FAST FOURIER TRANSFORM AND AUTOREGRESSIVE ANALYSIS OF MITRAL VALVE DOPPLER SIGNALS AND CLASSIFICATION BY FUZZY CLUSTERING

N. Barışçı[1], N. Bayrakcı[2]
[1]Kırıkkale University, Kırıkkale, Turkey; [2]Turk Telecom, Kırıkkale, Turkey

Corresponding Author: N. Barışçı, Kırıkkale University, Dept. of Computer Engineering
Kampus, 71450 Kırıkkale, Turkey; nbarisci@kku.edu.tr

Abstract. Ultrasonic Doppler blood flow measurement systems are widely used and highly successful in non-invasive techniques to detect heart diseases. In addition to estimates of stroke volume, the peak velocity and maximum acceleration of aortic blood may yield additional information on the inotropic state of myocardium. Continuous wave Doppler devices are most commonly used to provide velocity, not blood flow. With a pulsed Doppler system, the measurement of both arterial diameter and flow velocity are possible.
Today, quickly and trusty diagnosis of various diseases has great important. For this purpose, an additional diagnosis tool by artificial intelligence is developed to the aid of expert medical staff. In this work, cardiac Doppler signals recorded from mitral valve of 75 patients were transferred to a personal computer by using a 16 bit sound card. The fast Fourier transform (FFT) and Autoregressive (AR) method analysis was applied to the recorded signal from each patient and obtained FFT and AR parameters. Further these parameters were classified by using Fuzzy clustering algorithm. Thus, an additional diagnosis tool is developed for the aid of expert medical staff.
Our finding demonstrated that 90.67% correct classification rate was obtained from fuzzy clustering with FFT parameters and 94.67% correct classification rate was obtained from fuzzy clustering with AR parameters.

1 Introduction

Doppler blood flow ultrasonography (US) methods have been widely used in several medical practices effectively. Both pulse wave and continuous wave Doppler devices transmit ultrasonic waves into blood and receive some part of them which are reflected by red blood cells. There is a direct relationship between flow velocity of the blood and Doppler difference frequency which is obtained after the demodulation of transmitted signal with reflected signal. Some spectral analysis methods are applied to the Doppler signal in order to obtain medical information by taking into consideration the relationship between the Doppler signal and flow velocity of the blood. As a result, the sonograms that show the change of Doppler spectrum by time are obtained [1].

Heart consists of four cardiac chambers, namely two atriums and two ventricles. The mitral valve, examined in this study, is between the left atrium and the left ventricle. Mitral valve is the part of heart which prevents blood from flowing backward during systole, acting passively. The abnormality named as the mitral failure is a result of anatomic disorder of mitral valve. This abnormality causes heart valve not to open and close properly. As a result, the volume of blood that is pumped to the extremities decreases and it causes some other problems to occur, such as stenosis. In the case of mitral insufficiency, the blood in ventricle is released backward even if the mitral valve is closed.

Doppler signals obtained from heart valves have been analyzed by using different spectral analysis methods such as FFT, autoregressive, moving average, autoregressive-moving average and wavelet methods [2–5]. The output of a spectral analyzer is usually represented as a sonogram which illustrates the time variation of the spectral characteristics of the Doppler signal. A number of parameters related to the blood flow may be extracted from the sonogram and these are of high clinical value.

In the artificial intelligence domain, the studies related with the classification of heart data have gained great importance. In the literature, there are also studies related with the classification of several parameters of heart diseases with fuzzy algorithm and artificial neural network [6-11].

Furthermore, there are some medical applications of fuzzy clustering algorithm in the literature [12-14].

In this study, mitral valve Doppler signals obtained from 75 patients, are examined by taking into consideration of their spectral parameters (25 of these patients have mitral insufficiency, 25 have mitral stenosis, and the rest is normal). The FFT and AR parameters of mitral valve Doppler signals obtained from FFT and Burg AR analysis are recognized in fuzzy clustering algorithm and in this way diagnosis of mitral valve Doppler signals are performed correctly and rapidly. The purpose of this study is to enable fuzzy clustering algorithm aid to the Physician in diagnosis.

2 Material and Method

Ultrasonic Doppler system is based on the frequency change (f_d) of an emitted ultrasonic wave after its backscattering by moving structures. The frequency change is a function of the emission frequency (f_0), the velocity of moving structure (v) and the angle (θ) between the ultrasonic beam and the direction of blood flow:

$$f_d = 2 f_0 v Cos\theta / c \tag{1}$$

where c is the mean propagation velocity of ultrasound within tissues.

The measurement system consists of five functional blocks as shown in Figure 1. These are 2.50 and 3.75 MHz ultrasound transducers, an analog Doppler unit, Sony recorder, an analog/digital interface board (Sound Blaster Pro-16 bit), and a personal computer. The analog Doppler unit is capable of operating in both continuous and pulse wave modes. The Doppler unit is also equipped with an imaging facility that makes it possible to focus the sample volume to a desired location within the cardiac chambers. In this study, the Doppler unit is operated in the pulse mode and the sample volume is located at the ventricular side of the valve orifice. The measurement angle is taken as 60˚.

In this research, mitral valve Doppler signals were recorded from 75 subjects. The group consisted of 35 females and 40 males with ages ranging from 26 to 73 years and a mean age of 52±1 years. Data was obtained from Kırıkkale High Expertise Hospital in Turkey.

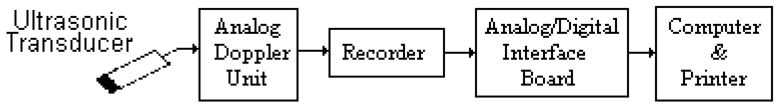

Figure1. Block diagram of measurement system

The signal at the output of the analog Doppler unit is recorded by a Sony recorder. The recorded signal is then sampled and digitized into 16-bit data packets by using an A/D interface board. The digital data are then stored on the hard disk of the PC. The interface board offers a range of sampling frequencies. The data stored as a sound file on the hard disk of the PC are converted to a text file. FFT and AR parameters are then obtained from this text file using spectral analysis software developed in MATLAB. Then, FFT and AR parameters are introduced to fuzzy clustering algorithm.

2.1 Fast Fourier transform method

In order to take the FFT of finite cardiac Doppler signal, an integer power of 2, it must be framed such as 64, 128, 256. Windowing is applied to the frequency spectrum of the frame. Windowing prevents the nonexisting frequency components from appearing in the spectrum. In addition, zero padding is applied to the same signal after the windowing process. This creates a certain overhead on the process, although it increases readability of the spectrum [5,15].

Discrete Fourier transform of a discrete time periodic signal is defined as the following:

$$X_k = \sum_{n=0}^{N-1} x(n) e^{(-jkn(2\pi/N))} \tag{2}$$

where X_k is expressed as discrete Fourier coeffcients, N is the frame size and $x(n)$ is the input signal on time domain. To obtain the frequency spectrum of this signal, logarithmic values of the squares of absolute values of X_k are found as shown below

$$P(k) = 10\log\left|X_k\right|^2 \tag{3}$$

The frequency content of the signal will determine the sampling rate to be used, since the maximum frequency (f_m) analyzed is half of the sampling frequency (f_s). Since f_m is usually not known a priori and depends upon the vessel on which the measurements are performed, it is useful to implement the system such that it is capable of operating at various sampling frequencies. In this work, measurements were performed at a sampling frequency of 10 kHz, so that the frequency aliasing, in the case of stenosis, would be avoided.

The number of signal samples required to form a frame depends heavily on the stationary condition of the signal. In general, the Doppler signal is non stationary. The signal may be assumed to be stationary for 10 ms or greater time periods if the flow is laminar and the velocity is not very high. However, this assumption is not valid for high velocity turbulence flows, such as the flow encountered at the aortic arch, where the Doppler spectrum changes very rapidly. In this case frame length should be shortened to validate the above assumption. On the other hand, very short frame lengths may yield statistically poor spectral resolution. Therefore, selection of frame length is an important factor in Doppler spectral analysis. The frame length used in this study is 128.

2.2 AR method

The model-based (parametric) methods are based on modeling the data sequence *x(n)* as the output of a linear system characterized by a rational system. In the model-based methods, the spectrum estimation procedure consists of two steps. Given the data sequence *x(n),* $0 \le n \le N-1$, the parameters of the method are estimated. Then, from these estimates, the PSD estimate is computed. AR method is the most frequently used parametric method because estimation of AR parameters can be done easily by solving linear equations. Since Burg AR method is computationally efficient and yields stable estimates, PSD estimates of mitral valve Doppler signals are obtained by using Burg AR method. The method is based on the minimization of the forward and backward prediction errors and on estimation of the reflection coefficient. The forward and backward prediction errors for a p^{th}-order model are defined as:

$$\hat{e}_{f,p}(n) = x(n) + \sum_{i=1}^{p} \hat{a}_{p,i} x(n-i), \qquad n = p+1,\ldots,N \tag{4}$$

$$\hat{e}_{b,p}(n) = x(n-p) + \sum_{i=1}^{p} \hat{a}^{*}_{p,i} x(n-p+i), \; n = p+1,\ldots,N \tag{5}$$

The AR parameters are related to the reflection coefficient $\hat{k}_p$ by

$$\hat{a}_{p,i} = \begin{cases} \hat{a}_{p-1,i} + \hat{k}_p \hat{a}_{p-1,p-i}, & i = 1,\ldots,p-1 \\ \hat{k}_p, & i = p \end{cases} \tag{6}$$

Burg method considers the recursive-in-order estimation of $\hat{k}_p$ given that the AR coefficients for order $p-1$ have been computed. The reflection coefficient estimate is given by

$$\hat{k}_p = \frac{-2\sum_{n=p+1}^{N} \hat{e}_{f,p-1}(n)\hat{e}^{*}_{b,p-1}(n-1)}{\sum_{n=p+1}^{N}\left[\left|\hat{e}_{f,p-1}(n)\right|^2 + \left|\hat{e}_{b,p-1}(n-1)\right|^2\right]} \tag{7}$$

The Burg method for estimating the parameters of AR method is computationally efficient and yields a stable AR method. On the other hand, the Burg method is suboptimal in that it estimates the *n* reflection coefficients by decoupling an *n*-dimensional minimization problem into the *n* one-dimensional minimizations. This is in contrast to the least squares AR method, in which AR coefficients are found by an *n*-dimensional minimization [15-17].

One of the most important aspects of the use of AR method is the selection of the order *p*. Much work has been done by various researchers on this problem and many experimental results have been given in the literature such as the papers presented by Akaike. One of the better known criteria for selecting the model order have been proposed by Akaike [18], called the Akaike information criterion (AIC), is based on selecting the order that minimizes

$$AIC(p) = \ln \hat{\sigma}^2_{wp} + 2p/N \tag{8}$$

where $\hat{\sigma}^2_{wp}$ is the estimated variance of the linear prediction error. Note that the term $\hat{\sigma}^2_{wp}$ decreases and therefore $\ln \hat{\sigma}^2_{wp}$ also decreases as the order of AR method is increased. However, $2p/N$ increases with an increase in *p*. In this situation, a minimum value is obtained for some *p*[18]. In this study, model order is taken as 10 for mitral valve signals by using Equation (8).

2.3 Fuzzy clustering

Clustering is the method of classifying data or objects within a data universe into different subgroups based on the similarities and the dissimilarities between objects. Clustering algorithms essentially perform two basic functions. They try to keep all objects, which are very close to each other in one subgroup or cluster. At the same time they try to maximize the dissimilarities between the clusters. Fuzzy c-means clustering (FCM) utilizes an objective function approach to allow formation of clusters in multi-dimensional space, where a data point is allowed to belong to more than one cluster, with different membership degree [19-21].

Let the data set to be partitioned be $X = \{x1, x2, \cdots, x_m\} \subset R^n$, which was required to be partitioned into c partitions, c є {2, 3, · · ·, (m-1)}. The partitions were described by a partition matrix U of size $c\times m$ and each of its elements μ_{ik}, i=1, 2, · · ·, c, k=1, 2, · · ·,m represented the membership value of $x_k \in X$ in the ith cluster. Hence FCM utilized the set of fuzzy partitions, given as:

$$M_{FCM} = \{U \in [0,1]^{cm}\} \tag{9}$$

$$\sum_{i=1}^{c} \mu_{ik} = 1, \qquad k = 1,2,...,m \tag{10}$$

$$\sum_{k=1}^{m} \mu_{ik} > 0, \qquad i = 1,2,...,c \tag{11}$$

Hence FCM was defined as the following problem: given the data set X, any norm ||•|| on R^n, and a fuzziness parameter $q \in (1,\infty)$, minimized the objective function:

$$J_{FCM}(U,V;X) = \sum_{k=1}^{m}\sum_{i=1}^{c} \mu_{ik}^{q} \left\|x_k - v_i\right\|^2 \tag{12}$$

where $U \in M_{FCM}$ and $V = \{v_1, v_2, \ldots, v_c\} \subset R^n$, was the set of the cluster centers. The cluster centers were taken to be point prototypes. More the value of q, more the clusters became fuzzy in nature and less the value of q, the clusters became increasingly harder.

The objective function J is minimized when high membership values are assigned to pixels whose intensities are close to the centroid of its particular class, and low membership values are assigned to them when the pixel data are far from the centroid.

Taking the first derivatives of J in Eq.(12) with respect to μ_{ik}, v_i we can obtain the following necessary conditions to minimize the objective function J:

$$\mu_{ik}^{r+1} = \frac{1}{\sum_{j=1}^{c} (d_{ik}^2 / d_{jk}^2)^{1/(m-1)}}, 1 < m < \infty \tag{13}$$

$$d_{ik}^2 = (x_k - v_i)^T A (x_k - v_i) \tag{14}$$

$$v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m x_k}{\sum_{k=1}^{N} \mu_{ik}^m} \tag{15}$$

After initialization of the centroids, μ_{ik} and v_i are iteratively calculated until some stop criteria are reached. Finally, the segmentation can be obtained by the principle of maximum membership.

The FCM clustering algorithm is quite robust and has been previously demonstrated to provide accurate clustering result, even if the initial fuzzy partition matrix $U^{(0)}$ is poorly constructed.

3 Results and Discussion

In this study, mitral valve cardiac Doppler signals are obtained from 75 patients. 25 of these patients have mitral stenosis, 25 have mitral insufficiency and the rest is normal. These signals are analysed with FFT and AR methods for obtain parameters. After this analysis, 65 FFT parameters and 10 AR parameters are obtained. Then these parameters are applied to fuzzy clustering algorithm for classification of mitral valve Doppler signals. After classification process, results for FFT method are shown in Table 1.

Table 1 demonstrated the true and false classification values of mitral valve Cardiac Doppler signals. Here, 22 mitral stenosis were classified as true, 3 mitral stenosis were classified as false, and 23 mitral insufficiencies and normal were classified as true, 2 mitral insufficiencies and normal were classified as false.

Class	*Numbers of Class*	*Numbers of True Classified Class*
Mitral Stenosis	25	22
Mitral Insufficiency	25	23
Normal	25	23

Table 1. Results for FFT Method

The simplest classification problem is that of separating one-dimensional feature vectors into two groups. In this situation the only choice that needs to be made is where to locate the decision threshold. If there is no overlap between the magnitudes of the vectors obtained from patients belonging to the two classes, the threshold can simply be chosen to separate the classes completely. In general, the results from the two classes do overlap and so depending on where the threshold is placed some signals from normal subjects will be adjudged abnormal and/or some signals from abnormals will be adjudged normal. The best choice of threshold will then depend on a number of factors. There are two important measures of the performance of a diagnostic test; sensitivity (or true positive fraction) and specificity (or true negative fraction) which are defined as:

Sensitivity = Number of true positive decisions / Number of actually positive cases

and

Specificity = Number of true negative decisions / Number of actually negative cases

These measures are dependent since they are both affected by the position of the decision threshold and as the threshold is moved to increase sensitivity, so specificity decreases.

To measure the classification performance, Table 1 should be evaluated statistically. As it is seen Table 2, the specificity is 92 %, sensitivity is 90 % and the true classification rate is 90.67 % for FFT method in the statistical results of data.

Statistical parameters	Value (%)
Specificity	92
Sensitivity	90
True Classification rate	90.67

Table 2. Statistical Parameters for FFT Method

After the FFT parameters, AR parameters obtained from mitral valve Doppler signals were applied to fuzzy clustering algorithm. Classification results for AR method are shown in Table 3. As it is seen Table 3, 23 mitral stenosis were classified as true, 2 mitral stenosis were classified as false, and 24 mitral insufficiencies and normal were classified as true,1 mitral insufficiency and normal was classified as false.

To measure the classification performance for AR method, Table 3 should be evaluated statistically. As it is seen Table 4, the specificity is 94 %, sensitivity is 96 % and the true classification rate is 94.67 % for AR method in the statistical results of data.

Class	*Numbers of Class*	*Numbers of True Classified Class*
Mitral Stenosis	25	23
Mitral Insufficiency	25	24
Normal	25	24

Table 3. Results for AR Method

Statistical parameters	Value (%)
Specificity	94
Sensitivity	96
True Classification rate	94.67

Table 4. Statistical Parameters for AR Method

4 Conclusion

In this work, cardiac Doppler signals recorded from mitral valve of 75 patients were used to FFT and Burg AR spectral analysis. The classification systems were developed with the application of FFT and AR parameters of mitral valve Doppler signals in to fuzzy clustering algorithm.

Our findings demonstrated that 90.67% correct classification rate was obtained from fuzzy clustering with FFT parameters and 94.67% correct classification rate was obtained from fuzzy clustering with AR parameters.

These results demonstrate that the classification performance of fuzzy clustering with FFT parameters is lower than fuzzy clustering with AR parameters.

5 Acknowledgement

This work is supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK Project number 107E195). We thanks to Dr. Özgür Yaşar Akbal for supplying cardiac Doppler data.

6 References

[1] Evans, D.H., Mcdicken, W.N., Skidmore, R., Woodcock J.P.: *Doppler Ultrasound: Physics, Instrumentation and Clinical Applications*. John Wiley & Sons, Chicester, 2000.

[2] Güler, İ., Hardalaç, F., Müldür, S.: *Determination of aorta failure with the application of FFT, AR and wavelet methods to Doppler technique.* Comput. Biol. Med., 31 (2001), 229–238.

[3] Güler, İ., Kara, S., Güler, N.F., Kiymik, M.K.: *Application of AR and FFT spectral analysis to tricuspid and mitral valve stenosis.* Comput. Methods Programs Biomed, 49 (1996), 29–36.

[4] Güler, İ., Kara, S.: *Detection of mitral stenosis by pulsed Doppler flowmeter and autoregressive spectral analysis*, In: First Regional Conference on IEEE Engineering, Medical and Biological Society and Biomedical Engineering Society, India, 1995, 291–292.

[5] Kay, S.M., Marple, S.L.: *Spectrum analysis—a modern perspective*. Proc. IEEE, 69 (1981), 1380–1419.

[6] Akay, M, Akay, Y.M., Welkowitz, W.: *Neural networks for the diagnosis of coronary artery disease*. In: International Joint Conference on Neural Networks. ijcnn, Baltimore, 1992, 419-424.

[7] Turkoglu, I., Arslan, A., Ilkay, E.: *An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks.* Computers in Biology and Medicine, 33 (2003), 319-331.

[8] Machado, L.O., Musen, M.A.*: Sequential versus standard neural networks for pattern recognition: An example using the domain of coronary heart disease*. Computers in Biology and Medicine, 27 (1997), 267-281.

[9] Güler, I., Hardalac, F., Barışçı, N.: *Application of FFT analyzed cardiac Doppler signals to fuzzy algorithm*. Comput Biol Med., 32 (2002), 435-444.

[10] Güler, I., Hardalac, F., Ergün, U., Barışçı, N.: *Classification of Aorta Doppler signals using variable coded-hierarchical genetic fuzzy system*. Expert Systems with Applications, 26 (2004), 321-333.

[11] Hardalac, F., Barışçı, N., Ergün, U.: *Classification of aorta insufficiency and stenosis using MLP neural network and neuro-fuzzy system.* Physica Medica, 20 (2004), 155-165.

[12] Masulli, F., Schenone, A.: *A fuzzy clustering based segmentation system as support to diagnosis in medical imaging.* Artifical Intelligence in Medicine, 16 (1999), 129-147.

[13] Yang, M., Hu, Y., Lin, K., Lin, C.: *Segmentation techniques for tissue differentation in MRI of Ophthalmology using fuzzy clustering algorithms*. Magnetic Resonance Imaging, 20 (2002), 173-179.

[14] Assen, H., Danilouchkine, M., Frangi, A., Ordas, S., Westenberg, J., Reiber, J., Lelieveldt, B.: *SPASM: A 3D-ASM for segmentation of Sparse and arbitrarily oriented cardiac MRI data*. Medical Image Analysis, 10 (2006), 286-303.

[15] Proakis, J.G., Manolakis, D.G.: *Digital Signal Processing Principles Algorithms and Applications*. Prentice-Hall, Englewood Cliffs, NJ., 1996.

[16] Stoica, P., Moses, R.: *Introduction to Spectral Analysis*. Prentice-Hall, Englewood Cliffs, NJ., 1997.

[17] Kay, S.M.: *Modern Spectral Estimation: Theory and Application.* Prentice-Hall, Englewood Cliffs, NJ, 1988.

[18] Akaike, H.: *A new look at the statistical model identification*. IEEE Trans. Automat. Control AC, 19 (1974), 716–723.

[19] Bezdek,J. :P*attern recognition with fuzzy objective function algorithms*, Plenum, Newyork, 1981.

[20] Ross, T.J.: *Fuzzy logic with engineering applications. International Editions*, McGraw-hill Inc.; 1997.

[21] Runkler, T.A., Bezdek J.C.: *Alternating cluster estimation: a new tool for clustering and function approximation.* IEEE Fuzzy Syst, 7 (1999), 377-393.