Vol 14 No 11, November 2005
1009-1963/2005/14(11)/2181-08

Chinese Physics

© 2005 Chin. Phys. Soc.
and IOP Publishing Ltd

Neural Volterra filter for chaotic time series prediction*

Li Heng-Chao(李恒超)[a)†], Zhang Jia-Shu(张家树)[a)], and Xiao Xian-Ci(肖先赐)[b)]

[a)] *Sichuan Province Key Laboratory of Signal and Information Processing, Southwest Jiaotong University, Chengdu 610031, China*

[b)] *Department of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China*

(Received 31 March 2005; revised manuscript received 11 May 2005)

A new second-order neural Volterra filter (SONVF) with conjugate gradient (CG) algorithm is proposed to predict chaotic time series based on phase space delay-coordinate reconstruction of chaotic dynamics system in this paper, where the neuron activation functions are introduced to constraint Volterra series terms for improving the nonlinear approximation of second-order Volterra filter (SOVF). The SONVF with CG algorithm improves the accuracy of prediction without increasing the computation complexity. Meanwhile, the difficulty of neuron number determination does not exist here. Experimental results show that the proposed filter can predict chaotic time series effectively, and one-step and multi-step prediction performances are obviously superior to those of SOVF, which demonstrate that the proposed SONVF is feasible and effective.

Keywords: chaotic time series, adaptive neural Volterra filter, conjugate gradient algorithm
PACC: 0545

1. Introduction

With the development of chaos theory and research on its application technique, modeling and prediction of chaotic time series have received considerable attention in the area of chaotic signal processing,[1] and have been applied to economy,[2] hydrology[3] and biomedicine[4] fields etc. successfully. The key problem of predicting chaotic time series lies in the formation of prediction model.

In recent years, neural networks such as multilayer perceptron (MLP),[5,6] radial basis function network (RBFN),[7,8] wavelet network (WN),[9] fuzzy neural network (FNN),[10] and recurrent neural network (RNN)[11] have been used to predict chaotic time series successfully due to their favourable nonlinear approximation capability, parallel information processing and fault tolerance ability. However neural network also presents many weaknesses:[13] (i) the numbers of hidden layers and neurons are difficult to determine; (ii) it suffers from over-fitting phenomenon and slow convergence characteristics; and (iii) it is prone to be trapped in local minima.

Compared to neural networks, adaptive Volterra filters synthetically utilize linear and nonlinear terms and essentially belong to the category of parameter identification model, so they have better prediction capability.[13] But computation complexity of adaptive Volterra filters will increase exponentially with the increase of their orders, which restricts the nonlinear order of Volterra filters to no larger than two in general. Furthermore, second-order Volterra filters (SOVF) have some limitations in predicting high dimensional chaotic time series, because the capability of approximating nonlinear functions is limited.[13–15]

It is well-known that nonlinear approximation of neural networks mainly comes from neuron activation function, so a new second-order neural Volterra filter (SONVF) is proposed to predict chaotic time series in this paper, where the neuron activation functions are introduced to constrain the Volterra series terms for improving the nonlinear approximation of second-order Volterra filter. Thus not only can the prediction accuracy be improved, there also does not exist the difficulty of the determination of the numbers of hidden layers and neurons.

This paper is organized as follows. In section 2 we

*Project supported by the National Natural Science Foundation of China (Grant No 60276096), the National Ministry Foundation of China (Grant No 51430804QT2201).

†E-mail: lihengchao_78@163.com

briefly review the prediction and modeling of chaotic time series. SONVF and its adaptive algorithm are formulated at length in Section 3. In Section 4, we present some simulation and results. Section 5 contains a brief summary and discussion.

2. Prediction and modeling of chaotic time series

Suppose a scalar time series$\{x(t), t = 1, 2, ..., n\}$ is obtained from a measurement on the chaotic dynamical system in the state space. If the embedding dimension m and the time delay τ have been chosen for the phase-space reconstruction, a reconstructed vector can be written as follows:

$$X_i = \{x(i), x(i-\tau), ..., x(i-(m-1)\tau)\}. \quad (1)$$

According to Taken's embedding theorem or its extensions, there exists a function F such that:

$$x(t+T) = F(X_t), \quad (2)$$

where integer T is the look-ahead time. Let the reconstructed attractor be denoted by S, it is necessary to fit F only in the domain S by the prediction model. A model of chaotic time series can be obtained by finding a functional form $\hat{F}$ for the training set pairs $(X(t), d(t)), t = 1, \cdots, n$, where $d(t) = x(t+T)$ is the output of the model, and $X(t)$ the input of the model. $\hat{F}$ maps $X(t)$ to $d(t)$:

$$x(t+T) = \hat{F}(X_t) + \varepsilon(t), \quad (3)$$

where $\varepsilon(t)$ represents the noise with finite length and fitting error.

3. Adaptive neural Volterra filter

3.1. Filter structure

Nonlinear filters based on Volterra series expansions provide a good representation to deal with nonlinear signals and systems.[16] Research results have shown that SOVF can adaptively track current chaotic fine trajectory and effectively predict some chaotic time series through using prediction errors to update filter coefficients. Presently, SOVF and its variant types have been widely used to predict some chaotic time series.[13−15,17,18]

Considering the second-order M-sample memory truncated Volterra filter described by the input-output relationship

$$x(n+T) = h_0 + \sum_{i=0}^{M-1} h_1(i)x(n-i) + \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} h_2(i,j)x(n-i)x(n-j), \quad (4)$$

where $x(n)$ and $x(n+T)$ are the input and output of SOVF, and M denotes the memory size of nonlinear system. Also, $h_1(i)$ and $h_2(i,j)$ represent the linear and quadratic filter coefficients, respectively. Without loss of generality, the quadratic coefficients are assumed to be symmetric: i.e. $h_2(i,j) = h_2(j,i)$. We rewrite Eq.(1) as

$$x(n+T) = h_0 + \sum_{i=0}^{M-1} h_1(i)x(n-i) + \sum_{i=0}^{M-1}\sum_{j=i}^{M-1} h_2(i,j)x(n-i)x(n-j). \quad (5)$$

However, the capability of SOVF to approximate nonlinear functions is limited, especially for predicting high dimensional chaotic time series.[13−15] It is well-known that nonlinear approximation of neural networks mainly comes from neuron activation function. In order to improve the ability of approximating nonlinear functions, SONVF is proposed to predict chaotic time series, where the neuron activation functions are introduced to constrain Volterra series terms. The proposed filter can be described as follows:

$$x(n+T) = h_0 S(1) + \sum_{i=0}^{M-1} h_1(i)S(x(n-i)) + \sum_{i=0}^{M-1}\sum_{j=i}^{M-1} h_2(i,j)S(x(n-i)x(n-j)), \quad (6)$$

where $S(\bullet)$ is neuron activation function such as $\dfrac{1}{1+\mathrm{e}^{-\beta x}}$ or $\dfrac{1-\mathrm{e}^{-\beta x}}{1+\mathrm{e}^{-\beta x}}$. Equation (6) can be viewed as a representation of neural network. Its hidden layers and neurons are determined by the Volterra series directly. The adaptive SONVF filer structure is shown in Fig.1.

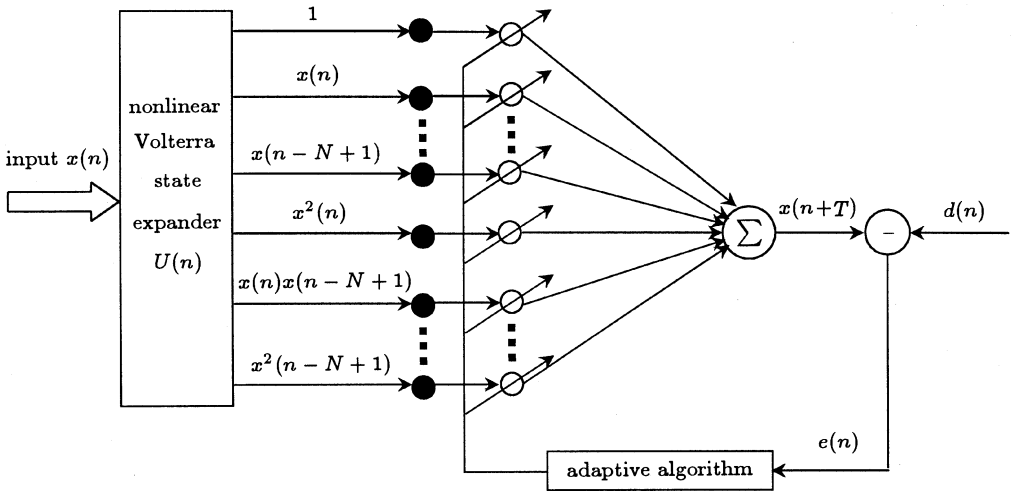


Fig.1. The structure of second-order neural Volterra filter.
● — Neuron activation function, ○— Coefficient of Volterra filter.

3.2. Adaptive algorithm

Adaptive Volterra filters can be developed using linear techniques by embedding the nonlinear problem into a linear multivariate problem, in which Volterra filter's coefficients can be updated by adaptive linear filtering algorithms. The corresponding SONVF in Eq.(6) can be embedded as follows:

$$x(n+T) = H^{\mathrm{T}}(n)X(n), \tag{7}$$

where superscript T denotes the transpose operator, the input vector $X(n)$ and corresponding adaptive coefficient vector $H(n)$ at time index n are defined as:

$$\begin{aligned} X(n) =&[S(1), S(x(n)), \cdots, S(x(n-M+1)),\\ &S(x^2(n)), \cdots S(x(n)x(n-M+1)),\\ &S(x^2(n-1)), \cdots,\\ &S(x(n-1)x(n-M+1)), \cdots,\\ &S(x^2(n-M+1))]^{\mathrm{T}}, \end{aligned} \tag{8}$$

and

$$\begin{aligned} H(n) =&[h_0, h_1(0), \cdots, h_1(M-1),\\ &h_2(0,0), \cdots, h_2(0,M-1),\\ &h_2(1,0), \cdots, h_2(1,M-1), \cdots,\\ &h_2(M-1,M-1)]^{\mathrm{T}}. \end{aligned} \tag{9}$$

Of the numerous algorithms that can be used for adaptive linear filters, the conjugate gradient (CG) algorithm has many advantages over the normalized least mean square (NLMS) and recursive least square (RLS) algorithm,[19,20] such as: (i) the CG algorithm has much faster rate of convergence than the NLMS algorithm; and (ii) it has lower computational complexity without inverting the auto-correlation matrix R and better numerical stability than the RLS algorithm. Therefore the CG algorithm is applied to adaptive SONVF in this paper. Based on the SONVF's construction of Fig.1, the CG algorithm for SONVF is presented as follows:

Initialization

$$\begin{aligned} &H(0) = [0, 0, \cdots, 0]^{\mathrm{T}},\\ &g(0) = [-1, 0, \cdots, 0]^{\mathrm{T}},\\ &p(1) = g(0), (\lambda_{\mathrm{f}} - 0.5) \le \eta \le \lambda_{\mathrm{f}},\ n = 1. \end{aligned}$$

Algorithm

$$R(n) = \lambda_{\mathrm{f}} R(n-1) + X(n)X^{\mathrm{T}}(n), \tag{10}$$

$$\alpha(n) = \eta \frac{p^{\mathrm{T}}(n)g(n-1)}{p^{\mathrm{T}}(n)R(n)p(n)}, \tag{11}$$

$$H(n) = H(n-1) + \alpha(n)p(n), \tag{12}$$

$$\begin{aligned} g(n) =&\lambda_{\mathrm{f}} g(n-1) - \alpha(n)R(n)p(n)\\ &+ X(n)(d(n) - W^{\mathrm{T}}(n-1)X(n)), \end{aligned} \tag{13}$$

$$\beta(n) = \frac{(g(n) - g(n-1))^{\mathrm{T}} g(n)}{g^{\mathrm{T}}(n-1)g(n-1)}, \tag{14}$$

$$p(n+1) = g(n) + \beta(n)p(n), \tag{15}$$

where $\lambda_{\mathrm{f}} (0 << \lambda_{\mathrm{f}} \le 1)$ is the forgetting factor, $g(n)$ the residual vector, $p(n)$ the direction vector and $R(n)$ the auto-correlation matrix. $\alpha(n)$ is the step size that minimizes the cost function $f(X) = \frac{1}{2}X^{\mathrm{T}}AX + b^{\mathrm{T}}X$, $\beta(n)$ provides R-conjugacy for the direction vector. The parameter η in Eq.(11) controls the convergence of the algorithm and must be set within the range $(\lambda_{\mathrm{f}} - 0.5) \le \eta \le \lambda_{\mathrm{f}}$.[20]

4. Simulation and results

To illustrate the effectiveness of the SONVF for making predictions, Lorenz system and a four-dimensional nonlinear dynamical system are chosen as examples. Each example uses the corresponding 2000 values after omitting initial transient points. The experimental time series are scaled appropriately in the range [–1, 1] and divided into two parts: the first 1000 values are the training set used to reconstruct the model, and the rest are treated as the test set. The mean-square-error (MSE) and the relative mean-square-error (RMSE) are used to evaluate the prediction accuracy, where:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}[x(i) - \hat{x}(i)]^2, \tag{16}$$

$$\mathrm{RMSE} = \sum_{i=1}^{n}[x(i) - \hat{x}(i)]^2 \Big/ \sum_{i=1}^{n} x^2(i). \tag{17}$$

In the following numerical experiments, the embedding dimension of Lorenz system and four-dimensional nonlinear dynamical system chaotic time series are 3 and 4, respectively, and the time delay τ is 1 in order to reconstruct the phase space of chaotic time series given above. The forgetting factor λ_{f} and parameter η are selected to be $\lambda_{\mathrm{f}} = 0.99$ and $\eta = 0.5$, and $\dfrac{1 - \mathrm{e}^{-\beta x}}{1 + \mathrm{e}^{-\beta x}}$ is implemented for neuron activation function in this research. To show clearly the prediction results of the SONVF for making one-step and multi-step prediction of chaotic series, only one is plotted for each ten points compared with the actual values in Fig.2(a) to Fig.5(a).

4.1. Prediction performance

Example 1 Lorenz chaotic system with integral step 0.01:

$$\begin{cases} \dot{x} = \sigma(y - x), \\ \dot{y} = -xz + rx - y, \\ \dot{z} = xy - bz, \end{cases} \tag{18}$$

$$\sigma = 10, r = 28, b = 8/3.$$

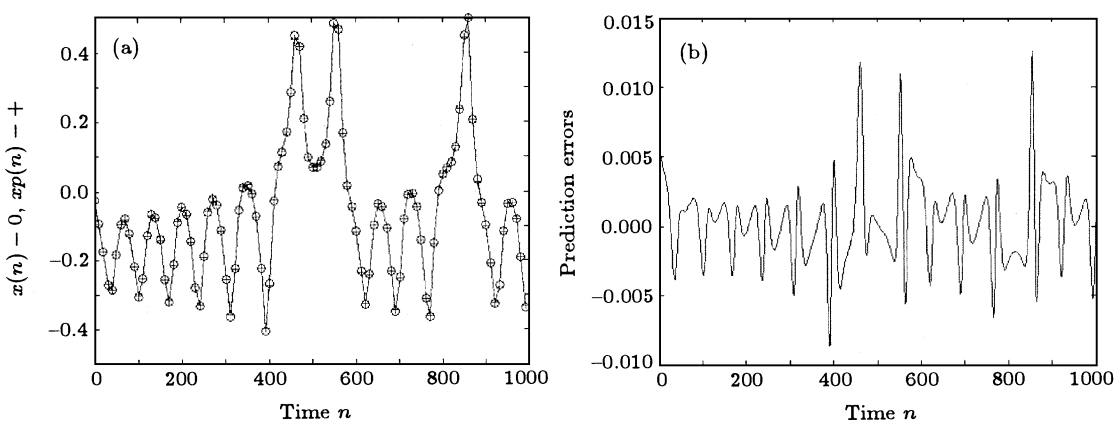


Fig.2. (a) Result of the original and predicted Lorenz system (true values $x(n)$–$\circ$, prediction values $xp(n)$–+), (b) Prediction errors. (MSE=8.4359×10^{-6}, RMSE=0.00018754, m=3, T=1, β=2).

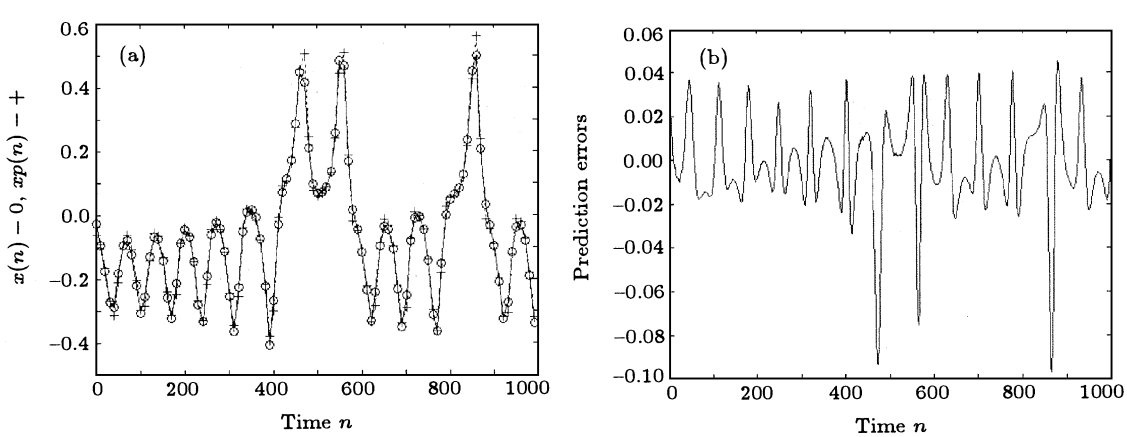


Fig.3. (a) Result of the original and predicted Lorenz system (true values $x(n)$–$\circ$, prediction values $xp(n)$–+), (b) Prediction errors. (MSE=4.5784×10^{-4}, RMSE=0.010178, m=3, T=6, β=4).

Example 2 Four-dimensional nonlinear dynamical system with integral step 0.1:

$$\begin{cases} \dot{x} = ax - y, \\ \dot{y} = x - yz^2, \\ \dot{z} = -b_1 y - b_2 z - b_3 w, \\ \dot{w} = z + cw, \end{cases} \tag{19}$$

$$a = 0.56, b_1 = 1.0, b_2 = 1.0, b_3 = 6.0, c = 0.80.$$

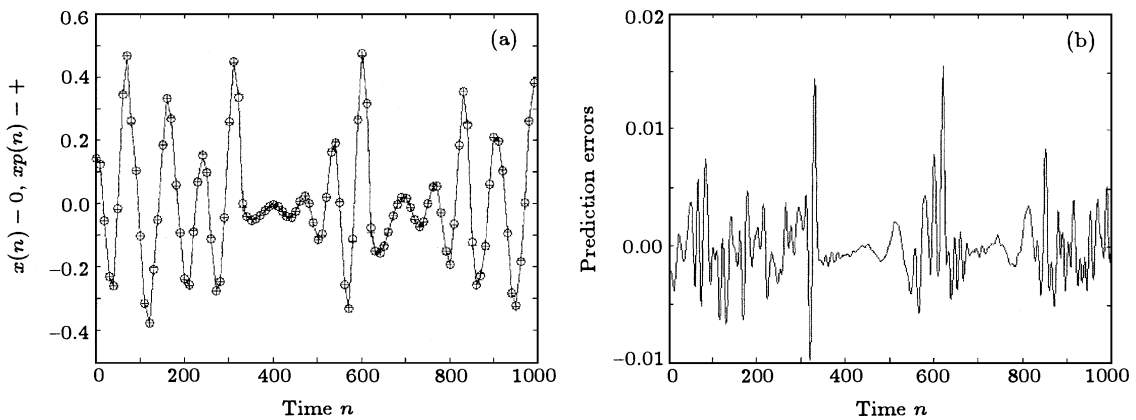


Fig.4. (a) Result of the original and predicted four-dimensional nonlinear dynamical system (true values $x(n)$–$\circ$, prediction values $xp(n)$–+), (b) Prediction errors. (MSE=8.1804×10^{-6}, RMSE=0.00022839, m=4, T=1, β=1).

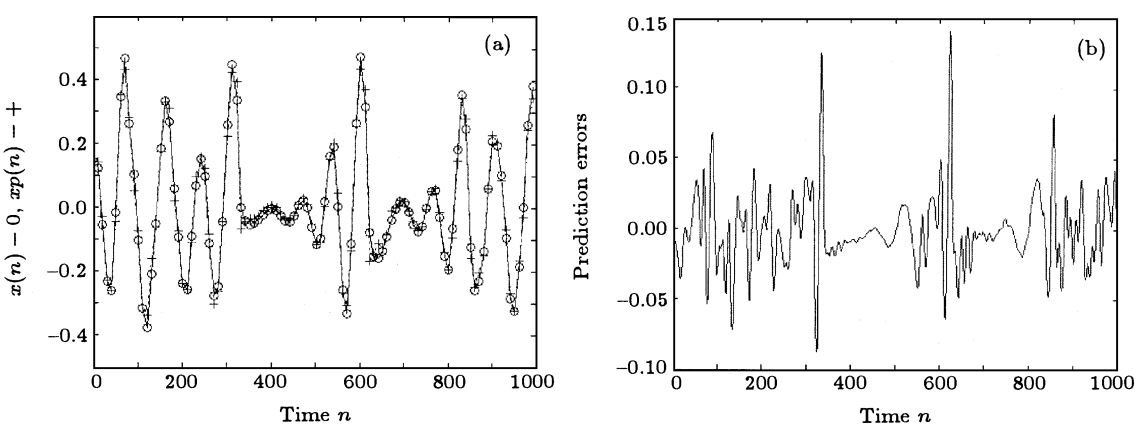


Fig.5. (a) Result of the original and predicted four-dimensional nonlinear dynamical system (true values $x(n)$–$\circ$, prediction values $xp(n)$–+), (b) Prediction errors. (MSE=6.9469×10^{-4}, RMSE=0.019395, m=4, T=6, β=2).

From the experimental results listed above, it can be seen clearly that the predicted values for the case of one step agree well with true values and that the predicted results for six steps are also good, though not as good as those of one step. The MSEs of two kinds of chaotic time series are less than 8.500×10^{-6} for one-step prediction, and the MSEs of the six-step predictions are below 7.0000×10^{-4}. Obviously, this neural Volterra filter with conjugate gradient algorithm is capable of capturing the underlying dynamics from chaotic time series efficiently.

In order to further discuss the prediction accuracy of the proposed filter, the performance comparison of SOVF and SONVF is explored using the same chaotic time series given above. The SOVF with NLMS algorithm and SONVF using two different adaptive algorithms (NLMS and CG algorithms) are implemented respectively in Matlab software. Table 1 presents the comparative results of three kinds of Volterra filters.

Table 1. The comparative results of three kinds of Volterra filters (MSE).

Prediction model	Lorenz system		Four-dimensional nonlinear dynamical system	
	One step	Six step	One step	Six step
SOVF (NLMS)	0.00025712	0.0011883	0.00057483	0.0039122
SONVF(NLMS)	4.4136×10^{-5}	0.00069956	3.4407×10^{-5}	0.00079312
SONVF(CG)	8.4359×10^{-6}	4.5784×10^{-4}	8.1804×10^{-6}	6.9469×10^{-4}

The results presented in Table 1 illustrate that one-step and six-step prediction accuracy of the SONVF is obviously superior to those of SOVF, and the performance of CG algorithm is also better than that of NLMS algorithm, which demonstrate that the SONVF proposed in this paper is feasible and effective. At the same time, this proposed filter overcomes the scenario that the capability of SOVF to approximate nonlinear functions is limited, and the numbers of hidden layers and neurons are difficult to be determined in neural networks.

4.2. Slope β of neuron activation function

Neuron activation function is the core of neural networks, the nonlinear approximation of neural networks depends on neuron activation function to a large extent besides that it is related to the network structure and learning algorithm. Through adjusting the slope of neuron activation function, the saturation of hidden neuron in neural networks can be avoided so as to quicken the learning process and improve the accuracy of nonlinear approximation. Consequently, the influence on prediction performance of SONVF caused by the slope β of neuron activation function is discussed. One-step prediction is used as an example in this experiment, and the slope β varies in the range [–20 20] and $\beta \neq 0$, when other parameters are chosen the same as those given in prediction performance. A summary of the performed experiments is given in Fig.6.

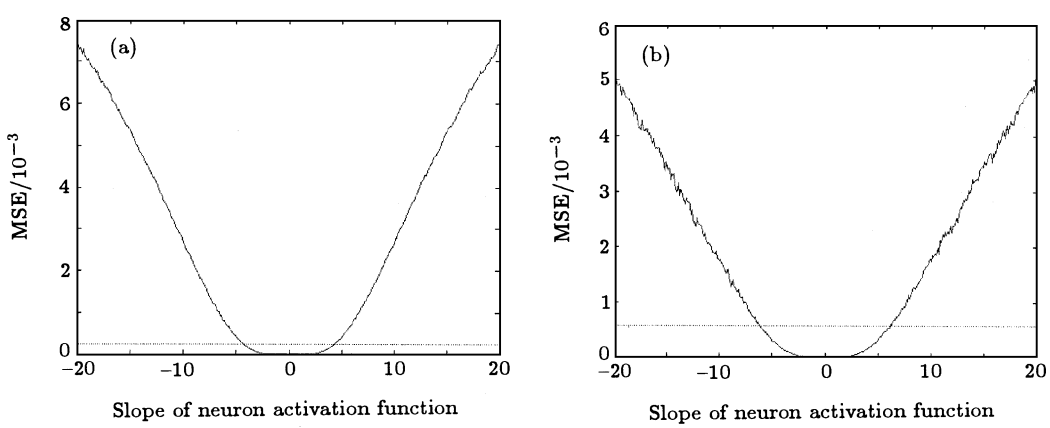


Fig.6. The relationship of MSE and the slope of neuron activation function: (a) for Lorenz system, (b) for four-dimensional nonlinear dynamical system (SONVF-solid line, SOVF-dashed line).

Figure 6(a) and 6(b) show that the influence on prediction performance of SONVF caused by the slope β of neuron activation function is symmetrical in positive and negative interval, and there is a confidence interval of the slope β of neuron activation function for SONVF compared to SOVF. And the confidence interval is different for different chaotic time series. As long as the slope value of neuron activation function is selected in the confidence interval, the prediction performance of SONVF is better than that of SOVF. At the same time, the selection is very convenient and simple.

4.3. Anti-noise performance

When chaotic dynamic system is located in a noisy environment, the collapse of chaotic attractor will take place so that chaotic dynamic system will lose its intrinsic characteristics and it is impossible to predict chaotic time series. If the predicted time series can resume its primary attractor to some extent, the model of this proposed filter is very close to dynamics characteristics of actual system and has certain anti-

noise performance.[21]

In the following numerical experiment, one-step prediction is also taken as an example, and each experimental time series are normalized and added by one percent of the additive Gaussian white noise which is generated by randn-function in Matlab and subject to normal distribution N (0,1).[22] Simulation parameters are selected the same as those given in prediction performance. Figures 7 and 8 present the attractors of noise-free and predicted noisy chaotic time series, respectively.

As shown in Fig.7 and Fig.8, the attractors of predicted noisy chaotic time series are very similar to those of noise-free chaotic time series in the outline and shape. Consequently the proposed SONVF has anti-noise performance to some extent.

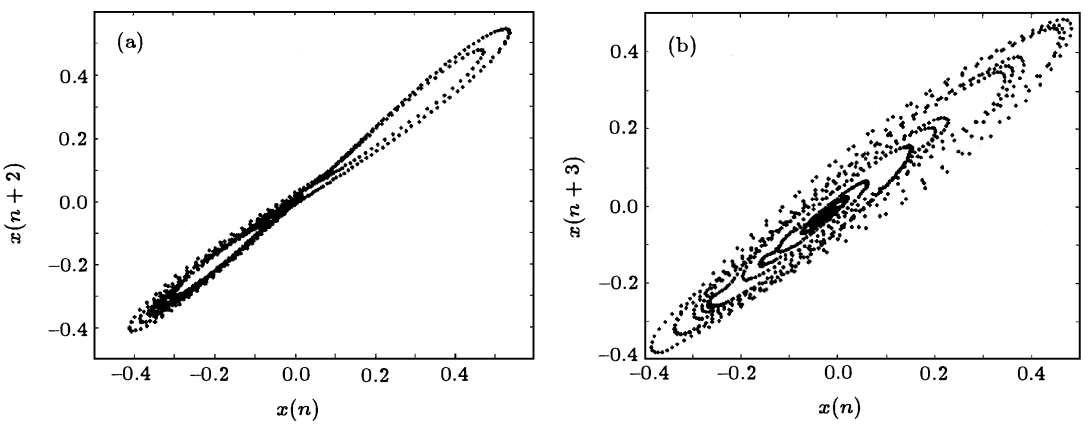


Fig.7. Attractor of noise-free chaotic time series: (a) for Lorenz system, (b) for four-dimensional nonlinear dynamical system.

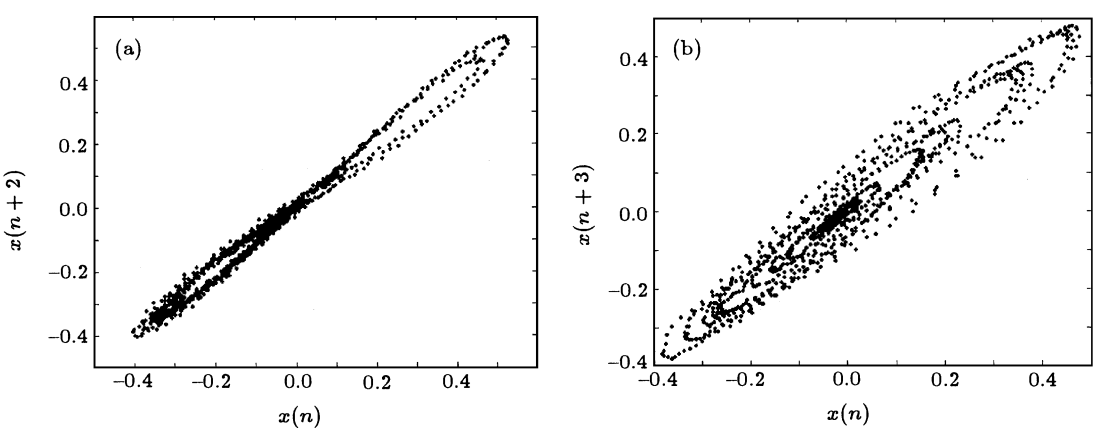


Fig.8. Attractor of predicted noisy chaotic time series: (a) for Lorenz system, (b) for four-dimensional nonlinear dynamical system.

5. Conclusion

In this paper, a new second-order neural Volterra filter (SONVF) with conjugate gradient algorithm is proposed to predict chaotic time series, where the neuron activation functions are introduced to constrain the Volterra series terms for improving the prediction accuracy of second-order Volterra filter (SOVF). From the above experimental results, we can see that this proposed filter has the following advantages. Firstly, the SONVF can predict chaotic time series effectively, and one-step and multi-step prediction performances are obviously superior to those of SOVF. Secondly, there does not exist the problem that the numbers of hidden layers and neurons are difficult to determine, and the computation complexity of this proposed SONVF does not increase. Thirdly, it has anti-noise performance to some extent.

References

[1] Kantz H and Schreiber T 1997 *Nonlinear Time Series Analysis* (Cambridge: Cambridge University Press)

[2] Saad E W, Prokhorov D V and Wunsch D C 1998 *IEEE Trans. Neural Networks* **9** 1456

[3] Sivakumar B 2002 *J. Hydrology* **258** 149

[4] Brandt M E, Ademoglu A and Pritchard W S 2000 *Int. J. Bifur. Chaos* **10** 123

[5] Haykin S and Li X B 1995 *Proc. IEEE* **83** 95

[6] Zhang J S and Xiao X C 2000 *Chin. Phys.* **9** 408

[7] He X D and Lapedes A 1993 *Physica* D **70** 151

[8] Casdagli M 1989 *Physica* D **35** 335

[9] Cao L Y, Hong Y G, Fang H and He G W 1995 *Physica* D **85** 225

[10] Maguire L P, Roche B, McGinnity T M and McDaid L J 1998 *Information Science* **112** 137

[11] Zhang J S and Xiao X C 2000 *Chin. Phys. Lett.* **17** 88

[12] Bishop C M 1995 *Neural Networks for Pattern Recognition* (Oxford: Clarendon)

[13] Zhang J S and Xiao X C 2000 *Acta Phys. Sin.* **49** 2333 (in Chinese)

[14] Zhang J S and Xiao X C 2001 *Chin. Phys. Lett.* **18** 337

[15] Wei B L, Luo X S, Wang B H *et al* 2002 *Acta Phys. Sin.* **51** 2205 (in Chinese)

[16] Mathews V J 1991 *IEEE Signal Processing Magazine* **8** 10

[17] Zhang J S and Xiao X C 2000 *Acta Phys. Sin.* **49** 403 (in Chinese)

[18] Zhang J S, Li H C and Xiao X C 2004 *Acta Phys. Sin.* **53** 710 (in Chinese)

[19] Boray G K and Srinath M D 1992 *IEEE Trans. Circuits and Systems* **39** 1

[20] Chang P S and Willson A N 2000 *IEEE Trans. Signal Processing* **48** 409

[21] Gan J C and Xiao X C 2003 *Acta Phys. Sin.* **52** 1102 (in Chinese)

[22] Gan J C and Xiao X C 2004 *Chin. Phys.* **13** 312