**Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006**

FrA03.1

Multiple ECG Beats Recognition in the Frequency Domain Using Grey Relational Analysis

Chia-Hung Lin, Yi-Chun Du, Yung-Fu Chen, and Tain-Song Chen

Abstract— **This paper proposes a method for multiple ECG beats recognition using novel grey relational analysis (GRA). Converts each QRS complex to a Fourier spectrum from ECG signals, the spectrum varies with the rhythm origin and conduction path. The variations of power spectra are observed in the range of 0Hz-20Hz in the frequency domain. According to the frequency-domain parameters, GRA performs to recognize the cardiac arrhythmias including the supraventricular ectopic beat, bundle branch ectopic beat, ventricular ectopic beat, and fusion beat. The method was tested on MIT-BIH arrhythmia database. The results demonstrate the efficiency of the proposed non-invasive method, and also show high accuracy for detecting electrocardiogram (ECG) signals.**

Keywords—Grey Relational Analysis (GRA), Frequency Domain, Cardiac Arrhythmia.

I. Introduction

ECG signal is the non-invasive measurement for reflecting the internal status of heart and myocardium electric activity. By placing electrodes on the body surface, 12-lead electrocardiograph is used to record the electrical activity. The sequence of electrical signals provides symptomatic information for identifying cardiac arrhythmias. The measurement devices (Holter ECG) can record large amounts of signals. However, they do not automatically classify abnormalities and require off-line analysis from the record data. Designing non-invasive methods, signal processing, signal recognition, decision support, and human computer interface for stationary portable monitored devices has become an aided function for pattern-recognition tasks [1]-[2]. To ensure accurate detection, the diagnostic procedure requires automatic classification, high-performance computing, and easy to implement for heartbeats recognition.

In the literature, diagnostic methods have been applied to detection in conjunction with time domain, frequency domain, and time-frequency domain techniques. The QRS complex in ECG signals varies with the origination and the conduction path of the activation pulse. Various features from each heartbeat are extracted to detect arrhythmia waveforms. In the time domain, these features are amplitude parameters (QRS, ST), duration parameters (QRS, QT, PR), and combined parameters (Q/R ratio, S/R ratio) [3].

Chia-Hung Lin is with the Department of Electrical Engineering, Kao-Yuan University, Lu-Chu Hsiang, Kaohsiung 821, Taiwan.

Tain-Song Chen and Yi-Chun Du are with the Institute of Biomedical Engineering, National Cheng-Kung University, Tainan 70101, Taiwan.

Yung-Fu Chen is with Department of Health Services Management, China Medical University, Taichung 40402, Taiwan.

When the activation pulse does not travel through the normal conduction path, the QRS complex becomes wide, and the high-frequency components are the range of 0-20Hz frequencies. The spectrum has a maximum amplitude at 4Hz in ventricular tachycardia attenuated.

In the frequency domain, power spectra of QRS complexes are Fourier Transformed and are found in (VT), and its amplitude decreases as the frequency increases [4]. The frequencies of ventricular fibrillation (VF) are concentrated between 4Hz-7Hz [5]. In the time-frequency technique, wavelet transform (WT) has applied to extract the features of cardiac arrhythmias. VT and VF occupy different time-frequency representations. These techniques are robust to time-varying signal analysis, but it is not capable of recognition. Applying these symptomatic features, artificial intelligent (AI) approaches have been proposed to improve the classification of cardiac abnormalities including wavelet neural networks [1], artificial neural network (ANN) [3]-[4], [7]-[8], and Fuzzy hybrid neural networks [9]-[10].

To develop an aided tool, diagnostic algorithm is necessary to easy implement in the virtual instrument technology and small hardware device. The morphology of QRS complex varies in both normal and abnormal rhythms. Accurate diagnose is limited by the number of amplitude parameters. The WT is robust to time-varying signal analysis and it can point out occurrence time, but it is not capable of recognition. In order to cardiac arrhythmia classification, ANNs are applied in this research. The ANNs are well known for its learning and recognition ability. However, the limitations of ANNs are training process, determining a possible architecture, and network parameters assignment in the clinical environments. Considering these limitations, fast Fourier transform (FFT) is used to estimate frequency spectra. GRA is studied and proposed for heartbeat signals recognition. For adaptation application, the property of the GRA has a function of mathematical operation for processing numerical data or binary data, flexible pattern mechanism with add-in and delete-off features, and expandable or reducible without adjusting any parameter.

In this study, diagnostic procedure consists of two stages: first, the frequency spectra are computed by FFT; subsequently, the GRA based classifier is used to classify normal beat and six cardiac arrhythmias. The results show computational efficiency and accurate recognition.

II. Problem Description

An ECG signal represents the changes in electrical potential during the heartbeat as recorded with non-invasive electrodes on the limbs and chest; a typical ECG signal consists of the P-wave, QRS complex, and T-wave. QRS complex in ECG signals varies with origin and conduction path of the activation pulse in the heart. When the activation pulse originates in the atrium and travels through the conduction path, the QRS complex has a sharp and narrow deflection. Converts each QRS complex to a Fourier spectrum, the spectrum contains high-frequency components [4]. The QRS complex becomes broad and distortion due to the activation pulse originate in the ventricle and doesn't through the conduction path.

Frequency-domain analysis applies high-frequency and low-frequency ranges to discriminate ventricular rhythm, atrial rhythm, parasympathetic and sympathetic activity signals. For example, spectral analysis is the linear transform used to diagnose ventricular tachyarrhythmia. Power spectra of individual QRS complex are found significant differences from 0Hz to 20Hz frequencies. The spectrum in VT has maximum amplitude at 4Hz [4]. Spectral analysis is also used in the analysis of heart rate variability (HRV) signals. Power spectra in the 0.15Hz ~ 0.5Hz ranges reflect respiratory sinus arrhythmias and cardiac vagal activity, baro-receptor control is mediated by vagal and sympathetic systems in the 0.04Hz~0.15Hz ranges, and very low-frequency (≤0.04Hz) is related to thermoregulatory, vascular mechanisms, and rennin-angio tension systems [6]-[7].

Fourier analysis (FA) is a method for data analysis, and breaks up a signal into sinusoidal waves of various frequencies. For sampled vector data, FA performs the discrete Fourier transform (DFT). FFT is an algorithm for computing the DFT of a sequence. It is particularly useful in areas including signal and image processing, where its uses range from filtering, convolution, and frequency analysis to power spectrum estimation. In this paper, frequency-domain analysis has been used in the applications of signal process. The amplitudes of power spectra are computed by FFT. Frequency-domain parameters are used to multiple ECG beats recognition.

III. Grey Classification Procedure

The grey relational analysis (GRA) includes local relation and global relation analysis. GRA is a method to determine the relation of a discrete data to other sequence data. Based on similarity and dissimilarity, the relation is the relational measurement on attribute in different sequences [11]-[12]. For certain window duration, each QRS complex is extracted as V_{QRS}=[v_1, v_2, v_3, …, v_p, …, v_P], P is the number of sampled points, p=1, 2, 3, …., P. Frequency spectra are computed by the function $fft(\bullet)$ [15]

$$X=[x_1 \quad x_2 \quad \cdots \quad x_i \quad \cdots \quad x_n]=fft(V_{QRS}) \tag{1}$$

The discrete Fourier transform (DET) is found by taking the n-point fast Fourier transformation (FFT). If X is complex, compute the amplitude of the FFT of a sequence by the function $abs(\bullet)$

$$A=[a_1 \quad a_2 \quad \cdots \quad a_i \quad \cdots \quad a_n]=\frac{abs(X)}{\max[abs(X)]} \tag{2}$$

Each element a_i of A is the absolute value of the corresponding element of X. The element a_i, i=1, 2, 3…, n, is the selected amplitude of frequency spectrum. Each spectrum is normalized with the maximum amplitude. Let a sequence be the test sequence for comparison to other sequences, where the test sequence is $a_i(0)$, i=1, 2, 3, …, n, and K comparative sequence is $A(k)$=[$a_1(k)$, $a_2(k)$, $a_3(k)$, …, $a_i(k)$, …, $a_n(k)$], k=1, 2, 3, …, K, can be represented as

$$A_{test}=[a_1(0) \quad a_2(0) \quad \cdots \quad a_i(0) \quad \cdots \quad a_n(0)] \tag{3}$$

$$A_{comp}=\begin{bmatrix} A(1) \\ A(2) \\ \vdots \\ A(k) \\ \vdots \\ A(K) \end{bmatrix}=\begin{bmatrix} a_1(1) & a_2(1) & \cdots & a_i(1) & \cdots & a_n(1) \\ a_1(2) & a_2(2) & \cdots & a_i(2) & \cdots & a_n(2) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_1(k) & a_2(k) & \cdots & a_i(k) & \cdots & a_n(k) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_1(K) & a_2(K) & \cdots & a_i(K) & \cdots & a_n(K) \end{bmatrix} \tag{4}$$

Compute the absolute deviation of test sequence A_{test} and k comparative sequence $A(k)$ by

$$\Delta d_i(k)=|a_i(0)-a_i(k)| \tag{5}$$

The deviation matrix ΔD can be represented as

$$\Delta D=\begin{bmatrix} \Delta d_1(1) & \Delta d_2(1) & \cdots & \Delta d_i(1) & \cdots & \Delta d_n(1) \\ \Delta d_1(2) & \Delta d_2(2) & \cdots & \Delta d_i(2) & \cdots & \Delta d_n(2) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \Delta d_1(k) & \Delta d_2(k) & \cdots & \Delta d_i(k) & \cdots & \Delta d_n(k) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \Delta d_1(K) & \Delta d_2(K) & \cdots & \Delta d_i(K) & \cdots & \Delta d_n(K) \end{bmatrix} \tag{6}$$

The grey grades $f(k)$ can be calculated as [13]

$$f(k)=\exp[-\xi(\frac{\sqrt{\sum_{i=1}^{n}(\Delta d_i(k))^2}}{\Delta d_{max}-\Delta d_{min}})^2], \quad \xi\in(0,\ 5) \tag{7}$$

$$\begin{cases} \Delta d_{mim} = \underset{\forall k}{mim}[\underset{\forall i}{mim}\,\Delta d_i(k)] & (8) \\ \Delta d_{max} = \underset{\forall k}{max}[\underset{\forall i}{max}\,\Delta d_i(k)] & (9) \end{cases}$$

Δd_{min} and Δd_{max} are the minimum and maximum values of the matrix ΔD respectively. Parameter ξ is the recognition coefficient with interval (0,10), ξ=5 was chosen in this study. The grey grades are used the Euclidean distances (ED) to measure relationship between the reference sequence data and comparative sequences data. The grey grades $f(k)$ are inversely proportional to the distances as $ED\to\infty$, $f(k)\to 0$ and $ED\to 0$, $f(k)\to 1$. If the test vector X_{test} is similar to any comparative vector $X(k)$, the grade $f(k)$ will be a maximum value. Then find the maximum grade, which can be represented as

$$f_{max}=f(k^*)=max[f(1),\ f(2),\ \ldots,\ f(k),\ \ldots,\ f(K)] \quad (10)$$

$$\gamma(k)=\begin{cases}1, & f(k)=f_{max}\\ 0, & f(k)\neq f_{max}\end{cases} \quad (11)$$

where k^* is the criterion index in the K comparative sequences; index $\gamma(k)\in\{0,1\}$. For m classes classification, the associated class for X_{test} could be expressed as weighting factor $w_{kj}\in\{0,1\}$, where m is the total number of possible classes, j=1, 2, 3, …, m. If X_{test} belongs to class j, the weighted factor w_{kj} equals to one, and the rest factors are zero as equation (12). The final grey grade g_j that an unknown vector X_{test} belongs to class j can be represented by equation (13)

$$w_{kj}=\begin{cases}1, & k\in Class\ \ j\\ 0, & k\notin Class\ \ j\end{cases},\quad j=1,\ 2,\ 3,\ \ldots,\ m \quad (12)$$

$$g_j=[\gamma(1)\quad \gamma(2)\quad \cdots\quad \gamma(k)\quad \cdots\quad \gamma(K)]\cdot\begin{bmatrix}w_{1j}\\ w_{2j}\\ \vdots\\ w_{kj}\\ \vdots\\ w_{Kj}\end{bmatrix}=\sum_{k=1}^{K}\gamma(k)w_{kj} \quad (13)$$

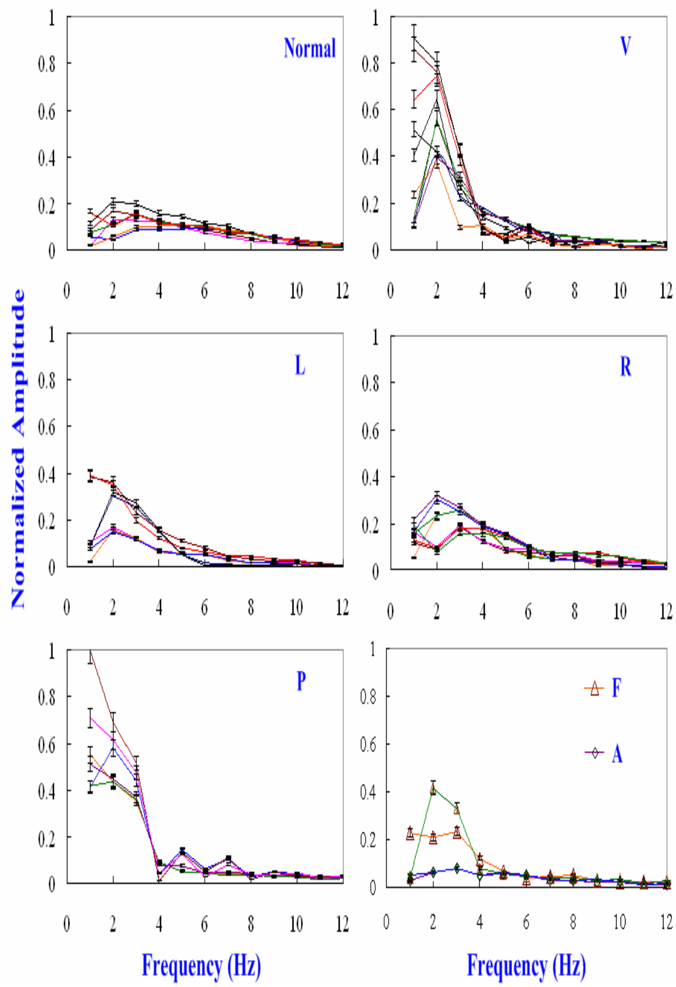
The dimension of grey relational vector $\Gamma=[f(1), f(2), f(3), \ldots, f(k)], \ldots, f(K)]$ can be reduced from K-dimension to m-dimension. The output can be represented

$$\boldsymbol{G}=[g_1,\ g_2,\ \ldots,\ g_j,\ \ldots,\ g_m],\quad g_j\in\{0,1\} \quad (14)$$

The GRA based classifier is used to m class problem (m=7 in our study). These seven classes include the normal beat (•), premature ventricular contraction (**V**), atrial premature beat (**A**), right bundle branch block beat (**R**), left bundle branch block beat (**L**), paced beat (**P**), and fusion of paced and normal beat (**F**).

IV. Frequency-domain features

The QRS complex of the ECG is an important information in heart-rate monitoring and cardiac diseases diagnosis. The R-waves are detected by a peak detection algorithm, which begins by scanning for local maxima in the absolute value of ECG data. For certain window duration, the searching continues to look for a larger value. If this search finishes without finding a larger maximum, the current maximum is assigned as the R peak [4]. Centered on the detected R peak, the QRS complex portion is extracted by applying a window of 280ms, and P-wave and T-wave are excluded by this window duration. Based on 360 sampling rate, 100 samples can be acquired around the R peak (Sampling point n=100, 50 points before and 50 points after). After sampling and analog-to-digital conversion, individual QRS complexes are extracted. Then frequency spectra of each QRS complex are computed by FFT. The spectrum varies with different cardiac arrhythmias, and power spectra are observed in the frequency range from 1Hz to 20Hz. The spectra are plotted and analyzed as shown in Figure 1, and all amplitudes are normalized with maximum amplitude. The amplitudes decrease as the frequency increases, and rapidly vanishes above 12Hz. Frequency components from 1Hz to 12Hz are selected for multiple ECG beats recognition. The frequency spectra are not disturbed by high frequency components above 20Hz such as power-line interference (50Hz/60Hz) and muscle noise, and very low-frequency components (<1Hz) such as baseline drift and breath [4].



Patient Number: Record 103, 107, 109, 111, 118, 119, 124, 200, 202, 209, 212, 214, 217, 221, 231, 232, 233.

Fig. 1. Power spectra of typical arrhythmia heartbeats in the frequency domain.

In this study, the dataset of QRS complexes typically for 7 heartbeat classes are taken from the MIT-BIH arrhythmias database [14]. From these records, a total of 43 QRS complexes (**ML II Signal**) are picked up and classified into seven types including **normal**, **V**, **A**, **L**, **R**, **P**, and **F**. FFT are applied to ECG signals for power spectrum estimation to construct various patterns. The frequency-domain features of seven classes are produced as shown in Figure 1. To quantify the differences among various classes, the comparative sequences for each class are created as $X(k)=[x_1(k), x_2(k), \ldots, x_i(k), \ldots, x_n(k)]$, $i=1, 2, 3, \ldots, n$, $k=1, 2, 3, \ldots, K$ ($n=12$ and $K=43$ in our study). The numbers of averaged patterns from the same class are 7-, 11-, 2-, 7-, 8-, 6-, and 2-set data respectively.

V. Experimental Test

The proposed diagnostic procedure was developed on a PC Pentium-IV 3.0GHz with 248MB RAM and MATLAB workspace, based on the MIT-BIH arrhythmias records including patient numbers 103, 107, 109, 111, 118, 119, 124, 200, 202, 209, 212, 214, 217, 221, 231, 232, and 233. Two study cases are chosen for demonstration, as detailed below.

Table 1. The test results of cardiac arrhythmias

Record		Number of Arrhythmias							CPU Time (sec)	Accuracy (%)
		●	V	A	L	R	P	F		
119	Actual	75	25	0	0	0	0	0	—	—
	Test1	73	25	0	0	1	0	1	1.344	98
	Test2	73	25	0	0	1	0	1	1.374	98
221	Actual	84	16	0	0	0	0	0	—	—
	Test1	82	16	0	1	0	0	1	1.359	98
	Test2	82	16	0	1	0	0	1	1.343	98
217	Actual	0	3	0	0	0	94	3	—	—
	Test1	0	6	0	0	4	87	3	1.408	93
	Test2	0	6	0	0	4	87	3	1.124	93

A. Single Cardiac Arrhythmia

The QRS complexes are extracted within the movable window with each shift in time. P-wave and T-wave can be removed in this window duration. Frequency-domain features are estimated by FFT technique. Using 100 heartbeats (about 1.5min long) of the patient numbers 119 and 221 containing normal beats (●) and pattern **V**s, the results of the Test 1 show that the overall accuracies are greater than 90% as shown in Table 1. In ECG measurement, signals may be disturbed by noise such as power line interference or quantification error. The proposed method must be robust enough to handle noisy environments. Test 2 shows the results of single cardiac arrhythmia involving noisy interference (50Hz and 60Hz). The results confirm that the major class is premature ventricular contraction. Expect sensitivity 92% for class **V**, specificity 100% for normal beat, and positive predictivity 100% (>80%) are obtained to quantify the performance of proposed method without/with a noisy background.

B. Multiple Cardiac Arrhythmias

One hundred heartbeats of the patient number 217 containing multiple cardiac rhythms **V**, **P**, and **F** are selected to test. The proposed procedure recognizes 94 paced beats with 7 failures. The overall accuracy is also greater than 90%. The results of Test 1 and Test 2 confirm that the major types are paced beat (**P**) as shown in Table 1. Positive predictivity 92% (>80%) for major class **P** is also obtained to quantify the performance of proposed method. The proposed method can also recognize the multiple cardiac arrhythmias with high confidence.

VI. Conclusion

The diagnostic procedure based on novel GRA is presented to recognize cardiac arrhythmias. FFT technique is used to estimate the frequency-domain features. Frequency components from 1Hz to 12Hz are selected for multiple ECG beats recognition. The spectrum varies with different class, origination and conduction path especially form 1Hz to 6Hz. Novel GRA then uses these features to identify the cardiac arrhythmias. For both recorded and unrecorded data, the experimental results demonstrate the efficiency of the proposed method. Designing a virtual medical instrument, measurement, data storage, signal processing, signal classification, decision support, and human computer interface have become aided functions for disease diagnosis. The proposed diagnostic algorithm is easy to implement in the PC-based virtual instrument. Matlab-Excel Link is a software add-on to integrate Excel and window-based Matlab computing environment. Excel Link provides data management including create, append, overwrite, or delete with data from the Excel workspace and the computing command from Matlab workspace. The proposed method can be further used as an aided tool for ECG beats recognition, and be integrated in the monitoring device.

References

[1] Hartmut Dickhaus and Hartmut Heinrich, "Classifying Biosignals with Wavelet Networks-A Method for Noninvasive Diagnosis, IEEE Engineering in Medicine and Biology, September/October 1996, pp. 103-111.

[2] Shuren Qin, Zhong Ji, and Hongjun Zhu, "The ECG Recording Analysis Instrumentation Based on Virtual Instrument Technology and Continuous Wavelet Transform, "Proceedings of the 25th Annual International Conference of the IEEE EMBS Cancun, Mexico, September 17-21, 2003, pp. 3176-3179.

[3] Giovanni Bortolan, Christian Brohet, and Sergio Fusaro, "Possibilities of Using Neural Networks for ECG Classification, "Journal of Electrocardiology, Vol. 29 (Supplement), pp. 10-16.

[4] Kei-ichiro Minami, Hiroshi Nakajima, and Takesshi Toyoshima, Real-Time Discrimination of Ventricular Tachyarrhythmia with Fourier-Transform Neural network, "IEEE Transactions on Biomedical Engineering, Vol. 46, No. 2, February 1999, pp. 179-185.

[5] Khadra L, Al-Fahoum A, and Binajjaj S, "A New Quantitative Analysis Technique for Cardiac Arrhythmia Classification Using Bispectrum and Bicoherency, "Proceeding of the 26th Annual International Conference of the IEEE EMBS, san Francisco, CA, USA, September 1-5, 2004, pp.13-16.

[6] Kamen, P. W., Krum, H., and TONKIN, A. M. (1996), “Poincare Plot of Heart Rate Variability Allows Quantitative Display of Parasympathetic Nervous Activity, ”Clin. Sci., 91, pp. 201-208.
[7] R. Acharya U, A. Kumar, P. S. Bhat, C. M. Lim, and S. S. Lyengar, “Classification of Cardiac Abnormalities Using Heart Rate signals, ”Medical & Biological Engineering & Computing, Vol. 42, No. 3, May 2004, pp. 288-293.
[8] Rosaria Silipo and Carlo Marchesi, “Artificial Neural Networks for Automatic ECG Analysis, ”IEEE Transactions on Signal Processing, Vol. 46, No. 5, May 1998, pp. 1417-1425.
[9] Yang Wang, Yi-Sheng Zhu, Nitish V. Thakor, and Yu-Hong Xu, “A Short-Time Multifractal Approach for Arrhythmias Detection Based on Fuzzy Neural Network, ”IEEE Transactions on Biomedical Engineering, Vol. 48, No. 9, September 2001, pp. 989-995.
[10] Stanislaw Osowski and Tran Hoai Linh, “ECG Beat Recognition Using Fuzzy Hybrid Neural Network, ”IEEE Transactions on Biomedical Engineering, Vol. 48, No. 11, November 2001, pp. 1265-1271.
[11] John H. Wu and Chie-Bein Chen, “An Alternative Form for Grey Relational Grades, ”The Journal of Grey System, Vol. 1, 1999, pp. 7-12.
[12] Wei-Che Chang, “A Comprehensive Study of Grey Relational Generating, ”Journal of the Chinese Grey Association, Vol. 1, 2000, pp. 53-62.
[13] Whei-Min Lin, Chine-Ming Chen, Long-Wei Wang, and Chia-Hung Lin, “Emergency Operation Strategy for Power System Restoration with Grey Relational Analysis, ”The 10th National Conference on Grey System Theory and Application, Ping-dong, Taiwan, December 2 2005, pp. 392-398.
[14] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, lvanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, Physio Toolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic signals. Circulation 101(23):e215-e220[Circulation electronic Pages; http://circ/cgi/content/full/101/23/e215]; 2000 (June 13).
[15] Duhamel, P. and M. Vetterli, “Fast Fourier Transforms: A Tutorial Review and a State of the Art, ”Signal Processing, Vol. 19, April 1990, pp. 259-299.