Performance Comparison Between DWT and CT Transforms for CBIR of Face Images

Dr. W.S. Besbas *Dr. M. R. Sunni* *Dr. R.M, Salman*
Computer Department- Faculty of science,
Zawia University

Abstract:

Applications of Content Based Image Retrieval (CBIR) are wide, such as: Crime Prevention, Security Check, Medical Diagnosis, Intellectual Property. This paper investigates the performance of CBIR of face images in transform domain. Face images that are similar to the query image are retrieved from the face database. Features of the face image are extracted in the spectrum domain of Discrete Wavelet Transform (DWT), and Contourlet Transform (CT). For the performance analysis of features selection methods, two face images databases are used. These are "*Indian face Database*", and "***G****eorgia* ***T****ech (GT) face database*". The City block distance measure

is analyzed to evaluate the performance of the retrieval process. The investigation concludes that the retrieval rate is database dependent. But in general, the DWT transform is better than CT transform.

Keywords: Face image retrieval, Content Based Image Retrieval (CBIR), Image retrieval in transform domain, Features selection for CBIR, Discrete Wavelet Transform (DWT), Contourlet Transform (CT).

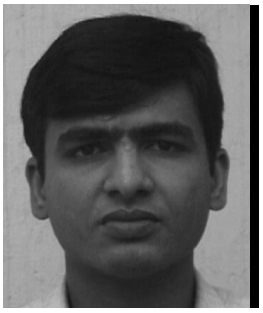
1. Introduction:

Face image retrieval is a process for finding a predefined number of images in a database that are similar to the query face image. Face image retrieval techniques can be used in applications such as *Crime prevention*, *Security Check*, etc. [1].

There are many ways of specifying what kind of images a user wishes to retrieve from the image database. C*ategory browsing*, *query by concept*, *query by sketch*, and *query by example* are commonly used query formations [1]. The first two types of queries (text based) are related to the semantic description of images. The other two types of queries (content based) are to provide an example image or draw a sketch from which images with similar visual features will be extracted from the database.

In query by example the system converts the example image provided by the user into an internal representation of features. And then images, with similar features, stored in the database are searched. Query by example can be further classified into query by external image example, if the query image is not in the database, and query by internal image example, if otherwise. In query by sketch the user draws a sketch of an image with a graphic editing tool, or free hand drawing, and then scanned. In most cases, a coarse sketch is sufficient, as the query can be refined based on retrieval results. Figure1 shows

an example face image and face sketch that can be used as query images.



Face image **Face sketch image**

Figure1. Query face images.

Content Based Image Retrieval (CBIR) methods can be assigned to one of two major approaches, spatial or transform domain techniques. Spatial domain techniques are mostly based on color, shape, or texture features that are extracted directly from images [2]. Transform domain methods utilize global information from images to perform image retrieval. The global information of an image is fundamentally represented by a small number of features derived from the spectrum of the image after transforming it from the special domain to the transform domain, using transforms such as Discrete Cosine Transform (DCT) [3,4], Walsh Hadamard Transform (WHT) [5], Discrete Wavelet Transform (DWT) [6,7], and Contourlet Transform (CT) [8], and. Different sizes of the feature vectors are analyzed, using City block and Euclidean distance measures.

This research is intended to analyse the performance of face image retrieval in DWT and CT transform domain. The proposal idea comes from the problem of searching for digital face images in a large face database. The content based face image retrieval system is illustrated below.

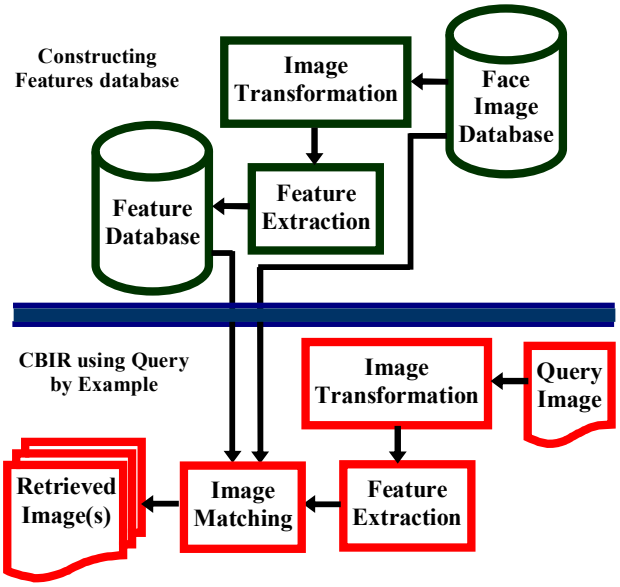


Figure2: Illustration of CBIR using Query by Example

2. Content Based Image Retrieval

The two main steps of Content Based Image Retrieval are:

- **Feature Extraction:** Extracting image features to a distinguishable extent.
- **Matching:** The second step involves matching these features to yield a result that is visually similar.

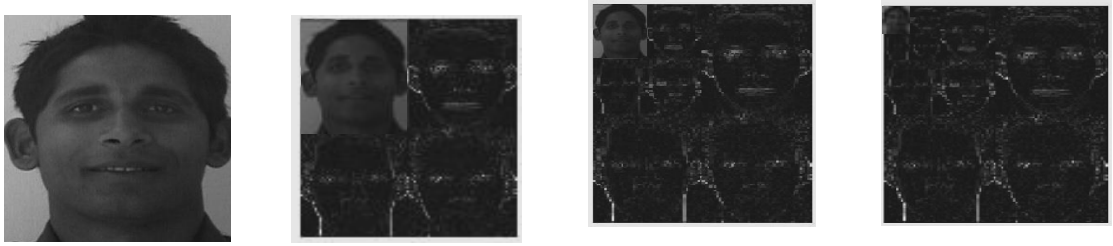
2.1. Constructing the Feature Vector

To investigate face image retrieval, two features selection methods are utilized using DWT and CT.

2.1.1 Discrete Wavelet Transform (DWT)

DWT is widely used for multi-scale image analysis [6]. It decomposes an image into four sub-bands: an approximated image, and horizontal (D^h), vertical (D^v), and diagonal (D^d) detailed images. The detailed images measure variations along the columns (horizontal edges), rows (vertical edges), and diagonals (diagonal edges) respectively.

More than one decomposition level may be utilized for face recognition task to give reduced but meaningful information describing face image [6]. The approximated image is decomposed again to wavelet sub-bands. Two or three decomposition levels may be used. The final resultant approximated image is used as a feature vector. Figure3 (a, b, c, d) shows an original image, level one, level two, and level three of decomposition respectively.



a: Original image | b: Level one of decomposition | c: Level two of decomposition | d: Level three of decomposition

Figure3: Wavelet Decomposition

2.1.2 Contourlet Transform

Contourlet Transform (CT) addresses the problem of representing the images with smooth contours in different directions, by providing two additional properties which are directionality and anisotropy, comparing with wavelets [8].

Directionality: The representation should contain basis elements oriented at a variety of directions, much more than the few directions that are offered by wavelets.

Anisotropy: To capture smooth contours in images, the representation should contain basis elements using a variety of elongated shapes with different aspect ratios.

The features (Edge and texture orientations) are captured by using CT decomposition with a 4 level LP decomposition. At each level, the

number of directional sub-bands are (3, 4, 8 and 16) respectively. For LP decomposition and directional sub-band decomposition the 'pkva' filters are used. These parameters results in a 32 dimensional feature vector (n=32). The normalized standard deviation (SD), which is computed on each directional sub-band of the CT decomposed image, is used to form the image feature vector. These normalized feature vectors are used for the creation of the features database. Figure5 shows the CT decomposition of a given image.

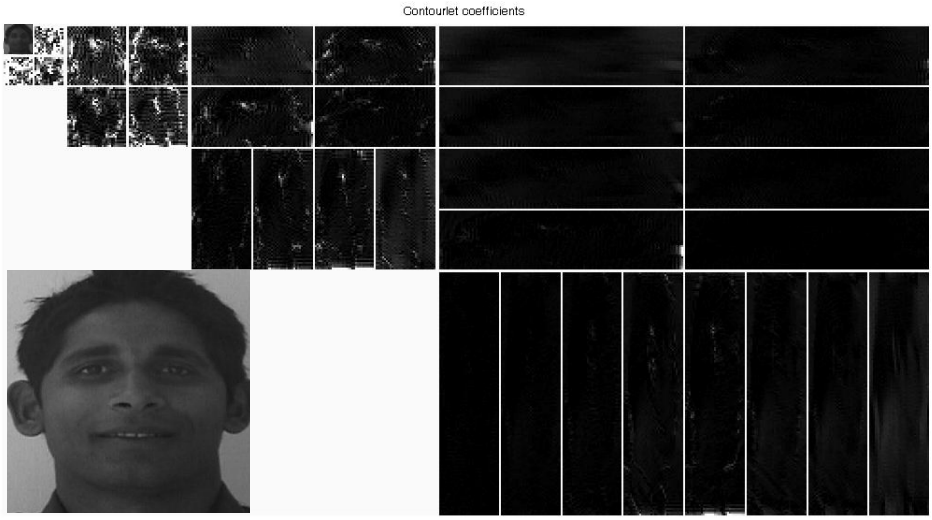


Figure4. Original and CT decomposed image (Using 4-level LP & 'pkva' filter)

2.2. Distance Measure:

The most common method for comparing two images in content based image retrieval is by using an image distance measure, which compares the similarity of two images in various dimensions such as color, texture, shape, etc..

The Euclidean and City block distance measures are most common measures which are used to compare the similarity of any two images.

Here in our research City block distance is implemented, and results are analyzed.

The distance measures are always greater than or equal to zero. The measurement would be zero for identical points and increases as the points show dissimilarity.

Euclidean Distance: It is the most common used distance measure. Euclidean distance or simply 'distance' examines the *root of square differences* between coordinates of a pair of objects. The general formula for this measure is given below:

$$d(x \quad , \quad y) \quad = \quad \sqrt{\sum_{i=1}^{n}(x_i \quad - \quad y_i)^2}$$

City block Distance: It is also known as *Manhattan*, *boxcar*, and *absolute value* distance. It represents the distance between points in a city road grid. It examines the *absolute differences* between coordinates of a pair of objects. The general formula for this measure is given below:

$$d(x \quad , \quad y) \quad = \quad \sum_{i=1}^{n}|x_i \quad - \quad y_i|$$

2.3. Performance Measure:

The most common method for measuring the performance in the field of CBIR is the "Percentage of the average retrieval rate". In some researches the "Recall and Precision" measure is used.

In our investigation the "Percentage of the average retrieval rate" is used to enable the comparison between the obtained results. The following equation shows how this measure is obtained.

$$\textit{Correctly retrieved images \%} = \frac{\textit{Number of relevant images retrieved}}{\textit{Total number of images retrieved}} \times 100$$

Recall and Precision: They are defined in terms of a set of retrieved images, such as the list of face images retrieved, and a set of the relevant images in the database (the list of all images on the database that belong to the same person). The standard definitions of these two measures are given by following equations.

$$Precision = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved}$$

$$Recall = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ relevant\ images\ in\ the\ database}$$

3. Images Databases:

For the performance analyses of the features selection methods two face images databases are used. They are "***G**eorgia **T**ech (GT) face database*"[9], and "*Indian face Database*"[10]. The images size is 640 x 480 pixels. All images in these databases are converted to gray scale and resized to 256 x 256 pixels as a preprocessing step.

The ***Georgia Tech*** (GT) face database consists of 750 colorful face images with 50 subjects and 15 images per subject. The ***Indian Face Database*** contains 440 face images of 40 people with eleven different poses for most of the individuals, with the exception that some of individuals have 9, 10, or 12 poses. The images size is 640 x 480 pixels. Some samples from these databases are shown next.

Figure5a: Sample images from GT Face database.

Figure5b: Sample images from Indian Face database.

4. Experimental Results:

In this section the selected face databases are used to demonstrate the performance of the CBIR system. The first database is used to compare our results with an existing result, given in [8], the second database is downloaded to compare, mainly, the performance of the CT with DWT transforms. For each subject (a class of *N* images, for the same person, in the database), every image is considered as query image and the algorithm retrieves *N/5="20%", N/3="33.3%", ……, and N="100%"* images that have the smallest distance between their vector of coefficients and the vector of coefficients for the query image. The average retrieval rate is then calculated.

4.1 CBIR from Indian Faces Database:

Table1 and Figure6 clearly show that the DWT produces the best retrieval rate comparing with CT transforms. The percentage of retrieving the same number of images, as the images of the class, using third level DWT approximation is 52.19, and 32.84 using CT.

Table1: CBIR - Indian faces database

Methods	Retrieving a percentage of number of top matches equivalent to the number of images in the query image class				
	20% 3 of 15	33.3% 5 of 15	53.3% 8 of 15	66.6% 10 of 15	100% 15 of 15
CT (Proposed in [8])	74.44	56.46	46.21	42.18	32.84
DWT 3rd Level	89.64	78.30	69.66	64.73	52.19

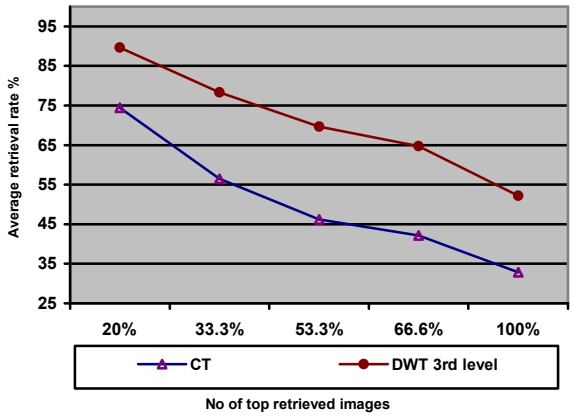


Figure6: CBIR – using Indian faces database

4.2 CBIR from Georgia Tech (GT) Faces Database:

This database is chosen for the purpose of comparison between the DWT and the results taken from [8]. The obtained results are given in table2 and Figure7. The CT transform results, for the GT database, are taken from [8]. The results show that the third level of decomposition DWT gives better retrieval rate than CT, when retrieving a number of images equivalent to the number of the images in the class of the query image. It is also found that, even the forth level of decomposition gives better retrieval rate than CT. On the other hand, when retrieving a subset of the class, the CT transform gives better results than DWT.

Table2: CBIR – using GT faces database

Methods	Retrieving a percentage of number of top matches equivalent to the number of images in the query image class				
	20% 3 of 15	33.3% 5 of 15	53.3% 8 of 15	66.6% 10 of 15	100% 15 of 15
Results from [8]					
GF	98.81	96.71	90.27	84.75	68.17
ZM	98.66	95.37	88.50	82.71	66.31
GF+ZM	98.96	96.88	90.22	85.11	67.44
CT (Euclidean distance)	98.47	96.12	90.62	85.17	68.97
CT (Manhattan distance) PROPOSED	99.99	99.25	98.03	96.78	72.65
DWT 3rd Level	99.47	97.92	92.35	87.35	**73.33**
DWT 4th Level	99.56	97.87	92.23	87.19	**73.27**

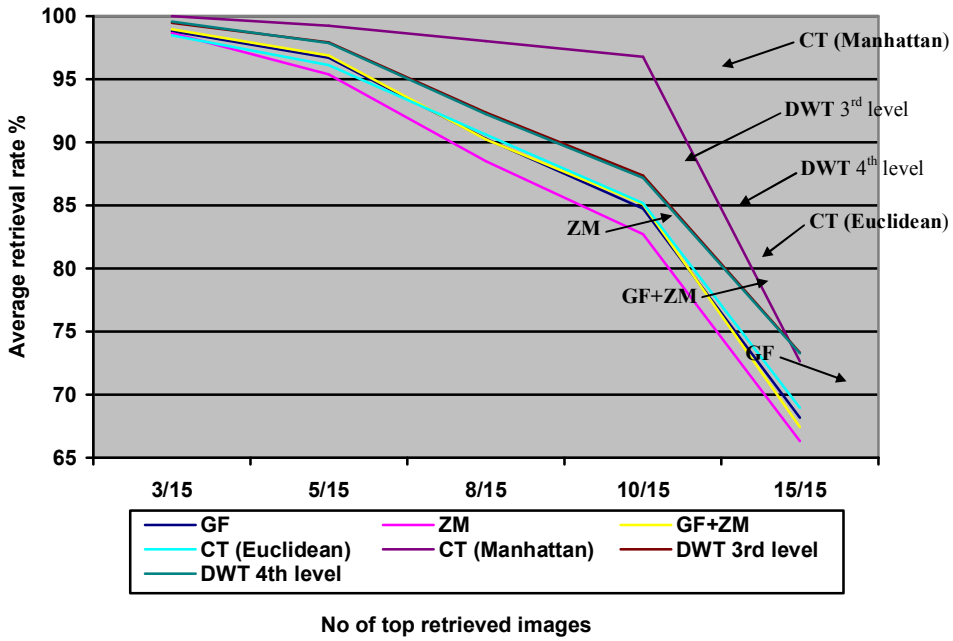


Figure7: CBIR – using GT faces database

An example of the retrieved images from the Indian face database is given in Figure8. The original images in the class of the query image are shown in Figure8.a. The retrieved images using DWT (third level approximation) and CT transforms are shown in Figure8 (b, c) respectively. The image in the top left corner is an exact copy of the query image. The query face image has 11 images in its class. The 11 images with the smallest distance between their vector of coefficients and the vector of coefficients for the query image are retrieved.

Figure8.a- A query image and all images in its class.

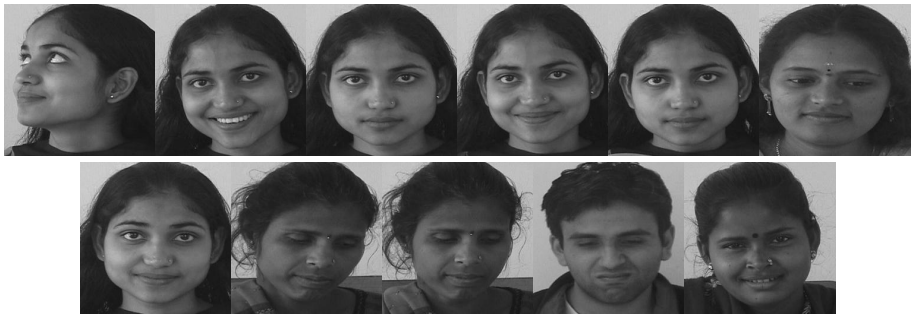


Figure8.b- Retrieved images using DWT features

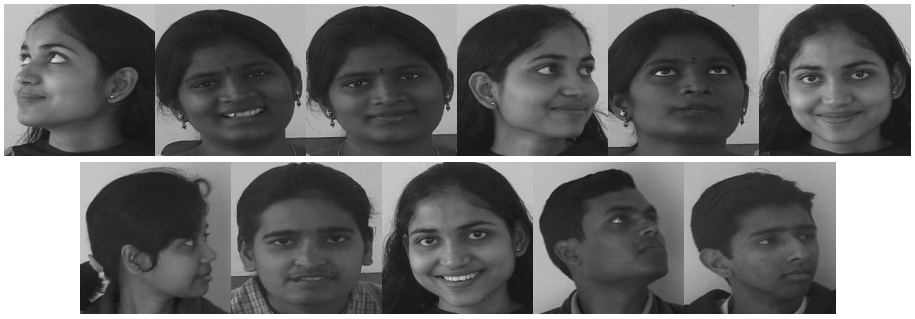


Figure8.c- Retrieved images using CT features

5. Conclusion:

A comparison between DWT and CT transforms for CBIR of face images is presented. The comparison is carried out with respect to the average retrieval rate. Two face databases are used to evaluate the performance of the CBIR system. These databases are "***Indian***", and "***GT***" face databases.

From the analysis of the results, it is found that the retrieval rate is database dependent. To retrieve all the images of a given class in the Indian face database, the average retrieval rate is 32.84% using CT and 52.19% using DWT. For the other face database (GT face database), the average retrieval rate is 72.65% using CT and 73.33% and 73.27% using the third and fourth level of DWT decomposition respectively. In general, the DWT gives better results than CT, but when retrieving a subset of the images from the GT face database the CT gives better results.

References:

1. *Long F., Zhang H, Feng D. D., "**Fundamentals of Content Based Image Retreival**".*

2. *Eakins J., Graham M.," Content-based Image Retrieval", University of Northumbria at Newcastle, JISC Technology Applications programme, Report no: 39, October 1999.*

3. *Mohamed A. S. S., Weng Y., Jiang J., Ipson S., "**An Efficient Face Image Retrieval through DCT Features",** Signal and Image Processing (SIP 2008), Kailua-Kona, HI, USA, August 18 – 20, 2008*

4. *Ramesh B. D. C., et. al. "**A Generic Approach to Content Based Image Retrieval Using DCT and Classification Techniques**",*

(IJCSE) International Journal on Computer Science and Engineering, Vol. 2, No. 6, 2010.

5. *Besbas W. S., Artemi M. A., Salman R. M., AL_Rjebi M. M., **"Content Based Face Image Retrieval in Walsh Hadamard Transform (WHT) Domain"**, ICCNDT 2012, Gulf University, Bahrain, Nov. 11-13, 2012*

6. *Lamard M, Cazuguel G, Quellec G, Bekri L, Roux C, Cochener B ,**"Content Based Image Retrieval based on Wavelet Transform Coefficients Distribution"**, Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, Lyon, France, August 23-26, 2007.*

7. *Hiremath P.S., Shivashankar S., and Pujari J., **"Wavelet Based Features for Color Texture Classification with Application to CBIR"**, International Journal of Computer Science and Network Security (IJCSNS), Vol. 6 No. 9A, September 2006*

8. *Rao C. S., kumar S. S., Chatterji B. N., **"Content Based Image Retrieval using Contourlet Transform"**, GVIP Journal, Vol 7, Issue 3, November 2007.*

9. *Georgia Tech (GT) Face Database Available at. ftp://ftp.ee.gatech.edu./pub/users/hayesfacedb*

10. *Indian face database website: http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase Indian Institute of Technology Kanpur.*