Comparative Analysis of Conventional, Real and Complex Wavelet Transforms for Face Recognition

Ankit A. Bhurane Dept. of Electronics and Telecom Engineering, SGGS IE &T, Nanded, India. Sanjay N. Talbar Dept. of Electronics and Telecom Engineering, SGGS IE &T, Nanded, India. Preeti N. Gophane Dept. of Electronics and Telecom Engineering, SGGS IE &T, Nanded, India.

ABSTRACT

Face recognition has been studied for many years in the context of biometrics and is one of the most successful applications of image analysis and understanding. Various methods, approaches and algorithms for recognition of human faces were proposed. In this paper, independent, comparative study of conventional discrete wavelet transform (DWT), real dual-tree discrete wavelet transform (R-DT-DWT), and complex dual-tree discrete wavelet transform (C-DT-DWT) based features for face recognition is carried out. In 2005, Delac et al. [26] presented an independent comparative study of PCA, ICA, and LDA on the FERET data set where it was concluded that no particular distance-metric combination is the best. In this paper we intend to bring further conclusions. Unlike the contribution by Delac et al., our conclusions are in context of DWT, R-DT-DWT, and C-DT-DWT. Moreover, these approaches are tested on nine different databases at different levels and under three different distance metrics, which allowed us to compare their performance independently. Our simulation results show that no particular distance-metric combination is the best across all standard benchmark face databases. However, the overall performance for city block distance measure was found to be better as compared to the Euclidean and cosine distance. Also, the performance for R-DT-DWT and C-DT-DWT based features were found equivalently efficient in many cases. So taking redundancy into consideration, it may be suggested to opt for R-DT-DWT for face efficient recognition.

Keywords

Real Dual-tree Discrete Wavelet Transform (R-DT-DWT), Complex dual-tree discrete wavelet transform (C-DT-DWT), Face Recognition Technology (FERET).

1. INTRODUCTION

Face recognition has always been a hot research topic for last several years. It is greatly motivated by the needs of military, public security, scene surveillance including commercial and law enforcement applications [1]. *Turk* and *Pentland* used the principal component analysis (PCA) technique to develop the first successful and well known Eigenface scheme for face recognition which is widely used due to its simplicity and effectiveness. Later various methods for the face recognition were proposed. Over the last decade wavelets have become powerful and flexible tools for computation and data reduction. For good direction selectivity, real dual-tree discrete wavelet transform (C-DT-DWT) were introduced in [2], [3].

The aim of this paper is to provide an independent, comparative study of conventional discrete wavelet transform (DWT), real dual-tree discrete wavelet transform (R-DT-DWT), and complex dual-tree discrete wavelet transform (C-DT-DWT) with accompanied distance metrics and in completely equal working conditions. In order to perform a fair comparison, same input images are the input into all three algorithms. In this paper, all the three algorithms are tested on three different distance metrics. Distance metrics used are L1 (City block), L2 (Euclidean), and cosine. This paper is organized as follows: Section II gives short introduction to DWT, R-DT-DWT, and C-DT-DWT. In Section III experimental results are presented. Finally, conclusions are drawn in Section IV.

2. PRELIMINARIES

A simple face recognition system is as shown in the Figure 1. It includes feature extraction, similarity measurement and decision making. Given a stored database of facial images one has to train an automated system to identify or verify a person.





2.1 Conventional two dimensional wavelet discrete transform (2D-DWT)

A conventional two dimensional wavelet discrete transform (2D-DWT) can be considered as equivalent to filtering the input image with a bank of filters, whose impulse responses are all approximately given by scaled versions of a mother wavelet. The separable (row-column) implementation of the 2D-DWT is characterized by three wavelets,

$$\psi_1(x, y) = \phi(x)\psi(y)$$
 (LH wavelet)
(1)

$$\psi_2(x, y) = \psi(x)\phi(y)$$
 (HL wavelet)
(2)

$$\psi_3(x, y) = \psi(x)\psi(y)$$
 (HH wavelet)
(3)

The LH wavelet is the product of the low-pass function $\phi(\cdot)$ along the first dimension and the high-pass (actually a bandpass) function $\psi(\cdot)$ along the second dimension. Similarly the HL and HH wavelets can be treated. In this paper, the filters given by *Abdelnour* et al. [4] were used.

2.1.1 Problems with Conventional Wavelets

The problems with conventional wavelets are as follows:

- *Shift Invariance:* Real DWT are very sensitive to shifts. Small shifts in the input signal can cause major variations in the distribution of energy between DWT coefficients.
- *Poor Directional Selectivity:* DWT coefficients reveal only three spatial orientations (horizontal, vertical and diagonal).
- Lack of Phase: DWT analysis of real signals lacks the phase information that accurately describes nonstationary signal behavior.

2.1.2 Comparison with Gabor approach

Apart from many techniques proposed in the literature, Gabor wavelet-based representation of face images is an excellent solution covering most of the above properties. However, Gabor wavelet-based representation is computationally very complex and the memory requirement for storing Gabor features is very high. These problems were tried to overcome using sub-Gabor [15], simplified Gabor wavelets [16], optimal sampling of Gabor features [17], etc. But none of these approaches succeeded effectively. An approach to both shift invariance and directional selectivity was given by Simoncelli et al. [18]. However, the dual-tree complex wavelet transforms approach proposed by Kingsburry [19], [20], gives perfect reconstruction and greater directional selectivity. Thus, it provides an excellent alternative to Gabor wavelets with the potential to overcome the above-mentioned shortcomings of the Gabor wavelets.

Sankaran et al. [21] and Celik et al. [22] used the Complex Dual-Tree Discrete Wavelet Transform (C-DT-DWT) and Gabor wavelets for facial feature extraction. Complex Wavelet Transform based face recognition was also discussed in [23]. In [24] and [25], orthogonal neighborhood preserving projections (ONPPs) and supervised kernel ONPP with C-DT-DWT was used for face recognition.

2.2 2-D Dual-Tree Discrete Wavelet Transform (DT-DWT)

The 2-D dual-tree discrete wavelet transform (DT-DWT) of an image is implemented using two critically-sampled separable 2-D DWT's in parallel as shown in Figure 2. One of the advantages of the dual-tree DWT (DT-DWT) over separable 2D DWT is that, it can be used to implement 2D wavelet transforms that are more selective with respect to orientation. The filters used in the first stage of the dual-tree DWT are different from the filters used in the remaining stages [3].

Here, we have used *Farras* filters for the first stage and *Kingsbury's* Q-shift filters for remaining stages.



Fig 2: 2D Dual-tree discrete wavelet transform.

2.2.1 2-D Real Dual-Tree Discrete Wavelet Transform (R-DT-DWT)

In real dual-tree discrete wavelet transform (R-DT-DWT), the sum and difference for each pair of subbands gives rise to wavelets in six different directions. At each level, the wavelets are strongly oriented at angle of $\pm 15^{\circ}$, $\pm 45^{\circ}$, and $\pm 75^{\circ}$ as shown in Figure 3.



Fig 3: Six Wavelets associated with the real 2D dual-tree DWT.

2.2.2 2-D Complex Dual-Tree Discrete Wavelet Transform (C-DT-DWT)

The 2-D complex dual-tree discrete wavelet transform (C-DT-DWT) has twice as many wavelets as that of R-DT-DWT (two wavelets in each direction). The wavelets are oriented in the same six directions as those of the 2-D R-DT-DWT. In each direction, one of the two wavelets can be interpreted as the real part of a complex-valued 2D wavelet, while the other wavelet can be interpreted as the imaginary part of a complex-valued 2D wavelet as shown in the Figure 4. The extra six wavelets are obtained by swapping the low pass and high pass filter coefficients of real and imaginary trees.

An illustration of R-DT-DWT operated on a synthetically generated image is as shown in the Figure 5.



Fig 4: Impulse responses for 2-D C-DT-DWT: First row is interpreted as the real part and the second row as imaginary

part of the complex wavelet. The third row shows the magnitude response.



(a) A synthetically generated image



(b) Thresholded R-DT-DWT subbands at third level



(c) Thresholded C-DT-DWT subbands at level = 3.

Fig 5 : R-DT-DWT and C-DT-DWT subbands of a synthetically generated image.

3. EXPERIMENTAL RESULTS

3.1 Simulation Software

The software for face recognition was built using MATLAB. This environment was chosen because it easily supports image processing, image visualization, and linear algebra. The software was tested against wide range of databases

3.2 Benchmark Face Databases and Settings

In order to evaluate the proposed face recognition system, our experiments were performed on following nine benchmark face databases:

3.2.1. Face Recognition Technology (FERET) Database

The FERET database [6] contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images. In our experiment 50 subjects were chosen randomly, each having 11 frontal face images varying in illumination, expression, and pose.

3.2.2. Olivetti Research Laboratory (ORL) face database

The ORL face database [7] contains ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open/ closed eyes, smiling/ not smiling) and facial details (glasses/ no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). There are 10 images per subject.

3.2.3. Yale face database

The Yale face database [8] contains 165 images of 15 subjects. There are 11 images per subject, one for each of the following facial expressions or configurations: center-light, with glasses, happy, left-light, without glasses, normal, right-light, sad, sleepy, surprised, and wink.

3.2.4. Japanese Female Facial Expression (JAFFE) database

The JAFFE database [9] contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects. In our experiment, 20 face image of each subject were chosen.

3.2.5. University of Bern face database

The University of Bern face database [10] contains frontal views of 30 people. For each person 10 gray level images with slight variations of the head positions (1, 2 right into the camera, 3,4 looking to the right, 5,6 looking to the left, 7,8 downwards, 9,10 upwards.

3.2.6. Indian Institute of Technology Kanpur (IITK) face database

The IITK database [11] consists of 11 different images of each of 40 distinct subjects. For some subjects, some additional photographs are included. All the images have a bright homogeneous background with the subjects in an upright, frontal position. The following orientations of the face are included: looking front, looking left, looking right, looking up, looking up towards left, looking up towards right, looking down. Available emotions are: neutral, smile, laughter, sad/disgust.

In our experiment, 20 male and 20 female data sets were chosen and were treated independently.

3.2.7. Computer Vision Science Research (CVSR) -Grimace face database

The CVSR database [12] contains a sequence of 20 images of 18 different individual, using a fixed camera. During the sequence the subject moves his/her head and makes grimaces which get more extreme towards the end of the sequence.



Fig 6: Database samples: (a) FERET, (b) ORL, (c) Yale, (d) JAFFE, (e) University of Bern, (f) IITK [Male], (g) IITK [Female], (h) CVSR [Grimace], (i) Georgia Tech., (j) S.G.G.S.

3.2.8. Georgia Tech face database

The Georgia Tech. database [13] contains 15 different images of each of 50 distinct people taken in two or three sessions at the Center for Signal and Image Processing at Georgia Institute of Technology. The pictures show frontal and/or tilted faces with different facial expressions, lighting conditions and scale. In our experiments we had used cropped database.

The benchmark images were treated on "as is" basis. All the images were initially resized to dimension 64×64 for our experiments. No other preprocessing was done. For each subject of each database, about half of the images were used for training purpose and remaining others for testing.

3.2.9. S.G.G.S. Face Database

The SGGS database [5] contains 25 different images of each of 30 distinct people taken in one session at Shri Guru Gobind Singhji (S.G.G.S.) Institute of Engineering and Technology, Nanded, India. For some subjects, some additional photographs are included. All the images were taken in plain background, varying the lighting, pose, facial expressions and facial details (glasses/ no glasses) in robust environment.

Figure 6 shows samples from each of the above databases.

3.3 Feature Extraction

At first level, a face image of size $N \times N$ is decomposed into four subband image: LL, LH, HL, and HH. Higher level decompositions are on LL subband image at previous level. The feature extraction is done as follows: Input image is initially resized to an $N \times N$ image. Further, the DWT, R- DT-DWT, C-DT-DWT are applied independently up to level J. At J^{th} level, the subband images will have size $\left(\frac{N}{2^J} \times \frac{N}{2^J}\right)$. Then, the feature vector can be formed by

concatenating the rows or columns of all the subbands at J^{th} level. For example, the feature vector size for conventional DWT, R-DT-DWT, and C-DT-DWT at J = 4 will be

$$4 \times \left(\frac{N}{2^{J}} \times \frac{N}{2^{J}}\right), \qquad \left[2 \times 4 \times \left(\frac{N}{2^{J}} \times \frac{N}{2^{J}}\right)\right] \qquad \text{and} \\ 2 \times \left[2 \times 4 \times \left(\frac{N}{2^{J}} \times \frac{N}{2^{J}}\right)\right] \qquad \text{respectively. For our}$$

experiments, the input images are initially scaled to 64×64 . Independent analysis of DWT, R-DT-DWT and C-DT-DWT was done at third and fourth level.

The R-DT-DWT and C-DT-DWT subbands of a face image are shown in Figure 7.

3.4 Similarity Measures

The similarity measures used in our experiments to evaluate the efficiency of different representation and recognition methods include Euclidean distance, city block distance, and cosine distance. If,

X: *mx*-by-*n* data matrix treated as *mx* $(1 \times n)$ row vectors x_1 , x_2 , ..., x_{mx} , and

Y: my-by-n data matrix treated as my $(1 \times n)$ row vectors y_{l} ,

*y*₂, ..., *y*_{my},

then, the similarity measures between the vectors x_s and y_t are defined as follows:

A. City Block distance (L1 distance):







(b) R-DT-DWT subbands of input image



(c) C-DT-DWT subbands of input image

Fig 7: R-DT-DWT and C-DT-DWT subbands of a face image.

B. Euclidean distance (L2 distance):

$$d_{st}^{2} = (x_{s} - y_{t})(x_{s} - y_{t})'$$
(5)

C. Cosine distance:

$$d_{st} = \left(1 - \frac{x_{s} y_{t}^{'}}{\sqrt{(x_{s} x_{s}^{'})(y_{s} y_{s}^{'})}}\right)$$
(6)

3.5 Results

For each subject of each database, about half of the images were used for training purpose and remaining others for testing. The results are tabulated in Table 1. The efficiency of recognition can be calculated as follows,

$$Efficiency = \frac{correctly \ recognized \ face \ images}{total \ number \ of \ input \ face \ images} \times 100\%$$

The results show that no particular distance-metric combination is the best across all standard benchmark face databases. However, overall performance for city block distance measure is found to be better as compared to the Euclidean and Cosine distance in many cases. Also, the performance of R-DT-DWT and C-DT-DWT based features are found to be about equally efficient in many cases.

4. CONCLUSIONS

This paper presented an independent, comparative study of DWT, R-DT-DWT and C-DT-DWT based face recognition methods with their accompanied three distance metrics (L1, L2 and cosine), at different levels, and in completely equal working conditions.

Our comparative research shows that no particular distancemetric combination is the best across all standard benchmark face databases. This verifies the contribution made by Delac *et al.* [26]. In addition, however, overall performance for city block distance measure was found to be better as compared to the Euclidean and cosine distance in many cases. Also, the performance of R-DT-DWT and C-DT-DWT based features were found to be about equally efficient in many cases. Although C-DT-DWT has the benefit of being both oriented and approximately analytic, the overall performance of R-DT-DWT and C-DT-DWT based features are found to be about equally efficient in many cases for face recognition. So taking redundancy into consideration, it may be suggested to opt for R-DT-DWT for efficient face recognition.

4. ACKNOWLEDGMENTS

The authors would like to thank the teaching, non-teaching staff members and M.Tech students from Electronics department of Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded (Maharashtra, India) for their cooperation in creating the SGGSIET face database. The authors would like to thank those researchers who have kindly provided the public face databases such as FERET, ORL, Yale, JAFFE, IITK, CVSR, University of Bern, and Georgia Tech. Authors would also like to thank Dr. Kingsbury (Cambridge University) and Prof. Selesnick (Polytechnic University, NY, USA) for answering the queries.

Table 1	:	Face	Reco	gnitio	n	Results.
---------	---	------	------	--------	---	----------

	Database		Efficiency (%)						
Distance			J = 3			J = 4			
Measures			DWT	R-DT- DWT	C-DT- DWT	DWT	R-DT-DWT	C-DT- DWT	
	FERET		74.33	73.33	73.33	73.33	73.66	73.66	
	ORL		93.50	93.50	93.50	92.50	94.00	94.00	
	Yale		87.77	88.88	88.88	85.55	85.55	85.55	
	JAFFE		92.00	91.00	91.00	90.00	91.00	91.00	
Euclidean	Unive	ersity of Bern	92.00	91.33	91.33	90.66	92.00	92.00	
	ПТК	IITK (Male)	74.16	70.00	70.83	72.50	73.33	73.33	
	IIIK	IITK (Female)	86.66	86.66	85.83	87.50	87.50	88.33	
	CVSR (Grimace)		96.66	96.66	96.66	96.66	96.66	96.66	
	Georgia Tech.		71.00	71.50	71.50	70.50	69.75	70.00	
	SGGS		67.43	66.92	66.92	62.05	62.56	63.07	
	P	OSTECH	72.88	73.33	73.11	63.77	64.00	64.00	
Average Efficiency			82.58	82.10	82.08	80.45	80.91	81.05	
	FERET		77.66	79.33	78.33	74.66	75.33	75.33	
		ORL	95.00	96.00	96.50	95.00	95.55	95.50	
	Yale		94.44	95.55	95.55	93.33	93.33	94.44	
	JAFFE		95.00	97.00	97.00	97.00	96.00	96.00	
	Unive	ersity of Bern	94.00	94.66	95.33	94.66	95.33	95.33	
Cityblock	IITV	IITK (Male)	75.00	77.50	77.50	74.16	76.66	76.00	
	IIIK	IITK (Female)	86.66	88.33	88.33	89.16	90.00	89.16	
	CVSR (Grimace)		96.66	96.66	96.66	96.66	96.66	96.66	
	Georgia Tech.		72.75	72.75	73.00	72.00	72.25	72.75	
	SGGS		77.94	77.69	77.17	71.53	74.35	74.61	
F		OSTECH	82.44	82.66	83.33	75.11	72.88	73.77	
Average Efficiency			86.14	87.10	87.15	84.84	85.30	85.41	
Cosine	FERET		84.00	84.33	84.33	83.00	82.00	82.33	
	ORL		98.83	92.00	92.00	93.00	83.50	94.00	
	Yale		91.11	92.22	92.22	87.77	90.00	90.00	
	JAFFE		96.00	96.00	96.00	98.00	98.00	97.00	
	University of Bern		90.00	90.66	90.66	88.00	88.00	88.00	
	IITK	IITK (Male)	69.16	67.50	67.50	70.83	70.00	69.16	
		IITK (Female)	87.50	87.50	87.50	89.16	89.16	89.16	
	CVSR (Grimace)		96.66	96.66	96.66	96.66	96.66	96.66	
	Georgia Tech.		65.75	65.75	65.75	67.75	68.50	68.75	
	SGGS		72.56	71.53	71.53	69.74	69.48	69.23	
POSTECH		72.44	72.66	72.66	62.22	63.33	63.55		
Average Efficiency		84.00	83.34	83.34	82.37	81.69	82.53		

5. REFERENCES

- W. Zhao, R. Chellappa, A. Rosenfeld, P. J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys, 35(4): 399-458, 2003.
- [2] N. Kingsbury. The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters. *IEEE Digital Signal Processing Workshop*, DSP 98, paper no. 86, August 1998.
- [3] I. Selesnick, R. Baraniuk, and N. Kingsbury. The dualtree complex wavelet transform. *IEEE Signal Process*. *Mag.*, 22(6):123–151, Nov. 2005.
- [4] A. F. Abdelnour and I. W. Selesnick. Nearly symmetric orthogonal wavelet bases. In Proceedings of IEEE International Conference on Acoustic, Speech, Signal Processing (ICASSP), May 2001.
- [5] A. A. Bhurane. "Face Recognition using Dual-Tree Discrete Wavelet Transforms". M.Tech Thesis, S.G.G.S. Institute of Engineering and Technology, Nanded, Maharashtra, India, July 2011.

- [6] The color FERET database, USA. Website. http://face.nist.gov/colorferet/.
- [7] AT & T: The database of faces (formerly: The ORL database of faces). Website. http://www.cl.cam.ac.uk/Research/DTG/attarchive/faced atabase.html.
- [8] The Yale face database. Website. http://cvc.yale.edu/projects/yalefaces/yalefaces.html.
- [9] The Japanese Female Facial Expression (JAFFE) database. Website. http://www.kasrl.org/jaffe.html.
- [10] The University of Bern face database. Website. http://www.ph.tn.tudelft.nl/PRInfo/data/msg00010.html.
- [11] The IITK face database. Website. http://viswww.cs.umass.edu/~vidit/IndianFaceDatabase/.
- [12] The CVSR Grimace database. Website. http://cswww.essex.ac.uk/mv/allfaces/index.html.
- [13] The Georgia Tech. face database. Website. http://www.anefian.com/research/GTdb_crop.zip.

- [14] The SGGS face database. Website. http://sggs.ac.in.
- [15] L. Nanni and D. Maio, "Weighted sub-Gabor for Face Recognition," *Pattern Recognition Letters*, vol. 28, no. 4, pp. 487–492, 2007.
- [16] W.-P. Choi, S.-H. Tse, K.-W. Wong, and K.-M. Lam, "Simplified Gabor Wavelets for Human Face Recognition," *Pattern Recognition Letters*, vol. 41, no. 3, pp. 1186–1199, 2008.
- [17] D.-H. Liu, K.-M. Lam, and L.-S. Shen, "Optimal Sampling of Gabor Features for Face Recognition," *Pattern Recognition Letters*, vol. 25, no. 2, pp. 267–276, 2004.
- [18] E P Simoncelli, W T Freeman, E H Adelson and D J Heeger, "Shiftable Multiscale Transforms", *IEEE Transactions on Information Theory*, 38(2), pp 587-607, March 1992.
- [19] N G Kingsbury, "The Dual-Tree Complex Wavelet Transform: A New Technique for Shift Invariance and Directional Filters", Proc. 8th IEEE DSP Workshop, Bryce Canyon, Aug 1998.
- [20] N G Kingsbury: "The Dual-Tree Complex Wavelet Transform: A New Efficient Tool for Image Restoration and Enhancement", Proc. EUSIPCO 98, Rhodes, Sept 1998.
- [21] H. E. Sankaran, A. P. Gotchev, K. O. Egiazarian, and J. T. Astola, "Complex Wavelets versus Gabor Wavelets for Facial Feature Extraction: A Comparative Study," in *Image Processing: Algorithms and Systems IV*, vol. 5672 of *Proceedings of SPIE*, pp. 407–415, San Jose, Calif, USA, January 2005.

- [22] T. Celik, H. Ozkaramanli, and H. Demirel, "Facial Feature Extraction using Complex Dual-Tree Wavelet Transform," *Computer Vision and Image Understanding*, vol. 111, no. 2, pp. 229–246, 2008.
- [23] A. Eleyan, H. Ozkaramanli, and H. Demirel, "Complex Wavelet Transform-Based Face Recognition," *EURASIP Journal on Advances in Signal Processing*, 2008, Article ID 185281.
- [24] Y. Sun and M. Du, "DT-CWT Feature Based Classification using Orthogonal Neighborhood Preserving Projections for Face Recognition," in Proceedings of the International Conference on Computational Intelligence and Security (CIS '06), pp. 719–724, Guangzhou, China, November 2006.
- [25] Y. Sun, "DT-CWT Feature based Face Recognition using Supervised Kernel ONPP," in *Proceedings of the International Conference on Computational Intelligence* and Security Workshops (CISW '07), pp. 312–315, Harbin, China, December 2007.
- [26] Delac K., Grgic M., and Grgic S. Independent comparative study of PCA, ICA, and LDA on the FERET data set. Wiley Periodicals, 15:252–260, 2005.
- [27] Bhurane A.A. and Talbar S.N. "Vision-based authenticated robotic control using face and hand gesture recognition," *International Conference on on Electronics Computer Technology (ICECT 2011)*, vol V-1, pp 64–68. ICECT, IEEE, April 2011.