Content Based Image Retrieval using Ripplet Transform and Kullback-Leibler Distance

A. Ambika [1], Dr.J.Jennifer Ranjani[2], K.B. Srisathya[3], P. Deepika[4],
[1, 2]Dept. of Computer Science and Engineering, Vivekanandha College of Engineering for Women
[3] Dept. of Information Technology, Vivekanandha College of Engineering for Women
Tiruchengode – 637205, India
[4] Logistics and service integration Division, Caterpillar Logistics India Pvt. Ltd. Bangalore, India
email: j.jenniferranjani@yahoo.co.in, ambikaalaguvel14@gmail.com

Abstract - **Content-based image retrieval (CBIR), also known as query by image content is one of the applications of computer vision. In CBIR, the contents derived from the image like color, shapes, and textures are analyzed rather than the metadata such as keywords, tags, and/or descriptions associated with the image. In this paper, the texture features of the image are modelled using Generalized Gaussian Distribution (GGD) and Ripplet Transform. Ripplet transform has the capability of capturing structural information along with the curves compared to the traditional discrete wavelet transform. The similarity between query image and the training database is found using Kullback-Leibler Distance between GGDs. The proposed system provides greater accuracy and flexibility in capturing texture information. Experimental results on a large image database demonstrate the efficiency and effectiveness of the proposed CBIR system in the image retrieval paradigm.**

Keywords – Content Based Image Retrieval, Ripplet Trasnsform, Kullback – Leibler Distance.

I. INTRODUCTION

Content based image retrieval (CBIR) is the task of retrieving relevant images from a large image database, by measuring similarities between the query image and the database images. CBIR is mainly based on automatically derived features like colour, texture, and shape etc which has become a potential area of research. There is a growing interest in CBIR because of the limitations inherent in metadata-based systems, as well as the large range of possible uses for efficient image retrieval. Textual information about images can be easily searched using existing technology, but requires humans to personally describe every image in the database. It is impractical for very large databases, or for images that are generated automatically, e.g. from surveillance cameras.

Image retrieval procedures can be roughly divided into two approaches: annotation-based image retrieval (ABIR) and CBIR. In ABIR, images are often annotated by words, such as time, place, or photographer. To access the desired image data, the seeker can construct queries using homogeneous descriptions, such as keywords, to match these annotations. Although ABIR potentially offers the most accurate information when images are well-named or annotated, it has drawbacks such as manual image annotation is time-consuming and costly while human annotation is subjective. Some images could not be annotated because it is difficult to describe their content with words. Unlike ABIR, CBIR tends to index and retrieve images based on their visual content. CBIR avoids many problems associated with traditional ways of retrieving images by keywords.

The similarity measures are used to determine the similarity or dissimilarity between two objects. Different similarity measures may be used for different feature representations. For example, Euclidean is used for comparing vector-based representations while Histogram Intersection is used for comparing colour histogram representations The technique in [1] has attracted the CBIR and is still an active research topic nowadays because, in contrast to text/document retrieval, judging the relevance of an image for a user is an almost instantaneous task

A CBIR system is required to effectively harness information from these image repositories. It is characterized by the ability of the system to retrieve relevant images based on the visual and semantic contents of images. In [3], image attributes are grouped into two categories: extrinsic and intrinsic. Extrinsic attributes denote the characteristics of an image or a domain object in the image which can only be obtained externally. Extrinsic attributes cannot be derived from the image itself. Intrinsic attributes are those that can be extracted from the image manually, automatically, or by a combination of manual and automatic methods.

An image retrieval system based on dual tree complex wavelet transforms and support vector machines is proposed in [6]. The images that a user needs through query image are similar to a group of images with the same conception. A CBIR exploits implicit relevance feedback during a search session. In [7], PINVIEW contains several novel methods that infer the intent of the user. From relevance feedback, such as eye movements or clicks, and visual features of images PINVIEW, in learns a similarity metric between images which depend on the current interests of the user. It then retrieves images with a specialized reinforcement learning algorithm that balances the trade between exploring new images and exploiting the already inferred interests of the user.

Improved retrieval results are obtained over existing CBIR algorithm based on Stacked Euler Vector (SERVE) combined with Modified Moment Invariants. Visual content based image retrieval, a method which has proved effective over the contemporary text based image retrieval system has

978-1-4799-2352-6/14/$31.00 ©2014 IEEE

been proposed in [8]. CBIR systems based on shape using invariant image moments, viz., Moment Invariants and Zernike Moments are available in the reference [9]. Superiority of the CBIR system is observed over other moment based methods, viz., Moment Invariants and Zernike Moments in terms of retrieval efficiency and retrieval time. The classification efficiency is improved by employing Support Vector Machine (SVM) classifier.

In [19], Image feature representation method, called multi-texton histogram (MTH), for image retrieval is proposed. MTH integrates the advantages of co-occurrence matrix and histogram by representing the attribute of co-occurrence matrix using histogram. It can be considered as a generalized visual attribute descriptor but without any image segmentation or model training. The proposed MTH method is based on Jules's texton theory, and it works directly on natural images as a shape descriptor.

In this paper, the parameters of the Generalized Gaussian Distribution are determined from the Ripplet Transforms (RT) coefficients at different scales and different directions and these parameters are used as features. The similarity between the GGD parameters of the query image and the image database are determined using KL distance and top N images are retrieved. The rest of the paper is organized as follows. Section 2, presents the proposed CBIR system. Experimental results are given in section 3, and the conclusion is drawn in section 4.

II. PROPOSED CBIR SYSTEM USING RIPPLET TRANSFORM

In the proposed CBIR system, the input image is given as a query into the CBIR system. The features using ripplet transform are extracted for the query image as well as the images in the database. The mean and standard deviation of the Ripplet coefficients are utilized for feature extraction. Then Kullback-Leibler Distance (KLD) is utilized for finding the similarity between the multiple images database array values and the query features. The top 20 relevant images are been retrieved as a resultant image.

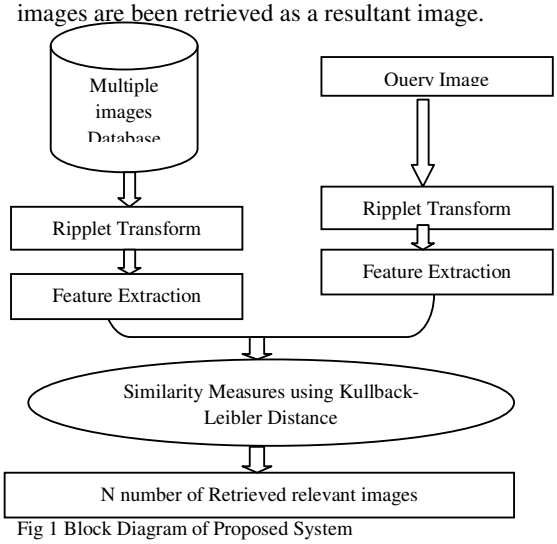


Fig 1 Block Diagram of Proposed System

A. *Ripplet Transform*

To overcome the limitations in the Wavelet Transform, Jun Xu et al. [13] proposed a new MGA-tool called RT. RT is a higher dimensional generalization of the Curvelet Transform (CVT), capable of representing images or 2*D* signals at different scales and different directions. To achieve anisotropic directionality, CVT uses a parabolic scaling law. From the perspective of micro local analysis, the anisotropic property of CVT guarantees resolving 2*D* 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. 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*.

Ripplet-II transform has the following properties, such as, Localization Directionality, and Flexibility. Ripplet-II representation of images with different degrees is presented. It can be observed that different degrees present different levels of sparsity. The magnitude of transform coefficients in a decreasing order for wavelet, ridgelet, ripplet-II and orthogonal ripplet-II transforms; the magnitude of coefficients of each transform is normalized by the largest coefficient of the corresponding transform. It can be observed that ripplet II has the fastest decay in coefficients, compared to wavelet and ridgelet. This is the reason why ripplet-II transform can provide sparser representation for images with edges than wavelet and ridgelet. Orthogonal ripplet-II demonstrates faster decay than ripplet-II, which indicates that orthogonal ripplet-II transform can provide sparser representation of functions than ripplet-II.

Besides the aforementioned properties, ripplet-II transform can provide rotation invariance. If an image $f_1(x; y)$ as well as its rotated version $f_2(x; y)$ rotated by an angle α, i.e.

$$f_2(x,y) = f_1(x\cos(\alpha) + y\sin(\alpha), -x\sin(\alpha) + y\cos(\alpha)) \quad (1)$$

In the polar coordinate system, therefore

$$f_2(\rho,\phi) = f_1(\rho,\phi-\alpha) \quad (2)$$

So ripplet-II transform of f_2 is

$$R_{f_2}(a,b,\mathrm{d},\theta) = \iint \psi_{a,b,\mathrm{d},\theta}\alpha(\rho,\phi) f_2(\rho,\phi)\rho\,\mathrm{d}\,\rho\,\mathrm{d}\,\phi \quad (3)$$

$$= \iint \psi_{a,b,\mathrm{d},\theta}\alpha(\rho,\phi) f_1(\rho,\phi-\alpha)\rho\,\mathrm{d}\,\rho\,\mathrm{d}\,\phi \quad (4)$$

$$= \iint \psi_{a,b,\mathrm{d},\theta} - \alpha(\rho,\phi) f_1(\rho,\phi)\rho\,\mathrm{d}\,\rho\,\mathrm{d}\,\phi \quad (5)$$

$$= R_{f_1}(a,b,\mathrm{d},\theta-\alpha) \quad (6)$$

Applying 1D Fourier transform on both sides of eq. (6) with respect to θ, we have

$$\mathrm{F}\{R_{f_2}(a,b,\mathrm{d},\theta)\} = \int R_{f_1}(a,b,\mathrm{d},\theta-\alpha)e^{-i\omega\theta}d\theta \quad (7)$$

$$= e^{-i\omega\alpha} F\{R_{f_1}(a,b,\mathrm{d},\theta)\} \quad (8)$$

Obviously, $\left|F\{R_{f_1}(a,b,\mathrm{d},\theta)\}\right| = \left|F\{R_{f_2}(a,b,\mathrm{d},\theta)\}\right|$ i.e., the magnitude of 1D Fourier transform w.r.t ripplet-II transform is rotation invariant. Hence, ripplet-II transform can provide rotation invariant features. Since there is no explicit direction parameter in orthogonal ripplet-II coefficient, orthogonal ripplet-II transform does not have the rotation invariant property.

B. Feature Extraction using GGD

A statistical view of the texture retrieval problem by combining the two related tasks, namely feature extraction (FE) and similarity measurement (SM), into a joint modelling and classification scheme.

Using a consistent estimator of texture model parameters for the FE step followed by computing the Kullback–Leibler distance (KLD) between estimated models for the SM step is asymptotically optimal in term of retrieval error probability. The statistical scheme leads to a new wavelet-based texture retrieval method that is based on the accurate modelling of the marginal distribution of wavelet coefficients using generalized Gaussian density (GGD) and on the existence a closed form for the KLD between GGDs. The proposed method provides greater accuracy and flexibility in capturing texture information, while its simplified form has a close resemblance with the existing methods which uses energy distribution in the frequency domain to identify textures. Digital image libraries are becoming more widely used as more visual information is put in digital form as well as on-line. To improve human access, however, there must be an effective and precise method for users to search, browse, and interact with these collections and to do so in a timely manner.

The marginal distribution of the ripplet image coefficients is generally non-Gaussian, symmetric, and sharply peaked around zero with heavy tails. Experiments depict that Generalized Gaussian density (GGD) is a symmetric function and it provides a good approximation to the non-Gaussian, symmetric and heavy tail characteristics of the ripplet coefficients. The GGD is a symmetric function and is characterized by two parameters, which can be easily estimated even when few samples are available.

The marginal distribution of the ripplet coefficients are modeled using a GGD which is characterized by two parameters, α and β is given by

$$p(x;\alpha,\beta) = \frac{\beta}{2\alpha\Gamma\left(1/\beta\right)} e^{-(|x|/\alpha)^{\beta}} \quad (9)$$

where the Gamma function $\Gamma(L) = \int_0^{\infty} t^{L-1} e^{-t}\, dt$, α is the standard deviation, sometimes called as the scale parameter. β is the shape parameter and it is inversely proportional to the decreasing rate of the peak. In this paper, the GGD parameters are estimated from the directional sub-bands at all scales.

Parameter estimation can be done using several strategies, such as Maximum Likelihood (ML) methodology, or the "method of moments" (MoM) . Given the i^{th} ripplet detail sub-band with a set of n ripplet coefficients $x_i = (x_{i1}, x_{i2}, ..., x_{in})$ for $1 \le i \le 3J$, because 3 directional subbands from each scale are considered for computing the dominant texture features. The mean and variance of the i^{th} subband are computed as $m_i^1 = \frac{1}{N}\sum_{j=1}^{N}\left|x_{i,j}\right|$, $m_i^2 = \frac{1}{N}\sum_{j=1}^{N} x_{i,j}^2$ respectively.

The GGD model parameters α and β can be computed from the mean and variance as

$$\alpha_i = m_i^1 \frac{\Gamma(1/\beta)}{\Gamma(2/\beta)} \text{ for } 1 \le i \le 3J$$

$$\beta_i = F^{-1}\left(\frac{(m_i^1)^2}{m_i^2}\right) \text{ for } 1 \le i \le 3J$$

where $F(x) = \dfrac{\Gamma^2(2/x)}{\Gamma(3/x)\Gamma(1/x)}$

The feature vector for an image *I* in the database is constructed as

$$z = [\alpha_1, \beta_1, \alpha_2, \beta_2, ..., \alpha_{3J}, \beta_{3J}]$$

C. Similarity measurement using Kullback-Leibler Distance

In probability theory and information theory, the Kullback–Leibler divergence (also information divergence, information gain, relative entropy, or KLIC) is a non-symmetric measure of the difference between two probability distributions P and Q.

For discrete probability distributions P and Q, the K–L divergence of Q from P is defined to be

$$D_{KL}(P \,||\, Q) = \sum_i \ln\left(\frac{p(i)}{q(i)}\right) p(i) \quad (10)$$

In words, it is the expectation of the logarithmic difference between the probabilities P and Q, where the expectation is taken using the probabilities P.

For distributions P and Q of a continuous random variable, KL-divergence is defined to be the integral,

$$D_{KL}(P \,||\, Q) = \int_{-\infty}^{\infty} \ln\left(\frac{p(\mathrm{x})}{\mathrm{q}(\mathrm{x})}\right) p(\mathrm{x})\, \mathrm{dx}, \quad (11)$$

where p and q denote the densities of P and Q.

Given the GGD model, the PDF of ripplet coefficients in each subband can be completely defined via two parameters α and β . KL distance between two joint PDFs $p(X,Y)$ and $q(X,Y)$ is given by

$$D(p(X,Y) \,||\, q(X,Y)) = D(p(X) \,||\, q(X)) + D(p(Y \mid X) \,||\, q(Y \mid X))$$

The closed form for KLD between two GGDs is obtained by substituting the GGD joint PDF (9) in (11)= and is given by

$$D(p(.;\alpha_1,\beta_1)\,\|\,p(.;\alpha_2,\beta_2)) = \log\left(\frac{\beta_1\alpha_2\Gamma(1/\beta_2)}{\beta_2\alpha_1\Gamma(1/\beta_1)}\right)+\left(\frac{\alpha_1}{\alpha_2}\right)^{\beta_2}\frac{\Gamma((\beta_2+1)/\beta_1)}{\Gamma(1/\beta_1)}-\frac{1}{\beta_1} \quad (12)$$

Therefore, the similarity measurement between two wavelet subbands can be computed very effectively using the model parameters. Furthermore, with the reasonable assumption that wavelet coefficients in different subbands are independent, the overall similarity distance between two images is precisely the sum of KLDs between corresponding pairs of subbands as,

$$D(I_1,I_2)=\sum_{j=1}^{B} D(p(.;\alpha_1^{(j)},\beta_1^{(j)})\,\|\,p(.;\alpha_2^{(j)},\beta_2^{(j)})) \quad (13)$$

where B is the number of analyzed subbands, $\alpha_i^{(j)}$, $\beta_i^{(j)}$ denote the texture features extracted from the wavelet subband *j* of the image I_i. Thus the KLD theory provides with a justified way of combining distances into an overall similarity measurement, and no normalization on the extracted features is needed. The distance function defined in (12) is a function of three variables: the ratio of two scales α_1/α_2 and two shape parameters β_1 and β_2.

In retrieval experiments, the KLD values are calculated between the query image and all images in the training set. Similar images are retrieved when their KLD value with the query is the minimum among the remaining images in the training set.

III. EXPERIMENTAL RESULTS

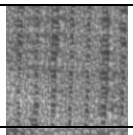
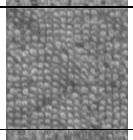
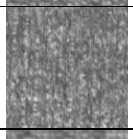
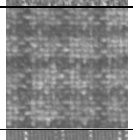
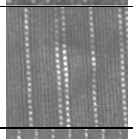
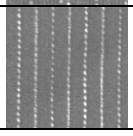
The retrieval performance of the proposed algorithm is tested on a complex Outex Database. Six classes of images are considered for the experiments and 160 images in each class are used to form the image database. The retrieval performance based on average recall is tabulated in Table 2. The proposed algorithm is compared with retrieval algorithms like gray level co-occurrence matrix, Statistical features like mean and variance calculated on steerable coefficients, and statistical features on ripplet coefficients. From table 2, it can be verified that the proposed algorithm yields maximum recall performance compared to the state of the art algorithms. From, the table 2, it is also clear that the images possess similarities between features across different classes and only an efficient algorithm can discriminate the images in each class. Thus, the proposed algorithm possess better discriminating capability compared to other retrieval algorithms.

The top 20 retrieved images of the proposed algorithm for the query from class 2 and class 5 are displayed in fig. 2 and fig. 3 respectively. The main advantage of the proposed system is its higher efficiency in representing edges and textures. Ripplet-II transform has the capability of capturing structure information along the arbitrary curves by tuning the scale, location, orientation, and degree parameters. Ripplet-II transform is invertible and orthogonality is also possible. Localization, directionality and flexibility are higher in Ripplet-II. Feature extraction with GGD parameters of the ripplet coefficients significantly improves the performance of the retrieval algorithm in comparison with the algorithm that utilized the mean and variance features of the ripplet coefficients.

IV. CONCLUSION

Thus, Ripplet transform based image representation is suitable for representing low level features of the images. The proposed CBIR system based on RT-II features is able to improve the accuracy of the retrieval performance at reduced computational cost. It also uses the idea of exploring the relationship among all images in the database and measures the relevance between the database and the query image. The retrieval performance is improved further by using GGD parameters as features and KLD for determining the similarity between these parameters. The experimental results show that Ripplet Transform-II outperforms than other transforms when applied to texture retrieval problem.

TABLE I
RETRIEVAL EVALUATION USING AVERAGE RECALL

CLASSES		Steerable filter with Statistical features	GLCM	RT with Statistical Features	RT with GGD
1		0.5375	0.6688	0.7534	0.8935
2		0.5797	0.6734	0.2534	0.775
3		0.4484	0.6922	0.6031	0.8243
4		0.7109	0.7156	0.2891	0.7815
5		0.5672	0.4672	0.7012	0.9351
6		0.5531	0.5734	0.6527	1.000

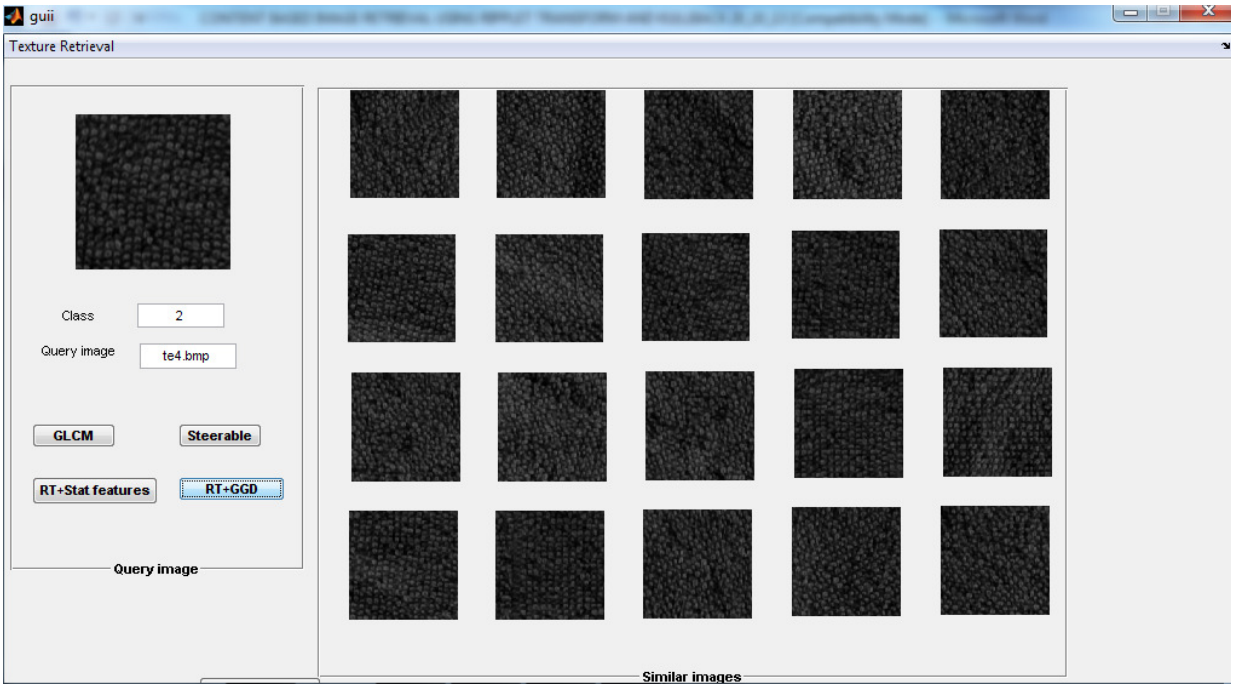


Fig. 2. Top 20 retrieved images using the proposed algorithm on a class 2 query image

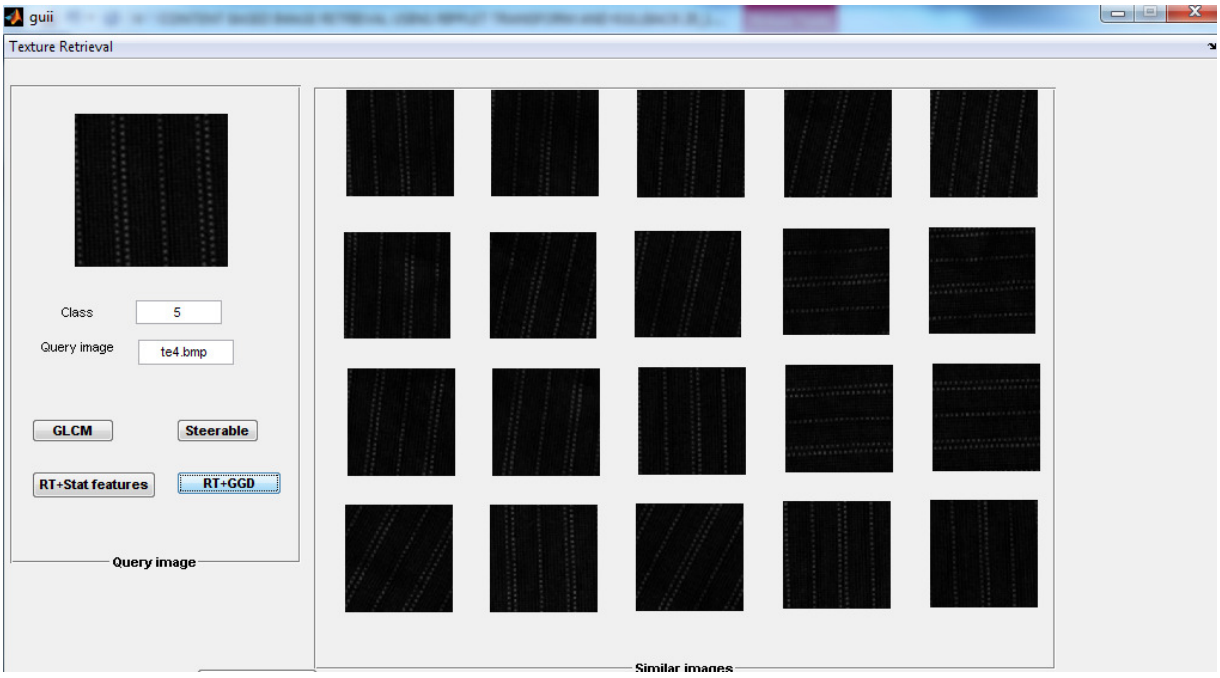


Fig. 3. Top 20 retrieved images using the proposed algorithm on a class 5 query image

REFERNCES

[1] Y. Rui, T. S.Huang, M. Ortega and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval", *IEEE Transactions on Circuits and Video Technology*, 1998.

[2] E.J. Candes and D. L. Donoho, "The curvelet transform for Image denoising", *IEEE Transactions on Image processing*, vol. 11, 2002.

[3] V. N Gudivada," Relevance Feedback in Content-based Image Retrieval", *Workshop Applied Imagery Pattern Recognition*, 2004.

[4] Y. Rui, T. S. Haung and S. - F. Chang, "Image Retrieval: Current Techniques, Promising Directions and open issues". *ACM Computing Surveys*, 2008.

[5] R. Rahmani, S. A. Goldman, H. Zhang, S. R. Cholleti, and J. E. Fritts, " Localized Content Based Image Retrieval". *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, Special Issue, 2008.

[6] Latika Pinjarkar, Manisha Sharma, Kamal Mehta, "Comparison and Analysis of Content Based Image Retrieval Systems Based On Relevance Feedback", *Journal of Emerging Trends in Computing and Information Sciences*, 2009.

[7] Peter Auer, Zakria Hussain, Samuel Kaski, Arto Klami, Jussi Kujala, Jorma Laaksonen, Alex P. Leung, Kitsuchart Pasupa, John Shawe-Tavlor, "Pinview: Implicit Feedback in Content-Based Image Retrieval". *In JMLR: Workshop and Conference Proceedings*, 2010.

[8] Dr. H. B. Kekre, Sudeep Thepade, Archana Athawale, Anant Shah, Prathamesh Verlekar, Suraj Shirke, "Walsh Transform over Row Mean and Column Mean using Image Fragmentation and Energy Compaction for Image Retrieval". *In International Journal on Computer Science and Engineering (IJCSE)*, 2010.

[9] Ch.Srinivasa Rao, S.Srinivas Kumar and B.Chandra Mohan, "Content Based Image Retrieval using Exact Legendre Moments and Support Vector Machine*". In The International journal of Multimedia and its applications (IJMA)*, 2010.

[10] Dorota G lowacka, John Shawe-Taylor, "Content-based Image Retrieval with Multinomial Relevance Feedback". *In JMLR: Workshop and Conference Proceedings and 2nd Asian Conference on Machine Learning (ACML)*, 2010.

[11] V.S.V.S. Murthy, E.Vamsidhar, J.N.V.R. Swarup Kumar, P.Sankara Rao Content Based Image Retrieval using Hierarchical and K-Means Clustering Techniques *in International Journal of Engineering Science and Technology*, 2010.

[12] Mann-Jung Hsiao, Yo-Ping Huang, Tienwei Tsai and Te-Wei Chiang, "An Efficient and Flexible Matching Strategy for Content-based Image Retrieval". *In Life Science Journal,* Vol 7, No 1, 2010.

[13] Jun Xu and Dapeng Wu,"Ripplet-II Transform for Feature Extraction", *In IET*, 2011.

[14] Minh N. Do and Martin Vetterli, "Finite Ridgelet Transform for Image Representation", *In IEEE Transactions and proceedings*, 2011.

[15] Patheja P.S., Waoo Akhilesh A. and Maurya Jay Prakash, "An Enhanced Approach for Content Based Image Retrieval". *In Research Journal of Recent Science*,Vol. 1(ISC) 2011.

[16] Minh N. Do, "The Contourlet Transform: An Efficient Directional Multi resolution Image Representation". *In International Journal of Computer Science and Application Issue*, 2011.

[17] Samuel Rota Bullo, Massimo Rabbi and Marcello Pelillo, "Content-Based Image Retrieval with Relevance Feedback using Random Walks". *In IEEE Transactions for Pattern Recognition*, 2011.

[18] Manish Chowdhury, Sudeb Das, and Malay Kumar Kundu, "Interactive Content Based Image Retrieval Using Ripplet Transform and Fuzzy Relevance Feedback". *In Springer-Verlag Berlin Heidelberg*, 2012.

[19] Latika Pinjarkar, Manisha Sharma and Kamal Mehta, "Relevance Feedback Techniques in Content Based Image Retrieval: A Survey". *In International Journal of Engineering Research and Development* ISSN: 2278-067X, Volume 1, Issue 9, 2012.

[20] Manimala Singha and K.Hemachandran, "Content Based Image Retrieval using Color and Texture". *In Signal & Image Processing: An International Journal (SIPIJ)*, 2012.

[21] Manish Chowdhury, Sudeb Das and Malay Kumar Kundu, "Novel CBIR System Based on Ripplet Transform Using Interactive Neuro-Fuzzy Technique". *In Electronic Letters on Computer Vision and Image Analysis*, 2012.