Impulsive Noise Filtering In Biomedical Signals With Application of New Myriad Filter

Tomasz Pander [1]
[1] Division of Biomedical Electronics, Institute of Electronics,
Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland,
tpander@polsl.pl

Abstract. The noise cancellation should be the first step of each signal processing device. The most difficult type of noise which should be suppressed is an impulsive type of noise. The group of nonlinear robust filter successfully supress such noise. The aim of this work is to present a new approach to myriad filter computation and its application in biomedical signal processing. This new method is based on 2nd order polynomial approximation. The special conditions are proposed for proper used of this new myriad filter. The effectiveness of such approach is shown for ECG signal filtering in the presence of an impulsive noise. This kind of noise is simulated using Symmetric-α-Stable distribution and the level of noise is controlled with Geometric-SNR. Additionally, the adaptive estimation of the linear parameter K of the myriad filter is proposed.

1 Introduction

Noises which appear during recording of biomedical signal are seldom characterized by the ideal Gaussian distribution, whereas a noise can have an impulsive nature. An application of traditional filtering method at such assumption can lead to introducing a distortion to a filtered signal. The process of noise suppression should be the first step of each biomedical signal processing system. The accuracy of farther signal processing steps like detection, measurement or classification depend on the quality of reduction noise algorithms. In the biomedical systems the first step of the processing of biomedical signals is very important. All later activities depend on the quality of the initial step [13]. For the purpose of this work the electrocardiogram (ECG signal) was chosen as a representation of biomedical signals.

A white Gaussian noise is usually used to model the muscle noise in ECG processing. This noise shows frequently an impulsive nature, and it means that the Gaussian model may fail. Another model which very likely describes some cases of the muscle noise is an application of the symmetric α-stable distribution [15]. One of the most disadvantage of linear filters are their sensitivity to outliers samples. More appropriate should be a non-linear group of filters which are robust to outliers in a filtered signal. The myriad filter has recently been proposed as a robust, non-linear filtering and an estimation technique in impulsive environments. It represents a wide class of maximum likelihood type estimators (M-estimators) of location [9, 10].

The main purpose of this paper is an application of the non-linear new myriad filter in biomedical signal processing (for example an ECG signal). This new myriad filter is based on the 2nd order polynomial approximation of data samples in the filter window. The reference filter is the myriad filter with the fixed point searching method. The plan of this paper is the following. It begins with introduction a myriad filter theory. Presentation of the new method of myriad filter output estimation is introduced in the next section. Finally, a number of results and conclusions are presented.

2 The M-estimator and the myriad filter

The M-estimator technique is one of the most popular methods to solve the problem of robust parameter estimation. The principle of M-estimators can be formulated in the following way. Let us choose a set of N data samples $x_1, x_2, \ldots, x_N$, where $x_i = \beta_i + v_i$ and $1 \leq i \leq N$, the problem is to estimate the β_i location parameter under noise component v_i. This parameter identifies the position of the probability density function (*pdf*) on the real line of data samples. The distribution of v_i in not assumed to be exactly known. The only basic assumption is that $v_1, \ldots, v_N$ obey a symmetric, independent, identical distribution (symmetric i.i.d.) [7, 12].

The M-estimate or maximum likelihood-type estimator (MLE) was originally proposed to improve robustness of statistical estimators subject to the small deviations mentioned above. The M-estimator of $\hat{\beta}$ is defined as the minimum of a global energy function:

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{N} \rho(x_i - \beta). \tag{1}$$

The function $\rho(\cdot)$ is usually known as the cost function. The term „cost" is explained from the engineering interpretation that a „penalty" with value $\rho(x_i - \beta)$ shall be paid for the estimator to be away from sample x_i. Under this point of view, the M-estimator is the point β with the minimum sum of costs [5]. An M-estimator of location is defined as the parameter $\hat{\beta}$ that minimizes the expression in (1). The behavior of the M-estimator is entirely characterized by the shape of ρ function [2, 6, 12]. The different types of M-estimators and corresponding filters are presented in the table 1 [1]. In each row of the table 1, the filter output is the value which minimizes the associated cost function

filter	cost function	filter ouput
mean	$\sum_{i=1}^{N}(x_i - \beta)^2$	mean$\{x_i\|_{i=1}^{N}\}$
median	$\sum_{i=1}^{N}\|x_i - \beta\|$	median$\{x_i\|_{i=1}^{N}\}$
myriad	$\sum_{i=1}^{N}\log[K^2 + (x_i - \beta)^2]$	myriad$\{x_i\|_{i=1}^{N}; K\}$
meridian	$\sum_{i=1}^{N}\log[\delta + \|x_i - \beta\|]$	meridian$\{x_i\|_{i=1}^{N}; \delta\}$
weighted mean	$\sum_{i=1}^{N} w_i(x_i - \beta)^2$	mean$\{w_i \cdot x_i\|_{i=1}^{N}\}$
weighted median	$\sum_{i=1}^{N} w_i\|x_i - \beta\|$	median$\{x_i \diamond w_i\|_{i=1}^{N}\}$
weighted myriad	$\sum_{i=1}^{N}\log[K^2 + w_i(x_i - \beta)^2]$	myriad$\{w_i \circ x_i\|_{i=1}^{N}; K\}$
weighted meridian	$\sum_{i=1}^{N}\log[\delta + w_i\|x_i - \beta\|]$	meridian$\{w_i \star x_i\|_{i=1}^{N}; \delta\}$

Tab. 1: M-estimator cost functions and filter outputs for various filter families

Let $\rho(z)$ be a function of a single variable $z \equiv (x_i - \beta)$. Given the Cauchy distribution with scaling factor $K > 0$,

$$f(z) = \frac{K}{\pi}\frac{1}{K^2 + z^2}, \tag{2}$$

its associated cost function has a form:

$$\rho(z) = \log(K^2 + z^2). \tag{3}$$

The maximum likelihood estimator of location (M-estimator) associated with this cost function is called a *sample myriad*. The scaling factor K is called *linearity parameter* and controls the impulse-resistance (outlier-rejection capability) of the estimator. This parameter offers a rich class of operation that can be easily controlled by changing value of K. Tuning K parameter adapts the behavior of the myriad filter from impulse-resistant mode-type estimators (small values of K) to the Gaussian-efficient sample mean (large values of K). For details see in [3, 6, 8, 9, 10, 11].

The myriad filter is also defined as the weighted myriad filter by assigning weights to the sample in the maximum likelihood of the location estimation. The weights reflect the different levels of reliability of the observed samples [9, 10, 11]. Let us consider a set of observations

$\{x_i\}_{i=1}^N$ and a set of filter weights $\{w_i\}_{i=1}^N$. The output of the weighted myriad filter is defined as

$$\begin{aligned}\hat{\beta}\hat{=}myriad(x_1 \circ w_1, x_2 \circ w_2, \ldots, x_N \circ w_N; K) = \\ = \arg\min_\beta \prod_{i=1}^N [K^2 + w_i(x_i - \beta)^2] = \\ = \arg\min_\beta \sum_{i=1}^N \log[K^2 + w_i(x_i - \beta)^2]\end{aligned} \quad (4)$$

where $x_i \circ w_i$ denotes the weighting operation and the weights are restricted to be non-negative: $w_i \geq 0, i = 1, \ldots, N$, N is the window length.

It is important to realize that the location estimation problem being considered in this paper is related to the problem of filtering a time-series $x(n)$ using a sliding window. The output $y(n)$ can be interpreted as an estimate of location based on input samples:

$$y(n) = myriad\{x(n - N_1), \ldots, x(n), \ldots, x(n + N_2); K\} \quad (5)$$

where N is window length (if N is even then $N_1 = \frac{N}{2}$ and $N_2 = N_1 - 1$, if N is odd then $N_1 = \frac{N-1}{2}$ and $N_2 = N_1$). In this work, to make it more simple, $w_i = \frac{1}{N}$ for $i = 1, \ldots, N$ and a further analysis of the weights value is not considered in the present work.

As table 1 shows, it is trivial to compute the mean or the weighted mean filter. The median and the weighted median can also be estimated directly, however it requires sorting the input samples, making algorithm more complicated. But the others two filters are not available to computer in explicit form [11]. The main algorithm of estimation of the myriad filter output is the fixed-point method presented in [9, 11]. Recently in [14] a new way of estimation of the myriad filter output using the branch-and-bound algorithm was presented.

3 The myriad filter output estimation applying 2nd order approximation polynomial

The method proposed in this paper is based on the simple idea of application the least squares approximation with the 2nd order polynomial.

The main problem in the least squares numerical approximation is finding the best fit function $f(x)$, that is close to the data samples: (x_1, y_1), (x_2, y_2), ..., (x_n, y_n).

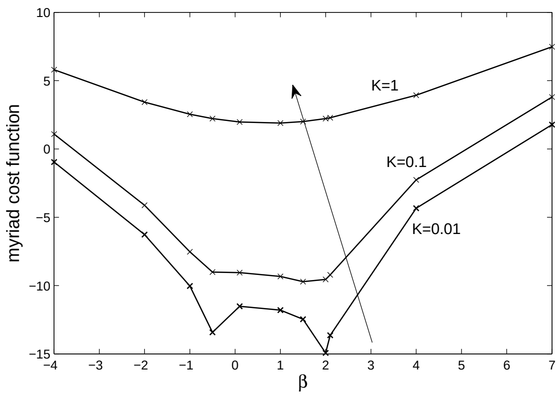

Fig. 1: Degree of non-convexity in the slide window for the myriad filter cost function calculated for data vector (-2 1.5 1 2.1 -0.5 -1 4 2 -4 2 -.5 0.1 7). The plot shows the sorted x_i data and corresponding values of the myriad cost function.

The figure 1 presents the sorted data **x** and corresponding the myriad cost function values **y** from eq. (3). According to eq. (4) β is such a value for which the expression at right side of eq. (4) reaches the minimum value. The main problem is that there can exist many local minima. The fixed-point method from [3, 9, 11] calculates iteratively this optimal β value, but requires at least 10 iterations to find it and this method can not be applied in on-line applications.

The idea of the proposed method is based on the observation that the sorted data samples in the filter window and their corresponding myriad cost function values are placed on a parabolic. Then coefficients of the 2nd order polynomial are calculated using the least square approximation. And the location of the peak of parabola is estimated. This value corresponds to the optimal β from eq. (4). The proposed algorithm can be described as follows [15]:

1. sort in ascending order samples $(x_1, x_2, \ldots, x_N)$ in the sliding window of the myriad filter $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(N)}$.
2. calculate the cost function of the myriad filter for each sample x_i in the window:

$$y_i = \sum_{j=1}^{N} \log[K^2 + w_j(x_j - x_i)^2]. \tag{6}$$

3. assure the approximation method from outliers in the filter window, the cut-off process is made in the following way:

$$\forall_{i=1,\ldots,N} |x_{(i)}| < \varepsilon. \tag{7}$$

   If condition (7) is true, then the filter window is cut down on a sample $x_{(i)}$ and a new filter window is created which consists of samples $(x_{(1)}, x_{(2)}, \ldots, x_{(N_1)})$, where N_1 is the new filter window length. Thus, N has a new value $N = N_1$. Similarly, the set of the cost function values is also cut down on a sample $y_{(i)}$, hence the set of values is the following: $(y_{(1)}, y_{(2)}, \ldots, y_{(N_1)})$. Value of ε is chosen in an empirical way and in this work, for ECG signal filtering, $\varepsilon = 10$.
4. the coefficients a_0, a_1, a_2 of the 2nd order polynomial are estimated using the least square approximation:

$$f_{myrappr}(\beta) = a_0 + a_1\beta + a_2\beta^2. \tag{8}$$

To solve this problem the following matrix **S** and vector **T** are created:

$$\begin{bmatrix} S_0 & S_1 & S_2 \\ S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} T_0 \\ T_1 \\ T_2 \end{bmatrix} \tag{9}$$

where $S_i = \sum_{j=1}^{N} x_j^i$ and $i = 0, 1, \ldots, 2m$, $T_k = \sum_{j=1}^{N} x_j^k y_j$ and $k = 0, 1, ..., m$, and for quadratic polynomial approximation $m = 2$. The Gauss elimination is applied to find coefficients a_0, a_1 and a_2.

The coordinate x of the parabola top corresponds to β value for which eq. (4) gets the minimum value:

$$\hat{\beta} = -\frac{a_1}{2a_2}. \tag{10}$$

In order to ensure the condition that solution belongs to the range $x_{(1)} \leq \hat{\beta} \leq x_{(N)}$, it is checked if for the obtained value $\hat{\beta}$ inequalities $\hat{\beta} < x_{(1)}$ or $\hat{\beta} > x_{(N)}$ hold true. If these last inequalities are true then the solution $\hat{\beta}$ is selected from the set of samples $x_{(i)}$:

$$\hat{\beta} \equiv x_{(i)}|_{y_{min}=\min(y_1, y_2, \ldots, y_N)} \tag{11}$$

The second condition is formulated in the following form:

1. find $y_{min} = \min(y_1, y_2, \ldots, y_N)$,
2. from eq. (8) calculate $f_{myrappr}(\hat{\beta})$,
3. check the following inequality $y_{min} < f_{myrappr}(\hat{\beta})$, if it holds true, then $\hat{\beta} \equiv x_{(i)}|_{y_{min}}$.

This last conditions can lead to the selected myriad filter.

4 Filtering evaluation procedure

The presented in this paper method is evaluated through a computer simulation procedure involving the filtering the ECG signal corrupted by a noise distributed with the symmetric α-stable distribution. A single ECG cycle is used in this experiment. The signal is sampled 2000 times per second and consists of 1560 samples. In order to obtain the signal with high geometric signal-to-noise ratio (GSNR), the signals are modeled by a linear combination of the Hermite functions. These signals are corrupted with the α-stable distributed noise for different values of α. The noisy ECG cycles are obtained by adding each ECG cycle to a noise with the various GSNR (see fig. 2). The variance of random variable of the α-stable distribution doesn't exist, that's why the GSNR is used and this factor is defined as [4]:

$$\mathrm{GSNR} = \frac{1}{2C_g}\left(\frac{A}{S_0}\right)^2, \tag{12}$$

where $C_g = e^{C_e} \approx 1.78$ is the exponential of the Euler constant, A is the amplitude of a modulated signal in an additive-noise channel with noise geometric power S_0. The normalization constant $2C_g$ is used to ensure that the definition of the GSNR corresponds to that of the standard SNR if the channel noise is Gaussian [4]. The geometric power definition is the following $S_0 = \left(\prod_{i=1}^{M}|x_i|\right)^{1/M}$ [4].

Filtering a signal in the time-domain results in a change of the signal's original spectral component. This change usually consists in decreasing of unwanted components of the input signal. The filtering process should not deform the signal, but there exists a group of filters which may introduce inadmissible deformations of the signal. The nonlinear filters belong to this group. For those reasons, the presented filters are evaluated using the mean square error (MSE) defined as $\mathrm{MSE} = \frac{1}{M}\sum_{i=1}^{M}(s_i - y_i)^2$, where: M - is the signal's length, y_i is the output of the evaluated filter, s_i is the deterministic part of signal, without a noise. The MSE factor allows to measure the mean square error of the additive residual distortions introduced by the nonlinear filtering. Signals y_i and s_i are aligned and they have the same time index.

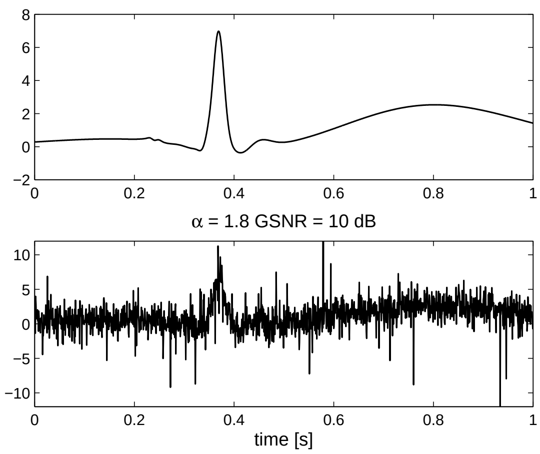

Fig. 2: An example of ECG signal corrupted with simulated impulsive noise modeled with SαS process for $\alpha = 1.8$ and GSNR$= 10$ dB (upper plot - clean signal, lower plot - noisy signal).

The results are obtained for one value of GSNR, i.e. 10 dB. For each value of α (from 1.7 to 2 with step 0.1 - this range corresponds to the impulsiveness of muscle noise [15]) and $\gamma = 1$, 100 different realizations of impulsive noise were generated with known value of GSNR. Afterwards the averaged MSE$_{ave}$ value is calculated as $\mathrm{MSE}_{ave} = \frac{1}{100}\sum_{i=1}^{100}\mathrm{MSE}_i$. Linear parameter K of the myriad filter is varying (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5) and the MSE factor is calculated for each value of K.

Two variants of the myriad filtering were analyzed:

1. the original myriad filter with the fixed-point method denoted as **FP** myriad filter, this is the reference filter,
2. the proposed 2nd order approximation myriad filter, denoted as **Appr** myriad filter.

The efficiency of the proposed method to filter ECG signals in impulsive environment was stated for the constant filter's window length and $N = 21$.

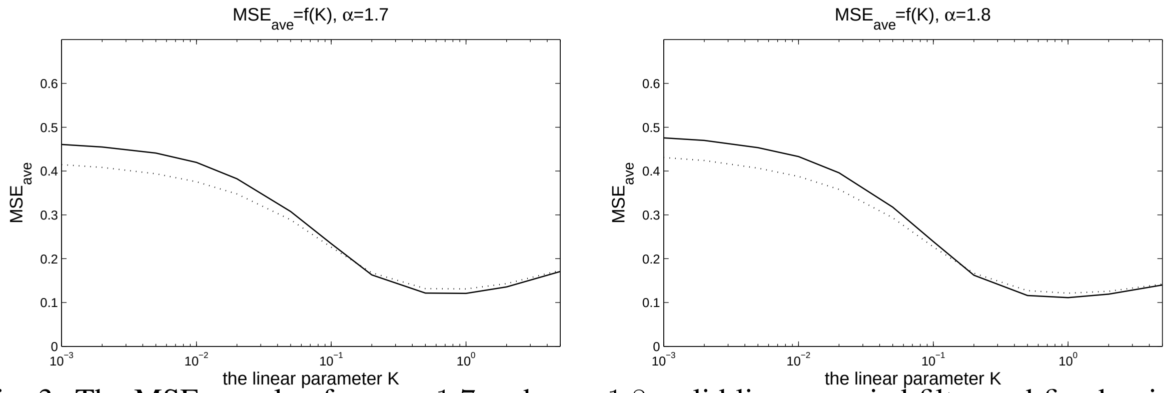

Fig. 3: The MSE$_{ave}$ value for $\alpha = 1.7$ and $\alpha = 1.8$, solid line – myriad filter and fixed-point method, dotted line – myriad filter and approximated method

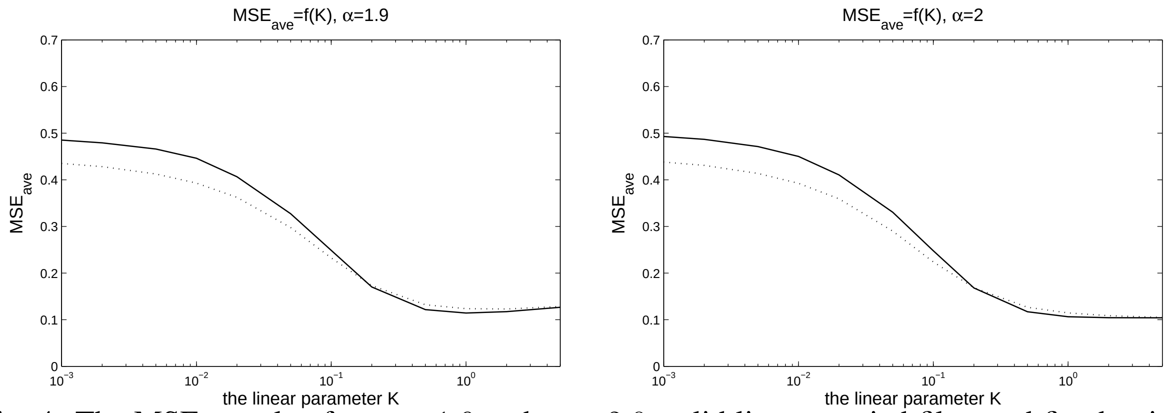

Fig. 4: The MSE$_{ave}$ value for $\alpha = 1.9$ and $\alpha = 2.0$, solid line – myriad filter and fixed-point method, dotted line – myriad filter and approximated method

4.1 Selection of the linear parameter K

The myriad filter operation is controlled by tunning the linearity parameter K. This parameter is necessary to proper work of the myriad filter. The idea of estimation method the K parameter comes from [6]. In order to make the estimation of myriad filter output more simpler, the following method of estimation the linear parameter K is proposed. This can be describe as follows:

1. for each sliding window of the myriad filter, find

$$K_n = \max_{i \neq j} |x_i - x_j| \,|_{i,j=1}^{N} \tag{13}$$

   where N - is the filter length.
2. then calculate K_{adapt} as:

$$K_{adapt} = \frac{1}{\min K_n \,|_{n=1}^{M}} \tag{14}$$

   where M - is the signal length.

Now, having K_{adapt} the myriad filter samples can be calculated according the proposed method. The effectiveness of this method is presented in the figure 5. In order to compare this method K_{opt} is calculated as $K_{opt} = \min \text{MSE}(K)$ where $K = (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5)$ and for α from the range $\langle 1, 2 \rangle$.

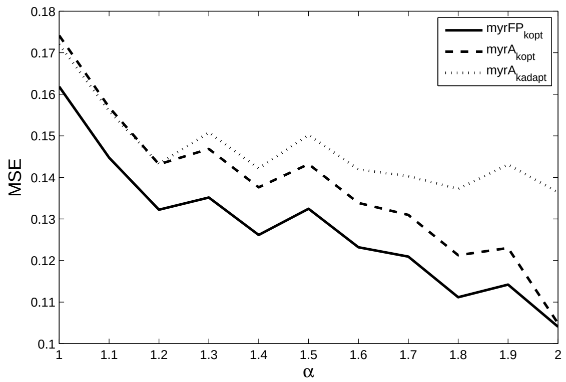

Fig. 5: Comparison of the proposed method of estimation the linear parameter K, solid line - myriad filter (FP) with K_{opt} parameter, dashed line - myriad filter (Appr) with K_{opt} parameter and dotted line - myriad filter (Appr) with K parameter calculated according eq. (13) and (14).

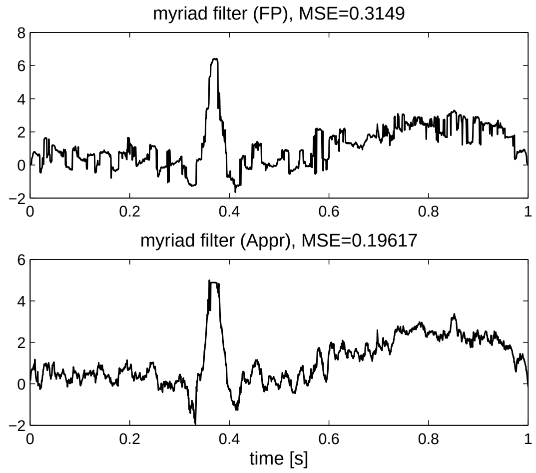

Fig. 6: The example of filtering signal with the FP myriad (upper plot) and the Appr myriad (lower plot) filters.

4.2 Results of filtering

The proposed algorithm for a myriad filter output estimation leads to obtaining the comparative or even better results with the reference myriad filter calculated with fixed-point method. For all values of α and K investigated filters lead to obtaining nearly the same value of MSE. When $K \leq 0.2$ the myriad filter Appr leads to obtaining the smallest value of MSE in comparison to the FP myriad filter. It can be seen in figures 3 and 4. In these cases a Appr filter shows the highest accuracy of filtering. When $K \geq 0.5$, then MSE values obtained for the Appr myriad filter are a little greater than MSE values obtained for the FP myriad filter.

The selection of K parameter leads to obtain comparative results when $1 \leq \alpha \leq 1.2$, but when $\alpha \leq 1.3$ the proposed method leads to obtain greater values of MSE than for K_{opt} estimations.

5 Conclusions

The new method of computation of the myriad filter output is proposed. This method is based on the least square approximation with the 2nd order polynomial. Knowing the coefficients of the approximation polynomial the top of parabola is calculated. Obtained coordinate x corresponds to the output of myriad filter in the running window of the length N. In order to prevent algorithm from outliers the filter window length is reduced. And two conditions are described which should be checked for correct use of the proposed computation method. This conditions allow to find an accurate value of the output of running window.

The experiments performed show that introduced filters are effective in suppression of muscle noise in ECG signal. The proposed polynomial approach to computation of the myriad filter output can effectively operate in wide range of the linear parameter K and the characteristic exponent α. The proposed method offers the comparative or better results of filtering signals in an impulsive environment than the reference myriad filter with the fixed-point searching method.

References

[1] Aysal TC, Barner KE. Meridian Filtering for Robust Signal Processing. IEEE Transactions on Signal Processing 2007; 55:3949–3962.

[2] Chan S, Zou Y. A Recursive least M-estimate algorithm for robust adaptive filtering in impulsive noise: fast algorithm and convergence performance analysis. IEEE Trans. on Signal Processing 2004; 52:975–991.

[3] Gonzalez JG, Arce GR. Optimality of the myriad filter in practical impulsive-noise environments. IEEE Trans. on Signal Processing 2001; 49:438–441.

[4] Gonzalez JG, Paredes JL, Arce GR. Zero-order statistics: a mathematical framework for the processing and characterization of very impulsive signals, IEEE Trans. on Signal Processing 2006; 54:3839–3851.

[5] Gonzalez JG, Lau DL, Arce GR. Towards a general theory of robust nonlinear filtering: selection filters. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97), Munich, Germany. 21-24 April 1997; 3837–3840.

[6] Gonzalez JG, Arce GR. Statistically-Efficient Filtering in Impulsive Environments: Weighted Myriad Filters. EURASIP Journal on Applied Signal Processing. 2002; 1:4–20.

[7] Hong X, Chen S. M-estimator and D-optimality model construction using orthogonal forward regression. IEEE Trans. on Systems, Man and Cybernetics 2005; 35:1–7.

[8] Heng Siong Lim, Teong Chee Chuah, Hean Teik Chuah, On the Optimal Alpha-*k* Curve of the Sample Myriad. IEEE Signal Processing Letters 2007; 14:545–548.

[9] Kalluri S, Arce GR. Robust frequency-selective filtering using weighted myriad filters admitting real-valued weights. IEEE Transactions on Signal Processing 2001; 49:2721–2733.

[10] Kalluri S, Arce GR. Adaptive weighted myriad filter algorithms for robust signal processing in α-stable noise environments. IEEE Transactions on Signal Processing 1998; 46:322–334.

[11] Kalluri S, Arce GR. Fast algorithms for weighted myriad computation by fixed point search. IEEE Trans. on Signal Processing 2000; 48:159-171.

[12] Li SZ. Robustizing robust M-estimation using deterministic annealing. Pattern Recognition 1996; 29:159–166.

[13] Łęski J. Robust Weighted Averaging. IEEE Trans. on Biomed. Eng. 2002; 8:796–804.

[14] Núñez RC, Gonzalez JG, Arce GR. Fast and Accurate Computation of the Myriad Filter via Branch-and-Bound Search. IEEE Transactions on Signal Processing 2008; 56:3340–3346.

[15] Pander T. An application of a weighted myriad filter to suppression an impulsive type of noise in biomedical signals. TASK Quartarly. 2004; 2:199–216.

[16] Pander T. New polynomial approach to myriad filter computation. Signal Processing 2010; 90:1991–2001.