Review and Comparison of Various Feature Extraction Techniques in CBIR

Aditi Mehta
Mtech Student
Computer Engineering Department,MPSTME
2nd SVKM's NMIMS (Deemed-to-be University),
Mumbai, India

Paulami Shah
Assistant Professor
Computer Engineering Department,MPSTME
2nd SVKM's NMIMS (Deemed-to-be University),
Mumbai, India

ABSTRACT

In fields such as medical, art galleries, museums, archaeology, medical imaging, trademark databases, criminal investigations, images especially the digital images grow in quantities of thousands and sometimes even lakhs every year. Content based image retrieval is required from such large databases. This paper compares various CBIR techniques based on difference in generating feature vectors in the transform as well as the non-transform domain. Euclidean Distance is used for the purpose of similarity measure. Four performance evaluation parameters namely, precision, recall, LIRS and LSRR are used.

Keywords

Content Based Image Retrieval(CBIR); Discrete Cosine Transform (DCT); Discrete Sine Transform (DST); Walsh Transform; Row Mean(RM); Column Mean(CM); Row Column Mean (RCM); Forward Diagonal Mean (FDM); Backward Diagonal Mean (BDM); Forward Backward Diagonal Mean (FBDM); Fractional Coefficients; Kekre Transform; Euclidean distance; Precision; Recall; Length of Initial Relevant String of images(LIRS); Length of String required to Recover the Relevant images (LSRR).

1. INTRODUCTION

The very large numbers of images are being generated from a variety of sources (digital camera, digital video, scanner, the internet etc.) which have posed technical challenges to computer systems to store/transmit and index/manage image data effectively to make such collections easily accessible. Image compression deals with the challenge of storage and transmission, where significant advancements have been made [1,4,5]. The challenge to image indexing is studied in the context of image database [2,6,7], which has become one of the promising and important research area for researchers from a wide range of disciplines like computer vision, image processing , image database and recognition systems.

The thirst of better and faster image retrieval techniques is increasing day by day. The ambiguity in text based retrieval emphasizes the need of a better and faster retrieval system. That is why CBIR becomes more important. Some of important applications for CBIR technology could be identified as art galleries, museums, archaeology [3], architecture design [8,13], geographic information systems [5], weather forecast [5], medical imaging [5], trademark databases, criminal investigations, image search on the Internet.

2. CONTENT BASED IMAGE RETRIEVAL

In literature the term content based image retrieval (CBIR) has been used for the first time by Kato et.al. [4], to describe his experiments into automatic retrieval of images from a database by colour and shape feature. The typical CBIR system performs two major tasks. The first one is extraction of feature vector which consists of various feature components. It is generated to represent the content of each image in the database with accuracy and uniqueness. The second task is similarity measurement (SM), where a distance between the feature vector of the query image and the feature vector of each image in the database is measured, compared and this is used to retrieve the top "closest" images.

For feature extraction in CBIR there are mainly two approaches [5] feature extraction in spatial domain and feature extraction in transform domain. The feature extraction in spatial domain includes the CBIR techniques based on histograms [5], BTC [1,2]. The transform domain methods are widely used to extract image features. Many current CBIR systems use Euclidean distance [1-3] on the extracted feature set as a similarity measure. The Direct Euclidian Distance between image P and query image Q can be given as equation 1, where Vpi and Vqi are the feature vectors of image P and Query image Q respectively with size 'n'.

$$ED = \sqrt{\sum_{i=1}^{n} (V_{pi} - V_{qi})_2} \qquad (1)$$

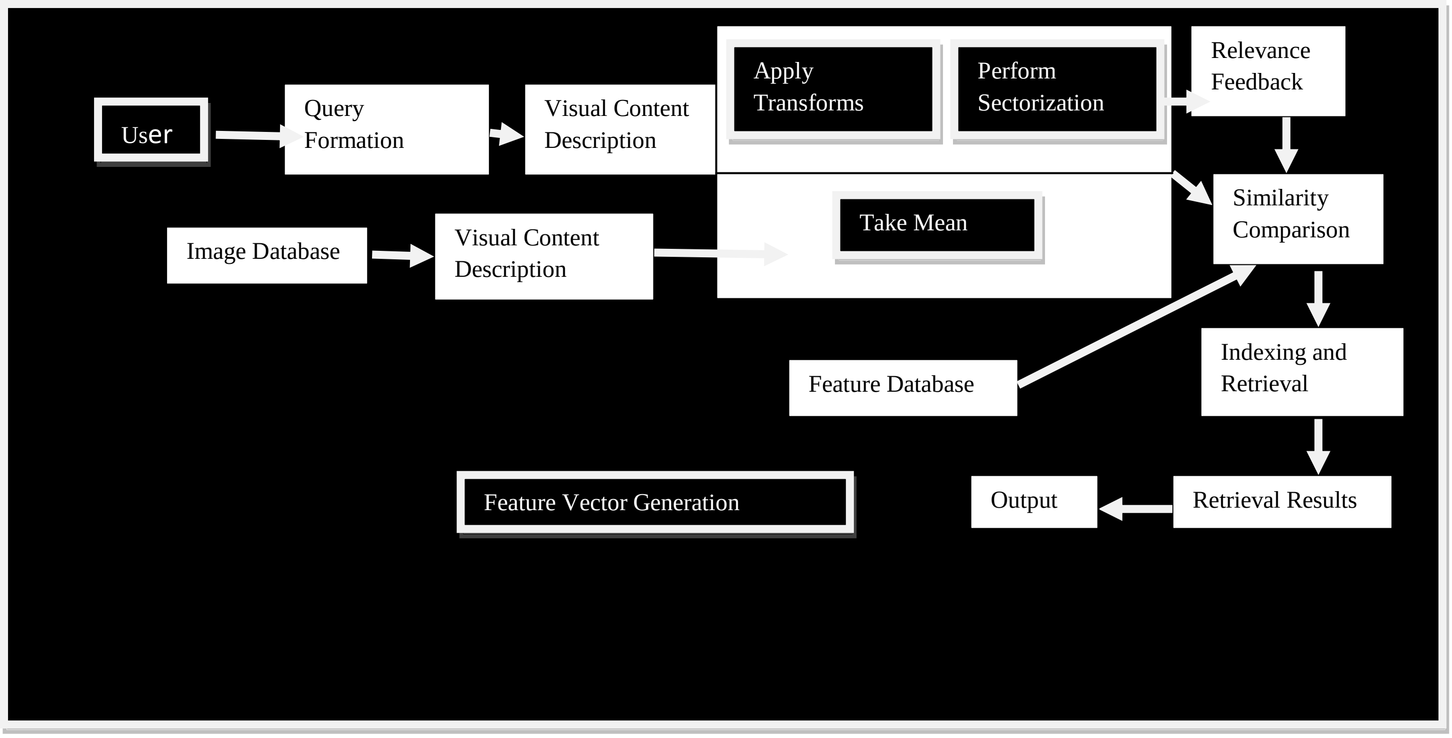

Figure 1: CBIR System

In a typical content-based image retrieval system as shown in figure above, the visual contents of the images in the database are extracted and described by feature vectors. The feature vectors of the images in the database form a feature database. To retrieve images, users provide the retrieval system with example images. The system then changes these examples into its internal representation of feature vectors. The similarities /distances between the feature vectors of the query example and those of the images in the database are then calculated and retrieval is performed. This provides an efficient way to search for the image database. Recent retrieval systems have also began with taking the feedback from users for making further improvements in the retrieval results.

3. COMPARISON OF TECHNIQUES

Various CBIR systems were studied and all of these systems discuss different techniques of feature vector generation. All of these techniques are discussed briefly in this section.

3.1 All Image Co efficient [11]

In this method all image pixels are considered as feature vector and Euclidean distance is used to find the best match. Precision and Recall has been calculated to check the retrieval performance of the system.

3.2 Row Mean of Image (RM) [11]

In this method, Feature Vector is generated by calculating row mean of all pixels in the row of image matrix for all three planes i.e. R,G,B. Euclidean distance is used as the similarity measure and the performance is evaluated using precision and recall.

3.3 Column Mean of Image (CM) [11]

In this method, Feature Vector is generated by calculating column mean of all pixels in the column of image matrix for all three planes i.e. R,G,B. Euclidean distance is used as the similarity measure and the performance is evaluated using precision and recall.

3.4 Row & Column Mean of Image (RCM) [11]

In this method, Feature Vector is generated by calculating row and column mean of all pixels in the column of image matrix for all three planes i.e. R,G,B. Euclidean distance is used as the similarity measure and the performance is evaluated using precision and recall.

3.5 Backward Diagonal Mean of Image (BDM) [11]

In this method, Feature Vector is generated by calculating backward diagonal mean of all pixels in the column of image matrix for all three planes i.e. R,G,B. Euclidean distance is used as the similarity measure and the performance is evaluated using precision and recall.

3.6 Forward Diagonal Mean of Image (FDM) [11]

In this method, Feature Vector is generated by calculating forward diagonal mean of all pixels in the column of image matrix for all three planes i.e. R,G,B. Euclidean distance is used as the similarity measure and the performance is evaluated using precision and recall.

3.7 Forward & Backward Diagonal Mean of Image (FBDM) [11]

In this method, Feature Vector is generated by calculating forward and backward diagonal mean of all pixels in the column of image matrix for all three planes i.e. R,G,B. Euclidean distance is used as the similarity measure and the performance is evaluated using precision and recall.

3.8 T-Gray [12]

'T' is the type of transform applied. It could be DCT or Walsh in our case.

In this method, the feature vector space of each image of size NxN has NxN number of elements. This is obtained using following steps of T-Gray

1. Extract Red, Green and Blue components of the color image.

2. Take average of Red, Green and Blue components of respective pixels to get gray image.

3. Apply the Transform 'T' on gray image to extract feature vector.

4. The result is stored as the complete feature vector 'T-Gray' for the respective image.

Thus the feature vector database for DCT and Walsh transform are generated as DCT-Gray and Walsh-Gray respectively.

3.9 T-RGB [12]

In this method, the feature vector space of each image of size NxNx3 has NxNx3 number of elements. This is obtained using following steps of T-RGB:

1. Extract Red, Green and Blue components of the color image.
2. Apply the Transform 'T' on individual color planes of image to extract feature vector.
3. The result is stored as the complete feature vector 'T-RGB' for the respective image.

Thus the feature vector database for DCT and Walsh transform are generated as DCT-RGB and Walsh-RGB respectively. Here the size of feature database is NxNx3.

Query Execution for 'T-RGB' CBIR and CBIR using 'Fractional-T-RGB' is similar as T-gray image as mentioned above.

3.10 Row Mean of Transformed Column Image [13]

This technique of CBIR can be explained using following steps:

1. Apply transform T on the column of image of size NxN (INxN) to get column transformed image of the same size (cINxN)

cINxN(column transformed) = [TNxN] [INxN]

2. Calculate row mean of column transformed image to get feature vector of size N (instead of N2)
3. The feature vector is considered with and without DC component to see variations in results. Then Euclidean Distance is applied to obtain precision and recall.

3.11 DCT Sectorization [8,14]

The rows in the discrete cosine transform matrix have a property of increasing sequency. Thus zeroeth and all other even rows have even sequencies whereas all odd rows have odd sequency. To form the feature vector plane the combination of co-efficient of consecutive odd and even co-efficient of every column is taken and even co-efficient is put on x axis and odd co-efficient on y axis of logical plane; thus taking these components as coordinates a point in x-y plane is obtained. An even-odd component plane for feature vector generation taking mean value of all the vectors in each sector with sum of absolute difference and Euclidean distance as similarity measures is taken. In addition to these; the feature vectors are augmented by adding two components which are the average value of zeroeth and the last row and column respectively. Thus for 4, 8, 12 & 16 DCT sectors 8, 16, 24 and 32 feature components along with augmentation of two extra components for each color planes i.e. R, G and B are generated. Thus all feature vectors are of dimension 30, 54, 72 and 102 components.

3.11.1 Four Sectors [8]

In this method, the DCT of the color image is calculated in all three R, G and B planes. The even rows/columns components of the image and the odd rows/columns components are checked for positive and negative signs and the values are assigned to respective sectors accordingly.

3.11.2 Eight Sectors [8]

In this method, each quadrant formed in the previous obtained 4 sectors are individually divided into 2 sectors each considering the angle of 45 degree. In total we form 8 sectors for R,G and B planes separately.

3.11.3 Twelve Sectors [8]

In this method, each quadrants formed in the previous section of 4 sectors are individually divided into 3 sectors each considering the angle of 30 degree. In total we form 12 sectors for R,G and B planes separately.

3.11.4 Sixteen Sectors [8]

In this method, sixteen sectors are obtained by dividing each one of eight sectors into two equal parts.

3.12 DCT-DST Sectorization [9,15]

This technique of CBIR can be explained using following steps:

1. Formation of the DCT-DST plane.
2. Sectorization of the DCT-DST Plane.
3. The average value of each segment of a particular sector were taken as feature vector components

To form the DCT-DST plane; the DCT co-efficient from column 2 to column N of a row wise DCT transformed image is extracted where N is number of columns in the image; and DST co-efficient from the column 1 to column N-1 of row wise DST transformed version of the image is extracted. The combination of these extracted co-efficient values are used to form the DCT-DST plane where DCT coefficient values are on X axis and DST co-efficient on Y axis. Now taking these components as coordinates, a point in X-Y plane is obtained. The average value of each segment of these sectors of all three planes i.e. R, G and B is considered to generate the feature vector. The average of co-efficient values of first column of DCT transformed image and the average magnitude of last column co-efficient values of DST transformed of the same image is then taken for the augmentation to the feature vectors. For 4 sectors, there are 8 feature components and 2 augmentation components per plane, i.e., 8+2 feature vectors per plane. Thus the feature components along with the augmentation of extra components for each color planes i.e. R, G and B generates feature vectors of dimension 30, 54, 72 and 102 components for 4,8,12 and 16 sectors respectively.

Sectorization here is done similarly as mentioned in the sections 3.14.1., 3.14.2., 3.14.3. and 3.14.4. respectively of Section 3.14.

3.13 Density Distribution in Walsh Transform [10,16]

The Walsh transform is considered by multiplying all sal functions by j and combining them with real cal functions of the same sequency. Thus it is possible to calculate the angle by taking tan inverse of sal/cal. However the values of tan are periodic with the period π radians and hence it can resolve these values in only two sectors. To get the angle in the range of 0-360 degrees, these points are divided into four sectors. These four sectors are further divided into 8, 12 and 16 sectors. The density distribution of sal and cal in each sector is considered for feature vector generation. Each Walsh sector is represented by single percentage value of sal-cal distribution in particular sector with respect to all sectors for feature vector generation; and is calculated as follows:

$$\frac{\text{(Total number of sal in particular sector)}}{\text{(Total number of sal in all sectors)}} \times 100 \quad (2)$$

In this technique, Sectorization is done similarly as mentioned in the sections 3.14.1., 3.14.2., 3.14.3. and 3.14.4. respectively.

Thus for 8 and 12 & 16 Walsh sectors 8 and 12 feature components for each color planes i.e. R, G and B are generated. In all feature vector is of dimension 24, 36 and 48 components.

4. IMPLEMENTATION

4.1 Databases

The CBIR techniques (3.1-3.13) are tested on two image databases. One is: Generic Image Database which contains 1000 variable size images spread across 11 categories of human being, animals, natural scenery and manmade things; Other is: Coil Image Database which consists of 1080 images of size 128x128x3. There are 15 different categories consisting of 72 images in each category. [11-13]

The CBIR techniques (3.14-3.16) are tested on the image database of 1055 variable size images spread across 12 different classes such as Flower, Sunset, Barbie, Tribal, Puppy, Cartoon, Elephant, Dinosaur, Bus, Parrots, Scenery, and Beach. [8-10]

4.2 Performance Evaluation Parameters

To assess the retrieval effectiveness, precision, recall, LIRS and LSRR are used as statistical comparison parameters for the proposed CBIR techniques. The standard definitions for these measures are given by following equations.

$$\text{Precision} = \frac{\text{Number of Relevant Images Retrieved}}{\text{Total Number of Images Retrieved}} \quad (3)$$

$$\text{Recall} = \frac{\text{Number of Relevant Images Retrieved}}{\text{Total Number of Relevant Images in Database}} \quad (4)$$

$$\text{LIRS} = \frac{\text{Length of Initial Relevant String of images}}{\text{Total Number of Relevant Images in Database}} \quad (5)$$

$$\text{LSRR} = \frac{\text{Length of string to recover all images}}{\text{Total Images in Database}} \quad (6)$$

4.3 Results

Figure 2 compares techniques (3.1-3.10) discussed in previous sections. It compares CBIR techniques in transform as well as non-transform domain. It is seen that RM-CD which is a non-transform domain gives the best result amongst both transform and the non-transform domain.

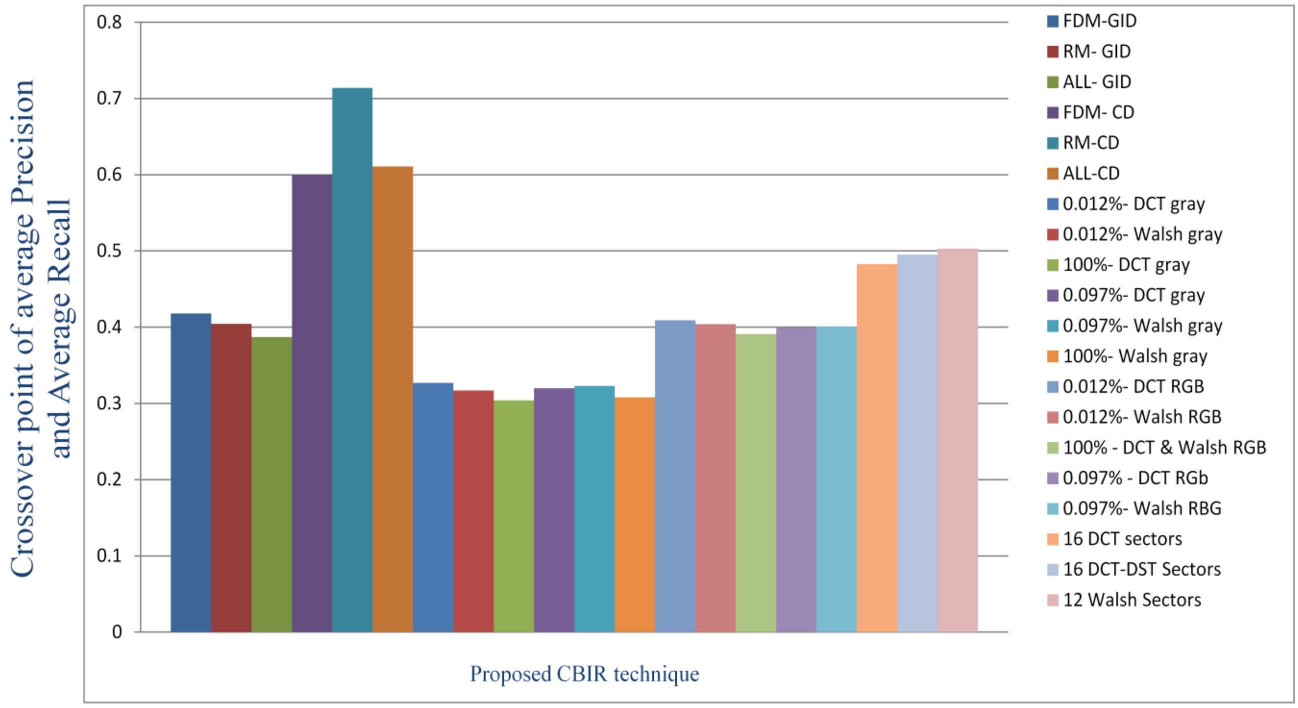

Figure 2: Graph of CBIR techniques (3.1-3.10) with highest precision & recall crossover values in transform & non transform domain with different methods for feature extraction.

Figure 3 compares techniques (3.11-3.13). It is seen that 12 sectors in Walsh transform gives the best result amongst all the transforms.

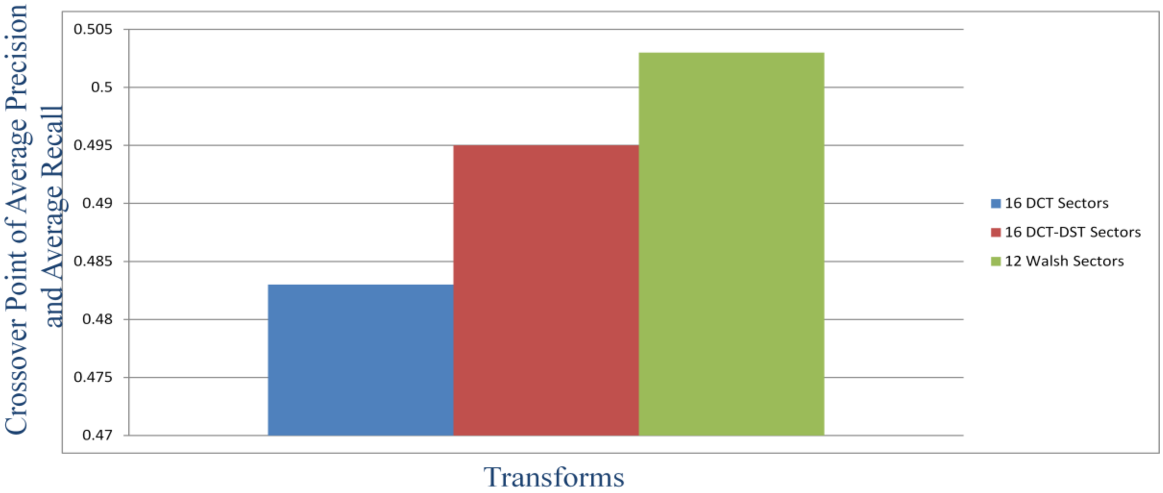

Figure 3: Graph of precision/recall crossover point versus the transform techniques (3.11-3.13).

5. CONCLUSION

5.1 Complexity

It can be seen from table of complexities [11, 12] that, when using all the co-efficient of the image for feature extraction; it has the maximum complexity i.e., maximum number of additions and multiplications. In the non-transform domain, RM/CM, RCM & FDM/BDM have approximately almost equal complexity. In the transform domain, Walsh has the least complexity and least number of computations and no additions at all. DCT however proves better in image compression since the values get concentrated on the upper left corner of the image and the rest of the pixel values can be ignored. As compared to the above techniques, full image taken as a feature vector has the highest complexity and number of computations, followed by DCT, Walsh and the non-transform techniques (RM, CM, RCM, etc).
techniques in the non-transform domain have minimum complexity and also provide the best results for image retrieval.

5.2 Space

Table 1: The transforms with their respective size of feature and augmentation components, image size for the respective sectorization.

Transforms	Sectors	Feature Components	Augmentation components	Size of each image (R,G,B) planes
DCT	4	8	2	8*3+2*3=30
DCT-DST	4	8	2	8*3+2*3=30
Walsh	4	4	-	4*3=12
DCT	8	16	2	16*3+2*3=54
DCT-DST	8	16	2	16*3+2*3=54
Walsh	8	8	-	8*3=24
DCT	12	24	2	24*3+2*3=78
DCT-DST	12	24	2	24*3+2*3=78
Walsh	12	12	-	12*3=36

From table 1, we can infer that a Walsh transformed image utilizes minimum space since it has no augmentation components and also the number of feature components doesn't double the size of the sector unlike the other transforms. Also, as the number of images increases, the space saved by Walsh proves to be better. Therefore, in large databases, Walsh transform is most suitable.
Thus, from the section 5.1 and 5.2, it can be seen that for CBIR, row mean in the non-transform domain and Walsh in the transform domain give the best results. And in the overall comparison of both the transform and the non-transform domain, Row Mean outperforms Walsh in terms of complexity as well as space. With respect to sectorization [8-10], optimal results were found at 16 sectors for DCT and DCT-DST; and at 12 sectors for Walsh Transform. Further Sectorization i.e., 32 sectors, 64 sectors and so on can also be done and is a matter of research.

6. ACKNOWLEDGMENTS

I would like to thank my mentor Mrs. Paulami Shah, my HOD, Dr. Dhirendra Mishra, and Dr. H. B. Kekre, for their constant support and motivation.

7. REFERENCES

[1] H.B.Kekre, Sudeep D. Thepade, "Boosting Block Truncation CodingusingKekre's LUV Color Space for Image Retrieval", WASETInternational Journal of Electrical, Computer and System Engineering (IJECSE), Volume 2, Number 3, pp. 172-180, Summer 2008.

[2] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using AugmentedBlock Truncation Coding Techniques", ACM International Conferenceon Advances in Computing, Communication and Control (ICAC3-2009), pp. 384-390, 23-24 Jan 2009, Fr. ConceicaoRodrigousCollegeofEngg., Mumbai. Is uploaded on online ACM portal.

[3] H.B.Kekre, Sudeep D. Thepade, "Scaling Invariant Fusion of ImagePieces in Panorama Making and Novel Image Blending Technique",International Journal on Imaging (IJI), www.ceser.res.in/iji.html,Volume 1, No. A08, pp. 31-46, Autumn 2008.

[4] Hirata K. and Kato T. "Query by visual example – content-based image retrieval", In Proc. of Third International Conference on H Extending Database Technology, EDBT'92, 1992, pp 56-71

[5] H.B.Kekre, Sudeep D. Thepade, "Rendering Futuristic Image RetrievalSystem", National Conference on Enhancements in Computer,Communication and Information Technology, EC2IT-2009, 20-21 Mar2009, K.J.Somaiya College of Engineering, Vidyavihar, Mumbai-77

[6] Minh N. Do, Martin Vetterli, "Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance", IEEE Transactions On Image Processing, Volume 11, Number 2, pp.146-158, February 2002.

[7] B.G.Prasad, K.K. Biswas, and S. K. Gupta, "Region – based image retrieval using integrated color, shape, and location index", International Journal on Computer Vision and Image Understanding Special Issue: Colour for Image Indexing and Retrieval, Volume 94, Issues 1-3, April- June 2004, pp.193-233.

[8] Dr.H.B.Kekre, Dr.Dhirendra Mishra, " DCT Sectorization for Feature Vector Generation in CBIR",International Journal of Computer Applications (IJCA) Vol.9(1) November 2010, ISSN 0975–8887 available online at http://www.ijcaonline.org/volume9/number1/pxc3871820.pdf

[9] Dr.H.B.Kekre, Dr.Dhirendra Mishra, "DCT-DST Plane sectorization of Row-wise Transformed color Images in CBIR",International Journal of Engineering Science and Technology (IJEST) Vol.2(12) 2010, ISSN 7234-7244available online at http://nmims.edu/wp-content/uploads/2012/p3/MPSTME/Direndra,DCT-DSTPlanesectorization.pdf

[10] Dr.H.B.Kekre, Dhirendra Mishra, " Density distribution in WalshTransform sectors as feature vectors for image retrieval", published in international journal of compute applications (IJCA) Vol.4(6) 2010, 30-36 ISSN 0975-8887 available online at http://www.ijcaonline.org/archives/volume4/number6/829-1072

[11] Dr.H.B.Kekre, SudeepDThepade, AkshayMaloo, "Query by Image Content Using Colour Averaging Techniques", International Journal of Engineering Science and Technology (IJEST), Volume 2, Issue 6, 2010.pp.1612-1622 (ISSN: 0975-5462) Available online at http://www.ijest.info

[12] Dr.H.B.Kekre, SudeepDThepade, AkshayMaloo, "Image Retrieval using Fractional Coefficients of Transformed Image using DCT and Walsh Transform", International Journal of Engineering Science and Technology (IJEST), Volume 2, Issue 4, 2010.pp.362-371 (ISSN: 0975-5462) Available online at http://www.ijest.info

[13] Dr.H.B.Kekre, SudeepDThepade, AkshayMaloo, "Performance Comparison of Image Retrieval using Row Mean of Transformed Column Image", International Journal of Engineering Science and Technology (IJEST), Volume 2, Issue 5, 2010.pp.1908-1912 (ISSN: 0975-3397) Available online at http://www.ijest.info

[14] Dr.H.B.Kekre, Dhirendra Mishra, Chirag Thakkar, "Column wise DCT plane sectorization in CBIR", International Journal of Computer Science and Information Technologies, Volume 3 Issue 1 , 2012,(ISSN: 3229-3235)

[15] Dr.H.B.Kekre, Dhirendra Mishra, "Discrete Sine Transform Sectorization for Feature Vector Generation in CBIR", Universal Journal of Computer Science and Engineering Technology , Volume 1 Issue 1, Oct. 2010. (ISSN: 2219-2158).

[16] Dr.H.B.Kekre, Dhirendra Mishra, "Sectorization of Full Walsh Transform for Feature Vector Generation in CBIR", International Journal of Computer Theory and Engineering, Volume 3, Issue 2, April 2011 (ISSN: 1793-8201).