

Interactive Content Based Image Retrieval Using Ripplet Transform and Fuzzy Relevance Feedback

Manish Chowdhury, Sudeb Das, and Malay Kumar Kundu

Machine Intelligence Unit, Indian Statistical Institute, Kolkata 700 108, India
{st.manishc,to.sudeb}@gmail.com, malay@isical.ac.in

Abstract. In this article, a novel content based image retrieval (CBIR) system based on a new Multiscale Geometric Analysis (MGA)-tool, called Ripplet Transform Type-I (RT) is presented. To improve the retrieval result, a fuzzy relevance feedback mechanism (F-RFM) is also implemented. Fuzzy entropy based feature evaluation mechanism is used for automatic computation of revised feature's importance and similarity distance at the end of each iteration. Experimental results on a large image database demonstrate the efficiency and effectiveness of the proposed CBIR system in the image retrieval paradigm

Keywords: Ripplet Transform, Relevance Feedback, Content Based Image Retrieval, Fuzzy logic.

1 Introduction

The tremendous growth in quantity of images stored in digital form in the internet, have raised many demanding issues regarding effective digital data management. The task of retrieving relevant images from a large image database (DB), by measuring similarities between the query image and the database images, based on automatically derived features like color, texture, shape etc. has become a potential area of research. The performance of a CBIR system, strongly depends both on the availability of suitable features, for proper representation of the semantic aspects of the images automatically, and also the effectiveness of the used similarity measure [1]. High retrieval efficiency and less computational complexity, are the desired characteristics of an efficient CBIR system.

Wavelet transform (WT) based low level features, provide a unique representation of an image, and are highly suitable for characterizing textures of the image [2]. Many WT based CBIR systems have been proposed in the literature [3,2]. But WT is inherently non-supportive to directionality and anisotropy. As a result, CBIR system based on WT provide limited performance. To overcome these limitations of WT, recently a theory called Multi-scale Geometric Analysis (MGA) for high-dimensional signals has been developed, and several MGA tools were proposed like Ridgelet, Curvelet, Bandlet and Contourlet etc.[4,5,6].

Relevance feedback mechanism (RFM) has been used as an effective tool to provide significant performance boost in CBIR systems through continuous

learning and interaction with the end-users. Most of the RFM, employ two approaches namely, query vector moving technique and feature re-weighting technique to improve the retrieval results [7]. Feature re-weighting technique utilizes both the relevant and the irrelevant result's information, to obtain more effective result [2,8]. But in all these cases time complexity per iteration are high and the accuracy of the relevant images are low.

In this paper, we have used the type-I Ripplet Transforms (RT) coefficients to represent the images at different scales and different directions. RT coefficients of intensity and chromaticity planes are used as the primitive features for computation of mean and standard deviation. To improve the performance of the retrieval system, a simple fuzzy based relevance feedback mechanism (F-RFM) with Euclidian similarity measure is used.

The rest of the paper is organized as follows. RT is described in section 2. Section 3, presents the proposed CBIR system. Experimental results and comparisons are given in section 4, and the conclusion is drawn in section 5.

2 Ripplet Transform Type-I (RT)

Conventional transforms like (Fourier Transform) FT and Wavelet transform (WT) suffer from discontinuities such as edges and contours in images. To address this problem, Jun Xu et al. proposed a new MGA-tool called RT. RT is a higher dimensional generalization of the Curvelet Transform (CVT), capable of representing images or $2D$ signals at different scales and different directions. To achieve anisotropic directionality, CVT uses a parabolic scaling law. From the perspective of microlocal analysis, the anisotropic property of CVT guarantees resolving $2D$ singularities along C^2 curves. On the other hand, RT provides a new tight frame with sparse representation for images with discontinuities along C^d curves [9].

There are two questions regarding the scaling law used in CVT: 1) Is the parabolic scaling law optimal for all types of boundaries? and if not, 2) What scaling law will be optimal? To address these questions, Jun Xu et al. intended to generalize the scaling law, which resulted in RT. RT generalizes CVT by adding two parameters, i.e., support c and degree d . CVT is just a special case of RT with $c = 1$ and $d = 2$. The anisotropy capability of representing singularities along arbitrarily shaped curves of RT, is due to these new parameters c and d .

2.1 Discrete Ripplet Transform (DRT)

As digital image processing needs discrete transform instead of continuous transform, here we describe the discretization of RT [9]. The discretization of Continuous RT (CRT) is based on the discretization of the parameters of ripplet functions. The scale parameter a is sampled at dyadic intervals. b (position parameter) and θ (rotation parameter) are sampled at equal-spaced intervals. a_j , \vec{b}_k and θ_l substitute a , \vec{b} and θ respectively, and satisfy that $a_j = 2^{-j}$, $\vec{b}_k = [c \cdot 2^{-j} \cdot k_1, 2^{-j/d} \cdot k_2]^T$ and $\theta_l = \frac{2\pi}{c} \cdot 2^{-\lfloor j(1-1/d) \rfloor} \cdot l$, where $\vec{k} = [k_1, k_2]^T$,

and $j, k_1, k_2, l \in \mathbb{Z}$. $(\cdot)^T$ denotes the transpose of a vector. $d \in \mathbb{R}$, since any real number can be approximated by rational numbers, so we can represent d with $d = n/m, n, m \neq 0 \in \mathbb{Z}$. Usually, we prefer $n, m \in \mathbf{N}$ and n, m are both primes. In the frequency domain, the corresponding frequency response of ripplet function is in the form

$$\hat{\rho}_j(r, \omega) = \frac{1}{\sqrt{c}} a^{\frac{m+n}{2n}} W(2^{-j} \cdot r) V\left(\frac{1}{c} \cdot 2^{-\lfloor j \frac{m-n}{n} \rfloor} \cdot \omega - l\right) \tag{1}$$

where W and V are the ‘radial window’ and ‘angular window’, respectively and satisfy the following admissibility conditions:

$$\sum_{j=0}^{+\infty} |W(2^{-j} \cdot r)|^2 = 1 \tag{2}$$

$$\sum_{l=-\infty}^{+\infty} |V\left(\frac{1}{c} \cdot 2^{-\lfloor j(1-1/d) \rfloor} \cdot \omega - l\right)|^2 = 1 \tag{3}$$

given c, d and j .

The ‘wedge’ corresponding to the ripplet function in the frequency domain is

$$H_{j,l}(r, \theta) = \{2^j \leq |r| \leq 2^{2j}, |\theta - \frac{\pi}{c} \cdot 2^{-\lfloor j(1-1/d) \rfloor} \cdot l| \leq \frac{\pi}{2} 2^{-j}\} \tag{4}$$

The DRT of an $M \times N$ image $f(n_1, n_2)$ will be in the form of

$$R_{j, \vec{k}, l} = \sum_{n_1=0}^{M-1} \sum_{n_2=0}^{N-1} f(n_1, n_2) \overline{\rho_{j, \vec{k}, l}(n_1, n_2)} \tag{5}$$

where $R_{j, \vec{k}, l}$ are the ripplet coefficients and $\overline{(\cdot)}$ denotes the conjugate operator.

The image can be reconstructed through Inverse Discrete Ripplet Transform (IDRT)

$$\tilde{f}(n_1, n_2) = \sum_j \sum_{\vec{k}} \sum_l R_{j, \vec{k}, l} \rho_{j, \vec{k}, l}(n_1, n_2) \tag{6}$$

3 Proposed CBIR System

The proposed CBIR system consists of two main parts. First part is the feature extraction using RT and second part is the F-RFM. In the following subsections each of the aforementioned parts are explained thoroughly.

3.1 Feature Extraction Using Ripplet Transform Type-I

RT is a new MGA-tool, which satisfies the multiresolution, good localization, high directionality, general scaling and support, anisotropy and fast coefficient decay properties. This makes RT as an effective tool to represent images or 2D signals, and motivated us to use it, to represent the images in the DB.

The images in the DB prior to RT based decomposition, are transformed to YCbCr color space. This ensures that the textural characterization of the image are independent of the color characterization. RT decomposition over the intensity plane (Y) characterizes the texture information, while the RT decomposition

over chromaticity planes (Cb & Cr) characterize color. Texture and color information are extracted by using RT on each color plane with a 4 level (1, 2, 4, 4) decomposition. This decomposition configuration provides 11 (= 1 + 2 + 4 + 4) subbands for each image of the DB, for each color plane. As, there are 3 color planes, so altogether we get 33 (= 3 × 11) subbands for each image of the DB. For each subband S_j , its mean (f_{mean}^j) and standard deviation (f_{std}^j) are computed, and used as the representative features of the image (I):

$$f_{mean}^j(I) = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N |S_j^I(m, n)| \tag{7}$$

$$f_{std}^j(I) = \sqrt{\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (S_j^I(m, n) - f_{mean}^j(I))^2} \tag{8}$$

where, $S_j^I(m, n)$ represents coefficient at (m, n) of the subband S_j of size $M \times N$ of the image I and $j = 1, 2, \dots, 33$. The final feature vector of an image I in the DB is as follows:

$$f_{vec}(I) = [f_{mean}^1(I), f_{mean}^2(I), \dots, f_{mean}^{33}(I), f_{std}^1(I), f_{std}^2(I), \dots, f_{std}^{33}(I)]$$

3.2 Fuzzy Based Relevance Feedback

The retrieval result of the first pass of the proposed CBIR system is obtained by measuring the Euclidean distance between the RT based features of the query image (provided by the user) and the images stored in the database (DB). If the user are not satisfied with the result, then the user will select a set of relevant and irrelevant images from the top 20 ranked displayed images.

The features of the image I of the DB, is computed using RT and is represented by a set of features, $F = \{f_q\}_{q=1}^N$, where f_q is the q^{th} feature component in the N dimensional feature space. We have used Euclidean distance with weight w_q for measuring the similarity between the query image I_{qr} and other images I in the DB:

$$D(I, I_{qr}) = \sum_{q=1}^N w_q \|f_q(I) - f_q(qr)\| \tag{9}$$

where $\|f_q(I) - f_q(qr)\|$ is the Euclidean distance between the q^{th} feature component and w_q is the weight assigned to the q^{th} feature component. The weights should be adjusted such that, the features have small variation over the relevant images and large variation over the irrelevant images. Let k similar images $I_s = \{I_1, I_2, \dots, I_k\}$ where, $I_k \in I_s$, are returned to the user. Let I_r be the set of relevant images and I_{ir} be the set of irrelevant images as marked by the user. $\{I_r = I_j | I_j$ relevant, for $I_j \in I_s\}$ and $\{I_{ir} = I_j | I_j$ irrelevant, for $I_j \in I_s\}$. The information from relevant images (I_r) and irrelevant images (I_{ir}) are combined to compute the relative importance of the individual features, by using fuzzy feature evaluation index (FEI) of pattern classification problems [10].

The FEI is defined from interclass and intraclass ambiguities as follows : let $C_1, C_2, \dots, C_j \dots C_m$ be the m pattern classes in an N dimensional $(f_1, f_2, f_q, \dots, f_N)$ feature space where class C_j contains n_j number of samples. The features values along the q^{th} co-ordinate along class C_j are assigned as standard S-type membership function between 0 and 1 [11] . The entropy of a fuzzy set gives a measure of *intrasets ambiguity* along the q^{th} co-ordinate axis in C_j and is computed as follows:

$$H(A) = \left(\frac{1}{n_j \ln 2}\right) \sum_i S_n(\mu(f_{iqj})); i = 1, 2, \dots, n_j \quad (10)$$

where the Shannon's function is,

$$S_n(\mu(f_{iqj})) = -\mu(f_{iqj}) \ln(\mu(f_{iqj})) - \{1 - \mu(f_{iqj})\} \ln\{1 - \mu(f_{iqj})\} \quad (11)$$

Entropy (H) is dependent on the absolute values of membership (μ). $H_{min} = 0$ for $\mu=0$ or 1, $H_{max} = 1$ for $\mu=0.5$.

The criteria of a good feature is that, it should be nearly invariant within class, while emphasizing differences between patterns of different classes [10]. The value of H would therefore decrease, after combining the class C_j and C_k , as the goodness of the q^{th} feature in discriminating pattern classes C_j and C_k increases. The measure denoted as H_{qjk} is called "*intersets ambiguity*" along q^{th} dimension between classes C_j and C_k . Considering these two types of ambiguities, the FEI for the q^{th} feature is given:

$$FEI_q = \frac{H_{qjk}}{H_{qj} + H_{qk}}. \quad (12)$$

Lower value of FEI_q , indicates better quality of importance of the q^{th} feature in recognizing and discriminating different classes. The precision of retrieval can be improved with these values.

In the proposed CBIR system, the number of classes are two. The relevant/positive images constitute the 'intra-class' and the irrelevant/negative images constitute the 'inter-class' image features. To evaluate the importance of the q^{th} feature, the q^{th} component of the retrieved images is considered. i.e., $I^{(q)} = \{I_1^{(q)}, I_2^{(q)}, I_3^{(q)}, \dots, I_k^{(q)}\}$. H_{qj} is computed from $I_r^{(q)} = \{I_{r1}^{(q)}, I_{r2}^{(q)}, I_{r3}^{(q)}, \dots, I_{rk}^{(q)}\}$. Similarly H_{qk} is computed from the set of images, $I_{ir}^{(q)} = \{I_{ir1}^{(q)}, I_{ir2}^{(q)}, I_{ir3}^{(q)}, \dots, I_{irk}^{(q)}\}$. H_{qkj} is computed combining both the sets. Images are ranked according to Euclidean distance. The user marks the relevant and irrelevant set from 20 returned images, for automatic evaluation of (FEI). The weight w_q is a function of the evaluated (FEI_q) as shown as,

$$w_q = F_q(FEI_q) \quad (13)$$

In the first pass, all features are considered to be equally important. The feature space of the relevant images are therefore altered in a similar fashion after updating the components with w_q . As a result, the ranks of the relevant images are not affected much. For irrelevant images, one feature component may be very

close to the query, whereas other feature component may be far away from the query feature. But the magnitude of the similarity vector may be close to the relevant ones. Multiplying by w_q increases the feature separation between the irrelevant component, such that due to the combined effect the irrelevant image may be pulled down.

4 Experimental Results and Comparisons

To prove the effectiveness of the proposed system, extensive experiments were performed on 1000 images of SIMPLiCity database [3] with 10 different classes (African, Ocean, Building, Bus, Dinosaurs, Elephant, Flower, Horse , Mountain, Food) of images. The retrieval results obtained using the proposed CBIR system, are compared with some of the existing retrieval systems [12,13,14]. The experiments were carried out on a Dell Precision T7400 PC with 4GB RAM and was implemented using MATLAB R2008a.

Two commonly used statistical measures were computed to assess the proposed system’s performance, namely precision and recall, which are defined:

$$Precision (P) = \frac{N_{RIR}}{N_{RIR} + N_{IRIR}} \tag{14}$$

$$Recall (R) = \frac{N_{RIR}}{T_{RID}} \tag{15}$$

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

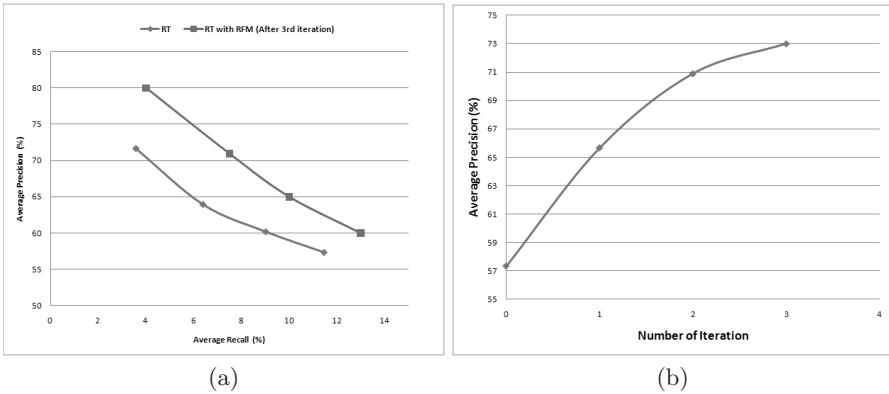


Fig. 1. (a) Average Precision vs. Average Recall (b) Average Precision vs. Number of Iteration

If the user is satisfied with the retrieved results without using RFM then the retrieval process is ended in the first pass. If, however, the user is not satisfied, he/she can select top query based relevant images as positive feedbacks and the remaining as negative examples for updating the weight parameters and revising the features. Using these feedbacks, FEIs and weights of the features are computed as described in Section.3.2. Then, all the images are re-sorted based on the weighted similarity metric. If the user is still not content with the result, he/she can repeat the process.

The graph of the Fig.1(a), shows the performance (in terms of average precision vs. average recall) of the proposed system. It clearly shows that, the performance of RT based features with F-RFM (considering only 3rd iteration's results) is better than the RT based features without F-RFM. The graph of the Fig.1(b), shows the improvement in performance with each iteration of the F-RFM.

As an example, using one of the image from the 'Flower' class of SIMPLiCity DB as the query image, the retrieval results are shown in Fig.2(a). There is one irrelevant image in the retrieval result at the first pass. But with every iteration of the F-RFM as shown in Fig.2(b), the ranking of the retrieved images got changed, and more and more similar images came up as the improved retrieval result.

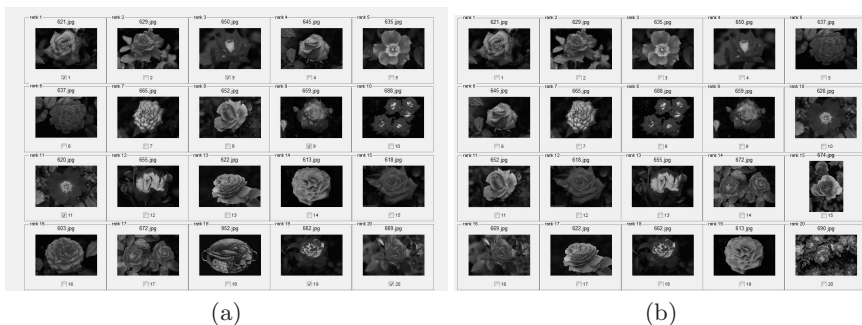


Fig. 2. Performance of the proposed CBIR system (top left side image is the query image) (a) First pass of the retrieval result (b) First Iteration

Table.1, shows the performance comparisons between our proposed CBIR system with some of the existing CBIR methods. From Table.1, it can be easily seen, that the proposed CBIR system provide improved retrieval performance over other existing CBIR algorithms namely, SIMPLiCity [12], FIRM [13], using salient points (salient points detected by Harris Corner Detector (SP by HCD), color salient points (CSP)) [14]. The 'Bold' values indicate the highest retrieval performance.

Table 1. Comparison with other existing CBIR systems in terms of Average Precision

Class	Wang[12]	FIRM[13]	SP by HCD[14]	CSP[14]	Proposed Method
Africans	0.48	0.47	0.40	0.48	0.49
Sea	0.32	0.35	0.31	0.34	0.40
Building	0.35	0.35	0.32	0.33	0.39
Buses	0.36	0.60	0.44	0.52	0.58
Dinosaur	0.95	0.95	0.92	0.95	0.96
Elephant	0.38	0.25	0.28	0.40	0.50
Flower	0.42	0.65	0.58	0.60	0.75
Horse	0.72	0.65	0.68	0.70	0.80
Mountain	0.35	0.30	0.32	0.36	0.40
Food	0.38	0.48	0.44	0.46	0.51
Average	0.47	0.51	0.47	0.51	0.55

5 Conclusions

From our experiments, we have noticed that ripple transform based image coding is suitable for representing low level features (color, texture, edge etc.) of the images. The proposed CBIR system based on RT features is able to improve the accuracy of the retrieval performance and to reduce the computational cost. The retrieval performance is improved further using fuzzy based RFM within 2 to 3 iterations. The proposed mechanism could be tested for video retrieval as future scope of research.

Acknowledgment. We would like to thank Jun Xu and Depeng Wu (Dept. of Electrical and Computer Engineering, University of Florida, USA) for helping us in the implementation of Ripple Transform.

References

1. Heesch, D.: A survey of browsing models for content based image retrieval. *Multimedia Tools Application* 40, 1380–7501 (2008)
2. Kundu, M.K., Chowdhury, M., Banerjee, M.: Interactive Image Retrieval with Wavelet Features. In: Kuznetsov, S.O., Mandal, D.P., Kundu, M.K., Pal, S.K. (eds.) *PRMI 2011*. LNCS, vol. 6744, pp. 162–172. Springer, Heidelberg (2011)
3. Wang, J.Z., Li, J., Wiederhold, G.: SIMPLicity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 947–963 (2001)
4. Do, M.N., Vetterli, M.: The finite ridgelet transform for image representation. *IEEE Transactions on Image Processing* 12, 16–28 (2003)
5. Candes, E., Donoho, D.: Continuous curvelet transform: I. resolution of the wavefront set. *Applied and Computational Harmonic Analysis* 19, 162–197 (2005)
6. Do, M.N., Vetterli, M.: The contourlet transform: An efficient directional multiresolution image representation. *IEEE Transactions on Image Processing* 14 (2005)

7. Rui, Y., Huang, T.S., Mehrotra, S.: Relevance feedback: a power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology* 8, 644–655 (1998)
8. Jin, Z., King, I., Li, X.: Content-Based Image Retrieval by Relevance Feedback. In: Laurini, R. (ed.) *VISUAL 2000*. LNCS, vol. 1929, pp. 521–529. Springer, Heidelberg (2000)
9. Xu, J., Yang, L., Wu, D.: Ripplet: A new transform for image processing. *Journal of Visual Communication and Image Representation* 21, 627–639 (2010)
10. Pal, S.K., Chakraborty, B.: Intraclass and interclass ambiguities (fuzziness) in feature evaluation. *Pattern Recognition Letters* 2, 275–279 (1984)
11. Pal, S.K., Majumder, D.D.: *Fuzzy Mathematical Approach To Pattern Recognition*. Willey Eastern Limited, New York (1985)
12. Li, J., Wang, J.Z., Wiederhold, G.: IRM: Integrated region matching for image retrieval. In: *Proceeding of the 8th ACM Int. Conf. on Multimedia*, pp. 147–156 (2000)
13. Chen, Y., Wang, Z.: A Region-based fuzzy feature matching approach to content-based image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 1252–1267 (2002)
14. Hiremath, P.S., Pujari, J.: Content based image retrieval using color boosted salient points and shape features of an image. *International Journal of Image Processing (IJIP)* 2, 10–17 (2008)