is decomposed using DT-CWT and DT-RCWF up to third level and two different sets of features were computed as follows.

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. 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 with combination of these two feature parameters always outperformed that using these features individually [18]. The Energy (E_k) and (σ_k) Standard Deviation of k^{th} subband is computed as follows

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

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

Where $W_k(i,j)$ is the k^{th} wavelet-decomposed subband, $M \times N$ is the size of wavelet decomposed subband, and μ_k is the mean of the k^{th} subband. The resulting feature vector using energy and standard deviation are $\bar{f}_E = [E_1 \quad E_2 \quad \ldots \quad E_n]$ and $\bar{f}_\sigma = [\sigma_1 \quad \sigma_2 \quad \ldots \quad \sigma_n]$ respectively. So combined feature vector is

$$\bar{f}_{\sigma E} = [\sigma_1 \quad \sigma_2 \quad \ldots \quad \sigma_n \quad E_1 \quad E_2 \quad \ldots \quad E_n] \tag{7}$$

3.2 Image Matching

We have randomly selected any one of the 1856 images as a query image from texture images. Query image is further processed to compute the feature vector as given in section 3.1. 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

$$\text{Canb}(x, y) = \sum_{i=1}^{d} \frac{|x_i - y_i|}{|x_i| + |y_i|} \tag{8}$$

4 Experiments

To evaluate the performance of a proposed system, we have used the Brodatz texture photographic album. The experiments were conducted using MATLAB 7.0 with Intel core2Duo, 1 GB RAM machine.

4.1 Image Database

The texture database used in our experiment consists of 116 different textures [18]. We used 108 textures from Brodatz texture photographic album, seven textures from USC database and one artificial texture. Size of each texture image is 512×512. Each 512×512 image is divided into sixteen 128×128 non overlapping sub images, thus creating a database of 1856 texture images.

4.2 Performance Measures

For experimental results, it is significant to define a suitable metric for performance evaluation. We have used Average accuracy and it is defined as the percentage of relevant images of retrieved images among all relevant images in the database.

Experimental results are evaluated on 116 queries randomly selected from the texture database. The reported results of average accuracy are obtained by taking an average over the 116 queries texture database.

For each experiment, one image was selected at random as the query image from each category and thus the retrieved images were obtained. Then, the users were asked to identify those images that are related to their expectations from the retrieved images. These selected images were used as feedback images for next iteration. Finally, we have computed the average accuracy of all the categories in the database. Each image category contains 16 images. The feedback processes were performed 4 times.

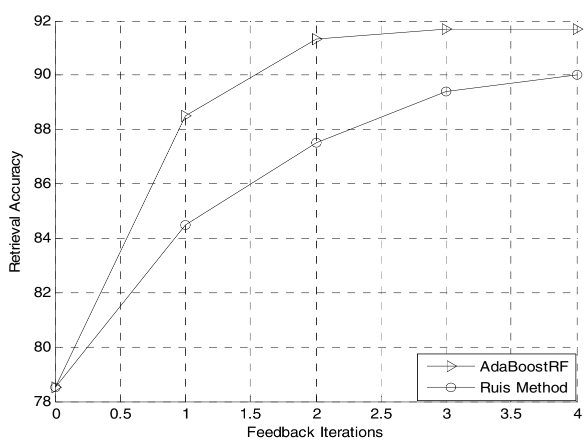

Fig. 3. Average accuracy versus iteration curves for texture images

Fig.3 describes detailed comparison of the average retrieval accuracy obtained using Rui[5] and AdaBoost on every feedback iteration of the randomly selected image from each category of texture database. The main observation of proposed RF using AdaBoost gives better retrieval performance comparing with Rui. method. From the Fig. 3, we observed that, there is a rapid increase in retrieval performance with each feedback iteration of proposed RF using AdaBoost learning algorithm. Retrieval performance is improved from 78.5% to 91.70% from first iteration to the fourth iteration using AdaBoostRF

We illustrated these observations using graph in Fig. 3 and results are also tabulated in table 1.

Table 1. Average accuracy on each feedback iterations for Texture database

Approach	CBIR	1st iteration	2nd iteration	3rd iteration	4th iteration
AdaBoostRF	78.5	88.52	91.32	91.70	91.70
Rui method	78.5	84.50	87.52	89.42	90.01

4.3 Image Retrieval Examples

We use an example to illustrate the performance improvement of the proposed approach in Fig. 4(a)-4(c) for texture database. Fig.4(a) is the result of CBIR using combined features (RCWT+DT-CWT), in which among top 20 images, 8 images belongs to the desired category (i.e images 1-6 and image 16, 20) and remaining 12 belongs to irrelevant category. So we got 50.0% retrieval precision from CBIR. Fig.4 (b)-(c) shows performance improvement of the proposed approach using the AdaBoost for texture database. From Fig. 4(b) to 4(c), we can observe that retrieval accuracy increasing from 81.25% to 93.75% from first iteration to second iteration of relevance feedback and it remains same in further iterations

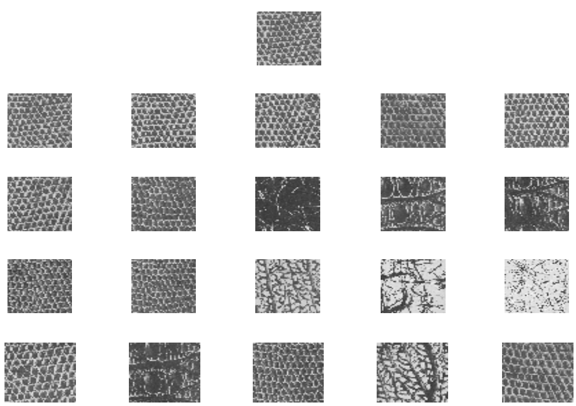

Fig. 4. (a) Result of CBIR using Combined features (RCWF +DT_CWT) (8/16)

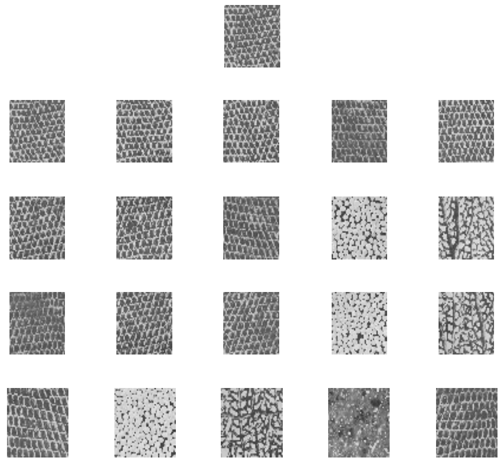

Fig. 4. (b) Result after first feedback iteration using AdaBoostRF(13/16)

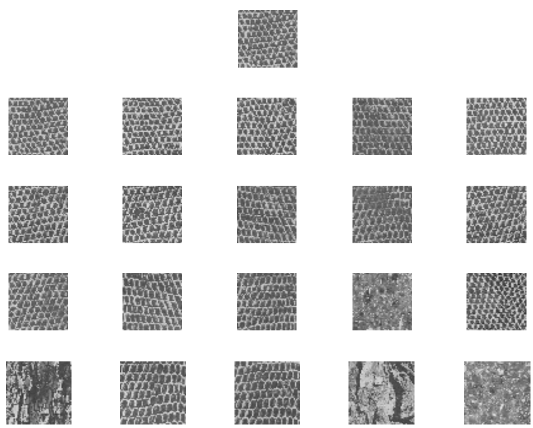

Fig. 4. (c) Result after second feedback iteration using AdaBoostRF (15/16)

5 Conclusion

In this paper, an active relevance feedback framework has been proposed to handle the small training data problem in RF and optimizing the testing set in order to reduce the retrieval time. The proposed relevance feedback framework is tested using AdaBoost with texture features. The RF framework is tested on large scale standard texture image database. The framework has demonstrated very promising retrieval accuracy. From experimental results, we found that RF using AdaBoost with combined texture features RCWF and DT-CWT gives better retrieval performance.

References

[1] Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image Retrieval: Ideas, Influences, and Trends of the New Age. ACM Computing Surveys 40(2), article 5, 5:1–5:60 (2008)
[2] Rui, Y., Hung, T.S., Chang, S.F.: Image retrieval: Current Techniques, Promising Directions and Open Issues. J. Visual Comm. and Image Representation 10, 39–62 (1999)
[3] Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R.: Content –based Image Retrieval at the End of the Early Years. IEEE Trans. Pattern Anal. Machine Intell. 22(12), 1349–1380 (2000)
[4] Kokare, M., Chatterji, B.N., Biswas, P.K.: A Survey on Current Content-based Image Retrieval Methods. IETE J. Res. 48(3&4), 261–271 (2002)
[5] Rui, Y., Huang, T., Ortega, M., Mehrotra, S.: Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval. IEEE Transactions on Circuits and Systems for Video Technology 8(5), 644–655 (1998)
[6] Rui, Y., Huang, T.S., Mehrotra, S.: Content-based Image Retrieval with Relevance Feedback in MARS. In: Proc. IEEE Int. Conf. on Image Proc. (1997)
[7] MacArthur, S.D., Brodley, C.E., Shyu, C.R.: Relevance Feedback Decision Trees in Content-based Image Retrieval. In: Proc. IEEE Workshop Content-based Access of Image and Video Libraries, pp. 68–72 (2000)
[8] Cox, I.J., Miller, M.L., Minka, T.P., Papathomas, T.V., Yianilos, P.N.: The Bayesian Image Retieval System, PicHunter: Theory, Implementation and Psychophysical Experiments. IEEE Tran. on Image Processing 9(1), 20–37 (2000)
[9] Su, Z., Zhang, H., Li, S., Ma, S.: Relevance Feedback in Content-based Image Retrieval: Bayesian framework, Feature Subspaces, and Progressive Learning. IEEE Trans. Image Process. 12(8), 924–936 (2003)
[10] Tong, S., Chang, E.: Support Vector Machines Active Learning for Image Retrieval. In: Proc. ACM Multimedia (2001)
[11] Tieu, K., Viola, P.: Boosting image retrieval. In: Proc. IEEE Conf. Computer Vision Pattern Recognition, pp. 228–235 (2003)
[12] Zhou, X.S., Huang, T.S.: Relevance Feedback in image retrieval: A Comprehensive review. Multimedia Systems 8(6), 536–544 (2003)
[13] Zhou, Z.-H., Chen, K.-J., Dai, H.-B.: Enhanced Relevance Feedback in Image Retrieval Using Unlabeled Data. ACM Trans. on Informations Systems 24(2), 219–244 (2006)
[14] Ferecatu, M., Boujemaa, N., Crucianu, M.: Semantic interactive image retrieval combining visual and conceptual content description. ACM Multimedia Systems Journal 13(5-6), 309–322 (2008)
[15] Hoi, S.C.H., Lyu, M.R., Jin, R.: A Unified Log-Based Relevance Feedback Scheme for Image Retrieval. IEEE Trans. on Knowledge and Data Engineering 18(4) (2006)
[16] Tong, S., Chang, E.: Support vector machine active learning for image retrieval. In: Proceedings of the 9th ACM International Conference on Multimedia, pp. 107–118. ACM Press, New York (2001)
[17] Tong, S., Koller, D.: Support vector machine active learning with applications to text classification. In: Proceedings of ICML-2000, 17th International Conference on Machine Learning, pp. 999–1006. Morgan Kaufmann, San Francisco (2000)
[18] Kokare, M., Biswas, P.K., Chatterji, B.N.: Texture Image retrieval using New Rotated Complex Wavelet Filters. IEEE Trans. on Systems, Man, and Cybernetics-Part B: Cybernetics 35(6) (2005)
[19] Husu, C.-T., Li, C.-Y.: Relevance Feedback Using Generalized Bayesian Framework with Region Based optimization Learning. IEEE Trans. on Image Processing 14(10) (2005)

[20] Ion, A.L., Stanescu, L., Burdescu, D.: Semantic Based Image Retrieval using Relevance Feedback. In: International Conference on Computer as a Tool, Warsaw, pp. 303–310 (2009)
[21] Kingsbury, N.G.: Image processing with complex wavelet. Phil. Trans. Roy. Soc. London A 357, 2543–2560 (1999)
[22] Kingsbury, N.G.: Complex wavelets for shift invariant analysis and filtering of signals. J. App. Comput. Harmon. Anal. 10(3), 234–253 (2001)
[23] Selesnick, I., Baraniuk, R., Kingsbury, N.: The dual-tree complex wavelet transform. IEEE Signal Process. Mag. 22(6), 123–151 (2005)
[24] Freund, Y., Schapire, R.E.: A decision –theoretic generalization of online learning and an application to boosting. J. Comput. Syst. Sci. 55(1), 119–139 (1997)