Complex Wavelet Transform with Vocabulary Tree for Content Based Image Retrieval

Anil Balaji Gonde
Ph. D. Student
Department of Electrical Engineering
IIT Roorkee
+91-9410508980
abgonde@gmail.com

R. P. Maheshwari *
Professor
Department of Electrical Engineering
IIT Roorkee
+91-9837229406
rpmaheshwari@gmail.com

R. Balasubramanian
Assistant Professor
Department of Mathematics
IIT Roorkee
+91-9412912081
balarfma@iitr.ernet

ABSTRACT

In this paper, combinations of spatial orientation tree (SOT), two-dimensional complex wavelet transform (CWT) and vocabulary tree (VT) is used for feature collection and retrieval of the images from natural as well as texture image database. SOT represents the parent-offspring relationship among the wavelet coefficients in multi-resolution wavelet sub-bands. Similarly, CWT captures directional information more accurately as compared to discrete wavelet transforms (DWT). SOT gives the set of descriptor vectors for each image which are further indexed by using vocabulary tree. The proposed method is tested on Corel 1000 and texture image database (Brodatz and USC) and the retrieval results have demonstrated a significant improvement in average precision, average recall and average rank compared to complex wavelet transform (CWT), optimal quantized wavelet correlogram (OQWC), Gabor wavelet correlogram (GWC).

Keywords

Complex wavelet transform, CBIR, Spatial orientation tree, Vocabulary tree, Gabor wavelet correlogram.

1. INTRODUCTION

1.1 Motivation

With advances in the internet and new digital image sensor technologies, large image database are being created by scientific, educational, medical, industrial and other applications. These large volumes of the images make difficult for a user to browse through the entire database. Therefore, an efficient and automatic procedure is a need for indexing and retrieving images from databases [1].

Traditionally, two approaches are used to retrieve the images: text based and content based approaches. In text based approaches, images are first annotated either manually or with the help of machine and then retrieved using traditional text retrieval techniques. Manual annotation is a cumbersome and expensive task for large image databases and also is often subjective in nature. Similarly, the first hurdle in machine annotation is the proper segmentation of image itself. As a result, it is difficult for the traditional text-based methods to retrieve a variety of images from database. In order to resolve this problem, a new technique known as content based image retrieval (CBIR) evolved. CBIR is a technique which uses visual contents of an image such as color, shape and texture to search images from large image databases. This technique finds application in advertising, medicine, crime detection, entertainment, and digital libraries. A comprehensive and recent literature survey on CBIR is available in [2, 3, 4, 5].

1.2 Related Works

Several low level features (color, shape, texture) have been extracted from images to aid the retrieval process. Texture identification plays a vital role in CBIR. No doubt in most of the applications of image retrieval, color features are used for image retrieval but in certain case single color feature is not sufficient to differentiate the image. Such as color of tiger and leopard looks similar but can be discriminated on the basis of black spots and strip on their skin thereby, depicting the importance of texture feature in CBIR. Texture contains important information about the structural arrangement of surface and their relationship to the surroundings. Varieties of techniques are developed for texture analysis. Haralick et al. [6] suggested the use of gray-level co-occurrence matrices (GLCM) which has become one of the most well-known and widely used texture features. The wavelet transform provides a multi-resolution approach to texture analysis and classification. Studies of human visual system support a multi-scale texture analysis approach, since researchers have found that the visual cortex can be modeled as a set of independent channels, each tuned to a particular orientation and spatial frequency band. Therefore wavelet transforms are found to be useful for texture feature extraction. Manjunath and Ma [7] proposed the Gabor transform for CBIR on Bordatz texture database and feature vectors were generated using the mean and standard deviation of four scale and six directions of Gabor transform coefficients. Mandal et al. [8] used Daubechies' wavelet in three scales to obtain transformed data. Further, histograms of wavelet coefficients in each sub-band were computed and stored to construct indexing feature vectors. Zhang et al. [9] used the Gabor wavelets transform for textured and natural images retrieval. The combination of multi resolution image decomposition and color correlation histogram is proposed by Moghaddam et al. [10, 11]. They used a combination of multiresolution analysis and color correlation histogram of the image to develop an algorithm where wavelet coefficients of the image are computed using Daubechies wavelets. A quantization step is applied before computing one-directional autocorrelograms of horizontal and vertical wavelet coefficients. Finally, index vectors are constructed using these wavelet correlogram. In later works, Saadatmand et al. [12, 13] presented an enhanced version of wavelet correlogram method using optimal quantization thresholds (OQWC). Moghaddam et al. [14] used Gabor transform in place of wavelet transform and that shows the significant improvement in rank, precision and recall compared to the wavelet correlogram (WC) and optimal quantized wavelet correlogram (OQWC). Birgale et al. [15] and Subrahmanyam et al. [16] used the combination of color histogram and Gabor wavelet transform features for CBIR. Gonde et al. [17] used the multiscale ridgelet transform features for retrieving the images from Corel database. Kim and Udpa [18] proposed a new set of two-dimensional (2-D) wavelet filters that are non-separable and oriented for improved characterization of diagonally oriented textures. Kokare et al. [19, 20] included the directional information of image using rotated wavelet filters, dual tree complex wavelet transform (DT-CWT) and dual tree rotated complex wavelet filters for texture image retrieval. Mohammad et al. [21] used the multiscale sub image matching method together with the discrete wavelet frames for retrieving the texture images without segmentation.

1.3 Main Contributions

2-D discrete wavelet transform decomposed the image into four sub-bands (*LL*, *LH*, *HL* and *HH*). *HH* sub band gives the diagonal information of the image but it is difficult to distinguish whether it has the direction of 45^0 or 135^0. In this paper authors suggest a new approach to conquer this problem. Flowchart of proposed method is as shown in figure 1:

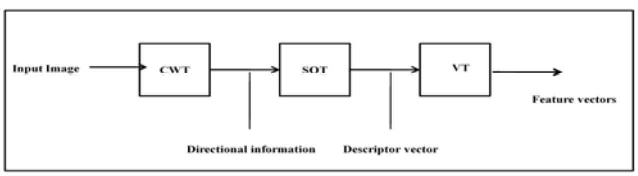

Figure 1: Flowchart of proposed method (CWT-VT)

In the flowchart, CWT provides directional information $\left(\pm 15^0, \pm 45, \pm 75\right)$. Next, descriptor vectors are generated from different sub-bands by using spatial orientation tree (SOT).Finally, the feature vectors are formed by mapping the set of descriptor vectors of image by using vocabulary tree. In most of the CBIR systems, mean, standard deviation and histogram of the transform coefficients are used as the features vectors [7, 15, 16, 17, 19, 20]. In proposed algorithm, vocabulary tree is used for indexing the image features [22]. The proposed method (CWT-VT) is tested on Corel and texture image database and the retrieval results shows a significant improvement in average precision, average recall rate, average rank compared to CWT, optimal quantized wavelet correlogram (OQWC) and Gabor wavelet correlogram (GWC).

The organization of the paper is as follows: In section 2, a brief review of CWT and SOT is given. Section 3 presents a concise review of vocabulary tree. A feature extraction and similarity measurement are presented in section 4. Experimental results and discussion can be observed in section 5 and finally section 6 is devoted to concluding remarks.

2. WAVELET THEORY

2.1 Dual Tree Complex Wavelet Transforms (DT-CWT)

Wavelet transform is used widely in the field of signal and image processing for analysis and reconstruction. But it suffers from two problems: The first one is the lack of shift invariant property and second is the poor directional selectivity for diagonal features. The shift invariant problem is resolved in *a' trous* wavelet transform using undecimated from of dyadic filter tree by paying a cost of high redundancy and increase in computation requirement. Kingbury [23] in his distinguished work introduced dual tree complex wavelet transform which overcome the problems of DWT. The properties of DT-CWT are summarized as

- Approximate shift invariance
- Good directional selectivity in m-dimensions where $m > 2$
- Perfect reconstruction using short linear phase filters
- Limited redundancy, independent of the number of scales, $\left(2^m : 1\right)$ for *m* dimensions
- Efficient order N-computation- 2^m times for *m* dimensions

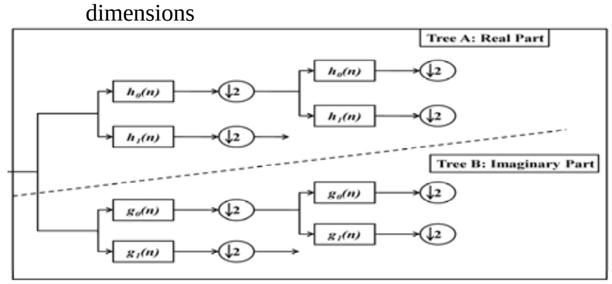

Figure 2: 1-D dual tree complex wavelet transforms

1-D DT-CWT is implemented using two real DWTs which operate in parallel on the same data as shown in figure 2. The first DWT gives the real part of the transform while the second DWT gives the imaginary part of complex wavelet transform. Let $h_0(n)$ and $h_1(n)$ denote the low pass/high pass filter pair on Tree A and $g_0(n)$ and $g_1(n)$ denote the low pass/high pass filter pair for the Tree B. Similarly the two real wavelets associated with each of the two real wavelet transforms are denoted as $\psi_h(t)$ and $\psi_g(t)$, then the complex wavelet $\psi_c(t)$ is given as $\psi_c(t) = \psi_h(t) + j\psi_g(t)$; where $\psi_g(t)$ is approximately the Hilbert transform of $\psi_h(t)$.The details of DT-CWT are available in [24, 25, 26]. The different subbands generated in Tree A and Tree B are further processed by spatial orientation tree technique as explained in section 2.2.

2.2 Spatial Orientation Tree (SOT)

In wavelet transform, an image is decomposed into subbands, such that lower subbands corresponds to higher image frequencies and higher subbands corresponds to lower image frequencies where most of the energy is concentrated. Similarly detailed coefficient values decreases as we move from highest to the

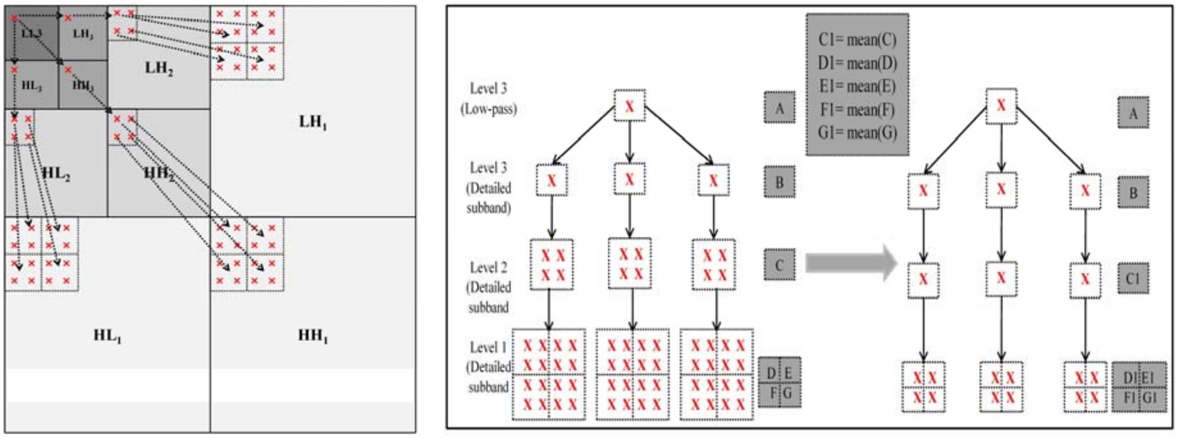

Figure 3: (a) Parent offspring relationship among the subbands, (b) Descriptor vectors construction

lowest levels of the subband pyramid [27]. It has been observed that there is a spatial self-similarity between subbands and the coefficients are expected to be of better magnitude order, if one moves downward in the pyramid following the same spatial orientation. A tree structure, called spatial orientation tree, defines the spatial relationship on the hierarchical pyramid. The tree is defined in such a manner that each directional subbands node has four offspring, which always forms a group of 2 x 2 adjacent pixels that has been shown clearly in figure 3 (a). The set of immediate descendants (offspring) of directional subbands are denoted by $O(i, j) = \{(2i, 2j), (2i-1, 2j), (2i, 2j-1), (2i-1, 2j-1)\}$.

The formation of descriptor vectors for three levels decomposition is shown in figure 3(b). For instance, descriptor vectors $d_{1,1}$ is given as

$$d_{1,1} = A, B_{lh}, B_{hl}, B_{hh}, C_{1_{lh}}, C_{1_{hl}}, C_{1_{hh}}, ..., G_{1_{lh}}, G_{1_{hl}}, G_{1_{hh}}. \quad (1)$$

Same procedure is repeated for rest of the descriptors vectors. For the image of size $X \times X$ and L level of decomposition, numbers of descriptor vectors (N_D) and the vector length of each descriptor vector (F_L) is given as

$$N_D = X_L \times Y_L$$

$$F_L = 4 + \sum_{i=2}^{L} 3 \times 4^{(i-2)} \quad (2)$$

where X_L and Y_L is the dimension of L^{th} level decomposition matrix

These descriptor vectors are further mapped on vocabulary tree (explained in section 3.2) for the formation of feature vectors of respective image.

3. VOCABULARY TREE

The vocabulary tree gives a hierarchical quantization threshold (cluster centers) that is built by hierarchical *k*-means clustering. Its algorithm contains three major steps:

- Unsupervised training of the tree using *k*-means clustering
- Mapping of image descriptor vectors on this tree
- Normalization of features

3.1 Unsupervised Training of The Tree Using *K*-means Clustering

The first task in VT building is that of collection of training vectors and for this purpose descriptor vectors of the different images are selected as the training vectors. Further, these training vectors are partitioned into different groups using *k*-means clustering technique. The algorithm can be observed as under:

- Let T numbers of the descriptor vectors are selected as training vectors.
- In level 1 clustering, T numbers of training vectors are grouped into k clusters using *k*-means clustering techniques.
- Store the respective cluster centres and cluster members.
- The same process is then recursively applied to descriptor vectors of each clusters so that the cluster members of respective group are further partitioned into *k*-groups. This process continues up-to a maximum L levels.

Figure 4 shows the flow chart of above algorithm.

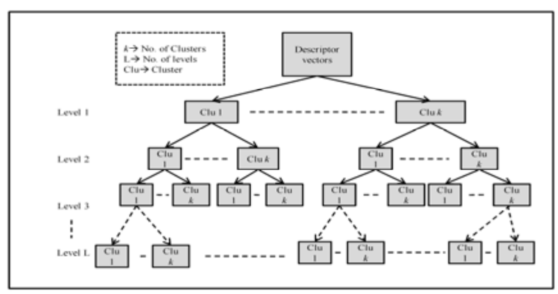

Figure 4: Vocabulary tree

By the end of this process, we will have the cluster centers values at different levels which are used in mapping process.

3.2 Mapping

After training a vocabulary tree, next step is to map the descriptor vectors of each image on the vocabulary tree using cluster centers of each group at different levels. The algorithm is as follows:

- Let an image has P numbers of descriptor vectors.

- Propagate each descriptor vectors down the tree by at each level comparing it to the k cluster centres and choosing the closest one using L_2 distance. At the same time, set a counter at every leaf nodes those count the number of descriptor vectors pass through it. This process is set at all levels so that we can collect an image feature at different levels. After exhausting all descriptor vectors, the counter values at different leaf nodes are further process at section 3.3.
- The above two steps are repeated for entire images of database.

3.3 Feature Vector Generation and Normalization

The feature vector of each image is defined at different levels. At L^{th} level and i^{th} leaf nodes, feature vector F_{Li} is given as

$$F_{Li} = m_i w_i \tag{3}$$

$$w_i = \ln \frac{N}{N_i} \tag{4}$$

where m_i –number of descriptor vectors of database image pass through the leaf node i
w_i – weight of leaf node i
N – Total number of images in the database
N_i – Number of images in the database with at least one descriptor vector pass through node i.

Similarly, normalization is used to achieve the fairness between database images with different descriptor vectors. Therefore, the normalized feature vectors at level L and node i is given as

$$F_{Nor} = \frac{F_{Li}}{\|F_{Li}\|} \tag{5}$$

This normalized feature vector is used in the formation of feature database.

4. FEATURE EXTRACTION AND SIMILARITY MEASURE

The performance of proposed method is compared with some existing methods. The brief review of these methods and feature collection procedure is explained bellow:

4.1 Complex Wavelet Transforms (CWT)

A 2-D DT-CWT operation is performed on each image of texture database and feature vector is formed by mean and standard deviation of each sub band [19]. For n numbers of sub bands, feature vector length is result in to $(n \times 2)$. Manjunath and Ma [7] proposed Gabor transform for CBIR on Bordatz texture database and feature vectors were generated using the mean and standard deviation of four scale and six directions of Gabor transform coefficients In retrieval phase, query image feature is compared with image database feature vectors using normalized Euclidean distance metric as a similarity measure, which is given by

$$NED(x,y) = \sum_m \sum_n d_{mn}(x,y)$$

where

$$d_{mn}(x,y) = \left| \frac{\mu_{mn}^x - \mu_{mn}^y}{\sigma(\mu_{mn})} \right| + \left| \frac{\sigma_{mn}^x - \sigma_{mn}^y}{\sigma(\sigma_{mn})} \right| \tag{6}$$

Where m and n are scale and orientation indices, μ_{mn} and σ_{mn} are the mean and standard deviation of the coefficients of wavelet subbands, $\sigma(\mu_{mn})$ and $\sigma(\sigma_{mn})$ are the standard deviations of the respective features over the entire image database. For Corel database, we used the procedure as given in [19] on detailed coefficients with Euclidean metrics.

4.2 Optimal Quantized Wavelet Correlogram (OQWC)

Moghaddam et al. [10, 11] proposed wavelet correlogram (WC) method where pyramidal three scale decomposition algorithm is used on each image. Next, they applied four levels quantization on *LH* and *HL* sub bands of DWT coefficients. Further, these quantized sub bands are passed through autocorrelogram operation that results into feature vectors of respective image. The more detailed information is available in [10, 11]. Saadatmand et al. [12, 13] optimized the quantization coefficients of wavelet sub bands in WC by using genetic algorithm. Rest of the feature collection procedure is same as WC. The final feature length for a given image is 96(3 x 2 x 4 x 4). Detailed information is available in [13].

4.3 Gabor Wavelet Correlogram (GWC)

Moghaddam et al. [14] proposed Gabor wavelet correlogram, where they used three scales and four orientations Gabor transform on each image results into twelve sub bands. These sub bands are quantized using four different levels. Further, autocorellogram operation with four distances is performed on quantized sub bands which results into feature vector of 192 lengths (3 x 4 x 4 x 4) for each image.

4.4 Proposed Method (CWT-VT)

In proposed method, each image is decomposed either three levels (texture database) or four levels (Corel database) and features are extracted as per the procedure given in section 2.2 and section 3. Following three sets of features are computed

- *Feature Set I*: Complex wavelet transform with vocabulary tree (CWT-VT) using mean operation at SOT.
  First step is to apply the 2-D dual tree complex wavelet transform on each image and decompose it up to four levels. Second step is to perform the SOT operation on these different subbands of Tree A and Tree B as per procedure given in the section 2.2. Output of SOT is the descriptor vectors which are further mapped on the vocabulary tree. The leaf nodes of vocabulary trees are used as the features vectors for respective image. For practical implementation, vocabulary tree is trained using ten cluster centers (k=10) and four level (L=4). Numbers of leaf nodes are 10,100, 1000 and 10000 are generated at level 1, 2, 3 and 4 respectively. In experiment work, leaf nodes at fourth level are used as a feature vector.
  $$F_{CWT-VT_{mean}} = f_{mean} \tag{7}$$
  where f_{mean} is the values available at leaf nodes of the fourth level of respective vocabulary trees.
- *Feature Set II*: Complex wavelet transform with vocabulary tree (CWT-VT) using standard deviation operation at SOT.
  The procedure of generation of feature vector is same as used in Feature Set I only standard deviation operation is performed among the coefficients of subbands at SOT

$$F_{CWT-VT_{std}} = f_{std} \quad (8)$$

where f_{std} is the values available at leaf nodes of the fourth level of respective vocabulary trees.

- *Feature Set III*: Combination of Feature Set I and Feature Set II

  This feature is obtained by combining equation (7) and (8). The final feature vector is given as

$$F_{CWT-VT} = F_{CWT-VT_{mean}} + F_{CWT-VT_{std}} \quad (9)$$

In experimental work, we used F_{CWT-VT}.

4.5 Similarity Measure

In the proposed method, a scoring is used as a similarity measure which calculate the score between query Q_i and database image I_i as given in following equation

$$Score = \sum_{i=1}^{Z} F_{Q_i} \times F_{I_i} \quad (10)$$

where F_{Q_i} - Feature vector of query image

Z -Number of leaf nodes

F_{I_i} - Feature vector of image in the database

Highest scoring database image is considered as more relevant image to query one.

5. EXPERIMENTAL ANALYSIS

In this section we discuss the effectiveness of combination of CWT and VT for CBIR. During the offline feature extraction stage, feature vectors are computed for all the images in the database using the procedure as explained in section 4.4. In online stage, the query image feature is compared with the feature vectors of images in the given database using the scheme mentioned in section 4.5. After all the database images have been compared, the images are then retrieved with decreasing similarity scores. Experiments are carried on two separate databases. The first consist of Corel 1000 database which include 1000 natural images [28]. The second database is the texture image database which comprises the Brodatz [29] and University of Southern California (USC) [30] texture image database. The evaluation takes into account several important factors, including precision, average recall rate and average rank.

5.1 Corel 1000 Database

Corel 1000 is the sub set of Corel database contains 1000 natural images of various categories ranging from animals and outdoor sports to natural images. These images are pre classified into different categories of size 100 by domain professionals. Corel 1000 database contains 1000 images of 10 different categories and each category has 100 images of sizes either 256 x 384 or 384 x 256. These categories are listed in Table 1.

For evaluation of the proposed method, query images are selected from a 1000 image subset of the Corel database. In advance, we know whether any two images are of the same category. In particular, a retrieved image is considered a match if and only if it is in the same category as the query. Each query results are display in descending order according to the scores between indexing vectors of the query and retrieved images. Following different parameters are used to quantify the performance of proposed method:

Precision as well as average precision P is given as

$$Precision = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved} \quad (11)$$

$$P = \frac{1}{N_q} \sum_{k=1}^{N_q} P(I_K) \quad (12)$$

Where N_q represents the number of queries and I_K represents the k^{th} image in the database. For global averaging, N_q =1000 and for averaging in a specified category, N_q =100.

Similarly recall, group recall and average recall is given as:

$$Recall(R) = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ relevant\ images} \quad (13)$$

$$Group\,Recall\,(GR) = \frac{1}{N_q} \sum_{i=1}^{Nq} R_i \quad (14)$$

$$Average\,Recall(AR) = \frac{1}{G} \sum_{j=1}^{G} GR \quad (15)$$

where

G- Number of groups in the database.

The rank of every query image q is computed over all matches using the following equation:

$$R_a(q) = \frac{1}{N_A} \sum_{i=1}^{N_A} rank(I_i)\ \forall I_i \in A \quad (16)$$

where N_A represents the number of images in the same category as the query image. The average rank are defined as

$$R_a = \frac{1}{N_q} \sum_{k=1}^{N_q} R_a(I_k) \quad (17)$$

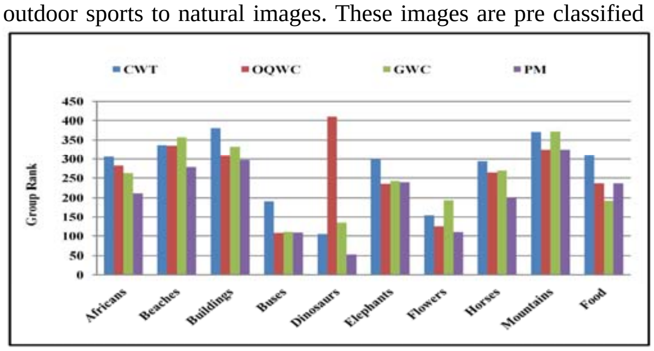

Figure 5: Comparative demonstration: Group rank vs. categories

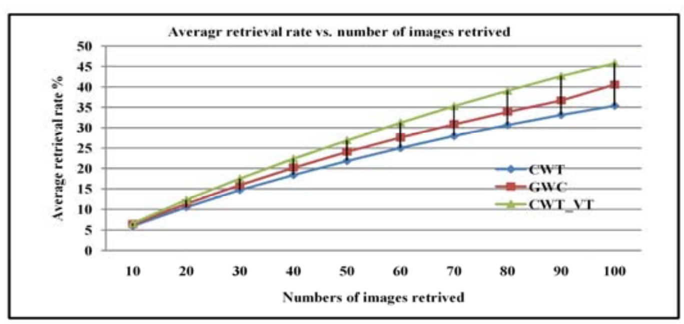

Figure 6: Average retrieval rate vs. number of images retrieved

Table 1: Comparative results of CWT-VT with existing methods

S. No.	Category	CWT		OQWC		GWC		CWT_VT	
		P *(N=10)%*	R_{ave} *(%)*	*P* *(N=10)%*	R_{ave} *(%)*	*P* *(N=10)%*	R_{ave} *(%)*	*P* *(N=10)%*	R_{ave} *(%)*
1	Africans	47.00	28.45	57.7	31.1	52.9	33.2	63.50	42.69
2	Beaches	45.00	27.78	49.3	28.6	42.0	26.2	50.00	31.70
3	Buildings	42.40	21.45	50.9	30.5	47.8	26.5	47.30	27.46
4	Buses	82.20	51.61	87.1	64.0	88.3	65.1	73.10	57.55
5	Dinosaurs	96.60	67.78	74.6	28.8	96.2	65.0	99.90	94.99
6	Elephants	49.20	25.89	55.7	30.7	65.9	37.0	58.10	35.52
7	Flowers	87.20	53.25	84.3	65.3	75.5	50.4	90.60	63.35
8	Horses	61.50	29.32	78.9	39.9	73.0	39.5	88.90	43.36
9	Mountains	32.60	19.93	47.2	25.1	35.2	20.1	33.40	23.54
10	Food	50.90	28.12	57.1	36.4	63.2	43.1	55.00	38.01
	Average	***59.46***	***35.36***	***64.30***	***38.0***	***64.1***	***40.6***	***65.98***	***45.82***

An ideal CBIR system, precision value is 1 and average rank is $N_A/2$, where N_A is the number of images in a specified category [11]

Table 1 represents the comparative results of CWT, OQWC and GWC with the proposed method (CWT-VT). Following different parameters are used for comparison purpose:

- Average precision
- Average recall

Following points are observed from Table 1

- Average precision of proposed method (65.98%) is more than CWT (59.46%), OQWC (64.30%) and GWC (64.10%).
- Average recall rate of proposed method (45.82%) is far better than CWT (35.36%), OQWC (38.0%) and GWC (40.60 %) method.

Figure 5 shows the comparative study between the proposed method and existing methods on the basis of group rank. It shows that out of ten categories selected, proposed method's group rank is superior in seven categories, equal in two categories and marginally less in one category. Also, the average rank of proposed method (206) is less than CWT (275), OQWC (263) and GWC (247).

5.2 Average Recall Rate Verses Number of Relevant Images Retrieved

Figure 6 presents the comparative results of proposed method (CWT-VT) with CWT and GWC on the basis of number of relevant images retrieved which further shows that the proposed method is superior to CWT and GWC. For top 10 images, average recall rate of proposed method is 6.60 % which further increases to 45.82% for top 100 images that is more than CWT (5.96%, 35.36%) and GWC (6.41%, 40.60%) respectively.

5.3 Retrieval Example

Figure 7 gives the sample retrieved result for query images of number 55 and 788. These images are selected among Corel database as the query image. For query image number 55, CWT-VT retrieved 10 images, GWC retrieved 6 images and CWT retrieved 4 images of same category among the first 10 retrieved images. Similarly, for query image 788, proposed method retrieved 9 images and GWC retrieved 7 and CWT retrieved 2 images of same category among the first 10 retrieved images.

5.4 Texture Database

Texture database used in this experiment consists of 116 different texture images with size of 512 x 512. Actually, this database is nothing but collection of 109 images of Brodatz texture Photographic album and 7 textures of USC database. Further, each image in the database is divided into sixteen 128 x 128 non overlapping sub images results into 1856 (116 x 16) texture mages. The performance of CWT-VT is measured in term of average retrieval rate (equation 15) and is compared with DWT, GWT, and dual tree complex wavelet (DT-CWT) [19]. Manjunath and Ma [7] used Gabor wavelet transform (GWT) on same texture database with four scale and six directions. Mean and standard deviation of the magnitude of Gabor transform coefficients were used as feature vector of image. Kokare et al. [19] used the same database for their experimental work and used energy and standard deviation of the magnitude of transform coefficient for feature vector construction.In proposed work, we use the three scale decomposition of image and further SOT and VT are employed for feature vector generation.

Table 2 shows the comparative results of CWT-VT with existing methods. Average retrieval rate of PM (76.60%) is far better than DWT (69.17%), GWT (74.19%), DT-CWT (74.73%), and DT-RCWT (71.17 %).

6. CONCLUSION

In this paper, authors propose a new method using complex wavelet transform with SOT and vocabulary tree. Two dimensional (2-D) DT-CWT captured six orientations information $\left(\pm 15^0, \pm 45, \pm 75\right)$ separately and SOT gives a new direction to arrange the descriptor vectors for CBIR application. Also the size and dimension problem of descriptor vectors is easily solved by using vocabulary tree method otherwise in such case low order statistics (mean and standard deviation) of the transformed images could not handle large numbers of descriptor vectors. The proposed method shows improved average recall rate (45.82%) over CWT (35.36%), OQWC (38.0%) and GWC (40.60%). Also, the average rank (216) of proposed method is less as compared to CWT (275), OQWC (263) and GWC (247). Similarly, proposed method proves its superiority on texture images. New image is added on the fly; no further retraining of VT is essential. Additional feature of CWT-VT is that it collects the local feature of image in transform domain. This is possible only because of the combination of SOT and VT. We are trying to explore this idea for detecting the particular object in image which is not possible in [13] and [19, 20]. Retrieval rate can further be improved if more orientation information is present in the future vector. An effort in this regard is currently underway and preliminary results are very encouraging.

Table 2: Comparative results of CWT-VT with existing methods on the basis of average retrieval rate

Level of decomposition	DWT (%)	GWT (%)	DT-CWT (%)	DT-RCWT (%)	CWT-VT (%)
3	69.17	74.19	74.73	71.17	76.60

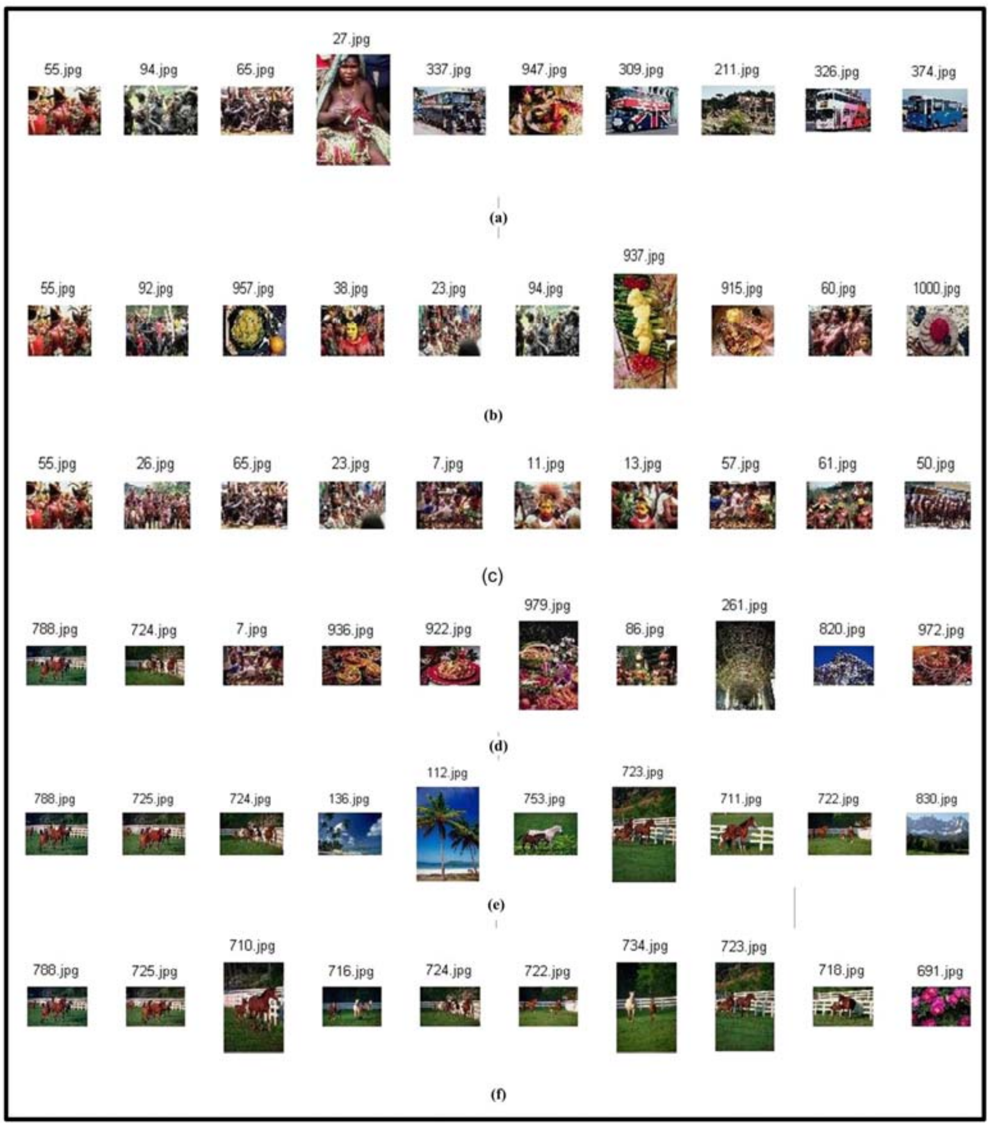

Figure 7: Query results of images 55 and 788 using: (a) & (d) CWT; (b) & (e) GWC; (c) & (f) CWT-VT

7. REFERENCES

[1] A. K. Jain and A. Vailaya. Image Retrieval Using Color and Shape. *Pattern Recognition*, 29(8), 233-244, 1996.

[2] Y. Rui and T. S. Huang. Image retrieval: Current techniques, promising directions and open issues," J. Visual Communication and Image Representation vol. 10 (1), 39-62, 1999.

[3] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain. Content bases image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Machine Intell.* 22 (12), 1349-1380, 2000.

[4] Y. Liu, D. Zhang, G. Lu and W. Y. Ma. (2007) A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, vol. 40, pp. 262-282, 2007.

[5] R. Datta, D. Joshi., J. Li and J. Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computer. Surv*., 40 (2), 1-66,2008.

[6] R. M. Haralick R, K. Shanmugam and I. Dinstein. Texture features for image Classification. *IEEE Trans. on Sys. Man and Cyb.,* 3 (6), 1450-1461, 1993.

[7] B. S. Manjunath and W.Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Anal. Mach. Intell*., 18 (8), 837–842, 1996.

[8] M. K. Mandal, S. Panchanathan and T. Aboulnasr. Fast wavelet histogram techniques for image indexing. *J. of Computer Vision Image understanding,* 75 (1/2), 99–110, 1999.

[9] D. Zhang, A. Wong, M. Indrawan and G. Lu. Content-based Image Retrieval using Gabor Texture Features. *IEEE International Symposium on Multimedia Information Processing,* Sydney, Australia, 392-395, 2000.

[10] H. A. Moghaddam, A. H. Rouhi and T Taghizadeh Khajoie. A new algorithm for image indexing and retrieval using wavelet correlogram. *IEEE International Conference on Image Processing (ICIP),* Spain, 3, 497–500, 2003.

[11] H. A. Moghadam, T. Taghizadeh Khajoie, A. H. Rouhi and M. Saadatmand-T. Wavelet correlogram: a new approach for image indexing and retrieval. *J. Pattern Recognition*, 38, 2506 – 2518, 2005.

[12] M. Saadatmand-T. and H. A. Moghadam. Enhanced Wavelet Correlogram Methods for Image Indexing and Retrieval. *IEEE International Conference on Image Processing (ICIP*), Genoa, Italy, 1, 11-14, 2005.

[13] M. Saadatmand-T. and H. A. Moghaddam. A Novel Evolutionary Approach for Optimizing Content-Based Image Indexing Algorithms. *IEEE Transactions On Systems, Man, And Cybernetics—Part B: Cybernetics,* 37(1), 139-153, 2007.

[14] H. A. Moghaddam and T. M. Saadatmand. Gabor wavelet Correlogram Algorithm for Image Indexing and Retrieval. *18th Int. Conf. Pattern Recognition (ICPR),* K. N. Toosi Univ. of Technol., Tehran, Iran, 925-928, 2006.

[15] L. Birgal, M. Kokare, D. Doye. Color and texture features for content based image retrieval. *International Conf. Computer Graphics, Image and visualization,* Washington, DC, USA, 146-149, 2006.

[16] M. Subrahmanyam, A. B. Gonde, R. P. Maheshwari Color and Texture features for Image Retrieval. *IEEE International Advance Computing conference, Thapar Institute of Engineering and Technology*, Patiala, India, pp. 1411-1416. 2009.

[17] A. B. Gonde,. R.P. Maheshwari and R. Balasubramanian. Multiscale Ridgelet Transform for Content Based Image Retrieval. *IEEE International Advance Computing conference (IACC 10),* Thapar Institute of Engineering and Technology, Patiala, India, 139-144, 2010.

[18] N. Kim and S Udpa. Texture classification using rotated wavelet filters. *IEEE Trans. Syst., Man and Cybernet Part A: Systems and Humans,* 30 (6), 847-852, 2000.

[19] M. Kokare., P.K. Biswas and B. N. Chatterji. (2005), Texture image retrieval using new rotated complex wavelet filters. *IEEE Trans. Systems, Man and Cybernetics*, 35(6), 1168-1178, 2005.

[20] M. Kokare, P. K .Biswas and B. N. Chatterji. (2007) Texture image retrieval using rotated wavelet filters. Pattern *recognition letters*, 28, 1240-1249, 2007.

[21] M. F. Ahmad Fauzi and P. H. Lewis. A multiscale approach to texture-based image retrieval. *Pattern Anal. Application*, 11, 141-157, 2008.

[22] D. Nister and H. Stewenius. Scalable Recognition with a Vocabulary Tree. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2161-2168, 2006.

[23] N. G. Kingsbury. Image processing with complex wavelet. *Phil. Trans. Roy. Soc. London*, 357,2543-2560, 1999.

[24] N. G. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signal. *J. of Applied and Computational Harmonic Analysis*, 10(3), 234-253, 2001.

[25] I. W. Selesnick. The design of approximate Hilbert transforms pairs of wavelet bases. *IEEE Trans. Signal Process.* , 50(5), 1144-1152, 2002.

[26] I.W. Selesnick, R.G. Baraniuk and N. G. Kingsbury. The dual-tree complex wavelet transform. IEEE Signal Processing Magazine, 123-151, 2005.

[27] Said and W. A. Pearlman. A new, fast and efficient Image codec based on set partitioning in hierarchical trees. *IEEE Trans. on circuit and systems for video tech.*, 6 (3), 1996.

[28] http://wang.ist.psu.edu/docs/related.shtml.

[29] P. Brodatz . Textures: A Photographic Album for Artists and Designers. New York: Dover, 1966.

[30] University of Southern California, Signal and Image Processing Institute. [Online]. Available: http://sipi.usc.edu/services/database/Database.html