2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

Recurrent Based Modular Neural Network

Jung-Heum Yon, Woo-Kyung Choi, *Yong-Min Kim and Hong-Tae Jeon
School of Electronic and Electrical Engineering, Chung-Ang Univ, Korea
*School of Computer, Chung Cheong College, Korea
221 Heuksuk-Dong, Donjak-Gu, Seoul, 156-756, Korea
E-mail : chwk001@hanmail.net

Abstract - **In this paper, we propose modular network to solve difficult and complex problems that are seldom solved with Multi-Layer Neural Network(MLNN). The structure of Modular Neural Network(MNN) in researched by Jacobs and jordan is selected in this paper. Modular network consists of several Expert Networks(EN) and a Gating Network(GN) which is composed of single-layer neural network(SLNN) or multi-layer neural network. We propose modular network structure using Recurrent Neural Network(RNN), since the state of the whole network at a particular time depends on aggregate of previous states as well as on the current input. Finally, we show excellence of the proposed network compared with modular network.**

I. Introduction

Robots has been upgraded very quickly so that they do simple and repeated works with industrial fields, disaster site and outer space instead of human. Nowadays, however, our final goal is to make a human-friendly robot that can understand what human say and learn by itself and have internal emotion. Being applied technology that imitates human ability of image and voice cognition etc, robot show up gradually as to human. However, we can not simply say a human-friendly robot that imitated human function simply.

Emotion is derived from a thought process of the being that has intelligence and they are inseparably related with reason. In other word, prior to rational sense is affected by emotion[1]. If we add the factor of emotion into robots, we may get more intelligent and human-friendly robots.

Then the intelligent method that is applied with consideration and adaptation of human has been proposed to solve the difficulty of mathematical modeling. Representative method is a soft computing, this is a method similar to solving ability of human in algorithm. MLNN and Radial Basis Function Network(RBFN) applicable in various field belong to it[2][3][4].

Neural network, which uses and train given information, have superior ability to extract character of system. However, temporal cross talk which forgets mapping result of prior learning shows up in neural network. On performing approximation of global mapping structure, neural network have defect of local minimum convergence. So we interested in the design method of modular. Modular network obtains an advantage of local mapping method and global mapping method respectively, and complicate problem is divided into several sub-task using MNN[2]. Specially, MNN which researched by Jacobs and Jordan performs superior ability in sight of recomposition, training and recombination of work. MNN consists of EN and GN[2][5]. EN divides a complex problem into several simple sub-tasks and then, learn a sub-task at each module. And so GN combines the outputs from EN and play a role of switching in order to determine final output. Most of MNN consist of several EN and a GN which is composed of SLNN or MLNN until now.

In learning, the state of the whole network at a particular time depends on an aggregate previous states as well as on the current input. So RNN applies to MNN structure. Through recurrent loop, RNN can represent situation and can perform role of feedforward network as small network

In this paper, we used training data pattern based on emotional assessment and studied by applying RNN to MNN which is proposed by Jocobs and Jordan. We compared and analyzed the consequences of learning to verify the excellence of the structure.

II. Information Process With Emotion

Since 1870, the skill of brain has been revealed through the clinical demonstration of skill of cerebrum started at Germany. In that process a sea horse is conscious memorial system and an amygdala forms emotion with sentiment and sentimental learning. The two organs have a great effect an human memorization and behavior

According to a medical report, an emotional factor, unlike a common mechanical control system, influences the control of a human brain system considerably. So learning data were obtained applying the emotional factor made of a primary environmental factor. The learning of intelligent robot uses a common factor among complicated and individual emotions.

There are several kind of information which affect an emotion. In this paper, we measured temperature, a status of smoke, brightness and the distance to an obstacle by sensors and used the result as an environmental information input. This environmental information input can be divided into 3 layers and influences an emotional factor by 9 different kinds of mode. An emotional factor can be divided into 3 things - astonishment, a state of mood and a rate of satisfaction. Astonishment is expressed in the will of the current pause of movement pattern and the requirement about a new pattern by a sudden environmental change. A state of mood represent rate of an emotional favor and hatred according to the environmental change. A rate of satisfaction is determined by a combination of these factors and degree of sensitivity about respective environmental information. Astonishment (E_a) and state of mood(E_h) are determined by formula (1).

$$E_{a,h}(n) = \sum_{i=1}^{m}(< x_i(n-1), x_i(n) > \times s_i) \quad (1)$$

Here $x_i(n)$ is an environmental factor, s_i is rate of sensitivity about environmental factor, $< x_i(n-1), x_i(n) >$ is a favor rate according to the change of environmental factor.

Satisfaction(S) rate is determined by amaze and a state of mood consists of satisfaction rate(S_{emo}) according to the change of mood and satisfaction rate(S_{emo}) according to a result of behavior pattern. Satisfaction rate(S) can be obtained formula (2).

$$S = F(S_{emo}) \times \lambda_{emo} + F(S_{act}) \times \lambda_{act} \quad (2)$$

$$S_{emo} = \lambda_{emo} \times \Delta E_h{}^{n}_{n-1}, \quad S_{act} = \lambda_{act} \times \Delta E_h{}^{n}_{n-1}$$

$\lambda_{emo}, \lambda_{act}$: Weight index of satisfaction rate

$n-T$: Time of determination of action pattern

$F(\cdot)$ is a sigmoid function.

Behavior pattern can be judged by amaze(E_a), a state of mood(E_h) and satisfaction rate(S) that are able to be determine.

III. MNN Sturcture Using MLNN

This paper used gaussian mixture modular neural network which display a great ability item of operational reconstruction and recombination of learning. MNN divided into different module each other and consists of EN which compete with each other respectively and GN which play a role of arbitration of EN. EN and GN for MNN used MLNN and was expressed by figure 1.

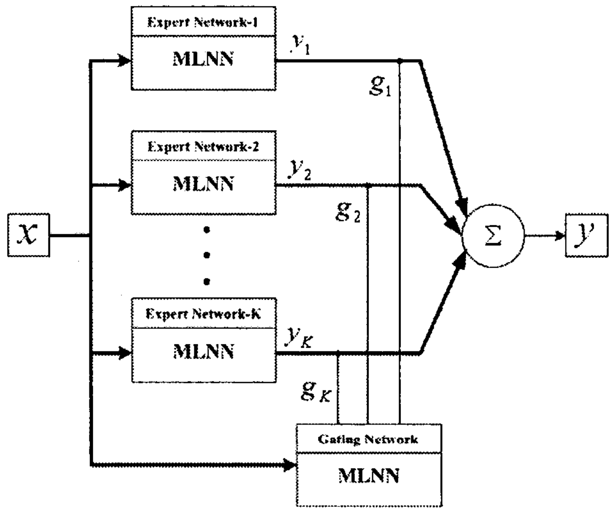

Figure 1. MNN structure using MLNN.

Input vector is P dimension by $x = [x_1\ x_2\ \cdots\ x_P]^T$, final output vector is $y = [y_1\ y_2\ \cdots\ y_Q]^T$, output vector of kth EN is $y_k = [y_k^1\ y_k^2\ \cdots\ y_k^Q]^T$ and desired vector is $d = [d_1\ d_2\ \cdots\ d_Q]^T$.

Final output of MNN is a formula (3).

$$y = \sum_{i=1}^{M} g_i y_i \quad (3)$$

If a output of GN is u_i, final output of GN which includes priori probability can be shown like equation (4) using soft-max function

$$g_i = \frac{exp(u_i)}{\sum_{j=1}^{M} exp(u_j)} \quad (4)$$

IV. MNN Structure Using RNN

MNN using RNN(MRNN) equal to structure which are MNN using MLNN and GN structure. But what is difference is using RNN in EN, this will have great effect on network output.

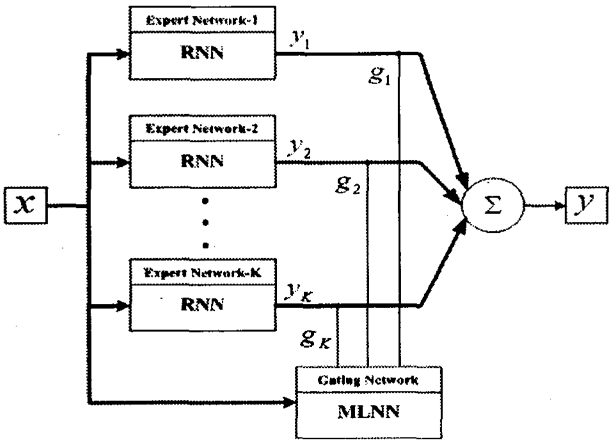

Figure 2. MNN structure using MNN.

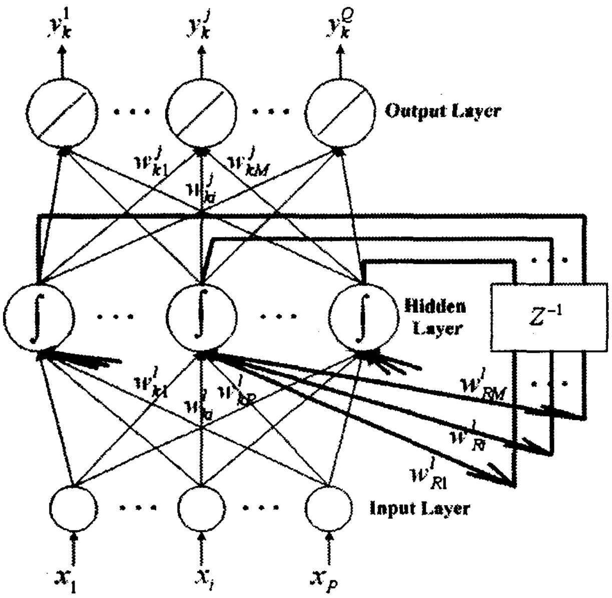

Figure 3. EN structure composed by RNN.

EN used Elman network structure which researched and published by Elman. In other word, it is said to simple recurrent network(SRN) Elman network. Hidden neuron clarify the statement of network, and hidden layer output recurrent hidden layer through delay unit bank. Input layer consists of recurrent node and input node. This network influenced and connected with exterior environment through input node[2]. RNN structure of Elman is like as figure 3.

About RNN of EN, we used bipolar sigmoid function in hidden layer, and used linear function in output layer. It consisted of Input P units, hidden layer M units and output layer Q unit nodes

lth hidden layer output of kth EN is like the next formula.

$$\begin{aligned} s_k^l(n+1) &= f\left(\sum_{i=1}^{P} w_{ki}^l\, x_i(n) + \sum_{i-1}^{M} w_{kRi}^l r_i(n)\right) \\ &= f(\boldsymbol{w}_k^l \boldsymbol{x}(n) + \boldsymbol{w}_{KR}^l \boldsymbol{r}(n)) = f(net_k^{l(e)}) \end{aligned} \quad (5)$$

Function $f(\cdot)$ is a bipolar sigmoid function.

jth output of kth EN is showed in formula (6).

$$\begin{aligned} y_k^j(n+1) &= f\left(\sum_{l=1}^{M} w_{kl}^j\, f\left(\sum_{i=1}^{P} w_{ki}^l x_i(n) + \sum_{i=1}^{M} w_{kRi}^l r_i(n)\right)\right) \\ &= f(\boldsymbol{w}_k^j \boldsymbol{s}_k(n+1)) = f(net_k^{j(e)}) \end{aligned} \quad (6)$$

$f(\cdot)$ is a linear function.

V. Learning Algorithm

Final goal of algorithm which used to learn MNN and MRNN is a modeling of probability distribution about training pattern. Learning algorithm to do it use a stochastic-gradient learning algorithm. EN and GN are trained using the difference between priori probability and posterior probability by input vector x. EN and GN are applied by a stochastic-gradient ascent learning algorithm and chain rule.

Renewal rule of weight w_{kl}^j in output layer and hidden layer of EN.

$$w_{kl}^j(t+1) = w_{kl}^j(t) + \eta \cdot h_k(t)\frac{e_k^j(t)}{\sigma_k^2(t)} \cdot s_k^l(t) \quad (7)$$

Renewal rule of weight w_{KRi}^l in hidden layer and input layer of EN.

$$\begin{aligned} w_{KRi}^l(t+1) &= w_{KRi}^l(t) + \Delta w_{KRi}^l(t) \\ &= w_{KRi}^l(t) + \eta \delta_k^l X_i \end{aligned} \quad (8)$$

Renewal rule of weight a_k^m in output layer and hidden layer of GN.

$$a_k^m(t+1) = a_k^m(t) + \eta(h_k - g_k)v_m \quad (9)$$

Renewal rule of weight a_m^i in hidden layer and input layer of GN.

$$a_m^i(t+1) = a_m^i(t) + \eta \cdot \delta_m^g \cdot x_i \qquad (10)$$

VI. Trial Experimentation

We have learned data pattern for emotion which get by environment information and experimented 100 times with simulator. Among experimentation, data which have high satisfaction rate and much experimentation are tested.

MNN used 4 EN and MLNN in each EN consisted of 6 hidden nodes. And so MRNN composed 4 network which had different size each other. MRNN-1, MRNN-2 and MRNN-4 have 4 EN and have 4, 5, 6 units of hidden layer respectively. Then MRNN-3 consists of 3 EN and 6 hidden layer. Activation function used bipolar sigmoid function in hidden layer and linear function in output layer. Trial experimentation tested that the gradient of activation are all 1, learning rate is 0.01, learning iteration have 10000 times and bias is 1. We arranged learning result on table 1.

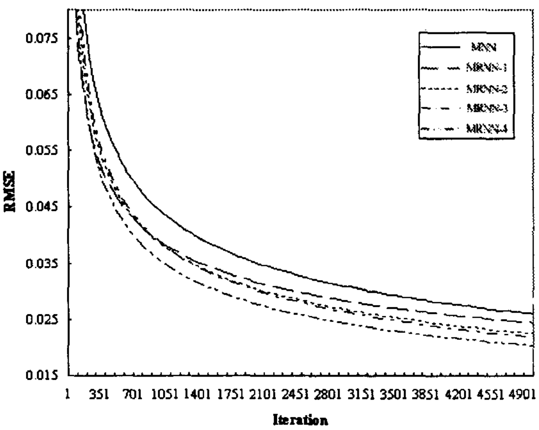

Figure 4. Training result comparison of MNN and MRNN.

Table 1. Training result of networks

Network	constitution	RMSE
MNN	Expert Net : 4 Hidden node : 6	0.0324
MRNN-1	Expert Net : 4 Hidden node : 4	0.0293
MRNN-2	EN : 4 Hidden node : 5	0.0271
MRNN-3	Expert Net: 3 Hidden node : 6	0.0266
MRNN-4	Expert Net : 4 Hidden node : 6	0.0238

VII. Conclusion

It is considered emotion like human to make robot that have similar behavior and thought like human, though there are many kind of data and pattern to study. So MNN is proposed to help defects of MLNN. MNN divided into several sub-task to solve complicate problem, and EN in MNN applied RNN that can represent situation and can perform role of feedforward network as small network.

MRNN was able to have similar effect or superiority in network which do not have more hidden node of neural network and the number of module than MNN. Through trail test, this had few learning error than MNN, through network down sizing, we saved the times. It has impressed us as RNN influence.

The later questions are to compose accurate simulator and to find optimal structure of MRNN and to show system that are more kindly between human and robot from addition of emotion.

Acknowledgement

This research was supported by the Brain Neuro-informatics Research Program. under the Korea Ministry of Science and Technology.

Reference

[1] Richard S. Lazarus, Bernice N. Lazarus, *Passion & Reason*, Oxford University Press, Inc, 1994.

[2] Simon Haykin, *Neural Networks - A Comprehensive Foundation*, Macmillian College Publishing Company Inc., 1994.

[3} CHIN-TENG LIN and C. S. GEORGE LEE, *Neural Fuzzy Systems - A Neuro Fuzzy Synergism to intelligent Systems*, Prentice Hall PTR, 1996.

[4] H. Kang and K. B. Sim, *Intelligent Information System*, Daeyoung Pub., 2000.

[5] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive Mixtures of Local Experts", Neural Computation, Vol. 6., pp. 181-214, 1994.

[6] C. Lee Giles and Sandiway Fong, "Natural Language Grammatical Inference with Recurrent Neural Networks", *IEEE Trans. on Knowledge and Data Engineering*, Vol. 12, No. 1, pp. 126-140.