Neural Network Aided Adaptive Minimum L_1-Norm Filter for EP Estimation

Guo Wenqiang[1,2], Qiu Tianshuang[1], and Li Fan[2]

1. Dalian University of Technology, Dalian , China , 116024
2. Xinjiang Institute of Finance and Economics, Urumchi 830012,China

Tel: +86-0411-84706009-3038

Summary:

Evoked potentials (EPs) have been widely used to quantify neurological system properties. Traditional EP analyses are developed under the condition that the background noises in EP are Gaussian distributed. Alpha stable distribution, a generalization of Gaussian, is better for modeling impulsive noises than Gaussian distribution in biomedical signal processing. This paper presents a new minimum *L1*-norm filter algorithm for EP estimation based on neural network. Simulation results show that the new algorithm is very effective under both Gaussian and lower order alpha stable noise condition and is closer to optimal fashion in EP signal estimation compared with traditional methods.

Keywords:

Evoked potentials (EPs), neural network, L_1-norm optimization, alpha stable distribution.

1. Introduction

As the biomedical signals generated by neural system, EP is the special regular reaction of the neural system to voice, light and electronic impulse in outside. It includes such abundant information about each part of neutral transportation channel that the conditions and change of neurological system can be reflected. The checking and analytic technology of EP is an important measure in practical medical to diagnose injury or pathological changes in neural system[1][2]. In this technology, the signal variation caused by physical pathological changes is used to diagnose or analysis the neural system illness and injury. EP signal can be classified into several category according to different stimulations outside: somatosensory evoked potentials (SEP), auditory evoked potentials (AEP), and visual evoked potentials (VEP).

Inevitably, the checking process of EP check is infected by electroencephalogram (EEG) signals which is a random signals generated in neural system. The EEG which is accompanied with the EP signals is regarded as noise and the signal-to-noise ratio(SNR) is less than 0dB generally. The measure extracted EP are used to include coherent average, weighted average, matching filter and parameter model etc. It is equal to low-pass filter for noise that to coherent average or coherent weighted average with the EP signals acquired. But the average measure always result in losing real EP information. The new EP extraction measure involve adaptive filter, wavelet transform, neural network, Higher-order cumulants, independent component analysis(ICA)，unscented Kalman filter etc.[3][4][5]. The algorithms based on second-order statistics are popular to gain the EP signals because EEG mixed in EP signals is always assumed to be Gaussian distribution. Recently, some study showed that the EEG have the non-Gaussian character which can result in performance degradation of conventional methods based on second-order statistics. Therefore, it is necessary to develop novel analysis method of EP in impulsive noise[6][7].

In many conventional signal analysis and handling measures which based on second-order linear theory, the system noise are always be assured as the Gaussian noise have finite second-order statistics. Nevertheless, the practical application in underwater acoustic, radar, communications and medical signal processing fields, many random signals and noises are non-Gaussian distributions, such as ocean circumstance noises, the instant peak in circuit lines, atmosphere noises, speech signals and biomedical signals and many kinds of noises man made, in which exist dramatic peak noises. The Non-Gaussian distribution signal and the Gaussian distribution signal are illustration in figure 1.

Alpha stable distributions[6][7][8] is a sort of

Gaussian distribution in generalized. The conception for alpha stable distributions is presented by Levy firstly in 1925 when he studied generalized center limit theory (GCLT). The statistic characters for Alpha stable distributions are determined by four parameters in the characters function. If a random variable is alpha stable distribution if and only if its characteristic function has the form [8] as

$$\Phi(t) = \exp\left\{j\mu t - \gamma|t|^{\alpha}\left[1 + j\beta\,\mathrm{sgn}(t)\varpi(t,\alpha)\right]\right\} \qquad (1)$$

where, $\omega(t,\alpha) = \tan\frac{\alpha\pi}{2}(\alpha \neq 1)$ or $\frac{2}{\pi}\log|t|(\alpha = 1)$, $-\infty < \mu < \infty$, $\gamma > 0$, $0 < \alpha \le 2$, $-1 \le \beta \le 1$, α is the characteristic exponent, it controls the thickness of the tail in the distribution. The Gaussian process is a special case of stable processes with $\alpha = 2$. The dispersion parameter γ is similar to the variance of Gaussian process and β is the symmetry parameter. If $\beta = 0$, the distribution is symmetric and the observation is referred to as the $S\alpha S$ (symmetry α -stable) distribution, i.e., it is symmetrical about μ. μ is the location parameter. When $\alpha = 2$ and $\beta = 0$, the stable distribution becomes the Gaussian distribution and $\gamma = \sigma^2/2$ .If the character index of the random signals in stable distribution is alpha only the statistic whose rank less than alpha are infinite[7][8].

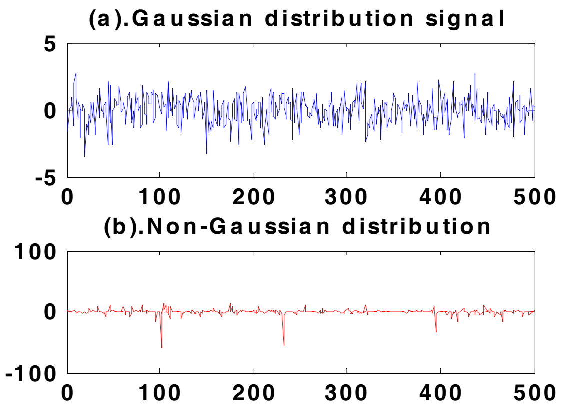

Fig.1. The stable distributions samples

In this paper, we present a new neural network to solve the L_1-norm optimization problem in alpha stable distribution environments, and use the Neural Network Aided Adaptive Minimum L_1-Norm Filter for EP estimation. The performance of this novel adaptive EP estimation scheme is achieved to outperform that of the former algorithms that use the least-squares (L_2-norm) model. Simulation and real data analysis showed that this new method is effective in tracking EP variations across trials and allows fast EP measurement in many time-critical circumstances.

The remainder of this paper is organized as follows. The L_1-norm optimization model and its neural network implementation are detailedly introduced in Section II. The experiences and results analysis is given in Section III, and the conclusions are given in Section IV.

2. L_1-Norm Optimization Model and Its Neural Network Implementation

Due to its excellent properties, the L_1 -norm optimization model has been extensively studied. However, with the increase of the model scale, these numerical algorithms are not adequate for solving real-time problems. One possible and promising approach to real-time optimization is to apply neural networks. Because of the inherent massive parallelism, the neural network-based approach can solve optimization problems within a time constant of the network.

2.1 Minimum L_1-norm Optimization Model

In fact, many models can be mathematically abstracted as the following over-determined system of linear equations[9]:

$$\mathrm{x} = \mathrm{As} - \mathrm{e} \qquad (2)$$

where $\mathbf{A} = \{a_{ij}\} \in R^{M\times N}\,(M > N)$ is the model matrix derived from a given data set, $\mathbf{s} = [s_1, s_2, ..., s_N]^{\mathrm{T}} \in R^N$ is the unknown vector of the parameters to be estimated, $\mathbf{x} = [x_1, x_2, ..., x_M]^{\mathrm{T}} \in R^M$ is the vector of observation or measurements containing errors or artifacts, $\mathrm{e} \in R^M$ is the alpha stable distribution error or noise vector. Define the L_1-norm of error vector as follows:

$$\|\mathrm{e}\|_1 = \|\mathrm{As} - \mathrm{x}\|_1 \qquad (3)$$

Then, the parameter vector $\mathbf{s}$ can be solved via solving the following unconstrained optimization model:

$$\mathbf{s}_{\text{opt}} = \min_{\mathbf{s}} \| \mathbf{As} - \mathbf{x} \|_1 \tag{4}$$

This model is called L_1-norm optimization model, which is generally difficult to be solved because of discontinuous derivatives. Using the following Proposition 1, we turn the problem described in (4) into another form, which is easier to be solved.

Proposition 1*:* The optimization model described in (4) is equivalent to the following optimization model:

$$\mathbf{s}_{\text{opt}} = \min_{\mathbf{s}} \left\{ \max_{\mathbf{y}} \left(\mathbf{y}^{\mathrm{T}} (\mathbf{As} - \mathbf{x}) \right) \right\} \tag{5}$$

where $\mathbf{y} = [y_1, y_2, ..., y_M]^{\mathrm{T}} \in R^M, |\, y_i \,| \le 1, i = 1,2,...,M$.

Proof: Let $\mathbf{u} = (\mathbf{As} - \mathbf{x}) \in R^M$, then for any $\mathbf{y}$, we have

$$\mathbf{y}^{\mathrm{T}} (\mathbf{As} - \mathbf{x}) = \mathbf{y}^{\mathrm{T}} \mathbf{u}$$

$$= \sum_{i=1}^{M} y_i u_i \le \sum_{i=1}^{M} |\, y_i \,||\, u_i \,| \le \sum_{i=1}^{M} |\, u_i \,|$$

$$= \| \mathbf{As} - \mathbf{x} \|_1 \tag{6}$$

Thus

$$\max_{\mathbf{y}} \left(\mathbf{y}^{\mathrm{T}} (\mathbf{As} - \mathbf{x}) \right) = \| \mathbf{As} - \mathbf{x} \|_1 \tag{7}$$

The proof of Proposition 1 is completed. □

2.2 The L1-norm Neural Network Implementation

Now we propose a neural network for solving the problem in (5) whose model is described by the following dynamic system:

$$\begin{cases} \dfrac{d\mathbf{s}}{dt} = -\mathbf{A}^T \mathbf{P}(\mathbf{y} + \mathbf{As} - \mathbf{x}) \\ \dfrac{d\mathbf{y}}{dt} = -(\mathbf{AA}^T + \mathbf{I})\mathbf{y} + \mathbf{P}(\mathbf{y} + \mathbf{As} - \mathbf{x}) \end{cases} \tag{8}$$

Where $\mathbf{P}(\mathbf{v}) = [P(v_1), P(v_2), ..., P(v_M)]^T \in R^M$, $P(v_i)$ is defined as a projection operator :

$$P(v_i) = \frac{1}{2}(|\, v_i + 1 \,| - |\, v_i - 1 \,|) = \begin{cases} 1 & if \quad v_i > 1 \\ v_i & otherwise \\ -1 & if \quad v_i < -1 \end{cases} \tag{9}$$

The proposed L_1-norm neural network (L_1N-NN) described in (8) is shown in figure 2.

Since the neural network described in (8) is a continuous-time network governed by a set of ordinary differential equations, It can be real-time implemented. In such implementation, the projection operator of $P(v_i)$ is actually a simple limiter with a unit threshold. The matrix or vector multiplications are actually the synaptic-weighting and summing operations and hence can be implemented via a number of adders with a weighting function [10]. And the rest are a number of simple integrators.

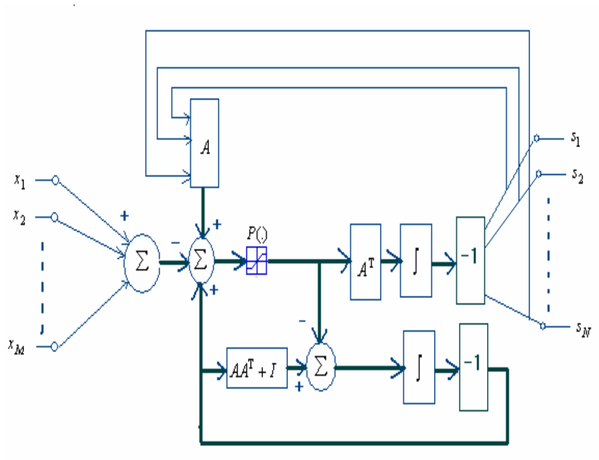

Fig.2. The L_1-norm neural network

In the following, we will prove that the neural network described in (8) and figure 2 globally converges to the exact solution to problem (5), or equivalently to problem (4).

Let $L(\mathbf{s}, \mathbf{y}) = \mathbf{y}^{\mathrm{T}} (\mathbf{As} - \mathbf{x})$, according to K-T theorem[7], If $\mathbf{s}^* \in R^N$ is a solution to the problem in (5), we know that $(\mathbf{s}^*, \mathbf{y}^*)$ is a solution if and only if there exists a saddle point of model (5), and $L(\mathbf{s}^*, \mathbf{y}) \le L(\mathbf{s}^*, \mathbf{y}^*) \le L(\mathbf{s}, \mathbf{y}^*)$. Thus we can easily have

$$(\mathbf{y}\text{-}\mathbf{y}^*)^{\mathrm{T}} (\mathbf{As}^* - \mathbf{x}) \le 0 \tag{10}$$

$$(\mathbf{y}^*)^{\mathrm{T}} (\mathbf{As}^* - \mathbf{x}) \le (\mathbf{y}^*)^{\mathrm{T}} (\mathbf{As} - \mathbf{x}) \tag{11}$$

Then there exists $\mathbf{y}^*$ satisfying

$$\begin{cases} \mathbf{A}^T \mathbf{y}^* = 0 \\ \mathbf{y}^* = \mathbf{P}(\mathbf{y}^* + \mathbf{A}\mathbf{s}^* - \mathbf{x}) \end{cases} \quad (12)$$

For any $\mathbf{y}$, the following inequality holds:

$$\{\mathbf{P}(\mathbf{y} + \mathbf{A}\mathbf{s} - \mathbf{x}) - \mathbf{y}^*\}^T (\mathbf{x} - \mathbf{A}\mathbf{s}^*) \geq 0 \quad (13)$$

The solution set of

$$\begin{cases} \dfrac{d\mathbf{s}}{dt} = -\mathbf{A}^T \mathbf{P}(\mathbf{y} + \mathbf{A}\mathbf{s} - \mathbf{x}) = 0 \\ \dfrac{d\mathbf{y}}{dt} = -(\mathbf{A}\mathbf{A}^T + \mathbf{I})\mathbf{y} + \mathbf{P}(\mathbf{y} + \mathbf{A}\mathbf{s} - \mathbf{x}) = 0 \end{cases} \quad (14)$$

is just the equilibrium point set of dynamic system (8).

Let $\mathbf{E}_1 = \mathbf{A}^T \mathbf{y}$, $\mathbf{E}_2 = \mathbf{y} - \mathbf{P}(\mathbf{y} + \mathbf{A}\mathbf{s} - \mathbf{x})$, then

$$\begin{cases} \dfrac{d\mathbf{s}}{dt} = -\mathbf{E}_1 + \mathbf{A}^T \mathbf{E}_2 \\ \dfrac{d\mathbf{y}}{dt} = -\mathbf{A}\mathbf{E}_1 - \mathbf{E}_2 \end{cases} \quad (15)$$

Let $(\mathbf{s}^*, \mathbf{y}^*)$ denote the solution set of model (5). By (12), we can get that when $(\mathbf{s}, \mathbf{y}) = (\mathbf{s}^*, \mathbf{y}^*)$, $\dfrac{d\mathbf{s}}{dt} = \mathbf{0}$ and $\dfrac{d\mathbf{y}}{dt} = \mathbf{0}$. Thus we can give the relationship between the solution set of model (5) and the equilibrium point set of dynamic system (8).

3. Numerical Testing Results

To test the performance of our method, two examples are given in this section. The simulations were carried out in Matlab6.5 program on a Pentium IV PC with 1.2-GHz CPU and 256-MB memory. The first example is used to compare convergence speed and approximation accuracy to the nonlinear function. Because the comparison of EP estimation performance between minimum L_1-norm and minimum L_2-norm has been widely made, here we just place emphasis on the difference between L_1-norm and L_2-norm filtering. In the second example, we give the correlation coefficient based on L_1-norm and based on L_2-norm.

3.1 Experiment one

The two signals passed through an unknown admixture matrix are a segment of evoked potential signal (the sample frequency for 1000 Hz) and a simulated impulsive EEG signal, traditional method and new methods of this paper are adopted respectively, to extract the admixture signal. The relative mean square error (MSE) was used to measure the effectiveness of the method. The MSE is defined as

$$\mathrm{MSE} = \frac{\mathrm{E}[(\mathrm{s} - \mathrm{y})^2]}{\mathrm{E}[\mathrm{s}^2]}$$

The convergence curve is depicted in figure 3.

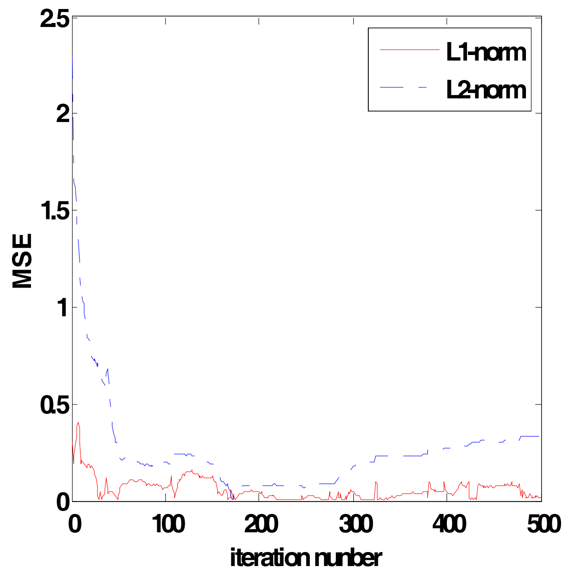

Fig.3. Solutions mix the square error margin curve diagram of matrix

When demixture matrix converge, make use of the solution of demixture matrix to carry on the EP signal separation extraction, the result is shown in figure 4. Among them (a) is pure EP signal; (b) is a mixture signal of the EP signal and the EEG signal; (c) for according to the extraction EP signal of the traditional method; (d) for according to the extraction EP signal of this paper method.

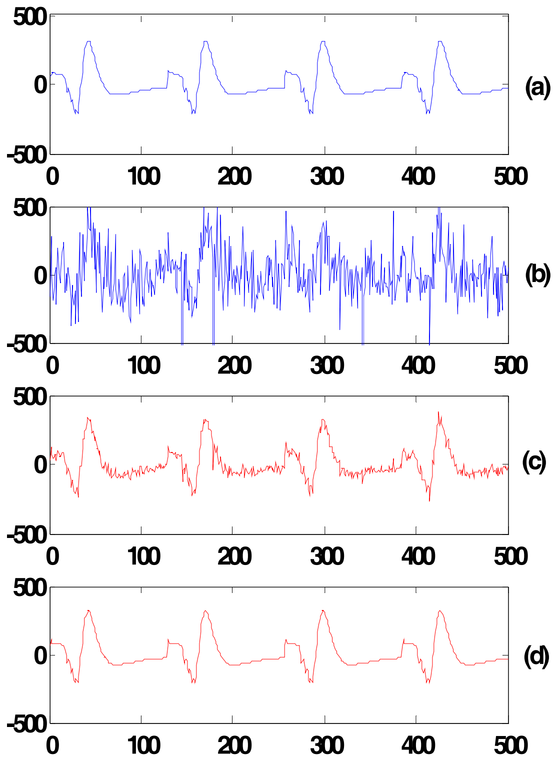

Fig. 4. EPs and the extraction result of the pulse EEG

Compared with the conventional method, the EP signals extracted method in this paper has more perfect stability and convergence as well as more good separated results.

3.2 Experiment two

Pass a segment of evoked potential signal (period for 128 points, the sample frequency for 1000 Hz) and impulsive EEG signal through the unknown matrix admixture. In order to evaluate the extracting results quantificationally, we take the correlation coefficient of extract signals and resource signals as inspection index, where

$$\tau(s_i, y_j) = \frac{\left|\sum_{n=1}^{N} s_i(n) y_j(n)\right|}{\sqrt{\sum_{n=1}^{N} s_i^2(n) \sum_{n=1}^{N} y_j^2(n)}}.$$

Extract EP signals by using the demixture matrix generated in different iterative times by using two kinds of method respectively, and get the correlation coefficient as figure 5, and table 1 show.

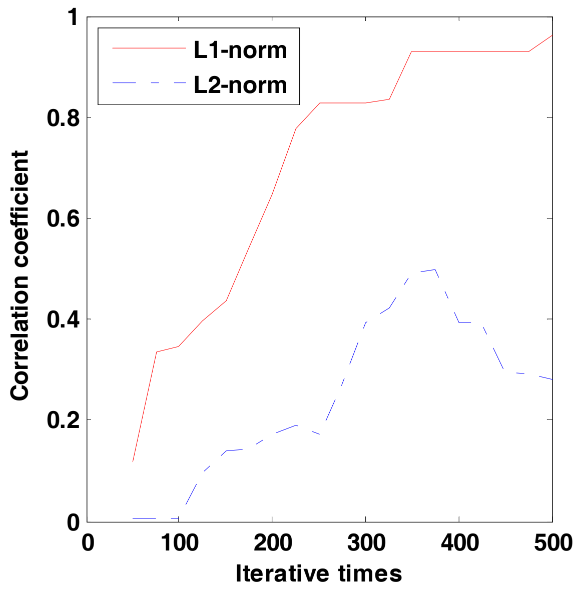

Fig.5. Correlation coefficient and iterative times of EP signals

Table 1. The EP signal correlation coefficient compare Base on different iterative times

Iterative times	EP correlation coefficient (L_1-norm) EP	EP correlation coefficient (L_2-norm) EP
50	0.1264	0.0044
100	-0.3450	-0.0050
150	0.4378	0.1378
200	0.6469	0.1716
250	-0.8291	-0.1711
300	-0.8290	-0.3937
350	-0.9293	-0.4893
400	0.9295	0.3945
450	0.9299	0.2941
500	-0.9612	-0.2816

It can be seen that the correlation coefficient based on based on L_1-norm is clearly bigger than that based on based on L_2-norm.

4. Conclusion

Alpha stable distribution can be used to describe the random signals and noises more perfect in practice, if the signals or noises have the prominent impulse character. Not as other statistic model, alpha stable distribution do not have close identical density function and their second-order statistic is not existed. This paper introduces the statistic character for stable distribution simply, and then presents a novel algorithm based on minimum L_1-norm which is suitable for EP separation. The simulation experiments and theoretical analysis show that this algorithm is a good extraction method of EP signal with better robustness under the condition of Gaussian and fraction low-order alpha stable distribution noise，especially suitable for EP signals blind extraction under circumstance of stable distribution pulse EEG noise and performance good in convergence and separation results.

Acknowledgement

This work was supported by the National Natural Science Foundation of China(60372081, 30570475, 30170259), the Ph.D. Programs Foundation of Ministry of Education of China(No. 20050141025).The authors would like to think the anonymous reviewers for their detailed comments and questions which improved the quality of the presentation of this paper.

References

[1] R.R.Gharieb, A.Cichocki, "Noise reduction in brain evoked potentials based on third-order correlations," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 5, pp. 501-512, 2001.

[2] C.E.Davila, R.Srebro, and I.A.Ghaleb, "Optimal detection of visual evoked potential," *IEEE Transactions on Biomedical Engineeri*ng, vol. 45, no.6, pp. 800-803, 1998.

[3] R.Mutihac, Van Hulle, "PCA and ICA neural implementations for source separation: a comparative study," *Proceedings of the International Joint Conference on Neural Networks*, vol. 1 , pp.20-24, 2003.

[4] J.Julier, K.Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401-422, 2004.

[5] Z.Ronghui, W.Jianwei, "Neural Network-Aided Adaptive Unscented Kalman Filter for Nonlinear State Estimation," *IEEE Signal Processing Letters*, vol. PP, no. 99,pp.1-4, 2006.

[6] N.Hazarika, A.C.Tsoi, A.A.Sergejew, "Nonlinear considerations in EEG signal classification," *IEEE Transactions Signal Processing*, vol. 45, no. 4, pp. 829-836, 1997.

[7] X.Ma, C.L.Nikias, "Joint estimation of time delay and frequency delay in impulsive noise using fractional lower-order statistics," *IEEE Transactions Signal Processing*, vol. 44, no. 11, pp.2669-2687, 1996.

[8] P. G. Georgiou, P. Tsakalides, and C. Kyriakakis, "Alpha-Stable Modeling of Noise and Robust Time-Delay Estimation in the Presence of Impulsive Noise," IEEE Transactions on Multimedia, vol. 1, no. 3, pp. 291–301, September 1999.

[9] C. A. Zala, I. Barrodale, and J. S. Kennedy, "High-resolution signal and noise field estimation using the L1 (least absolute values) norm," IEEE J. Oceanic Eng., vol. OE-12, pp. 253–264, 1987.

[10] Y. S. Xia and J. Wang, "A general methodology for designing globally convergent optimization neural networks," IEEE Trans. Neural Networks, vol. 9, pp. 1331–1343, 1998.

Guo Wenqiang: received the B.S. degree and the M.S. degree in electrical engineering from Dalian University of Technology, Dalian, China, in 1998, and in 2003. He served on the Xinjiang Institute of Finance and Economics, since 1998 through 2000.He is currently a Ph.D candidate in Dalian University of Technology. His research interests include biomedical signal processing.

Qiu Tianshuang: received the B.S. degree from Tianjin University, Tianjin, China, in 1983, the M.S. degree from Dalian University of Technology, Dalian, China, in 1993, and the Ph.D. degree from Southeastern University, Nanjing, China, in 1996, all in electrical engineering. He worked as a post doctoral fellow in the Department of Electrical Engineering at Northern Illinois University, USA, since 1996 through 2000. He is currently a professor in Dalian University of Technology. His research interests include adaptive digital signal processing, biomedical signal processing, and non-Gaussian signal processing.