

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Methods of Weighted Averaging with Application to Biomedical Signals

Alina Momot

*Silesian University of Technology, Institute of Computer Science
Poland*

1. Introduction

During the analysis of real biomedical signals it can almost always be seen noise that distorts the image. The presence of interference is associated with the specific acquisition of these signals. For example in the case of bioelectric signals, disturbances may come from the hardware retrieves those signals, the powerline or the bioelectric activity of body cells. The bioelectric signals, which are widely used in most fields of biomedicine, are generated by nerve cells or muscle 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 (Bruce, 2001).

There are many approaches to the noise reduction problem while preserving the variability of the desired signal morphology. One of the possible methods of noise attenuation is low-pass filtering such as arithmetic mean. The classical band-pass filtering is very simple method but also very ineffective because the frequency characteristics of signal and noise significantly overlap. Therefore there are developed other methods of noise attenuation based on transforming the input space of signal and creating a new space with the help of discrete cosine transform (Paul et al., 2000) or wavelets transform (Augustyniak, 2006), based on fuzzy nonlinear regression (Momot et al., 2005), nonlinear projective filtering (Kotas, 2009), higher-order statistics at different wavelet bands (Sharma et al., 2010) or extreme points determination by mean shift algorithm and dynamical model-based nonlinear filtering (Yan et al., 2010).

In the case of repeatable biomedical signals, another possible method of noise attenuation is the synchronized averaging (Jane et al., 1991). The method assumes that the biomedical signal is quasi-cyclic and the noise is additive, independent and with zero mean. Averaging could be performed by simple arithmetic mean or its generalization, namely weighted mean where the weights are tuned by some adaptive algorithm.

Recently there have been published several works concerning different approaches to the problem of determining the weights. The algorithm of adaptive estimation of the weights is described in (Bataillou et al., 1995). In (Leski, 2002) there is described method of estimation of the weights based on criterion function minimization. Application of Bayesian inference to the weights estimation problem is presented in (Momot et al., 2007a) and (Momot, 2008b). Weighted averaging method based on partition of input data set in time domain is described in (Momot et al., 2007b). The generalization of the method is presented in (Momot,

2008a). In (Momot, 2009a), there is presented comparative study of performance of weighted averaging methods using Bayesian inference and criterion function minimization. There were also published works describing algorithms based on fuzzy approaches to the problem of determining the weights. In (Momot, 2009b), there is presented the new weighted averaging method incorporating Bayesian and empirical Bayesian inference and its extension using fuzzy systems with fuzzy partitioning of the single repeatable input signal in the time domain. In (Momot & Momot, 2009c), there is described the weighted averaging method, where the weighting coefficients are fuzzy numbers instead of classical real numbers. Moreover in (Momot, 2010), there is presented an application of Bayesian weighted averaging to digital filtering 2D images.

Most of the mentioned above methods of weighted averaging were presented with application to noise suppression in ECG signal, the typical biomedical signal with a quasi-cyclical character. In the case of ECG signal the averaging in the time domain is only one of many methods of noise attenuation, but in some medical applications the averaging is the only method taken into account. For example in order to evaluate some clinical indexes based on the ST depression, such as ST versus HR (ST/HR) diagram or ST/HR hysteresis, the averaged ECG beats are necessary to compute (Bailon et al., 2002). Another example of biomedical repeatable signal in which noise attenuation can be performed using averaging methods is the high resolution electrocardiographic signal (HRECG). Methods of averaging help in detecting very low amplitude waves (which are called late potentials) originating in the ventricles of abnormal heart conditions patients (Laciar & Jane, 2001). Signal averaging methods are also required in the case of evoked potentials (EP), the electrical potentials recorded from the nervous system of patients, which are effects of a stimulation (in contrast to the spontaneous potentials as in the case of electroencephalography (EEG) or electromyography (EMG)) (Davila, 1992).

Traditional arithmetic averaging technique can be used when the noise is stationary (it is constant level of power noise) throughout the averaging period. If this condition is violated, increasing the input noise level during the averaging process results in increasing the residual noise and deteriorating quality of the averaged signal. Unfortunately, the physiological noise level often varies in tests, even if strict artifact rejection is applied (Elberling & Don, 2006). 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 aim of this study is to reveal the fundamental differences among the mentioned versions of the weighted averaging methods and to present how these differences affect the quality of the averaged signal. In section 2, the necessary details of the compared methods are reminded. The differences among them and their applications are studied based on numerical experiments in section 3. The final conclusions are formulated in section 4.

2. Fundamental of weighted averaging

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

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

where N is the number of cycles to be averaged, and L is the length of the single cycle. Typical assumption states that each signal cycle $x_i(j)$ is the sum of the signal $s(j)$, which is deterministic and invariant from cycle to cycle, and the random noise $n_i(j)$ with zero mean and variance cycle σ_i^2 (the noise remains stationary within each evolution, but its variance

may vary from one cycle to the other). The assumptions and symbols are based on (Laciar & Jane, 2001).

The weighted averaged cycle can be expressed as:

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

where w_i is the weight for i th signal cycle. Usually, there is taken the assumption that the weights sum up to one ($\sum_{i=1}^N w_i = 1$), which leads to the unbiased estimation. Depending on the choice of the weights, different types of the signal averaging methods can be defined. The simplest method is arithmetical averaging, where all weights are the same, equal to M^{-1} . The classical procedure assumes that the weights are proportional to the inverses of corresponding variances Johnson & Bhattacharyya (2009):

$$w_i = \sigma_i^{-2} \left(\sum_{k=1}^N \sigma_k^{-2} \right)^{-1} \quad i \in \{1, 2, \dots, N\}, \quad (3)$$

which leads to obtaining the arithmetical averaging weights if the noise power is the same in all cycles. However, in practice the variability of noise power is observed and it is impossible to measure the variances directly. Thus there are employed different methods to estimate the noise variances or to compute the optimal weights without direct estimation of the noise variance. These two approaches will be presented below. The Bayesian methods incorporate estimation of the noise variations and methods using criterion function minimization usually lead to direct computation of the weights.

2.1 Methods based on criterion function minimization

The noise variance, which appears in formula (3), can be estimated according to the formula (Laciar & Jane, 2001):

$$\hat{\sigma}_i^2 = \frac{1}{L} \sum_{j=1}^L (\hat{n}_i(j) - \bar{n}_i)^2 \quad \text{with} \quad \bar{n}_i = \frac{1}{N} \sum_{j=1}^L \hat{n}_i(j), \quad (4)$$

where L is a length of the averaging window and \bar{n}_i is the mean value of the estimated noise component of i th cycle in the averaging window.

Assuming that the signal is deterministic and invariant from cycle to cycle and the noise has zero mean, the estimated noise component can be described by:

$$\hat{n}_i(j) = x_i(j) - \bar{x}(j) \quad j \in \{1, 2, \dots, L\} \quad (5)$$

and the formula (4) may be written in simplified form:

$$\hat{\sigma}_i^2 = \frac{1}{L} \sum_{j=1}^L (x_i(j) - \bar{x}(j))^2, \quad (6)$$

where $\bar{x}(j)$ is the averaged cycle in the analysis window.

It is worth noting that formula (6) contains $\bar{x}(j)$ defined by (2) and (3), thus the use an iterative determination of these values could improve the estimation. Generalization of this approach may lead to presented below Weighted Averaging method based on

Criterion Function Minimization (WACFM) (Leski, 2002). Another method described below is Weighted Averaging method based on Partition of input data set in time domain and using criterion function Minimization (WAPM) (Momot et al, 2007b) and generalization of the method (Momot, 2008a).

2.1.1 Weighted averaging method based on criterion function minimization

The main idea of the Weighted Averaging method based on Criterion Function Minimization (WACFM) (Leski, 2002) is minimization the following scalar criterion function:

$$I_m(\mathbf{w}, \bar{\mathbf{x}}) = \sum_{i=1}^N w_i^m \rho(\mathbf{x}_i - \bar{\mathbf{x}}), \quad (7)$$

where $m \in (1, \infty)$ is a weighting exponent parameter and $\rho(\cdot)$ is a measure of dissimilarity for vector arguments, i.e. $\mathbf{x}_i = [x_i(1), x_i(2), \dots, x_i(L)]^T$, and $\bar{\mathbf{x}} = [\bar{x}(1), \bar{x}(2), \dots, \bar{x}(L)]^T$. The measure of dissimilarity could be for example the quadratic function $\rho(\mathbf{t}) = \|\mathbf{t}\|^2 = \mathbf{t}^T \mathbf{t}$ and then the formula (7) can be expressed as:

$$I_m(\mathbf{w}, \bar{\mathbf{x}}) = \sum_{i=1}^N \left(w_i^m \sum_{j=1}^L (x_i(j) - \bar{x}(j))^2 \right). \quad (8)$$

Minimization the criterion function with respect to the weights vector $\mathbf{w} = [w_1, w_2, \dots, w_N]^T$ yields:

$$w_i = \frac{\rho(\mathbf{x}_i - \bar{\mathbf{x}})^{1/(1-m)}}{\sum_{k=1}^N \rho(\mathbf{x}_k - \bar{\mathbf{x}})^{1/(1-m)}} \quad i \in \{1, 2, \dots, N\}, \quad (9)$$

and for the quadratic function ρ it can be expressed as:

$$w_i = \frac{\left(\sum_{j=1}^L (x_i(j) - \bar{x}(j))^2 \right)^{1/(1-m)}}{\sum_{k=1}^N \left(\sum_{j=1}^L (x_k(j) - \bar{x}(j))^2 \right)^{1/(1-m)}}. \quad (10)$$

It is worth noting that for the parameter $m = 2$ this formula is equivalent to the formula (3) with variance estimated by the formula (6). However, the obtained for the quadratic function ρ averaged signal is given by:

$$\bar{\mathbf{x}} = \frac{\sum_{i=1}^N (w_i)^m \mathbf{x}_i}{\sum_{i=1}^N (w_i)^m}, \quad (11)$$

which is not exactly equivalent to the formula (2).

The optimal solution for minimization (7) with respect to \mathbf{w} and $\bar{\mathbf{x}}$ is a fixed point of (10) and (11) and it could be obtained from the Picard iteration. Therefore the algorithm can be described as follows, where ε is a preset parameter.

1. Fix $m \in (0, \infty)$. Initialize $\bar{x}^{(0)}$ as the arithmetically averaged signal. Set the iteration index $k = 1$.
2. Calculate $w^{(k)}$ for the k th iteration using the formula (10).
3. Update the averaged signal for the k th iteration $\bar{x}^{(k)}$ using the formula (11) and $w^{(k)}$.
4. If $\|w^{(k-1)} - w^{(k)}\| > \varepsilon$ then $i \leftarrow i + 1$ and go to 2.

It is suggested to set parameter $m = 2$, because if parameter m tends to one, then the trivial solution is obtained where only one weight is equal to one and for large m the weights are similar to each other, like in arithmetic averaging (Leski, 2002).

2.1.2 Weighted averaging method based on partition of input data set

Below it is described the Weighted Averaging method based on Partition of input data set in time domain and using criterion function Minimization (WAPM) (Momot et al, 2007b) and generalization of the method (Momot, 2008a). The main idea of the WAPM is minimization the following scalar criterion function:

$$I(w_1, w_2) = \|x^1 w_1 - x^2 w_2\|^2 = (x^1 w_1 - x^2 w_2)^T (x^1 w_1 - x^2 w_2), \quad (12)$$

where the input set $x = [x_1, x_2, \dots, x_N]$ ($x_i = [x_i(1), x_i(2), \dots, x_i(L)]^T$) is divided into two disjoint subsets x^1 and x^2 and w_1 and w_2 are the weights vectors, respectively.

Taking into account the constraints $w_1^T \mathbf{1} = 1$ and $w_2^T \mathbf{1} = 1$, which mean that sum of weights for each vector is equal to one, minimization (12) with respect to the weights vectors yields:

$$w_1 = \left((x^1)^T x^1 \right)^{-1} (x^1)^T x^2 w_2 + \frac{1 - \mathbf{1}^T \left((x^1)^T x^1 \right)^{-1} (x^1)^T x^2 w_2}{\mathbf{1}^T \left((x^1)^T x^1 \right)^{-1} \mathbf{1}} \left((x^1)^T x^1 \right)^{-1} \mathbf{1} \quad (13)$$

and

$$w_2 = \left((x^2)^T x^2 \right)^{-1} (x^2)^T x^1 w_1 + \frac{1 - \mathbf{1}^T \left((x^2)^T x^2 \right)^{-1} (x^2)^T x^1 w_1}{\mathbf{1}^T \left((x^2)^T x^2 \right)^{-1} \mathbf{1}} \left((x^2)^T x^2 \right)^{-1} \mathbf{1}. \quad (14)$$

The optimal solution for minimization (12) with respect to w_1 and w_2 is a fixed point of (13) and (14) and it could be obtained from the Picard iteration, which leads to the averaged signal given by:

$$\bar{x} = \frac{N_1 x^1 w_1 + N_2 x^2 w_2}{N}, \quad (15)$$

where N_1 and N_2 are the cardinalities of the two disjoint subsets x^1 and x^2 , i.e. $N_1 + N_2 = N$. Although described above method involves partitioning of input set into two disjoint subsets, it can be generalized by increasing the number of disjoint subsets (Momot, 2008a). The generalized WAPM algorithm can be described as follows, where ε is a preset parameter.

1. Determine partition of input set x into disjoint subsets x^c with cardinalities N_c , where $c \in \{1, 2, \dots, C\}$, $N_1 + N_2 + \dots + N_C = N$ and $C \geq 2$. Calculate the following values, which remain constant during the whole iteration procedure:

$$\begin{aligned}
X_{1,1}^{-1} &= ((\mathbf{x}^1)^T \mathbf{x}^1)^{-1}, \\
X_{2,2}^{-1} &= ((\mathbf{x}^2)^T \mathbf{x}^2)^{-1}, \\
&\dots \\
X_{C,C}^{-1} &= ((\mathbf{x}^C)^T \mathbf{x}^C)^{-1}, \\
X_{1,C} &= (\mathbf{x}^1)^T \mathbf{x}^C, \\
X_{2,1} &= (\mathbf{x}^2)^T \mathbf{x}^1, \\
&\dots \\
X_{C,C-1} &= (\mathbf{x}^C)^T \mathbf{x}^{C-1}.
\end{aligned} \tag{16}$$

Initialize weights $\mathbf{w}_C^{(0)}$ as in the case of arithmetical averaging (all the same and equal to N_C^{-1}). Set the iteration index $k = 1$.

2. Calculate $\mathbf{w}_1^{(k)}$ for the k th iteration using

$$\mathbf{w}_1^{(k)} = X_{1,1}^{-1} X_{1,C} \mathbf{w}_C^{(k-1)} + \frac{1 - \mathbf{1}^T X_{1,1}^{-1} X_{1,C} \mathbf{w}_C^{(k-1)}}{\mathbf{1}^T X_{1,1}^{-1} \mathbf{1}} X_{1,1}^{-1} \mathbf{1}. \tag{17}$$

3. Calculate $\mathbf{w}_c^{(k)}$ for the k th iteration using

$$\mathbf{w}_c^{(k)} = X_{c,c}^{-1} X_{c,c-1} \mathbf{w}_{c-1}^{(k)} + \frac{1 - \mathbf{1}^T X_{c,c}^{-1} X_{c,c-1} \mathbf{w}_{c-1}^{(k)}}{\mathbf{1}^T X_{c,c}^{-1} \mathbf{1}} X_{c,c}^{-1} \mathbf{1}, \tag{18}$$

for $c \in \{2, \dots, C\}$.

4. If $\sum_{c=1}^C \|\mathbf{w}_c^{(k-1)} - \mathbf{w}_c^{(k)}\| > \varepsilon$ then $k \leftarrow k + 1$ and go to 2.
5. Calculate averaged signal

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{c=1}^C N_c \mathbf{x}^c \mathbf{w}_c. \tag{19}$$

It is suggested to use this method with equal in number of elements subsets and interlaced partitioning, i.e. to divide the input set into subsets with cardinalities equal N/C , where each of the subset indexes was equal to one plus remainder in division cycle index by C ($\mathbf{x}^c = \{\mathbf{x}_c, \mathbf{x}_{c+C}, \mathbf{x}_{c+2C}, \dots, \mathbf{x}_{c+N-C}\}$ for $c = 1, 2, \dots, C$), to obtain the best performance (Momot, 2008a).

2.2 Methods based on statistical inference

Below there are presented weighted averaging methods, which incorporate Bayesian inference and the expectation-maximization technique: the Empirical Bayesian Weighted Averaging algorithm (EBWA) (Momot et al., 2007a) and the Empirical Bayesian Weighted Averaging using Cauchy distribution algorithm (EBWA.C) (Momot, 2008b). There is also presented the Simplified Empirical Bayesian Weighted Averaging algorithm (SEBWA) using method of moments to estimate the unknown parameters of signal and noise distributions (Momot, 2009b).

All the Bayesian methods are based on the assumption that the random noise $n_i(j)$, which appears in signal cycle $x_i(j)$ (see formula (1)), is zero-mean Gaussian with variance for

the i th cycle σ_i^2 and the second component of the sum, i.e. the useful signal $\mathbf{s} = [s(1), s(2), \dots, s(L)]$, has also Gaussian distribution with zero mean and covariance matrix $B = \text{diag}(\eta_1^2, \eta_2^2, \dots, \eta_N^2)$. The zero-mean assumption for the useful signal expresses no prior knowledge about the real distance from the signal to the baseline.

From the Bayes rule it could be calculated the posterior distribution over the useful signal and the noise variance, which is proportional to

$$p(\mathbf{s}, \boldsymbol{\alpha} | \mathbf{x}, \boldsymbol{\beta}) \propto \prod_{i=1}^N \alpha_i^{\frac{L}{2}} \exp \left(-\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^L (x_i(j) - s(j))^2 \alpha_i \right) \prod_{j=1}^L \beta_j^{\frac{1}{2}} \exp \left(-\frac{1}{2} \sum_{j=1}^L (s(j))^2 \beta_j \right), \quad (20)$$

where $\alpha_i = \sigma_i^{-2}$ and $\beta_j = \eta_j^{-2}$.

The main idea of the Bayesian method is to maximize this posterior distribution. The values \mathbf{s} and $\boldsymbol{\alpha}$, which maximize it, can be calculated by setting the derivative of the logarithm of the posterior distribution to zero with respect to α_i and with respect to $s(j)$ respectively. The values can be expressed as:

$$\alpha_i = \frac{L}{\sum_{j=1}^L (x_i(j) - s(j))^2}, \quad i \in \{1, 2, \dots, N\}, \quad (21)$$

and

$$s(j) = \frac{\sum_{i=1}^N \alpha_i x_i(j)}{\beta_j + \sum_{i=1}^N \alpha_i}, \quad j \in \{1, 2, \dots, L\}. \quad (22)$$

Unfortunately it is impossible to measure β_j directly and the following methods estimate these values in different ways.

2.2.1 Empirical Bayesian weighted averaging algorithm

The Empirical Bayesian Weighted Averaging algorithm (EBWA) assumes the gamma prior for β_j with scale parameter λ and shape parameter p for all $j \in \{1, 2, \dots, L\}$ and exploits the iterative expectation-maximization technique (Momot et al., 2007a). Conditional expected value of β_j is given by:

$$E(\beta_j | s(j)) = \frac{2p + 1}{(s(j))^2 + 2\lambda}, \quad j \in \{1, 2, \dots, L\}. \quad (23)$$

Assuming that p is a positive integer, the estimate $\hat{\lambda}$ of hyperparameter λ can be calculated based on first absolute sample moment:

$$\hat{\lambda} = \left(\frac{\Gamma(p)(2p-1)}{(2p-1)!!} 2^{p-\frac{3}{2}} \frac{\sum_{j=1}^L |s(j)|}{L} \right)^2, \quad (24)$$

where $(2p-1)!! = 1 \cdot 3 \cdot \dots \cdot (2p-1)$, or based on third absolute sample moment:

$$\hat{\lambda} = \left(\frac{\Gamma(p)(2p-3)}{(2p-3)!!} 2^{p-\frac{7}{2}} \frac{\sum_{j=1}^N |x(j)|^3}{N} \right)^{\frac{2}{3}}, \quad (25)$$

however in this case assumption that p is greater than 1 is required.

Summarizing, the Empirical Bayesian Weighted Averaging (EBWA) algorithm can be described as follows, where ε and p are preset parameters.

1. Initialize $\mathbf{s}^{(0)} \in R^L$ as in the case of arithmetical averaging (all the same and equal to N^{-1}) and set iteration index $k = 1$.
2. Calculate the hyperparameter $\lambda^{(k)}$ using (24) in the case of EBWA.1 (or using (25) in the case of EBWA.3, but only for $p > 1$), next $\beta_j^{(k)}$ using (23) for $j \in \{1, 2, \dots, L\}$ and $\alpha_i^{(k)}$ using (21) for $i \in \{1, 2, \dots, N\}$.
3. Update the signal $\mathbf{s}^{(k)}$ using (22), $\beta_j^{(k)}$ and $\alpha_i^{(k)}$.
4. If $\|\mathbf{s}^{(k)} - \mathbf{s}^{(k-1)}\| > \varepsilon$ then $k \leftarrow k + 1$ and go to 2, else set the averaged signal $\bar{\mathbf{x}} = \mathbf{s}^{(k)}$ and stop.

It is suggested to use this method with parameter $p = 1$ (hence EBWA.1), because performed numerical experiments indicate that increasing values of p usually did not improve performance of the method (Momot, 2009a).

2.2.2 Empirical Bayesian weighted averaging using Cauchy distribution algorithm

The presented above EBWA method requires assumption that certain parameter p is a positive integer. The observation that increasing values of p usually did not improve performance of the method has become the motivation to extension of the algorithm for some values of $p < 1$. It can be observed that for $p = \frac{1}{2}$, function $p(s(j)|\lambda)$ is Cauchy probability distribution function:

$$p(s(j)|\lambda) = \frac{\sqrt{2\lambda}}{\pi (s(j)^2 + 2\lambda)} \quad (26)$$

with the scale parameter equal to $\sqrt{2\lambda}$ and the location parameter equal to 0. All absolute moments of the Cauchy distribution are infinite, but the first and third quartiles are the linear functions of scale parameter:

$$Q1 = -\sqrt{2\lambda}, \quad Q3 = \sqrt{2\lambda}. \quad (27)$$

Thus the hyperparameter λ can be estimated based on sample interquartile range:

$$\hat{\lambda} = \frac{(\hat{Q}3 - \hat{Q}1)^2}{8}. \quad (28)$$

Therefore the Empirical Bayesian Weighted Averaging using Cauchy distribution algorithm (EBWA.C) can be described as follows, where ε is a preset parameter (Momot, 2008b).

1. Initialize $\mathbf{s}^{(0)} \in R^L$ as in the case of arithmetical averaging (all the same and equal to N^{-1}) and set iteration index $k = 1$.
2. Calculate the hyperparameter $\lambda^{(k)}$ using (28), next $\beta_j^{(k)}$ using (23) for $j \in \{1, 2, \dots, L\}$ and $\alpha_i^{(k)}$ using (21) for $i \in \{1, 2, \dots, N\}$.
3. Update the signal $\mathbf{s}^{(k)}$ using (22), $\beta_j^{(k)}$ and $\alpha_i^{(k)}$.
4. If $\|\mathbf{s}^{(k)} - \mathbf{s}^{(k-1)}\| > \varepsilon$ then $k \leftarrow k + 1$ and go to 2, else set the averaged signal $\bar{\mathbf{x}} = \mathbf{s}^{(k)}$ and stop.

2.2.3 Simplified empirical Bayesian weighted averaging algorithm

Below there is presented the Simplified Empirical Bayesian Weighted Averaging algorithm (SEBWA) which does not use hierarchical probabilistic model and does not require the determination of parameter p . In this method the unknown parameters of signal and noise distributions are estimated using method of moments (Momot, 2009b).

Giving assumption as described previously, but $\beta = \eta_1^{-2} = \eta_2^{-2} = \dots = \eta_N^{-2}$, the posterior distribution of signal (see formula 20) can be calculated from the Bayes rule explicitly as Gaussian distribution with mean vector \mathbf{m} :

$$\forall j \in \{1, 2, \dots, L\} \quad m(j) = \frac{\sum_{i=1}^N \alpha_i x_i(j)}{\beta + \sum_{i=1}^N \alpha_i} \quad (29)$$

and covariance matrix equal to γ^{-1} multiplied by the identity matrix of dimension L , where

$$\gamma = \beta + \sum_{i=1}^N \alpha_i. \quad (30)$$

Therefore the original signal \mathbf{s} can be estimated as $\hat{\mathbf{m}}$ using (29) and the unknown parameters α_i for $i \in \{1, 2, \dots, M\}$ and β can be estimated using method of moments (the estimated parameters of random distribution are expressed in terms of its moments which are substituted by the corresponding sample moments). Values α_i^{-1} are noise variations in each cycle and taking into account the mean equals zero

$$\hat{\alpha}_i = \frac{L}{\sum_{j=1}^L (x_i(j) - s(j))^2}, \quad i \in \{1, 2, \dots, N\}. \quad (31)$$

Value β^{-1} is variation of the original signal and taking into account the zero-mean assumption

$$\hat{\beta} = \frac{L}{\sum_{j=1}^L (s(j))^2}. \quad (32)$$

Therefore the Simplified Empirical Bayesian Weighted Averaging algorithm (SEBWA) can be described as follows, where ε is a preset parameter.

1. Initialize $\mathbf{s}^{(0)} \in R^L$ as in the case of arithmetical averaging (all the same and equal to N^{-1}) and set iteration index $k = 1$.
2. Calculate $\beta^{(k)}$ using (32) and $\alpha_i^{(k)}$ using (31) for $i \in \{1, 2, \dots, M\}$.
3. Update the signal $\mathbf{s}^{(k)}$ using (29), $\beta^{(k)}$ and $\alpha_i^{(k)}$, assuming $\mathbf{s}^{(k)} = \mathbf{m}$.
4. If $\|\mathbf{s}^{(k)} - \mathbf{s}^{(k-1)}\| > \varepsilon$ then $k \leftarrow k + 1$ and go to 2, else set the averaged signal $\bar{\mathbf{x}} = \mathbf{s}^{(k)}$ and stop.

It is suggested to use this method with fuzzy partition of signal cycle (this extension is presented in the next subsection), because the numerical experiments indicate that the simplified Bayesian method gives worse results with compare to the EBWA method, however using this method with fuzzy partition gives much better results with compare to the EBWA method (even EBWA with fuzzy partition) (Momot, 2009b).

2.3 Fuzzy extensions to weighted averaging methods

This subsection presents two aspects of possible fuzzy extensions applied to weighted averaging methods: fuzzy partition of signal cycle proposed in (Momot, 2009b) and using the fuzzy numbers as coefficients of weight vector instead of classical real numbers (applied to described above WACFM method in (Momot & Momot, 2009c)).

2.3.1 Fuzzy partition of signal cycle

The algorithms of weighted averaging can be extended by partition each signal cycle of the length L . The new idea of signal partition differs from previously presented in subsection 2.1.2 that earlier the set of N cycles was divided into disjointed subsets with cardinalities N_c and now the partition concerns each cycle separately, i.e. the length of averaging window changes.

The partition may be performed by using traditional (sharp) or fuzzy membership function. When the input signal is divided into K parts:

$$x_i^k(j) = \begin{cases} x_i(j), & j \in \{(k-1)L/K + 1, \dots, kL/K\} \\ 0, & j \in \{1, \dots, L\} - \{(k-1)L/K + 1, \dots, kL/K\} \end{cases} \quad (33)$$

for $k \in \{1, 2, \dots, K\}$, this partition will be called sharp. Taking into account Gaussian membership function with location parameter equal $a^k = (k - 0.5)L/K$ (for $k \in \{1, 2, \dots, K\}$) and scale parameter $b = 0.25L/K$, defined by:

$$\mu_{(a^k, b)}(j) = \exp \left\{ - \left(\frac{j - a^k}{b} \right)^2 \right\}, \quad k \in \{1, 2, \dots, K\}, \quad (34)$$

it is possible to divide the input signal into K fuzzy parts:

$$x_i^k(j) = \frac{x_i(j) \cdot \mu_{(a^k, b)}(j)}{\sum_{k=1}^K \mu_{(a^k, b)}(j)}, \quad k \in \{1, 2, \dots, K\}. \quad (35)$$

In both cases i is the cycle index $i \in \{1, 2, \dots, N\}$ and j is the sample index in the single cycle $j \in \{1, 2, \dots, L\}$ (all cycles have the same length L). The idea of this extension is to perform K times the averaging for $k \in \{1, 2, \dots, K\}$ input data and then sum the weighted averages.

Although the partition may be performed by using traditional (sharp) or fuzzy membership function, the numerical experiments presented in (Momot, 2011) indicate that in the case of sharp partition incorrectly chosen number of parts may result in even worse results than the one obtained for the arithmetic averaging. Therefore it is suggested to use the fuzzy partition rather than the sharp partition (especially for the signals with unknown characteristics).

2.3.2 Using fuzzy numbers as coefficients of weight vector

Another aspect of possible fuzzy extensions applied to weighted averaging methods is using the fuzzy numbers as coefficients of weight vector instead of classical real numbers. This idea was presented for described above WACFM method in (Momot & Momot, 2009c), where the coefficients were replaced with symmetrical triangular fuzzy numbers. Consequently, the weighted average was a vector containing triangular fuzzy numbers and as the necessary to compute distance, between the input signal (vector of real numbers) and the averaged signal, was taken the distance between the real number and the α -cut of the corresponding fuzzy number.

The fuzzy membership function of a symmetrical triangular fuzzy number A can be expressed as:

$$\mu_A(x) = \max \left\{ 1 - \frac{|x - m_A|}{r_A}, 0 \right\}, \quad (36)$$

where m_A is the center point of the fuzzy number and r_A is its radius. Thus, the α -cut of the fuzzy number A ($\alpha \in [0, 1]$), defined as the ordinary subset $\{x \in \mathcal{R} : \mu_A(x) \geq \alpha\}$, is given by:

$$(A)_\alpha = [m_A - r_A(1 - \alpha), m_A + r_A(1 - \alpha)]. \quad (37)$$

The distance between a real number x and the α -cut of the symmetrical triangular fuzzy number A can be written explicitly as:

$$\rho_\alpha(x, A) = \max \{|x - m_A| - r_A, 0\} \quad (38)$$

and when the arguments of distance function $\rho_\alpha(\cdot, \cdot)$ are N -dimensional vectors, the formula can be expressed as:

$$\rho_\alpha(\mathbf{x}, \mathbf{A}) = \sum_{i=1}^N (\rho_\alpha(x_i, A_i))^2. \quad (39)$$

Therefore, the Fuzzy Weighted Averaging algorithm based on Criterion Function Minimization (FWACFM) can be described as follows, where ε is a preset parameter (Momot & Momot, 2009c). It is assumed that values of parameters r (the radius of all symmetrical triangular fuzzy numbers) and α (the cutting level) remain constant during all iterations.

1. Determine parameters r and α . Initialize centers of fuzzy weights $\mathbf{w}^{(0)}$ setting all the same values equal to N^{-1} . Set the iteration index $k = 1$.
2. Calculate vector of centers of fuzzy weights $\mathbf{w}^{(k)}$ as:

$$w_i^{(k)} = \left(\rho_\alpha(\mathbf{x}_i, \bar{\mathbf{x}})^{\frac{1}{1-m}} \right) / \left(\sum_{k=1}^N \rho_\alpha(\mathbf{x}_k, \bar{\mathbf{x}})^{\frac{1}{1-m}} \right), \quad (40)$$

for $i \in \{1, 2, \dots, N\}$.

3. Calculate the averaged signal as:

$$\bar{x}^{(k)} = \left(\sum_{i=1}^N (W_i^{(k)})^m x_i \right) / \left(\sum_{i=1}^N (W_i^{(k)})^m \right), \quad (41)$$

where $W_i^{(k)}$ is a symmetrical triangular fuzzy number with center given by (40) and radius r .

4. If $\|w^{(k-1)} - w^{(k)}\| > \varepsilon$ then $k \leftarrow k + 1$ and go to 2.

5. Calculate the final averaged signal as:

$$\bar{x} = \left(\sum_{i=1}^N (w_i^{(k)})^m x_i \right) / \left(\sum_{i=1}^N (w_i^{(k)})^m \right). \quad (42)$$

It is worth noting that during the iteration process both vectors: the averaged signal and the weights are treated as vectors of fuzzy numbers. At the end of the procedure the fuzzy averaged signal is defuzzified.

This algorithm is generalization of the original WACFM method because for radius equal zero both methods are equivalent. The numeric experiments presented in (Momot & Momot, 2009c) indicate that for some positive values of radius parameter such generalization of WACFM method can outperforms the original method. However, further research for method of automatic determining this parameter is needed because even small increasing its value could rapidly increase the root mean square error.

Summarizing, taking into account higher computational complexity the Fuzzy WACFM algorithm than the original algorithm and difficulties in proper choosing the parameters r and α , usefulness of this method is rather limited. Therefore, the Fuzzy WACFM algorithm can be treated as an interesting theoretical study of the problem of using the fuzzy numbers as coefficients of weight vector instead of classical real numbers.

2.4 Adaptation of weighted averaging methods to 2D images

In the case of linear spatial filtering the response of the filter is given by a sum of products of the filter coefficients and the corresponding image pixels in the area spanned by the filter mask. For arithmetic mean filtering all the coefficients are the same and sum up to one. Mean filtering is often used to reduce noise in images due to its simplicity. It is an efficient method for reducing the amount of intensity variation between one pixel and the next. Mean filtering minimizes the influence of pixel values which are unrepresentative of their surroundings.

Like the mean filter, the median filter considers each pixel in the image, looking at its nearby neighbors, to decide whether or not it is representative of its surroundings. It replaces the pixel value with the median of neighboring pixel values instead of replacing it with the mean of those values. Mean filtering is special case of weighted averaging filtering where the filter mask coefficients are nonnegative and sum up to one (often the pixel at the center of the mask is multiplied by a higher value than any other, thus giving this pixel more importance in the calculation of the average). In this context, median filtering can be treated as an adaptive weighted averaging filtering. For median filtering the mask coefficients are not always constant, there is only one non-zero coefficient (equal one) and which coefficient is non-zero depends on result of the sorting operation.

When the mask coefficients are not constant there is a need of procedure how to compute the coefficients. Below there is presented a method for computing the values of the mask

coefficient based on the described above weighted averaging method created originally for noise reduction in biomedical signal. The technique of adaptation of weighted averaging methods to digital filtering 2D images was introduced in (Momot, 2010) where the author used the Simplified Empirical Bayesian Weighted Averaging algorithm (SEBWA), which is described above in 2.2.3.

Given the radius R of the square mask and the input image f of $X \times Y$ dimension, the output image g dimension is $(X - 2R) \times (Y - 2R)$. For each pixel $f(x, y)$, i.e. $x \in \{R + 1, R + 2, \dots, X - R\}$ and $y \in \{R + 1, R + 2, \dots, Y - R\}$, there is computed sum based on the neighborhood of the pixel which determines the pixels $g(x, y)$ of the output image:

$$g(x, y) = \sum_{r=-R}^R \sum_{s=-R}^R w_{rs} f(x + r, y + s). \quad (43)$$

Each value $g(x, y)$ could be calculated by one of the described above iterative algorithms. In the case of SEBWA method the algorithm is given by following procedure:

1. Initialize $g(x, y)^{(0)}$ as the arithmetic average:

$$g(x, y)^{(0)} = \frac{1}{(2R + 1)^2} \sum_{r=-R}^R \sum_{s=-R}^R f(x + r, y + s). \quad (44)$$

If the sample variance of the neighborhood of the pixel:

$$\sigma^2(x, y) = \frac{1}{(2R + 1)^2} \sum_{r=-R}^R \sum_{s=-R}^R \left(f(x + r, y + s) - g(x, y)^{(0)} \right)^2 \quad (45)$$

is greater than zero, set the iteration index $k = 1$ else stop.

2. Calculate the hyperparameter $\alpha_{rs}^{(k)}$ as:

$$\alpha_{rs}^{(k)} = \left(f(x + r, y + s) - g(x, y)^{(k-1)} \right)^{-2}, \quad r, s \in \{1, 2, \dots, R\} \quad (46)$$

and $\beta^{(k)}$ according to the formula:

$$\beta^{(k)} = \left(g(x, y)^{(k-1)} \right)^{-2}. \quad (47)$$

3. Update the average $g(x, y)^{(k)}$ for k th iteration as:

$$g(x, y)^{(k)} = \frac{\sum_{r=-R}^R \sum_{s=-R}^R \alpha_{rs}^{(k)} f(x + r, y + s)}{\beta^{(k)} + \sum_{r=-R}^R \sum_{s=-R}^R \alpha_{rs}^{(k)}}. \quad (48)$$

4. If $\left(g(x, y)^{(k)} - g(x, y)^{(k-1)} \right)^2 > \epsilon$ then $k \leftarrow k + 1$ and go to 2, else stop.

The algorithm assumes that the values $f(x, y)$ are in interval $[0, 1]$. Thus parameter $\beta^{(k)}$ is always positive, although for some values r and s the parameter $\alpha_{rs}^{(k)}$ could be undefined because of dividing by zero (the pixel represented by $(x + r, y + s)$ is equal the average $g(x, y)^{(k-1)}$ in k th iteration). In such case the parameter $\alpha_{r's'}^{(k)}$ should be set to a value significantly greater than other parameters $\alpha_{rs}^{(k)}$.

Numerical experiments presented in (Momot, 2010) evaluate this method with comparison to traditional arithmetic average filtering (mean filtering) and median filtering for synthetic and real images in presence of salt-and-pepper and Gaussian noise. Analyzing results of these methods in the case of salt-and-pepper noise (appearing as white and black dots superimposed on an image), it can be stated that using the new method gives results the same or slightly worse than the median filter. The mean filter for such type of noise gives poor results as expected, but the good results of the new method which originates from mean filtering is worth emphasizing. In the case of Gaussian noise analysis of the results shows that the best results are obtained for mean filtering as expected but the results for the new method are only slightly worse (median filter for such type of noise gives the worst results as expected).

Because in reality noise is often characterized by mixture of these two types, the hypothesis that the new method will give the best results in such cases was suggested. Nevertheless the conducted so far numerical experiments do not confirm such hypothesis and taking into account high computational complexity of this method its usefulness seems to be rather limited. Therefore, similarly as in the case of described above the Fuzzy WACFM algorithm in subsection 2.3.2, this method could be treated as an interesting theoretical study of the adaptation of weighted averaging methods to 2D images.

3. Numerical experiments

In this section there is presented performance of the described above methods. In all experiments, using weighted averaging, calculations were initialized as the means of disturbed signal cycles and the parameter ε was equal to 10^{-6} . For the computed averaged signal the performance of tested methods was evaluated by the root mean-square error (RMSE) between the deterministic component and the averaged signal. The maximal absolute difference between the deterministic component and the averaged signal (MAX) was also computed.

The simulated signal cycles were 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 (Zywietz et al., 2001), designed to reproduce the typical ECG waveform with 60 bpm (beats per minute) heart rate.

This section contains results of several numerical experiments, which show the differences among the weighted averaging methods and the possibility of their applications. First subsection presents the influence of the number of cycles to be averaged on the results of the averaging procedure, next shows the impact of changes in the amplitude and type of noise on the performance of the investigated methods and last subsection describes results obtained when the fuzzy and sharp partition of the ECG signal is applied.

3.1 Influence of the averaged cycles number on the results of the averaging procedure

A series N of ECG cycles was generated with the same deterministic component and zero-mean white Gaussian noise with different standard deviations or real muscle noise with different amplitude. The amplitude of noise was constant during each cycle. The parameter N was equal 20, 40, 60, 80 or 100. For the first, second, third and fourth $N/4$ cycles, the noise standard deviations were respectively 0.1s, 0.5s, 1s, 2s, where s is the sample standard deviation of the deterministic component. Figure 1 presents the signal to be averaged for $N = 60$. The amplitude of the signal is expressed in μV and the length of the deterministic signal L is equal 1000.

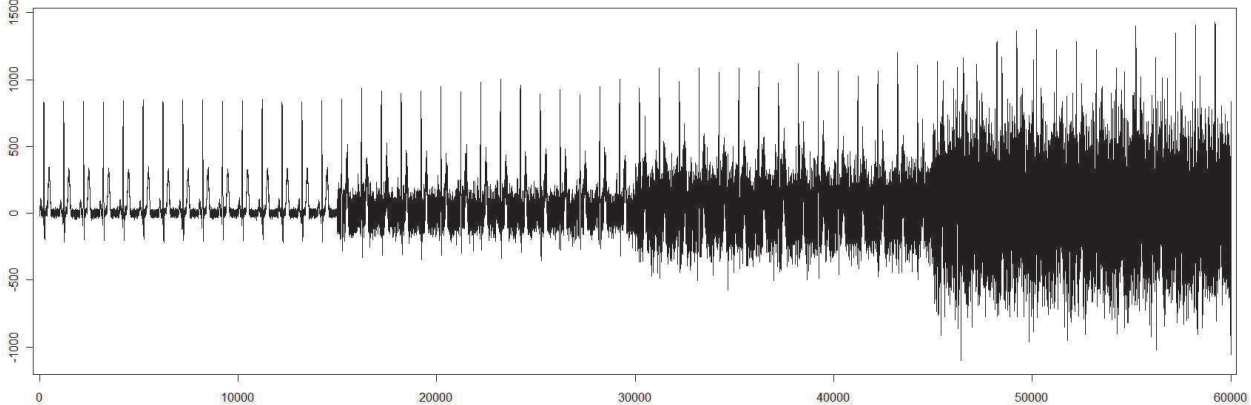


Fig. 1. ANE20000 with added Gaussian noise.

Table 1 presents the results of averaging 20 cycles obtained as the root mean-square error (RMSE) between the deterministic component and the averaged signal. The maximal absolute difference between the deterministic component and the averaged signal (MAX) is also shown in the table. The lower index G characterizes the results obtained for the Gaussian noise and the M indicates the muscle noise.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	36.58372	107.12475	29.13075	91.23949
WACFM	13.76586	43.15516	17.73509	66.74813
WAPM.2	6.215809	23.274264	8.224947	27.706856
WAPM.3	6.472495	24.816161	7.955619	23.171022
WAPM.4	6.399116	22.656614	8.112725	23.318876
SEBWA	6.138603	22.835092	7.672932	24.281101
EBWA.1	6.084599	23.163887	7.542227	24.861945
EBWA.C	5.933567	23.092718	7.167177	24.713475

Table 1. Results for averaging 20 cycles.

The evaluated methods are described by following abbreviations:

- AA** – the traditional Arithmetic Averaging;
- WACFM** – the Weighted Averaging method based on Criterion Function Minimization (the required parameter m is set to 2, it is the value suggested by author of the method in (Leski, 2002));

WAPM – the Weighted Averaging method based on Partition of input data set in time domain and using criterion function Minimization (number after dot describes number of disjoint subsets of input data set);

SEBWA – the Simplified Empirical Bayesian Weighted Averaging method;

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 in (Momot, 2009a));

EBWA.C – the Empirical Bayesian Weighted Averaging method using Cauchy distribution.

The results presented in table 1 show that all the weighted averaging methods give the similar values of $RMSE$ or MAX except the WACFM method which gives the more than twice as worse results, although better than one obtained using traditional arithmetic averaging (AA). The best method in this case seems to be the EBWA.C and this applies to both Gaussian and muscle noise.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	25.48282	90.57738	18.04031	53.39217
WACFM	4.459371	14.307712	12.65819	43.43847
WAPM.2	4.364517	14.906725	5.530065	18.993917
WAPM.3	4.347873	14.549218	4.775429	16.448935
WAPM.4	4.324780	14.256036	5.176974	16.936097
SEBWA	4.260886	14.473659	4.688686	14.488379
EBWA.1	4.228607	14.336413	4.646224	14.305369
EBWA.C	4.153096	13.956814	4.541089	13.865779

Table 2. Results for averaging 40 cycles.

The results of averaging 40 cycles contained in table 2 show that in the case of Gaussian noise all the weighted averaging methods give the similar values of $RMSE$ or MAX . Although in the case of muscle noise the WACFM method once again gives the more than twice as worse results, although still better than one obtained using traditional arithmetic averaging (AA). The best method in this case seems to be the EBWA.C and this applies to both Gaussian and muscle noise.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	20.35780	71.61070	15.71738	48.69969
WACFM	3.825289	12.139730	12.44298	54.07430
WAPM.2	3.729522	12.222603	4.967362	14.089614
WAPM.3	3.684434	12.966805	4.198599	16.084199
WAPM.4	3.723764	12.982499	4.357827	15.039165
SEBWA	3.634302	12.292184	4.198425	16.198985
EBWA.1	3.61344	12.20644	4.168981	15.990323
EBWA.C	3.569731	11.982379	4.102664	15.675835

Table 3. Results for averaging 60 cycles.

The results of averaging 60 cycles contained in table 3 show that all the weighted averaging methods give the similar values of $RMSE$ or MAX both Gaussian and muscle noise (except of the WACFM method). Although the best method in this case seems to be the EBWA.C. As

expected the results indicate that increasing number of cycles to be averaged decreasing the root mean square errors as well as the maximal absolute difference.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	18.37609	77.90424	14.24187	47.40299
WACFM	3.281969	9.794294	4.030636	12.747331
WAPM.2	3.276318	10.421047	4.218003	12.839638
WAPM.3	3.232206	9.746258	3.691453	11.585205
WAPM.4	3.261982	9.962358	3.80167	10.47284
SEBWA	3.186362	9.657737	3.479337	11.553013
EBWA.1	3.177097	9.607379	3.460802	11.482392
EBWA.C	3.151826	9.463573	3.424195	11.293873

Table 4. Results for averaging 80 cycles.

The results of averaging 80 cycles contained in table 4 are similar to the ones in the table 3 (the 60 cycles averaging procedure). For Gaussian and muscle noise all the weighted averaging methods give the similar values of $RMSE$ or MAX and the slightly better method seems to be the EBWA.C. The same pattern can be observed in table 5 contained the results of averaging 100 cycles.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	16.38897	57.06793	13.50228	39.93048
WACFM	2.829289	8.738847	3.726777	11.334553
WAPM.2	2.892244	8.895864	4.02018	12.07286
WAPM.3	2.806800	9.118727	3.561343	10.197009
WAPM.4	2.775190	8.569817	3.638291	11.357989
SEBWA	2.729694	8.378217	3.207977	9.226357
EBWA.1	2.721339	8.511705	3.193742	9.175142
EBWA.C	2.700835	8.544868	3.164265	9.031031

Table 5. Results for averaging 100 cycles.

The obtained results (tables 1 - 5), which can be compared because of the same the noise characteristics in all conducted experiments, indicate that increasing number of cycles to be averaged N decreases the root mean square errors as well as the maximal absolute difference. Thus it seems to be obvious that the better results will be obtained when the N is larger. Although in real application this conclusion may not be held because of problem of time alignment which is critical in the analysis of repetitive signals. Attempt of improvement of signal quality by increasing number of cycles to be averaged increases also the risk of misalignment caused by both noise and signal nonstationarity.

3.2 Influence of noise amplitude changes on the results of the averaging procedure

In the previous subsection the noise amplitude changes between the cycles (constant during each cycle) were described by the following function, where N is the number of cycles to be averaged:

$$A_0(i) = \begin{cases} 0.1, & i \in \{1, \dots, N/4\} \\ 0.5, & i \in \{N/4 + 1, \dots, N/2\} \\ 1, & i \in \{N/2 + 1, \dots, 3N/4\} \\ 2, & i \in \{3N/4 + 1, \dots, N\} \end{cases} \tag{49}$$

multiplied by s , the sample standard deviation of the deterministic component. Now by changing the function A_0 , which determines the noise amplitude changes within each cycle, the influence of noise level on the results of the averaging procedure is investigated. The number of cycles to be averaged is constant in the next experiments and equal 60. Figure 2 presents the cycles of the signal with added nonstationary Gaussian noise. The noise amplitude is described by following function:

$$A_1(i) = \begin{cases} 0.1, & i \in \{1, \dots, 6\} \\ 0.1 + (i - 6)/18, & i \in \{7, \dots, 42\} \\ 2, & i \in \{43, \dots, 54\} \\ (61 - i)/3, & i \in \{55, \dots, 60\}. \end{cases} \quad (50)$$

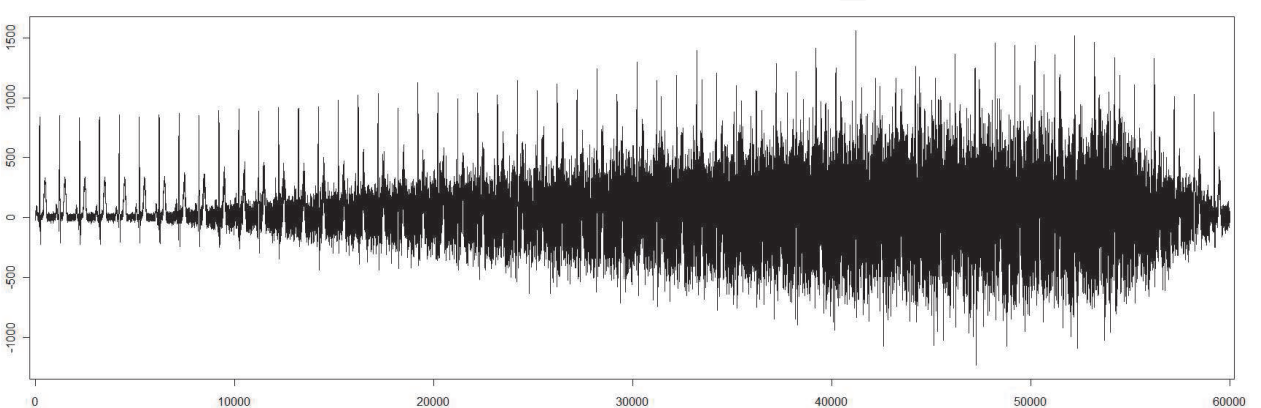


Fig. 2. ECG signal to be averaged with the noise amplitude described by function $A_1(\cdot)$.

Table 6 presents the results of averaging 60 cycles obtained as $RMSE$ (the root mean-square error between the deterministic component and the averaged signal) and MAX (the maximal absolute difference between the deterministic component and the averaged signal). Used abbreviations are the same as in the previous subsections. Comparing the results to the ones presented in table 3, it can be seen that despite the same range of noise amplitude level (from 0.1 to 2), now the results are worse in both cases: Gaussian and muscle noise. All the weighted averaging methods give the similar values of $RMSE$ or MAX except the WACFM method which gives the more than twice as worse results, although better than one obtained using traditional arithmetic averaging (AA). It is worth noting that the WACFM method in the case of rapidly changing the noise amplitude showed numerical instability.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	26.64724	77.79896	19.75347	60.37421
WACFM	14.24589	49.80394	17.84805	80.87108
WAPM.2	5.443441	17.119353	8.052216	26.413713
WAPM.3	5.407122	16.690267	6.933108	20.198006
WAPM.4	5.391607	16.598576	7.314647	27.541761
SEBWA	5.338154	16.257010	6.849912	25.334237
EBWA.1	5.304617	16.036901	6.745417	25.654471
EBWA.C	5.190697	16.181674	6.506906	25.451207

Table 6. Results in the case of the noise amplitude described by function $A_1(\cdot)$.

All the presented above results show that all the weighted averaging methods give usually the similar values of $RMSE$ or MAX , the methods based on statistical inference are slightly better than the methods based on criterion function minimization and the EBWA.C method seems to be the best. The next experiment show that this conclusion may not be held. Figure 3 presents the cycles of the signal with added nonstationary Gaussian noise to be averaged where the noise amplitude is described by following function:

$$A_2(i) = \begin{cases} i/12, & i \in \{1, \dots, 24\} \\ 2, & i \in \{25, \dots, 36\} \\ (61-i)/12, & i \in \{37, \dots, 60\}. \end{cases} \tag{51}$$

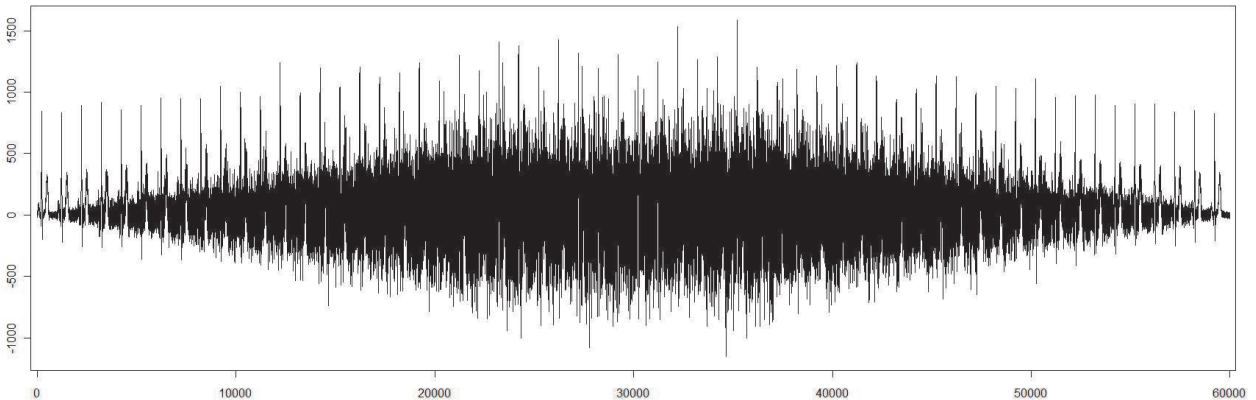


Fig. 3. ECG signal to be averaged with the noise amplitude described by function $A_2(\cdot)$.

Table 7 presents the results of averaging procedure and it can be seen that in this case the best seems to be the WAPM methods, especially WAPM.2 where the input set is divided into two disjoint subsets. In the case of Gaussian noise this method gives the best results both for $RMSE$ and MAX . In the case of real muscle noise this method gives the best result only for MAX and minimal $RMSE$ is obtained for WAPM.3.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	24.95884	86.02304	22.18302	63.14914
WACFM	12.12508	47.95216	8.338565	51.949784
WAPM.2	6.97295	23.53681	7.664572	24.181065
WAPM.3	7.129491	25.233519	7.443705	30.669492
WAPM.4	7.566451	26.480298	7.986691	26.380763
SEBWA	12.12508	47.95216	8.338565	51.949784
EBWA.1	12.12508	47.95216	8.338565	51.949784
EBWA.C	12.12508	47.95216	8.338565	51.949784

Table 7. Results in the case of the noise amplitude described by function $A_2(\cdot)$.

Another interesting fact is equal results for WACFM and all the methods based on Bayesian inference caused the possibility of only one non-zero weight. In this case the methods find the least disturbed cycle and set the corresponding weight equal to 1. In the case of WAPM method the least number of non-zero weights is equal to the number of disjoint subsets. Here this property proved to be effective.

In the next experiment the cycles of the signal are contaminated by added nonstationary Gaussian noise with the amplitude described by following function:

$$A_3(i) = \begin{cases} (25-i)/12, & i \in \{1,\dots,24\} \\ 1/12, & i \in \{25,\dots,30\} \\ (i-30)/15, & i \in \{31,\dots,60\}. \end{cases} \tag{52}$$

Figure 4 presents the cycles of the signal and table 8 shows the results of averaging procedure in this case. Like previously all the weighted averaging methods give the similar values of *RMSE* or *MAX* (except the WACFM method), the methods based on statistical inference are slightly better than the methods based on criterion function minimization and the EBWA.C method seems to be the best.

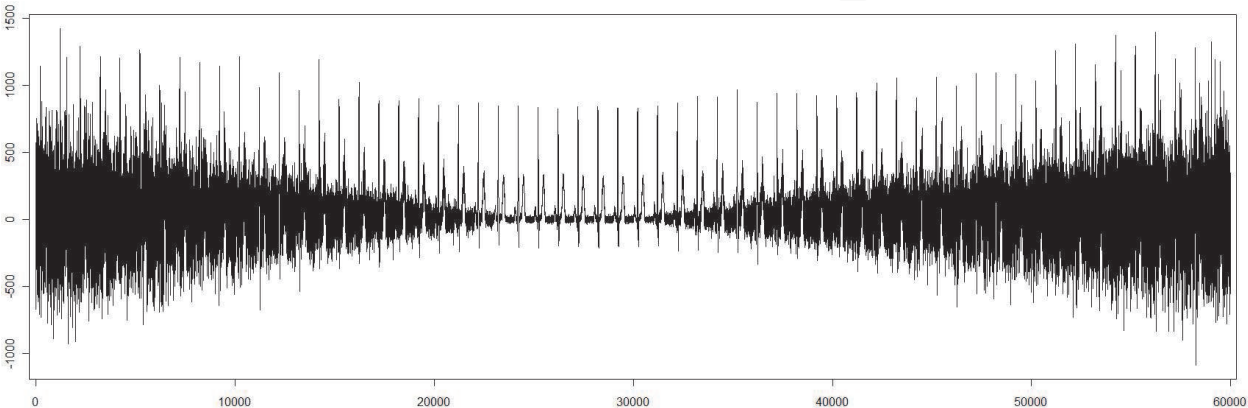


Fig. 4. ECG signal to be averaged with the noise amplitude described by function $A_3(\cdot)$.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	22.18602	75.84577	20.08812	69.75064
WACFM	6.70973	21.06666	9.59511	30.20457
WAPM.2	3.530179	12.508598	3.84337	12.16353
WAPM.3	3.939755	12.860902	3.828325	12.001807
WAPM.4	3.393327	12.382687	3.773901	11.322390
SEBWA	3.048759	13.518930	3.747146	10.741282
EBWA.1	3.024732	13.315327	3.729486	10.967476
EBWA.C	2.993796	13.191613	3.686413	11.094574

Table 8. Results in the case of the noise amplitude described by function $A_3(\cdot)$.

The last experiment study the results with presence of noise with amplitude characteristics described by:

$$A_4(i) = i/30, \quad i \in \{1,\dots,60\}. \tag{53}$$

The signal to be averaged is presented in figure 5 and table 9 shows the results of averaging procedure. In this case the best seems to be the WAPM.2 method. All the weighted averaging methods give the similar values of *RMSE* or *MAX* and equals for WACFM and all the methods based on Bayesian inference caused finding the least disturbed cycle and set the corresponding weight equal to 1.

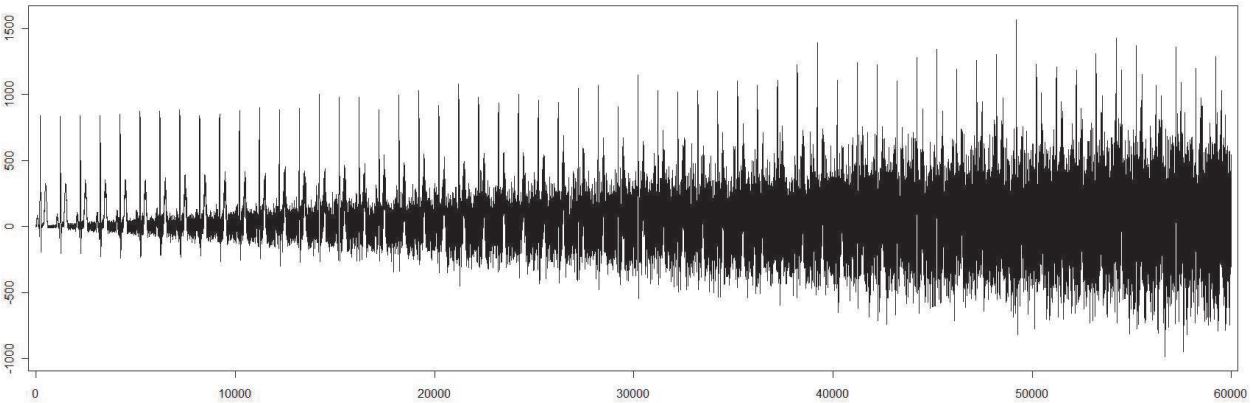


Fig. 5. ECG signal to be averaged with the noise amplitude described by function $A_4(\cdot)$.

Method	$RMSE_G$	MAX_G	$RMSE_M$	MAX_M
AA	22.19804	80.13893	16.19718	51.03785
WACFM	4.563198	15.679271	8.915569	28.417666
WAPM.2	4.220262	12.790181	6.802766	17.392957
WAPM.3	4.942776	17.736005	6.660585	23.792032
WAPM.4	5.453174	20.018136	7.315934	21.823084
SEBWA	4.563198	15.679271	8.915569	28.417666
EBWA.1	4.563198	15.679271	8.915569	28.417666
EBWA.C	4.563198	15.679271	8.915569	28.417666

Table 9. Results in the case of the noise amplitude described by function $A_4(\cdot)$.

3.3 Influence of the partition of the signal on the results of the averaging procedure

Below there are presented results of numerical experiments which investigate how partition of the input signal affects the results of the averaging procedure. In the experiments it is studies both sharp and fuzzy partitions described in subsection 2.3.1 and because of numerical instability of WACFM algorithm the method is omitted. The number of parts K is in $\{2, 3, 4, 5\}$ and for $K = 1$ the original method is used. Similarly to the previous subsection the number of cycles to be averaged is constant in the next experiments and equal 60.

First experiment studies influence of the partition on the root mean square error in the case presented in figure 1, 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 s , i.e. the sample standard deviation of the deterministic component. Results presented in table 3 show the RMSE equal 20.35780 in the case of the traditional arithmetic averaging method and the RMSE of the weighted averaging methods range from 3.569731 (EBWA.C) to 3.825289 (WACFM). This time the results are slightly different because of randomness of the noise and the RMSE for the traditional arithmetic averaging method is equal 22.09798. Detailed results for the weighted averaging methods are presented in table 10.

The root mean square errors presented in table 10 are computed in both type of partitions: sharp and fuzzy, with taking into consideration varying number of parts K . Obviously for $K = 1$ the results obtained for sharp and fuzzy partitions are equal (the signal in the single cycle is not divided). Analyzing the results presented in table 10 it is easy to conclude that without the partition all method gives similar RMSE although the methods based on Bayesian inference (SEBWA, EBWA.1, EBWA.C) are slightly better. The partition in the case of

methods based on criterion function minimization (WAPM.2, WAPM.3, WAPM.4) results in deterioration of the RMSE (the same pattern may be seen for the WACFM method although the numerical instability of the algorithm makes it difficult and it is the reason that the results are not presented).

Method	type	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$
WAPM.2	sharp	3.671434	3.735696	3.830139	3.981567	4.08268
	fuzzy		3.733926	3.808319	3.934622	4.046011
WAPM.3	sharp	3.588868	3.735696	3.830139	3.981567	4.08268
	fuzzy		3.733926	3.808319	3.934622	4.046011
WAPM.4	sharp	3.603876	3.624834	3.67803	3.706631	3.766816
	fuzzy		3.627341	3.665056	3.681049	3.757504
SEBWA	sharp	3.563763	3.570995	2.943464	3.112997	2.762172
	fuzzy		3.574001	2.884799	3.084993	2.777102
EBWA.1	sharp	3.547627	3.437116	3.022750	3.185023	3.024058
	fuzzy		3.478558	2.967808	3.166432	3.039432
EBWA.C	sharp	3.50404	2.823939	3.184680	3.562267	3.363044
	fuzzy		2.955845	3.110675	3.242250	3.655863

Table 10. RMSE for zero-mean Gaussian noise with amplitude described by function $A_0(\cdot)$.

Next experiment studies influence of the partition on the root mean square error in the case presented in figure 3, where the cycles of the signal were disturbed by nonstationary Gaussian noise with the noise amplitude described by function $A_2(\cdot)$. It is interesting case because here the best seems to be the WAPM method and using the Bayesian methods results in the same value of RMSE. In this experiment the RMSE for the traditional arithmetic averaging method is equal 25.25697 and detailed results for the weighted averaging methods are presented in table 11.

Method	type	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$
WAPM.2	sharp	6.581978	6.878741	7.114349	7.615843	7.950328
	fuzzy		6.842354	7.083266	7.499244	7.751919
WAPM.3	sharp	6.913927	6.878741	7.114349	7.615843	7.950328
	fuzzy		6.842354	7.083266	7.499244	7.751919
WAPM.4	sharp	7.348039	7.425919	7.52601	7.751074	7.85626
	fuzzy		7.411294	7.532835	7.740782	7.784718
SEBWA	sharp	11.50770	11.66183	9.444625	10.08938	9.07164
	fuzzy		11.31708	9.30354	9.51357	8.564705
EBWA.1	sharp	11.50770	11.66183	9.555289	10.20881	9.443687
	fuzzy		11.31708	9.305964	9.642958	8.869182
EBWA.C	sharp	11.50770	8.726462	5.347378	5.509787	5.98894
	fuzzy		8.695319	5.481904	6.983782	6.328657

Table 11. RMSE for zero-mean Gaussian noise with amplitude described by function $A_2(\cdot)$.

The results presented in table 11 show that the partition in the case of methods based on statistical inference improves the results, even significantly for the EBWA.C method (the RMSE decreases more than twice).

The improvement or deterioration of results obtained using the partition are not expected in the cases when the amplitude of noise is constant within each cycle. The difference can be explained by randomness of the noise level in each part of the divided signal cycle despite the same value of standard deviation of random variable with Gaussian distribution which characterize the noise. Utility of the partition seems to be obvious in the cases when the assumption of noise amplitude constancy within each cycle is not hold. Presented below results of numerical experiments where the the amplitude of noise is not constant within each cycle show improvement of the root mean square errors for all tested methods. Figure 6 presents the ECG signal to be averaged in this experiment, which is disturbed by Cauchy noise. The location parameter of Cauchy distribution is equal to 0 and the scale parameter is set to 0.01s, where s is the standard deviation of the deterministic component, i.e. the original ANE20000 signal.

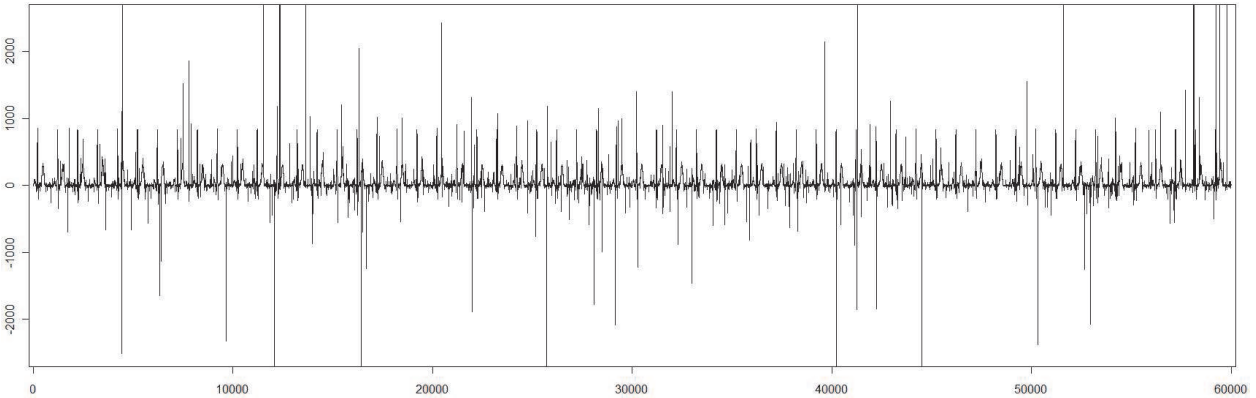


Fig. 6. ECG signal disturbed by Cauchy noise with scale parameter 0.01s.

The RMSE for the traditional arithmetic averaging method is equal 38.25505 and detailed results for the weighted averaging methods are presented in table 12.

Method	type	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$
WAPM.2	sharp	4.293157	3.590644	5.802226	4.875072	3.05765
	fuzzy		3.782668	3.981781	3.457431	3.059209
WAPM.3	sharp	4.336904	3.590644	5.802226	4.875072	3.05765
	fuzzy		3.782668	3.981781	3.457431	3.059209
WAPM.4	sharp	4.124156	3.46319	2.910611	2.551464	2.364648
	fuzzy		3.560695	3.018473	2.807478	2.468142
SEBWA	sharp	4.044945	3.195667	2.195166	1.917149	1.644676
	fuzzy		3.396122	2.191550	2.215406	1.721535
EBWA.1	sharp	4.027516	3.093357	2.410249	2.033164	1.840781
	fuzzy		3.328734	2.384026	2.307351	1.923214
EBWA.C	sharp	3.970909	2.622315	2.405902	1.896787	1.818165
	fuzzy		2.966283	2.475908	2.14492	1.918657

Table 12. RMSE in the case of Cauchy noise with scale parameter 0.01s.

Analyzing the results it can be observed the improvement of the root mean square errors for all tested methods at least for some values of K . As was mentioned in the subsection 2.3.1 in the case of sharp partition incorrectly chosen number of parts may cause the deterioration of the results.

Figure 7 presents the ECG signal to be averaged in the last experiment, which is disturbed by Cauchy noise with the location parameter 0 and the scale parameter 0.05s, the standard deviation of ANE20000 signal. As can be seen the input signal is very distorted, particularly visible are many random impulse values. The RMSE for the traditional arithmetic averaging method is equal 2143.182 and detailed results for the weighted averaging methods are presented in table 13.

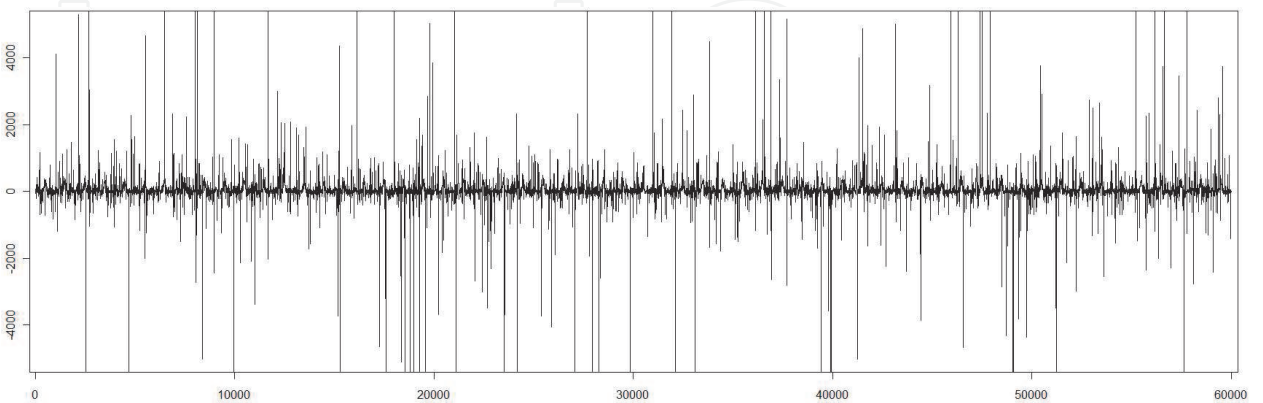


Fig. 7. ECG signal disturbed by Cauchy noise with scale parameter 0.05s.

Method	type	$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$
WAPM.2	sharp	26.33396	23.59549	18.82174	16.77062	15.80005
	fuzzy		22.58925	19.25907	17.36819	15.80776
WAPM.3	sharp	23.19694	23.59549	18.82174	16.77062	15.80005
	fuzzy		22.58925	19.25907	17.36819	15.80776
WAPM.4	sharp	25.47250	20.02517	14.65732	13.77551	12.44631
	fuzzy		20.07785	15.25705	14.32465	12.49090
SEBWA	sharp	22.60814	17.82965	10.86278	10.33561	8.407054
	fuzzy		17.94750	10.86471	10.68507	8.42861
EBWA.1	sharp	21.00193	15.65491	10.68224	9.928945	8.798122
	fuzzy		16.40589	10.69901	10.31931	8.69519
EBWA.C	sharp	30.63047	24.34970	16.48281	12.77003	10.85225
	fuzzy		26.39931	16.21782	15.39638	11.52634

Table 13. RMSE in the case of Cauchy noise with scale parameter 0.05s.

Such great distortion of the signal in real cases obviously causes rejection of the disturbed signal, however this example shows that having the a priori information about the number of cycles and the length of the single cycle allows using some weighted averaging methods which recover the data with the relatively small error (compare the RMSE 2143.182, for the arithmetic averaging, and 8.407, for the SEBWA method with the sharp partition for $K = 5$).

4. Summary

This chapter presents several methods of weighted averaging: based on criterion function minimization (WACFM, WAPM) and on statistical inference (EBWA.1, EBWA.3, EBWA.C, SEBWA) together with fuzzy extensions, which use the fuzzy partition the signal cycle as well as using fuzzy numbers as coefficients of weight vector. The adaptation of SEBWA method to

filtering of 2D images is also presented. This study reveals the fundamental differences among the weighted averaging methods and presents, through several numerical experiments, how these differences affect the quality of the averaged signal.

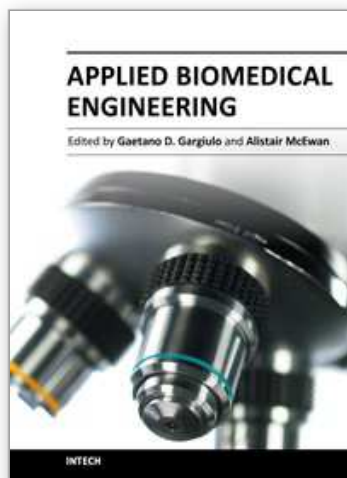
The most frequently used method, due to its simplicity, is the arithmetical averaging. The improvement of results obtained by the method can be reached rejecting the very noisy cycles. Averaging with rejecting very noisy cycles can be treated as weighted averaging method where the weights corresponding these cycles are set to zero. The crucial problem is how to find these cycles. The presented weighted averaging methods implement the automatic determining the weights, such that the most noisy cycles have the smallest weights (even zero in particular) and the least noisy ones receive the greatest weights, which increase their influence on the resulting averaged signal. Analyzing the presented results of performed numerical experiments, it is difficult to determine the best method, because for various power and type of noise accompanying the signal, different methods appear to give the best results.

This research was partially supported by Polish Ministry of Science and Higher Education as Research Project N N518 406438.

5. References

- Augustyniak, P. (2006). Adaptive wavelet discrimination of muscular noise in the ECG, *Computers in Cardiology*, Vol. 33, pp. 481–484, ISSN: 0276-6574.
- Bailon, R.; Olmos, S.; Serrano, P.; Garcia J. & Laguna P. (2002). Robust Measure of ST/HR Hysteresis in Stress Test ECG Recordings, *Computers in Cardiology*, Vol. 29, pp. 329–332, ISSN: 0276-6574.
- Bataillou, E.; Thierry, E.; Rix, H. & Meste, O. (1995). Weighted averaging using adaptive estimation of the weights, *Signal Process.* Vol. 44, pp. 51–66, ISSN: 0165-1684.
- Bruce, E.N. (2001). *Biomedical signal processing and signal modeling*, Wiley, ISBN: 0-471-34540-7, New York.
- Davila, C.E. & Mobin M.S. (1992). Weighted averaging of evoked potentials. *IEEE Trans. Biomed. Eng.*, Vol. 39, No. 4, pp. 338–345, ISSN: 0018-9294.
- Elberling, C. & Don, M. (2006). Detecting and assessing synchronous neural activity in the temporal domain, In: *Auditory Evoked Potentials - Basic principles and Clinical Application*, Burkard, R.F.; Eggermont, J.J.; Don, M. (Eds.), pp. 102–123, Lippincott Williams & Wilkins, ISBN: 978-0-7817-5756-0, Philadelphia.
- Jane, R.; Rix, H.; Caminal, P. & Laguna, P. (1991). Alignment methods for averaging of high-resolution cardiac signals: a comparative study of performance, *IEEE Trans. Biomed. Eng.*, Vol. 38, No. 6, pp. 571–579, ISSN: 0018-9294.
- Johnson, R.A.; Bhattacharyya, G.K. (2009). *Statistics: Principles and Methods*, John Wiley & Sons, ISBN: 978-0-470-40927-5.
- Kotas, M. (2009). Nonlinear projective filtering of ECG signals, In: *Biomedical engineering*. Mello C.A.B., (Ed.), pp. 433–452, InTech, ISBN: 978-953-307-013-1.
- Laciar, E. & Jane, R. (2001). An Improved Weighted Signal Averaging Method for High-Resolution ECG Signals, *Computers in Cardiology*, Vol. 28, pp. 69–72, ISSN: 0276-6574.
- Leski, J. (2002). Robust weighted averaging, *IEEE Trans. Biomed. Eng.*, Vol. 49, No. 8, pp. 796–804, ISSN: 0018-9294.

- Momot, A.; Momot M. & Leski J. (2005). The Fuzzy Relevance Vector Machine and its Application to Noise Reduction in ECG Signal, *J. Med. Inform. Technol.*, Vol. 9, pp. 99–106, ISSN: 1642-6037.
- Momot, A.; Momot, M. & Leski, J. (2007a). Bayesian and empirical Bayesian approach to weighted averaging of ECG signal, *Bull. Pol. Acad. Sci., Technol. Sci.*, Vol. 55, No. 4, pp. 341–350, ISSN: 0239-7528.
- Momot, A.; Momot, M. & Leski, J. (2007b). Weighted Averaging of ECG Signals Based on Partition of Input Set in Time Domain, *J. Med. Inform. Technol.*, Vol. 11, pp. 165–170, ISSN: 1642-6037.
- Momot, A. (2008a). Weighted Averaging of ECG Signal Using Criterion Function Minimization, In: *Information Technologies in Biomedicine, Advances in Soft Computing*, Vol. 47, Pietka, E.; Kawa, J. (Eds.), pp. 267–274, Springer-Verlag, ISBN: 978-3-540-68167-0, Berlin Heidelberg.
- Momot A., Momot M. (2008b). Empirical Bayesian Approach to Weighted Averaging of ECG Signal Using Cauchy Distribution. In: *Information Technologies in Biomedicine, Advances in Soft Computing*, Vol. 47, Pietka, E.; Kawa, J. (Eds.), pp. 275–282, Springer-Verlag, ISBN: 978-3-540-68167-0, Berlin Heidelberg.
- Momot, A. (2009a). Methods of Weighted Averaging of ECG Signals Using Bayesian Inference and Criterion Function Minimization, *Biomed. Signal Process. Control*, Vol. 4, No. 2, pp. 162–169, ISSN: 1746-8094.
- Momot, A. (2009b). Fuzzy Weighted Averaging of Biomedical Signal Using Bayesian Inference, In: *Man-Machine Interactions, Advances in Intelligent and Soft Computing*, Vol. 59, Cyran, K.A.; Kozielski, S.; Peters, J.F.; Stanczyk, U.; Wakulicz-Deja, A. (Eds.), pp. 133–140, Springer-Verlag, ISBN: 978-3-642-00562-6, Berlin Heidelberg.
- Momot, A. & Momot, M. (2009c). Fuzzy Weighted Averaging Using Criterion Function Minimization, In: *Man-Machine Interactions, Advances in Intelligent and Soft Computing*, Vol. 59, Cyran, K.A.; Kozielski, S.; Peters, J.F.; Stanczyk, U.; Wakulicz-Deja, A. (Eds.), pp. 273–280, Springer-Verlag, ISBN: 978-3-642-00562-6, Berlin Heidelberg.
- Momot, A. (2010). Application of Adaptive Weighed Averaging to Digital Filtering of 2D Images, In: *Information Technologies in Biomedicine (vol.2), Advances in Soft Computing*, Vol. 69, Pietka, E.; Kawa, J. (Eds.), pp. 33–44, Springer-Verlag, ISBN: 978-3-642-13104-2, Berlin Heidelberg.
- Momot A. (2011). On Application of Input Data Partitioning to Bayesian Weighted Averaging of Biomedical Signals, *Expert Systems*, in press (accepted 28.01.2011).
- Paul, J.S.; Reddy, M.R.; Kumar, V.J. (2000). A transform domain SVD filter for suppression of muscle noise artefacts in exercise ECG's, *IEEE Trans. Bimed. Eng.*, Vol. 47, No. 5, pp.654–663, ISSN: 0018-9294.
- Sharma, L. N.; Dandapat, S. & Mahanta, A. (2010). ECG signal denoising using higher order statistics in Wavelet subbands, *Biomed. Signal Process. Control*, Vol. 5, No. 3, pp. 214–222, ISSN: 1746-8094.
- Yan, J.; Lu, Y.; Liu, J.; Wub, X. & Xu, Y. (2010). Self-adaptive model-based ECG denoising using features extracted by mean shift algorithm, *Biomed. Signal Process. Control*, Vol. 5, No. 2, pp. 103–113, ISSN: 1746-8094.
- Zywietz, C.; Alraun, W. & Fischer, R. (2001). Quality assurance in biosignal processing - procedures and recommendations for evaluation for electrocardiological analysis systems, *Computers in Cardiology*, Vol. 28, pp. 201–204, ISSN: 0276-6574.



Applied Biomedical Engineering

Edited by Dr. Gaetano Gargiulo

ISBN 978-953-307-256-2

Hard cover, 500 pages

Publisher InTech

Published online 23, August, 2011

Published in print edition August, 2011

This book presents a collection of recent and extended academic works in selected topics of biomedical technology, biomedical instrumentations, biomedical signal processing and bio-imaging. This wide range of topics provide a valuable update to researchers in the multidisciplinary area of biomedical engineering and an interesting introduction for engineers new to the area. The techniques covered include modelling, experimentation and discussion with the application areas ranging from bio-sensors development to neurophysiology, telemedicine and biomedical signal classification.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Alina Momot (2011). Methods of Weighted Averaging with Application to Biomedical Signals, Applied Biomedical Engineering, Dr. Gaetano Gargiulo (Ed.), ISBN: 978-953-307-256-2, InTech, Available from: <http://www.intechopen.com/books/applied-biomedical-engineering/methods-of-weighted-averaging-with-application-to-biomedical-signals>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen