# Best Wavelet Function for Face Recognition Using Multi-Level Decomposition

Nadir Nourain Dawoud
Department of Electrical and Electronic Engineering
Universiti Teknologi PETRONAS
Tronoh, Perak, Malaysia
nadirnourain@gmail.com

Brahim Belhaouari Samir
Department of Fundamental and Applied Science
Universiti Teknologi PETRONAS
Tronoh, Perak, Malaysia
brahim_belhaouari@petronas.com.my

*Abstract*-**The selection of appropriate wavelets is an important target for any application. In this paper, wavelets functions are examined in order to choose the best wavelet for face classification process and for finding the optimal number of levels of decomposition. Seven wavelet functions namely Symelt, Daubechig, Coiflets, Mayer Discrete, Biorthogonal, Reverse Biorthogonal and Haar were tested with different number of decomposition levels and different number of biggest coefficients is selected to reduce the huge feature dimension, and then the Euclidean Distance Method (EDM) was used for classification process. Also a statistical method has been proposed to produce new metric of features coefficients, the experiments brought about 40% improvements in comparison to the method that accounts the biggest coefficients from the four levels of decompositions. The experiments have been performed on Olivetti Research Laboratory database (ORL) and Yale University database (YALE). The result showed the effect of wavelets proprieties on classification process and the Symelt wavelets are the optimum wavelets for the face classification with four levels.**

***Keywords-wavelet transform; multi-level decomposing; Euclidean distance method;***

## I. Introduction

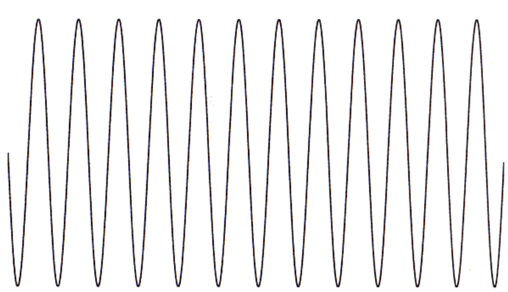
The appropriate mother wavelet should be chosen in order to use wavelet transform effectively for any application and the details of particular application should be taken into account. This is attributed to characteristics of the wavelet transform result which will be determined by the mother wavelet that it produced for all wavelet functions used in the transformation through translation and scaling. Many numbers of basic functions can be used as mother wavelet for wavelet transform. In other words, these wavelet functions affect the accuracy what so ever the purpose or the application would be. Many researchers used wavelet transform for different purposes; in [1] they used multi-level of decomposition to increase the diagnostic accuracy as well as the reproducibility of mammographic explanation and they used four different levels of decomposition based on three different wavelet functions, which are (db8), (sym8) and (bior3.7). In [2] they examined the visual texture classification in order to investigate how well it could be used for haptic texture search engine. In classifying the visual texture the image features are extracted and then using the wavelet decomposition the transformation coefficients were obtained. The wavelet function used in the analysis was Daubechies wavelet. In [3] they proposed a new idea about texture modeling. From orthogonal wavelet decomposition such as Haar or Daubechies wavelet, they obtained non-redundant sub-bands in the different levels (scales) and directions. Also on the starting point of introducing the wavelet decomposition, Principal Component Analysis (PCA) and Support Vector Machine (SVM), were used for face recognition as proposed by [4] and according to [4] (db4) wavelet functions were used to extract the facial features. A new recognition system was proposed in [5] this system is based on Hidden Markov Model (HMM) and wavelet decomposition and the wavelet coefficients were computed from (Haar) wavelet for each sub-image extracted from original image. In the other hand, the relationship between wavelets functions and digital filter banks can be found in [6] and [7]. The functions (families) were organized according to the numerous properties that enable to make differentiate among them.

The reduction of wavelet transform coefficients which represent the face in the image is the most significant problem due to the redundancy of features which are not required in Discrete Wavelet Transform (DWT) and these features in turn draws a negative effect on the classification process. This problem refers to the properties of wavelet functions (families) which are used in particular application to produce the coefficients to represent the features. In the case of face classification, some of these coefficients don't have face information that leads to increase the error rate of classification. For this problem, it's important to reduce the coefficients by choosing those coefficients that contain face information and ignoring the remaining. In [1] they solve the problem of coefficients' redundancy by selecting the biggest hundred coefficients from each level of decomposition and then passing these coefficients to EDM classifier.

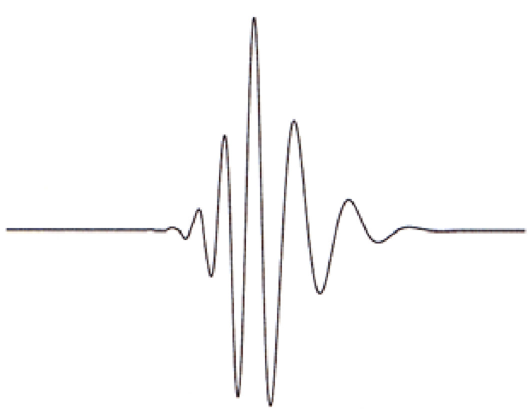
The rest of the paper is organized as follows. In Section II the Wavelet Transform is briefly resumed, and in Section III the system introduced in [1] is described. Section IV the proposed statistical model and error probability are derived. Section V reports a brief introduction of the wavelet families, and in Section VI the comparison between wavelet families for face recognition is discussed. Finally in Section VII, conclusions are drawn.

## II. Wavelet Transform (WT)

From [8] the transform of a signal is considered as a different way of representing the signal. And it save the information content present in the signal. Wavelets make use of different sets of basis functions to permitting the decomposition of continuous and discrete signals. Wavelet Transform offers a time-frequency representation of the signal. It was built up to rise above the short coming of the Short Time Fourier Transform (STFT), which can also be used to analyze non-stationary signals. While STFT provides a constant resolution at all frequencies, the Wavelet Transform utilizes multi-resolution technique by which different frequencies are analyzed with different resolutions. A wave is fluctuating function of time or space and is cyclic. Unlike the wavelets are different and localized waves. Also anther difference, wavelets have their energy concentrated in time or space and are appropriate to analysis of temporary signals. While Fourier Transform and STFT utilize waves to analyze signals, the Wavelet Transform (WT) uses wavelets of finite energy. The signal which expects to analyze is multiplied with wavelet function and then the transform is calculated for each part created and that is not like the STFT, furthermore in the wavelet the width of the wavelet function vary with each spectral component. At high frequencies, wavelet transform gives good time resolution and poor frequency resolution, while at low frequencies; gives good frequency resolution and poor time resolution. "Fig. 1" illustrates the differentiation between the wave and wavelet:



(a)



(b)

Figure 1. Demonstration of (a) a Wave and (b) a Wavelet.

From [8], in equation (1), The Continuous Wavelet Transform (CWT) is provided where x (t) is the signal to be analyzed. Ψ (t) is the mother wavelet or the basis function. The mother wavelet is considered as source of all the wavelet functions used in the transformation during translation (shifting) and scaling (dilation or compression).

$$X_{WT}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t).\Psi^{*}\left(\frac{t-\tau}{s}\right)dt \qquad (1)$$

All the basic functions are generated from the mother wavelet used and designed based on some most wanted characteristics associated with that function. The two parameters, translation parameter τ and scale parameter s are defined as relates to the position of the wavelet function as it is shifted through the signal and corresponds to the time information in the Wavelet Transform for the parameter τ also scale parameter s is defined as |1/frequency| and corresponds to frequency information. The scaling either dilates (expands) or compresses a signal. Large scales expand the signal and gives detailed information disappear in the signal, whereas small scales compress the signal and gives general information about the signal. In the other approach the Wavelet Series is a sampled version of CWT and its calculation may need large amount of time and resources, which it based on the resolution requested. The Discrete Wavelet Transform (DWT), which is based on sub-band coding, is found in [8] and is not like (CWT) its fast computation of Wavelet Transform. Also it is easy to execute and reduces the computation time and resources requested. When we want to compare between CWT and DWT in hand of signal analysis we found the signals are analyzed using a set of basic functions in CWT which relate to each other by simple scaling and translation. But in the case of DWT, we can obtain the time-scale representation of the digital signal using digital filtering techniques. The signal we want analyzed is passed through filters with different cutoff frequencies at different scales.

In signal processing the filters can be consider as one of the most widely used. If we repeat the filters with rescaling we can realize the Wavelets. The amount of the information in the signal is measuring by resolution of signal which is determining by the filter operation; also the scale has been determined by the up-sampling and down-sampling (sub-sampling) operations. The DWT is calculated by in a row low-pass and high-pass filtering of the discrete time-domain signal as shown in Fig. 2. This is called the Mallat algorithm or Mallat-tree decomposition. Its worth is in the approach it join the continuous-time multi-resolution to discrete-time filters. From the "Fig. 2", we can observe that the sequence x[n] represent the signal, where n is an integer. G0 represent the low pass filter as well as H0 represent the high pass filters. The detail information is producing by the high pass filter at each level; d[n], while uncouth approximations, a[n] produces by the low pass filter associated with scaling function.

## III. USING WAVELET FUNCTION FOR INCREASING THE DIAGNOSTIC ACCURACY

In this section the method proposed in [1] is detailed. In this method the system consists of three stages: preprocessing, feature extraction and classification process. In processing stage a set of images from MIAS database are used with (1024 X 1024) pixels and almost 50% of the whole image consisted of the background with a lot of noise and they used two techniques to enhance the mammograms (image pruning and histogram equalization) to cut off undesired portion of the image. In the second stage the features are extracted from previous stage based on the wavelet decomposition process. The features extracted from the coefficients that were produced by the wavelet analysis decomposition. Four different level of decomposition based on three different wavelet functions which are (db8), (sym8) and (bior3.7) wavelet function. Then they extracted biggest 100 coefficients from each level of decomposition and pass these to the next stage. Finally in the last stage the Euclidean Distance Method is used for classification process. The Euclidean Distance is used to design the classifier. In "Fig. 3" a sample of MIAS data base and the Figure show the original image and cropping image which is obtained by applying cropping operation on the original image to eliminate the noise and black background.

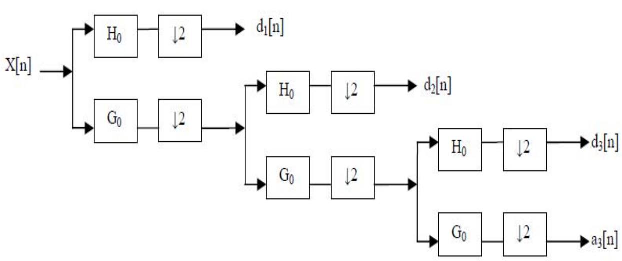
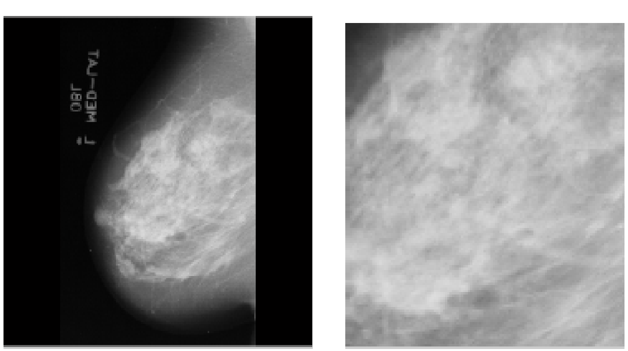



Figure 2. Three-level wavelet decomposition tree.

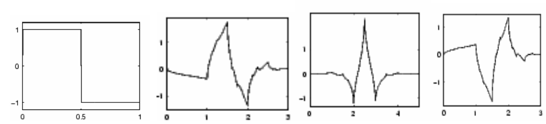


(a)The original image (b) the cropped image

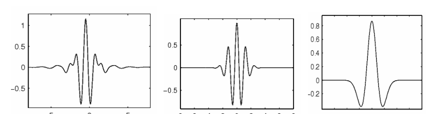
Figure 3. (a) the original image (1024x1024) pixels and (b) the cropped image (128x128) pixels

## IV. THE WAVELET FUNCTIONS (FAMILIES)

In wavelet transformation many basic functions can be consider as mother wavelet. While the all wavelet functions are producing by the mother the functions it will be used in the transformation through translation and scaling, furthermore the characteristics of the resulting Wavelet Transform determine by these function. Therefore, in order to using wavelet transform effectively the most important issue is the details of the particular application must be taken into account and the appropriate mother wavelet should be selected. "Fig. 4" shows some of the commonly used wavelet functions. One of the oldest and simplest wavelet is Haar wavelet. Therefore, any discussion of wavelets starts with the Haar wavelet. Daubechies wavelets are the most popular wavelets. They represent the basics of wavelet signal processing and are used in several applications. Due to their characteristic of frequency responses have maximum smoothness at frequencies 0 and π these are also called Maxflat wavelets. This is a very desirable property in some applications. The Haar, Daubechies, Symlets and Coiflets are considering as compactly supported orthogonal wavelets. These wavelets as well as Meyer wavelets are capable of perfect reconstruction. The Meyer, Morlet and Mexican Hat wavelets are symmetric in shape. The wavelets are chosen based on their shape and their ability to analyze the signal in a particular application.



(a) (b) (c) (d)



(e) (f) (g)

Figure 4. Wavelet families (a) Haar (b) Daubechies4 (c) Coiflet1 (d) Symlet2 (e) Meyer (f) Morlet (g) Mexican Hat.

## V. EXPERIMENTS AND RESULT AND DISCUSSION

In this section, firstly four levels of decomposing and seven different wavelet functions (Symelt, Daubechig, Coiflets, Mayer Discrete, Biorthogonal, Reverse Biorthogonal and Haar) were selected to choose the best among these functions for face recognition. The experiment have been conducted on ten images for randomly chosen 15 individuals out of a total of 40 individuals in ORL database while eleven images were chosen for each of the 15 individuals in YALE database. These databases are presented with different images per individual: one can notice that this databases is characterized by illumination, pose and expression changes between images of the same subject, in addition to all these characteristics of images, YALE data base suffer from deep shadow more than what is found in ORL database. A sample of YALE and ORL data base which were used during the experiment are shown in “Fig. 5”. The next step of the experiment is aimed at reducing the dimensions of the coefficients of the functions choosing the coefficients which have information that is important for the classification process. From [1], the different sets of the biggest coefficients are extracted from each level of decomposition. Besides, we modified the method in [1] by extracting the biggest coefficients after all features dimensions are aggregated together. Later, the functions which gave good results for each database were used again but with different number of decomposition levels to select the best number of levels. Finally, the simplest classifier, Euclidean Distance Algorithm (EDA), was used for classification process.

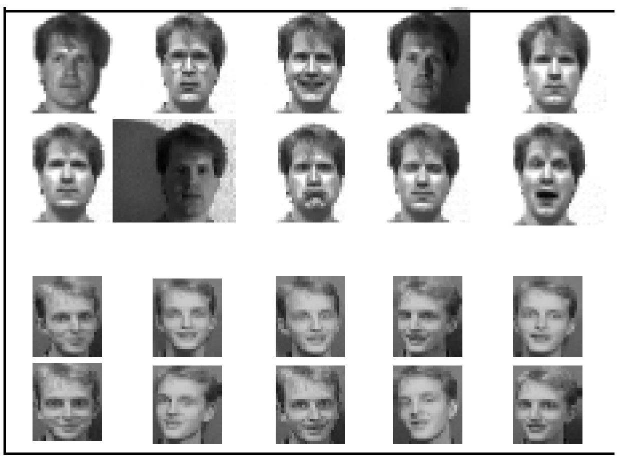


Figure 5. Face images in the Yale face database (Top) and the ORL face database (Bottom)

According to the main four properties of wavelets functions [7], the experiments results were analyzed based on these properties namely, orthogonality or biorthogonality, existence of associated filters, real or complex wavelets and compact or not compact support. In “Fig. 6” for ORL database and “Fig. 7” for YALE database, the seven different wavelets function (Symelt, Daubechig, Coiflets, Mayer Discrete, Biorthogonal, Reverse Biorthogonal, Haar) were examined and the biggest 200 coefficients from each level of decomposition and also from all levels were extracted as features vectors. Furthermore, the comparison between the two types of coefficient has been done and result shows that for ORL data base the Reverse Biothogonal wavelets gave better accuracy and this is associated to properties like compact support, symmetry, arbitrary number of zero moments, capability to continuous and discrete transformation among which the most important property is the biorthogonal analysis. In addition, the other two functions (biorthogonal and symelt) gave result which is close to the previous function. While in the case of YALE data base, according to properties of images like change in illumination, pose and expressions, it was observed that Mayer Discrete wavelets gave better results and this is linked to properties of infinite regularity and orthogonal and bi-orthogonal analysis but whose effects are appreciated after relatively longer time. In “Fig. 6-7”, the blue lines represent the method of selecting the biggest coefficients from each level, and red line represents the method used to extract the biggest coefficients from the all features dimensions.

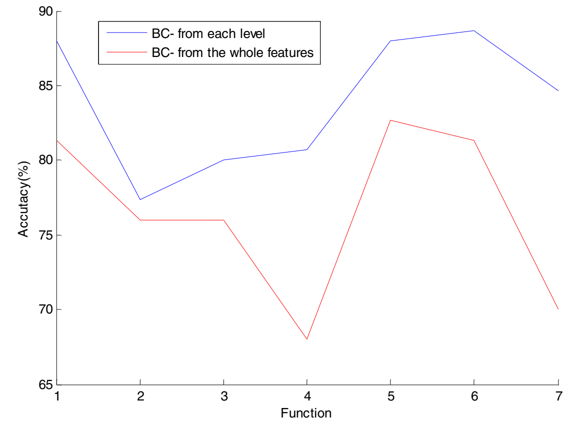



Figure 6. Comparison between the wavelet functions for the two methods on (ORL) database using 200 of biggest coefficients with 4 levels decomposition.

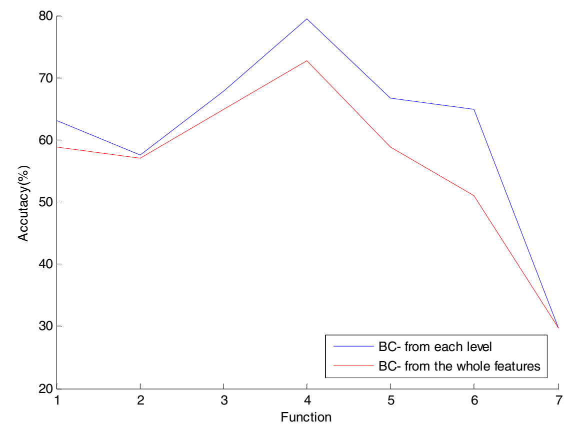



Figure 7. Comparison between the wavelet functions for the two methods on (YALE) database using 200 of biggest coefficients with 4 levels decomposition.

In “Fig. 8” for ORL database and “Fig. 9” for YALE database the two different functions (Reveres Biorthogonal) for ORL data base and (Meyer Discrete) were tested again based on their best results but with different number of the biggest features coefficients by using the two methods. During the experiments, we found that the maximum coefficient we can extract in ORL data base is 272 coefficients therefore the dimensions of level 4 did not allow us to extract more coefficients. The increase in the accuracy was observed until the number of extracting coefficients equals to number of coefficients of level 4 of decomposition. While in YALE data base, due to redundancy in features, it was possible to extract more number of coefficients with significant effects on the classification process. First an increase in accuracy with subsequent decrease was observed when taking more than 180 coefficients. In "Fig. 8-9", the blue lines represent the method of selecting the biggest coefficients from each level, and red line represents the method used to extract the biggest coefficients from the all features dimensions.

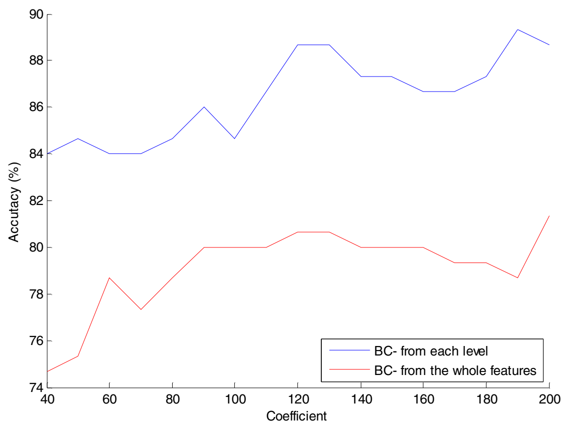



Figure 8. Comparison between two functions using two methods and different coefficients number with 4 level of decomposition (ORL) database.

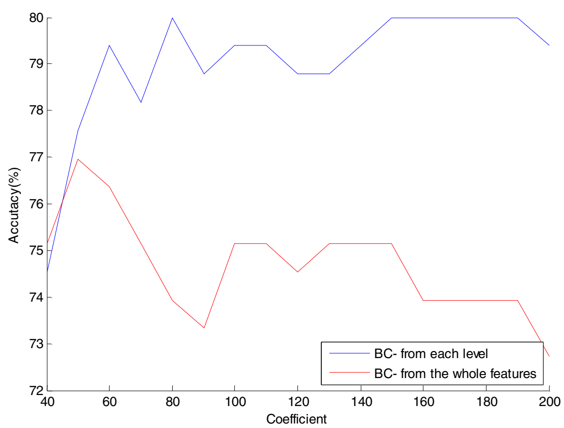



Figure 9. Comparison between two functions using two methods and different coefficients number with 4 level of decomposition (YALE) database.

In "Fig. 10" for ORL database and "Fig. 11" for YALE database during the experimental, Symelt functions gave results close to the results in the previous test and when the numbers of the extracted biggest coefficients were changed for the two methods of selecting the coefficients. The main reason of this test is due to the best result of the Symelt within suitable time. In "Fig. 10-11", the blue lines represent the method of selecting the biggest coefficients from each level, and red line represents the method used to extract the biggest coefficients from the all features dimensions. .

In "Fig. 12" for ORL database and "Fig. 13" the result showed that when 4 levels of decomposition were used the classification gave best results for ORL data base and above average result for YALE data base. In addition, from the experiments when the coefficients were extracted from approximation coefficients only the accuracy decreased. "Fig. 6" and "Fig. 7" showed the relation between Reveres Biorthogonal wavelets and Symelt wavelets in term of accuracy and the number of coefficients. In "Fig. 12-13", the blue lines represent the method of selecting the biggest coefficients from each level, and red line represents the method used to extract the biggest coefficients from the all features dimensions.

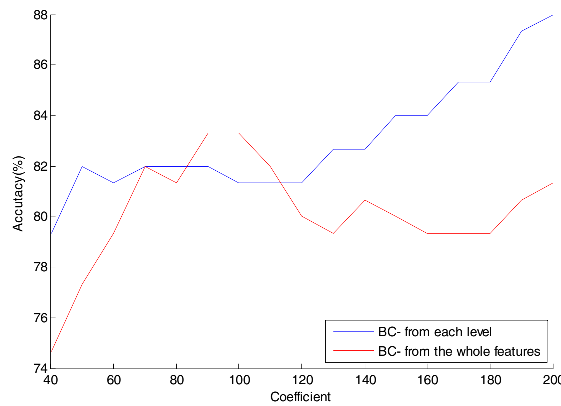



Figure 10. Comparison between two methods using Symelt function for different coefficients number with 4 levels of decomposition (ORL) database.

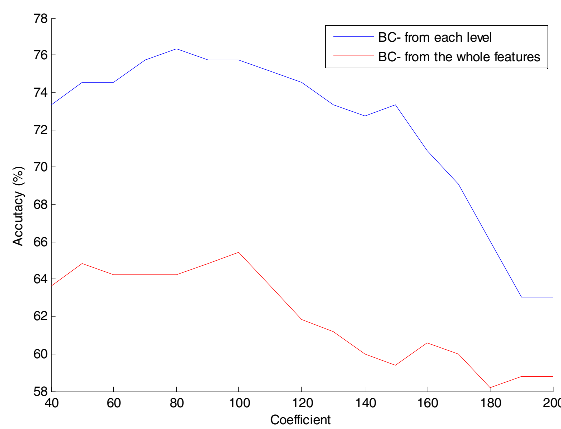



Figure 11. Comparison between two methods using Symelt function for different coefficients number with 4 levels of decomposition (YALE) database

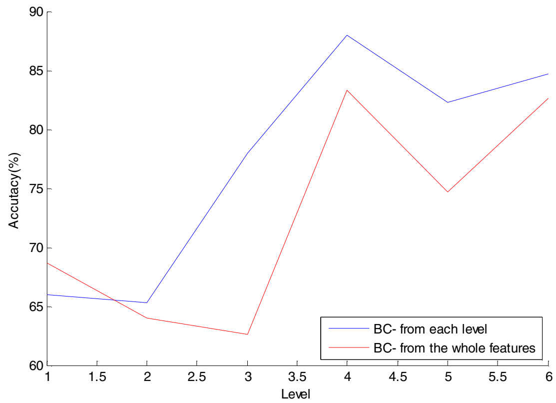



Figure 12. Comparison between two methods using Symelt function with fixes number of coefficients but with different numbers of decomposition levels (ORL) database.

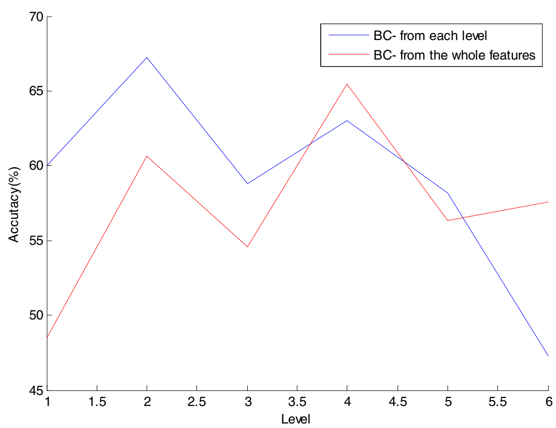



Figure 13. Comparison between two methods using Symelt function with fixes number of coefficients but with different numbers of decomposition levels (ORL) database.

## VI. CONCLUSION AND FUTURE WORK

In this paper, different seven wavelet functions were examined and different numbers of biggest coefficients were extracted as features vectors two times, firstly, from each level of decomposition, and secondly, by extracting the features from the all features dimensions to reduce the features dimensions. The experiments showed that Symelt wavelets produced best coefficients for classification process within suitable time due to their properties and it's appropriate for applications in face classification. Besides approximation coefficients were the only coefficients chosen for the classification process. Furthermore, it can be depicted from the experiments that the best number of decomposition levels is four. A set of images from ORL database and YALE data base were used in the experiments. Future works in connection to this paper will concentrate on finding other method to reduce the features dimensions since the redundancy in the features is unwanted in discrete wavelet transform (DWT) for classification process.

## REFERENCE

[1] M. Meselhy. Eltoukhy, I. Faye, and B. Belhaouari. Samir, "Breast cancer diagnosis in digital mammogram using multi-scale curvelet transform," Computerized Medical Imaging and Graphics, sciencedirect, November 2009.

[2] S. Adi, W. Sulaiman, "Texture classification using wavelet extraction: An approach to haptic texture searching," IEEE Trans, Monash, pp. 434-439, July 2009.

[3] L. Wang, J. Liu, Z. S. Li, "Texture classification using wavelet decomposition with Markov random field models," IEEE Trans, Brisbane. Qld, vol. 2, pp. 1613-1615, August 1998.

[4] S. Bo. Liu, Z. Y. Yuan, J.H. Zhao, X. Li. Wang, "An approach to face recognition based on wavelet decomposition SPCA and SVM," IEEE Trans, Baoding, vol. 2, pp. 889-893, July 2009.

[5] M. Bicego, U. Castellani, V. Murino, "Using Hidden Markov Models and Wavelets for face recognition," IEEE Trans, Verona Univ. Italy, pp. 52-56, Sept. 2003.

[6] Geetha. S. Rao,"Wavelet Analysis and Applications," New Age International, pp. 97, 2004.

[7] M.Misiti, Y. Misiti, G. Oppenheim and J. M. Poggi,"Wavelets and their Applications," Digital Signal and Image Processing Series, 2006.

[8] Gilbert. G. Walter, X. Shen, "Wavelets and Other Orthogonal Systems," Library of Congress Cataloging –in-publication data, Vol 2, pp. 28-48, 2000.