Composite Relevance Feedback for Image Retrieval

Pushpa B. Patil, Manesh B. Kokare

Abstract—This paper presents content-based image retrieval (CBIR) framework with relevance feedback (RF) based on combined learning of support vector machines (SVM) and AdaBoosts. The framework incorporates only the most relevant images obtained from both the learning algorithms. To speed up the system, it removes irrelevant images from the database, which are returned from SVM learner. It is the key to achieve the effective retrieval performance in terms of time and accuracy. The experimental results show that this framework had significant improvement in retrieval effectiveness, which can finally improve the retrieval performance.

Keywords—Image retrieval, relevance feedback, wavelet transform.

I. INTRODUCTION

RECENTLY there is a rapid growth in digital image technology, due to this, huge collection of images and its management has been done in several applications, such as digital libraries, medicine, and biodiversity information systems. With these large size collections and development of internet, there is need of efficient and effective mechanisms to retrieve images. The process of image searching, browsing and retrieval tools are the need of the users from various area, including remote sensing, fashion, crime prevention, publishing medicine etc. For this purpose, it is essential to have efficient image retrieval systems. This is the goal of the so called Content-Based Image retrieval. There are different ways to retrieve the images in CBIR. References [1]-[4] had presented 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, there is a big challenge in CBIR to reduce the 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 period and it will remains an active research area, due to ambiguities in interpretation of images than words which makes the necessity of user interaction and judgment of an image is faster than the document [7].

Recently, most of the work in RF is based on SVMs [18]-[20] because they minimize a measure of error on the training set while simultaneously maximizing the margin between relevant and irrelevant images. A SVM is highly effective mechanism for avoiding the 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.

A. 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 [9], Bayesian learning [10], [11], support vector machines [12], boosting [13], genetic algorithms and genetic programming’s are used to optimize the query pattern [7], [14], [15] and so on. There is good review on RF in [7], [16]. The process of learning is very difficult task in RF [17], [18], 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 too much imbalance between relevant and irrelevant images.
3. In RF, for every iteration we have to perform both training and testing online, which takes much real time.

Recently, of the work in RF is based on SVMs [18]-[20] because they minimize a measure of error on the training set while simultaneously maximizing the margin between relevant and irrelevant images. A SVM is highly effective mechanism for avoiding the 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.

B. Our Approach

To improve the retrieval performance, in this paper, we have designed a new relevance feedback framework using support vector machine and AdaBoost supervised learning techniques. The proposed framework uses the linear combination of the high ranking images obtained from both learning and removing the low ranking images from the database. Hence firstly it reduces testing process of each iteration. Since relevance feedback is online process, hence in every iteration of relevance feedback, testing and training has to perform online. Secondly, it increases the training data, so more the training yields more the accuracy. We used image descriptors, combination of dual tree complex wavelets (DT-CWT) and rotated complex wavelet filters (RCWF) [21]. The efficiency of proposed method is compared with other relevance feedback technique [8], [26], [27].

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 II and composite RF frame work in Section III. In Section IV, experimental results are discussed. Finally, the conclusion is given in Section V.

II. IMAGE DESCRIPTORS

In this paper we have used our earlier recent work [21] to extract more compact and 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, which gives details in twelve different directions. For image matching, we used Canberra distance measure.

A. Dual Tree Complex Wavelet Transforms

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 [22] Kingsbury proposed complex wavelet transform, which provides perfect reconstruction along with providing the other advantages of complex wavelet. The DT-CWT gives orientations details 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 scale L 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. See [22], [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

$$\psi(x) = \psi_h(x) + j\psi_g(x)$$

where $\psi_h(x)$ and $\psi_g(x)$ are both real-valued wavelets.

B. Dual Tree Rotated Complex Wavelet Filters

Recently we have designed 2D- rotated complex wavelet transform [22]. Directional 2D RCWF are obtained by rotating the directional 2D DT-CWT filters by $45^\circ$ so that decomposition is performed along new direction, which are apart from decomposition $45^\circ$ directions of CWT. The size of a filter is $(2N - 1) \times (2N - 1)$, where $N$ is the length of the 1-D filter. The decomposition of input image with 2-D RCWF followed by 2-D down sampling operation is performed up to the desired level. The computational complexity associated with RCWF decomposition is the same as that of standard 2-D CWT, if both are implemented in the 2-D frequency domain. The set of RCWF s retains the orthogonality property. The six sub bands of 2D DT-RCWF gives information strongly oriented at $(30^\circ,90^\circ, -30^\circ, 60^\circ, 90^\circ, 120^\circ)$. The mechanism of the DT-RCWF is not covered here. See [21] for a comprehensive explanation of the transform and details of filter design for the trees. Thus, the 2D DT-CWT and RCWF provide us with more directional selectivity in the directions

$$(+15^\circ, +45^\circ, +75^\circ, -15^\circ, -45^\circ, -75^\circ),$$
$$(0^\circ, +30^\circ, +60^\circ, +90^\circ, 120^\circ, -30^\circ)$$

than the DWT whose directional sensitivity is in only three directions $(0^\circ, \pm 45^\circ, 90^\circ)$.

III. COMPOSITE RELEVANCE FEEDBACK

In this paper we present the composite relevance feedback, which uses the two learning methods SVM and AdaBoost. The linear combination of the high ranking images obtained from these two learning methods gives better results.

A. AdaBoost

In AdaBoost, the base classifiers are trained in sequence, and each base classifier is trained using weighted form of the data set in which the weighting coefficient associated with each data point depends on the performance of the previous classifiers. In particular, points that are misclassified by one of the base classifiers are given greater weight when used to train the next classifier in the sequence. Once all the classifiers have been trained, their predictions are then combined through a weighted majority voting scheme.

Consider a two-class classification problem, in which the training data comprises input vectors $x_1, \ldots, x_N$ along with corresponding binary target variables $t_1, t_2, t_3, \ldots, t_N$ 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 performance of the previously trained classifier so as to give greater weight 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 [25].

B. Support Vector Machines

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

$$y(X) = W^T \psi(X) + b$$

(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 $X_1, X_2, \ldots, X_N$, with corresponding target values $t_1, t_2, \ldots, t_N$ and new data points are classified according to the sign of $y(X)$.

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

$$
\min_{w,b,\xi} \frac{1}{2}w^TW + C \sum_{i=1}^{l} \xi_i
$$

Subject to

$$
t_i(W^T\phi(X) + b) \geq 1 - \xi_i
$$

where training vectors $X_i$ are mapped into a higher dimensional space by the function $\phi$ SVM finds a linear separating hyper plane with maximal margin in this higher dimensional space. Furthermore, $k(X,X') = \phi(X)^T\phi(X')$ is called the kernel function.

C. Proposed Method

The proposed framework uses linear combination of SVM and AdaBoost returned high ranking images. First advantage of SVM over other learning algorithms lies in its high generalization performance without the need to add a priori knowledge. Second advantage is it can work for small training sets [16]. The effective use of labeled samples helps to learn a query concept faster and more accurately. Thus we have used first SVM to learn the query concept to get the more accurate results. Next we used AdaBoost to learn the query concept. Though the AdaBoost is slower, it gives better results. Moreover, to speed up the retrieval performance, we considered high ranking images returned from learners. Due to removal of low ranking images obtained from SVM, few of the relevant images are filtered. To avoid this we used AdaBoost. Thus linearly combined results give better retrieval performance. The composite RF framework is compared with individual framework. The algorithm 1 depicts the proposed composite RF framework.

Let DB be the collection of images in the database and q be the query image. Let P and N are the relevant and irrelevant images respectively marked from the feedback of the user.

**Algorithm 1: Composite RF**

1. **Input:** $q$: user query
   - DB: Image database
   - P: Relevant images
   - N: Irrelevant images
2. **Output:** Result
3. **Begin**
4. Result= CBIR (DB, q);
5. **Repeat until user satisfaction or result remains same**
   - (P, N)=Labeling (Result);
6. **End**

At the beginning of the retrieval process, user provides the query image $q$. With this query image CBIR shows the initial results to the user (line 1). Hence, with this initial results, the user labels the relevant (i.e. P) and irrelevant (i.e. N) images (line 6). Each iteration involves the following steps: user labels relevant and irrelevant images; the learning of user perception by using SVM (line 8) and AdaBoost (line 9). Here, we have considered high ranking images from both learning methods. Thus, combining the high ranking images obtained from both learning approaches to get better retrieval performance (line 10). Finally, Canberra distance (eq. (5)) is used for similarity measure (line 11-13) in order rank the database images by sorting distance vector (line 14). Effectiveness of the proposed method is evaluated based on Top 20 images (line 15). To speed up the retrieval time, we have removed the irrelevant images obtained from SVM Learner from the database (line 16). This procedure is repeated until user satisfaction or a result remains same.

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

IV. Experiments

In our experiments, with known number of images in each category and number of category in the database, we have simulated our RF framework to obtain the user feedback automatically. In this simulation, all images belonging to the same category of the query image are considered relevant. Experiments were performed six iterations for each query. The retrieval performance of the proposed method is evaluated by considering top 20 images for each iteration. To evaluate the performance of a proposed system, we have used the Brodatz texture photographic album [21] and Corel Image Database [27]. The experiments were conducted using MATLAB 7.0 with Intel core2Duo, 1 GB RAM machine.

A. Texture Image Database

The texture database used in our experiment consists of 116 different textures [21]. 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 x 512. Each 512 x 512 image is divided into sixteen
B. Corel Image Database

It contains 1000 color photographs of resolution 384x256 pixels, covering a wide range of semantic categories, from natural scenes to artificial objects [26]. The database is partitioned into ten categories, each with 100 photographs.

TABLE I

<table>
<thead>
<tr>
<th>Approach</th>
<th>CBIR</th>
<th>1st iteration</th>
<th>2nd iteration</th>
<th>3rd iteration</th>
<th>4th iteration</th>
<th>5th iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVMRF</td>
<td>78.50</td>
<td>89.27</td>
<td>91.75</td>
<td>92.18</td>
<td>92.29</td>
<td>92.29</td>
</tr>
<tr>
<td>AdaBoostRF</td>
<td>78.50</td>
<td>88.52</td>
<td>91.32</td>
<td>91.70</td>
<td>91.70</td>
<td>91.70</td>
</tr>
<tr>
<td>CompositeRF</td>
<td>78.50</td>
<td>88.84</td>
<td>92.99</td>
<td>94.18</td>
<td>94.61</td>
<td>94.66</td>
</tr>
</tbody>
</table>

Fig. 1 Average accuracy versus iteration curves for texture images

C. Performance Measures

For a retrieval task, it is significant to define a suitable metric for performance evaluation. So we have used average accuracy and precision. 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 6 times.

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

Fig. 1 describes detailed comparison of the average retrieval performance obtained using SVMRF, AdaBoostRF and CompositeRF on every feedback iteration of the randomly selected image from each category of texture image database. The proposed framework is compared to SVMRF [8] and AdaBoostRF [27]. We can observe from the Table I and Fig. 1, that the proposed composite RF framework is superior to SVMRF and AdaBoostRF. However SVMRF is slightly superior to an AdaBoostRF. Similarly, Fig. 2 describes the detailed comparison average retrieval performance obtained using SVMRF, AdaBoostRF and CompositeRF on every feedback iteration for Corel Images. We can observe from the Table II, the proposed method yields better retrieval performance than the Single_RBF and RBFGaussFunction proposed by Rongtao et al. [26].

TABLE II

<table>
<thead>
<tr>
<th>Approach</th>
<th>CBIR</th>
<th>1st iteration</th>
<th>2nd iteration</th>
<th>3rd iteration</th>
<th>4th iteration</th>
<th>5th iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>CompositeRF</td>
<td>57.2</td>
<td>76.1</td>
<td>86.0</td>
<td>91.6</td>
<td>95.5</td>
<td>96.4</td>
</tr>
<tr>
<td>Single_RBF</td>
<td>65.2</td>
<td>86.5</td>
<td>88.4</td>
<td>90.4</td>
<td>91.5</td>
<td>92.3</td>
</tr>
<tr>
<td>RBFGaussFunction</td>
<td>65.2</td>
<td>79.2</td>
<td>81.9</td>
<td>82.3</td>
<td>83.1</td>
<td>84.6</td>
</tr>
<tr>
<td>SVMRF</td>
<td>57.2</td>
<td>78</td>
<td>86.9</td>
<td>92.2</td>
<td>94</td>
<td>94.6</td>
</tr>
<tr>
<td>AdaBoostRF</td>
<td>57.2</td>
<td>75.4</td>
<td>91.32</td>
<td>91.70</td>
<td>91.70</td>
<td>92.5</td>
</tr>
</tbody>
</table>
Fig. 2 Average precision versus iteration curves for Corel images

D. Image Retrieval Examples

Figs. 3 (a)-(g) demonstrate the some snapshots of the retrieval results during relevance feedback of images and use an example to illustrate the performance improvement of the proposed approach in Fig. 3 for texture database. Fig. 3 (a) is the result of CBIR using combined features (RCWT+DT-CWT), in which among top 20 images, 7 images belong to the desired category and remaining 13 belongs to irrelevant category. So we got 43.75% retrieval accuracy from CBIR. From Figs. 3 (b), (c), we can observe that, there is rapid increase in performance using SVMRF [8]. Fig. 3 (d), (e) shows performance improvement of the approach [27] using AdaBoost for texture database. From Figs. 3 (f), (g), we can observe that retrieval accuracy increasing from 75% to 81.25% from first iteration to second iteration of the proposed composite relevance feedback. Here we have shown first to second iteration results of all approaches and their results remain same in further iterations.
In this paper, we presented a new relevance feedback based content-based image retrieval framework. It uses composite learning method using SVM and AdaBoost to learn the user perspective. High ranking images obtained from these learning algorithms are linearly combined and low ranking images, which are returned from SVM are removed from the database to improve retrieval performance. It is tested on large scale texture image database and natural image database; the framework has demonstrated very promising retrieval accuracy.

V. CONCLUSION

REFERENCES


Pushpa B. Patil completed her B. E., and M.Tech. from the Department of Computer Science and Engineering, Karnataka University Darawad, Visvishwaraya 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, Maharashtra, 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, Maharashatra, India, in 1993 and 1999 respectively, and Ph.D. from the Department of ECE, Indian Institute of Technology, Kharagpur, India, in 2005. Since June 1993 to Oct 1995, 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.