robust filtering, myriad filters
neuro-fuzzy systems, fuzzy rules extraction

Robert CZABAŃSKI*, Tomasz PANDER*

PARAMETERS ESTIMATION FOR DIGITAL NON-LINEAR FITLERS USING NEURO-FUZZY SYSTEM

Abstract: The biomedical signals such as an electrical activity of the heart are commonly recorded with noise. One of the most difficult types of disturbances that need to be removed from ECG records is an electrical activity of muscles (EMG). This paper introduces the research on possibility of myriad filters application to suppress the EMG type of noise in ECG signals. The proper myriad filtering operation requires the knowledge about the statistical properties of the signal that allows choosing suitable values of the filter linearity parameter. Unfortunately, standard methods of an estimation of the myriad linearity parameter are not accurate enough to assure appropriate filtering quality. In this work, an application of Artificial Neural Network Based on Fuzzy Inference System is introduced to solve this task. To show usefulness of the proposed algorithm two numerical experiments are provided. The first one concerns filtering of an ECG signal corrupted with a simulated impulsive noise modelled by symmetric α-stable distributions. In the second experiment a real muscle noise is used.

1. INTRODUCTION

The biomedical signals are usually recorded with noise. Many different kind of biomedical signals exist, but for the purpose of this work an ECG signal was chosen. The ECG signal arises as a result of an electrical activity of the heart and it is almost always disturbed by a noise. The muscle noise (EMG) is the most difficult to suppress because its frequency spectrum agrees with the spectrum of ECG signal for a wide range of frequency and it shows frequently an impulsive nature. The white Gaussian noise is usually used to model the EMG disturbances, but this traditional model may fail. In this work the muscle noise is modelled using the α-stable distribution model. The impulsive noise requires robust methods that can suppress it [7, 10]. Therefore, non-linear filters can be applied. To a group of non-linear filters belongs a myriad filter for which the filtering performance can be controlled by the so called linearity parameter k which plays the fundamental role in the theory of myriad filters [6]. In this paper we applied a myriad filter to suppress an impulsive type of noise in ECG signal. The novelty of this non-linear filtering method consists in application of the Artificial Neural Network Based on Fuzzy Inference System (ANNBFIS) for estimation of the unknown parameter k to assure improvement of the filtering quality.

2. IMPULSIVE NOISE

* Silesian University of Technology, Institute of Electronics, Akademicka Str. 16, 44-100 Gliwice, Poland

The muscle noise may be modelled with the Symmetric α-Stable distribution (SαS) [8]. The class of SαS is characterized by their distribution using a characteristic function $\varphi(t) = \exp\left(jt\mu - \gamma|t|^{\alpha}\right)$, where α is the characteristic exponent restricted to the range $0 < \alpha \leq 2$, μ is the real-valued location parameter, γ is the dispersion of the distribution. The most important parameter of α-stable distributions is the characteristic exponent α, because it controls the heaviness of the distribution tails [10]. The knowledge of the parameters of the impulsive noise is required for the proper working of a myriad filter. For an estimation of the α-stable distribution several methods have been proposed, see [4, 10] for details. In this work, the method described in [4] is used. Let $Y = \log|X|$, where X denotes random variable of the α-stable distribution. The first and second moment of Y can be written as:

$$E(Y) = C_e\left(\frac{1}{\alpha} - 1\right) + \frac{1}{\alpha}\log(\gamma);\ Var(Y) = E\{[Y - E(Y)]^2\} = \frac{\pi^2}{6}\left(\frac{1}{\alpha^2} + \frac{1}{2}\right), \tag{1}$$

where: $C_e = 0.57721566\ldots$ is Euler constant, $E(\cdot)$ is the expected value. From (1) the characteristic exponent α and the dispersion γ are calculated.

3. THE NON-LINEAR MYRIAD FILTER

For a given set of observations $\{x_i\}_{i=1}^{N}$, where $x_i = \theta + n_i$ ($i = 1,\ldots,N$; θ is a location parameter and $\{n_i\}_{i=1}^{n}$ is a sequence of i.i.d. zero-mean noise components, N is the signal length), a maximum likelihood estimate (M-estimator) of location is given as [5]: $\hat{\theta} \triangleq \arg\min_{\theta} \sum_{i=1}^{N} \rho(x_i - \theta)$, where $\rho(\cdot)$ is called the cost function of the M-estimator and N is the filter length. If $\rho(x) = \log(k^2 + x^2)$ then the simple myriad filter is obtained [5, 6]. Using the fact that $\log(\cdot)$ is a strictly increasing function, the output of myriad filter can be determined from:

$$\hat{\theta}_k = \mathrm{myriad}\left(\mathrm{x}_1, x_2, \ldots, x_N; k\right) = \arg\min_{\theta} \sum_{i=1}^{N} \log\left[k^2 + (x_i - \theta)^2\right], \tag{2}$$

where k is a tuneable constant called the linearity parameter of the filter. The class of myriad filters includes a rich variety of filtering operations which can be controlled by simply adjusting the linearity parameter k. The case $k \to 0$ leads to highly robust selection filter called the weighted mode-myriad filter. The other case takes place, if $k \to \infty$ and $\alpha \to 2$, then the output of the myriad filter behaves like the output of moving average filter [5, 6]. We use an ANNBFIS for estimation of unknown values of the linearity parameter for myriad filters in the following considerations.

4. NEURO-FUZZY SYSTEM WITH PARAMETERIZED CONSEQUENTS

Artificial Neural Network Based on Fuzzy Inference System is a neuro-fuzzy system with parameterized consequents that generates inference results based on fuzzy if-then rules. In ANNBFIS fuzzy sets of rule antecedents have Gaussian membership function defined using two parameters: center $c_j^{(i)}$ and dispersion $s_j^{(i)}$. The linguistic connective "and" of multi-input rule predicates is represented by algebraic product t-norm. Consequents of ANNBFIS fuzzy rules have symmetric triangular membership functions. They can be defined using two parameters: width of the triangle base $w^{(i)}$ and center of gravity location determined by linear combinations of fuzzy system inputs: $y^{(i)}(\mathbf{x}_0) = \mathbf{p}^{(i)T}\mathbf{x}_0^{'}$, where $\mathbf{x}_0^{'} = [1, x_{01}, x_{02}, \ldots, x_{0t}]^T = [1, \mathbf{x}_0]^T$ is the extended input vector, $\mathbf{p}^{(i)} = [p_o^{(i)}, p_1^{(i)}, \ldots p_t^{(i)}]^T$ is a vector of parameters and t denotes a number of inputs. This dependency formulates so called parameterized (moving) consequent [2]. The neuro-fuzzy system with parameterized consequents allows both conjunctive and logical interpretations of fuzzy if-then rules. We assume conjunctive interpretation using Larsen's product in the following considerations. Assuming additionally a normalized arithmetic mean as an aggregation operator and modified indexed center of gravity [2] as a defuzzifier, we can evaluate the final crisp output value of the system from the following formula:

$$y_0 = \sum_{i=1}^{I} \frac{w^{(i)} F^{(i)}(\mathbf{x}_0)}{\sum_{j=1}^{I} w^{(j)} F^{(j)}(\mathbf{x}_0)} y^{(i)}(\mathbf{x}_0) = \sum_{i=1}^{I} G^{(i)}(\mathbf{x}_0) y^{(i)}(\mathbf{x}_0), \tag{3}$$

where I denotes a number of fuzzy if-then rules and $F^{(i)}(\mathbf{x}_0)$ is the firing strength of the i-th fuzzy rule. The fuzzy system with parameterized consequents can be treated as a radial basis function neural network [2]. Consequently, the unknown neuro-fuzzy system parameters can be estimated using learning algorithms of artificial neural networks [2, 3]. In this work, a hybrid learning procedure, which connects deterministic annealing and least square methods is presented [3]. For next considerations let us assume that we have N_u examples of input vectors $\mathbf{x}_0(n) \in \Re^t$ and the same number of known output values $t_0(n) \in \Re$ which formulate the training set. Our goal is the extraction of a set of fuzzy if-then rules that represents the knowledge of the phenomenon under consideration. The extraction process consists of an estimation of membership function parameters of antecedents as well as consequents. The number of rules I is also unknown. We assume that it is pre-set arbitrarily. The number of inputs t is defined by the size of input training vector directly.

To increase ability to avoid many local minima that traps steepest descent method used in original ANNBFIS learning algorithm, we employ the technique of deterministic annealing [9] adapted for the sake of learning the neuro-fuzzy system with parameterized consequents [3]. The equation (3) defines the neuro-fuzzy system as a mixture of experts (models). Its global output is expressed as a linear combination of I outputs $y^{(i)}(\mathbf{x}_0)$ of local models, each represented by a single fuzzy conditional statement. A randomness of the association between data and local models can be measured using the Shannon entropy S. In deterministic annealing method the objective is minimization of the cost function defined as a squared-error E while simultaneously controlling the entropy level of a solution. The deterministic annealing optimization problem is formulated as a minimization procedure of the Lagrangian $L = E - TS$, where T is the Lagrange multiplier [9]. A connection between the equation presented above and the annealing of solids is essential here. The quantity L can be identified as the Helmholtz free energy of physical system with "energy" E, entropy S and "temperature" T [9]. The procedure involves a series of iterations while the entropy level is reduced gradually. To allow achievement of the cost global optimum the simulated annealing method framework is used. The algorithm starts at a high level of pseudo-temperature T_{max} and tracks the solution for lowered values of T. The pseudo-temperature reduction procedure is determined by the annealing schedule function. We use the following decremental rule in the next considerations: $T \leftarrow qT$, where $q \in (0,1)$ is a pre-set parameter. At each level of temperature we minimize the Lagrangian iteratively using gradient descent method in L over the parameter space. In the original ANNBFIS learning method parameters of linear equations in consequents $\mathbf{p}^{(i)}$ are estimated using the least square (LS) method [2]. It accelerates the learning convergence [2]. For the same reason we use the LS algorithm in next considerations. The parameters $\mathbf{p}^{(i)}$ are adjusted using LS procedure and then tuned using the deterministic annealing algorithm. To avoid the matrix inverse operation the recurrent LS method [2] can be applied. For decreasing the computational burden of the learning procedure the deterministic annealing method with "freezing" phase (DAF) can be applied [3, 9]. The "freezing" phase consists of the calculation of $\mathbf{p}^{(i)}$ using LS procedure after every decreasing step of pseudo-temperature value while keeping $c_j^{(i)}$, $s_j^{(i)}$ as well as $w^{(i)}$ constant. Another problem is an estimation of initial values of membership functions for antecedents. It can be solved by means of preliminary clustering of the input training data [2].

5. NUMERICAL EXPERIMENTS

To validate the introduced method of an estimation of a linearity parameter for myriad filters two numerical experiments concerning filtering of ECG signals corrupted with noise were conducted. In the first, the simulated disturbances modelled by symmetric α-stable distributions are used ($\alpha \in [1.4, 2]$). In the second, a real muscle noise was applied. In order to evaluate a filtering performance the normalized mean absolute error (NMAE) was used $\text{NMAE} = \sum_{i=1}^{N} |y(i) - s(i)| / \sum_{i=1}^{N} |x(i) - s(i)| \cdot 100\%$, where: $y(i)$ is the output of the myriad filter, $s(i)$ is the deterministic part of signal without noise and $x(i)$ is the noisy signal. The noisy ECG cycles were obtained by adding a noise with five values (5, 10, 20, 30 [dB]) of Generalized Signal-to-Noise Ratio (GSNR) to a “clean” ECG cycle modelled by a linear combination of Hermite functions. The GSNR replaces the standard SNR, because a variance for α-stable noise does not exist. The GSNR is defined as: $\text{GSNR} = log_{10}\left[\sigma_s^2 / (a\gamma)\right]$ where: σ_s^2 is the variance of a “clean” signal, γ is the dispersion of an impulsive noise [4] and a is a scaling factor. As a reference we used: a moving average filter (MA), a median filter (MED) and a myriad filter with constant value of the linearity parameter k set to 1. The first objective was to create a learning set for the ANNBFIS neuro-fuzzy system. The application of four different parameters defining the impulsive noise was tested: a characteristic exponent α, a dispersion γ, a kurtosis K and a value of parameter defining the area under a curve of cumulative distribution of a signal $A = \sum_{i=1}^{N} F_X(x) / \sum_{i=1}^{N} x(i)x(i)$, where $F_X(x)$ is the cumulative distribution. We examined the following definitions of input vectors: (i) $\mathbf{x}_0 = [\alpha, \gamma]^T$; (ii) $\mathbf{x}_0 = [\alpha, \gamma, K]^T$; (iii) $\mathbf{x}_0 = [\alpha, \gamma, A]^T$; and (iv) $\mathbf{x}_0 = [\alpha, \gamma, K, A]^T$. Generally the best filtering results (the lowest values of NMAE error) were obtained if all four parameters were employed in the learning process. As an output t_0 we used the optimal values of linear parameter k_{opt} calculated on the basis of a comparison of a signal without noise, a disturbed signal and a signal after filtering process. The optimal linearity parameter was defined as an argument for which the normalized mean square error function $\mathrm{NMSE} = \sum_{i=1}^{N}[y(i)-s(i)]^2 / \sum_{i=1}^{N}[s(i)]^2$ reached the minimum, i.e. $k_{opt} = \min_{k} \mathrm{NMSE}(k)$. To get reasonable learning results, we set $t_0 = -1$ for a very large values of k_{opt} (k_{opt} >1000) for which the myriad filter behaves like the moving average filter. The learning of the ANNBFIS system was carried out for the number of if-then rules I changed form 2 to 6. For the deterministic annealing procedure the following parameters' values were applied: $T_{max} = \{10^{-3}, 10^{-2}, \dots, 10^{3}\}$, $T_{min} = 10^{-5} T_{max}$, $q = 0.95$. The initial values of the learning step for steepest descent procedure η_{ini} were changed in the range from 0.01 to 0.10 with step 0.01. The initial values of antecedents parameters were estimated using FCM [1] clustering results. In the first numerical experiment we used the data set which consists of N_u = 2200 learning pairs (an input vector and an output value), with parameters defining the impulsive noise calculated for a single ECG cycle corrupted with the artificial noise. In the filtering process we applied a set of if-then rules for which the lowest value of the learning error was achieved. Values of the learning parameters (η_{ini}, T_{max}) that led to the best learning quality are shown in Table 1. To test the filtering quality with ANNBFIS system we used noisy ECG signals which were not present during the learning phase. If the output value of ANNBFIS was negative then the myriad filter was replaced with the moving average filter. The obtained results (NMAE index values) are tabulated in Table 1.

Table 1. Results of filtering of ECG signals corrupted with the artificial noise (GSNR = 10dB; N = 21)

α	MA	MED	k_{opt}	k = 1	I = 2 (0.08,10^3)	I = 3 (0.04,10^3)	I = 4 (0.07,10^3)	**I = 5 (0.09,10^3)**	I = 6 (0.08,10^2)
1.4	39.3144	16.8898	15.3770	26.4440	16.8265	16.9035	17.1557	**16.6248**	16.7366
1.5	36.0899	19.0976	16.9918	26.7446	18.3767	18.7428	19.3897	**18.0572**	18.6419
1.6	31.7250	21.6381	18.9836	26.7620	20.5439	20.6203	20.6347	**20.2619**	20.6304
1.7	28.9597	23.5308	20.2122	25.9391	23.1324	22.8896	22.7643	**22.2388**	23.0031
1.8	27.1148	25.4678	21.5024	25.0588	23.9395	23.4304	23.556	**23.2550**	23.8289
1.9	25.3288	27.0498	22.5279	24.2956	24.2501	23.9697	24.0551	**23.9993**	24.2947
2.0	23.8336	28.4619	22.8925	23.0969	24.041	23.8193	23.9526	**24.0846**	24.1794

The obtained results confirm that the application of the ANNBFIS neuro-system for estimation of the myriad filter parameter leads to improvement of the filtering quality in comparison to the linear MA and the nonlinear MED filter. Only for α = 2 we did not have the decrease of the NMAE index in comparison to the myriad filter with constant linearity parameter value. For α = 2 application of the MA filter leads to good filtering quality also. It is due to the Gaussian characteristic of the noise. Generally, the best filtering results were obtained for I = 5 if-then rules, η_{ini} = 0.09 and T_{max} = 1000. In the second numerical experiment we used a data set which consists of N_u = 800 learning pairs with values of parameters defining the impulsive noise calculated for a single ECG cycle corrupted with real muscle disturbance samples. The specification of the learning algorithm was defined the same. Again, to test the filtering quality we used noisy ECG signals which were not present during the learning phase. Similar to the previous example, if the output value of ANNBFIS was negative then the myriad filter was replaced with the moving average filter. The NMAE values of filtering together with values of the learning parameters (η_{ini}, T_{max}) that led to the best learning quality are shown in Table 2. For small values of GSNR the MA filter and myriad filters lead to smallest values of NMAE. If the GSNR level is high then a model of noise characterized by a Laplacian distribution should be assumed for muscle samples. Therefore, the best filtering results are obtained for MED filter which is optimized for such distribution.

Table 2. Results of filtering of ECG signals corrupted with the real noise (N = 21)

GSNR [dB]	MA	MED	k_{opt}	$K = 1$	$I = 2$ (0.01,10^{-2})	$I = 3$ (0.09,10^{-1})	$I = 4$ (0.08,10^{-2})	$I = 5$ (0.09,10^{-1})	**$I = 6$ (0.09,10^{0})**
5	48.0590	54.1424	48.7218	49.1515	48.9696	48.6602	48.7301	48.5281	**48.5753**
10	50.5882	54.6228	49.6216	50.4933	51.0340	50.7071	50.5829	50.6469	**50.4418**
20	74.0827	60.5868	58.5727	64.9127	67.0474	61.9575	61.4167	61.5438	**61.7357**
30	128.4960	71.6625	78.8215	97.6757	101.0326	81.7831	81.6493	81.1815	**81.1762**

Nevertheless, the application of ANNBFIS neuro-system for estimation of the myriad filter parameter allows filtering performance improvement of the ECG signals corrupted with the real muscle noise in comparison to the myriad filter with constant k value. The proposed filtering method leads to satisfactory filtering quality for low as well as high GSNR level. Generally, the best filtering results were obtained for I = 6 if-then rules, η_{in} = 0.09 and T_{max} = 1.

6. CONCLUSIONS

In this paper the research on possibility of myriad filters application to suppress the EMG type of noise in ECG signals was presented. To find the unknown values of the linearity parameter for myriad filters the Artificial Neural Network Based on Fuzzy Inference System was applied. In the proposed learning algorithm of the neuro-fuzzy system, parameters of fuzzy sets from antecedents and consequents of fuzzy if-then rules were adjusted separately by means of the deterministic procedure and the least squares method respectively. Experimentation shows the usefulness of the myriad filtering of ECG signal corrupted with the simulated impulsive disturbances as well as the real muscle noise, when the linearity parameter is estimated using the ANNBFIS.

BIBLIOGRAPHY

[1] BEZDEK J.C., Pattern recognition with fuzzy objective function algorithms. Plenum, New York, 1982.
[2] CZOGAŁA E., ŁĘSKI J., Fuzzy and neuro-fuzzy intelligent systems. Springer-Verlag, Heidelberg, 2000.
[3] CZABAŃSKI R., Automatic fuzzy if-then rules extraction from numerical data, Ph.D. Theses, Gliwice, 2003.
[4] GEORGIOU P.G., TSAKALIDES P., KYRIAKAKIS Ch., Alpha-stable modeling of noise and robust time-delay estimation in the presence of impulsive noise. IEEE Trans. on Multimedia, Vol.1, pp. 291-301, 1999.
[5] GONZALEZ J.G., ARCE G.R.: Optimality of the myriad filter in practical impulsive-noise environments, IEEE Trans. on Signal Processing, Vol. 49, No. 2, pp. 438-441, 2001.
[6] KALLURI S., ARCE G.R., Adaptive weighted myriad filter algorithms for robust signal processing in α-stable noise environments. IEEE Transactions on Signal Processing, Vol. 46, pp. 322-334, 1998.
[7] LEE Y.H., KASSAM S.A., Generalized median filtering and related nonlinear filtering techniques. IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. 33, pp. 672-683, 1985.
[8] PANDER T., Application of weighted myriad filters to suppress impulsive noise in biomedical signals, TASK Quarterly, Vol. 8, No. 2, pp. 199-216, 2004.
[9] RAO A.V., MILLER D., ROSE K., GERSHO A., A deterministic annealing approach for parsimonious design of piecewise regression models. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 2, pp. 159-173, 1999.
[10] SHAO M., NIKIAS Ch.L., Signal processing with fractional lower order moments: stable processes and their applications. Proceedings of IEEE, Vol. 81, pp. 986-1009, 1993.