BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, November 2014. Vol. **11**(Spl. Edn. 1), p. 377-380

Signal Waveform Extraction in the Presence of Regular and Random Noise

Diana Konstantinovna Avdeeva, Oleg Nikolaevich Vylegzhanin, Mariya Aleksandrovna Yuzhakova, Sergey Anatol'yevich Rybalka, Grigoriev Michael Georgievich and Nikita Vladimirovich Turushev

National Research Tomsk Polytechnic University, Russia, 634050, Tomsk, Lenin Ave., 30.

doi: http://dx.doi.org/10.13005/bbra/1489

(Received: 27 September 2014; accepted: 10 October 2014)

The paper focuses on the problem of the signal waveform extraction in the presence of random and regular noise. The principal component analysis has been proposed to extract the waveform. Assuming that the analyzed signal in the recorded sequence is repeated with a certain periodicity, several portions containing the analyzed signal can be extracted using the "caterpillar" method. The obtained matrix is then subjected to singular value decomposition. It is shown that the waveform is defined by the first left singular vector. Mathematical modeling demonstrates the possibility to extract the waveform of the analyzed signal in the presence of random and regular noise. The model calculations prove the possibility to extract the signal waveform in case the level of random noise and the correlation of the extracted signal and regular noise change within a wide range.

Key words: Signal analysis, regular and random noise, Waveform recovery, principal component analysis.

The problem of the waveform recovery for signals in the presence of various types of noise can be found in a variety of applications. Different filtering techniques are typically used to solve the problem[1, 2]. However, some a priori information on the features of the extracted signal and noise is essential to develop an effective filter.

The principal component analysis is considered to be an effective way to reduce the dimensionality of the problem or to identify the main factors influencing the response function[3-5]. We propose to apply the principal component analysis to extract the waveform of the unknown signal.

* To whom all correspondence should be addressed.

Signal forming

Let us assume that the time sequence with the analyzed signal repeated with a certain periodicity has been recorded. Define the portions of the recorded sequence containing the analyzed signal. Also, assume that each of the portions may be represented as:

$$c=e+\varepsilon$$

where c is a portion of the recorded sequence; e is the analyzed signal; [epsilon] is random noise.

Form the matrix with columns being the defined portions of the sequence. The proposed technique implemented by us in[6-10] is similar to the transformation of the time series into a matrix which is called "caterpillar" or Singular spectrum analysis (SSA)[8].

The matrix A can be represented[8] in the form of the singular value decomposition:

$$A = U\Lambda V^T$$

where U and V are unitary matrices of the right and left singular vectors, respectively; [Lambda] is a diagonal matrix of the singular values. The decomposition can be written as:

$$A = \sum_{i=1}^{n} U_i \lambda_i V_i^T$$

where U_i and V_i are i-th singular vectors; $[\text{lambda}]_i$ is a singular value.

Since norms U_i and V_i are equal to one, then the norm is

$$\left\| U_i \lambda_i V_i^T \right\| = \lambda_i$$

Hence:

$$\|A\| = \sum_{i=1}^{n} \lambda_i$$

The relation $\lambda_i \Big/ \sum_{i=1}^{n} \lambda_i$ determines the contribution of the i-th expansion term to A. If this contribution is large enough (~ 90%), the left singular vector corresponding to the maximum $[\text{lambda}]_i$ is close to vector e.

Let $U_i \lambda_i V_i^T$ be the i-th component of the decomposition. It is clearly seen that the j-th column of the matrix takes the form $U_i \lambda_i V_{ij}$ where v_{ij} is the j-th element of the vector V_i. Thus, each column of the singular value decomposition component is equal to vector U_i up to a constant factor. Hence, each column of the initial matrix can be represented as:

$$A_j = \sum_{i=1}^{k} U_i \lambda_i v_{ij}$$

i.e., as a linear combination of the left singular vectors of decomposition.

To validate the proposed algorithm, the model calculations have been carried out. Figure 1 shows the predetermined signal.

The matrix with columns containing the predetermined signal and random noise as a signal-to-noise ratio (SNR) calculated as:

$$\sqrt{\frac{e^T e}{(c-e)^T (c-e)}}$$

and equal to 4.17 is created. The initial matrix contains 20 columns. Each column contains 1700 discrete points.

The singular value decomposition of the obtained matrix is computed[10]. In this case, the first singular value is more than 84% of the sum of the singular values.

The extracted signal is formed from the singular decomposition components in the form:

$$e = U_1 \cdot \lambda_1 \cdot V_1 \cdot \bar{1}$$

where U_1 is the first left singular vector; $[\text{lambda}]_i$ is the first singular value; V_1 is the first right singular vector; $\bar{1}$ is a vector of ones.

Figure 2 shows the graphs of the predetermined and extracted signals.

As seen in Figure 2, the waveform of the extracted signal corresponds to the waveform of the predetermined signal. After processing, SNR is equal to 16.67.

Consider the performance of the proposed algorithm in the presence of random and regular noise. For this purpose, we create a matrix with

Table 1. Pair correlation coefficients of the predetermined signal and regular noise and the calculated values of the goodness-of-fit for the vector of the predetermined signal and normalized first left singular vector

Pair correlation coefficient	-0.035	0.031	0.097	0.162	0.228	0.294	0.359
Goodness-of-fit value χ^2 for a signal	0.0057	0.0058	0.0058	0.0057	0.0057	0.0056	0.0057

Table 2. The goodness-of-fit values for the vector of the predetermined signal and normalized first left singular vector at different noise levels

Noise coefficient	0.01	0.05	0.1	0.2	0.3	0.4	0.5
Goodness-of-fit value x^2 for a signal	0.0057	0.0058	0.0058	0.0058	0.0058	0.0058	0.0058

columns containing the other regular signal in addition to the predetermined signal and random noise. Figure 3 illustrates this additional signal.

The obtained matrix contains 10 columns with 1700 entries in each column. The norms of the predetermined signal, regular noise and random noise are 4770, 363 and 12.35, respectively. A pair correlation coefficient of the signal and regular noise is 0.035, that is, the vectors of these signals are substantially orthogonal. In the absence of random noise, after the decomposition of the matrix in singular values, the first left singular vector coincides with the original dominant signal up to a constant factor, and the second left singular vector coincides with the vector of the regular noise. The calculated values of goodness-of-fit X^2 are equal to 0.0057 and 0.0022, respectively.

A number of vectors of the regular noise with different pair correlation coefficients are created in order to assess the effect of the correlation of the predetermined signal and regular noise on the possibility to extract the predetermined

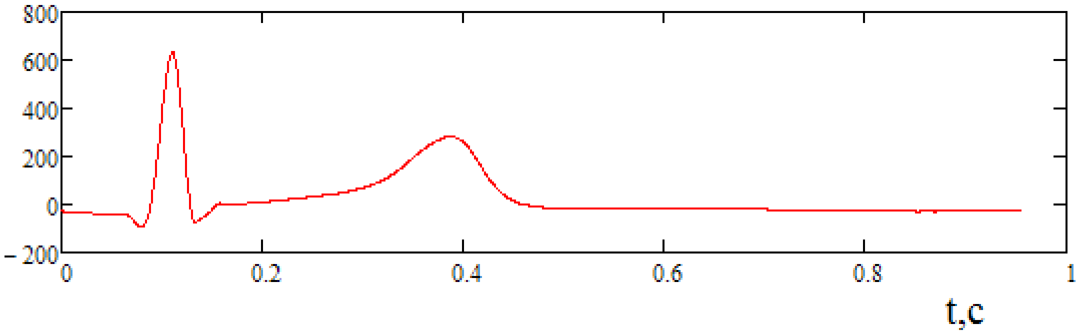

Fig. 1. Graph of the predetermined signal

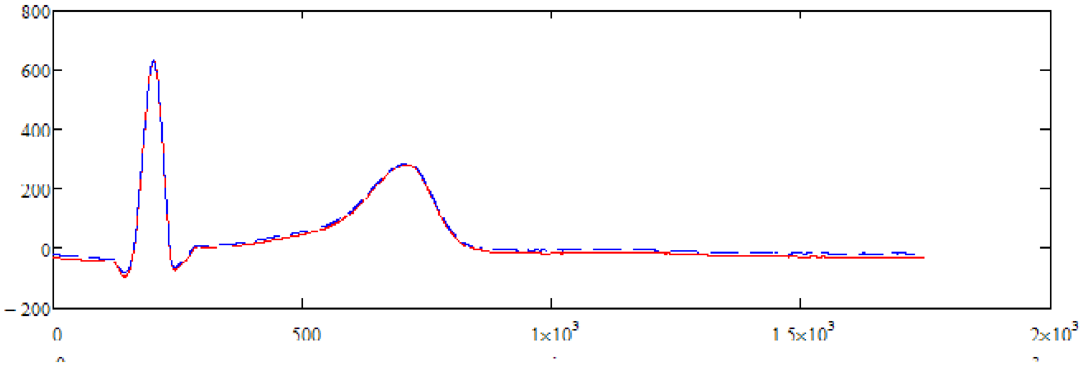

Fig. 2. The predetermined (firm line) and extracted (dash line) signals

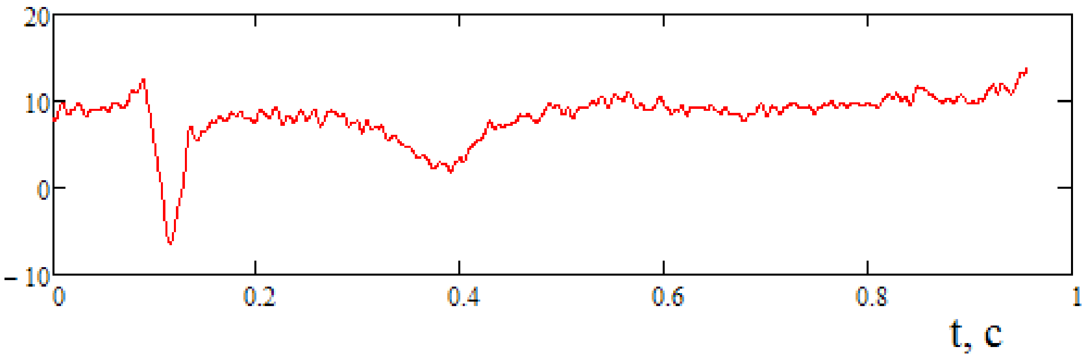

Fig. 3. Graph of the additional signal

signal using the described procedure. Table 1 presents the pair correlation coefficients of the predetermined signal and regular noise and the goodness-of-fit values for the vector of the predetermined signal and normalized first left singular vector.

As seen from Table 1, the increase in correlation of the predetermined signal and regular noise does not affect the proximity measure of signal and normalized first left singular vector.

The corresponding goodness-of-fit values for the dominant signal and first normalized singular vector are calculated at different levels of random noise to assess the effect of random noise on the proximity measure of the predetermined signal and first left singular vector. The level of random noise is controlled by multiplying the noise by a constant factor. The results are presented in Table 2

As seen from Table 2, the increase in the noise level by a factor of 50 does not affect the shape of the first singular vector. In all cases, the waveform of the predetermined signal can be easily recovered. The signal-to-noise ratio is 16.5-20.0.

RESULTS AND DISCUSSION

The obtained results show the possibility to define the waveform of the periodic signal by the left singular vector computed after the decomposition of the matrix, which is obtained from the recorded sequence using the “caterpillar” method, in the presence of regular and random noise. In this case, change of the level of random noise and correlation of the extracted signal and regular noise within a wide range do not have a substantial effect on the waveform of the extracted signal.

ACKNOWLEDGEMENTS

The research was financially supported by the Federal Target Program “Research and development on priority directions of scientific-technological complex of Russia in 2014-2020 years”, project #. 14.578.21.0032 “Development of an experimental model of the hardware-software complex for noninvasive recording of the heart micropotentials in a broad band without filtering and averaging in real time in order to detect the early signs of sudden cardiac death”. The unique identifier of the contract: RFMEF157814X0032.

REFERENCES

1. Poll D.S., F.E. Marchlinski, R.A. Falcone, M.E. Josephson and M.B. Simson, Abnormal signal-averaged electrocardiograms in patients with nonischemic congestive cardiomyopathy: relationship to sustained ventricular tachyarrhythmias. *Circulation*, 1985; **72**: 1308-1313.
2. Grigoriev, M.G., N.V. Turushev and D.K. Avdeeva, Human body electroneuromyographic diagnostics device. Journal of Radio Electronics, 4. Date Views 24.07.2014 jre.cplire.ru, 2014.
3. Laciar E. and L. L. Orosco, Analysis of ventricular late potentials in high resolution ecg records by time-frequency representations. *Latin American Applied Research,* 2009; **39**: 255–260
4. Latfullin I.A., Z.F. Kim, G.M. Teptin, Ventricular late potentials. *The Herald of Arrhythmology*, 2009; **53**: 44-55
5. Sung-Nien Yu, Kuan-To Chou, Selection of significant independent components for ECG beat classification. Expert Systems with Applications: *An International Journal,* 2009; **36**: 2088–2096
6. Zarzoso V.and P. Comon, Robast independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size. IEEE Trans.on neural networks, 2010; **21**: 248–261
7. Yoshida K., A. Chugh, E. Good *et al,* A critical decrease in dominant frequency and clinical outcome after catheter ablation of persistent atrial fibrillation. *Heart Rhythm,* 2010; **7**: 295–302
8. Simoliuniene R., A. Krisciukaitis, A. Macas, G. Baksyte, V. Saferis, R. Zaliunas, Principal Component Analysis Based Method for Detection and Evaluation of ECG T-Wave Alternans. *Computers in Cardiology* 2008; **35**: 757-760.
9. Martis R. J., U. Rajendra Acharya, K. M. Mandana, A. K. Ray, C. Chakraborty, Application of principal component analysis to ECG signals for automated diagnosis of cardiac health. *Expert Systems with Applications*. 2012; **39**: 11792-11800
10. Ikramov, Kh.D., Effective algorithms for decomplexifying a matrix by unitary similarities or congruences. Mathematical notes, 2012; **92**(6). Date Views 24.07.2014 link.springer.com.
11. Bakhvalov, N.S., Numerical methods. Moscow: Nauka, 1987; 636.