Journal of Naval Science and Engineering
2015, Vol. 11, No.1, pp.1-18

Optimizing MLP Classifier and ECG Features for Sleep Apnea Detection

Oğuz Han TİMUŞ[1],
Erkan KIYAK[2],

[1] *Turkish Navy Inventory Control Center, Golcuk, Kocaeli*
[1] *oguzhantimus@gmail.com.tr*
[2,] *Turkish Navy Inventory Control Center, Golcuk, Kocaeli*
[2] *ekiyak@hotmail.com.tr*

Abstract

The purpose of this study is to optimize multilayer perceptron (MLP) classifier and find optimal ECG features to achieve better classification for automated sleep apnea detection. k-fold cross-validation technique was employed for classification of apneaic events on the apnea database of the DREAMS project containing 12 whole-night Polysomnography (PSG) recordings previously examined by an expert. To achieve the best possible performance with MLP, the correlation feature selection method was utilized. The performance for apnea event diagnosis after optimization of the features and the classifier resulted almost 10% in accuracy, %7 in sensitivity and %13 in specificity.

MLP Sınıflandırıcısı ve EKG Özniteliklerinin Uyku Apnesi Tanısı için Optimizasyonu

Özetçe

Bu çalışmanın amacı otomatik uyku apnesi tanımlamasında daha iyi sınıflandırma sağlamak amacıyla çok katmanlı algılayıcı sınıflandırıcısı ile kullanılacak EKG özniteliklerinin optimizasyonunu gerçekleştirmektir. Uzman hekim tarafından değerlendirilmiş 12 hastanın bulunduğu DREAMS projesi veri tabanından elde edilen PSG kayıtlarındaki apne olgularının sınıflandırılmasında k-kat çapraz doğrulama algoritması kullanılmıştır. Çok katmanlı algılayıcı ile sınıflandırmada en iyi başarımı elde etmek için ilinti öznitelik seçim metodu kullanılmıştır. Apne olgularının tespitinde, sınıflandırıcı ve öznitelik optimizasyonu sonrasında doğrulukta yaklaşık %10, duyarlılıkta %7 ve kesinlikte %13 artış elde edilmiştir.

***Keywords**:* Electrocardiogram, Heart rate variability (HRV), sleep apnea, multilayer perceptron (MLP), classification, Correlation Feature Selection (CFS)
***Anahtar Kelimeler**:* Elektrokardiyogram, kalp atım hızı değişikliği, uyku apnesi, çok katmanlı algılayıcı, sınıflandırma, ilinti öznitelik seçimi

1. INTRODUCTION

Sleep is as important as these basic needs, since a human spends nearly one-third of his life sleeping. Sleep is essential for a healthy life because the regeneration of damaged body's cell, production of growth hormone, formation of the immune system, refreshing and resting the brain occurs during sleep. Sleeps disorders prevents humans from utilizing sleep function and causes decreased immune system, concentration problems, stress and physical activity insufficiency [1].

Sleep apnea is the most common trouble encountered among sleep disorders. It is defined as a decrease in or complete interruption of breathing for at least 10 seconds and observed as central, obstructive, and mixed forms [2]. In central sleep apnea, breathing is interrupted by the lack of respiratory effort; but in obstructive sleep apnea, breathing is interrupted by a physical block to airflow despite respiratory effort. In mixed sleep apnea, during the apnea events a transition from central to obstructive apnea features are observed [3]. Sleep apnea is a respiratory event, but it also affects the cardiovascular system. Therefore, the ECG can provide very valuable information about apneaic events [4].

Polysomnography is considered as the gold standard for sleep apnea diagnosis. It is performed in a hospital sleep units, while the patient is asleep and a wide range of physiological parameters is recorded; e.g. electro-encephalography (EEG), electrocardiography (ECG), electromyography (EMG), electrooculography (EOG), respiratory airflow, blood oxygen saturation (SpO2), respiratory effort) etc. After the patients sleep these recordings visually inspected and scored by a physician [5].

Various researches have been conducted for sleep apnea detection based on ECG. In 1984, Guilleminault proposed a test against the RR-interval derived from the ECG signal could be used for the screening of sleep apnea [6].

In 2004, Chazal et al. suggested an obstructive sleep apnea detection using ECG signal. The statistical measurement of variables derived from RR-intervals and ECG-derived respiratory signal (EDR) were the features used in and separated normal recordings from the apnoea recordings by linear discriminant analysis gave 100% accuracy rate [7].

In 2007, Mendez et al. [15] used a bivariate autoregressive model to evaluate beat-by-beat power spectral density for both the Heart Rate Variability (HRV) and the R peak areas. k-NN supervised learning classifier was employed for categorizing apnea events from normal ones, on a minute-by-minute basis for each recording and achieved more than 85% classification accuracy of apnea and non apnea minutes.

In 2009, Khandoker et al. [9] employed support vector machine (SVM) for automatic diagnosis of Obstructive Sleep Apnea Syndrome (OSAS) from ECG signal. They utilized wavelet based features extracted from HRV derived from R-R intervals and ECG-derived respiration (EDR) as inputs of the SVM for recognition of the OSAS. They obtained an accuracy rate of 92.85% with the Cohen’s κ value of 0.85.

In addition, various researches regarding to the classification of sleep apnea syndrome have been conducted based on using ECG signals.

In 2010, Yilmaz et al. conducted research for the sleep stages and apnea classification using a single lead ECG. The features used in the research were median, inter-quartile range (IQR), and the mean absolute deviation (MAD) values derived from RR-intervals in each epoch. For classification kNN (k Nearest Neighbor), Quadratic Discriminant Analysis (QDA), and SVM were used and classification based on QDA and SVM gave the best accuracy rate of 94.5% [3].

Sani et.al. [10] examined two approaches of feature selection and concluded that classification results using only 3 features as proposed by Yilmaz et al. [3] gives about 3.59% gain on overall classification accuracy (CA) and 7.5% gain on area under the receiver operating characteristic (ROC) - curve (AUC) than the classification accuracy using 8 features as proposed by Chazal et al. [7].

Mendez et al. [11] used a bivariate timevarying autoregressive model (TVAM) to evaluate beat-by-beat power spectral density for both RR intervals and the QRS complex areas. k-NN and neural network classifiers were employed for categorizing apnea minutes from normal ones, and both classfiers achieved 88% classification accuracy.

The aim of this study is to optimize multilayer perceptron (MLP) classifier and find optimal ECG features to achieve better classification for automated sleep apnea diagnosis. This study is based on 13 time domain features extracted from Heart Rate Variability (HRV) derived from RR-intervals. To achieve the best possible performance with MLP, Correlation Feature Selection (CFS) method was used. Then the MLP model was optimized based on the sensitivity, specificity and classification accuracy. For training the classifier k-fold cross-validation (k=10) technique was employed on the apnea database of the DREAMS project [12].

2. MATERIALS AND METHODS

The ECG recordings used in this study were collected from the DREAMS project aims for automatic sleep signal analysis of adults [12]. It consists of 12 whole-night PSG recordings, coming from patients with sleep apnea syndrome and annotated by an expert in respiratory events (central, obstructive and mixed apnea and hypopnea). The subject age in the dataset ranged between 41 to 79 years (53.93 ± 7.6), sleep duration ranged between 7.5 to 9.5 hours (8.63 ± 0.7) and AHI ranged 7.4 to 80.9 apnea/hypopnea per hour (46.52 ± 25.9).

A minute was defined as apneaic if one apnea or hypopnea event occurs, otherwise defined as normal. Schematic diagram of the methodology used in this study is shown in Fig.1. After importing the PSG recordings from apnea database to MATLAB, the ECG data for each subject were divided into 1 minute epochs for processing.

Then R-peaks were detected for each epoch. Before calculating RR-intervals to form HRV data, a correction algorithm to detect and eliminate missed R-peaks and false R-peaks was performed.

Later, time domain features were extracted from HRV data. To achieve the best possible performance with classifier the correlation feature selection (CFS) method was used.

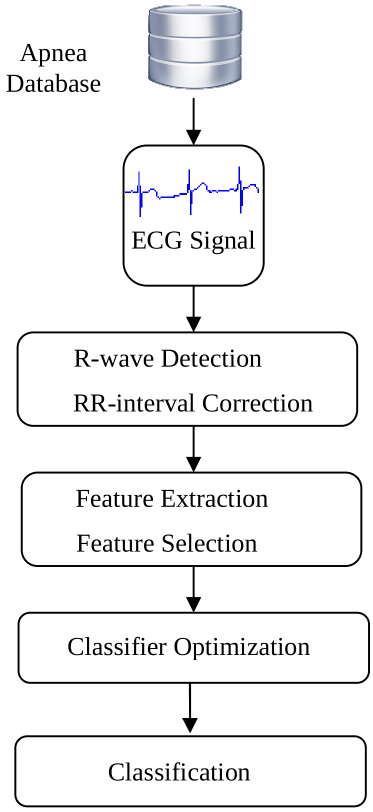

Fig. 1. Schematic diagram of the methodology

Before training the classifier optimal number of nodes in hidden layer (m=16) was calculated by observing sensitivity, specificity and classification accuracy measures using different number of nodes in the hidden layer of MLP.

For training the classifier 10-fold cross-validation technique was employed on the training data set.

2.1. The ECG signal and HRV Calculation

Since ECG is a simple and efficient technology for sleep apnea detection various researches have been conducted based on Heart Rate Variability (HRV) derived from RR-intervals. HRV is related to the presence of sleep apnea and consists of many features in both time and frequency domains. For calculation of these features there are different methods to analyze short term and long term by the length of raw ECG data in both time and frequency domains [13]. A sample of ECG signal is shown in Fig.2. In this study Teager Energy Operator (TEO) was used to detect R-waves.

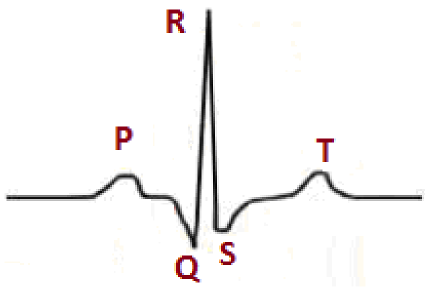

Fig. 2. The ECG Signal

P (the atrial wave), QRS (the ventricular depolarization) and T (the ventricular repolarization) are the basic waves in the ECG signal. The QRS complex is the central and most visually obvious part of the ECG signal and contains the most characteristic wave set. In order to calculate HRV, the peak value of R-wave (or, the peak value of the QRS complex) detection must be performed.

HRV is the time differences between two consecutive R-wave peaks (RR-intervals). HRV is calculated after detecting R-wave peaks in each epoch. Since RR-intervals depend on the number of heart beats, the number of HRV samples in each epoch differs.

TEO is a handy method for studying in continuous and discrete domains of a single signal derived from its energy. In continuous domain TEO is formulated as:

$$\psi_c[x(t)] = \dot{x}^2(t) - x(t)\ddot{x}(t) \quad (1)$$

For discrete domain TEO is applied over the signal off-line, which is sampled at a particular frequency and formulated as below:

$$\psi[x(n)] = x^2(n) - x(n-1)x(n+1) \quad (2)$$

Teager energy operator covering the three samples of the signal side by side, shows a very local feature of the signal [14].

After importing the PSG recordings to MATLAB, the ECG data for each subject were divided into 1 minute epochs for processing. In order to calculate the peak value of R-waves in each epoch, peak finder function was applied to each consecutive 120 samples for training group. Errors in the R-peak detection will result in errors in the calculation of the HRV. Therefore, a correction algorithm to detect and eliminate missed R-peaks and false R-peaks must be performed.

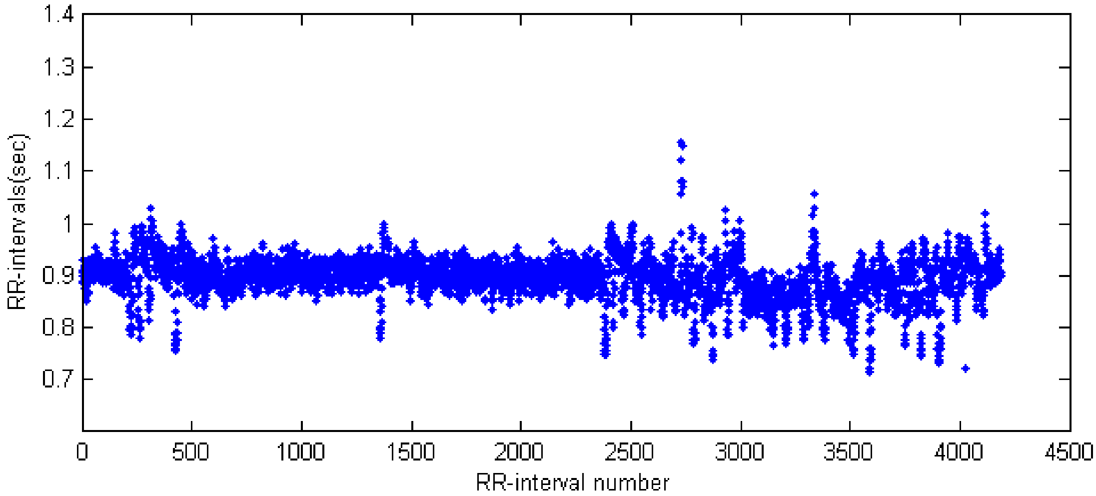

Fig.3. HRV obtained from a patient (duration one hour)

A common approach to eliminate missed R-peaks and false R-peaks is excluding RR-intervals less than 0.5 and greater than 1.5 seconds. Before applying this R-peak correction algorithm, min and max RR-interval values in each epoch were analyzed. Min RR-intervals were greater than 0.5 second (meaning no false R-peaks were detected) but some of the max RR-intervals were close to or greater than 1.5 seconds (meaning missed R-peaks were detected). For achieving more accurate results, RR-intervals exceeding 180% of the mean value in the epoch were replaced by the mean value of prior and next RR-intervals. The intervals between successive RR-intervals are taken as HRV. HRV obtained from a patient duration of one hour recording is shown in Fig.3.

2.2. Feature Extraction and Selection

The features in our study are analyzed in time domain and extracted from HRV data for each epoch. The best group of features in apnea classification consists of 9 features and indicated by asterisks (*) below:

- inter-quartile range (IQR*), the measure of statistical dispersion equal to the difference between the upper (75th) and lower quartiles (25th)
- mean absolute deviation (MAD*), the mean value of absolute difference of the data set from its mean
- median*, mean and standard deviation (SDNN) of RR-interval
- root mean square of successive differences (RMSSD), the square root of the mean of the sum of the squares of the successive differences between adjacent RR-intervals
- standard deviation of successive differences (SDSD), the standard deviation of the successive differences between adjacent RR-intervals
- NN50 count (variant 1) *, the number of all pairs of adjacent RR-intervals differing by more than 50 ms
- NN50 count (variant 2)' the number of pairs of adjacent RR-intervals differing by more than 50 ms in which the **first** interval is longer
- NN50 count (variant 3) *, the number of pairs of adjacent RR-intervals differing by more than 50 ms in which the **second** interval is longer

- pNN50 (variant 1) *, pNN50 (variant 2) * and pNN50 (variant 3) *, the proportion of each NN50 count divided by the total number of RR-intervals
- $\overline{HR}$*, the mean of the heart rate
- SDHR, the standard deviation of instantaneous HR values

The accuracy of the classification methods can be degraded by the presence of noisy or irrelevant features. To prevent *the course of dimensionality* and achieve the best possible performance with a learning algorithm, a feature selection method should be used. In our study, we used CFS method for the features extracted from training database.

CFS method basically evaluates subsets of features based on their correlation with the classification. Then selects the features highly correlated with the classification and unrelated to each other.

2.3. Multilayer Perceptron Classifier

The Multilayer Perceptron (MLP) is a feed forward artificial neural network and widely used supervised learning method (or classifier) that can distinguish data that are not linearly separable, capable of modeling complex functions and good at ignoring irrelevant inputs and noise. A MLP consists of multiple layers of nodes. Each node is a neuron (or processing element) with a nonlinear activation function (except the input nodes). The most common type of MLP consists of an input layer, one hidden layer (number of hidden nodes set by trial-and-error) and output layer where each layer is fully connected to the next one. The information moves in forward direction only, from the input nodes, through the hidden nodes and to the output nodes. A typical MLP model with n inputs, one hidden layer with m nodes and two output node is illustrated in Fig.4.

The hyperbolic tangent $y(v_i) = \tanh(v_i)$ (ranges from -1 to 1) and the logistic function (or sigmoid) $y(v_i) = (1 + e^{-v_i})^{-1}$ (ranges from 0 to 1) are the two main activation functions used in current applications. Here y_i is the output of the ith node (neuron) and v_i is the weighted sum of the input synapses.

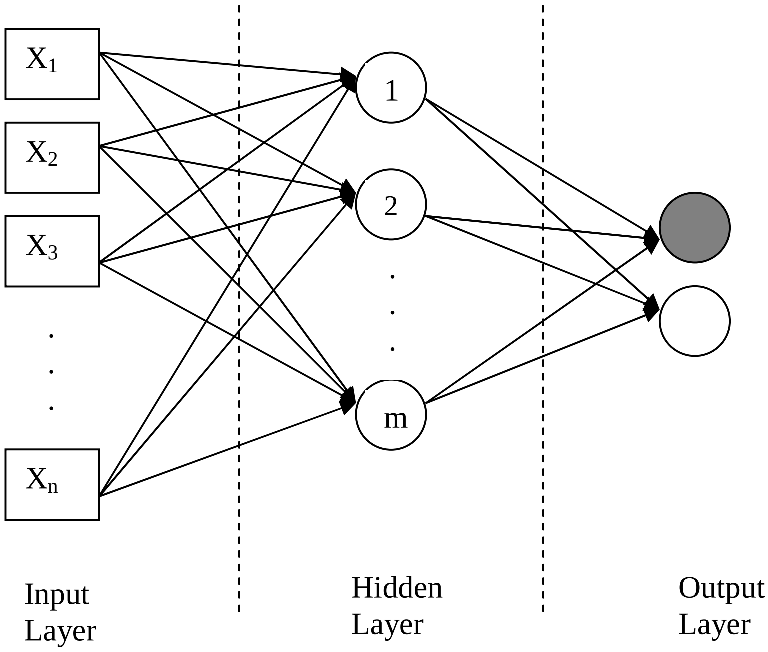

Fig.4. A typical MLP model

MLP model was optimized by adjusting the number of hidden nodes in the hidden layer, based on sensitivity, specificity and accuracy of the classifier performance using 10-fold cross-validation (see Fig.5).

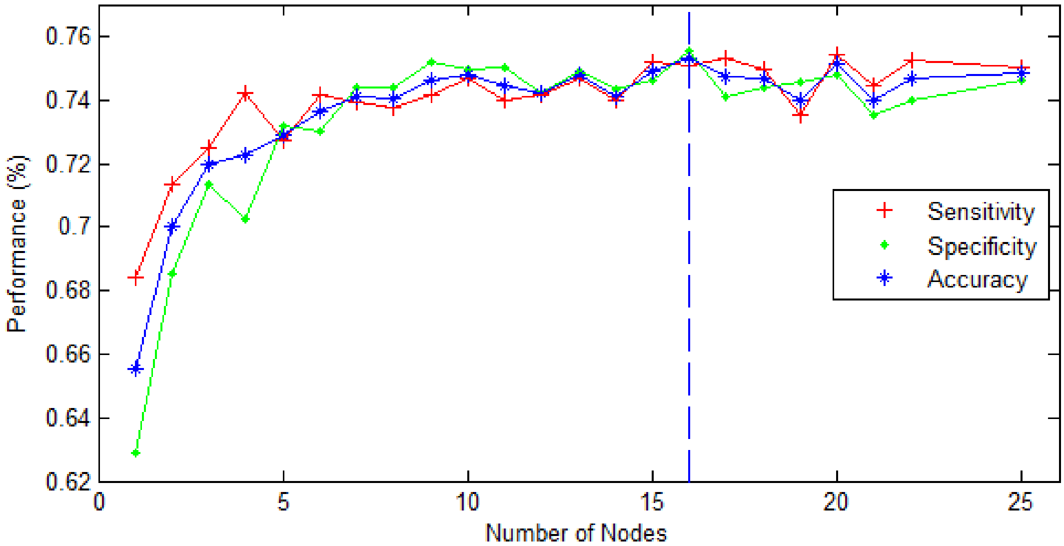

Fig.5. Classifier performance at different number of nodes in hidden layer

3. RESULTS

MLP classifier performance, with 9 features selected by the CFS method, 10-fold cross-validation was used to evaluate the performance in training data obtained from the apnea database contained 12 recordings, of which 2984 min annotated as apneaic (47.2%) and 3254 min labeled as normal (52.2%).

Table I shows the performance assessment on training data before and after optimization.

TABLE I
PERFORMANCE ASSESMENT ON TRAINING DATA

	Accuracy (%)	Sensitivity (%)	Specificity (%)
Before Optimization	65.72	68.35	62.48
After Optimization	75.29	75.09	75.53

4. DISCUSSION AND CONCLUSION

Based on the experimental results, MLP classifier was found to be a good alternative for apnea classification based on ECG signals. Using the features obtained from the CFS feature selection algorithm has improved the classification performance based on accuracy, sensitivity and specificity. The performance for apnea event diagnosis after optimization of the features and the classifier resulted almost 10% gain in accuracy, %7 gain in sensitivity, and % 13 gains in specificity.

The training set was lack of normal or a control group, it consisted subjects of apnea and mild/borderline apnea. Using normal or control group in training might lead to a better classification success.

In future studies, different classifiers such as Naïve Bayes, Neural Networks and Decision Trees can be employed and compared with each other to provide a reliable decision support system for the diagnosis and classification of the severity of sleep apnea.

REFERENCES

[1] Murali N.S., Svatikova A., Somers V.K.: “Cardiovascular physiology and sleep”, Frontiers in Bioscience, 2003, 8:636-652.

[2] Morgenthaler T.I., Kagramanov V., Hanak V., Decker P.A.: “Complex sleep apnea syndrome: is it a unique clinical syndrome?”, Sleep 2006, 29 (9): 1203-9.

[3] Yilmaz B., Asyali M.H., Arikan E., Yetkin S., Ozgen F.: “Sleep stage and obstructive apneaic epoch classification using single-lead ECG”, Biomed Eng Online, 2010 Aug 19;9 (1): 39.

[4] Mendez M.O., Bianchi A.M., Matteucci M., Cerutti S., Penzel T.: “Sleep apnea screening by autoregressive models from a single ECG Lead”, IEEE Transactions Biomedical Engineering, 2009, 56(12):2838-49.

[5] Otero A., Félix P., Barrob S., Zamarrónc C.: “A structural knowledge-based proposal for the identification and characterization of apnoea episodes”, Applied Soft Computing, 12 (1): 516-526, 2011.

[6] Guilleminault C., Connoly S., Winkle R., Melvin K., Tilkian A. “Cyclical Variation of the Heart Rate in Sleep Apnea Syndrome”, Lancet 1984;21:126–131

[7] Chazal P., Penzel T., and Heneghan C., "Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram", Physiological Measurement, 2004 July; 25: 967-983.

[8] Mendez M.O., Ruini D.D., Villantrieri O.P., Matteucci M., Penzel T., Cerutti S., Bianchi A.M.: "Detection of sleep apnea from surface ecg based on features extracted by an autoregressive model", Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite Internationale, Lyon, France, 2007.

[9] Khandoker, A. H.; Palaniswami, M.; & Karmakar, C. K.; "Automated scoring of
obstructive sleep apnea and hypopnea events using short-term electrocardiogram
recordings", IEEE Transactions On Information Technology In Biomedicine, Vol. 13, No.6, (November 2009), pp. 1057-1067, ISSN 1089-7771.

[10] Isa S., Fanany M., Jatmiko M., Murini A.: "Feature and Model Selection on Automatic Sleep Apnea Detection Using ECG", International Conference on Computer Science and Information Systems, ICACSIS 2010, pp 357-362, 2010.

[11] Mendez M.O., Bianchi A.M., Matteucci M., Cerutti S., Penzel T.: "Sleep apnea screening by autoregressive models from a single ecg lead", IEEE Transactions Biomedical Engineering, 56 (12), 2009, 2838-2849.

[12] Devuyst S., Dutoit D., Kerkhofs M., DREAMS apnea database, University of MONS - TCTS Laboratory and Université Libre de Bruxelles - CHU de Charleroi Sleep Laboratory Charleroi, http://www.tcts.fpms.ac.be/~devuyst/Databases/DatabaseApnea/ (Accessed Novemer 10, 2014).

[13] Camm A., J, Malik M., Bigger J. T., Breithardt G., Cerutti S., Cohen R. J., Coumel P., Fallen E. L., Kennedy H. L., Kleiger R. E., Lombardi F., Malliani A., Moss A. J., Rottman J.N., Schmidt G., Schwartz P. J., Singer D. H.: "Heart rate variability-Standards of measurement, physiological interpretation, and clinical use", European Heart Journal, Vol 17, No3, (March 1996), pp354-381.

[14] Kaiser J.F: "Some Useful Properties of Teager's Energy Operators, Acoustics, Speech and Signal Processing", ICASSP-93, IEEE International Conference, Vol3, Issue April 1993, pp149-152, ISSN 1520-6149.