Image Retrieval using DST and DST Wavelet Sectorization

Dr. H.B.Kekre  
Sr. Professor, Computer Engineering Dept.  
MPSTME, SVKM’s NMIMS (Deemed-to be University)  
Vile Parle West, Mumbai, INDIA

Abstract—The concept of sectorization of the transformed images for CBIR is an innovative idea. This paper introduces the concept of Wavelet generation for Discrete sine transform (DST). The sectorization of the DST transformed images and the DST wavelet transforms has been done into various sector sizes i.e., 4, 8, 12, and 16. The transformation of the images is tried and tested threefold i.e. row wise transformation, column wise transformation and Full transformation. We have formed two planes i.e. plane 1 and plane 2 for sectorization in full transformation. The performance of the all the approaches has been tested by means of the three plots namely average precision-recall cross over point plot, LIRS (Length of initial relevant string of images) plot, LSRR (Length of string to recover all relevant images in the database) plot. The algorithms are analyzed to check the effect of three parameters on the retrieval: First the way of transformation (Row, Column, Full), Second the size of sector generated, Third the type of similarity measures used. With the consideration of all these the overall comparison has been performed.

Keywords-CBIR, Feature extraction; Precision; Recall; LIRS; LSRR; DST; DST Wavelet.

I. INTRODUCTION

The digital world innovations has evolved itself to a very large extent. The result of which has increased the more and more dependency on the digital data and in turn on computer system. The information of any form i.e. multimedia, documents, images etc. everything has got its own place in this digital world. The computer system has been accepted to be the very powerful mechanism to use these digital data, for its secured storage and efficient accessibility whenever required.

Digital Images play a very good role for describing the detailed information about man, money, machine etc. almost in every field. The various processes of digitizing the images to obtain it in the best quality for the more clear and accurate information leads to the requirement of more storage space and better storage and accessing mechanism in the form of hardware or software. As far as the accessing of these images are concerned one needs to have the good mechanism of not only for accessing of the images but also for any other image processing to be done one needs to have the faster, accurate, efficient retrievals of these images. There are various approaches of proposing the methodologies of retrieving the images from the large databases consisting of millions of images stored.

Content Based Image Retrieval (CBIR) [1-4] is one of the evolving fields of image processing. CBIR needs to have the innovative algorithm to extract the perfect features to define identity of an image. It has been researched upon to use content of the image itself to draw out its unique identity. This unique identity can make one to differentiate the images with each other with better and accurate retrieval of images. There are mainly three contents i.e. shape, color and textures of the image as of now is being experimented by many researchers. These contents leads one to extract the exact feature of the image which can be well utilized to compare with all images available in the database by means of some similarity measures like Euclidean distance, sum of absolute difference etc. The tremendous use of images in the digital world of today has proved the CBIR as very useful in several applications like Finger print recognition, Iris Recognition, face recognition, palm print recognition, speaker identification, pattern matching and recognition etc.

There are various approaches which have been experimented to generate the efficient algorithm for image feature extraction in CBIR. These approaches advocate different ways of extracting features of the images to improve the result in the form of better match of the query image in the large database. Some papers discuss the variation in the similarity measures in order to have its lesser complexity and better match [5-15]. Methods of feature extraction using Vector Quantization [16], bit truncation coding [17,18], Walsh Transform [20,21] has also provided the new horizon to the feature extraction methodology. The method of sectorization has already been experimented on DCT [22], DST [23], DCT-DST Plane [24], Haar Wavelet [25] and Kekre’s Transform [26] earlier.

This paper proposes the use of sectorization of DST Wavelet for feature extraction in CBIR. The outcome of which has been compared with the DST sectorization performance.

II. DST WAVELET

A. Generation of DST Wavelet[4]

The wavelet analysis procedure is to adopt a wavelet prototype function, called an analyzing wave or mother wave. Other wavelets are produced by translation and contraction of the mother wave. By contraction and translation infinite set of functions can be generated. This set of functions must be orthogonal and this condition qualifies a transform to be a
wavelet transform. Thus there are only few functions which satisfy this condition of orthogonality. Generation of DST Wavelet transform matrix of size $N^2 \times N^2$ from DST matrix of size $NxN$ is given in [4]. However in this case we require DST matrix of size 128x128 where 128 is not a square. To resolve this situation, this paper proposes an algorithm to generate discrete sine wavelet transform from discrete sine transform of 8x8 and 16x16.

In this paper the DST Wavelet has been generated using the contraction and translation. Due to the size of images in the database is 128x128 we need the wavelet transform to be of size 128x128. The 128x128 Wavelet transform matrix generated from 16x16 orthogonal DST matrix and 8x8 DST matrix. First 16 rows of Wavelet transform matrix are generated by repeating every column of DST matrix of dimension 16x16, 8 times. To generate next 17 to 32 rows, second row of DST (8x8) is translated using groups of 8 columns with horizontal and downward shifts. To generate next 33 to 48 rows, third row of DST (8x8) matrix is used in the same manner. Like wise to generate last 113 to 128 rows, 8th row of transform DST (8x8) matrix is used. Note that by repeating every column of the basic transform 8 times we get global components. Other wavelets are generated by using rows of DST matrix of size 8x8 giving local components of the DST Wavelet.

III. FEATURE VECTOR GENERATION

The figure given below shows the formal steps of feature vector generation in brief.

A. Sectorization [7-13],[23-28]

The individual components of row and column wise transformed (DST and DST Wavelet separately) images are distributed into different co-ordinates of Cartesian coordinate system according to their sign change to form four sectors. The even rows/columns components of the transformed image and the odd rows/columns components of the transformed images are checked for positive and negative signs. The even and odd DST values are assigned to each quadrant.

The division of each of these 4 sectors into 2 partitions forms the 8 sectors which distributes the transformed image components into its appropriate sectors.

Continuing the same division concept further the 4 sectors already generated has been divided into 3 parts with consideration of 30 degree angle to generate sector sizes of 12.

Each sector of 8 sectors are individually divided into two to obtain 16 sectors. For each individual sector sizes the mean of each sectors are taken as the feature vector component. The feature vector components of each plane i.e. R, G and B has been calculated and concatenated together with the average of first row/column and last row/column to form the final feature vector for each method separately. The size of the feature vector and its component varies for each sector sizes the maximum sector size will have maximum feature vector components.

The feature database consists of feature vectors of all images. The features of the query image extracted are compared with the feature database using similarity measures. There are two similarity measures used in this experiment i.e. Euclidean distance (ED) and sum of absolute difference (AD) as given in the equation (1) and (2) shown below:

$$ED(P_,Q_) = \sum_{i=1}^{N} (Pi - Qj)^2$$  
$$AD(P_,Q_) = \sum_{i=1}^{N} |Pi - Qj|$$

The match of query image feature with the feature database with the minimum value of ED/AD gives the perfect match.

The performance measure of the algorithms proposed are done with the calculation of the precision-recall and LIRS (Length of initial relevant string of retrieval) and LSRR (Length of string to recover all relevant images in the database) refer equations (3) to (6).

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total Number of images retrieved}}$$  
$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in database}}$$  
$$\text{LIRS} = \frac{\text{Length of initial relevant string of images}}{\text{Total relevant images in the Database}}$$  
$$\text{LSRR} = \frac{\text{Length of string to recover all relevant images}}{\text{Total images in the Database}}$$

The performance of the proposed methods are checked by means of calculating the class wise average performance and overall average performance of each approach with respect to the transformation method applied, the way of applying the transformation i.e. row wise, column wise, full, sector sizes used and type of similarity measures used.

IV. EXPERIMENTAL RESULTS

A. Image Database

The sample Images of the augmented Wang database [29] consists of 1055 images having 12 different classes such as Cartoons, Flowers, Elephants, Barbie, Mountains, Horses,
Buses, Sunset, Tribal, Beaches, Monuments and Dinosaur shown in the Figure 2.

The class wise distribution of all images in its respective classes such as ther are 46 images of cartoon in the class, there are 100 images for flower and so on are shown in the Figure 3 below.

B. Sectorization of Row wise Transformed (DST) images.

The class wise precision-recall cross over point plot for sectorization of row wise DST transformed images has been shown in the Figure 4. The x axis of the plot denotes the class of images in the database and the y axis denotes the average precision-recall cross over point plots of five randomly selected images per class. These values are taken into percentage. Comparing all classes of images it is found that the dinosaur class has the best retrieval result of more than 80% for 12 and 16 sectors with sum of absolute difference as similarity measure. Other classes like flower, horse, sunset and elephants have the retrieval up to 68% (12 and 16 sectors with AD), 55% (4,8,12 sectors with ED), more than 50% (16 sectors with AD) and more than 50% (12 sectors with AD) respectively. Looking at the LIRS plot for the same (as shown in the Figure 5) to check the performances it has been found that the initial length of string containing the relevant images must be more. In this case the higher value of LIRS is very clearly visible for the best performer class i.e. dinosaur class. The LIRS varies for all other classes from 1% to 20% which indicates that the first image is always relevant image. The Figure 6 shows the LSRR Plot checks for the length of string containing all relevant images retrieved. It must be minimum which is very much seen for dinosaur class.

C. Sectorization of Row wise Transformed (DST Wavelet) images

This section of the paper discusses the performance of the DST Wavelet (row wise) sectorization with respect to all three performance measuring parameters namely average precision-recall cross over point plot (see Figure7), LIRS(see Figure8) and LSRR(see Figure9). Comparing the performance of all the classes within once again the dinosaur class outperforms all classes with 90% retrieval for 12 sectors with sum of absolute difference as similarity measure. This performance is better than the simple DST (row wise) sectorization as discussed in section 4.2. The flower class has retrieval of more than 65% in DST Wavelet whereas it is below 60% in the case of DST. Horse class has the result of 50%,The cartoon class has improved a lot goes up to 60% compared to only 45% in normal DST sectorization. The LIRS and LSRR shows its relevant plots in the Figure8 and Figure9.

D. Sectorization of Column wise Transformation (DST)

There are 12 classes of the images used in the database. The performance of the algorithm varies from class to class. The class wise average precision-recall cross over point plotted in the Figure10 shows the performance of the sectorization of DST (column wise) in various sectors.
Figure 7. Class wise Average Precision-Recall cross over point plot of DST Wavelet row wise sectorization for all sector sizes with respect to similarity measures i.e. Euclidian distance (ED) and sum of absolute difference (AD).

Figure 8. Class wise LIRS plot of DST Wavelet row wise sectorization for all sector sizes with respect to similarity measures i.e. Euclidian distance (ED) and sum of absolute difference (AD).

Figure 9. Class wise LSRR plot of DST Wavelet row wise sectorization for all sector sizes with respect to similarity measures i.e. Euclidian distance (ED) and sum of absolute difference (AD).

Looking at the class wise performance the Dinasaur class has the best retrieval close to 80%. Flower class reaches up to 70% (16 sectors with AD), sunset and horses class has the retrieval more than 50%.

The elephant class has the resultant retrieval rate close to 50% whereas the performance for the Barbie class is more than 40%. The LIRS and LSRR which keeps check on the performance evaluation of the method shown in the Figure 11 and Figure 12. The maximum f LIRS has been achieved for the dinosaur class for the combination of all sectors and the sum of absolute difference.
E. Sectorization of Column wise Transformation (DST Wavelet).

The sectorization of DST Wavelet has better performance of average precision-recall cross over points as it can be easily seen in the Figure 13. There is increase in the performance of the Dinosaur, Flowers, Cartoon, Barbie classes i.e. 85% (12 and 16 sector with AD), 75% (16 sectors with ED), 42% (16 sectors with AD), 50% (16 sectors with AD) respectively.

The LIRS performances of the method are very interestingly depicts the high rises for the better performance of precision-recall. The minimum value of LSRR for the dinosaur, flower, sunset classes depicts the good performance.

F. Overall Comparison of all Approaches.

The overall performance of all approaches gives very clear idea about the overall average performance of the retrieval rates as shown in the Figure 16 – Figure 18. The overall average precision-recall cross over point plot as shown in the Figure 16 depicts that on average performance of the retrieval for all methods proposed is 40%. It is observed that in most of the cases the sectorization with the sum of absolute difference as similarity measure has better retrieval than Euclidian distance. As far as sector sizes are concerned all have good performance except for 8 sectors in DST-WT (Row).
This paper discusses the new idea of generating the DST wavelet transform and sectoring the transformed image into 4, 8, 12 and 16 sectors. The results obtained clearly show that the performance of DST wavelet transform is better than DST using sectorization of full-Plane1 and full-plane2 transformations with the sum of absolute difference giving best results close to 45% (see the Figure 16). The similarity measure plays vital role in the applications of CBIR. Similarity measures provide the efficient and faster access and matching of the images in the database. We have employed and analyzed two similarity measures i.e. Euclidian Distance (ED) and sum of absolute difference (AD). The sum of absolute difference as shown in the equation (2) has lesser complexity of calculation and better retrieval rates in the approaches applied for all sector sizes (see figure 16). The performance measurement of all approaches has been checked by means of LIRS and LSRR which are very powerful tools to comment on the retrieval rate of any algorithm. The maximum LIRS and the Minimum LSRR are the best measures for it.

**REFERENCES**


AUTHORS PROFILE

H. B. Kekre has received B.E. (Hons.) in Telecomm. Engg. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D.(System Identification) from IIT Bombay in 1970. He has worked Over 35 years as Faculty and H.O.D. Computer science and Engg. At IIT Bombay, From last 13 years working as a professor in Dept. of Computer Engg. at Thadomal Shahani Engg. College, Mumbai. He is currently senior Professor working with Mukesh Patel School of Technology Management and Engineering, SVKM’s NMIMS University vile parle west Mumbai. He has guided 17 Ph.D.s 150 M.E./M.Tech Projects and several B.E./B.Tech Projects. His areas of interest are Digital signal processing, Image Processing and computer networking. He has more than 350 papers in National/International Conferences/Journals to his credit. Recently twelve students working under his guidance have received the best paper awards. Two research scholars working under his guidance have been awarded Ph. D. degree by NMIMS University. Currently he is guiding 10 PhD. Students. He is life member of ISTE and Fellow of IETE.

Dhirendra Mishra has received his BE (Computer Engg) degree from University of Mumbai. He completed his M.E. (Computer Engg) from Thadomal shahani Engg. College, Mumbai, University of Mumbai. He is PhD Research Scholar and working as Associate Professor in Computer Engineering department of, Mukesh Patel School of Technology Management and Engineering, SVKM’s NMIMS University, Mumbai, INDIA. He is life member of Indian Society of Technical education (ISTE), Member of International association of computer science and information technology (IACSIT), Singapore, Member of International association of Engineers (IAENG). He has more than 30 papers in National/International conferences and journals to his credit. His areas of interests are Image Processing, Operating system, Information Storage and Management.