Content Based Image Retrieval with Relevance Feedback using Riemannian Manifolds

Pushpa B.Patil
Dept. of Computer Science and Engg.
BLDEA's Dr. P. G. H. College of Engg. and Tech.
Bijapur, India
Email: pushpa_ms@rediffmail.com

Manesh B. Kokare
Dept. of Electronic and Tele. Commucation
S. G. G. S. Institute of Engg. and Tech.
Nanded, India
Email:mbkokare@yahoo.com

***Abstract*—In this paper we propose a novel approach for content-based image retrieval with relevance feedback, which is based on Riemannian Manifold learning algorithm. This method uses positive and negative (relevant/irrelevant) images labeled by the user at every feedback iteration. In this paper, we pre-computed the cost adjacency matrix and its eigenvectors corresponding to the smallest eigen values for effectiveness and efficiency of the retrieval system. Then we apply the Riemannian Manifolds learning concept to estimate the boundary between positive and negative images. Experimental results of the proposed method have been compared with earlier approaches, which show the superiority of the proposed method.**

I. Introduction

Due to tremendous growth of the digital electronic, there is an availability of huge collection of image and video archive for many applications (art galleries, picture and photograph archives, medical and geographical databases etc), which demands most effective and efficient advanced image retrieval/browsing to address the perceptual aspects of visual information. As a consequence of this, content based image retrieval (CBIR) system was introduced. Comprehensive and recent extensive literature survey on content based image retrieval is presented in [1]-[4]. The CBIR retrieves the images based on low level content like color, shape and texture etc. Due to semantic gap between the low level content and high level concepts, the CBIR system cannot retrieve the images based on user perception. To reduce this semantic gap, relevance feedback was introduced, which was basically designed for text document retrieval and later Rui et al.[5]-[6] introduced for content based image retrieval to reduce the semantic gap in CBIR. Comprehensive surveys in RF in CBIR can be found in [7]-[8].

A. Related Work

In 2004, He et al. [9] proposed a novel transductive learning framework named manifold-ranking based image retrieval (MRBIR). It makes use of a manifold ranking algorithm to explore the relationship among all the features in the feature space, and then ranks relevance between the query and all the images in the database using graph representation, which is different from traditional similarity metrics based on pair-wise distance. In relevance feedback, MRBIR uses both positive and negative images for feedback iteration. In 2005, the same author He et al.[10] proposed CBIR with Multiple Random Walk(MRW) to create two generative models using Markov random walks one for relevant and other for irrelevant to the query image. Then it refines the two random walks through expectation maximization (EM) like iterative procedure in order to get more accurate estimation of the likelihood functions. In 2008, Sabbi et al.[11] introduced a new Graph Laplacian which makes it possible to robustly learn the embedding, of the manifold enclosing the dataset, via a diffusion map. In 2012, Patil and Kokare[12] developed a graph based CBIR model for texture image retrieval by employing greedy Prim's algorithm. Comparing with the previous approaches, the proposed approach is better. Since EM like process needs a number of pre-estimation of iterative steps and they are dependent on the parameters. Hence we propose a novel approach to CBIR with relevance feedback, which is inspired by the Riemannian Manifolds (RM) learning algorithm for classification introduced by Niyogi and Belkin in [13].

The rest of the paper is organized as follows. In section II, discussed about image descriptors, In section III, we discussed proposed method using Riemannian Manifolds. In section IV, experimental results are discussed. Finally, the conclusion is given in section V.

II. IMAGE DESCRIPTORS

For image feature extraction we have used combined dual tree rotated complex wavelet filters (DT-RCWF) and dual tree complex wavelets (DT-CWT), which gives information in twelve directions. However results of CBIR and CBIR with relevance feedback will vary depending on visual features [14] and similarity distance [15]. The DT-CWT and DT-RCWF are briefly explained in the following section.

A. Dual-Tree Complex Wavelet Transforms

Real real desreate wavelet(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 [16] 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

978-0-7695-5100-5/13 $31.00 © 2013 IEEE
DOI 10.1109/ICSIP.2014.9

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 2^d. The total redundancy of the transform is and independent of$L$. The mechanism of the DT-CWT is not covered here. See [17] for a comprehensive explanation of the transform and details of filter design for the trees. A complex valued $\psi(x)$ 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.

B. *Dual-Tree Rotated Complex Wavelet Filters*

Recently Kokare et al. have designed 2D-rotated complex wavelet transform [17]. Directional 2D RCWF are obtained by rotating the directional $2D$ DT-CWT filters by $45°$ so that decomposition is performed along new direction, which are $45°$ apart 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 $(30°, 0°, -30°, 60°, 90°, 120°)$. The mechanism of the DT-RCWF is explained in our earlier work[17]. The 2D DT-CWT and RCWF provide us with more directional selectivity in the direction $\left\{ \begin{array}{c} (+15°, +45°, +75°, -15°, -45°, -75°) \\ (0°, +30°, +60°, +90°, 120°, -30°) \end{array} \right\}$ than the DWT whose directional sensitivity is in only four directions.

III. Riemannian Manifolads for CBIR with Relevance Feedback

We model the CBIR framework as graph based method. Let us define graph $G = (V, E)$, where V is the set of vertices (nodes) and E is the set of edges. In this framework vertices represents the images and edges represents the similarity distance d_{ij} between the image i and j. Then the graph is represented using cost adjacency matrix W. To compute the cost adjacency matrix, we used Canberra distance measure, hence we have NxN cost adjacency matrix for size N image database.

Given N images $X_1, X_2 \ldots X_N \in \mathbb{R}^l$, let us assume that s $(s < N)$ images are labeled by user with labels l_i, where $l_i \in \{-1, 1\}$ and the rest are unlabeled. Our goal is to label the unlabeled images. To do this, we computed the cost adjacency matrix W for the adjacency graph G using Canberra distance eq.5and then eigenvectors are estimated for it.

A. *Eigenvector Computation*

The basic theory behind the eigenvector system is to reduce the size of the images to be recognized from a high to a lower dimension. While lowering the dimensionality in the set, using the eigenvector method also highlights the variance within the set. Since eigenvectors corresponding to the largest eigen values represent the directions in $\mathbb{R}^l$ of the greatest variation among a set of images having that covariance. Therefore, the coordinates of an image along these eigenvector directions provide a useful set of parameters, or a feature vector characterizing the image.Hence, we compute p eigenvectors corresponding to the smallest eigen values for the eigenvector problem as in eq.2.

$$Lz = \lambda z \tag{2}$$

Matrix $L = W - D$ is the graph Laplacian for the adjacency graph. Here D is diagonal matrix of the same size as W, with row sums of W as entries, $D_{ii} = \sum_j W_{ji}$. Laplacian is a symmetric, positive semidefinite matrix which can be thought of as an operator on functions defined on vertices of the graph. The eigen functions can be interpreted as a generalization of the low frequency Fourier harmonics on the manifold defined by the data points[13].

$$\begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1N} \\ z_{21} & z_{22} & \cdots & z_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ z_{p1} & z_{p2} & \cdots & z_{pN} \end{bmatrix} \tag{3}$$

The classifier is built by computing the space of coefficients $a = (a_1, \ldots, a_p)^T$ using the eq.4 where p is the number of eigen functions we wish to employ

$$a = ({Z_{labeled}}^T Z_{labeled})^{(-1)} {Z_{labeled}}^T l \tag{4}$$

where $l = (l_1, \ldots, l_s)$ and $\begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1N} \\ z_{21} & z_{22} & \cdots & z_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ z_{p1} & z_{p2} & \cdots & z_{pN} \end{bmatrix}$ is the matrix of values of eigen functions on the labeled images. Finally, we classified the unlabeled images, if $X_i, i > s$ is an unlabeled image using eq.5.

$$l_i = \begin{cases} 1 & if \sum_{j=1}^{p} z_{ij} a_j \geq 0 \\ \\ -1 & if \sum_{j=1}^{p} z_{ij} a_j < 0 \end{cases} \tag{5}$$

B. *Proposed Algorithm*

The pseudocode of the proposed approach using Riemannian Manifolds(RM) is described in the Algorithm 1.

Algorithm 1 *proposed relevance feedback*

```
Input: Query image q, distance matrix W
```

```
Output: Top  T  Retrieved Images
Begin
1  Z=Compute eigenvectors
    for distance matrix W (eq. 1)
2  R=CBIR (W, q)
3  Let vector l=(l1,,lk ) be the labels
   of the retrieved images R
4  Repeat  until  user not satisfied
     with R or R remains same do
5  Zlabeled=Compute eigenvectors
    for labeled images
6  a=Building classifier
     using eq. (4)
7  [rel, irel]=Labeling the
           unlabeled images eq. (5)
8  S=Sort_relevant_images(rel)
9  T=DisplayTop20(S);
10 l =Labeling (T)
11 End

End.
```

The algorithm takes the input cost adjacency matrix W, which is a representation of graph G and where q is the query image. We pre-computed the eigenvectors Z of the cost adjacency matrix W (line 2). Initial retrieval results R are obtained using the traditional content-based image retrieval (line 2). Retrieval results are labeled with labels $l = (l_1, \ldots, l_s)$ as relevant/irrelevant by the user, where $l_i \in (-1, 1)$. Then number of feedback iterations is repeated until the user satisfaction or result remains same (line 4-line 11). We compute the eigen functions Zlabeled for labeled images (line 5). Unlabeled images are labeled as relevant and irrelevant using eq.4 (line 7). To speed up the retrieval system, we employ only relevant images for each feedback iteration. Finally, we sort the relevant images and retrieved first 20 images (line 8-line 9).

IV. Experimental Results

The experiments were conducted using MATLAB 7.0 with Intel core2duo, 1 GB RAM machine. For evaluation purposes we used Wang Dataset [18], which is a subset of the known Corel dataset consisting of 1000 images grouped into 10 categories (100 images per category). The sample image of the each category shown in Figure 1.

Figure 1. Sample Image of each category of Wang dataset.

We used the rotated complex filters and dual tree complex wavelet transforms jointly for feature extraction, which gives twelve different orientation information of the image [17]. The Canberra distance metric is used to measure the distance between the database image and query image using eq.6, where x and y are the feature vectors of an image from the image database and query image respectively of dimension d.

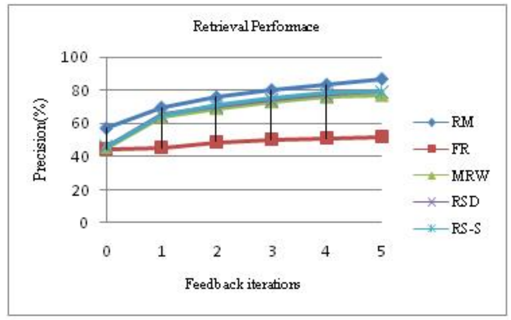

Figure 2. Retrieval Performance.

$$Canb(x, y) = \sum_{i=1}^{d} \frac{|x_i - y_i|}{|x_i| + |y_i|} \tag{6}$$

A linear normalization procedure has been performed, so that each feature takes values in the range between 0 and 1. From the database, each image is used as a query and the top twenty nearest neighbors are returned. Relevance feedback is performed by marking images belonging to the same class of the query as relevant, and all other images in the top twenty as non-relevant.

We compared proposed approach with following four earlier methods, feature re-weighting (FR) [19], which describes the significance of the features to represent the relevant images. Relevance Score (RS) [20], which computes the rank of each image by employing the similarity measure between the nearest positive (relevant) and nearest negative (irrelevant). Relevance Score Stabilized (RS-S) [21], which was the extended work of the relevance score (RS), in which Bayesian Query shift added to boost the retrieval performance and Multiple Random Walks (MRW) [10].

In Figure 2, we summarize the results obtained in terms of precision with scope size 20. It is evident from an observation that proposed method using Riemannian Manifolds (RM) outperforms the earlier approaches.

A. Retrieval Examples

Figure 3 shows results of CBIR with initial query image, in which among top-20 images, eleven images (red frame) belongs to the desired category and remaining nine belongs to irrelevant category. So we got 55% of retrieval precision from CBIR. Then based on the information fed-back by the user, the proposed system dynamically learns and much more relevant images are retrieved. The performance improvement of proposed approach is shown in Figure 4–5 and then we can observe that retrieval precision is increased from 55% to 80% in second iteration of relevance feedback.

Figure 3. Retrieval results without RF(55%)

Figure 4. Retrieval results after first feedback iteration(60%).

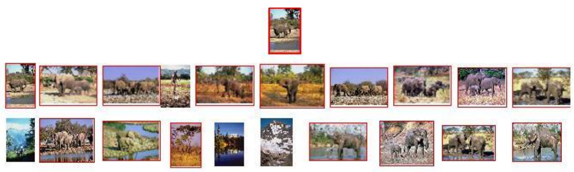

Figure 5. Retrieval results after second feedback iteration(80%).

V. Conclusion

This paper presents a novel CBIR with RF framework based on graphical method using Riemannian Manifold learning. This learning algorithm uses pre-computed eigen vectors. The images are ranked as relevant/non-relevant based on the user perception. Then each unlabeled image is finally ranked according to the classifier built using eigen functions. To speed up the retrieval system, we used only high ranking images for each feedback iteration.

References

[1] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang.,'Image Retrieval: Ideas, Influences, and Trends of the New Age,' *ACM Computing Surveys*, Vol. 40, No. 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*, pp. 39-62, Jan. 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), pp.1349-1380, 2000.

[4] Manesh B. Kokare, P.K. Biswas, and B.N.Chatterji, 'A Survey on Current Content-Based Image Retrieval Methods,' *IETE J. Res.* 48(3and4), pp. 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*, Vol. 8(5), pp. 644-655, 1998.

[6] Rui, Y., Huang, T.S., and Mehrotra,S. 'Content-based Image Retrieval with Relevance Feedback in MARS,' *in Proc. IEEE Int. Conf. on Image proc.*, pp. 815-818, 1997.

[7] Zhou, X. S. and Huang, T. S., 'Relevance Feedback in Image Retrieval: A Comprehensive Review,' *Multimedia systems*, 8, 6, pp. 536-544, 2003.

[8] Liu, Y., Zhang, D., Lu, G., and Ma, W.-Y.,'A Survey of Content-Based Image Retrieval with High-level Semantics,' *Pattern recognition*, 40(1), pp.262-282, 2011

[9] J. He, H. Tong, M. Li, W. Y. Ma, C. Zhang,'Manifold Ranking Based Image Retrieval,' *international conference of Multimedia*, pp. 9-16, 2004.

[10] J. He, H. Tong, M. Li, W. Y. Ma, C. Zhang, 'Multiple Random Walk and its Application in Content-Based Image Retrieval,' *Int. Workshop on Mult. Inf. Retrieval*, pp.151-158, 2005.

[11] H. Sabbi, P. Etyngier, J. Y. Audibert, R. Keriven, 'Manifold Learning using Robust Graph Laplacian for Interactive Image Search,'*IEEE Conf. Computer Vision and Patt. Recogn.(CVPR)*, pp.1-8, 2008.

[12] Pushpa B. Patil and Manesh B. Kokare, 'Texture Image Retrieval using Greedy Method,' *Proceedings of ICADC-2012*, pp. 885-891, 2012, Bangalore India.(Springer)

[13] Belkin and Niyogi, 'Semi-Supervised Learning on Manifolds,' Machine Learning, vol. 56, pp. 209-239, 2004.

[14] Pushpa B. Patil and Manesh B. Kokare, 'Interactive content-based texture image retrieval,'*in Porc.of 2nd Int. Conf. on computer and Communication Technology(ICCCT-11)*, pp. 71-76, MNNIT, Allahabad, 2011.

[15] Manesh B. Kokare, B. N. Chatterji and P. K Biswas, 'Comparison of similarity metrics for texture image retrieval,'*in IEEE TENCON*, pp. 571-575, 2003.

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

[17] G. W. J. Z. Wang, J. Li, 'SIMPLIcity: Semantics-sensitive Integrated Matching for Picture Libraries,' *IEEE Trans. Pattern Anal. Machine Intell.*, 23 (9), pp. 947-963, 2001.

[18] Manesh 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, pp. 1168-1178, 2005.

[19] G. Das, S. Ray, C. Wilson, 'Feature Re-weighting in Content-based Image Retrieval,' *Int. Conf. in Image and Video Retrieval*, pp. 193-200, 2006.

[20] G. Giacinto, F. Roli, 'Instance-Based Relevance Feedback for Image Retrieval,' *Adv. in Neural Inform. Process. Syst. (NIPS)*, vol. 17, pp. 489-496, 2005.

[21] G. Giacinto, 'A Nearest-Neighbor Approach to Relevance Feedback in Content Based Image Retrieval,' *in Int. Conf. in Image and Video Retrieval*, pp. 456-463, 2007.