

Classification of the Types of Defects in Steam Generator Tubes using the Quasi-Newton Method

Joonpyo Lee*, Nam H. Jo[†] and Youngsu Roh*

Abstract – Multi-layer perceptron neural networks have been constructed to classify four types of defects in steam generator tubes. Three features are extracted from the signals of the eddy current testing method. These include maximum impedance, phase angle at the point of maximum impedance, and an angle between the point of maximum impedance and the point of half the maximum impedance. Two hundred sets of these features are used for training and assessing the networks. Two approaches are involved to train the networks and to classify the defect type. One is the conjugate gradient method and the other is the Broydon-Fletcher-Goldfarb-Shanno method which is recognized as the most popular algorithm of quasi-Newton methods. It is found from the computation results that the training time of the Broydon-Fletcher-Goldfarb-Shanno method is much faster than that of the conjugate gradient method in most cases. On the other hand, no significant difference of the classification performance between the two methods is observed.

Keywords: Eddy current testing, Steam generator, Neural network, Quasi-Newton method

1. Introduction

Due to its diagnostic advantages, the eddy current testing (ECT) method has been extensively used for the periodic testing of steam generator (SG) tubes to ensure the safe operation of nuclear power plants[1]-[7]. ECT is one of popular non-destructive testing methods. The basic principle of ECT is based on a phenomenon of electromagnetic induction. When an alternating current is flowing in the ECT probe coil which is moving inside the electrically conductive SG tube, the magnetic field induced by the probe current interacts with the tube and generates an eddy current. The eddy current also creates a magnetic field which has an opposite direction to the primary field of the coil. One of the most important factors to determine the characteristics of the eddy current is the tube impedance. Since the impedance is directly related to the eddy current path, the impedance can be changed by the deformation of the path which is typically caused by the tube defect. It is possible, therefore, to predict the defects by measuring the characteristics of the current flowing in the probe coil.

There are general difficulties inspecting the tube defect by the ECT method. For example, the ECT signals may be distorted at the presence of noise or interfered by unwanted signals which are generated from the geometry variations of support structures and probe wobble. For the reliable inspection of the tube defect, therefore, it is extremely important to interpret the ECT signals with enhanced accuracy and consistency.

The ECT defect signals may be analyzed by a number of techniques to obtain information on the defect type and size. Among those techniques, an artificial neural network (ANN) is recognized as one of the most popular techniques. The methods to classify the defect type have been studied based on the ECT signals by means of a probabilistic neural network[6] and multi-layer perceptron (MLP) neural networks[7]. In [7], a conjugate gradient back-propagation (CGBP) algorithm has been used for training a MLP network to classify the defect type. It should be noted that the classification performance of the back-propagation (BP) algorithm is dependent on a set of initial guesses of weights[8]. For a reliable assessment of the classification performance of the network, therefore, it is necessary to train the network using sufficiently many sets of initial guesses.

It is also required to test the network based on various neuron numbers in the hidden layer because these numbers may have an effect on the classification performance. Consequently, a huge amount of computation time may be needed to train the network using the BP algorithm for the reliable evaluation of the defect classification performance. As will be discussed later, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is known as the most popular gradient-based method[9]. In the present article, therefore, BFGS algorithm is employed for training the neural network to improve the training time of the BP algorithm.

This paper is organized as follows. The next section involves a description of four axisymmetric defects and three feature vectors for classifying the defect type. In Sec. 3, the architecture of the MLP network is presented and CGBP and BFGS back-propagation (BFGSBP) algorithms to train the neural network are explained. In Sec. 4, the simulation results are discussed to compare training time and classifi-

[†] Corresponding Author: Dept. of Electrical Engineering, Soongsil University, Seoul, Korea. (nhjo@ssu.ac.kr)

* Dept. of Electrical Engineering, Soongsil University, Seoul, Korea.

cation performances of the both algorithms. Finally, the conclusion of the paper is presented in Sec. 5.

2. Axisymmetric Defects and Feature Vectors

2.1 Axisymmetric Defects

As shown in Fig. 1, the following four types of axisymmetric defects in the SG tube wall are considered in this paper: (1) the I-type defect in the inner wall (I-In), (2) the I-type defect in the outer wall (I-Out), (3) the V-type defect in the inner wall (V-In), (4) the V-type defect in the outer wall (V-Out). For the classification of these defect types using an ANN, a sufficient amount of ECT defect signal samples are required to train a MLP network. However, it is difficult to obtain such amount of defect signals from actual SG tubes. It has been revealed in [6] that the defect signals can be numerically generated based on the finite element model and these signals are almost identical to the signals acquired from actual SG tubes. In this paper, the ECT signals for the axisymmetric defects are also numerically generated using the same method of [6] and these signals are used for training the MLP network and assessing the classification performance as well. A total of 50 defect signals are created for each defect type by changing the defect size (depth and width) in various ways. The depth and the width are changed within a range of 0.1 ~ 1.0 mm with a step size of 0.1 mm and a range of 0.2 ~ 1.0 mm with a step size of 0.2 mm, respectively.

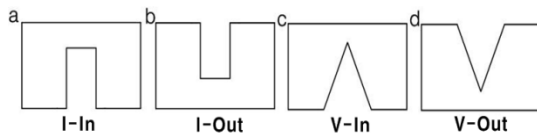


Fig. 1. Four axisymmetric defect types.

2.2 Feature Vectors

The types of feature vectors to be extracted from ECT signals play a crucial role determining the performance to classify the defect type. In [6], a total of 22 features, such as maximum resistance, phase angle at the point of maximum resistance, phase angle at the starting point of the signal, and so on, have been extracted in the impedance plane to classify the defect pattern. In [7], a significant improvement of the classification performance has been realized using only three features which are different from those in [6]. In this paper, therefore, these three features are also employed to classify the defect type. Fig. 2 depicts the three feature vectors. Here F1, F2, and F3 are defined as maximum impedance, phase angle at the point of maximum impedance, and an angle between the point of maximum impedance and the point of half the maximum impedance, respectively.

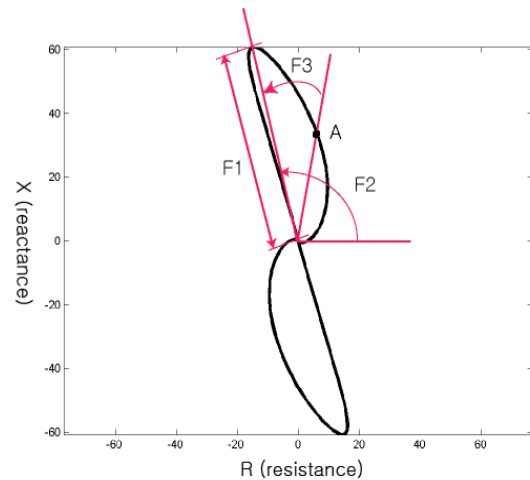


Fig. 2. Feature vectors extracted from the ECT signals in the impedance plane to classify four defect types.

3. Neural Network

Using the three feature vectors, four independent MLP neural networks are constructed to classify the four defect types. Each network consists of a hidden layer, three input neurons and an output neuron. As an activation function, the following log-sigmoid function is adopted in the hidden layer[8]:

$$\phi(v) = \frac{1}{1 + \exp(-v)} \quad (1)$$

In the MLP network for the I-In type defect, a classifier is trained by setting a target value of I-In defect to 1 and the target values of other defects to 0. For the classification of other defects, other classifiers are trained in a similar way as in the I-In defect classifier. In this paper, the MLP networks are trained by means of the CGBP algorithm as well as the BFGSBP algorithm. The simulation results of both algorithms will be examined in terms of training time and classification performance. Before proceeding to the discussion of details regarding the simulation results, a brief description of CGBP and BFGSBP algorithms will be presented to understand how the networks are trained by them.

3.1 Conjugate Gradient Algorithm

The basic BP algorithm adjusts the weights in the steepest descent direction (negative of the gradient). This is the direction in which the performance function is decreasing most rapidly. It turns out that, although the function decreases most rapidly along the negative of the gradient, this does not necessarily produce the fastest convergence. In the conjugate gradient algorithm a line search is performed along conjugate directions, which produces generally faster convergence than steepest descent directions. The CGBP

algorithm implemented in the MLP network of this paper is shown in Fig. 3. In step 1, N and ϵ denote the number of iteration and the tolerance for stopping criterion.

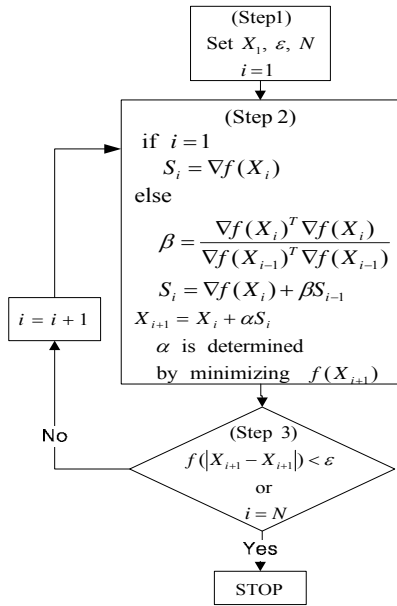


Fig. 3. Conjugate Gradient Algorithm.

3.2 BFGSBP Algorithm

Newton’s method locates the minimum of a quadratic function in one iteration, and usually produces faster convergence than steepest descent[9]. In Newton’s method, it is assumed that the function can be locally approximated as a quadratic Taylor expansion in the region around the optimum. Since Newton’s method requires the Hessian matrix of second derivatives of the function to find the stationary point, this method is not very effective for functions with large numbers of variable. In contrast, it is not necessary in quasi-Newton methods to compute the Hessian matrix at any stage. The Hessian matrix is approximated by a matrix A (so-called metric) using only first partial derivatives of the function. The approach involves starting with an initial metric matrix and updating and improving it at each iteration. There are two primary methods of this type: Davidon-Fletcher-Powell (DFP) and BFGS algorithms. The two algorithms differ in how they handle convergence issue. That is, the metric matrix in the former converges to the inverse Hessian matrix but that in the latter converges to the Hessian matrix. The BFGS method is generally recognized to be superior in most cases[9].

The BFGSBP algorithm to obtain the minimum of an objective function $f(X)$ consists of three main steps as shown in Fig. 4. In step 1, A_i denotes an initial metric matrix. From an initial guess of X_i and an initial metric A_i chosen in step 1, step 2 and step 3 are repeated until X converges to the target value. In step 2, the search direction S_i at the i -th stage is solved and a line search is performed to find α_i and update X_{i+1} . In step 3, the metric A_i at the i -th stage is updated by adding two matrices B and C .

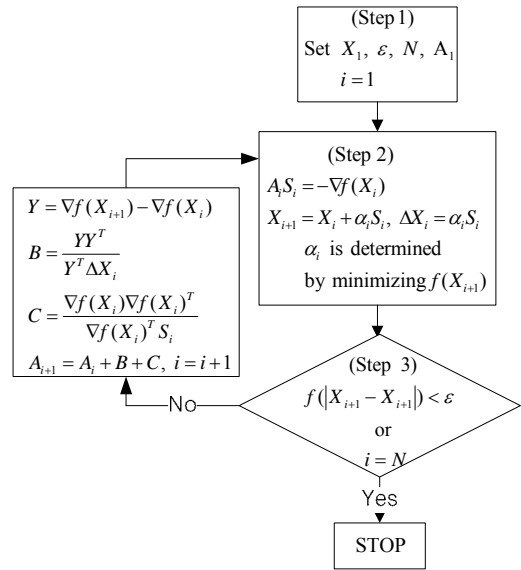


Fig. 4. BFGS Algorithm.

4. Computer Simulation

Fig. 5 illustrates the flowchart for computation. First of all, a total of 200 ECT defect signals are numerically generated. 200 sets of three feature vectors (F1, F2, F3) are extracted from these ECT signals and divided into two data groups. One group includes 100 sets of feature vectors extracted from the ECT signals of defects whose depths are 0.1, 0.3, 0.5, 0.7, and 0.9 mm. The other group contains the other 100 sets of feature vectors extracted from the ECT signals of defects whose depths are of 0.2, 0.4, 0.6, 0.8, and 1.0 mm. The former are the training data to optimize the weights and biases of the networks. The training data are also used to measure the training time of the two MLP classifiers based on CGBP and BFGSBP algorithms. The latter are the test data to assess the classification performance of the two trained classifiers.

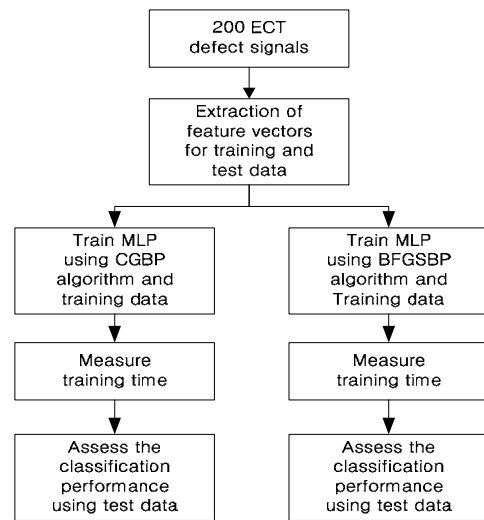


Fig. 5. Flowchart for computer simulation.

The purpose to train a neural network is to find a set of weight (and bias) values that make the output of the network to match the actual target values as closely as possible. In this paper, the MLP classifier is trained until the mean square error (MSE) reaches a tolerance for stopping criterion. Here MSE is expressed by

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - \tilde{x}_i)^2 \quad (2)$$

where N is the total number of sample data, x_i and \tilde{x}_i are the target value and the MLP output of the i -th data, respectively. MSE is also employed to compute the classification performance of the trained classifier. The two algorithms are implemented in a personal computer (core2 duo processor @ 2.33 GHz).

Since the number of neurons in the MLP hidden layer (which will be denoted by ‘NEU’, henceforth) is one of crucial factors to determine the classifier performance[8], it is important to examine the effect of NEU on the network performance. In this regard, a variety of NEUs (10, 20, 30, 40, 50 and 100) are involved in the computation of the training time and the classification performance. As a matter of fact, no significant difference among those NEUs has been observed in terms of the training time and the classification performance. To avoid redundant description, therefore, the computation results for two cases (NEU=30, 100) are presented in the following.

Fig. 6 shows the training time with respect to the defect type when NEU is 30. Here, dotted and solid lines denote training time of CGBP and BFGSBP, respectively. As mentioned previously, it is required to train the network as many as possible to reliably evaluate the performance of the network. In this paper, the networks are trained using 30 sets of initial guesses of weights which are randomly generated in a range of -0.5 and 0.5. Accordingly, 30 training time are included in the data line of each defect type. Each training time is calculated based on Eq. (2). In Fig. 6, the bar in the middle of the data line represents the mean value of the 30 training time. As can be seen, the training time of BFGSBP is faster than that of CGBP. It can be proved by calculating an average of training time for four

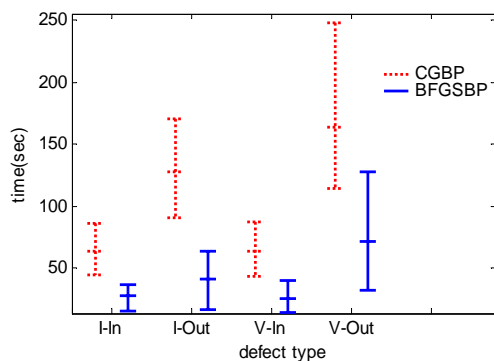


Fig. 6. Training time of CGBP and BFGSBP with respect to defect type (NEU=30).

defects that the average training time of BFGSBP is approximately 2.5 times faster than that of CGBP.

It is also possible to compare the training time of the two algorithms in terms of the convergence rate (MSE versus epoch) as shown in Fig. 7. Here the convergence rates are calculated for only V-Out defect training data that are in close proximity of the mean values in Fig. 6. It is clear in Fig. 7 that the convergence rate of BFGSBP is much faster than that of CGBP.

Once the training process of a classifier is completed, the test data are used to evaluate the classification performance of the trained classifier. The computation results regarding the classification performances of CGBP and BFGSBP are given in terms of MSE with respect to the defect type in Fig. 8. It seems that the MSE and variation width (length of data line) of CGBP is somewhat better than those of BFGSBP in the cases of I-out and V-Out defects. However, the difference is negligibly small. That is, no significant difference exists between classification performances of CGBP and BFGSBP.

Figs. 9 and 10 show the simulation results when NEU is 100. As in the case that NEU is 30, 100 sets of training data are used for each defect type in Fig. 9. It can be easily validated in Fig. 9 that average training time of BFGSBP and

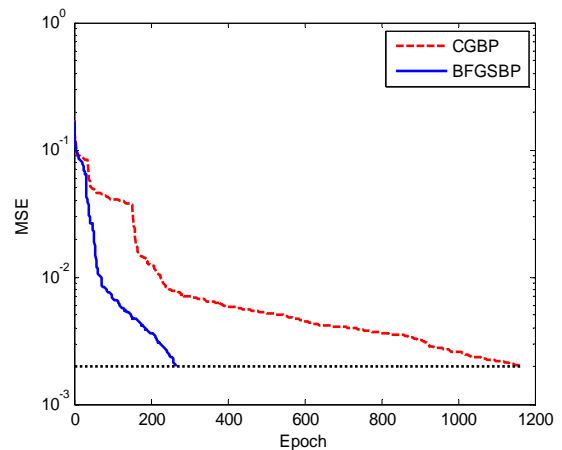


Fig. 7. Convergence rate of CGBP and BFGSBP for the classification of V-Out defects (NEU=30).

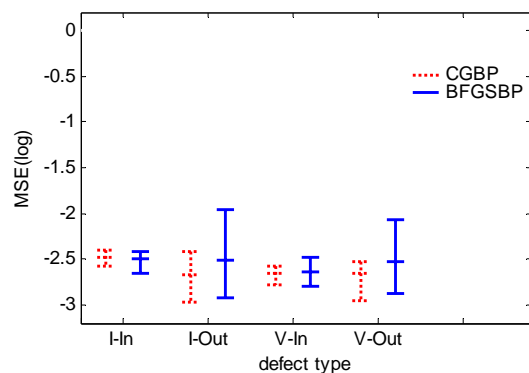


Fig. 8. Classification performances of CGBP and BFGSBP with respect to defect type (NEU=30).

CGBP for all defects are 90.4 and 261.6 sec, respectively. That is, the average training time of BFGSBP is approximately 2.9 times faster than that of CGBP.

The comparison of classification performances of CGBP and BFGSBP are given in Fig. 10. As in Fig. 8, it is difficult to point out the performance difference between the two algorithms. It may be addressed, therefore, that the classification performance of BFGSBP is almost identical to that of CGBP regardless of NEU.

As can be seen in Figs. 6 and 9, training time of the two algorithms change in a similar manner as NEU is increased from 30 to 100. This is quite an unexpected result because the size of the metric matrix A in the BFGSBP algorithm equals $[(n+m+1)NEU+m] \times [(n+m+1)NEU+m]$ where n and m are the numbers of input and output vectors, respectively [8]. That is, the size of A increases in proportion to the square of NEU as NEU increases. It should be noted that the computation of the inverse of A is required to find the search direction S_i in the step 2 in Fig. 4. In other words, an increase of NEU may give rise to a big increase of computation time of the inverse of A . Consequently, when the BFGSBP algorithm is employed in a neural network that has a large number of hidden neurons, the speed to train the network may become extremely slow compared to the CGBP algorithm. In order to quantitatively examine the effect of NEU on the training time, we compute the training time of the classifiers as various values such as 20, 30,

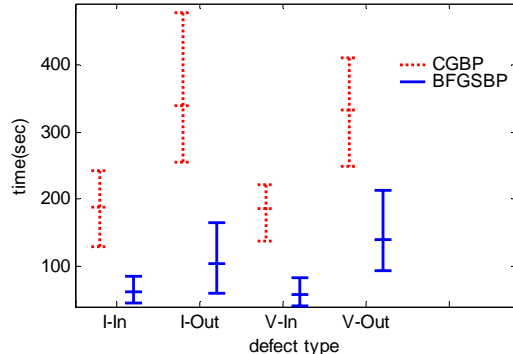


Fig. 9. Training time of CGBP and BFGSBP with respect to defect type (NEU=100).

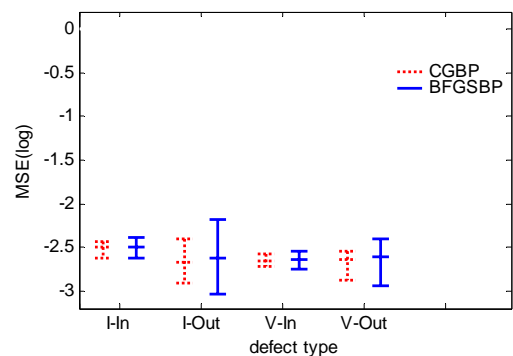


Fig. 10. Classification performances of CGBP and BFGSBP with respect to defect type (NEU=100).

40, 50, 100, 150, and 200 are selected to be NEU. Figure 11 shows an example of the computation results as to how the training time depends on NEU. Here only V-In defect classifier is tested to compute the training time. As can be seen, the average training time of BFGSBP is much better than that of CGBP when $NEU \leq 100$. However, such tendency is not sustained at $NEU=200$. In Fig. 11, average training time of BFGSBP and CGBP are computed to be 1278.1 and 698.4 sec, respectively. Unlike the case that $NEU \leq 100$, the average training time of BFGSBP is much slower than that of CGBP. Apparently, this change is related to the dependence of the size of A on NEU as explained previously.

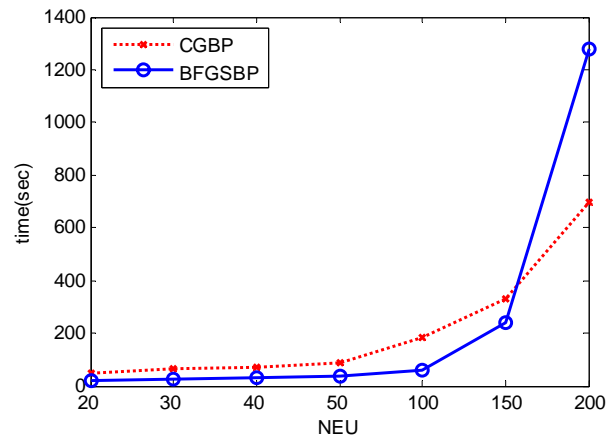


Fig. 11. Training time of CGBP and BFGSBP for V-In defect classifier with respect to NEU.

5. Conclusion

Since it is extremely important to reliably inspect the SG tube defect, MLP networks have been constructed to accurately classify four axisymmetric defects based on the ECT signals. The 100 sets of three feature vectors extracted from the ECT signals have been used to train the MLP networks. BFGSBP and CGBP algorithms have been employed in the training process. The other 100 sets of three feature vectors have been used to test the trained classifiers. According to the computation results, the BFGSBP algorithm is capable of training the MLP classifier in a high speed compared to CGBP when $NEU \leq 100$. Moreover, the classification performance of BFGSBP is similar to that of CGBP. Considering that NEU is typically less than 100 for a neural network to classify the SG tube defect pattern, it may be addressed that, for the classification of SG tube defects, the BFGSBP algorithm is a beneficial and effective optimization method compared to the CGBP algorithm.

References

- [1] E.E. Kriezis, T.D. Tsiboukis, S.M. Panas, J.A. Tego-poulos, "Eddy currents: Theory and applications,"

Proceedings of the IEEE, Vol. 80, No. 10, pp. 1559-1589, 1992.

- [2] H.B. Lee, and D.H. Kim, "Impedance calculation for a plate with crack in eddy current NDT using 3D indirect BIEM," *IEEE Trans. Magnetics*, Vol. 36, No. 5, pp. 3131-3133, 2000.
- [3] M. Rebican, Z. Chen, N. Yusa, L. Janousek, and K. Miya, "Shape reconstruction of multiple cracks from ECT signals by means of a Stochastic method," *IEEE Trans. Magnetics*, Vol. 42, No. 4, pp. 1079-1082, 2006.
- [4] C. Gangzhu, A. Yamaguchi, and K. Miya, "A novel signal processing technique for eddy-current testing of steam generator tubes," *IEEE Trans. Magnetics*, Vol. 34, No. 3, pp. 642-648, 1998.
- [5] H. Haoyu, and T. Takagi, "Inverse analyses for natural and multicracks using signals from a differential transmit-receive ECT probe," *IEEE Trans. Magnetics*, Vol. 38, No. 2, part 1, pp. 1009-1012, 2002.
- [6] S.J. Song and Y.K. Shin, "Eddy current flaw characterization in tubes by neural networks and finite element modeling," *NDT & E International*, Vol. 33, pp. 233-243, 2000.
- [7] N.H. Jo and H.B. Lee, "A novel feature extraction for eddy current testing of steam generator tubes," *NDT & E International*, vol. 42, pp. 658-663, 2009.
- [8] S. Haykin, *Neural Networks*, New Jersey: Prentice-Hall, 1999.
- [9] P. Venkataraman, *Applied Optimization with MATLAB Programming*, Wiley-Interscience, 2002.



Joon-Pyo Lee received his B.S Degree in Electric Engineering from SoongSil University, Korea in 2009 and currently, he is an M.S student in SoongSil University, Korea. His research interests include pattern recognition, and system control.



Nam H. Jo received the B. S., M. S., and Ph. D. degrees from School of Electrical Engineering, Seoul National University, Seoul, Korea, in 1992, 1994 and 2000, respectively. From 2001 to 2002, he worked as a Senior Research Engineer at Samsung Electronics, Suwon, Korea. Since March

2002, he has been with the Department of Electrical Engineering at Soongsil University, Seoul, Korea, where he is currently an associate professor. His research interests include nonlinear systems control theory and its applications to systems biology, and artificial intelligence with its application to defect pattern classification in steam generator tubes. Dr. Jo was a recipient of the Distinguished Teaching Award presented by Soongsil University.



Young-Su Roh received his B.S. and M.S. degrees in Electrical Engineering from Seoul National University in 1984 and 1986, respectively. He received a Ph.D. degree in Applied Science from the University of California, Davis in 2001. From 1988 to 1996, he worked at the Korea Electricity Research Institute.

He is currently an Associate Professor at the Department of Electrical Engineering at Soongsil University. His research interests include plasma physics, nuclear fusion, and electrical discharges.