Higher-Order-Statistics-Based Radial Basis Function Networks for Signal Enhancement

Bor-Shyh Lin, *Member, IEEE*, Bor-Shing Lin, Fok-Ching Chong, and Feipei Lai, *Senior Member, IEEE*

Abstract—In this paper, a higher-order-statistics (HOS)-based radial basis function (RBF) network for signal enhancement is introduced. In the proposed scheme, higher order cumulants of the reference signal were used as the input of HOS-based RBF. An HOS-based supervised learning algorithm, with mean square error obtained from higher order cumulants of the desired input and the system output as the learning criterion, was used to adapt weights. The motivation is that the HOS can effectively suppress Gaussian and symmetrically distributed non-Gaussian noise. The influence of a Gaussian noise on the input of HOS-based RBF and the HOS-based learning algorithm can be mitigated. Simulated results indicate that HOS-based RBF can provide better performance for signal enhancement under different noise levels, and its performance is insensitive to the selection of learning rates. Moreover, the efficiency of HOS-based RBF under the nonstationary Gaussian noise is stable.

Index Terms—Gaussian noise, higher order statistics (HOS), radial basis function (RBF) networks, signal enhancement.

I. INTRODUCTION

ADAPTIVE filtering techniques have been widely used for signal enhancement [1]–[3]. Signal enhancement can be considered as a mapping from the noisy input space to the noise-free output space. Different from parametric model estimation, they do not need prior information about the signal of interest (SOI) and noise to enhance signals effectively. Therefore, they are easily implemented to model the noise-free systems, in particular, when the complexities of SOI and noise increase. The original scheme of adaptive signal enhancement was proposed by Widrow *et al.* [1]. Here, an adaptive linear filter was used for signal enhancement. However, in practice, lots of SOIs are nonlinear. To effectively approximate nonlinear signals, a great number of filter taps are required. Thus, the linear filter is not relevant to enhance nonlinear signals.

In order to effectively enhance nonlinear signals, neural network filters become an alternative [4]–[9]. Basing on multilayer perceptron (MLP), artificial neural networks need far fewer filter weights and a small amount of training data to approximate any continuous nonlinear function. Watterson used MLP neural network to detect signal in band-limited Gaussian noise [4]. Gandhi and Ramamurti employed the MLP neural network for signal

detecting in non-Gaussian noise [7]. However, it is well known that the learning of the MLP is slow and global. Therefore, the training set for the MLP should contain a relatively complete coverage of input space, and it is just suitable for those conditions which do not have time restrictions. Moreover, the parameters of the MLP are highly nonlinear, and their estimates may become trapped at a local minimum during the learning procedure.

Radial basis function (RBF) network is another popular neural network. Cha and Kassam used RBF network for interference cancellation [8]. Billings and Fung used recurrent RBF networks for noise cancellation [9]. The design of RBF was first derived by Broomhead and Lowe [10]. It can be regarded as a special three-layer network. The only hidden layer performs a nonlinear transformation which maps the input space onto a new space. Its mechanism is equivalent to finding a surface in a multidimensional space that provides a best fit to training data. Park and Sandberg proved that RBF is a universal approximation [11], [12].

In 1989, Moody and Darken proposed a hybrid learning process for RBF [13]. This learning procedure separately uses different learning algorithms to train centers and weights in RBF. The positions and width of centers in Gaussian basis functions are determined by self-organized learning procedure [14], [15]. The weights between the hidden nodes and the output nodes are trained by supervised learning algorithm, such as a least mean square (LMS) algorithm [3], [16]. In 1990, Poggio and Girosi proposed a supervised approach which used gradient-descent method to adjust all centers and weights [17]. Despite the existence of other learning schemes, in practice, the hybrid learning process proposed by Moody and Darken is most frequently used to adapt RBF. Because of the strategy of learning desired input–output mapping, the hybrid learning scheme can achieve fast training. However, additional noise on the reference signal will influence the clustering of input space, in particular, when the signal-to-noise ratio (SNR) is poor. Further additional noise in the desired signal also directly influences the adaptation of weights by using LMS. Therefore, how to reduce the influence of additional noise on the learning procedure is important for improving the performance of RBF.

Recently, higher order statistics (HOS) techniques have been used for signal enhancement [18]–[24]. HOS have a natural tolerance to Gaussian and symmetrically distributed non-Gaussian noise [25], [26]. Therefore, HOS are useful for suppressing Gaussian noise and detection of asymmetrically distributed non-Gaussian noise. In this paper, a HOS-based RBF network is proposed and applied for signal enhancement. Higher order cumulants of the reference signal were used as the

B.-S. Lin, B.-S. Lin, and F.-C. Chong are with the Institute of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: borshyhlin@ntu.edu.tw; d88921027@ntu.edu.tw; fcchonh@cc.ee.ntu.edu.tw).

F. Lai is with the Department of Computer Science and Information Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: flai@ntu.edu.tw).
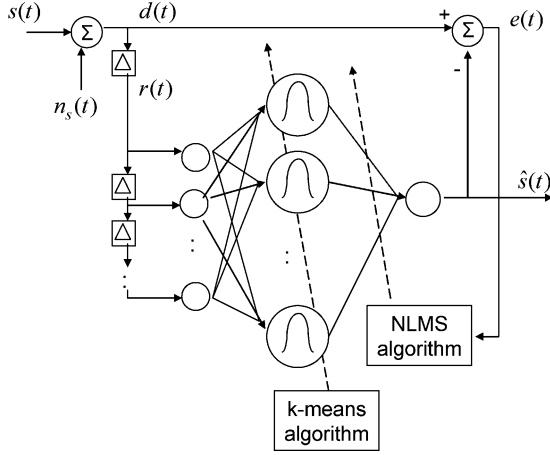
Fig. 1.   RBF networks for signal enhancement.

input of HOS-based RBF to reduce the influence of additional noise on clustering of the input space, such that an HOS-based supervised learning algorithm for weights is developed. The mean square error obtained from HOS of the desired signal and the system output was used as the learning criterion. Therefore, the influence of additional noise on adaptation of weights can be effectively reduced.

This paper is organized as follows. RBF network for signal enhancement is introduced in Section II. In Section III, the definition and several properties of HOS are introduced. HOS-based RBF is also introduced here. Sequentially, the implement of HOS-based RBF for the third-order statistics is illustrated. In Section IV, simulated results and discussions for the efficiencies of signal enhancement of RBF and HOS-based RBF are provided. Finally, conclusions are drawn in Section V.

II. RBF NETWORKS FOR SIGNAL ENHANCEMENT

Consider the measured signal $d(t)$ of the primary sensor, satisfying

$$d(t) = s(t) + n_s(t) \tag{1}$$

where $s(t)$ and $n_s(t)$, respectively, denote SOI and uncorrelated Gaussian noise at iteration t. The estimate $\hat{s}(t)$ of $s(t)$ can be obtained by using the scheme of signal enhancement derived by Widrow *et al.* [1], [3], shown in Fig. 1. In this scheme, RBF can be treated as a noise-free model. The delayed version of $d(t)$ is commonly used as the reference signal $r(t)$, i.e.,

$$r(t) = d(t - \Delta). \tag{2}$$

In general, the unit of sampling period is used as the prediction depth Δ [2], [3]. Other ways of selecting prediction depth Δ were discussed in [2].

Considering an RBF which has N_0 input nodes, N_1 hidden nodes, and a single output node, the estimate $\hat{s}(t)$ can be obtained by

$$\hat{s}(t) = \boldsymbol{\delta}(t)^T \mathbf{w}(t) \tag{3}$$

where $\mathbf{w}(t) = [w_1(t), w_2(t), \ldots, w_{N_1}(t)]^T$ denotes the weight vector connecting hidden nodes with the output node and $\boldsymbol{\delta}(t) = [\delta_1(t), \delta_2(t), \ldots, \delta_{N_1}(t)]^T$ is the output vector of hidden nodes. The output of Gaussian basis function in the kth hidden node can be defined by

$$\delta_k(t) = \exp\left(-\frac{\|\mathbf{r}(t) - \mathbf{v_k}(t)\|^2}{2\rho^2(t)}\right) \tag{4}$$

where $\|\cdot\|$ denotes the Euclidean norm, $\mathbf{v_k}(t)$ denotes the center in the kth hidden node, and $\mathbf{r}(t) = [r(t), r(t-1), \ldots, r(t - N_0 + 1)]^T$ is the input vector of RBF. Here, $\rho(t)$ denotes the width of centers and can be defined as the variance of the input vector [3], i.e.,

$$\rho^2(t) = \frac{1}{N_0}\left(\sum_{i=0}^{N_0-1} r(t-i) - \bar{\mathbf{r}}(t)\right)^2 \tag{5}$$

where $\bar{\mathbf{r}}(t)$ denotes the mean of $\mathbf{r}(t)$.

The center $\mathbf{v_k}(t)$ can be obtained by self-organized learning procedure, such as k-means clustering algorithm [14], [15], which is given in (6), shown at the bottom of the page.

Here, μ_v is the learning rate for centers. Because the reference vector $\mathbf{r}(t)$ contains $n_s(t)$, it is straightforward that the centers obtained by k-mean algorithm are directly influenced by $n_s(t)$, in particular, when the SNR of the input vector is poor.

The weight vector $\mathbf{w}(t)$ is commonly adapted by LMS algorithm [3]. In LMS algorithm, the mean square output error $d(t) - \hat{s}(t)$ is defined as the learning criterion. The normalized LMS algorithm is given by

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \frac{\mu_w}{\left(1 + \boldsymbol{\delta}^T(t)\boldsymbol{\delta}(t)\right)}\boldsymbol{\delta}(t)\left(d(t) - \hat{s}(t)\right) \tag{7}$$

where μ_w denotes the learning rate for weights. However, the system output noise $e(t)$ is also correlated with $n_s(t)$, i.e.,

$$e(t) = d(t) - \hat{s}(t) = s(t) + n_s(t) - \hat{s}(t). \tag{8}$$

From (8), we found that (7) will be affected directly by $n_s(t)$. Moreover, $n_s(t)$ will also be amplified when μ_w increases. It can be viewed as a gradient noise amplification problem. Thus, using LMS algorithm to adapt weights may become inefficient under poor SNR.

III. HOS-BASED RBF NETWORKS

A. Basic Scheme of HOS-Based RBF Network

Higher order cumulants are defined as the coefficients in the Taylor series expansion of the joint characteristic

$$\begin{cases} \mathbf{v_k}(t+1) = \mathbf{v_k}(t) + \mu_v\left[\mathbf{x}(t) - \mathbf{v_k}(t)\right], \\ \mathbf{v_i}(t+1) = \mathbf{v_i}(t), \end{cases} \quad \begin{aligned} &k = \arg\min\left[\|\mathbf{x}(t) - \mathbf{v_j}(t)\|, 1 \le j \le N_1\right] \\ &1 \le i \le N_1 \quad \text{and} \quad i \ne k \end{aligned} . \tag{6}$$
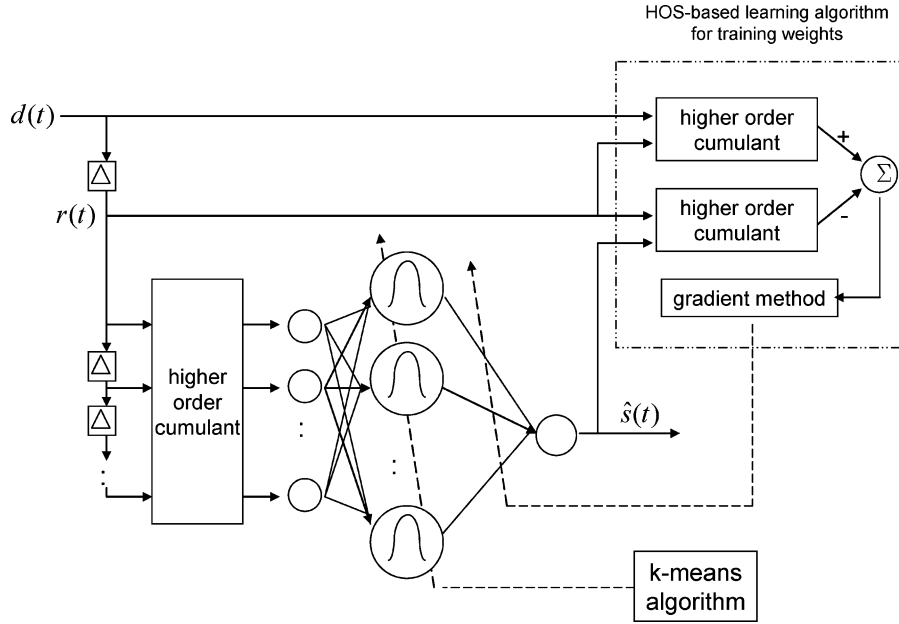
Fig. 2. Basic scheme of HOS-based RBF networks.

function of a random process [26]. For a set of n real variables $\{z_1, z_2, \ldots, z_n\}$, their joint cumulant of order $k = k_1 + k_2 + \ldots + k_n$ can be defined by

$$\text{Cumulant}\left[z_1^{k_1}, z_2^{k_2}, \ldots, z_n^{k_n}\right]$$
$$\equiv (-j)^k \left.\frac{\partial^k \psi(w_1, w_2, \ldots, w_n)}{\partial z_1^{k_1} \partial z_2^{k_2} \ldots \partial z_n^{k_n}}\right|_{w_1=w_2=\ldots=w_n=0} \quad (9)$$

where $\psi(w_1, w_2, \ldots, w_n) \equiv E[j(w_1 z_1 + w_2 z_2 + \ldots + w_n z_n)]$ is their joint characteristic function and $E[\cdot]$ denotes the expectation operation.

For a set of n real stationary variables $\{z_i(t)\}, i = 1, 2, \ldots, n$, the nth-order cross cumulant of $\{z_i(t)\}$ can be represented by

$$C_{z_1 z_2 \ldots z_n}(\tau_1, \tau_2, \ldots, \tau_{n-1})$$
$$\equiv \text{Cumulant}\left[z_1(t), z_2(t+\tau_1), z_3(t+\tau_2), \ldots, z_n(t+\tau_{n-1})\right]. \quad (10)$$

If the random variables $\{z_i(t)\}, i = 1, 2, \ldots, n$ can be divided into any two or more statistically independent groups, then $C_{z_1 z_2 \ldots z_n}(\tau_1, \tau_2, \ldots, \tau_{n-1}) = 0$. If $\{z_i(t)\}$ is a Gaussian process, then all cumulants of $\{z_i(t)\}$, whose order $n > 2$, are identical to zero [25], [26].

Since $s(t)$ and $n_s(t)$ are independent, under the assumption that there exists nth-order cumulant of the reference signal and it is not identically zero, it can be represented by

$$C_{rr \ldots r}(\tau_1, \tau_2, \ldots, \tau_{n-1}) = C_{ss \ldots s}(\tau_1, \tau_2, \ldots, \tau_{n-1})$$
$$+ C_{n_s n_s \ldots n_s}(\tau_1, \tau_2, \ldots, \tau_{n-1}). \quad (11)$$

If $n_s(t)$ is a Gaussian noise, then $C_{n_s n_s \ldots n_s}(\tau_1, \tau_2, \ldots, \tau_{n-1}) = 0$. Therefore, higher order cumulant of the reference signal can effectively close to that of SOI. By this property, the basic idea behind HOS-based RBF, shown in Fig. 2, is that the higher order cumulant of the reference

signal is used as the input of HOS-based RBF to reduce the influence of $n_s(t)$ on clustering. The output of the kth hidden node in HOS-based RBF can be represented by

$$\delta_k(t) = \exp\left(-\frac{\|\mathbf{C_{rr \ldots r}}(t) - \mathbf{v_k}(t)\|^2}{2\rho(t)^2}\right) \quad (12)$$

where $\mathbf{C_{rr \ldots r}}(t)$ denotes the matrix of higher order cumulants of $\mathbf{r}(t)$. Here, the width $\rho(t)$ of centers is defined as the variance of $\mathbf{C_{rr \ldots r}}(t)$.

B. HOS-Based Learning Algorithm for Weights

In the HOS-based learning algorithm, the mean square error of the two higher order cumulants of $d(t)$ and $\hat{s}(t)$ is used as the learning criterion ξ to reduce the influence of $n_s(t)$ on the adaptation of weights

$$\xi = \sum_{\tau_1} \sum_{\tau_2} \cdots$$
$$\sum_{\tau_{n-1}} \frac{1}{2}\left[C_{\hat{s}rr \ldots r}(\tau_1, \tau_2, \ldots, \tau_{n-1})\right.$$
$$\left. - C_{drr \ldots r}(\tau_1, \tau_2, \ldots, \tau_{n-1})\right]^2$$
$$= \sum \sum_{(\tau_1, \tau_2, \ldots, \tau_{n-1}) \in \Gamma} \cdots$$
$$\sum \frac{1}{2}\left[C_{\hat{s}ss \ldots s}(\tau_1 - \Delta, \tau_2 - \Delta, \ldots, \tau_{n-1} - \Delta)\right.$$
$$\left. - C_{sss \ldots s}(\tau_1 - \Delta, \tau_2 - \Delta, \ldots, \tau_{n-1} - \Delta)\right]^2$$
$$= \sum \sum_{(\tau_1, \tau_2, \ldots, \tau_{n-1}) \in \Gamma} \cdots$$
$$\sum \frac{1}{2}\left[\sum_{j=1}^{N_1} w_j C_{\delta_j rr \ldots r}(\tau_1, \tau_2, \ldots, \tau_{n-1})\right.$$
$$\left. - C_{drr \ldots r}(\tau_1, \tau_2, \ldots, \tau_{n-1})\right]^2. \quad (13)$$

TABLE I
SUMMARY OF LEARNING PROCEDURE OF HOS-BASED RBF WITH THE THIRD-ORDER STATISTICS

Definition:

N_0: the number of input nodes, N_1: the number of hidden nodes, λ: forgetting factor, μ_w: learning rate for weights, μ_v: learning rate for centers, $\mathbf{w}(t)$: weight vector, $\mathbf{v_k}(t)$: center of k-th hidden node, ρ: the width of centers

1. Initial conditions:

$\hat{C}_{r_i rr}(0;\tau_1,\tau_2) = 0,$ $\tau_1 = -m_1, (-m_1+1),...,(m_1-1),m_1,$ $\tau_2 = 1,2...,m_2,$ $i = 1,2,...,N_0$

$\hat{C}_{\delta_i rr}(0;\tau_1,\tau_2) = 0,$ $\tau_1 = -m_1, (-m_1+1),...,(m_1-1),m_1,$ $\tau_2 = 1,2...,m_2,$ $i = 1,2,...,N_1$

$\hat{C}_{drr}(0;\tau_1,\tau_2) = 0,$ $\tau_1 = -m_1, (-m_1+1),...,(m_1-1),m_1,$ $\tau_2 = 1,2...,m_2$

Randomly set $\mathbf{v_k}(0),$ $k = 1,2,...,N_1$

2. Computation, $t = 1,2,...$:

$$\hat{s}(t) = [\delta_1(t),\delta_2(t),...,\delta_{N_1}(t)]\mathbf{w}(t), \quad \text{where} \quad \delta_k(t) = \exp(-\frac{\|\mathbf{C_{rrr}}(t) - \mathbf{v_k}(t)\|^2}{2\rho^2})$$

2.1 Update centers:

$$\begin{cases} \mathbf{v_k}(t+1) = \mathbf{v_k}(t) + \mu_v[\mathbf{C_{rrr}}(t) - \mathbf{v_k}(t)] & , k = \arg\min[\|\mathbf{C_{rrr}}(t) - \mathbf{v_j}(t)\|, 1 \le j \le N_1] \\ \mathbf{v_i}(t+1) = \mathbf{v_i}(t) & , 1 \le i \le N_1 \text{ and } i \ne k \end{cases}$$

2.2 Estimate third order cumulants:

For $\tau_1 = -m_1, (-m_1+1),...,(m_1-1),m_1,$ $\tau_2 = 1,2...,m_2$

$\hat{C}_{drr}(t;\tau_1,\tau_2) = \lambda\hat{C}_{drr}(t-1;\tau_1,\tau_2) + (1-\lambda)d(t)r(t+\tau_1)r(t+\tau_1+\tau_2)$

For $\tau_1 = -m_1, (-m_1+1),...,(m_1-1),m_1,$ $\tau_2 = 1,2...,m_2,$ $i = 1,2,...,N_0$

$\hat{C}_{r_i rr}(t;\tau_1,\tau_2) = \lambda\hat{C}_{r_i rr}(t-1;\tau_1,\tau_2) + (1-\lambda)r(t-i+1)r(t+\tau_1)r(t+\tau_1+\tau_2)$

For $\tau_1 = -m_1, (-m_1+1),...,(m_1-1),m_1,$ $\tau_2 = 1,2...,m_2,$ $i = 1,2,...,N_1$

$\hat{C}_{\delta_i rr}(t;\tau_1,\tau_2) = \lambda\hat{C}_{\delta_i rr}(t-1;\tau_1,\tau_2) + (1-\lambda)\delta_i(t)r(t+\tau_1)r(t+\tau_1+\tau_2)$

2.3 Construct $\mathbf{C_{drr}}(t)$, $\mathbf{C_{\delta rr}}(t)$, and $\mathbf{C_{rrr}}(t)$:

$\mathbf{C_{drr}}(t) = [\hat{C}_{drr}(t;-m_1,0), \hat{C}_{drr}(t;-m_1,1),..., \hat{C}_{drr}(t;m_1,m_2-1), \hat{C}_{drr}(t;m_1,m_2)]^T$

$$\mathbf{C_{rrr}}(t) = \begin{bmatrix} \hat{C}_{r_1 rr}(t;-m_1,0) & \hat{C}_{r_1 rr}(t;-m_1,1) & ... & \hat{C}_{r_1 rr}(t;m_1,m_2) \\ \hat{C}_{r_2 rr}(t;-m_1,0) & \hat{C}_{r_2 rr}(t;-m_1,1) & & \hat{C}_{r_2 rr}(t;-m_1,m_2) \\ \vdots & & \cdot & \vdots \\ \hat{C}_{r_{N_0} rr}(t;-m_1,0) & \hat{C}_{r_{N_0} rr}(t;-m_1,1) & ... & \hat{C}_{r_{N_0} rr}(t;m_1,m_2) \end{bmatrix}^T$$

$$\mathbf{C_{\delta rr}}(t) = \begin{bmatrix} \hat{C}_{\delta_1 rr}(t;-m_1,0) & \hat{C}_{\delta_1 rr}(t;-m_1,1) & ... & \hat{C}_{\delta_1 rr}(t;m_1,m_2) \\ \hat{C}_{\delta_2 rr}(t;-m_1,0) & \hat{C}_{\delta_2 rr}(t;-m_1,1) & & \hat{C}_{\delta_2 rr}(t;-m_1,m_2) \\ \vdots & & \cdot & \vdots \\ \hat{C}_{\delta_{N_1} rr}(t;-m_1,0) & \hat{C}_{\delta_{N_1} rr}(t;-m_1,1) & ... & \hat{C}_{\delta_{N_1} rr}(t;m_1,m_2) \end{bmatrix}^T$$

2.4 Update weights:

$$\nabla_{\mathbf{w}}(t) \equiv \frac{\partial\xi}{\partial\mathbf{w}(t)} = -\mathbf{C}_{\delta rr}^T(t)[\mathbf{C_{drr}}(t) - \mathbf{C_{\delta rr}}(t)\mathbf{w}(t)]$$

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \frac{\mu_w}{tr(\mathbf{C}_{\delta rr}^T(t)\mathbf{C_{\delta rr}}(t))}\nabla_{\mathbf{w}}(t)$$

Equation (13) can be rewritten in a matrix form

$$\xi = \frac{1}{2}[\mathbf{C_{\delta rr...r}w} - \mathbf{C_{drr...r}}]^2. \tag{14}$$

Here, $\mathbf{C_{\delta rr...r}}$ and $\mathbf{C_{drr...r}}$ are, respectively, a $N_\Gamma \times N_1$ matrix and a $N_\Gamma \times 1$ column vector, and N_Γ denotes the number of

points in the set Γ. In order to minimize ξ, the gradient-descent method was used. The gradient of ξ is given by

$$\nabla_{\mathbf{w}}(t) \equiv \frac{\partial\xi}{\partial\mathbf{w}(t)} = -\mathbf{C}_{\delta rr...r}^T[\mathbf{C_{drr...r}} - \mathbf{C_{\delta rr...r}w}]. \tag{15}$$
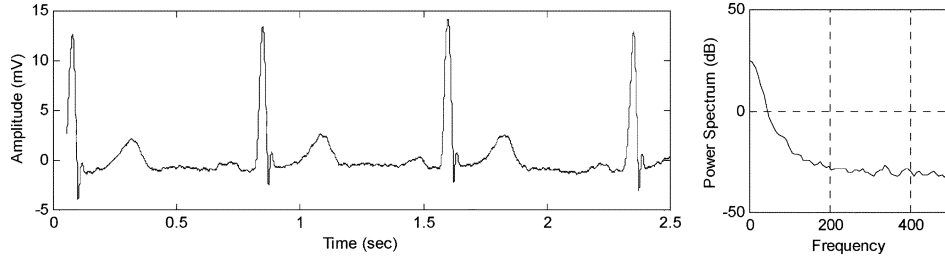
Fig. 3. Pattern of ECG in MIT/BIH database and its power spectrum.

Finally, the adaptation formula of HOS-based learning algorithm can be given by

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \frac{\mu_w}{\left(1 + tr\left(\mathbf{C}_{\boldsymbol{\delta}\mathbf{rr}...\mathbf{r}}^T \mathbf{C}_{\boldsymbol{\delta}\mathbf{rr}...\mathbf{r}}\right)\right)} \nabla_{\mathbf{w}}(t). \quad (16)$$

From (13), under the assumption of that $s(t)$ approximates to $\hat{s}(t)$, the learning criterion will be close to optimal. Moreover, from (13), the influence of $n_s(t)$ on $\mathbf{C}_{\mathbf{drr}...\mathbf{r}}$ can be eliminated. Therefore, in HOS-based learning algorithm, the gradient noise amplification problem can be effectively reduced.

C. Implementation of HOS-Based RBF Network

In this paper, HOS-based RBF with the third-order statistics was used for the signal enhancement. To implement HOS-based RBF, the estimate of the third-order cumulant [26] can be recursively computed by

$$\begin{aligned}
\hat{C}_{z_1 z_2 z_3}(t; \tau_1, \tau_2) = {} & \langle z_1(t) z_2(t + \tau_1) z_3(t + \tau_2) \rangle \\
& - \langle z_1(t) \rangle \langle z_2(t + \tau_1) z_3(t + \tau_2) \rangle \\
& - \langle z_2(t + \tau_1) \rangle \langle z_1(t) z_3(t + \tau_2) \rangle \\
& - \langle z_3(t + \tau_2) \rangle \langle z_1(t) z_2(t + \tau_1) \rangle \\
& + 2 \langle z_1(t) \rangle \langle z_2(t + \tau_1) \rangle \langle z_3(t + \tau_2) \rangle. \ (17)
\end{aligned}$$

If $\{z_i(t)\}$ is a zero-mean signal, the estimate of the third-order cumulants can be simplified to

$$\hat{C}_{z_1 z_2 z_3}(t; \tau_1, \tau_2) = \langle z_1(t) z_2(t + \tau_1) z_3(t + \tau_2) \rangle. \quad (18)$$

Here, the operation $\langle \cdot \rangle$ is defined as

$$\langle f(t) \rangle = \lambda \langle f(t-1) \rangle + (1 - \lambda) f(t) \quad (19)$$

where λ is a forgetting factor. The implementation of the HOS-based RBF with the third-order statistics is summarized in Table I.

IV. RESULTS AND DISCUSSION

A. For Stationary Gaussian Noises

In this section, several simulations are considered to compare the performance of the RBF with the normalized LMS (RBF-NLMS) algorithm [3] and HOS-based RBF with the third-order statistics for the signal enhancement under Gaussian noise. The electrocardiogram (ECG) pattern at the Massachusetts Institute of Technology/Boston's Beth Israel Hospital (MIT/BIH) database, shown in Fig. 3, was used as SOI. Its power spectrum distributes mainly in less than 150 Hz. To generate simulated trials,

the colored Gaussian noise distributed from 90 to 135 Hz was superimposed on SOI.

Fig. 4 is the result of the signal enhancement obtained by RBF-NLMS ($N_0 = 20$, $N_1 = 30$, $\mu_w = 0.05$, and $\mu_v = 0.5$) and HOS-RBF ($N_0 = 20$, $N_1 = 30$, $\mu_w = 0.05$, $\mu_v = 0.5$, $m_1 = 10$, $m_2 = 3$, and $\lambda = 0.9$) under the SNR of -10 dB. In this paper, the definition of SNR is given by

$$\text{SNR} = 20 \log_{10} \left(\frac{A_s}{A_n} \right) \quad (20)$$

where A_s and A_n, respectively, denote the magnitudes of SOI and noise. Fig. 4(a) is the noisy simulated trial and its power spectrum. It shows that the additional Gaussian noise heavily overlapped SOI. However, as we expected, Fig. 4(b) and (c) clearly shows that the HOS-based RBF can effectively enhance SOI under colored Gaussian noise. In this case, RBF-NLMS provides about 20-dB reduction for Gaussian noise, whereas HOS-based RBF can provide about 30-dB reduction. Obviously, the performance for signal enhancement can be effectively improved by HOS techniques.

To investigate the effect of using the higher order cumulants of the reference signal as the input of HOS-based RBF, comparison of performance for various methods with different learning rates μ_v was carried out. Here, the average of mean square error of 50 randomly selected results was used to present performance. As shown in Fig. 5, under a fixed learning rate μ_w, HOS-based RBF can provide the better performance. Moreover, the performance of HOS-based RBF is more insensitive to the selection of learning rate μ_v. This result indicates that using higher order cumulants of the reference signal as the input of HOS-based RBF can provide more stable clustering of training data under poor SNR.

Next, the effect of using the HOS-based learning algorithm for weights on signal enhancement was investigated. Fig. 6 shows the comparison of the performance of RBF-NLMS and HOS-based RBF with different learning rate μ_w. Under a fixed learning rate μ_v, it shows that HOS-based RBF can provide the better performance for signal enhancement, and its performance obviously is more insensitive to μ_w. The simulated result fits our expectation that the gradient noise amplification problem can be effectively improved by using HOS-based learning algorithm. The influence of additional noise on the LMS algorithm increases, in particular, when the learning rate μ_w is large. Although a small learning rate μ_w on both RBF-NLMS and HOS-based RBF can provide more stable performance, it also may cause heavy distortion of the estimate
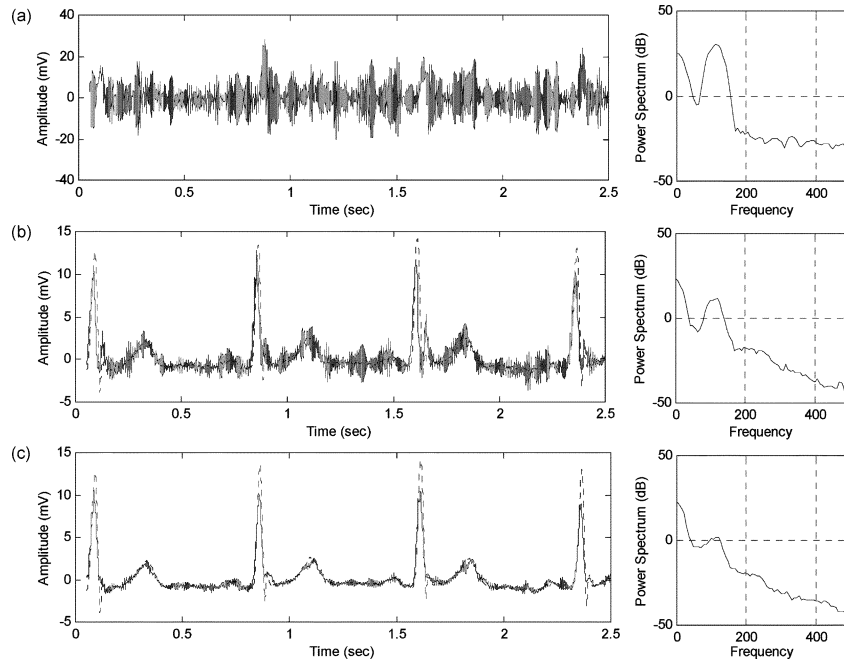
Fig. 4.   (a) Noisy simulated trial with SNR of -10 dB. (b) Result obtained by RBF-NLMS. (c) Result obtained by HOS-based RBF.
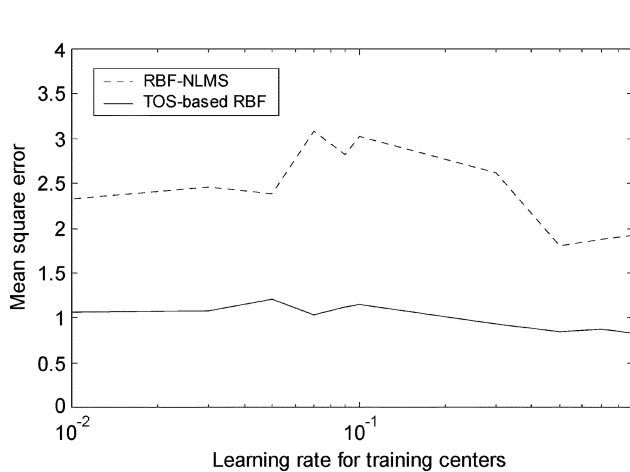

Fig. 5.   Comparison of performance of RBF-NLMS and HOS-based RBF with different learning rate μ_v.

Fig. 6.   Comparison of the performances of RBF-NLMS and HOS-based RBF with different learning rate μ_w.

of SOI due to the slow adaptation. Therefore, the performances of both RBF-NLMS and HOS-based RBF become poor when the learning rate μ_w is less than 0.05.

To investigate the efficiency of RBF-NLMS and HOS-based RBF in different noise levels, several simulated trials with different SNRs from -2.5 to -25 dB were generated. Fig. 7 is the comparison of the performances between RBF-NLMS and HOS-based RBF under different noise levels. It shows that HOS-based RBF provides a better performance against different noise levels. The performance of RBF rapidly becomes very poor, in particular, when the SNR is less than -10 dB. Although the performance of HOS-based RBF is also influenced by the variation of SNR of input signals, it is more insensitive than that of RBF-NLMS.
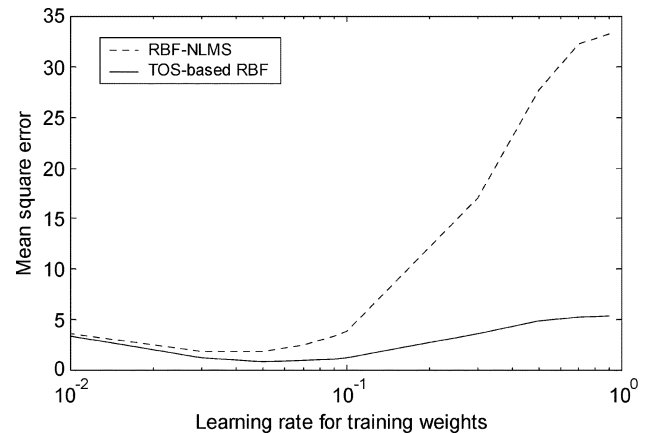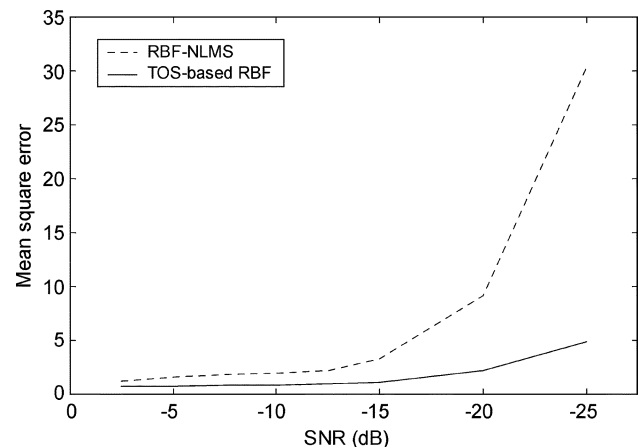

Fig. 7.   Comparison of the performance of various methods under different noise levels.
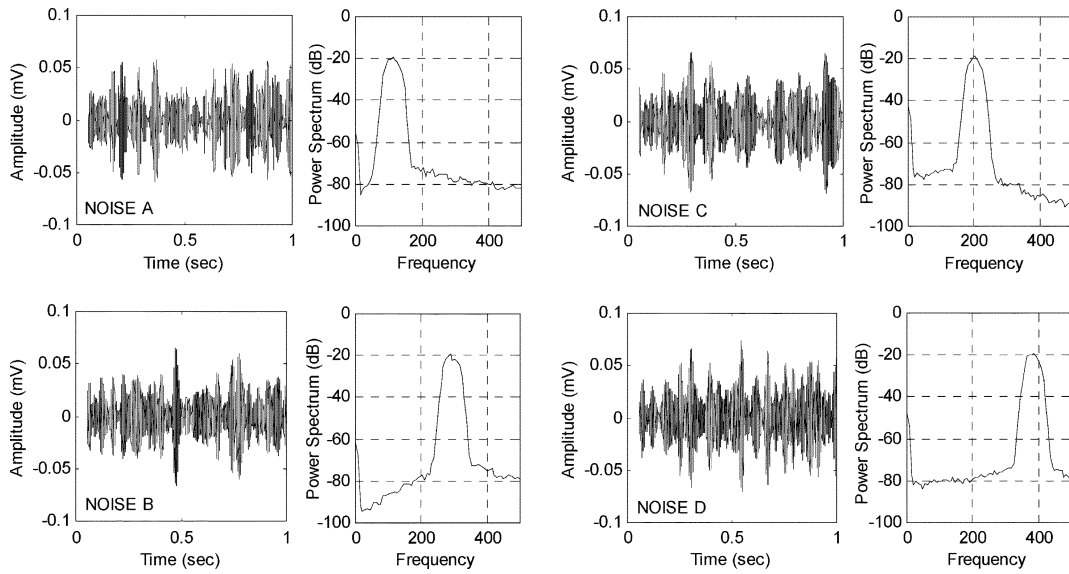
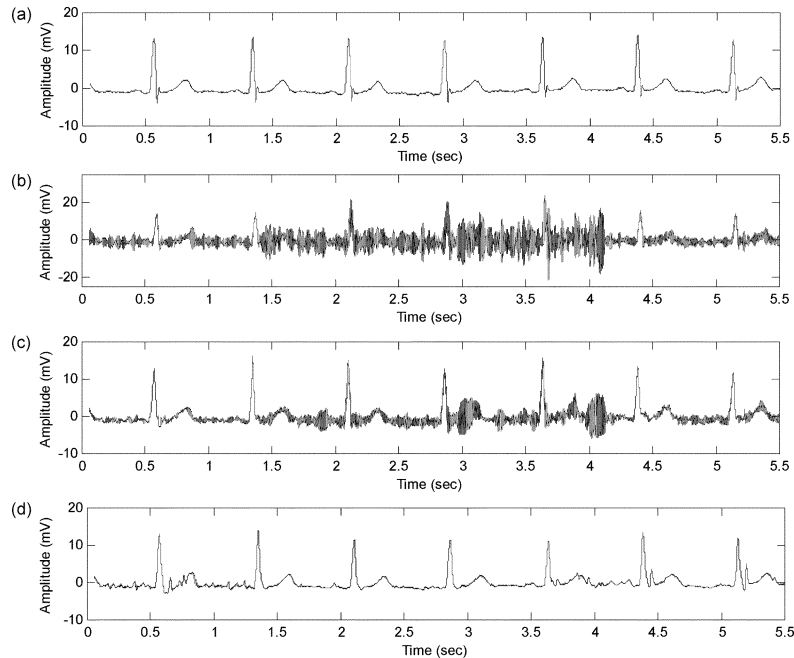Fig. 8. Four types of Gaussian noises and their power spectrum.

Fig. 9. (a) Noise-free pattern for case I. (b) Noisy simulated trial under nonstationary noises. (c) Result obtained by RBF-NLMS. (d) Result obtained by HOS-based RBF.

B. For Nonstationary Gaussian Noise

In this section, the performance of HOS-based RBF for signal enhancement under nonstationary Gaussian noise is investigated. We assume that both the magnitude and the frequency of the additional noise are time-varying. Four types of additional Gaussian noise distributed in different frequency ranges were generated, shown in Fig. 8. The SNR and noise type of additional Gaussian noise were randomly varied every 1.4 ms. The varying range of SNR of additional noise is from -2.5 to -10 dB. Figs. 9–11 are, respectively, the simulated results obtained by RBF-NLMS ($N_0 = 20$, $N_1 = 30$, $\mu_w = 0.1$, and $\mu_v = 0.5$) and HOS-based RBF ($N_0 = 20$, $N_1 = 30$, $\mu_w = 0.1$, $\mu_v = 0.5$, $m_1 = 10$, $m_2 = 3$, and $\lambda = 0.9$) for

the three types of nonlinear signals. It shows that HOS-based RBF can effectively enhance signals under nonstationary noise. The influence of the variation of frequency of Gaussian noise seems negligible to both RBF-NLMS and HOS-based RBF. Therefore, RBF also can be a good approach for a signal enhancement, except under poor SNR. The influence of the variation of magnitude of the additional noise on the performance of HOS-based RBF is small in these cases. The average of mean square error of 50 randomly selected results was used to present performance, and the performance of various methods under nonstationary Gaussian noise was listed in Table II. It shows that HOS-based RBF can provide better performance of signal enhancement under the nonstationary noise in these cases.
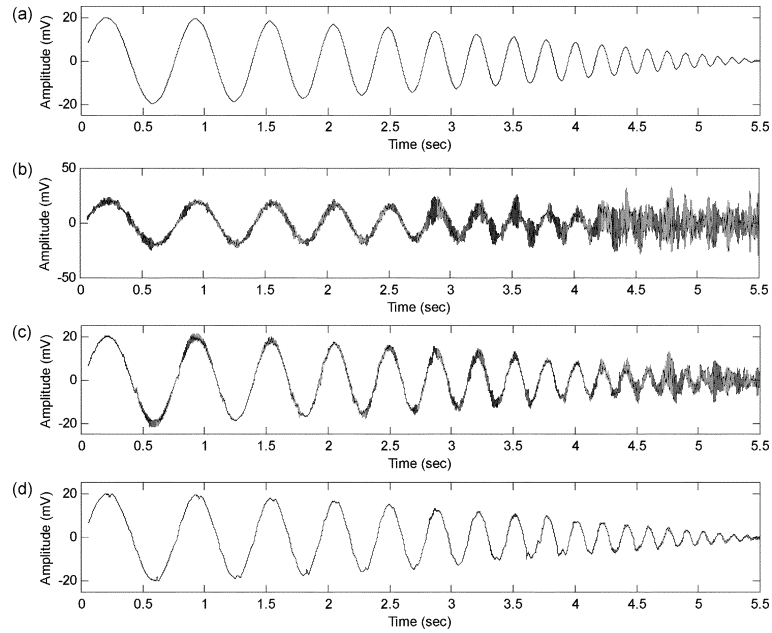
Fig. 10.   (a) Noise-free pattern for case II. (b) Noisy simulated trial under nonstationary noises. (c) Result obtained by RBF-NLMS. (d) Result obtained by HOS-based RBF.
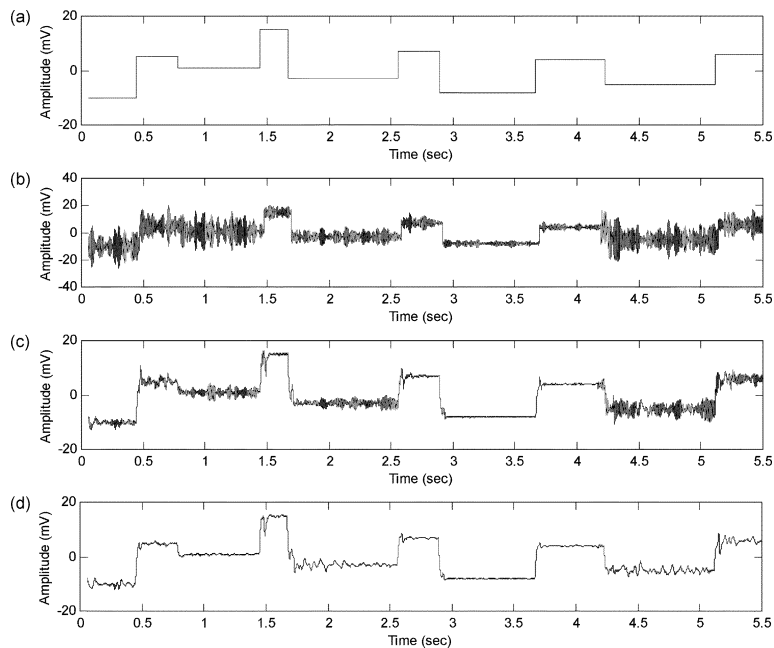

Fig. 11.   (a) Noise-free pattern for case III. (b) Noisy simulated trial under nonstationary noises. (c) Result obtained by RBF-NLMS. (d) Result obtained by HOS-based RBF.

TABLE II
COMPARISON OF PERFORMANCE BY USING RBF-NLMS AND HOS-BASED RBF FOR THREE DIFFERENT CASES

Set of tests	Performance	
	RBF-NLMS	**HOS-based RBF**
Case I	2.4092	1.0966
Case II	4.1583	1.2829
Case III	2.8423	1.6529

The computational complexity for an algorithm is an important aspect. The comparison of computational complexity for normalized LMS algorithm and HOS-based learning algorithm was shown in Table III. Here, the number of multiplications per iteration was used to present the computational complexity. From Table III, it shows that the computational complexity of HOS-based learning algorithm is very large. Therefore, the improvement of the performance of HOS-based RBF is achieved at the expense of more computations.

V. CONCLUSION

In this paper, HOS-based RBF networks have been presented and employed for signal enhancement. By using HOS to suppress Gaussian noise, higher order cumulants of the reference

TABLE III
COMPARISON OF COMPUTATIONAL COMPLEXITY OF NORMALIZED LMS
ALGORITHM AND HOS-BASED LEARNING ALGORITHM

Type of algorithm	Computational complexity
Normalized LMS algorithm	$3N_1 + 2$
HOS-based learning algorithm	$(2m_1 - 1)m_2(N_1^2 + N_1 + 9) + 1$

signal were used as the input of HOS-based RBF to reduce the influence of additional noise on clustering of input space. Simulations show that using HOS techniques provides more stable clustering on input space and an HOS-based supervised learning algorithm is developed to adapt weights. In this algorithm, the mean square error of two higher order cumulants of $d(t)$ and $\hat{s}(t)$ is defined as the learning criterion. As we expected, the influence of additional Gaussian noise on the adaptation of weights can be reduced and the gradient noise amplification problem on LMS algorithm can be effectively solved by using HOS-based learning algorithm. Moreover, the simulated results show that the performance by using HOS-based learning algorithm is stable and superior to that of LMS algorithm under different noise levels. In fact, RBF is also a good approach for signal enhancement under the nonstationary Gaussian noise, except under poor SNR. HOS-based RBF can effectively overcome this drawback, although HOS-based RBF needs more expensive computational complexity. Overall, HOS-based RBF is practicable for signal enhancement under the stationary and nonstationary Gaussian noise with different SNR.

REFERENCES

[1] B. Widrow, J. R. Glover, Jr., J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, Jr., and R. C. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proc. IEEE*, vol. 63, no. 12, pp. 1692–1716, Dec. 1975.

[2] J. R. Zeidler, "Performance analysis of LMS adaptive prediction filters," *Proc. IEEE*, vol. 78, no. 12, pp. 1781–1806, Dec. 1990.

[3] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[4] J. W. Watterson, "An optimum multilayer perceptron neural receiver for signal detection," *IEEE Trans. Neural Netw.*, vol. 1, no. 4, pp. 298–300, Dec. 1990.

[5] Z. Michalopoulou, L. Nolte, and D. Alexandrou, "Performance evaluation of multilayer perceptrons in signal detection and classification," *IEEE Trans. Neural Netw.*, vol. 6, no. 2, pp. 381–386, Mar. 1995.

[6] R. P. Lippmann and P. Beckman, "Adaptive neural net preprocessing for signal detection in non-Gaussian noise," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 1989, vol. 1, pp. 124–132.

[7] P. P. Gandhi and V. Ramamurti, "Neural networks for signal detection in non-Gaussian noise," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2846–2851, Nov. 1997.

[8] I. Cha and S. A. Kassam, "Interference cancellation using radial basis function networks," *Signal Process.*, vol. 47, pp. 247–268, 1995.

[9] S. A. Billings and C. F. Fung, "Recurrent radial basis function networks for adaptive noise cancellation," *Neural Netw.*, vol. 8, no. 2, pp. 273–290, 1995.

[10] D. S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive network," *Complex Syst.*, vol. 2, pp. 321–355, 1988.

[11] J. Park and I. W. Sandberg, "Universal approximation using radial basis function networks," *Neural Comput.*, vol. 3, pp. 246–257, 1991.

[12] ——, "Approximation and radial basis function networks," *Neural Comput.*, vol. 5, pp. 305–316, 1993.

[13] J. E. Moody and C. J. Darken, "Fast learning in networks of locally-tuned processing units," *Neural Comput.*, vol. 1, pp. 281–294, 1989.

[14] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.

[15] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," *IEEE Trans. Neural Netw.*, vol. 4, no. 4, pp. 570–578, Jul. 1993.

[16] M. Brown and C. Harris, *Neurofuzzy Adaptive Modelling and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[17] T. Poggio and F. Girosi, "Regularization algorithm for learning that are equivalent to multilayer networks," *Science*, vol. 247, pp. 978–982, 1990.

[18] B. M. Sadler, G. B. Giannakis, and K.-S. Li, "Estimation and detection in non-Gaussian noise using higher order statistics," *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2729–2741, Oct. 1994.

[19] H. M. Ibrahim, R. R. Gharieb, and M. M. Hassan, "A higher order statistics-based adaptive algorithm for line enhancement," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 527–532, Feb. 1999.

[20] A. N. Rajagopalan, "Image clustering using higher-order statistics," *Electron. Lett.*, vol. 38, pp. 122–124, 2002.

[21] K. Li, M. N. S. Swamy, and M. O. Ahmad, "An improved voice activity detection using higher order statistics," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 5, pp. 965–974, Sep. 2005.

[22] M. E. Hassouni, H. Cherifi, and D. Aboutajdine, "HOS-based image sequence noise removal," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 572–581, Mar. 2006.

[23] S. Mo and B. Shafai, "Blind equalization using higher order cumulants and neural network," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3209–3217, Nov. 1994.

[24] J.-S. Lee, J.-H. Kim, D.-K. Jee, J.-J. Hwang, and J.-H. Lee, "Blind equalization using RBF and HOS," *Intell. Data Eng. Autom. Learn.*, vol. 2690, pp. 442–446, 2003.

[25] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Process. Mag.*, vol. 10, no. 3, pp. 10–37, Jul. 1993.

[26] C. L. Nikias and A. P. Petropulu, *Higher Order Spectral Analysis: A Nonlinear Signal Processing Framework*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

Bor-Shyh Lin (S'02–M'03) was born in Changhua, Taiwan, in December 1975. He received the B.S. degree in electrical engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1997 and the M.S. and Ph.D. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1999 and 2006, respectively.

His research interests are in areas of medical informatics, and biomedical signal processing.

Bor-Shing Lin was born in Changhua, Taiwan, in December 1973. He received the B.S. degree in electrical engineering from National Cheng Kung University, Tainan City, Taiwan, in 1997 and the M.S. and Ph.D. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1999 and 2006, respectively.

His research interests are in areas of signal processing, medical informatics, wireless communication systems, embedded system, and RF IC design.

Fok-Ching Chong received the B.S. degree from the Electrical Engineering Department, National Taiwan University, Taipei, Taiwan, in 1971.

Since then he has been with the same university. He was the Head of the Operation Division of the Computer Center from 1976 to 1989. He was a Group Leader for the Medical Engineering Division of the Electrical Engineering Department. He is interested in biomedical signal processing, simulations, system analysis, and design.

Feipei Lai (S'84–M'87–SM'94) received the B.S.E.E. degree from the National Taiwan University, Taipei, Taiwan, in 1980 and the M.S. and Ph.D. degrees in computer science from the University of Illinois at Urbana-Champaign, in 1984 and 1987, respectively.

He is a Professor in the Graduate Institute of Biomedical Electronics and Bioinformatics, the Department of Computer Science and Information Engineering and the Department of Electrical Engineering at National Taiwan University. He is a Vice Superintendent of the National Taiwan University Hospital. He is the Chairman of the Taiwan Network Information Center. He was a Visiting Professor in the Department of Computer Science and Engineering, the University of Minnesota, Minneapolis. He was also a Guest Professor at the University of Dortmund, Dortmund, Germany and a Visiting Senior Computer System Engineer in the Center for Supercomputing Research and Development, the University of Illinois at Urbana-Champaign. He holds six Taiwan patents and three U.S. patents. He is a Co-Founder of the Institute of Information and Computing Machinery. His current research interests are SOC low-power computing and medical information system.

Dr. Lai is a member of Phi Kappa Phi, Phi Tau Phi, Association for Computing Machinery (ACM), and Chinese Institute of Engineers. He is the Chairman of Taiwan Internet Content Rating Foundation. He received the Taiwan Fuji Xerox Research Award in 1991. He was included in *Who's Who in Science and Engineering* and *Who's Who in the World*.