Improving the Tuning Capability of the Adjusting Neural Network

Yoichi Sugita
Hitachi Research Labolatory
Hitachi Ltd., Japan
6, Kanda-surugadai 4-choume
Chiyoda-ku, Tokyo
101-8010, Japan

Kotaro Hirasawa
Graduate School of
Information, Production and Systems
Waseda University
Hibikino 2-7, Wakamatsu-ku, Kitakyushu-shi
Fukuoka-ken, 808-0135, Japan

Abstract

We have proposed the adjusting neural network (AJNN), which is the extended multi-layered network for control model tuning in the process plant. The AJNN consists of 2 networks (conventional neural network (CNN) and error calculation neural network (ECNN)) and can calculate the accurate tuning value using the ECNN which output the error of the CNN. However, the output of the ECNN does not correspond to the error of the CNN where the input value is large, and decrease the accuracy at this region. In this paper, methodology to improve the accuracy of the AJNN by introducing the procedure to control the ECNN output properly is described and its effectiveness is shown by simulation results.

1 Introduction

In the control system for the plant such as the iron steel plant or chemical plant, which has large dead time and has restricted number of sensors, the model based control is generally applied[1], which predicts appropriate control commands to be inputted to the plant using the model which identifies the plant behavior precisely. In that case, the accuracy of model directly influents the control performance. Therefore, in such control system, model tuning approach is applied and the neural network, which has the high capability to identify the non-linear relationship, is considered to be one of the effective means[2][3].

The neural network establishes the relationship between the error of controlled output ($d\theta$) and the deviation of model parameter to be modified. We have introduced the AJNN in our previous papers [4]~[7], which is extended model of conventional multi layered neural network(CNN) to increase the tuning performance. The AJNN consists of 2 networks which has the same architecture, and one of them calculates the output error of the other network at $d\theta = 0$ (called ECNN; Error Calculation Neural Network). The AJNN generates final output by subtracting the output of ECNN from the others'. The AJNN can calculate the accurate deviation of model parameter in the neighborhood of $d\theta=0$, by eliminating the network error at $d\theta=0$. However, since the ECNN can calculate the network error only at $d\theta=0$ and the AJNN subtracts the output of the ECNN for all region of $d\theta$, accuracy of the output of the AJNN at the region far from $d\theta=0$ may decrease.

In this paper, we discuss the method to increase the AJNN capability where $d\theta$ is large using the non linear function which restrict the output of the ECNN we have proposed before[5]~[7] and describe the optimization of this function.

2 Model Tuning for Reheating Furnace Plant

Reheating furnace plant is one of the typical plant of which the model based control is effecive in the control. The pourpose of this system is to raise the slab temperature to approximetry 1200 centigrate for the rolling process. For this, appropriate furnace temperatures to obtain the target slab temperature are calculated using the control model, which estimate the plant behaivior precicely using the phisical equiations. The accuracy of the model directly effects the control quality, therefore, the model has to be modified corresponding to the plant behaivior change. Fig.1 shows the structure of the model tuning system for the reheating furnace plant. When the slab is extraced, a model tuner modifies one parameter (Φcg) included in the control model and correct the model behaivior using the difference between the detected temperature value and the target one ($d\theta_{out}$), and several state variables such as the furnace temperature. Since the relationship between the deviation of the Φcg ($\Delta\Phi cg$) and $d\theta_{out}$ is basicaly non-linear and the several variables affect this

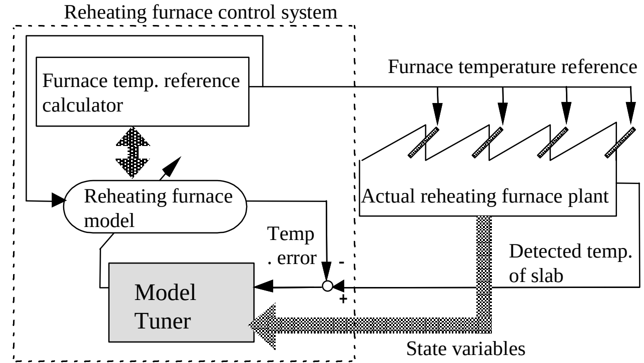

Fig.1. Control system for reheating furnace plant.

relationship, the CNN is considered to be effective to identify this relationship and realizes the model tuner.

3 The Adjusting Neural Network

Let the outputs of the model tuner by given by $F(d\theta_{out}, x_1, \ldots, x_n)$, where $x_1, \ldots, x_n$ are state variables. Since desired output of the model tuner is obviously zero when $d\theta_{out}$ is equal to zero, the model tuner must always satisfy the following equiation.

$$F(0, x_1, \ldots, x_n) = 0 \qquad (1)$$

This is a constrain condition for the accurate parameter tuning. However, if the CNN is utilized as the model tuner, it is quite difficult to realize the Eq.1 permanently because of infinite training data and the learning error. The output error at $d\theta_{out} = 0$ cause the steady state error when the tuning is converged.

Fig. 2 shows the structure of the AJNN. It has an ECNN(Error Calculation Neural Network) added in parallel to the CNN, which has the same structure of the CNN. The ECNN receive zero instead of $d\theta_{out}$ and calculate the output error of the CNN at $d\theta_{out}$. By subtracting the output of the ECNN from the one of the CNN, an output of the AJNN, $F_{AJNN}(d\theta_{out}, x_1, \ldots, x_n)$ is given by

$$F_{AJNN}(d\theta_{out}, x_1, \ldots, x_n) = F_{NN}(d\theta_{out}, x_1, \ldots, x_n) - F_{NN}(0, x_1, \ldots, x_n) \qquad (2)$$

Where, $F_{NN}(d\theta_{out}, x_1, \ldots, x_n)$ is the output of the CNN and $F_{NN}(0, x_1, \ldots, x_n)$ is the one of the ECNN.

When the $d\theta_{out}$ is zero, Eq.2 indicates the following equation, it is obvious the AJNN satisfies the Eq.1 permanently and completes the tuning without the steady state errors.

$$F_{AJNN}(0, x_1, \ldots, x_n) = 0 \qquad (3)$$

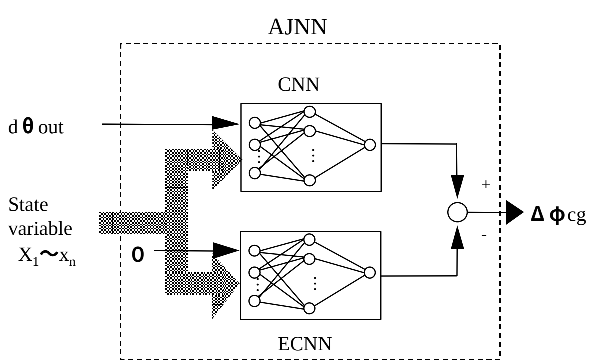

Fig.2. Architecture of AJNN.

4 A Problem of the AJNN

The output of the ECNN directly corresponds to the output error of the CNN at $d\theta_{out} = 0$, but not where $d\theta_{out}$ is not equal zero. Therefore, when the $d\theta_{out}$ is large, the output of the AJNN may possibly decrease its accuracy by subtracting the output of the ECNN which does not correspond to the error of the CNN. The problem of the AJNN is the decreasing accuracy where $d\theta_{out}$ is large cause the increase of tuning number compared to the CNN.

5 Non Linear Function

5.1 The AJNN with the Non Linear Function

To solve the problem discussed above, we have introduced the AJNN with the non linear function (see Fig.4). The proposed AJNN has the non linear function, which restrict the output of the ECNN corresponding to the volume of the $d\theta_{out}$, added at the output side of the ECNN and increase the accuracy of output where the $d\theta_{out}$ is large. The output of the AJNN with this function is given by

$$F_{AJNN}(d\theta_{out}, x_1, \ldots, x_n) = F_{NN}(d\theta_{out}, x_1, \ldots, x_n) - \alpha(d\theta_{out}) \cdot F_{NN}(0, x_1, \ldots, x_n) \qquad (4)$$

Where $\alpha(d\theta_{out})$ is the non linear function.

The non linear function should have the smooth shape, however, we assume the piecewise function to facilitate the further discussion.

$$\alpha(x) = \begin{cases} 0 & (x < -A/2 - T) \\ \frac{1}{T}(x + A/2 + T) & (-A/2 - T \leqq x < -A/2) \\ 1 & (-A/2 \leqq x \leqq A/2) \\ -\frac{1}{T}(x - A/2 - T) & (A/2 < x < A/2 + T) \\ 0 & (A/2 + T < x) \end{cases} \qquad (5)$$

where A is the size of active region and T is the size of transient region.

In the active region, the output of the ECNN is active and in the transient region, the output is decreased gradually. By utilizing this function, the AJNN can calculate the accurate tuning value independently of the volume of $d\theta_{out}$.

5.2 Optimization of the Non Linear Function

To maximize the performance of the AJNN, the non linear

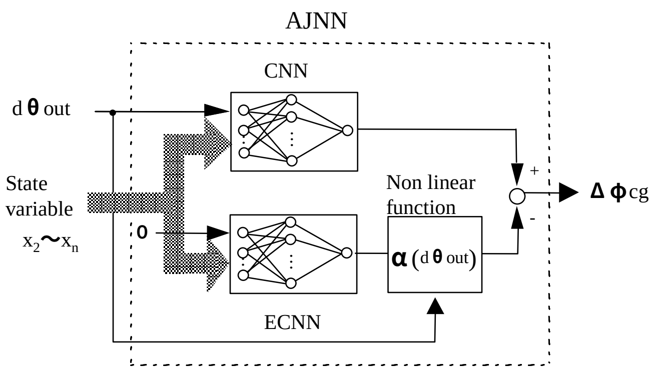

Fig.3. AJNN with a Non-linear Function.

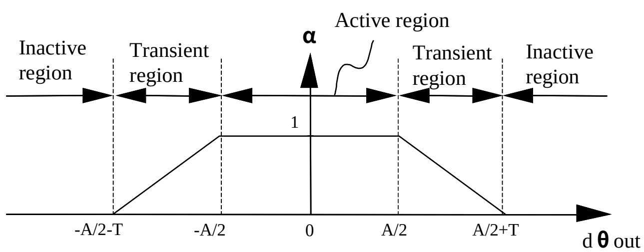

Fig.4. Shape of non-linear function

function has to be optimized, that is, methodology for determination of the size of active region has to be discussed. We propose the methodology to utilize the training data effectively. The determination procedure is described as follows.

Step1; Train the CNN using the training data with BP method.

Step2; Set the initial value(usually 1[℃]) for the size of active region.

Step3;For all training data, evaluate the error between the output of the AJNN with the non linear function and the training data.

Step4; Increment the size of active region by predetermined step width, and find a active region which minimizes the error described at Step 3.

This method is expected to produce the AJNN which minimizes the errors of the training data for all region of $d\theta_{out}$ and even decreases the errors compared to the CNN. That is, the AJNN with the non linear function decided by this procedure can identify the relationship between $d\theta_{out}$ and $\Delta\Phi cg$ described with the training data as much accurately as possible.

6 Simulation Results

6.1 Simulation Conditions

Inputs for the AJNN in this simulation are the current Φcg the control model has, $d\theta_{out}$ and six state variables (the four furnace temperature, an initial slab temperature, thickness of slab). Output of the AJNN in this simulation is the tuning ratio P for the current Φcg. The current Φcg is modified into $(1 + P)\ \Phi cg$ when the slab is extracted from the reheating furnace plant and the output temperature is detected. 672 training data to train the network are prepared by simulation.

6.2 Results and Discussion

Fig.5 shows an example of simulation results. A linear equation, which is a conventional tuning method and identify the relation between $d\theta_{out}$ and P by the linear equation. The linear equation, which can not identify the non linear relation , can not output appropriate tuning value and took 5 times to converge. The CNN can converge quickly but finished the tuning to remain the steady state errors because of the output error at $d\theta_{out} = 0$. On the other hand, the AJNN converged by 2 tuning numbers without any steady state error.

Fig.6 indicates the error behavior against the size of active region of the non linear function. The mean error in vertical axis is mean of |training data – output of the CNN|. The size of transient region is fixed at 20[℃]. A size of active region which minimizes the errors is 24[℃], and the error raises gradually corresponding to the increase of the size of active region. Effect of the ECNN is maximized at 24[℃] and after that, the ECNN decrease the accuracy of the output of the CNN gradually. As a result of this examination, 24[℃] is considered to be the best size of active region for the non linear function. Besides, the error at 24[℃] is detected to be lower than the one of the CNN and also the AJNN without the non linear function.

The comparison of tuning performance between the AJNNs which have different size of the non linear function are described in Fig. 7. 30 different conditions which were not used as the training data were applied. A tuning number at 24[℃], which is decided as the size of active region for the non linear function is smallest. Since the $d\theta_{out}$ was less than 200[℃] in this simulation, a tuning number at 200[℃] is exactly equal to the result of the AJNN without the non linear function. The tuning number of the proposed AJNN is obviously lower than the one of the conventional AJNN, and it tells the methodology discussed above can increase the performance of the AJNN and can determine the appropriate the non linear function.

7 Conclusion

We discussed the improvement of the tuning capability of the AJNN, which can modify the control model immediate and accurately for the process control. A methodology to optimize the size of non linear function is introduced, which is a procedure to found the size to minimize the training data error, and its effectiveness is evaluated and found to work effectively to decrease the tuning numbers by simulation result for the reheating furnace plant.

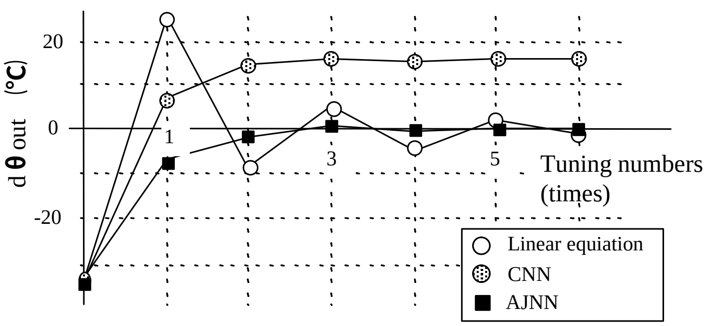

Fig.5. An example of simulation result.

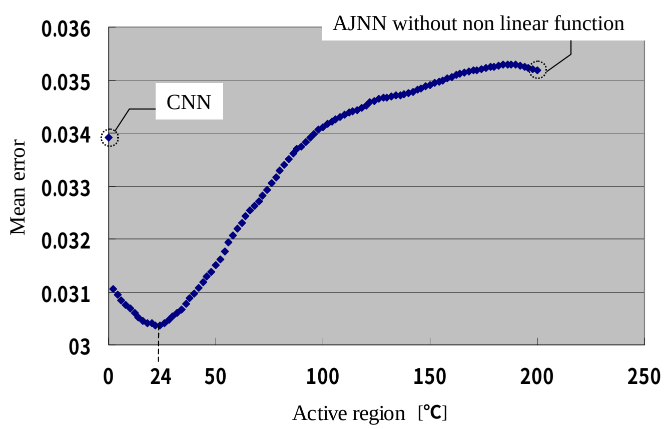

Fig.6. Error behavior against the size of active region

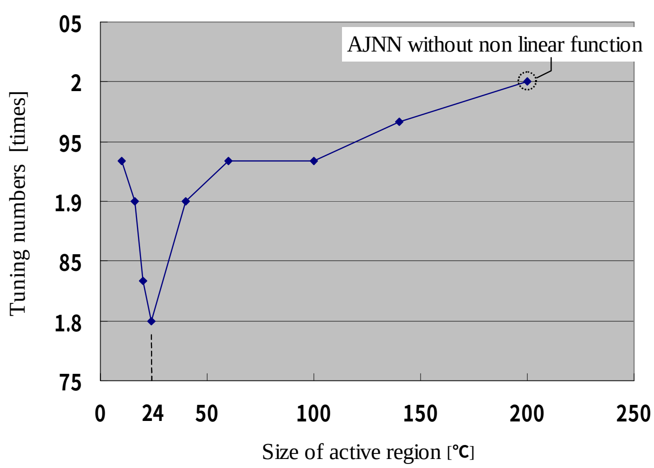

Fig.7. Comparison of tuning performance

References

[1] D.W.Clarke and P.S.Tuffs, "Properties of Generalized Predictive Control", Automatica, Vol.25, pp. 859-875, 1989.

[2] D. Rumelhart et al., "Parallel Distributed Processing", MIT Press, Cambridge, MA, 1986.

[3] W.T. Miller et al., "Neural Networks for Control", MIT Press, Cambridge, MA, 1990.

[3] Y. Sugita et al., "Architecture and Training Algorithm of the Adjusting Neural Network for Accurate Model Tuning", T.IEE of Japan, Vol.115-D,pp461-469, 1995.

Y. Sugita et al., "Extending architecture of the Adjusting Neural Network for Improving Tuning Capability", T.IEE of Japan, Vol.116-C,pp859-866, 1995.

[5] M. Kayama et al, "Adjusting Neural Networks for Accurate Control Model Tuning", FUZZ-IEEE/IFES'95, pp.1995-2000, 2001.

[6] Y. Sugita et al., "Improving Tuning Capability of the Adjusting Neural Network", IEEE IJCNN'97, pp.761-767, 1995.