Wavelet Analysis of Compressed Biomedical Signals

Andrey B. Stepanov
The Bonch-Bruevich Saint-Petersburg State University of Telecommunications
Saint-Petersburg, Russia
sabarticle@yandex.ru

Abstract—The paper proposes mathematical apparatus that can be used for wavelet analysis of compressed biomedical signals. As an example of biomedical signals, electrocardiogram and electroencephalogram are considered. A brief description of these signals is given. In the basis of the proposed algorithm of wavelet analysis of compressed biomedical signals lies the use of wavelet decomposition of the signal with the subsequent analysis of approximating coefficients of the set level with the use of continuous wavelet transform and synthesized wavelet. Below is suggested a brief description of the wavelet synthesis procedure for continuous wavelet transform as well as neural network and spline wavelet models proposed by the author. It has been practically proven that application of this algorithm allows us to compress electrocardiogram and electroencephalogram 8 times. In this case possibility to detect the target feature in biomedical signal based on the analysis results of the continuous wavelet transform. Noted, however, that the use of wavelet compression results in a loss of high frequency information in a signal. Therefore, the algorithm must not be applied in cases where the preservation of small fragments in a signal typical of high-frequency components is very important. This algorithm can be applied in the implementation of wavelet analysis of biomedical signals system on mobile devices, where it is important to reduce the amount of stored, transmitted and/or processed information.

I. INTRODUCTION

Wavelets are widely used in the analysis of one-dimensional signals, including biomedical. This is due to their main advantage – the possibility of use in time-frequency analysis. The widespread method in the analysis of one-dimensional signals is algorithms based on continuous wavelet transform (CWT). The results received from continuous signal wavelet transform allow to accurately detect the feature location in the signal, and to determine its type. However, the continuous wavelet transform has a serious drawback – it requires high computational costs. Even with the power of modern element base it can limit the implementation of such systems in the form of mobile devices or lead to their considerable cost.

Meanwhile, the mobile systems allowing to monitor the state of human health, at a high pace are becoming a part of our daily life. It can be a separate device or systems implemented as software installed on a smartphone or tablet. In the latter case, the user simply installs the application and, if necessary, connects the coupler.

The advantages of the biomedical signal analysis system on a smartphone or tablet are:

- Possibility to use computing resources and the interface device
- In some cases, cost reduction
- The ability to ensure user secrecy etc

Biomedical signals may include:

- Electroencephalogram
- Electrocardiogram
- Electrogastrogram
- Other signals

Let us focus on the electroencephalogram (EEG) and electrocardiogram (ECG). The choice of these two types of signals is related to the fact that they allow monitoring the status of two major systems of the human body: the central nervous system (one of its components – the brain) and cardiovascular system (the heart).

EEG analysis on a mobile device (a specialized device, smartphone or tablet) is much more complicated than ECG analysis. In particular, this problem is related to the correct placement of the electrodes on the patient's head. But this problem can be solved. It is necessary to change the concept of the electrodes. They can be performed in the form of small hairpins, grids, etc. Smart technologies can significantly reduce the number of EEG electrodes for express analysis.

Mobile phone occupies more and more space in human life. Transmission of voice and text messages, storage and music playback, books, movies, games, etc. A person subconsciously wants to have everything you need in one device. In this regard, the development of mobile biomedical signal analyzing system, including EEG, can be justified.

Analyzing the reference sources on wavelet analysis of biomedical signals, we can divide them into the following groups:

1) Research papers dedicated to wavelet compression of biomedical signals. Typically, in such studies, discrete or packet wavelet decomposition can be used. Their main goal is to obtain the highest possible signal compression rate. Some papers [1] show research that has made it possible to obtain a signal compression more than twentyfold. Such results can be achieved with the use of threshold processing of decomposition coefficients. In fact, often most of the decomposition coefficients contribute insignificantly to the signal and do not need to be used in its reconstruction. Removing such coefficients allows us to obtain significant signal compression. However, from the details of research, it is not clear how the
signal restoration should be carried out. As a rule, the appearance of such coefficients is random. Therefore, in order to recover the signal, it is also necessary to store information about the serial numbers of the deleted coefficients. This will lead to a significant reduction in the real gain when compressing and complicating the algorithm. More universal methods are those based on discrete wavelet decomposition with the removal of all coefficients of a given type. They allow us to get a small but fixed compression rate.

2) Research papers dedicated to the identification of information signs in a biomedical signal. In this case, as a rule, continuous wavelet transform and traditional wavelet families are used [2]–[4]. At the same time, the need to create new mother waves for continuous wavelet transform is often emphasized [5].

3) Research papers dedicated to the evaluation of characteristics of wavelet coefficients obtained after performing wavelet decomposition [1]. This research is focused on the specific nature of particular types of signals.

In most types of research with the use of continuous wavelet transform, the problem of its implementation on element base with a limited computational resource and the ways to solve this problem are not considered.

The author proposes a mathematical apparatus based on the use of continuous and discrete wavelet transform. It can be used for the implementation of biomedical signal analysis mobile system. This takes into account the basic requirements for such apparatus:

- High precision of biomedical signals analysis.
- Compression of the biomedical signal to enable compact storage or transmission to a remote computer for more detailed analysis (if necessary).

II. BIOMEDICAL SIGNALS

Electrocardiogram (ECG) is a multi-dimensional signal which can be recorded from the body surface and can serve as a basis for cardiac arrhythmias diagnosis [6].

Fig. 1, a) shows the basic form of a normal ECG.

![Diagram of ECG](image)

**Fig. 1. Electrocardiogram: the basic form of a normal ECG a), 10-second-length ECG fragment b)**

It includes [7]: P, Q, R, S, T oscillations which will also be called waves. Especially, you can distinguish the so-called QRS complex, which together form the respective waves.

Fig. 1, b) shows a 10-second-length ECG fragment. Signal sampling frequency is 250 Hz. The fragment was downloaded from the electronic library of biomedical signals [8]. It will be used in the future to demonstrate the possibilities of the proposed mathematical apparatus.

Electroencephalogram (EEG) is a signal that can be detected from the surface of a human head and is the result of brain neurons electrical activity [9]–[11].

In the EEG analysis, it is important to identify its basic rhythms, as well as a number of features. These features include artifacts and graphoelements typical of pathology. Artifacts are phenomena not directly related to the human brain activity. They can cause harm during the analysis. They must be identified and taken into account in the preparation of clinical judgment.

In order to identify the basic EEG rhythms, it is usually enough to use methods based on the Fourier Transform. Detecting special features in the signal is normally more difficult. To demonstrate the capabilities of the proposed mathematical apparatus, we are going to consider an ocular artifact as a special feature (Fig. 2, a).

Further, in the analysis, we are going to use 5-second-length ECG fragment, containing several ocular artifacts (Fig. 2, b).

Signal sampling frequency is 250 Hz.

III. MATHEMATICAL APPARATUS

A. Continuous wavelet transform

Wavelets are a generic name for special functions with a zero integral value localized along the time axis, able to shift along it and to scale [12]–[18].

Wavelets are widely applied in the analysis of one-dimensional signals, among which are biomedical signals discussed previously [1], [19]. Thus, most often continuous wavelet transform (CWT) is used in the analysis of signals to detect features in them.

![Diagram of EEG with artifact](image)

**Fig. 2. Electroencephalogram: EEG fragment with ocular artifact a), 5-second-length ECG fragment with ocular artifacts b)**
The formula of continuous wavelet transform function $f(t)$ is as follows:

$$ W(a, b) = \frac{1}{|a|^{1/2}} \int f(t) \psi\left(\frac{t-b}{a}\right) dt, $$

(1)

where $\psi(t)$ is wavelet, $a$ is scale, $b$ is shift parameter.

The main advantage of the continuous wavelet transform is that it can be used in algorithms to detect information signs in the signal.

Continuous wavelet transform results in a number of coefficients obtained for different values of scale and shift.

The key point during the continuous wavelet transform is a wavelet choice. The importance of such a choice follows from the analysis of the formula of continuous wavelet transform (1). The more accurate the wavelet will coincide with the imposition of a signal part, the more localized response to an information sign will be achieved.

The results of the continuous wavelet transform can be displayed as a graph. This graph is called wavelet spectrogram (Fig. 3). Based on wavelet spectrogram, time-frequency signal analysis can be performed. This is achieved by simultaneously obtaining information about the size of signal components (frequency components) and their position in time. Small scale values $a$ transmit information about high-frequency signal components (small components), and large scale values $a$ transmit information about low-frequency signal components (large components). In case when a wavelet with predetermined scale coincides with the signal component, it is displayed on spectrogram as a localized light area.

The main disadvantage of the continuous wavelet transform is a need for high computational cost to run it. This is due to some redundancy of CWT. Even with the limitation of scale levels, it can still have a negative impact on its application in signal processing algorithms with limited computing resource of element base.

Thus, we can conclude that the efficiency of the continuous wavelet transform in the analysis of biomedical signals based on the element with a limited computing resource is affected by the following factors:

- Selection of the wavelet, close to the type of features that provides a good localization on the wavelet-spectrogram and the accuracy of its detection in the signal
- Reducing the redundancy in the continuous wavelet transform

B. Models of wavelets for continuous wavelet transform

The studies conducted by the author [13] have shown that the traditional families of wavelets do not provide the variety of functions, which would allow to accurately detect all the features in biomedical signals.

The problem of selecting wavelet suitable for identifying a specific feature can be solved by means of its synthesis.

Let us formulate the main stages of the wavelet synthesis procedure for continuous wavelet transform:

1) Selecting a fragment that will serve as the basis for the wavelet. Such a fragment is called a sample.
2) Sample modification in order to provide all zero values on its edges.
3) Mathematical description of the sample to obtain a formal representation of the wavelet.
4) Functions check to make sure they satisfy the condition of admissibility for the wavelet:

$$ C = \int \|\psi\|^2 |\omega|^{-1} d\omega < \infty. $$

In practice, it is enough for the function to have zero integral value.

When the function satisfies this condition, it can be considered wavelet suitable for continuous wavelet transform.

5) If the resulting function does not satisfy the condition of admissibility, it is subjected to further modification. Then steps 2–4 are repeated until it allows us to obtain the corresponding function.

6) Wavelet valuation.

The key point in the wavelet synthesis is the choice of the method of sample mathematical description.

Using approximation in the synthesis of wavelets by algebraic polynomials [1] does not allow to obtain the necessary feature localization on wavelet-spectrogram.

The author has developed neural network and spline models [20], [21], which can be used in the synthesis of wavelets for continuous wavelet transform. These mathematical models allow to receive the wavelets that can be used in the continuous wavelet transform and provide localized mapping of the features on the wavelet-spectrogram.

Below is the general information and comparative analysis of the data for these models.
Neural network models are based on the use of the mathematical description of the sample approximation performed by means of artificial neural networks.

As is known [22], artificial neural networks can be regarded as universal approximators.

The study conducted by the author proves that multilayer perceptrons with one and two hidden layers, as well as artificial neural network based on radial basis functions can be successfully used in the synthesis of wavelets for continuous wavelet transform [20].

In order to demonstrate it, we chose ECG fragment as a sample (Fig. 4).

The minimum accuracy of approximation to the wavelet pattern was achieved using a multilayer perceptron with one hidden layer (Fig. 5, a). The standard deviation was 0.0215 mV. A significant deviation from the sample wavelet is a disadvantage. However, such an artificial neural network allowed us to obtain a model with a small number of parameters that can be considered its advantage.

The total number of parameters, taking into account the normalizing coefficient was 45. The network contains 21 neurons in the hidden layer and 1 in output layer. This number of neurons was obtained in a practical series of experiments and corresponds to the minimum standard deviation of the wavelet model for this type of neural network.

Hyperbolic tangent is used in the hidden layer neurons as the activation function. Linear activation function is used in the output layer neuron.

A more accurate approximation was obtained by using a multi-layer perceptron with two hidden layers (Fig. 5, b). The standard deviation of the wavelet sample was 0.0182 mV. However, the number of mathematical model parameters increased to 73.

Such neural network has two hidden layers and one output layer. Hyperbolic tangent is used in the hidden layer neurons as the activation function, linear activation function is used in the output layer neuron.

Maximum approximation accuracy of the obtained wavelet compared to the sample was achieved by using artificial neural network based on radial basis functions (Fig. 5, c). The standard deviation was $1.68 \times 10^{-12}$ mV. However, this model is complex and has 403 parameters.

Wavelet spline models for continuous wavelet transform are based on the use of cubic spline sample in the process of interpolation [21]. The use of splines allows us to obtain a formalized representation of the wavelet on the one hand, and on the other hand, it provides high accuracy of the wavelet approach to the original sample.

Fig. 6 shows the interpolation procedure. It is obvious that the resulting wavelet will exactly match the sample, since the interpolation cubic spline will pass through the interpolation points corresponding to the values of the sample frame. A slight approximation error that may occur is related to the modification of the sample at its edges, in accordance with the requirements of paragraph 2 of the synthesis algorithm. It cannot be eliminated.

High precision of wavelet approximation to the sample is the main advantage of spline models. The main disadvantage is their high complexity. This model has a significant number of parameters, which is 4 times the number of sample readings. For the sample examined, the number of parameters was 801.

Comparing the neural network models to spline models, one can note the high accuracy of the wavelet approximation to the sample by using artificial neural networks approximation based on radial basis functions and a guaranteed accuracy while using cubic spline interpolation. However, an important factor which greatly affects the possibility of application of these models in the implementation of continuous wavelet transform calculation algorithm, is the complexity of these models. The final choice of the model should be based on the computing power of an element base.
C. Wavelet analysis of compressed biomedical signals

The requirements for computing resources of the element base used can be reduced by decreasing the amount of information processed. Reducing the number of biomedical signal channels by reducing the number of leads may reduce the accuracy of ongoing research. Simple reduction of the number of readings by reducing the sampling rate would result in a loss of signal registration accuracy and impossibility of taking into account the impact of temporary reading.

In order to solve this problem, we need a method which would allow us to reduce the number of processed readings, as well as keep the possibility to analyze such a signal to detect information characteristics.

In this paper, we propose an algorithm for the analysis of biomedical signals in their compressed representation.

The basis of this algorithm is to use in the first stage the discrete wavelet decomposition of signal followed by continuous wavelet transform of approximating coefficients with the use of synthesized wavelets in the second stage.

Let us study the algorithm in more detail.

The algorithm of compressed biomedical signal wavelet analysis (Fig. 7) is a modification of the two-level algorithm of electroencephalogram wavelet analysis, developed earlier by the author [19]. This algorithm includes the following steps:

1) Registration of biomedical signal. At this step, the signal is recorded.

2) Signal is separated into fragments of fixed length. This step allows us to reduce the requirements for computing power of the element base on which the algorithm is implemented.

3) The wavelet is decomposed into approximating and detailing (if necessary) coefficients with a given number of decomposition levels. At this stage, discrete wavelet transform is used.

Fig. 8 shows a conventional single-level diagram of wavelet decomposition of the signal. The signal is fed into two decomposition branches. The upper branch contains a cascaded low-frequency decomposition filter (LFDF) and decimator with coefficient 2. The lower branch contains a cascaded high-frequency decomposition filter (HFDF) and decimator with coefficient 2. Decimator removes even count, and in the other cases – uneven count. Thus, the number of coefficients at the output of each branch is reduced by half. The coefficients on the output of the upper branch are called approximating coefficients. They describe low-frequency signal components. The coefficients on the output of the lower branch are called detailing coefficients. They describe high-frequency signal components.

Signal wavelet decomposition is reversible. The original signal can be reconstructed on the basis of approximating and detailing coefficients. For this purpose, restoration filters are used, and the missing coefficients are zero padded in the interpolation process.

Wavelet decomposition may be performed repeatedly. In such case approximating coefficients of the previous level of decomposition are subjected to decomposition. At every level the number of coefficients in each branch is reduced by half.

Studies have shown that for the majority of biomedical signals detailing coefficients of several decomposition levels do not contain useful information. They can be removed. And in case of signal recovery, the deleted detailing coefficients can be supplemented with zeros.

By removing detailing coefficients of selected levels, one can get a significant reduction in the volume of stored information.

Fig. 7. The algorithm of wavelet analysis of compressed biomedical signal

Fig. 8. Single-level wavelet decomposition scheme
Thus, for an electroencephalogram recorded at a sampling frequency of 250 Hz the signal was compressed by 8 times without significant loss of quality in the original signal. In this case detailing coefficients of three levels of decomposition were removed. The electrocardiograms recorded at a sampling frequency of 250 Hz was compressed 8 times. Thus detailing coefficients of three levels of decomposition were removed.

Obviously, wavelet compression is a compression with losses. Therefore, when selecting the required number of decompositions, it is important to consider the type of signal, its characteristics, including the sampling rate.

Retaining detailing coefficients is only necessary if we want the exact signal recovery. In case when the recovery is not required or there are no strict requirements for accuracy of recovery, the lower branch of the decomposition chain may be missing.

4) Allocation of approximating coefficients of a given level of wavelet decomposition. At this stage, approximating coefficients of the last level of decomposition are detected. They will be subjected to further analysis.

5) Primary analysis of the signal fragments using a continuous wavelet transform and traditional wavelet families.

At this level, "rough" wavelet analysis of biomedical signal using one of the wavelet is performed. Such a wavelet can be represented by wavelet "Mexican hat" (Fig. 9). Studies have shown [13] that this wavelet is well suited for the analysis of electroencephalogram. Due to its characteristics, it can be used for analysis of other smooth biomedical signals including electrocardiogram.

In addition, this wavelet has formalized representation which is a rare case for wavelets of traditional families. A formalized representation of the wavelet is a significant advantage in the implementation of algorithms for continuous wavelet transform. After performing a continuous wavelet transform with the use of this wavelet, the obtained wavelet coefficients are analyzed. In the simplest case, the threshold or multithreshold processing can be used. The result of analysis is the selection of signal fragments with original features and their primary classification.

6) Secondary (qualifying) signal fragments analysis using a continuous wavelet transform and synthesized wavelet.

After the primary analysis of biomedical signal, the secondary wavelet analysis of selected fragments is carried out in order to clarify the type of features. So in the continuous wavelet transform synthesized wavelets are used.

Fig. 9. "Mexican hat" Wavelet

Fig. 10 shows the results of continuous wavelet transform of electrocardiogram and electroencephalogram fragments compressed eightfold using wavelets synthesized on the respective samples. In the synthesis, spline wavelets models have been used. These wavelets have formalized representation.

As it follows from the figure, the data of wavelet-spectrogram has sufficient localizing capacity to detect features on them.

Since each feature has its own basic frequency, it is characterized by an appropriate response to the synthesized wavelet, which is most explicit for the coefficients of the scale, corresponding to the given frequency.

Thus, in the simplest case, the threshold processing may be used. For features having a complicated shape, more sophisticated methods of analysis can be used, including the ones based on application of learning systems.

It should also be noted that the used procedure of wavelet compression leads to losses of high-frequency component of the signal. Therefore, the proposed algorithm should not be used in the analysis of signals that requires the preservation of small fragments of the signal typical of the high-frequency signal components.

7) Forming a list of features by type and their location in the signal.

As a result of the two levels of analysis a list of features is formed with the obligatory indication of their type and location in the signal.

8) Making a report on biomedical signal analysis.

In the report on the analysis of biomedical signal its basic rhythms, features and other signal parameters, specific to its type, are taken into account. Comprehensive assessment of information received is possible. In this case, a brief report indicating the warning information for the patient can be obtained.
IV. CONCLUSION

The paper considers the mathematical apparatus which can be the basis for the algorithm of biomedical signals analysis. In practice, the following results were obtained:

1) The possibility to compress ECG 8 times. The sampling frequency of the original signal was 250 Hz.

2) The possibility to wavelet-analyze the eightfold compressed electrocardiogram with the use of synthesized wavelets. In the synthesis of wavelets spline wavelets model has been used. The fragment of the ECG has been selected as a sample.

3) The possibility to compress ECG 8 times. The sampling frequency of the original signal was 250 Hz.

4) The possibility to wavelet-analyze the eightfold compressed electroencephalogram with the use of synthesized wavelets. In the synthesis of wavelets spline wavelets model has been used. The fragment of the EEG with ocular artefact has been used as a sample.

5) Applying the proposed algorithm is not possible when it is needed to maintain high-frequency component of biomedical signal. In the process of wavelet decomposition when removing detailing coefficients, it results in the loss of information about small detail of signals typical of the high frequency components.

6) The choice of number of wavelet decomposition levels is carried out based on the type of signal, the sampling frequency and its upper frequency which must be preserved.

7) In the synthesis of wavelets for continuous wavelet transform, used in the second level of wavelet analysis, neural network and spline models can be used. The choice of a specific type of model is determined by its computing capabilities used in the implementation of element base. In the absence of such restrictions spline s can be used, allowing us to obtain a high accuracy of approximation of the synthesized wavelet to the sample. In case where it is impossible to use a large number of model parameters, neural network wavelet models can be used. In assessing the proximity of the synthesized wavelet to the sample qualitative and quantitative evaluation is used.

8) In evaluation of the results of continuous wavelet transform, qualitative assessment is used based on the subjective opinion of an expert.

ACKNOWLEDGMENT

The author thanks the City Epileptic Center of Saint-Petersburg for consultation on the study of the visual analysis of EEG features.

REFERENCES


