Mirror Image Learning for Autoassociative Neural Networks

Shusaku SHIMIZU, Wataru OHYAMA, Tetsushi WAKABAYASHI and Fumitaka KIMURA
Faculty of Engineering, Mie University
1515 Kamihama-cho, Tsu, Mie, 514-8507 Japan

Abstract

This paper studies on the mirror image learning algorithm for the autoassociative neural networks and evaluates the performance by handwritten numeral recognition test. Each of the autoassociative networks is first trained independently for each class using the feature vector of the class. Then the mirror image learning algorithm is applied to enlarge the learning sample of each class by mirror image patterns of the confusing classes to achieve higher recognition accuracy.

Recognition accuracy of the autoassociative neural network classifier was improved by the mirror image learning from 98.76% to 99.23% in the recognition test for handwritten numeral database IPTP CD-ROM1 [1].

1. Introduction

The mirror image learning was proposed in [2] and was studied to apply to a linear classifier and the subspace method [3][4][5]. The algorithm generates a mirror image of a pattern that belongs to one class of a pair of confusing classes and utilizes it as a learning pattern of the other class. The mirror image of a pattern is generated in respect to the mean vector of the linear classifier, and the minimum mean square error hyperplane (subspace) of the subspace method.

This paper studies on the mirror image learning algorithm for the autoassociative neural networks [6] and evaluates the performance by handwritten numeral recognition test. In the case of the autoassociative neural networks, the mirror image is generated in respect to the associated output of a neural network.

2. Autoassociative neural networks

Autoassociative neural network is a network having the same number of neurons in input and output layers, and the less in the hidden layers. The network is trained using the input vector itself as the desired output. This training leads to organize a compression/encoding network between the input layer and the hidden layer, and a decoding network between the hidden layer and the output layer (**Figure 1**).

Each of the autoassociative networks is trained independently for each class using the feature vector of the class. As a result, the squared error between an input and the output is generally minimized by the network of the class to which the input pattern belongs. This property enables us to classify an unknown input pattern. The unknown pattern is fed to all networks, and is classified to the class with minimum squared error [6].

In contrast with the pattern recognition using the mutual associative networks, each autoassociative network is organized independently for each class, and the training load of the networks can be distributed to handle large-scale pattern recognition problems.

Figure 2 shows how to classify an input pattern by the autoassociative networks. The number of neurons in the input layer, the hidden layer, and the output layer of each network is 2, 1, 2, respectively for a class. Two line segments in this figure represent the trace of the output that are obtained by sweeping the output of hidden layers from 0 to 1. Squares on the line segments represent the projections of the samples. These line segments correspond to the principal axes of the distributions. An input pattern (x_1, x_2) is given to the network of each class. The output (y_1, y_2) and (z_1, z_2) are the coordinates of Y, and Z on the projection line respectively, and the squared distance are given by $\|Y - X\|^2$, and $\|Z - X\|^2$ respectively. The input pattern X is classified to the class with the minimum distance, i.e.

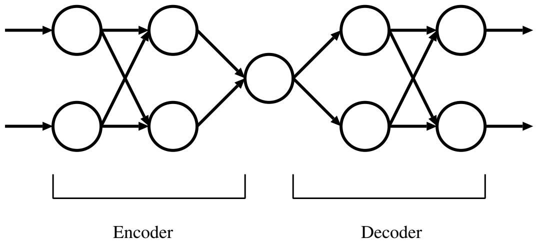


Figure 1. Autoassociative neural network.

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 IEEE

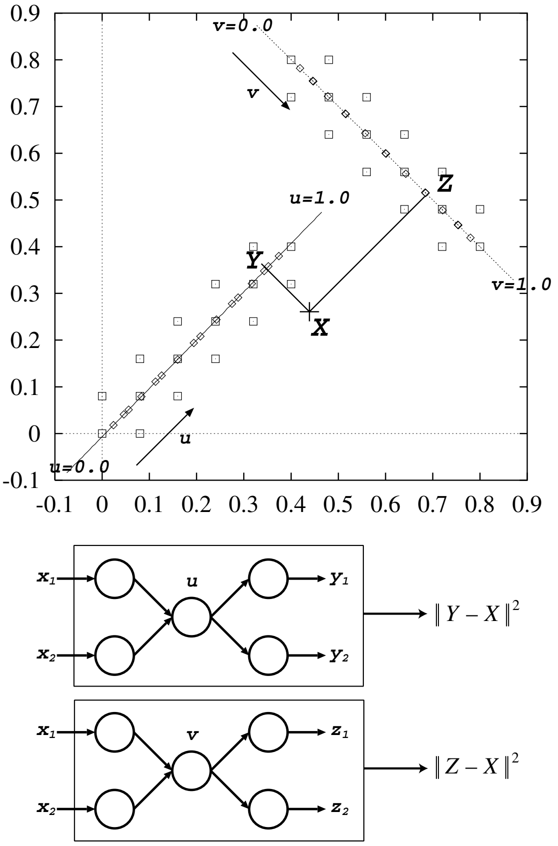


Figure 2. Discrimination of unknown pattern by three-layered autoassociative neural networks.

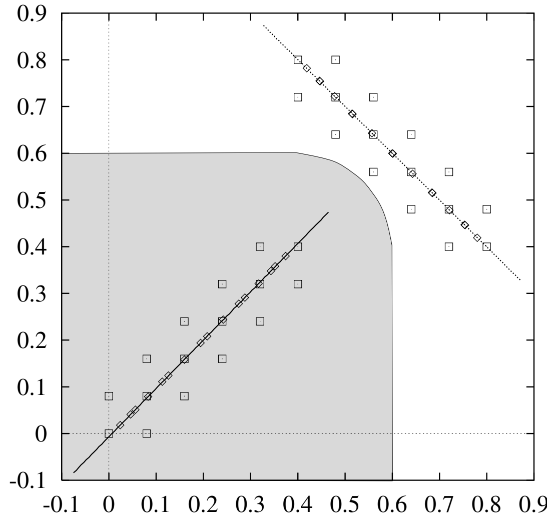


Figure 3. Decision boundary of three-layered autoassociative neural networks.

the class which minimizes the square error of the input and the output of the network, e.g. the left lower class in this figure. The output u and v from the hidden layers take the value from 0 to 1.

Figure 3 shows the decision boundary for this example.

In the rest of this paper, we focus on the mirror image learning for three layered autoassociative networks without loss of the extendability to five layered autoassociative networks.

3. Corrective learning by mirror image

3.1. Generation and learning of mirror images

When a pattern X of a class (class 1) is misclassified to a different class (class 2), the minimum mean square error hyperplane/hypersurface of the class 2 is kept away from the pattern X as follows (**Figure 4**).

The mirror image X' of the pattern X in respect to the hyperplane of class 2 is given by

$$X' \quad = \quad 2Y - X, \tag{1}$$

where Y is the output of the network, which is the projection of X to the hyperplane. This mirror image X' is used as a learning pattern of class 2 to keep the hyperplane away from the pattern X.

While the mirror image X' is used as a learning pattern of class 2, the pattern X itself is also used as a learning pattern of class 1 to move the hyperplane of class 1 toward the pattern X. As a result the decision boundary shifts to the side of class 2. After the pairs of X and X' for all misclassified patterns are successively used as learning patterns, then the same procedure is repeated until there is no change in the number of misclassifications.

3.2. Mirror image learning with margin

If the number of misclassified pattern is too small the mirror image learning quickly converges close 100% correct recognition for the learning sample, before the recognition rate for the test sample is significantly improved. In order to supply the lack of misclassified patterns, confusing patterns near to the decision boundaries are extracted and utilized to generate the mirror images (**Figure 5**).

A proximity measure μ of a pattern X to a decision boundary is defined by

$$\mu(X) \quad = \quad \frac{d_1(X) - d_2(X)}{d_1(X) + d_2(X)}, \tag{2}$$

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 IEEE

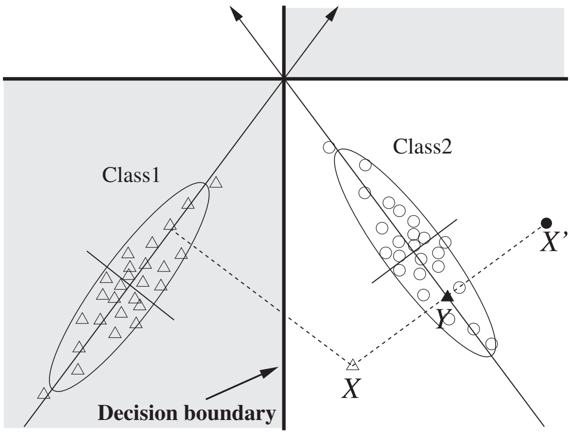


Figure 4. Generation and learning of mirror images.

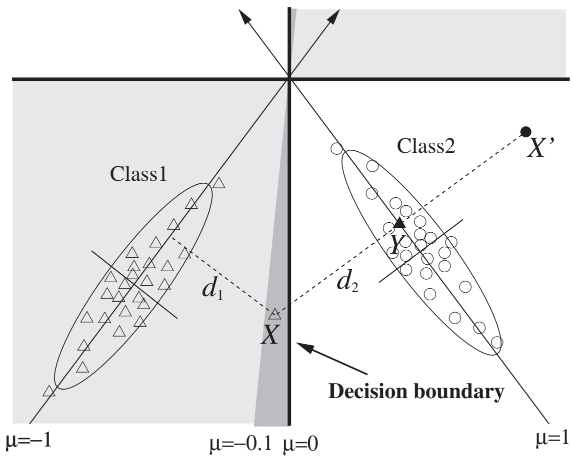


Figure 5. Mirror image learning with margin.

where $d_1(X)$ is the square error between X and the output Y from the network of its own class, and $d_2(X)$ is the minimum square error between X and the output of the other classes. The range of μ is $[-1, 1]$, and if μ is positive (negative) the classification is wrong (correct). For a pattern X on the decision boundary $\mu = 0$, and for the one on the hyperplane of class 1 (class 2) $\mu = -1$ ($\mu = 1$). Even if the μ is negative (correct classification) but is close to zero, the pattern is selected to generate the mirror image to enlarge the learning sample, i.e. a pattern X of class 1 is selected for mirror image generation if

$$\mu \geq \mu_t \qquad (-1 \leq \mu_t \leq 0) \tag{3}$$

for a threshold μ_t. The smaller the μ_t is, the more learning patterns are selected for the mirror image generation.

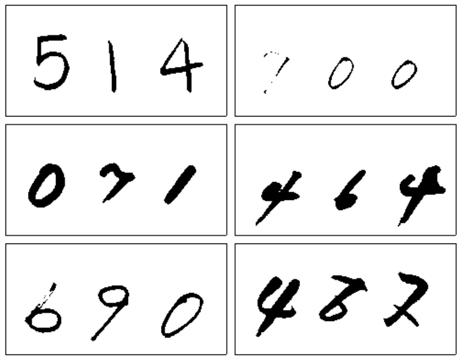


Figure 6. Examples of binary ZIP code image.

4. Performance evaluation

4.1. Used sample and experiment

The handwritten ZIP code database IPTP-CDROM1 [1] provided by Institute for Posts and Telecommunications Policy is used in this experiment. The CDROM contains three digit handwritten ZIP code images collected from real Japanese New Year greeting cards. The writing style and equipment have wide range of variation. The size of image is 240 dots x 480 dots in height and width respectively, and the gray scale is 8 bits (256 levels). **Figure 6** shows examples of binary images of ZIP codes. The total number of the images is 12,000 (36,000 numerals), and about a half is used for learning, and the rest for test. A series of preprocessing such as binarization and character segmentation [1] are first applied to generate binary numeral images of 120 dots x 80 dots in height and width respectively.

A feature vector of size 400 was extracted from each numeral image by the gradient analysis method [7], [8]. The feature extraction process consists of procedures for gradient calculation, feature vector generation, and dimension reduction of the feature vector. A gray scale image is generated from an input binary image by repeating 2 x 2 mean filtering, and the gradient is calculated by Roberts filter [9]. Obtained gradient image is divided into square blocks, and a gradient feature vector is composed of the strength of gradient accumulated separately in each quantized direction in each block to obtain the feature vector of size 400. Then the feature size was reduced to 64 by the principal component analysis. Ten autoassociative networks for ten classes (0 to 9) were created and trained. The back propagation method was used for training the networks. The number of neurons in each layer is 64, 16, 64, respectively. For the first k iterations of learning, the ordinary class independent learning is performed and after the correct classification rate is saturated, the mirror image learning is alterna-

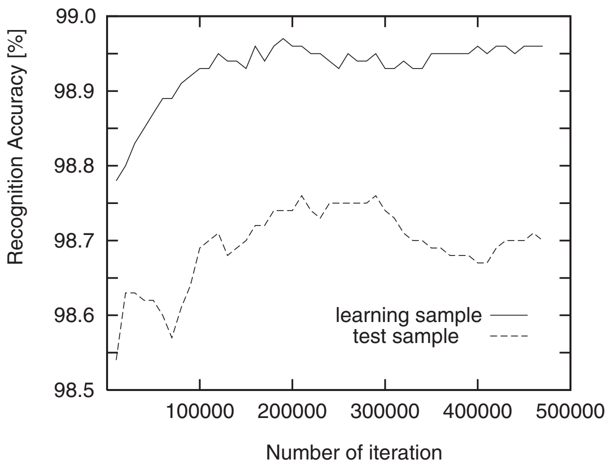


Figure 7. Recognition accuracy of the class independent learning.

tively performed together with the class independent learning. The mirror image learning is performed for the patterns for which $\mu \geq \mu_t$.

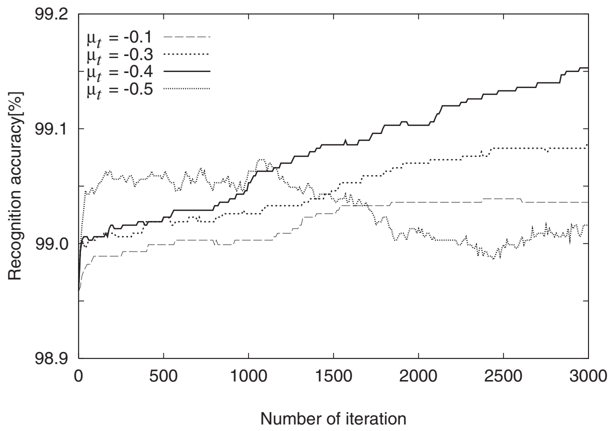
4.2. Experimental result of class independent learning

Figure 7 shows the recognition accuracy of the class independent learning. It is shown that the recognition accuracies are nearly saturated at 200,000 times of iterations and reach 98.96% and 98.76% for the learning sample and the test sample respectively.

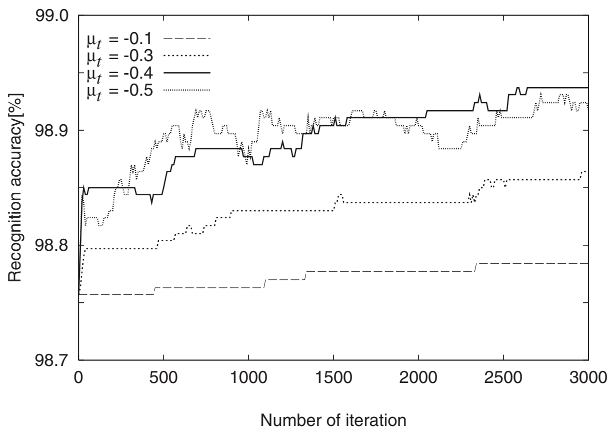
4.3. Experimental result of mirror image learning

The mirror image learning was performed after the class independent learning of 210,000 times of iteration. **Figure 8** shows the improvement of the recognition accuracy by the mirror image learning with various μ_t. When the margin with $\mu_t = -0.4$ was set up (thick solid line) the recognition accuracies for both the leaning and the test sample were best improved during the 3,000 times of iteration. For $\mu_t = -0.5$ the recognition accuracy for the learning sample peaked at smaller times of iterations and was deteriorated by further iterations. In these experiments the learning rate parameter α of the back propagation method was fixed to 0.1.

Figure 9 shows the improvement of the recognition accuracy by the mirror image learning for various learning rate parameter α. The margin was fixed to $\mu_t = -0.4$. The recognition accuracy for the test sample was improved from 98.76% to 99.23% for the learning rate parameter $\alpha = 1.5$ to 2.0.



(a)Learning sample



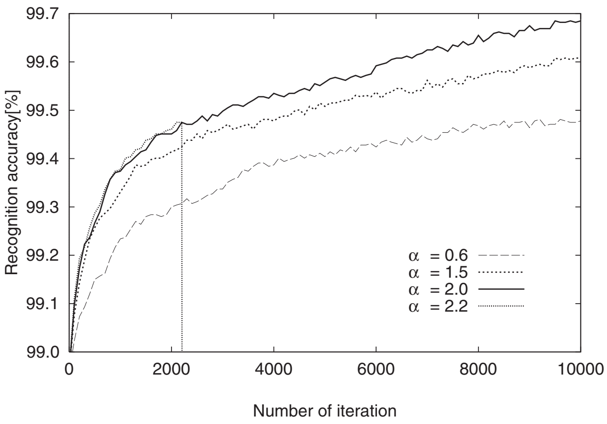
(b)Test sample

Figure 8. Accuracy improvement by mirror image learning ($\alpha = 0.1$).

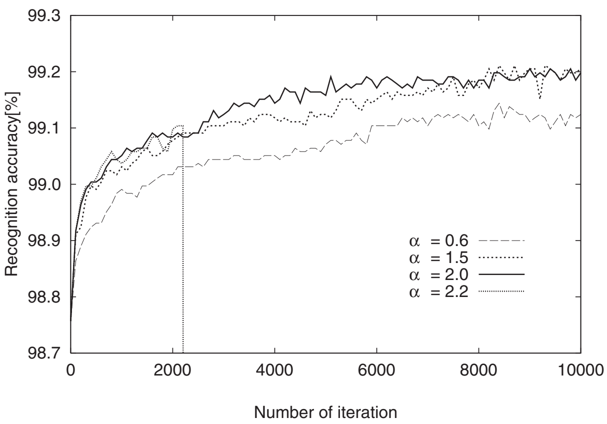
5. Conclusion

The recognition rate of the autoassociative neural network classifier was improved by the mirror image learning. The effect of the mirror image learning is enhanced by the margin that extracts the confusing patterns near to the decision boundary to generate the mirror images. The recognition rate was improved from 98.76% to 99.23% in the recognition test for handwritten numeral database IPTP CD-ROM1. Further studies on (1) the mirror image learning for feature vectors of higher dimensionality (> 64), (2) the mirror image learning for five layered autoassociative neural networks, and (3) application to other classification problems e.g. automatic text categorization, are remaining as future research topics.

Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)
0-7695-1960-1/03 $17.00 © 2003 **IEEE**



(a)Learning sample



(b)Test sample

Figure 9. Accuracy improvement by mirror image learning ($\mu_t = -0.4$).

Acknowledgement

The authors would like to acknowledge the support of Institute for Posts and Telecommunications Policy for providing us the IPTP CDROM1.

References

[1] K. Osuka, T. Tsutsumida, S. Yamaguchi, and K. Nagata. IPTP survey on handwritten numeral recognition. IPTP Research and Survey Report (English translation) R-96-V-02, Institute for Posts and Telecommunications Policy, Japan, June 1996.

[2] T. Wakabayashi, M. SHI, W. Ohyama, and F. Kimura. Accuracy improvement of handwritten numeral recognition by mirror image learning (http://www.info.mie-u.ac.jp/waka/mil.pdf). In *Proc. of 6th International Conference on Document Analysis and Recognition (ICDAR'01)*, pages 338–343, Seattle, 2001.

[3] S. Watanabe and N. Pakvasa. Subspace method of pattern recognition. In *Proc. 1st IJCPR*, 1973.

[4] E. Oja. *Subspace Methods of Pattern Recognition*. Research Studies Press, England, 1983.

[5] M. Ikeda, H. Tanaka, and T. Moto-oka. Projection distance method for recognition of hand-written characters (in Japanese). *Trans. IPS Japan*, 24(1):106–112, 1983.

[6] F. Kimura, S. Inoue, T. Wakabayashi, S. Tsuruoka, and Y. Miyake. Handwritten numeral recognition using autoassociative neural networks. In *Proc. of 14th International Conference on Pattern Recognition (ICPR'98)*, volume 1, pages 166–171, Brisbane, 1998.

[7] T. Wakabayashi, S. Tsuruoka, F. Kimura, and Y. Miyake. Increasing the feature size in handwritten numeral recognition to improve accuracy. *Systems and Computers in Japan (English edition), Scripta Technica*, 26(8):33–44, 1995.

[8] M. Shi, Y. Fujisawa, T. Wakabayashi, and F. Kimura. Handwritten numeral recognition using gradient and curvature of gray scale image. *Pattern Recognition*, 35(10):2051–2059, 2002.

[9] J. C. Russ. *The Image Processing Handbook — 2nd Edition*. CRC Press, Boca Raton, FL, 1995.