weighted averaging,
noise attenuation,
quasi-cyclic biomedical signal

Alina MOMOT[1] , Janusz JEZEWSKI[2], Janusz WROBEL[2]

DATA PARTITIONING BASED WEIGHTED AVERAGING FOR NOISE SUPPRESSION IN BIOMEDICAL SIGNALS

In the case of biomedical signals with a quasi-cyclic character, such as electrocardiographic signals, the high resolution electrocardiograms or electrical potentials recorded from the nervous system of patients (estimating brain activity evoked by a known stimulus), as a method of averaging in the time domain may be used for noise attenuation. In this paper there is presented input data partitioning applied to a few different methods of weighted averaging. This procedure usually leads to improve the quality of the resulting averaged signal, especially when fuzzy partitioning is used. Below it is presented the computational study of weighted averaging with traditional (sharp) and fuzzy partition of the input data in the presence of non-stationary noise. The performance of presented methods is experimentally evaluated for analytical signal of EN 60601-2-51 (2003), namely ANE20000 ECG record.

1. INTRODUCTION

The presence of interference in biomedical signals is associated with the specific acquisition of these signals. For example in the case of bioelectric signals, which are widely used in most fields of biomedicine, disturbances may come from the acquisition hardware, a powerline or the bioelectric activity of body cells. The electric field propagates through the tissue and can be acquired from the body surface, eliminating the potential need to invade the biosystem. However, using surface electrodes results in high amplitude of noise and the noise should be suppressed to extract a priori desired information [3]. Particularly difficult case is when fetal heart rate measurement is needed. One of the most commonly used techniques of fetal heart rate measurement is a pulsed Doppler ultrasound method, although the exact cardiac cycle can be measured only on a basis of electrical activity signal – the fetal electrocardiogram [8]. Recording of fetal electrocardiogram can be accomplished by non-invasive method where measuring electrodes are placed on maternal abdomen and then the main problem is the suppression of maternal electrocardiogram, many times exceeding the useful signal component [7, 9].

There are many approaches to the noise attenuation problem while preserving the variability of the desired signal morphology. One of the possible methods is low-pass filtering such as arithmetic mean. It is very simple method, the classical band-pass filtering, but also very ineffective because the frequency characteristics of signal and noise significantly overlap. Therefore there are developed other methods of noise suppression based on transforming the input space of signal and creating a new space with the help of discrete cosine transform [18] or wavelets transform [1], based on fuzzy nonlinear regression [13], nonlinear projective filtering [10], higher-order statistics at different wavelet bands [19] or extreme points determination by mean shift algorithm and dynamical model-based nonlinear filtering [20].

In the case of quasi-cyclic biomedical signals, such as electrocardiographic signals, the high resolution electrocardiograms or electrical potentials recorded from the nervous system of patients (estimating brain activity evoked by a known stimulus), another possible method of noise attenuation is the synchronized averaging [6]. The method assumes that the given biomedical signal is quasi-cyclic with the additive noise, which is independent and with zero mean. Performing averaging could be done by simple arithmetic mean or its generalization, namely weighted mean where the weights are tuned by some adaptive algorithm.

[1] Silesian University of Technology, Institute of Informatics, Akademicka St. 16, 44-100 Gliwice,
email: alina.momot.@polsl.pl.
[2] Institute of Medical Technology and Equipment, Biomedical Signal Processing Department, Roosevelt St. 118, 41-800 Zabrze, Poland.

Traditional arithmetic averaging technique can be used in the case of stationary noise (it is constant level of power noise throughout the averaging period). Unfortunately, the physiological noise level often varies in tests, even if strict artifact rejection is applied [4] and increasing the input noise level during the averaging process results in increasing the residual noise and deteriorating quality of the averaged signal. Thus, using methods of weighted averaging is motivated by the reason that most types of noise are not stationary and the variability of noise power can be observed.

The crucial point in the problem of weighted averaging is estimation of the weights, which may be solved by numerous approaches, such as methods based on the minimum energy principle [5], Kalman filter theory [11], adaptive estimation of the weights [2], criterion function minimization [12] or Bayesian inference [14]. In [15], there were published the results of experiments which show supremacy of the methods using Bayesian inference over the ones using criterion function minimization with regard to both the synthetic ECG signal and the real ECG signal as well as with regard to different noise types. In [16] there was presented modification of existing Empirical Bayesian Weighted Averaging method with extension by partitioning of the input data in the time domain and the numerical experiments show supremacy of the new method over the original Bayesian method [17].

This paper shortly describes a few weighted averaging methods and in details how the partition of the signal is made and how to combine the results in one averaged signal. It is shown that the partitioning may be performed by using traditional (sharp) or fuzzy membership function and the partition may be applied to all presented methods. The aim of the paper is to study the influence of the type of partition and the number of parts on the resulted signal. The performance of the methods is experimentally evaluated for analytical signal of EN 60601-2-51 (2003), namely ANE20000 ECG record, however, the presented methods may be applied not only to averaging of ECG signal but any quasi-repetitive and synchronized signal.

2. METHODOLOGY

2.1. WEIGHTED AVERAGING FUNDAMENTALS

The biomedical signal with repetitive patterns can be (after segmentation and synchronization) represented by:

$$x_i(j) = s(j) + n_i(j), \quad i \in \{1,2,...,N\},\ j \in \{1,2,...,L\}, \tag{1}$$

where N is the number of cycles to be averaged, and L is the length of the single cycle. Thus, each signal cycle $x_i(j)$ is the sum of the deterministic and invariant from cycle to cycle signal $s(j)$ and the random noise $n_i(j)$ with zero mean and variance cycle σ_i^2. The weighted averaged cycle can be expressed as:

$$x(j) = \sum_{i=1}^{N} w_i x_i(j), \quad j \in \{1,2,...,L\}, \tag{2}$$

where w_i is the weight for i-th signal cycle. The choice of the weights defines different types of the signal averaging methods. In the simplest case of arithmetical averaging, all weights are the same, summing up to one, equal to N^{-1}. If the noise power is the same in all cycles, the classical procedure which assumes that the weights are proportional to the inverses of corresponding variances:

$$w_i = \frac{\sigma_i^{-2}}{\sum_{k=1}^{N} \sigma_k^{-2}}, \quad i \in \{1,2,...,N\}, \tag{3}$$

leads to obtaining the arithmetical averaging weights. However, in practice the variability of noise power is observed and measuring the variances directly is impossible. Thus there are employed different methods to estimate the noise variances or to compute the optimal weights without direct estimation of the noise variance.

2.2. SELECTED METHODS OF WEIGHTED AVERAGING

For numerical experiments, presented in the next section, there were selected a few weighted averaging methods described in detail in [15] and the one presented in [16] with extension by partitioning of the input data in the time domain.

- **AA –** the traditional Arithmetic Averaging.
- **WAPM.n –** the Weighted Averaging method based on Partition of input data set in time domain and using criterion function Minimization, which is based on minimization a functional expressing the distance between the two (or more) averaged signals. The parameter of the method is set after dot and describes number of disjoint subsets of input signals.
- **SEBWA –** the Simplified Empirical Bayesian Weighted Averaging method [16].
- **EBWA.1 –** the Empirical Bayesian Weighted Averaging method with hyperparameter λ calculated based on first absolute sample moment (the required parameter *p* is set to 1, it is the value suggested by author of the method).
- **EBWA.C –** the Empirical Bayesian Weighted Averaging method using Cauchy distribution.

2.3. INPUT DATA PARTITIONING

The algorithms of weighted averaging can be extended by partition each signal cycle of the length *L*. The idea of signal partition differs from the one used in WAPM method that earlier the set of *N* cycles was divided into disjoined subsets and now the partition concerns each cycle separately, i.e. the length of averaging window changes.

The partition will be called sharp (traditional) when the input signal is divided into *K* parts (for $k \in \{1,2,...,K\}$):

$$x_i^k(j) = \begin{cases} x_i(j) & j \in \{(k-1)L/K+1,...,kL/K\} \\ 0 & j \in \{1,...,L\} - \{(k-1)L/K+1,...,kL/K\} \end{cases} \tag{4}$$

and fuzzy when the input signal is divided into *K* parts:

$$x_i^k(j) = \frac{x_i(j)\mu_{(a_k,b)}(j)}{\sum_{k=1}^{K}\mu_{(a_k,b)}(j)} \tag{5}$$

for Gaussian membership function with varying location parameter equal $a_k = (k-0.5)L/K$ and constant scale parameter $b = 0.25L/K$. In both cases i is the cycle index $i \in \{1,2,...,N\}$ and j is the sample index in the single cycle $j \in \{1,2,...,L\}$ (all cycles have the same length L). The idea of the partitioning is to perform K times the averaging for $k \in \{1,2,...,K\}$ input data and then sum the resulted signals.

3. NUMERICAL EXPERIMENTS

This section investigates how partition of the input signal affects the results of the averaging procedure. In the experiments both sharp and fuzzy partitions are studied and the number of parts *K* varies from 2 to 5 (for *K* = 1 the original method is used). The number of cycles to be averaged *N* is constant and equal 60. The simulated signal cycles are obtained as the same deterministic component with added independent realizations of random noise. As the deterministic component was taken ECG signal ANE20000, analytical signal compliant with the European Standard EN 60601-2-51 (2003). It is the standardized analytical ECG signal from the CTS database [21], designed to reproduce the typical ECG waveform with 60 bpm (beats per minute) heart rate.

First experiment studies influence of the partition on the root mean square error (RMSE) in the case where the signal is disturbed by zero-mean Gaussian noise with constant amplitude of noise during each cycle. For the first, second, third and fourth 15 cycles, the noise standard deviations were respectively 0.1, 0.5, 1, 2 multiplied by the sample standard deviation of the deterministic component. The RMSE for the traditional arithmetic averaging method is equal 22.098 and detailed results of RMSE for the weighted averaging methods, depending on the *K* parts, are presented in figure 1 (both sharp and fuzzy partitions).

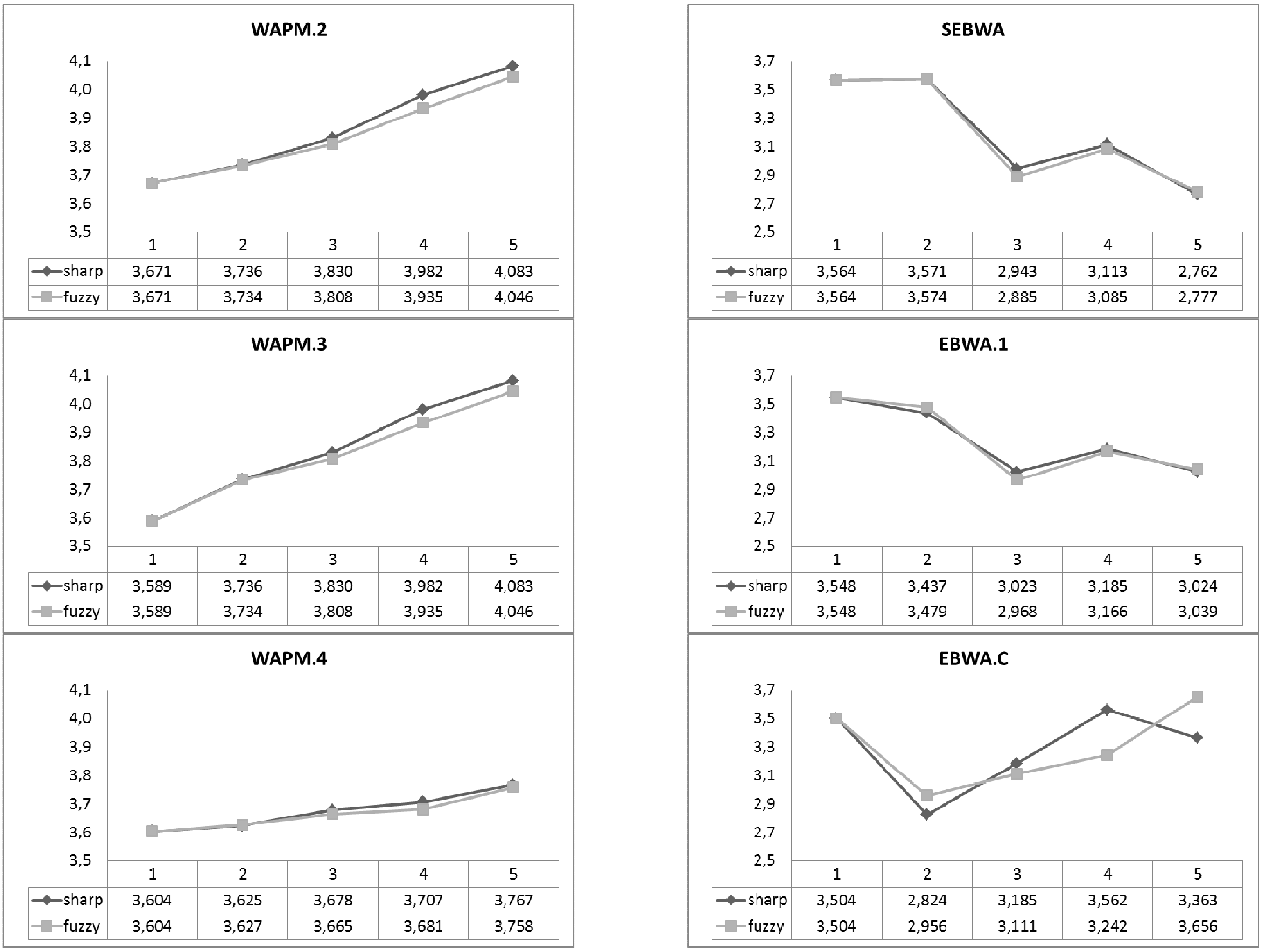


Fig. 1. RMSE for zero-mean Gaussian noise with stepwise changing the amplitude of the noise.

In the case of Bayesian methods the RMSE vary from 2.762 to 3.574 and for the different WAPM methods from 3.589 to 4.083, showing that the first ones outperform the other. The best result is obtained for SEBWA method, as can be seen increasing number of parts decreasing the RMSE more than in the case of original EBWA.1 method which confirms the conclusions presented in [17].

Next experiment studies influence of the partition on the root mean square error in the case where the cycles of the signal are disturbed by zero-mean Gaussian noise with the continuously changing noise amplitude from cycle to cycle, described by function:

$$A(i) = \begin{cases} i/12 & i \in \{1,2,...,24\} \\ 2 & i \in \{25,,26,...,36\} \\ (61-i)/12 & i \in \{37,38,...,60\} \end{cases}. \tag{6}$$

In this experiment the RMSE for the traditional arithmetic averaging method is equal 25.257 and detailed results of RMSE for the weighted averaging methods, depending on the *K* parts, are presented in figure 2 (both sharp and fuzzy partitions).

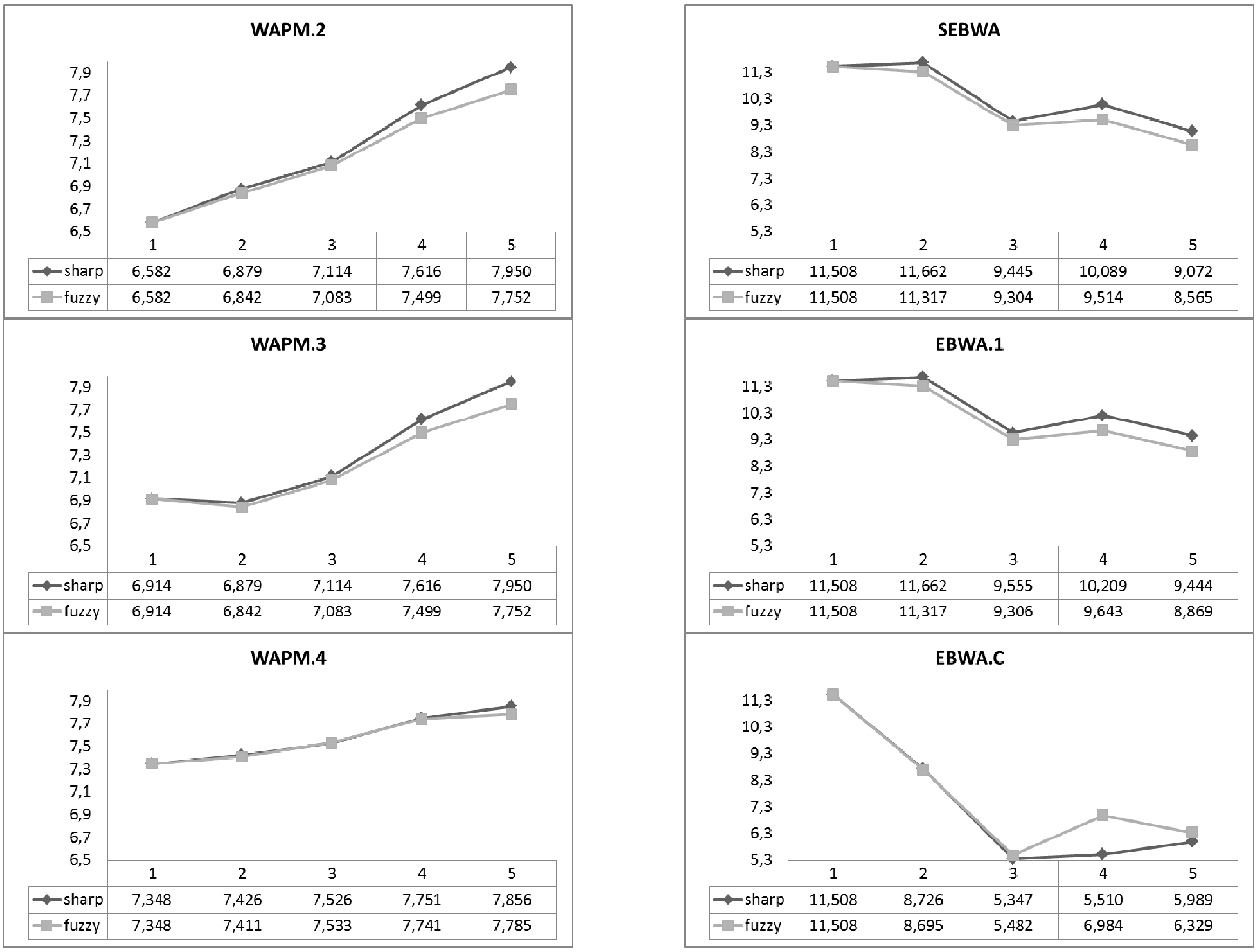
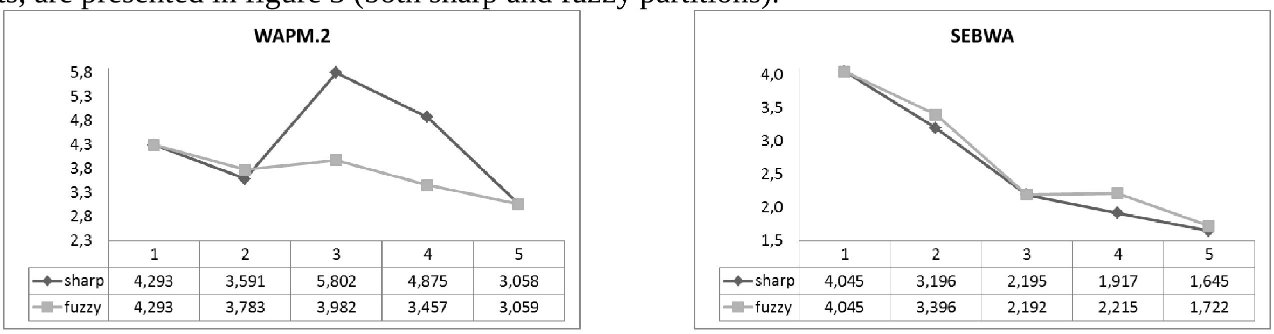


Fig. 2. RMSE for zero-mean Gaussian noise with continuously changing the amplitude of the noise.

Although maximum noise amplitude is the same (2 multiplied by the sample standard deviation of the deterministic component) errors in this experiment are significantly greater than in the previous one. In the case of Bayesian methods the RMSE vary from 5.347 to 11.662 and for the different WAPM methods from 6.582 to 7.950. The best result is obtained for EBWA.C method, and as in the previous experiment it can be seen that in most cases increasing number of parts decreasing the RMSE for Bayesian methods but increasing the RMSE for WAPM method.

Next experiment studies influence of the partition on the root mean square error in the case where the cycles of the signal were disturbed by Cauchy noise. The location parameter of Cauchy distribution is equal to 0 and the scale parameter is set to 0.01 multiplied by the standard deviation of the deterministic component, i.e. the original ANE20000 signal. The RMSE for the traditional arithmetic averaging method is equal 38.255 and detailed results of RMSE for the weighted averaging methods, depending on the *K* parts, are presented in figure 3 (both sharp and fuzzy partitions).



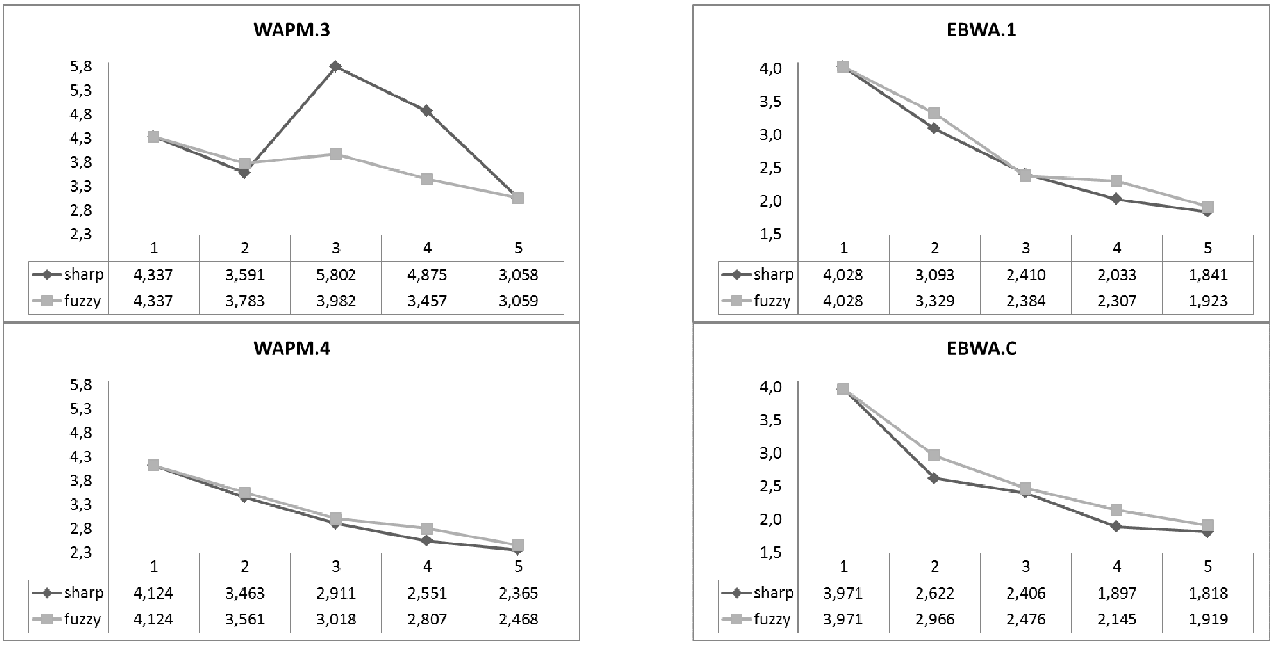


Fig. 3. RMSE for Cauchy noise.

In the case of Bayesian methods the RMSE vary from 1.645 to 4.045 and for the different WAPM methods from 2.365 to 5.802. The best result is obtained for SEBWA method.

The aim of the experiments was to compare the influence of the type of partition and the number of parts on the resulted signal for a few selected weighted averaging methods. Based on the obtained numerical results statistical test ANOVA was also conducted, which found statistically significant differences in the effectiveness of noise suppression by different methods. In the case of Gaussian noise the *p*-value has the order of magnitude 10E-6 and in the case of Cauchy noise the *p*-value is 0.003.

Presented experiments confirm the conclusions described in [17], where only SEBWA and EBWA.1 methods were analyzed, that the SEBWA method with fuzzy partition of the input data seems to be the best choice among the presented methods to use for reducing noise in repetitive signals, especially in the case of presence of impulse noise (in performed experiments it was simulated by Cauchy noise), and application of the input data partitioning leads to improve the quality of the resulting averaged signal, especially when fuzzy partition is used.

4. ACKNOWLEDGMENTS

This work was financed in part by the Polish National Science Centre.

BIBLIOGRAPHY

[1] AUGUSTYNIAK P., Adaptive wavelet discrimination of muscular noise in the ECG, Comput. Cardiol., 2006, Vol. 33, pp. 481-484.
[2] BATAILLOU E., THIERRY E., RIX H., MESTE O., Weighted averaging using adaptive estimation of the weights, Signal Process., 1995, Vol. 44, pp. 51-66.
[3] BRUCE E.N., Biomedical signal processing and signal modeling. Wiley, New York, 2001.
[4] ELBERLING C., DON M., Detecting and assessing synchronous neural activity in the temporal domain, In: BURKARD R.F., EGGERMONT J.J., DON M. (Eds.), Auditory Evoked Potentials - Basic principles and Clinical Application, Lippincott Williams &Wilkins, Philadelphia, 2006, pp. 102-123.
[5] FAN Z., WANG T., A weighted averaging method for evoked potential based on the minimum energy principle, Proc. IEEE EMBS Conf., 1991, Vol. 13, pp. 411-412.
[6] JANE R., RIX H., CAMINAL P., LAGUNA P., Alignment methods for averaging of high-resolution cardiac signals: a comparative study of performance, IEEE Trans. Biomed. Eng., 1991, Vol. 38, No. 6, pp. 571-579.
[7] JEZEWSKI J., HOROBA K., MATONIA A., et al., A new approach to cardiotocographic fetal monitoring based on analysis of bioelectrical signals, Proc. 25th IEEE/EMBS Int. Conf., Cancun, 2003, pp. 3145-3149.

[8] KUPKA T., JEZEWSKI J., MATONIA A., et al. Timing events in Doppler ultrasound signal of fetal heart activity, Proc. 26th IEEE/EMBS Int. Conf., San Francisco, 2004, pp. 337-340.
[9] MATONIA A., JEZEWSKI M., KUPKA T. et al., The influence of coincidence of fetal and maternal QRS complexes on fetal heart rate reliability, Med. Biol. Eng. Comput., 2006, Vol. 44, pp. 393-403.
[10] KOTAS M., Nonlinear projective filtering of ECG signals, In: MELLO C.A.B. (Ed.) Biomedical engineering, InTech, Rijeka, Croatia, 2009, pp. 433–452.
[11] LESKI J., New concept of signal averaging in time domain, Proc. IEEE EMBS Conf., 1991, Vol. 13, pp. 367-368.
[12] LESKI J., Robust Weighted Averaging, IEEE Trans. Biomed. Eng., Vol. 49, No. 8, pp. 796–804, 2002.
[13] MOMOT A., MOMOT M., LESKI J., The Fuzzy Relevance Vector Machine and its Application to Noise Reduction in ECG Signal, J. Med. Inform. Technol., 2005, Vol. 9, pp. 99–106.
[14] MOMOT A., MOMOT M., LESKI J., Bayesian and empirical Bayesian approach to weighted averaging of ECG signal, Bull. Pol. Acad. Sci., Technol. Sci., 2007, Vol. 55, No. 4, pp. 341–350.
[15] MOMOT A., Methods of weighted averaging of ECG signals using Bayesian inference and criterion function minimization, Biomed. Signal Process. Control, 2009, Vol. 4, pp. 162-169.
[16] MOMOT A., Fuzzy Weighted Averaging of Biomedical Signal Using Bayesian Inference, In: CYRAN K.A., et al. (Eds.), Man-Machine Interactions, Advances in Intelligent and Soft Computing, Springer-Verlag, Berlin Heidelberg, 2009, Vol. 59, pp. 133–140.
[17] MOMOT A., On application of input data partitioning to Bayesian weighted averaging of biomedical signals. Expert Systems (article first published online: 28 APR 2011) doi=10.1111_j.1468-0394.2011.00597.
[18] PAUL J.S., REDDY M.R., KUMAR V.J., A transform domain SVD filter for suppression of muscle noise artefacts in exercise ECG's, IEEE Trans. Bimed. Eng., 2000, Vol. 47, No. 5, pp. 654–663.
[19] SHARMA L. N., DANDAPAT S., MAHANTA A., ECG signal denoising using higher order statistics in Wavelet subbands, Biomed. Signal Process. Control, 2010, Vol. 5, No. 3, pp. 214–222.
[20] YAN J., LU Y., LIU, J., WUB X., XU Y., Self-adaptive model-based ECG denoising using features extracted by mean shift algorithm, Biomed. Signal Process. Control, 2010, Vol. 5, No.2, pp. 103–113.
[21] ZYWIETZ C., ALRAUN W., FISHER R., Quality assurance in biosignal processing - procedures and recommendations for evaluation for electrocardiological analysis systems, Comput. Cardiol., 2001, Vol. 28, pp. 201–204.