*International Journal of Computer Applications (0975 – 8887)*
*Volume 14– No.6, February 2011*

# Image Retrieval using 2D Dual-Tree Discrete Wavelet Transform

N S T Sai
Tech Mahindra Limited,
Mumbai, India.

R C Patil
Ph.D.Research Scholar,
Electronics Department,
Mukesh Patel School of Tech.
Mgmt. and Engineering, NMIMS
University, Mumbai, INDIA

## ABSTRACT
The large amount of image collections available from a variety of sources has posed increasing technical challenges to computer systems to store/transmit and index/manage the image data to make such collections easily accessible. Here to search and retrieve the expected images from the database we need Content Based Image Retrieval system. This paper proposes a new feature vector based on 2D Dual-tree Discrete Wavelet Transform. One of the advantages of the dual-tree complex wavelet transform is that it can be used to implement 2D wavelet transforms that are more selective with respect to orientation than is the separable 2D DWT. Most of the natural images have short span high frequencies and low frequencies extending for larger span. Hence, the design of our feature vector is such that it provides higher spatial localization and lower frequency resolution at higher frequencies and the reverse for lower frequencies. The energy and mean of the frequency content of the image at various sub bands and different spatial resolution (higher for higher frequency bands) is stored as feature vector. Thus, the given feature vector encodes high frequency information as well.

## General Terms
Content Based image retrieval aims at developing new effective techniques to search and browse similar images from the large image database by analyzing the image contents. With the rapid development of technology of multimedia, the traditional information retrieval techniques based on keywords are not sufficient, content - based image retrieval (CBIR) has been an active research topic.

## Keywords
CBIR; DDWT; Wavelet Transform; Precision; Recall; Euclidean Distance.

## 1. INTRODUCTION
With the rapid development of technology of multimedia, the traditional information retrieval techniques based on keywords are not sufficient, content - based image retrieval (CBIR) has been an active research topic.

Image retrieval finds its application in different domains such as multimedia, satellite image databases, medical imaging etc. Traditional exhaustive manual searches can no longer be done over the data set of thousands of images. Image retrieval techniques can be broadly classified as i) Text based, ii) Content based. Earlier image retrieval techniques were based on text. It has many disadvantages such as i) manual tagging of all the images with keywords, ii) inherent difficulty of describing certain aspects, iii) highly subjective nature [1]. This makes the manual approach inadequate for the increasing database. In order to make image retrieval more effective, Content-based Image Retrieval (CBIR) has been introduced [2]. CBIR aims at automatic extraction of features based on the mathematical characteristics and contents of the image [3] [4].

Typical CBIR involves two phases. In the first phase, some feature characterizing each image in the database is computed and stored as feature vectors. In the second phase, the same set of feature vector is calculated for the user given query image and it is compared with all the stored feature vectors using distance measure such as Euclidian distance as shown in Fig.1.

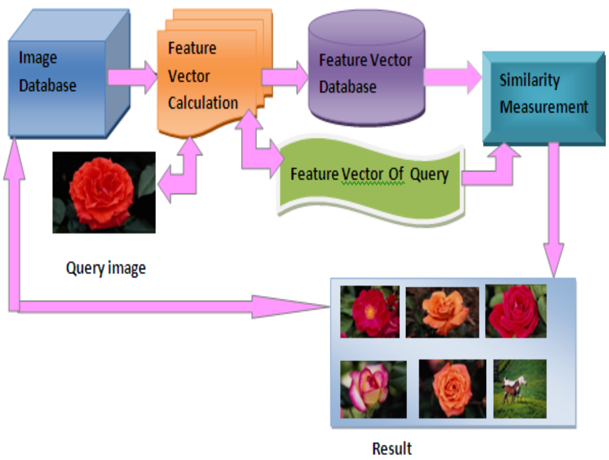



**Fig 1: Content Based Image Retrieval System.**

The images most similar to the query image are returned to the user. The features used should be effective in matching similar images and discriminating dissimilar ones. Color, shape and texture [5] [6] are the most widely used features. Color histogram is one of the widely and extensively used visual feature to index color images, but two images which are not semantically related may have same color histogram. Shape based indexing depends largely on the output of the shape detection algorithms and texture features doesn't work well for non-texture images.

Wavelet based features were first introduced in [14]. Jacobs et al. selects sixty four largest Haar wavelet coefficients in each of the

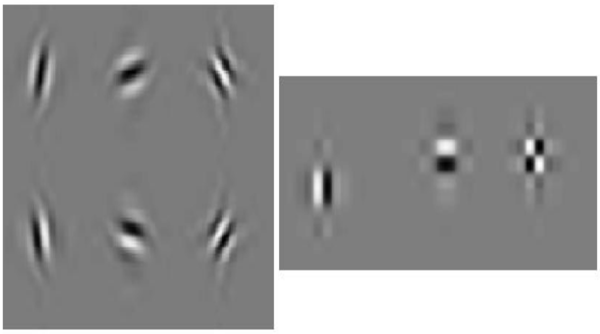
3 color band and stores them in feature vector as +1 or -1 along with their position in the transformation matrix. Low frequency coefficients tend to be more dominant than those of the high frequency coefficients and this makes this algorithm ineffective for images with sharp color changes. In addition to that, Haar wavelet basis is not suitable for natural images.

Wang et al. [15] have used Daubecheis wavelet for multi resolution feature vector. Few of the wavelet coefficients in the lowest frequency band and their variances are used as feature vector. To decrease the retrieval time, a crude selection is done based on the variances and further refined selection is based on 768 dimensional feature vectors. This also discards most of the high frequency information similar to [16] [17]. Two different images with same low frequency content and different pattern of sharp edges may look similar in this feature space. This feature space doesn't well quantify the texture information of the images. In [18], local features of each sub-band including the high frequency bands are computed as feature vector. They have used total energy of each sub-band as a feature to represent the content of each sub-band. This is an efficient feature space as it doesn't discard the high frequency information altogether. This feature vector has good frequency resolution, but no space resolution. This leads to dissimilar images being presented as similar image to the user.

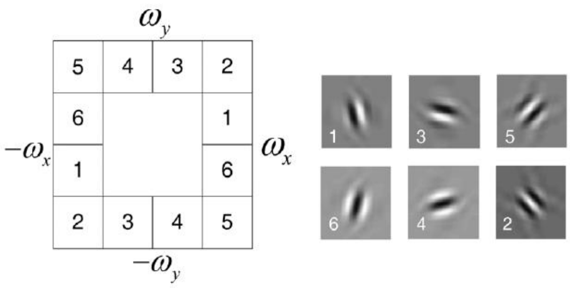
The rest of this paper is organized as follows: Section 2 gives a brief introduction of 2D DDWT. Proposed method in Section 3. Feature vector matching presented in Section 4 implementation and result in Section 5.followed by conclusion in section 6.

## 2. 2D DUAL-TREE DISCRETE WAVELET TRANSFORM [21][22]

DDWT is developed to overcome two main drawbacks of DWT: shift variance and poor directional selectivity [21]. With carefully designed filter banks, DDWT mainly has following advantages: approximate shift invariance, directional selectivity, limited redundancy, and similar computation efficiency as DWT. Either the real part or the imaginary part of DDWT [21][22] yields perfect reconstruction and thus can be employed as a stand-alone transform. We use magnitude of subbands to calculate feature vector.The implementation of DDWT is very straightforward. An input image is decomposed by two sets of filter banks, ($H_0^a$, $H_1^a$) and ($H_0^b$, $H_1^b$) separately, filtering the image horizontally and then vertically just as conventional 2D DWT does. Then eight sub bands are obtained: $LL_a$, $HL_a$, $LH_a$, $HH_a$, $LL_b$, $HL_b$, $LH_b$ and $HH_b$. Each high-pass subband from one filter bank is combined with the corresponding subband from the other filter bank by simple linear operations: averaging or differencing. The size of each subband is the same as that of 2D DWT [19][21][22] at the same level. But there are six highpass subbands instead of three highpass subbands at each level. The two lowpass subbands, $LL_b$ and $LL_a$, are iteratively decomposed up to a desired level within each branch. The basic functions of 2D DDWT and 2D DWT are shown in Fig. 2a and Fig. 2b respectively. Each DDWT basis function is oriented at a certain direction, including ±75º, ± 15º, and ±45º. However, the basis function of HH subband of 2D DWT mixes directions of ±45º together.



(a) (b)




(c)

**Fig 2: (a) Six basis functions of 2D DDWT (real part) at level 3 and (b) three basis functions of 2D separable DWT at the same level.(c) frequency tiling of 2-D DDWT and six wavelets of 2-D DDWT.**

Typical wavelets of six 2-D DDWT subbands together with their ideal spectrum supports are illustrated in Fig. 2(c). [19][20][21].

## 3. PROPOSED METHOD

A colour space is a model for representing colours in terms of intensity values. RGB colour space is fundamental colour space in imaging. Compared with the RGB colour space YCbCr most suitable for the human visual system and it is also used in the proposed colour layout descriptor in the international standard MPEG-7[13]. Hence during the decoding process we extract feature directly from the luminance (Y ) and chrominance (Cb,Cr).

As feature vectors includes information about the high frequency bands, it will be able to group similar images better than those feature vectors which considers only the low frequency information. The main drawback of this is the lack of spatial information in the feature vector. Spatial information is of critical importance in high frequency bands than that of lower frequency bands. This is because high frequencies mostly occur for a shorter span than that of low frequencies. High frequency spatial information is obtained using following steps.

- The images in the database are rescaled to 256x256 pixels and converted to YCbCr..
- The image is subjected to first level Dual-tree Discrete Wavelet decomposition resulting in 1*6+2 real and imaginary subbands.
- Take magnitude of subbands so resulting in 1*3+1 subband.
- $HL_1, LH_1, HH_1$ sub and divided into 16x16 sub block and compute the energy of each sub block and mean of $LL1$. The energy of the frequency content of each sub-image forms the elements (E1 to E192) and mean is E193 element of feature vector for only one color plane thus for whole color image it is (E1-E579) for first level decomposition as shown in Fig.3.
- $LL1$ gives low frequency information is further sub divided in second level decomposition $LL_2, HL_2, LH_2, HH_2$.
  $LL_2, HL_2, LH_2, HH_2, LL_1, HL_1, LH_1, HH_1$ contains high frequency information .As high frequencies require good spatial resolution, the bands of level 2 and 1 are sub-divided into sub-images of size 32x32 and 16x16 respectively and mean of low frequency sub band $LL_2$ .Make a feature vector (E1-E205) for each color space. So feature vector size for color image is E1-E615 for second level decomposition.
- Apply third level decomposition on $LL_2$ it gives $LL_3, HL_3, LH_3, HH_3$. The sub bands of level 2 and 1 are sub-divided into sub-images of size 32x32 and 16x16 respectively. Mean of $LL_3, HL_3, LH_3, HH_3$ sub band make a feature vector size E1-E624 for third level decomposition.
- Apply forth level decomposition on $LL_3$ it gives $LL_4, HL_4, LH_4, HH_4$. The sub bands of level 2 and 1 are sub-divided into sub-images of size 32x32 and 16x16 respectively. Mean $HL_3, LH_3, HH_3$ and mean of forth level sub band $LL_4, HL_4, LH_4, HH_4$ make a feature vector size E1-E633.
- Apply fifth level decomposition on $LL_4$ it gives $LL_5, HL_5, LH_5, HH_5$. The sub bands of level 2 and 1 are sub-divided into sub-images of size 32x32 and 16x16 respectively. Mean of $HL_3, LH_3, HH_3$ and mean of forth level sub band $HL_4, LH_4, HH_4$ and mean of $LL_5, HL_5, LH_5, HH_5$ make a feature vector size E1-E642.

Thus as decomposition level goes from first to fifth level then feature vector size goes on increasing and that size is given table 1 .

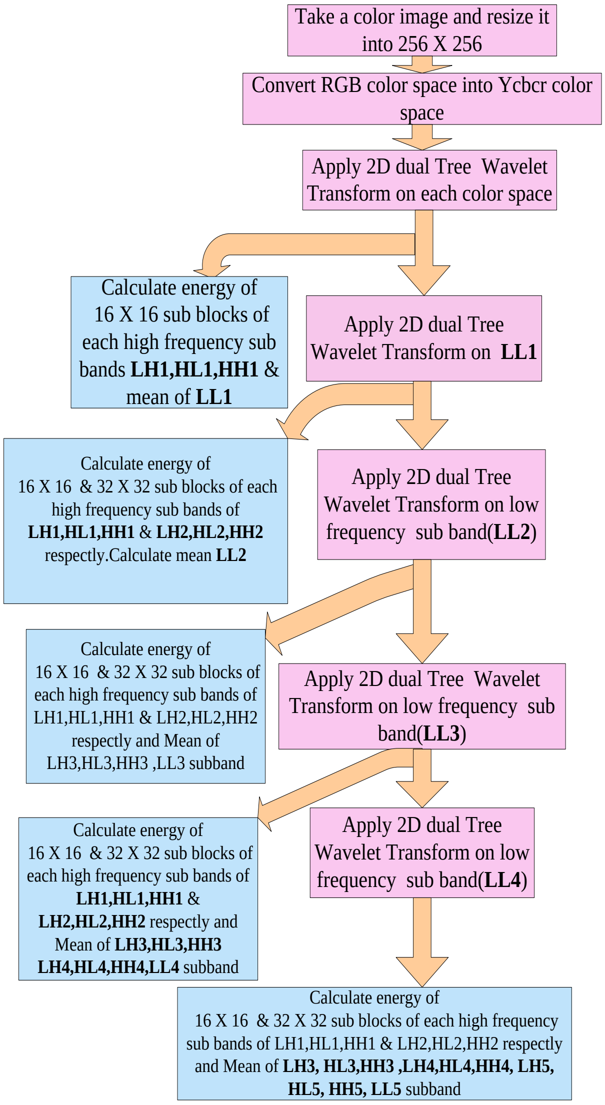



**Fig 3: Feature Vector generation.**

**Table 1. Feature vector size for color image having size 256 X 256**

| | 2 D-DT-CWT Decomposition Level | | | | |
|---|---|---|---|---|---|
| | First | Second | Third | Fourth | Fifth |
| Feature Vector Size | 579 | 615 | 624 | 633 | 642 |

## 4. FEATURE VECTOR

When a query image is submitted by a user, we need to compute the feature vector as before and match it to the precomputed feature vector in the database. This is shown in Figure 4. The simulation engine consists of feature extraction process, batch[3,4]. The feature extraction process is based upon the following .Which the batch feature extraction and storage process as described in the following steps.

a. Images are acquired from a collection one after another.
b. Feature extraction process is applied to them.

c. The resultant vector is saved in a database against the image name under consideration.

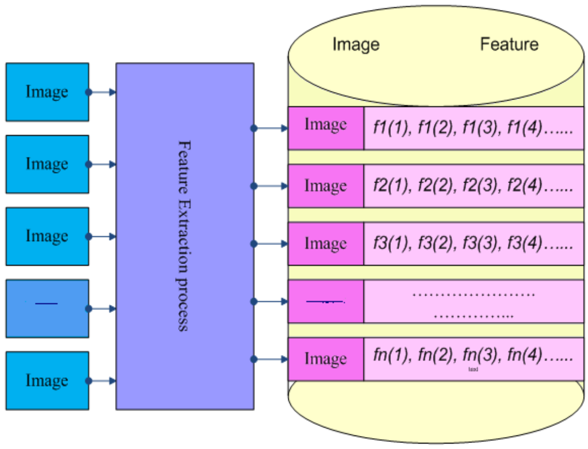



**Fig 4: Feature extraction and storage process for an image collection**

After that query image and database image matching is done using Euclidean distance. The different types of distances which are used by many typical CBIR systems are city block distance, chess board distance, intersection distance , the Earth mover's distance (EMD), Euclidian distance. Minkowski (Euclidean distance when r=2) distance is computed between each database image & query image on feature vector to find set of images falling in the class of query image.

$$Ed(Q,I)=\left(\sum_{M=0}^{M-1}\left|\mathrm{H}_Q-H_I\right|^r\right)^{1/r} \qquad (1)$$

Where Q-Query image
I-Database image.
$H_Q$-Feature vector query image.
$H_I$-Feature vector for database image.
M-Total no of component in feature vector.

## 5. IMPLIMENTATION AND RESULT

The implementation of CBIR technique is done in MATLAB 7.0 using a computer with Intel Core 2 Duo Processor T8100 (2.1GHz) and 2 GB RAM. The CBIR technique are tested on the image database of 800 variable size images include 8 categories of animals, buses, flowers ,bikes, beaches, Historical, Mountains etc[27].

Sample images from each category shown in Fig.5 For 40 query images (five from each category from database) the precision and recall is calculated for proposed methods and average recall and average precision is plotted against decomposition level for each category image. Database of 800 images of 8 different classes is used to check the performance of the algorithms developed. The query image and database image matching is done using Euclidean distance which is given equation 1.The average precision and average recall is calculating by using following equation 2,3.

**Fig 5: Sample images from database.**

The average precision for images belonging to the qth category (Aq) has been computed by:

$$\bar{P}_q=\sum_{k\in Aq} P(I_K)/\left|(A_q)\right|, q=1,2,....5 \qquad (2)$$

Finally, the average precision is given by:

$$\overline{P}=\sum_{q=1}^{5}\overline{P_q}/5 \qquad (3)$$

From Fig 6. to Fig. 13. gives average precision and recall plots against the 2 D Dual Tree wavelet transform decomposition level for each class of image. These two performance parameter judge which decomposition level suitable for this image retrieval system.

It is observed that the beach class image category and historical mountain image category average precision and recall performance is average compared to the remaining class of images. Another important point is noted that as decomposition level goes on increasing the precision value goes on increasing and recall value goes on decreasing.

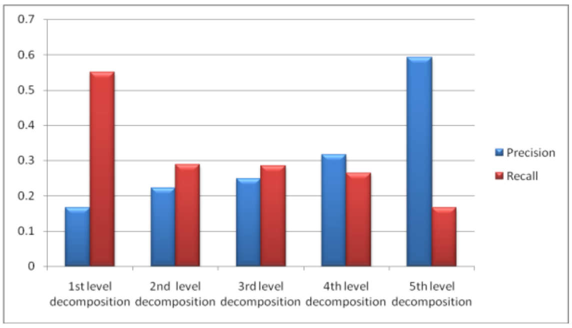



**Fig 6: Average Precision and Average Recall performance for Beach class of images.**

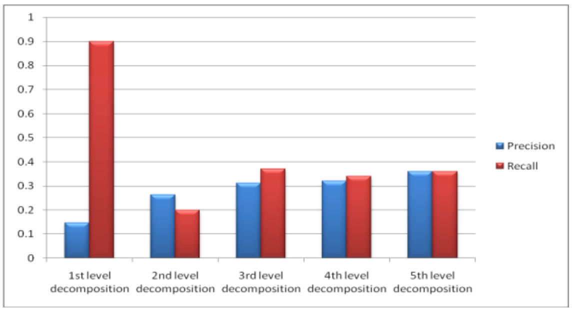



**Fig 7: Average Precision and Average Recall performance for Historcal Mountains class of images.**

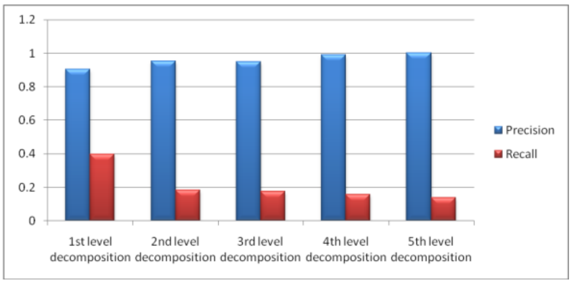



**Fig 8: Average Precision and Average Recall performance for Horse class of images.**

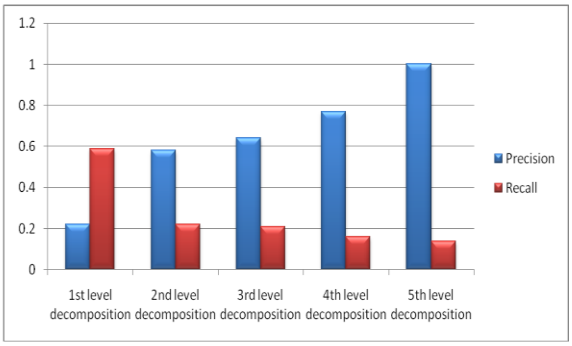



**Fig 9: Average Precision and Average Recall performance for Elephant class of images.**

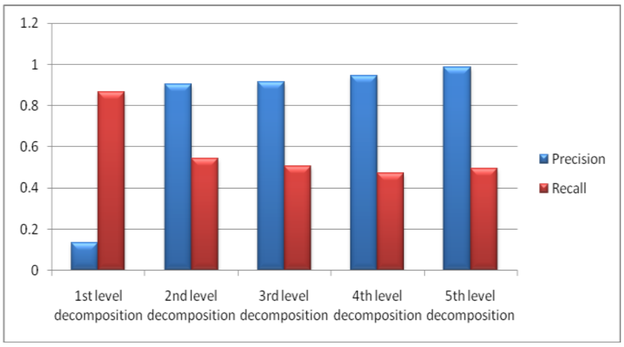



**Fig 10: Average Precision and Average Recall performance for Bike class of images**

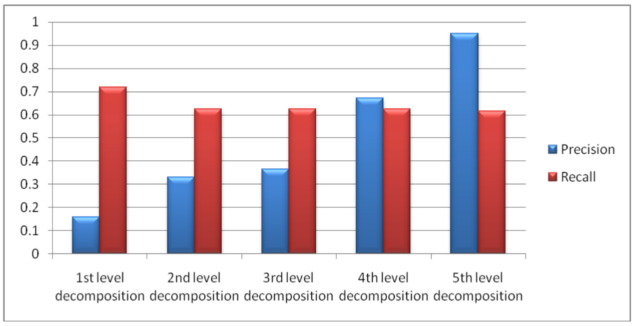



**Fig 11: Average Precision and Average Recall performance for Dinosaur class of images**

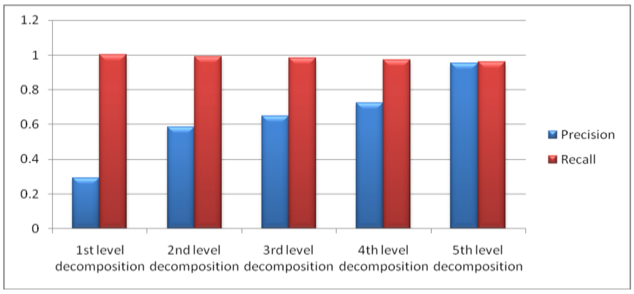



**Fig 12: Average Precision and Average Recall performance for Rose class of images**

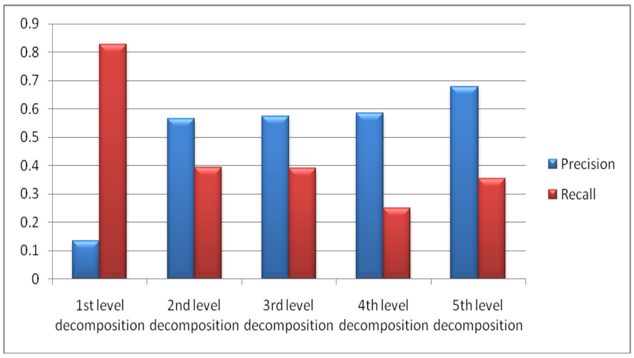



**Fig 13: Average Precision and Average Recall performance for Bus class of images**

Rose, bus, bike, dinosaur image class have good precision and recall performance.

**Table 2. Overall average recall and average precision for all decomposition level for all class of images.**

| Sr.No | DT-CWT Decomposition Level | Average Precision | Average Recall |
|---|---|---|---|
| 1. | 1st Level Decomposition | 0.236 | 0.731 |
| 2. | 2nd Level Decomposition | 0.55 | 0.431 |
| 3. | 3rd Level Decomposition | 0.599 | 0.44 |
| 4. | 4th Level Decomposition | 0.604 | 0.462 |
| 5. | 5th Level Decomposition | 0.8 | 0.4 |

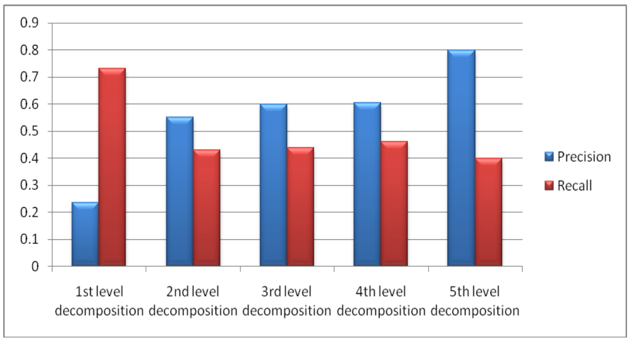



**Fig 14: Average Precision and Average Recall performance for Average Precision and Average Recall of all decompositon level for all class images shown in Fig 6 to Fig 13 .**

Table 2 gives overall average recall and average precision for all decomposition level for all class of images. Fig. 14 shows the plot of overall recall and overall precision against the decomposition level. On the plot it is observed that precision for 5$^{th}$ level decomposition level is very good and recall for 1$^{st}$ level is very good but precision value is not that much good.

## 6. CONCLUSION

The search for the relevant information in the large space of image databases has become more challenging. More précised retrieval techniques are needed to access the large image achieves being generated, for finding relatively similar images. The proposed methods are 2-D DDWT on color image.

We have proposed a new wavelet based feature which includes both low and high frequency information. It incorporates i) high frequency resolution and less spatial resolution in lower frequency sub-bands, ii) low frequency resolution and high spatial resolution in higher frequency bands. The precision value goes on increasing and recall value goes on decreasing as decomposition level increases. On the plot it is observed that 2$^{nd}$,3$^{rd}$ and 4$^{th}$ level decomposition of 2-D Dual discrete wavelet transform performance is good in terms of both precision and recall value.

## 7. REFERENCES

[1] NST Sai, Ravindra patil ,"Average Row and Column Vector Wavelet Transform for CBIR", Second international conference on Advance in Computer Vision and Information Technology (ACVIT2009),Aurangabad, India.

[2] NST Sai, Ravindra patil ,"New Feature Vector for Image Retrieval Walsh Coefficients", Second international conference on Advance in Computer Vision and Information Technology (ACVIT2009),Aurangabad, India.

[3] NST Sai, Ravindra patil ,"Image Retrieval using DCT Coefficients of Pixel Distribution and Average Value of row and Column Vector "IEEE International Conference on Recent Trends in Information ,Telecommunication and Computing(ITC2009),Kochi, Kerala, India.

[4] NST Sai, Ravindra patil," Moments of Pixel Distribution of CBIR" International Conference and Workshops on Emerging Trends in Technology (ICWET2010),Mumbai, India.

[5] NST Sai, Ravindra patil ,"New Feature Vector for Image Retrieval: Sum of the Value of Histogram Bins "IEEE Conference on Advance in Computing, Control & Telecommunication Technologies (ACT2009),Trivandrum, India.

[6] NST Sai, Ravindra patil,"Image Retrival usng Equalized Histogram Image Bins Moment" Inter national Joint Journal Conference in Engneering ,IJJCE,2010,Trivandrum,India.

[7] M. K. Mandal, T. Aboulnasr, and S. Panchanathan,, "Image Indexing Using Moments and Wavelets", IEEE Transactions on Consumer Electronics, Vol. 42, No. 3, August 1996.

[8] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC System", IEEE Computer, 28(9):23–32, Sept. 1995.

[9] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI), 22(12):1349–1380, Dec. 2000.

[10] K. Hirata and T. Kato, ªQuery by Visual Example," Advances in Database Technology EDBT '92, Third Int'l Conf. Extending Database Technology, 1992.

[11] W.Y. Ma and B.S. Manjunath, "Pictorial Queries: Combining Feature Extraction with Database Search," Technical Report 18, Dept. of Electrical Eng., Univ. of California at Santa Barbara, 1994.

[12] A. Gupta and R. Jain, ªVisual Information Retrieval," Comm. ACM, vol. 40, no. 5, 1997.

[13] W.Y. Ma and B.S. Manjunath, "Pictorial Queries: Combining Feature Extraction with Database Search," Technical Report 18, Dept. of Electrical Eng., Univ. of California at Santa Barbara, 1994.

[14] C.E. Jacobs, A. Finkelstein, and D.H. Salesin, "Fast Multiresolution Image Querying," Proc. SIGGRAPH 95, 1995.

[15] W.W.J.Z. Wang, G. Wiederhold, O. Firschein, and S.X. Wei, “WaveletBased Image Indexing Techniques with Partial Sketch Retrieval Capability,” J. Digital Libraries, 1997.

[16] Seung Jun-Lee, Yong-Hwan Lee, Hyochang Ahn, Sang Burm Rhee, “Color image descriptor using wavelet correlogram,” The 23rd international conference on Circuits/systems, computers and communication, 2008.

[17] Weibao Zou, and Yan Li, “Image classification using wavelet coefficients in low pass bands”, Proceedings of international joint conference on neural networks, Orlando, Florida, USA, aug 12-17, 2007.

[18] Elif Albuz, Erturk Kocalar, and Ashfaq A. Khokhar, “Scalable Color Image indexing and Retrieval Using Vector Wavelets”, IEEE Transaction on Knowledge and Data engineering, vol. 13, NO. 5, Sep/Oct 2001.

[19] N. G. Kingsbury and T. H. Reeves, "Redundant representation with complex wavelets: how to achieve sparsity", in Proc. Int. Conf. Image Processing, Barcelona, Sept. 2003.

[20] B. Wang, et al., “Video coding using 3-D dual-tree wavelet transforms”, in Proc. Int. Conf. on Acoustics, Speech, and Signal Processing, Philadelphia, Mar. 2005.

[21] N. G. Kingsbury, “Complex wavelets for shift invariant analysis and filtering of signals”, Applied Computational Harmonic Anal, vol. 10, no. 3, pp. 234-253, May 2001.

[22] G.Y. Chen, B. Kegl. Invariant Pattern Recognition using Dual-tree Complex Wavelets and Fourier Features, In: Conference on Image and Vision Computing (IVCNZ 05). Newzealand, 2005.

[23] H.B.Kekre, Ms Tanuja Sarode, Sudeep D. Thepade, “DCT Applied to Row Mean and Column Vectors in Fingerprint Identification”, In Proceedings of International Conference on Computer Networks and Security (ICCNS), 27-28 Sept. 2008, VIT, Pune..

[24] H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shaha, Prathmesh Verlekar, Suraj Shirke, “Image Retrieval using DCT on Row Mean, Column Mean and Both with Image Fragmentation”, (Selected), ACM-International Conference and Workshop on Emerging Trends in Technology (ICWET 2010), TCET, Mumbai, 26-27 Feb 2010, The paper will be uploaded on online ACM Portal.

[25] Guoping Qiu, “Colour Image Indexing Using BTC”, IEEE Transition on Image Processing, vol. No 12,.,Janauary 2003.

[26] Pdamshree Suresh,RMD Sundaram,Aravindhan Arumugam,” Feature Extraction in Compressed Domain for Content Based Image Retrieval ”,International Conference on Advanced Computer Theory and Engineering.2008.

[27] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, “Query by Image and Video Content: The QBIC System”, IEEE Computer, 28(9):23–32, Sept. 1995.

[28] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Content-based image retrieval at the end of the early years,” IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI), 22(12):1349–1380, Dec. 2000.