Proceedings of the 13th
KACC, October 1998

A multi-modal neural network using Chebyschev polynomials

Ikuo Yoshihara*, Tomoyuki Nakagawa*, Moritoshi Yasunaga**
and Ken-ichi Abe*

*Large Scale Systems Stage Eng. Lab.,Graduate School of Eng., Tohoku Univ.,
Aza-Aoba05, Aramaki, Aoba-ku, Sendai,980-8579, JAPAN
Tel:+81-22-217-7083; Fax:+81-22-263-9302; Email:yoshiha@largesys.ecei.tohoku.ac.jp

**Institute of Information Science and Electronics,University of Tsukuba,
Tennoudai, Tsukuba, 305, JAPAN
Tel:+81-298-53-5323; Fax:+81-298-535206; E-mail:yasunaga@is.tsukuba.ac.jp

Abstract

This paper presents a multi-modal neural network composed of a preprocessing module and a multi-layer neural network module in order to enhance the nonlinear characteristics of neural network. The former module is based on spectral method using Chebyschev polynomials and transforms input data into spectra. The latter module identifies the system using the spectra generated by the preprocessing module.

The omnibus numerical experiments show that the method is applicable to many a nonlinear dynamic system in the real world, and that preprocessing using Chebyschev polynomials reduces the number of neurons required for the multi-layer neural network.

Keywords : nonlinear, Chebyschev, neural network, multi-modal neural network

1. Introduction

The construction of models for pattern recognition, system identification, time series and control of initially unknown and potentially nonlinear systems is a difficult and fundamental problem of real world applications.

Neural network(NN) is a prevalent method to deal in nonlinear problems, and its representative roles are functional approximation and pattern recognition. However, if a system is strongly nonlinear, the conventional multilayer neural network is not always easy to use, because it requires many neurons. We want to develop neural network model building up nonlinear characteristics without using so many neurons[1]. Any function can be approximated using only a finite number of terms of orthogonal functions, for example a function is approximately expressed by a few terms of Taylor series. This means lower order components are more important than higher ones. Our basic idea is that giving the lower order components and neural network processes getting the components.

An intuitive expression of the significant property of the multi-modal model is that the neural network makes the discriminant surface combining curves yielded from the functional transformation, not straight lines.

We propose a multi-modal neural network consisting of multilayer neural network with a preprocessing module. The preprocessor is based on spectral method using Chebyschev polynomials and transforms input data into spectra which is used as input of the latter module.

2. A multi-modal neural network using Chebyschev polynomials

2.1 Multi-modal model

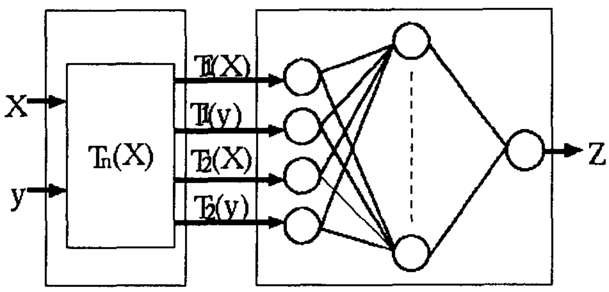

Fig.1 Multi-modal neural network

In general multilayer neural network can map a space of input data to that of output data imitating a given nonlinear function. To restrict increasing the number of neurons, we divide the mapping process into two. One is transforming input data into spectra, which are outputs of some orthogonal functions. Another is inputting the spectra into the successive neural network and it gives outputs. To address intuitively, the combined model maps partly by functional transform and partly by multilayer neural network.

A neural network model consisting of more than one neural network is often called multi-modal neural network, and a neural network involving some non-neural model is also called multi-modal neural network. Our model belongs to this category, because it consists of preprocessing module and multilayer neural network module.

2.2 Chebyschev polynomials

We select Chebyschev polynomials for preprocessing transform. Chebyschev polynomials are defined by

$$T_n(x)=\cos(n \cdot \cos^{-1} x). \quad (1)$$

First few polynomials are

$$T_0(x)=1, \quad T_1(x)=x,$$

$$T_2(x)=2x^2-1, \quad T_3(x)=4x^3-3x,\cdots. \quad (2)$$

Chebyschev polynomials have following properties,

- $T_n(x)$ is a polynomial in x of degree n.
- If n is even(odd), $T_n(x)$ is even(odd) function..
- $|T_n(x)| \leqq 1 \quad (\text{if } |x| \leqq 1)$ (3)

Most functions can be represented with orthogonal functions, *e.g.* Taylor series $f(x)= \Sigma\ a_k\ x^k$, and Chebyschev function series $f(x)= \Sigma\ a_k T_k(x)$.etc[3]. If we cut the series in only a few terms, Chebyschev approximation is more accurate than Taylor one, for the same function.[2] For a rapidly converging series, the error due to truncation is approximately given by the first term of the remainder, i.e., $a_n T_n(x)$.

To produce a viable mechanism for building a model based on Chebyschev polynomials with a limited degree is probably difficult. Our essential idea to avoid this is to combine Chebyschev transformation and multilayer neural network. It realizes no need of higher order of Chebyschev polynomials nor explicitly determining coefficient of them, because fitting to data is performed while learning of neural network.

3. Pattern recognition

Pattern recognition is one of the most representative applications of multilayer neural network. If a neural network can recognize a complicated pattern, it must discriminate a nonlinear hypersurface. To investigate the discriminant ability of the neural model, we make an experiment to recognize a region enclosed by curved lines.

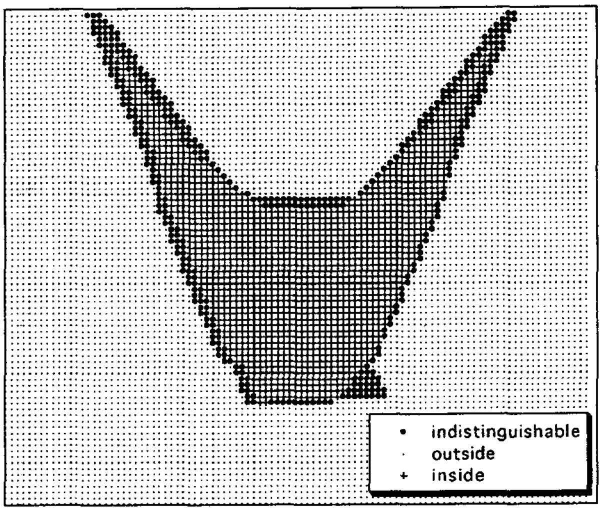

Fig2. Pattern recognition by multilayer neural network

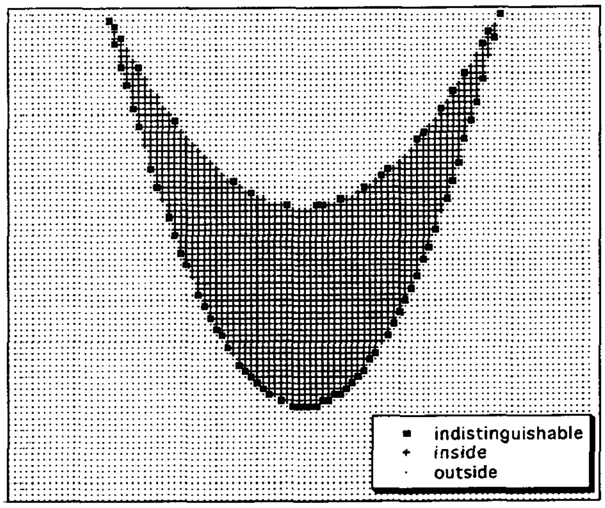

Fig3. Pattern recognition by multi-modal neural network

Fig.2 shows the result using multilayer neural network, whose structure is 2-6-4-1,*e.g.* it has two neurons in the input layer, 6 neurons in the second layer, 4 neurons in the third layer and one neuron in the output layer. Roughly speaking the region seems to be recognized, but the region is quarried by straight lines. There are many big dots(·) that denote indistinguishable points. Here we call "indistinguishable" is that the output from neural network is no less than 0.1 and no greater than 0.9. In the indistinguishable region, we cannot consider the points if they belong to inside or outside of the region.

Fig.3 shows the result using multi-modal neural network, whose structure is 4-6-2-1. This model involves 47 unknown parameters *e.g.* synapses and thresholds. Number of unknown parameters is less than that of the conventional model which requires 51 parameters. The region is recognized clearly and is quarried by softly curved lines. There are some big dots(·), but the number of them is far less than that in Fig.2. It means our model recognizes the region with only a few ambiguity.

4. Functional approximation

Functional approximation is also one of the most representative applications of multilayer neural network. To evaluate the approximating ability of the neural model, we make two types of experiment: nonlinear function approximation and time series prediction.

4.1 Non linear functions

The following is an example of test functions for functional approximation:

$$f(x) = -\log(2+x) + \sin(2\pi x) + \frac{1}{3}\cos(6\pi x) \qquad (4)$$

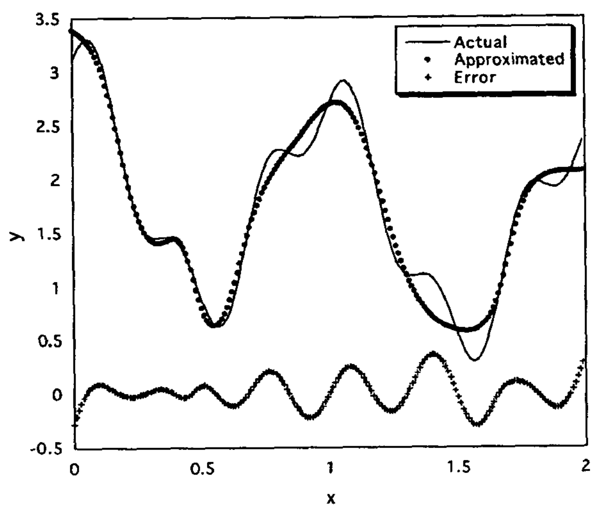

Fig4. Functional approximation by multilayer neural network

Fig.4 shows a functional approximation by multilayer neural network whose structure is 1-10-1. The conventional model can only fit outline of the curve. Small hillocks and gullys are almost ignored.

Fig.5 shows the functional approximation by multi-modal neural network whose structure is 3-6-1. Here 3 means $T_1(x)$, $T_2(x)$, and $T_3(x)$ are used for preprocessing. Seeing error curves of Fig. 4 and 5, approximation by multi-modal model is obviously better than conventional neural model.

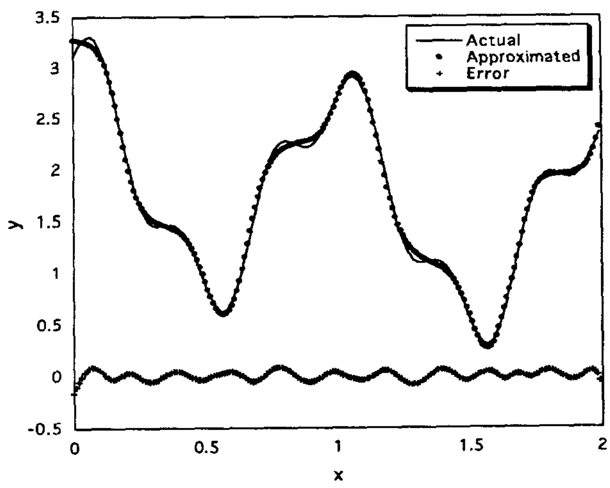

Fig5. Functional approximation by multi-modal neural network

4.2. Time series prediction

Analysis of time series data plays an important role in finance and marketing. For example, models predicting demand for product will be used to production schedule and direct capital allocation. A model derived from past data can help to predict future behavior of the market.

Time series problems are substantially function identification, and multi-modal model learns the characteristics from past data. Time series prediction is defined like that

$$Y^*_{n+1} = f(y_n, y_{n-1}, \cdots y_{n-t}) \qquad (5)$$

The test data of time series prediction is artificially generated. The values(y's) are drastically up and down like a stockprice. Prediction Y^*_{n+1} is calculated from three previous data y_n, y_{n-1}, and y_{n-3}.

Fig.6 shows predicted value using conventional neural network whose structure is 3-5-1. The prediction is errors are not so high.

Fig.7 shows predicted value using multi-modal neural network whose structure is 6-3-1. The prediction looks like almost same as Fig.6. However precise inspection reveals that the prediction error of multilayer model is about half of that of conventional model(Fig.8).

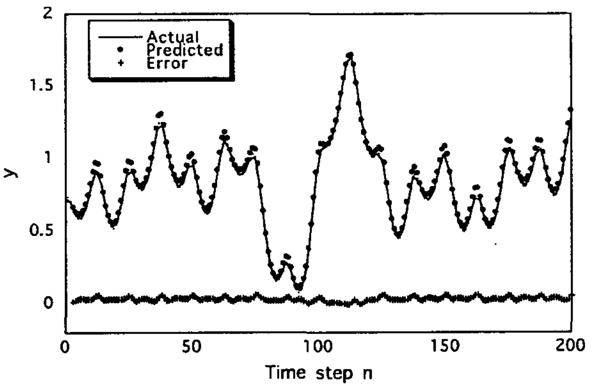

Fig.6 Time series prediction by neural network

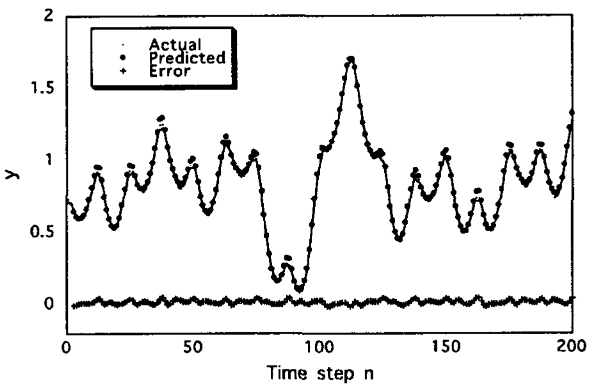

Fig.7 Time series prediction by Multi-modal neural network

5. Conclusions

We propose a malti-modal neural network model consisting of Chebyschev transform and multilayer neural network. Using Chebyschev polynomial we succeeded in enhancing the ability of multilayer neural network , for example pattern recognition and functional approximation.

Nunerical experiments result in the proposed method can find discriminant curves , and approximate strongly nonlinear functions.with small number of neurons.

We begin this research intending to build neural controller having strong nonlinearity. This application research is now going on by one of the coauthors.

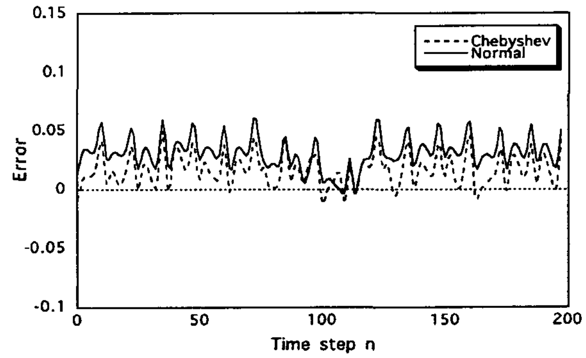

Fig.8 Comparision of prediction errors

Acknowledgement

We would like to address thankfullness to East Japan Railway Co., because Large Scale System Stages Engineering Chair in Tohoku University..is funded by them..

References

1) Vecci, L., Piazza, F., and Uncini, A.,:Learning and approximation capabilities of adaptive spline activation,Neural Networks,11,(1998)pp.259-270

2) Gerald, C.F., and Wheatly, P.,O.:"Applied Numerical Analysis", Menlo Park, CA: Addison-Wesley Pub. Co.,1992.

3) Foxand, L., and Parker, I.:"Chebyshev Polynomials in Numerical Analysis, New York,NY, 1968