Heartbeat Feature Extraction from Vowel Speech Signal Using 2D Spectrum Representation

Dmitriy E. Skopin
Al'Balqa Applied University, Amman, Jordan
m825453@yahoo.com

Sergey U. Baglikov
Corporation "Help MediCom Group", Kursk, Russia
baglikovs@bk.ru

ABSTRACT

ECG is a method using to measure the rate and regularity of heartbeats to detect any irregularity in a heart. ECG translates heart electrical activity into wave-line on paper or screen. Our investigations show that heartbeat modulation exists in human voice signal and can be extracted by transformations in frequency domain. In this paper, the 2D spectrum of human vowel speech will be used to extract the relevant ECG information, where 2D spectrum is a novel proposed method of signal feature extraction.

Key words: Electrocardiogram (ECG), voice signal, spectral transformation, signal feature extraction.

1. Introduction

The health of a population is a fundamental element contributing to progressive sustainable development in all regions of the world. Virtually all sciences contribute to the maintenance of human health and the practice of medicine. Medical physicists and biomedical engineers are the professionals who develop and support the effective utilization of this medical science and technology as their responsibilities to enhance human health care with the new development of the medical tools such as electrocardiogram (ECG). Heart disease is a broad term that includes several more specific heart conditions which are Coronary Heart Disease, Heart Attack, Ischemia, Arrhythmias, Cardiomyopathy, Congenital Heart Disease, Peripheral Arterial Disease (PAD).The most common heart condition is coronary heart disease, which can lead to heart attack and other serious conditions [1].

Electrocardiogram is a method used to measure the rate and regularity of heartbeats to detect any irregularity of the heart. An ECG translates the heart electrical activity into wave-line on paper or screen. The spikes and dips in the line tracings are called P wave, QRS complex, T wave, ST segment and RR interval (period between two sequences of R and R waves of ECG signal) [2].

The heart is a muscular pump made up of four chambers. The two upper chambers are called atria, and the two lower chambers are called ventricles. A natural electrical system causes the heart muscle to contract and pump blood through the heart to the lungs and the rest of the body [5,6].

The ECG detection, which shows the information of the heart and cardiovascular condition, is essential to enhance the patient living quality and appropriate treatment. It is valuable and an important tool in the diagnosing the condition of the heart diseases. In recent year, numerous research and algorithm have been developed for the work of analyzing and classifying the ECG signal.

The classifying method which have been proposed during the last decade and under evaluation includes digital signal analysis, Fuzzy Logic methods, Artificial Neural Network, Hidden Markov Model, Genetic Algorithm, Support Vector Machines, Self-Organizing Map, Bayesian and other method with each approach exhibiting its own advantages and disadvantages. The ECG features can be extracted in time domain [6][7] or in frequency domain [8][9]. Some of the features extraction methods implemented in previous research includes Discrete Wavelet Transform [5], Karhunen-Loeve Transform [10], Hermitian Basis and other methods [11].

All methods of ECG feature extraction based on electrocardiogram that is a noninvasive record of variation of the biopotential signal recorded from human skin surface. The noninvasive technique means that this signal can be measured without entering the body at all. Electrodes are placed on the user's skin area to detect the bioelectric potentials.

In current paper we have discover a new method of heartbeat signal extraction using human voice signal. In this case, we have a vowel sound as source signal that can be recorded by microphone and transmitted by mobile communications for further analysis and this method opens new fields of ECG applications such as heartbeat detection of patient that located in difficult access areas.

2. Establishing the Relation between Heart Activity and Speech Production

As numerous researches show [12], speech signals contain linguistic, expressive, organic and biological information. Acoustic speech output in humans is commonly considered to result from a combination of a source of sound energy (e.g. the larynx) modulated by a transfer (filter) function determined by the shape of the supralaryngeal vocal tract. This combination results in a shaped spectrum with broadband energy peaks. This model is often referred to as the "source-filter theory of speech production" and stems from the experiments of Johannes Müller (1848) in which a functional theory of phonation was tested by blowing air through larynges excised from human cadavers. The supralaryngeal vocal tract, consisting of both the oral and nasal airways, can serve as a time-varying acoustic filter that suppresses the passage of sound energy at certain frequencies while allowing its passage at other frequencies [13]. Formants are those frequencies at which local energy maxima are sustained by the supralaryngeal vocal tract and are determined, in part, by the overall shape, length and volume of the vocal tract [14]. Taking into account the fact, that larynx contains muscles covered by blood vessels connected to human circulatory system, heart beat should be related with dynamical variations of vocal cord parameters such as length and volume that should be directly related with acoustic properties of speech. In our point of view, it is possible to detect changing of speech properties according to heart activity by obtaining a spectrogram of vowel speech signal together with raw ECG data of the same patient located in one coordinate system. The structure of device that allows recording and analyzing speech and ECG signals represented in the figure below:

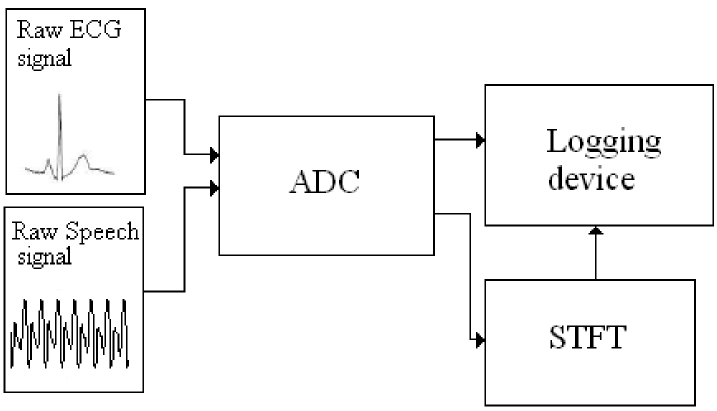

Figure 1. The block diagram of proposed system

The system contains a microphone as a source of raw speech signal and Cardiette AR600 Portable ECG recorder as source of raw EEG signal. Both signals where recorded from patient concurrently during 6 second. Given duration determined as a comfortable time for vowel speech generation by a patient. Analog-to-digital converter of designed system represented as a microphone input of computer soundcard, where left channel connected to the attenuated output of AR600 device, right one to the output of microphone. Recording of signals and processing according to the short time Fourier transform (STFT) was done by software based on Visual C++ 2005 and MatLab. Representations of both signals in time domain represented in the figure 2

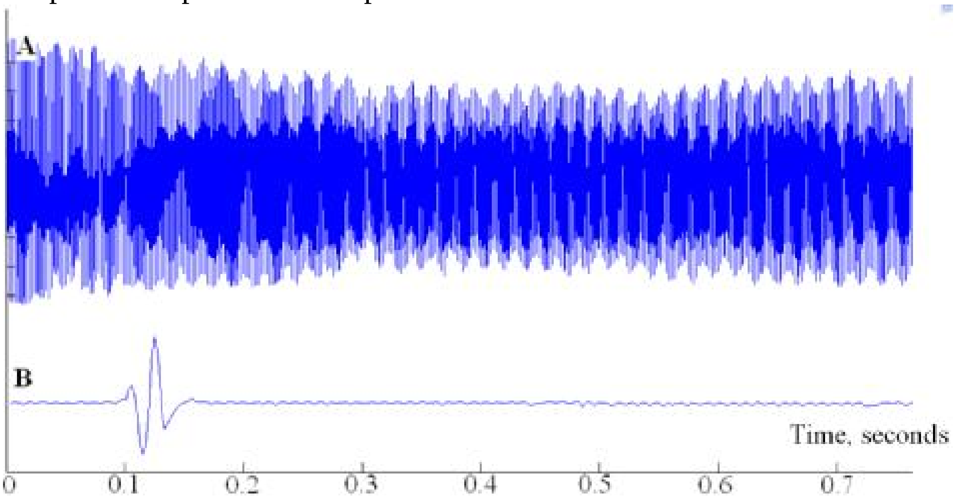

Figure 2. The raw signals in time domain: A – vowel speech signal (vowel /e/ like in the word "email", pronounced by a male), B – ECG signal of the same patient recorded at the same time

The P and T waves of ECG signal suppressed by microphone amplifier that contains low pass filter with cut-off frequency around 40 Hz. The figure 2 represents both signals, recorded with sampling frequency 44 KHz in time domain.

Further investigations of speech signal provided in time-frequency domain according to discrete short-time Fourier transform defined according to the equation 1.

$$X_m(w) = \sum_{n=-\infty}^{\infty} x(n)\omega(n - mR)e^{-jwn} \qquad (1),$$

Where x(n) is input vowel signal, m – size of window, ω - length m Hamming windowing function, R size of overlap.

Size of window should be less than RR interval of ECG signal and in current research (after the set of conducted examinations) was chosen to be equal 2048 signal samples, considering a value of sampling frequency equal 44 KHz gives a spectrum resolution 21Hz. The size of overlaps between signal windows was 1800 samples. The spectrogram of vowel speech sound (vowel /e/) represented in the figure 3.

Horizontal lines of the figure are time evolution of speech formants, it is clearly seen that the heart activity (a moment of R wave appearing) produces a frequency modulation of voice for all formants located within a band 2-6 KHz. The data, represented on the figure above allow to come to a conclusion that it is possible to extract an RR information (a time between 2 sequential R waves) directly, from spectrogram or by additional approach named "the 2D spectrum surface" discovered by us and represented in the chapter below

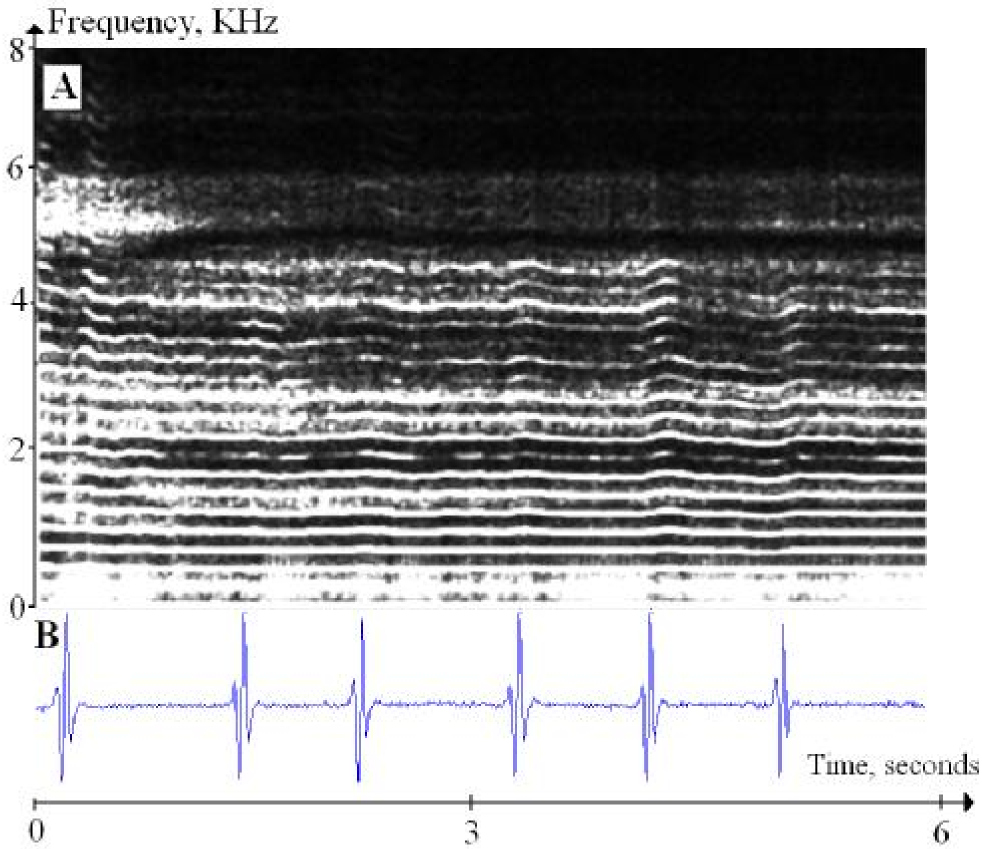

Figure 3. The spectrogram (A) of vowel speech signal represented together with ECG (B) of the same patient

.

3. Heartbeat feature extraction

To extract the heartbeat information from speech spectrogram we have designed a new approach that allows obtaining RR information in 2D spectrum surface. The algorithm of the method represented in the figure 4.

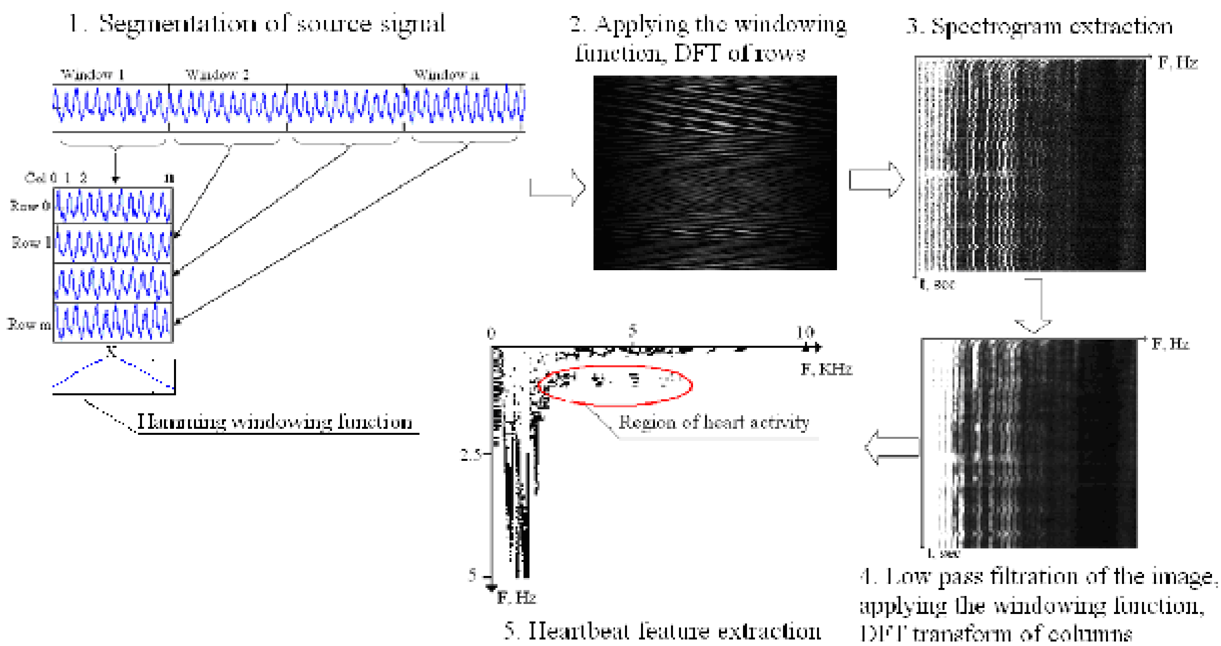

Figure 4. The algorithm of heartbeat feature extraction

The analysis of image in the figure 3 allows concluding that heart activity has an influence upon the speech formant properties, modulating them in frequency domain. Therefore, it is possible to say about a voice frequency modulation by the heartbeat. We have offered a method for heartbeat feature extraction with algorithm that adduced on the figure 4

1. The pickup signal is divided into consecution of windows with length equal 2048 signal samples, overlapped by 24%, forming a 2D matrix of source signal
2. The Hamming windowing function applied in each row of 2D matrix of source vowel speech signal.
3. The Discrete Fourier transform applied in each row of 2D matrix of source signal that created in the step 2 to extract a spectrogram of source signal. Time axis in this case directed to the bottom of computer screen (Y direction), the axis of signal frequency has horizontal (X) direction
4. During this stage, the Image processing was applied using low pass spatial 5X5 filter then result in each column was multiplied by Hamming windowing function. The filter allows to reduce a high frequency noise of the spectrogram without distorting low frequency components of heartbeat ECG.
5. The Discrete Fourier transform applied in each column of spectrogram forming a "spectrum of spectra" 2D surface. The heartbeats represented in this surface as clusters, which have horizontal (X) positions within a band 2-6 KHz that refer to the formants frequencies of vowel speech and vertical (Y) position within a band 0.8-1 Hz that refer to the heartbeat values equal 60-75 beats per minute.

4. Result

The data includes ECG and sound records of 15 patients were processed using designed by as software that implements proposed methods of heartbeat extraction. During examination of designed system, such parameters of system performance as accuracy and measuring error were tested.

$$\text{Accuracy} = \frac{\text{Number of beats correctly classified}}{\text{Total number of beats tested}}$$

Fragments were made by a medical expert by means of interactive analysis of records. During analysis of each ECG fragment, a medical expert visually determined an average value of patient's heart rate then compare results with data, provided by software. Experimental work shows that the average value of system accuracy equal approximately 92%, the measuring error determined as 15%.

5. Conclusion

The ECG detection and heartbeat feature extraction are widely known methods, but in current paper we have discover a new method of heartbeat signal extraction using human vowel speech that can be recorded by microphone and transmitted by mobile communications for further analysis. The heartbeat information may be extracted using proposed method of 2D spectrum representation that is extension of Short Time Fourier transform and also can be used for AM, FM, ultrasound, Doppler sonography, and other signals analysis. The system that implements proposed methods of heartbeat feature extraction shows a reasonable agreement of results of automatic heartbeat allocation with expert estimations.

Further development of suggested methods lies in solving of problems of automatic classification of heart pathology using a vowel speech signal. Technology of artificial neural networks can be effectively used for this problem solving.

References

[1.] Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN. "Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association" //Medicine and science in sports and exercise, ISSN: 0195-9131 2007 Aug;39(8):1423-34

[2.] Taylor-Robinson DC, Milton B, Lloyd-Williams F "Planning ahead in public health? A qualitative study of the time horizons used in public health decision-making"//BMC public health, ISSN: 1471-2458, 2008 Dec 18;8(1):415.

[3.] Sumner W 2nd, Schootman M, Asaro P. "Using county-level public health data to prioritize medical education topics"// The Journal of continuing education in the health professions, ISSN: 0894-1912 2008 Fall;28(4):197-204

[4.] Haskell WL, Lee IM, Pate RR, Powell KE "Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association."// Circulation, ISSN: 0009-7322, 2007 Aug 28;116(9):1081-93.

[5.] Tran Hoai Linh, Stanisaw Osowski and Maciej Stodolski, "On-Line Heart Beat Recognition Using Hermite Polynomials and Neuro-Fuzzy Network", Warsaw University of Technology, 2003.

[6.] Cuiwei Li, Chongxun Zheng, Changfeng Tai, "Detection of ECG Characteristic Points Using Wavelet Transforms", IEEE Transactions on Biomedical Engineering, Vol. 42, No. 1, January 1995.

[7.] Hu, Y.H., Palreddy, S., Tompkins, W., "A Patient Adaptable ECG Beat

Classifier Using A Mixture Of Experts Approach", IEEE Trans.Biomed.

Eng. (44), 891–900, 1997.

[8.] Moraes, J.C.T.B., Seixas, M.O., Vilani, F.N., Costa, E.V., "A Real Time QRS Complex Classification Method Using Mahalanobis Distance" Comput. Cardiol., 2002.

[9.] Costas Papaloukas, Dimitrios I. Fotiadis, Aristidis Likas, Lampros K. Michalis, "Automated Methods for Ischemia Detection in Long Duration ECGs", 2003.

[10.] Franc Jager, "Feature Extraction and Shape Representation of Ambulatory Electrocardiogram Using the Karhunen-Lo`eve Transform" Electrotechnical Review, Ljubljana, Slovenija, 2002.

[11.] David Cuesta-Frau, Juan C, Perez-Cortes, Gabriela Andrea Garcia, Daniel Navak, " Feature Extraction Methods Applied to the Clustering of Electrocardiographic Signals. A Comparative Stucy", Czech Technical University In IEEE Proceeding, 2002.

[12.] J. M. Pickett. "Acoustics of Speech Communication, The: Fundamentals, Speech Perception Theory, and Technology" Allyn & Bacon 1999, ISBN-10: 0205198872

[13.] Browman, C. P. & Goldstein, L. (1990). Representation and reality: Physical systems and phonological structure. Journal of Phonetics, 18, 411-424

[14.] Maton, Anthea; Jean Hopkins, Charles William McLaughlin, Susan Johnson, Maryanna Quon Warner, David LaHart, Jill D. Wright (1993). Human Biology and Health. Englewood Cliffs, New Jersey, USA: Prentice Hall. ISBN 0-13-981176-1.

[15.] J. B. Allen and L. R. Rabiner, ``A unified approach to short-time Fourier analysis and synthesis," Proc. IEEE, vol. 65, pp. 1558-1564, Nov. 1977.