THE USAGE OF STATISTICAL FEATURES IN THE APPROXIMATION COMPONENTS OF WAVELET DECOMPOSITION FOR ECG CLASSIFICATION: A CASE STUDY FOR STANDING, WALKING AND SINGLE JUMP CONDITIONS

Makbule Hilal MÜTEVELLİ[1], Semih ERGİN[2]

ABSTRACT

The purpose of this study is to classify electrocardiogram (ECG) signals with a high accuracy rate. The ECG signals used are obtained from the Physiobank archive. These signals are preprocessed to remove noise. Features with distinctiveness in classification are obtained both in the time domain and the frequency domain. The Discrete Wavelet Transform method is used for feature extraction in frequency domain. ECG signals are classified by the Naive Bayes method after the required features are extracted.

***Keywords*:** ECG, statistical feature extraction, discrete wavelet transform, ECG classification.

INTRODUCTION

ECG is electrical appearance of the heart and is recorded by placing electrodes on the patient's body. The noises such as electrode contact, motion artifacts, muscle contraction, line noise and instrumental noise can often disrupt ECG signals (Poungponsri & Yu, 2013). The ECG contains important information about the rhythm and function of the heart (Mitra, Mitra, & Chaudhuri, 2006). The analysis of ECG is commonly used in the diagnosis of many cardiac diseases (Mahmoodabadi, Ahmadian, & Abolhasani, 2005). ECG components can be categorized in a range of classes. Every ECG signal consists of three different waves described as P, QRS, and T (Fig. 1). It indicates the depolarization and repolarization of the heart muscle (Mitra et al., 2006; Jenkal et al., 2016). P wave is formed by atrial depolarization. QRS complex is formed by ventricular depolarization. T wave is formed by ventricular repolarization.

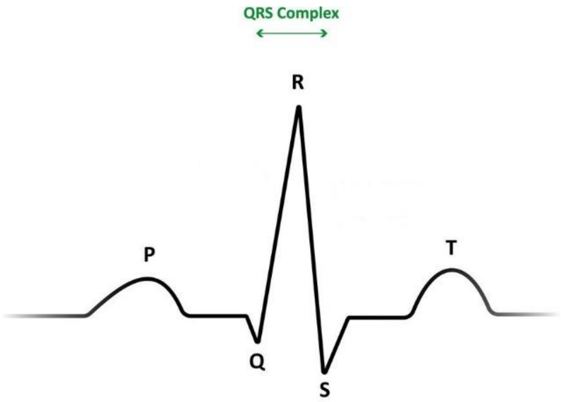

Fig. 1. The components of the ECG signal

Analysis of ECG signal give information about heart problems. ECG feature extraction is very importance because features with distinctiveness are used in the classification and the analysis process. To analyze the ECG signal correctly, many scientific studies have been carried out (Mitra et al., 2006;Martens, Rabotti, Mischi, & Sluijter, 2007;Datian & Xuemei, 1996;Zhao & Zhang, 2006;Saritha, Sukanya, & Murthy, 2008;Yeh & Wang, 2008;Islam, Haque, Tangim, Ahammad, & Khondokar, 2012;C.Bakır, 2015;Ceylan, Özbay, & Karlik, 2009). Different techniques such as Discrete Wavelet Transform (DWT) (Guyon & Elisseeff, 2006;Lei, Wang, & Liu, 2013;Mahmoodabadi et al., 2005) , Fuzzy Wavelet Packet Analysis, Time Domain (TD), Principal Component Analysis (PCA), Power Spectral Density (PSD) have been used for these purposes.

[1] Kastamonu Üniversitesi, hmutevelli@kastamonu.edu.tr
[2] Eskişehir Osmangazi Üniversitesi, sergin@ogu.edu.tr

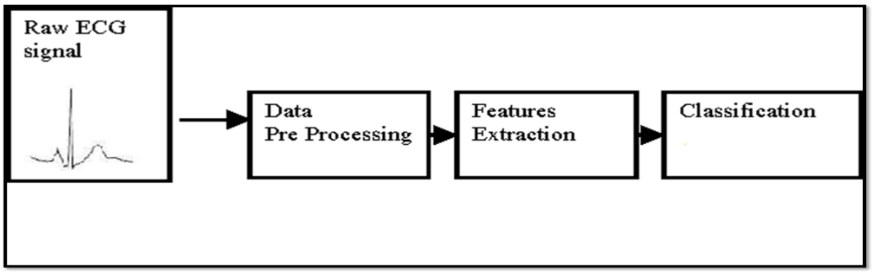

Fig. 2. Methods Block Diagram

In this work, it is aimed to correctly classify ECG signals which are recorded from a healthy 25-year-old male performing different physical activities. As seen in Fig. 2, firstly, preprocessing phase is performed. At this stage, the noises that can disrupt the signal such as muscle noise, noise from the electrode and line noise are eliminated. Then, time domain features and discrete wavelet transform features are extracted. Finally, classification is made by the Naive Bayes Method.

METHODS

Discrete Wavelet Transform

Wavelet transform represents signals in both time and frequency. Various wavelets are generated from the mother wavelet in wavelet transform. Wavelet transform analyzes non-stationary signal, such as ECG signals and separates the data into different frequency components (Yılmaz & Bozkurt, 2013;Poungponsri & Yu, 2013).

The discrete wavelet transform is a mathematical technique. It is usually used for the signal processing. This transform decomposes a signal into different resolutions using high pass and low pass filters. DWT decomposition equation (Eqns. (1)) is given as :

$$DWT(m,k) = \frac{1}{a}\sum_{n=0}^{N-1} s(n)g\left(\frac{k-b}{a}\right) \tag{1}$$

where $s(n)$ is the original signal, N is the number of samples in windowed signal, $g(\cdot)$ represents mother wavelet, m is decomposition level, a and b is called scaling and translation parameter, respectively (Lei et al., 2013).

Discrete wavelet transform provides sufficient information for analysis and synthesis of the original signal. DWT obtains hidden information from the ECG signal. It allows also the signal to be processed efficiently in the frequency domain. The DWT decomposes a signal into low and high frequency components. The approximation coefficients are obtained from the low frequency component and the detail coefficients from the high frequency component (Elhaj, Salim, Harris, Tian, & Ahmed, 2016;Joy, Acharya, & Choo, 2013).

Statistical Feature Extraction

To correctly classify ECG signals requires generation of the feature vector which contains features both in the time domain and the frequency domain (Zhao & Zhang, 2006). In this work, features in the time domain include mean, standard deviation, energy, maximum, minimum, kurtosis, skewness, and the difference between maximum and minimum values of ECG signals.

The Discrete Wavelet Transform method is used to extract the features in the frequency domain. In this work, wavelet packet analysis is used to extract DWT features. The advantage of wavelet packet analysis is that it decomposes both approximations and details at all levels to obtain full subband decomposition. The number of decomposition level is important in analysis of signals. 4-level wavelet packet decomposition is applied on each signal and 16 subbands are obtained. However, the approximation coefficients from the low frequency component are preferred and 8 of these subbands are used because the approximation coefficients represent the main characteristic of each heart signal. Then features such as mean, standard deviation, maximum, minimum, skewness, kurtosis, variance, entropy, median are extracted from each subband.

Mean is computed as:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \,. \tag{2}$$

Where x_i is the *i* th sample in the current window and N is the window size.

Standard deviation is obtained by:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \ . \tag{3}$$

Maximum and minimum are computed as :

$$Maximum = max\{x_i, 1 \le i \le N\} \tag{4}$$

and

$$Minimum = min\{x_i, 1 \le i \le N\} \ . \tag{5}$$

Skewness and kurtosis are obtained by:

$$S = \frac{1}{N}\frac{\sum_{i=1}^{N}(x_i - \mu)^3}{\sigma^3} \tag{6}$$

and

$$K = \frac{1}{N}\frac{\sum_{i=1}^{N}(x_i - \mu)^4}{\sigma^4} \ . \tag{7}$$

Variance is computed as:

$$var(X) = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \ . \tag{8}$$

Entropy is calculated using:

$$H(X) = \sum_{i=1}^{N} {x_i}^2 \log({x_i}^2) \ . \tag{9}$$

Median is computed as:

$$Median = \left\{\left(\frac{n+1}{2}\right)^{th} term, if\ N\ is\ odd\ \left(\frac{n}{2}\right)^{th} term + \left(\frac{n}{2} + 1\right)^{th} term, if\ N\ is\ even \ . \right. \tag{10}$$

EXPERIMENTAL STUDY

Database

All ECG data are obtained from the Physiobank archive that contains large collections of recorded physiologic signals. The data is Motion Artifact Contaminated ECG database. The ECG data are recorded from a healthy 25-year-old male performing different physical activities. Used ECG signals are short duration. Each recording contains four signals (ECG 1 to ECG 4). Each row of the recording contains the samples of one signal. There are twenty seven records. The data consists of totally 108 ECG signals and there are three distinct classes that have different physical activities. The classes are s=standing, w=walking, j=single jump. Each class contains 36 ECG signals.

Classification

ECG signals are classified by the Naive Bayes method. Two thirds of the data is used for training while one third of the data is used for testing. After feature extraction, new characteristic data to be used in the training are obtained and class labels are added to the data. These labels and known classes set a model. This model contains parameters used in training. Then the data to be used in testing are obtained. Labels for this data and model are predicted and classification accuracy rate is achieved 69,44 %.

FINDINGS

Table 1. Classification rates for different wavelet transform based subbands.

Wavelet Transform Subbands	Feature Dimension	Moving Average Filter Tap	Accuracy Rate
H,LH,LLH,LLLH,LLLLH,LLLLL	68	4	66.66 %
		9	58.33 %
LLLL,LLLH,LLHL,LLHH,LHLL,LHLH,LHHL,LHHH	79	4	63.88 %
		9	69.44 %
L,LL,LLL,LLLL,LLLL	58	4	52.77 %
		9	55.55 %

LL LLLH + LLHL + 0.5 x LLHH LLLH + LLHL + 0.1 x LLHH LLLH + LLHL LLLLLH + LLLLHL + 0.5 x LLLLHH LLLLLH + LLLLHL + 0.1 x LLLLHH LLLLLH + LLLLHL	78	4	55.55 %
		9	61.11 %
LH + HL + HH LLLH + LLHL + 0.5 x LLHH LLLH + LLHL + 0.1 x LLHH LLLH + LLHL LLLLLH+LLLLHL+LLLLHH	58	4	63.88 %
		9	61.11 %
H,LH,LLH,LLLH,LLLL	58	4	61.11 %
		9	61.11 %
H,LH,LLH,LLL	48	4	41.66 %
		9	55.55 %
H,LH,LL	38	4	44.44 %
		9	58.33 %
LLLLL,LLLLH,LLLHL,LLLHH,LLHLL,LLHLH,LLHHL,LLHHH LHLLL,LHLLH,LHLHL,LHLHH,LHHLL,LHHLH,LHHHL,LHHHH	152	4	47.22 %
		9	55.55 %
HLLLL,HLLLH,HLLHL,HLLHH,HLHLL,HLHLH,HLHHL,HLHHH, HHLLL,HHLHL,HHLHL,HHLHH,HHHLL,HHHLH,HHHHL,HHHHH	152	4	41.66 %
		9	50.00 %

In this section, the findings of the work are described and illustrated. Moving Average Filter used in the preprocessing is commonly used for smoothing an array of sampled signal. It takes M samples of input at a time and take the average of those M-samples and produces a single output point. When Moving Average Filter tap (M) is 9, results are generally better as seen in Table 1.

Different wavelet subbands are obtained and the classification rates corresponding to these subbands are presented in Table 1. 4-level DWT (LLLL,LLLH,LLHL,LLHH,LHLL,LHLH,LHHL,LHHH) performed for feature extraction achieve better performance than others because the approximation coefficients represent the main characteristic of each heart signal.

CONCLUSIONS

The preprocessing of ECG signals is performed. Time Domain and Discrete Wavelet Transform methods are used for feature extraction. It is important to obtain features with distinctiveness for the classification. Therefore, 4-level DWT is preferred in frequency domain. Appropriate usages of statistical features in the approximation components of Discrete Wavelet decomposition give better classification performance. In this work, it is seen that increase in the dimension of feature vector increases the accuracy rate in the classification when a suitable decomposition level is chosen.

REFERENCES

C.Bakır. (2015). "ECG Signals Classification with Neighborhood Feature Extraction Method." In *Medical Technologies National Conference (TIPTEKNO), 2015* (pp. 173–176). https://doi.org/10.1109/TIPTEKNO.2015.7374099

Ceylan, R., Özbay, Y., & Karlik, B. (2009). A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network. *Expert Systems with Applications*, *36*(3 PART 2). https://doi.org/10.1016/j.eswa.2008.08.028

Datian, Y., & Xuemei, O. (1996). Application of wavelet analysis in detection of fetal ECG. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings* (Vol. 3).

Elhaj, F. A., Salim, N., Harris, A. R., Tian, T., & Ahmed, T. (2016). Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. *Computer Methods and Programs in Biomedicine*, *127*, 52–63. https://doi.org/10.1016/j.cmpb.2015.12.024

Guyon, I., & Elisseeff, A. (2006). Feature Extraction, Foundations and Applications: An introduction to feature extraction. *Stud. Fuzziness Soft Comput.*, *207*, 1–25. https://doi.org/10.1007/978-3-540-35488-8_1

Islam, M. K., Haque, A. N. M. M., Tangim, G., Ahammad, T., & Khondokar, M. R. H. (2012). Study and analysis of ECG signal using MATLAB & LABVIEW as effective tools. *International Journal of Computer and Electrical*

Engineering, *4*(3), 404–408. https://doi.org/10.7763/IJCEE.2012.V4.522

Jenkal, W., Latif, R., Toumanari, A., Dliou, A., El, O., & Maoulainine, F. M. R. (2016). ScienceDirect An efficient algorithm of ECG signal denoising using the adaptive dual threshold filter and the discrete wavelet transform. *Integrative Medicine Research*, *36*(3), 499–508. https://doi.org/10.1016/j.bbe.2016.04.001

Joy, R., Acharya, U. R., & Choo, L. (2013). Biomedical Signal Processing and Control Technical note ECG beat classification using PCA , LDA , ICA and Discrete Wavelet Transform. *Biomedical Signal Processing and Control*, *8*(5), 437–448. https://doi.org/10.1016/j.bspc.2013.01.005

Lei, L., Wang, C., & Liu, X. (2013). Discrete Wavelet Transform Decomposition Level Determination Exploiting Sparseness Measurement. *International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, *7*(9), 691–694.

Mahmoodabadi, S. Z., Ahmadian, A., & Abolhasani, M. D. (2005). Ecg feature extraction using daubechies wavelets. In *Proceedings of the Fifth IASTED International Conference* (Vol. 2, pp. 343–348).

Martens, S. M. M., Rabotti, C., Mischi, M., & Sluijter, R. J. (2007). A robust fetal ECG detection method for abdominal recordings. *Physiological Measurement*, *28*(4), 373–388. https://doi.org/10.1088/0967-3334/28/4/004

Mitra, S., Mitra, M., & Chaudhuri, B. B. (2006). A rough-set-based inference engine for ECG classification. *IEEE Transactions on Instrumentation and Measurement*, *55*(6), 2198–2206. https://doi.org/10.1109/TIM.2006.884279

Poungponsri, S., & Yu, X.-H. (2013). An adaptive filtering approach for electrocardiogram (ECG) signal noise reduction using neural networks. *Neurocomputing*, *117*, 206–213. https://doi.org/10.1016/j.neucom.2013.02.010

Saritha, C., Sukanya, V., & Murthy, Y. N. (2008). ECG signal analysis using wavelet transforms. *Bulg. J. Phys*, *35*(1), 68–77. Retrieved from http://bjp-bg.com/papers/bjp2008_1_68-77.pdf%5Cnhttp://www.iiste.org/Journals/index.php/ISDE/article/viewFile/610/499

Yeh, Y. C., & Wang, W. J. (2008). QRS complexes detection for ECG signal: The Difference Operation Method. *Computer Methods and Programs in Biomedicine*, *91*(3), 245–254. https://doi.org/10.1016/j.cmpb.2008.04.006

Yılmaz, Z., & Bozkurt, M. R. (2013). Ayrık Dalgacık Dönüşümü Kullanarak Aritmilere Ait Özniteliklerin Çıkarılması, 23–26.

Zhao, Q., & Zhang, L. (2006). ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines. In *International Conference on Neural Networks and Brain, 2005. Icnn&b* (pp. 1089–1092).