A SYSTEM TO DIAGNOSE THE ATHEROSCLEROSIS USING WAVELET TRANSFORMATION, PRINCIPAL COMPONENT ANALYSIS AND ARTIFICIAL NEURAL NETWORK

Sadık Kara, Fatma Dirgenali

Erciyes University, Department of Electrical Engineering, Kayseri, TURKEY

kara@erciyes.edu.tr, fdirgenali@tse.org.tr

Abstract: In this study, Doppler ultrasound signals were acquired from carotid arteries of 82 patients with atherosclerosis and 95 healthy volunteers. We have employed discrete wave transform (DWT) of Doppler signals and power spectral density graphics of these decomposed signals using Welch method. After that, we have performed Principles component analysis (PCA) for data reduction and ANN in order to distinguish between atherosclerosis and healthy subjects. After the training phase, testing of the artificial neural network (ANN) was established. The overall results show that 97.9% correct classification was achieved, whereas 2 false classifications have been observed for the test group of 97 people. In conclusion we are proposing a complimentary system that can be coupled to software of the ultrasonic Doppler devices. The diagnosis performances of this study show the advantages of this system: it is rapid, easy to operate, non invasive, inexpensive and making a decision without hesitation.

Introduction

The narrowing and hardening of the arteries named atherosclerosis has dramatic effects on blood pressure, resistance and blood flow. Resistance increases when radius decreases, as friction of blood flow against vessel wall increases. Therefore the circulation of blood flow is reduced. Development of a plaque also deforms the endothelial wall, increasing turbulent flow and increasing resistance. The hardening of the arterial walls increases resistance to flow [1-3].

Recent advances in the Doppler imaging technique have made possible evaluation of the temporal and spatial flow characteristics in the different portions of the arterial system, such as aorta, coronary, carotid and peripheral arteries. [4].

Doppler devices work by detecting the change in frequency of a beam of ultrasound that is scattered from targets that are moving with respect to the ultrasound transducer [5]. A Doppler signal is not a simple signal. It includes random characteristics due to the random phases of scattering particles present in the sample volume. The following Doppler equation is given by:

$$\Delta f_D = \frac{2\, v\, f_t\, \cos\theta}{c} \qquad (1)$$

Where v is the velocity of the blood flow, ft equals the frequency of the emitted ultrasonic signal, c is the the velocity of sound in tissue, Δf_D is the measured Doppler frequency shift, and θ equals the angle of incidence between the direction of blood flow and the direction of the emitted ultrasonic beam [6].

Formerly in Doppler systems, spectral estimation was traditionally achieved with the Fast Fourier Transform (FFT) method, afterwards the Short-Time Fourier Transform (STFT) has been the commonly used method for generating time-frequency representations of Doppler blood flow signals [7, 8]. This method requires that the signal being analyzed is stationary during a short time interval. Additionally, the spectral components occurring in a large interval will be smeared in the time domain, resulting in a decreased resolution in time [9].

Because of these drawbacks of STFT, as a new alternative technique, to accurately quantify the Doppler blood flow signal for the diagnosis of cardiovascular disease, discrete wavelet transform (DWT) was investigated. DWT was designed for nonstationary signals since it incorporates the concept of scale into the transform, which gives better time-frequency resolution: a compressed wavelet for analyzing high frequency details and a dilated wavelet for detecting lower frequency underlying trends [10].

The data reduction properties of Principal Component Analysis (PCA) are well known in connection with multivariate, time independent measurements [11]. In studies of hemodynamic disturbances resulting from atherosclerotic diseases and intracranial pathology, Martin et al [12] and Evans and co-workers [13] have shown the utility of PCA in association with blood flow signals measurements made with Doppler ultrasound. The application of ANNs has opened a new area for solving problems not reasonable by other signal processing techniques [14].

In this study we have employed DWT of carotid artery Doppler signals and power spectral density graphics of these decomposed signals using Welch method. After that, we have performed Principles component analysis (PCA) the aim of data reduction of decomposed signals' power spectral density graphics

ISSN: 1727-1983 © 2005 IFMBE

and Artificial Neural Networks (ANN) in order to distinguish between atherosclerosis and healthy subjects.

The results were compared with the arteriographic findings and showed that DWT-PCA and ANN architecture represents a significant improvement in diagnostic accuracy when compared with other techniques [15-17].

Materials and Methods

Carotid arterial Doppler ultrasound signals were acquired from left carotid arteries of 82 patients and 95 healthy volunteers. The subjects had no clinical and echocardiographic evidence of valvular disease or heart failure. The patient group included 48 males and 34 females with an established diagnosis of the early phase of atherosclerosis through coronary or aortofemoropopliteal (lower extremity) angiographies (mean age, 56 years; range, 42-73 years). Healthy volunteers were young non-smokers who seem to not bear any risk of atherosclerosis, including 53 males and 42 females (mean age, 22 years; range, 18-26 years).

Doppler signal acquisition was conducted by Toshiba PowerVision 6000 Doppler ultrasound unit in the radiology departments of Erciyes university hospital and Dinar state hospital. The system hardware was composed of Digital Doppler ultrasound unit that can work in the pulsed mode, linear ultrasound probe, input-output card and a personal computer. A personal computer (PC) was used for storage, displaying and spectral analysis of the acquired Doppler data [18].

2.1 Discrete Wavelet Analysis of carotid artery Doppler signals

For many signals, the low-frequency content is the most important part. It is what gives the signal its identity. The high-frequency content, on the other hand, imparts flavor or nuance. In wavelet analysis, we often speak of approximations and details. The approximations are the high-scale, low-frequency components and the details are the low-scale, high-frequency components of the signal. The DWT analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail coefficients. These coefficients represent different frequency subbands. DWT is used for the extraction of frequency features from the Doppler signals by decomposing the signal into multiple frequency subbands. The filtering process, at its most basic level, looks like Figure (1).

In the present study, carotid artery Doppler signals were decomposed with 'db3' Daebauchies wavelets to the fourth level. Four detail signals and one approximation signal at the fourth level were reconstructed from the coefficient matrices.

The detail wavelet coefficients of carotid artery Doppler signals obtained from patients with atherosclerosis and healthy subjects are given in Figs. 2 and 3, respectively. The horizontal axis is the number of samples, whereas the vertical axis is the amplitude.

2.2 Welch Method for spectral analysis of carotid artery Doppler signals

FFT based Welch method is defined as classical (Nonparametric) method. In Welch method, signals are divided into overlapping segments and each data segment is windowed [19].

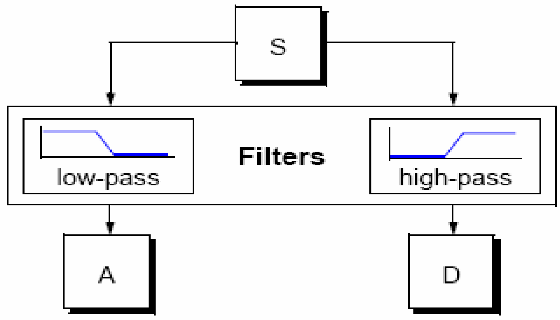

Figure 1. The filtering process

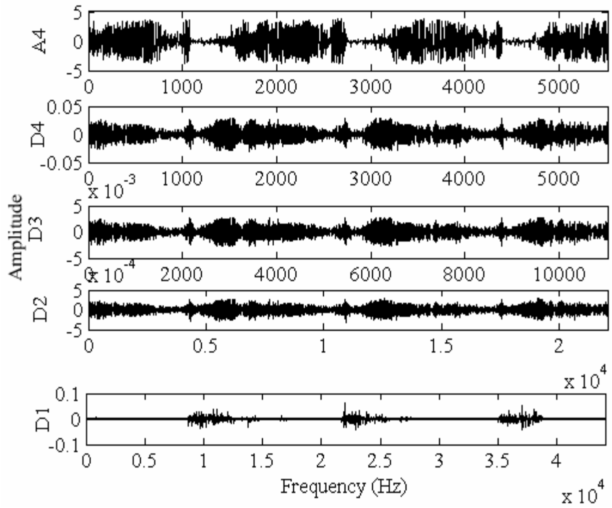

Figure 2. Decomposition of carotid artery Doppler signal for patient with atherosclerosis subject with 'db3'

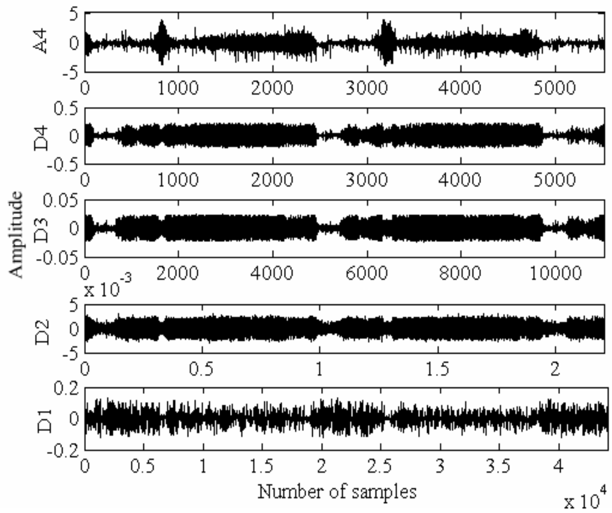

Figure 3. Decomposition of carotid artery Doppler signal for normal subject with 'db3' wavelet

ISSN: 1727-1983 © 2005 IFMBE

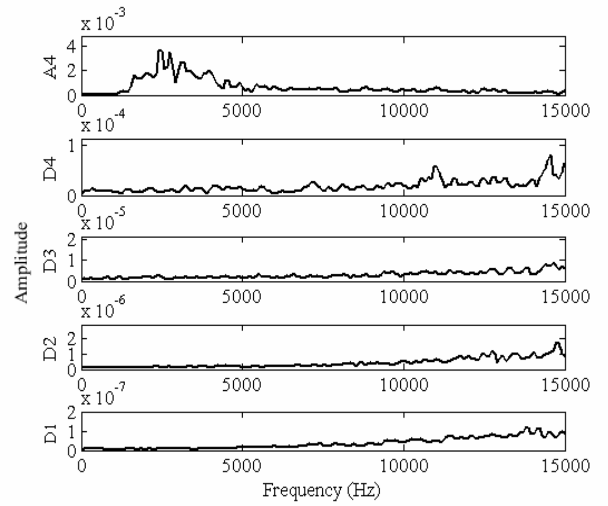

Figure 4. Power spectral density graphics of a decomposed atherosclerotic Doppler signal

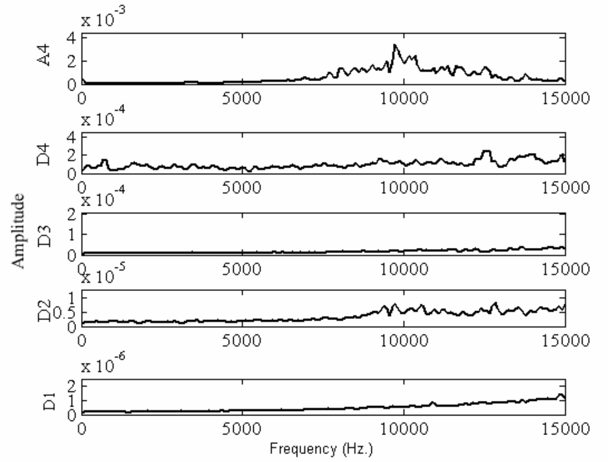

Figure 5. Power spectral density graphics of a decomposed normal Doppler signal

In this study power spectral density calculations for each of these signals were made using Welch method. Hanning window of 256 samples with an overlap of 128 samples are used in the Welch method. Power spectral density graphics for patients with atherosclerosis and healthy subjects can be seen in Figures 4 and 5 respectively.

The horizontal axis is the frequency, whereas the vertical axis is the amplitude of the power spectral density. All calculations were computed using Matlab software package.

2.3 Principal component analysis (PCA)

PCA was used to make an ANN system more effective. For this aim, before classifying with ANN, PCA method was used for data reduction of decomposed signals' power spectral density matrices. Therefore, each power spectral density matrices were represented as a vector consists of 5 samples.

PCA is based on the assumption that most information about classes is contained in the directions along which the variations are the largest. The most common derivation of PCA is in terms of a standardised linear projection which maximises the variance in the projected space [61, 62].

2.4 Artificial Neural Networks

An ANN is an information processing system where information spreads parallel on. An ANN can determine its conditions and adjust itself to provide different responses by using inputs and desired outputs, which are provided to the system. The most important thing about an ANN is that it works as a system which will eventually help the physicians on the decision making process about the existence of the disease. An ANN is trained with the available data samples to explore the relation between inputs and outputs, so that you can reach the proper and accurate outputs when you input some new data [21].

A Multilayer feed forward ANN was implemented in the MATLABsoftware package. This choice is appropriate for solving pattern classification problems where supervised learning is implemented with a Levenberg Marquard **(**LM) backpropagation algorithm. A LM back propagation neural network was used for the interpretation of PCA waveforms depicted from decomposed signals' power spectral density matrices. The advantage of using this type of ANN is the rapid execution of the trained network, which is particularly advantageous in signal processing applications. ANN training is usually formulated as a nonlinear least-squares problem.

ANN underwent supervised learning to perform successful pattern recognition of the carotid artery Doppler signals. During supervised learning, the ANN was trained on input vectors and the target output vectors with which it is required to associate the input vectors. With sufficient training, the ANN should be able to classify correctly previously unseen input vectors. In order to find the optimum structure that yields the minimum mean square error in testing, the network is iterated for single and double hidden layers with combinations of one to twenty neurons in each layer. For each layer combination, the target mean square error was set to $0.1 * 10^{-4}$ and the epoch number was taken as 20.

The train input data set consisted of 42 normal and 38 patients (80 sets $*$ 5 samples), while the test data set was made of 53 normal and 44 patients (97 sets $*$ 5 samples). The minimum training and testing errors were accomplished with the combination of two hidden layer consisting of 18-7 neurons. Hidden layer sigmoidal function and output layer linear function was used. The training parameters and structure of the MLP used in this study are as shown in Table 1.

The performance of the ANN algorithms were assessed by the following measures as seen Table 2 [22]

1) True Positive (TP): The ANN identifies an input as a patient with atherosclerosis diagnosed by the clinicians.

2) True Negative (TN): The ANN identifies an input as a normal that was labeled as a healthy by the expert clinicians.

Table 1. ANN architecture and training parameters

ARCHITECTURE		TRAINING PARAMETERS	
The number of layers	4	**Training Algorithm**	LM
The number of neuron on the layers	Input : 5 First Hidden: 18 Second Hidden: 07 Output : 1	**Mean-square error**	$5.80*10^{-5}$
		Mean absolute error	0.0010
Activation Functions	Tan-sigmoid	**Epoch number**	9

Table 2. Test results

GROUP TYPE	REAL NUMBER OF SUBJECTS IN THE TEST GROUP	TRUE	FALSE
Negative (Normal)	53	52(%98.1)	1(%1.9)
Positive (Subnormal)	44	43(%97.7)	1(%2.3)
TOTAL	97	95(%97.9)	2 (%2.1)
Sensitivity (SEN) = %97.7			
Specifity (SPE) = %98.1			
Average Detection Rate (ADR).= %97.9			

3) False Positive (FP): The detection of a atherosclerosis that was labeled as a normal by the expert.

4) False Negative (FN): The detection of a normal that was labeled as an atherosclerosis patient by the expert.

The performance of the classifier is also assessed in terms of sensitivity and specificity as follows [23].

1) Sensitivity (SEN): A measure of the ability of the classifier to detect atherosclerosis.

Sensitivity = (TP)/(TP+FN) %

2) Specificity (SPE): A measure of the ability of the classifier to specify normal.

Specificity = TN/(TN+FP) %

3) Average Detection Rate (ADR): The average of sensitivity and specificity.

Average Detection Rate (ADR) = (Sensitivity+Specificity)/2 %.

Results

After the training phase, testing of the LM backpropagation neural network was established. The data, which has not been used as an input to the network, was applied to the network for testing the network performance. After the process, normal (healthy) subjects are classified correctly with 98.1% and incorrectly with 1.9% (Table 2). Subjects having atherosclerosis are classified correctly with 97.7% and incorrectly with 2.3%.

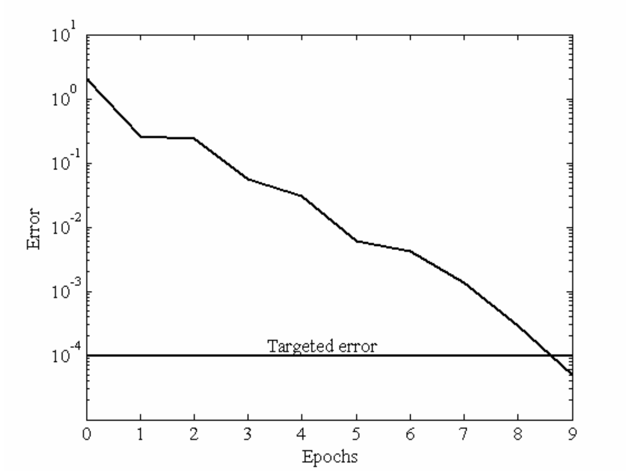

Figure 6. Variation of the error rate with respect to the epoch number in the ANN

The sizes of mean square error (MSE) and mean absolute error (MAE) can be used to determine how well the network output fits desired output, a testing MSE of 0.00103 and MAE of 0.00217 was observed for our optimized MLP feed forward network with a training MSE of $5.80*10^{-5}$ (Figure 6) and MAE of 0.0010. As seen in Table 2, 97.9% success rate of classification was accomplished with the designed feature extraction and the neural network structures. Final results were classified as normal and subnormal. There was one false classification in the negative group, while 52 subjects were correctly recognized as healthy. In the positive (patient) group, only one subject was misclassified, and 43 people were accurately classified as diseased. The overall results show that 97.9% correct classification was achieved, whereas 2 false classifications have been observed for the group of 97 people in total. With these results (the values of statistical parameters), this network has about 97.7% SEN, 98.1% SPE and ADR is calculated to be 97.9% (Table 2).

Discussion and Conclusion

The fuzzy appearance of the Doppler signals makes physicians suspicious about the existence of diseases and causes false diagnosis. Our technique gets around this problem using DWT-PCA-ANN architecture to decide and assist the physician to make the final judgment in confidence. Obtained results shown that, ANN classified Doppler signals successfully. However, training time and processing complexity were reduced.

We are proposing a complimentary system that can be coupled to software of the ultrasonic Doppler devices. The benefit of the system is to assist the physician to make the final decision without hesitation. The limitation of our proposed ANN structure is that the classification is realized based solely on the presence of abnormality with the carotid artery Doppler signals. However, we are projecting to also sort out the diseased group based on the source of the atherosclerosis problem. The other limitation of this system is that the position of the ultrasound probe, which used for data

ISSN: 1727-1983 © 2005 IFMBE

acquisition from the carotid artery, must be taken into consideration by physician.

In this study, we believe that this research developed a system for the interpretation of the carotid artery Doppler signals using ANN. The stated results show that the proposed method can make an effective interpretation. The diagnosis performances of this study show the advantages of this system: it is rapid, easy to operate, non invasive and inexpensive. This system is of the better clinical application over others, especially for earlier survey of population.

References

[1] HOSKINS P.R., MCDICKEN W.N., ALLAN P.L., (2000), 'Haemodynamics and blood flow'. In: ALLAN PL., DUBBINS PA., POZNIAK MA., MCDICKEN WN., (Ed.), *Clinical Doppler Ultrasound. London,*: pp.27-38

[2] SCHOEN F.J., COTRAN R.S., (1999), 'Blood vessels Pathologic basis of disease' *W.B. Saunders Company Press, (6th ed.)*, pp. 493-542.

[3] LIBLEY P., (2001), 'Prevention and treatment of atherosclerosis', In: by BRAUNWALD E, FAUCI AS, KASPER DL, HAUSER SL, LONGO DL, JAMESON JL (Ed), *Harrison's Principles of Internal Medicine. McGraw-Hill, (15th Ed)*, pp. 1382-1385

[4] EVANS D., (2000), 'Doppler Signal Analysis', *Ultrasound in Med. & Biol*., **26**, pp. 13–15.

[5] UBEYLI E., GÜLER I., (2004), 'Spectral analysis of internal carotid arterial Doppler signals using FFT, AR, MA, and ARMA methods', *Comp. in Biol. and Med.* **34**, pp. 293–306.

[6] SAINI V.D., NANDA N.C., MAULIK D., (1993), 'Basic principles of ultrasound and Doppler effect' *Doppler echocardiography, Philadelphia, London: Lea & Febiger.*

[7] CANNON S.R., RICHARDS K.L., ROLLWITZ W.T., (1982), 'Digital Fourier techniques in the diagnosis and quantification of aortic stenosis with pulse-Doppler echocardiography', *J. Clin. Ultrasound*, **10**, pp.101–107.

[8] KALMAN P.G., JOHNSTON K.W., ZUECH P., KASSAM M., POOTS K., (1985), 'In vitro comparison of alternative methods for quantifying the severity of Doppler spectral broadening for the diagnosis of carotid arterial occlusive disease', *Ultrasound Med Biol*, **11**, pp.435–40.

[9] GUO Z., DURAND LG., LEE H.C., (1994), 'Comparison of time-frequency distribution techniques for analysis of simulated Doppler ultrasound signals of the femoral artery', *IEEE Trans. Biomed. Eng.*

[10] LI J.H., PENG H., DU Y.Y., LI J.H., WANG W.H., (2000), 'Analysis of Doppler ultrasound blood flow signal basing on wavelet transform', *Northeastern University*, **21**, pp.487–489.

[11] MORRISON D.F., (1976), 'Multivariate statistical methods', *New York: Mc Graw – Hill*.

[12] SHERIFF S.B., BARBER D.C., MARTIN T.R.P., (1982) 'Use of principal component factor analysis in the detection of carotid artery disease from Doppler ultrasound', *Med. Biol. Eng. Comut*., **20**, pp.351-356

[13] EVANS D.H., ARCHER L.N.J., LEVENCE M.I., (1985), 'The detection of abnormal neonatal cerebral haemodynamics using principal component analysis of the Doppler ultrasound waveform', *Ultrasound. Med. Biol.*, **11**, pp. 441-449.

[14] BAXT W.G., (1990), 'Use of an ANN for data analysis in clinical decision making, the diagnosis of acute coronary occlusion', *Neural Comput.,* **2**, pp. 480-489.

[15] GUO Z., (1994), 'Comparison of Time-Frequency Distribution Techniques for Analysis of Simulated Doppler Ultrasound Signals of the Femoral Artery', *IEEE Transaction on Biomedical Engineering,* 41, pp. 332-342.

[16] ISABEL A., (1999), 'Artificial Neural Network Analysis Of Common Femoral Artery Doppler Shift Signals: Classification Of Proximal Disease', *Ultrasound in Med. & Biol.,* **24**, pp.735–743.

[17] VAN ASTEN W.N., Beijneveld W.J., Pieters B.R., et al., (1991), 'Assessment of aortoiliac obstructive disease by Doppler spectrum analysis of blood flow velocities in the common femoral artery at rest and during reactive hyperemia', *Surgery*, **109**, pp.633–639.

[18] KARA S., (1995), 'A study of Mitral and Tricuspid valve blood flows by Autoregressive spectral analysis method and Doppler Unit, Thesis of Doctorate', *Institute of Science of Erciyes University Press*.

[19] GIRAULT J.M., KOUAME D., OUAHABI A., PATAT F., (2000), 'Micro-emboli detection: an ultrasound Doppler signal processing viewpoint', *IEEE Trans. Biomed. Eng.*, **47**, pp.1431–1439.

[20] WANG X., PALIWAL K.K., (2003), 'Feature extraction and dimensionality reduction algorithms and their applications in vowel recognition', *The Journal of The Pattern Recognition Society*, **36**, pp. 2429-2439.

[21] SIMPSON P.K., (1989) 'Artificial Neural Systems', *Pergamon Press*.

[22] TARASSENKO L., KHAN Y. U., HOLT M. R. G., (1998), 'Identification of inter-ictal spikes in the EEG using neural network analysis', *Inst. Elect. Eng. –Proc. Sci. Meas. Technol.*, **145**, pp.270-278.

[23] PANG C. C., UPTON A. R., SHINE G., KAMATH M. V., (2003), 'A comparison of algorithms for detection of spikes in the Electroencephalogram', *IEEE Trans. Biomed. Eng.*, **50**, pp.521 – 526.

