Unification of Neural Network with a Hierarchical Pattern Recognition

Chang-Mock Park and Gi-Nam Wang
Department of Industrial Engineering
Ajou University, Suwon 441-749, Korea
(Fax) 0331-214-0942 (Tel) 0331-219-2425
(Email Add: gnwang@madang.ajou.ac.kr)

Abstract. Unification of neural network with a hierarchical pattern recognition is presented for recognizing large set of objects. A two-step identification procedure is developed for pattern recognition: coarse and fine identification. The coarse identification is designed for finding a class of object while the fine identification procedure is to identify a specific object. During the training phase a course neural network is trained for clustering larger set of reference objects into a number of groups. For training a fine neural network, expert neural network is also trained to identify a specific object within a group. The presented idea can be interpreted as two step identification. Experimental results are given to verify the proposed methodology.

I. INTRODUCTION

Automatic optical recognition of an object or object has a wide application in automated manufacturing systems of robotic control, parts inspection, postal code reading and automatic reading devices. In an automated manufacturing and inspection, parts and semi-assembled products are continuously moving on a conveyer line during visual inspection. It is an important and difficult problem to design an intelligent object recognition scheme. In optical object recognition, a object may have various shapes and different position, rotation and scale, therefore designing an automatic object recognition scheme is also a difficult problem. The problem is even harder when the objects have various multi-shape or their images are noisy.

This paper is concerned with a neural network-based approach to object recognition. The hierarchical search technique is employed for designing neural network architecture . and learning paradigm. A new two-step identification procedure is proposed for improving network generalization performance. The presented two-step identification procedure consists of a coarse identification and a fine identification. The coarse identification is to find appropriate group in which the object is included. Once , the coarse identification is completed, the fine identification is performed to identify the best probable object within a class. The presented coarse and fine procedure could be efficiently used for identifying a large set of patterns.

Neural network is based on a non-parametric classifier instead of a statistical classifier. A classical approach to pattern recognition relies on statistical classification. The Minimum Distance Classification and the Bayesian approach have served as basis for statistical pattern recognition. The Minimum Distance Classifier computes the distance between a pattern X of unknown classification and a prototype of each class, then assigns the pattern to the class to which it is closest in distance. The Bayesian approach minimizes the average cost of misclassification as well as yields the lowest probability of error. Shortcomings of the statistical method exist; these include limitation of probability distribution, uncertainty of the sample, and difficulty in determining prototype exemplars.

II. LITERATURE REVIEW AND BACKGROUND INFORMATION

Correlation is a basic approach employed in most pattern recognition algorithms. However, this is only practical for ideal situations. Correlators are very sensitive to distortions between the ideal patterns and input (unknown) patterns. In particular, scale, rotation, and position changes between the ideal reference and the target located in an input pattern will severely degrade the correlation and the ability to detect the target [24]. In light of these limitations, a more desirable feature space is often sought for accomplishing object detection and identification. Various techniques have been suggested for the position, rotation, and scale invariant pattern recognition [7,8,12,14]. Preprocessing is required to extract an invariant position, rotation and scale feature. Most of the existing algorithms involve extensive processing on the image before the features are extracted. This processing requires excessive computation. They also need a huge library set to store the extracted features.

(2). Neural Networks Approach

Neural Networks approach has the special objectistics that comprise learning. If a class membership is of interest, the system learns from observations of patterns that are identified by class and infers a discriminate for classification. The twin processes of generalization and specialization are all-important in neural networks.

Generalization enables a pattern-recognition system to function completely throughout pattern space, even though it has learned from observing only a limited body of examples. Specialization allows such a system to recover from error and to improve itself [40]. It can acquire the capability to generalize a specific or limited piece of input to produce an output solution. This capability is important because it allows the system to provide solution output even when it is given incomplete input information. After finishing the process of learning, pattern recognition is performed on the basis of similarity in shape between patterns. It is neither affected severely by deformation nor by changes in size, or by shift in the position of the input patterns. The learning time can be separate from the on-line computation, which results in reducing the on-line processing.

Pattern recognition tasks require the ability to match large amount of input information simultaneously and then generate categorical or generalized output. Neural networks possess these capabilities as well as the ability to learn and build unique structures for a particular problem. Fukushima [11] proposed the Neocognitron model for recognizing hand-printed objects. The local features of the input pattern are extracted by the cells of a lower stage, and they are gradually integrated into more global features. Finally, each cell of the highest stage integrates all the information of the input pattern, and responds only to one specific pattern. However, this is the most complicate network ever developed and usually requires a large number of processing elements and connections.

Carpenter and Grossberg simulated on an alphabet learning circuit based on Adaptive Resonance Theory (ART) utilizing a two-thirds rule to allow for self-stability of the network [6]. Carpenter/Grossberg algorithm can perform well with perfect input patterns but even a small amount of noise can cause problems[5]. With noise, there are problems of determining the vigilance threshold and capacity for the stored exemplars.

The incorporation of several neural memories, each coupled with a spatially filtered feature space, was presented by Richard A. Messner and Harold H. Szu [14]. They derived multiple bands of information from an input. The recognition accuracy is increased; however, it requires an excessive amount of time for processing. G. Eichmann and T. Kasparis presented a pattern classification using a linear associative memory. Hough transformation was used as a feature extraction method. Linear associative model replaces tedious clustering algorithms and similarity measure with a stored vector matrix multiplication [10].

Recently, a versatile object Recognition was developed by Rajavelu, Musavi and Shirvaiker. They designed a multishape object recognition system. The Walsh transformation was used to extract features from a object. Backpropagation [1] was applied to this problem. Recognition time and accuracy were considerably improved. However, there existed still some difficulties for recognizing a large amount of various objects (e.g. multi-shape, size, Chinese objects, and Korean objects).

Despite of several existing neural networks based approaches, few methods for determining a set of training exemplars has been suggested. Training set has a significant meaning in that it has a great effect on the networks generalization. The unsupervised clustering mechanism of ART and Kohonen is applied for choosing the exemplars. Automatic self-training procedure is suggested to improve the networks generalization performance.

III. PROPOSED METHODOLOGY

A typical pattern classification algorithm can be divided into two parts [10]; first, the patterns are extracted through preprocessing of input data. Second, these features are compared with each previously stored set of reference features (exemplars) until a match is found (see Figure 1). The proposed methodology consists of three components: preprocessing procedure for feature extraction, a training set generation method, and the two-step learning algorithm.

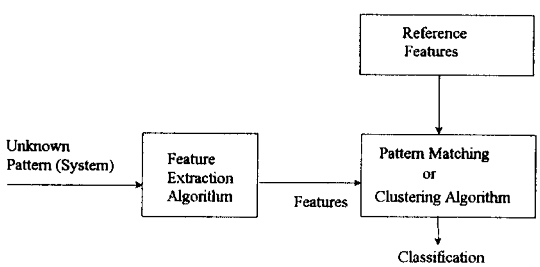


Figure 1. General Description of Pattern Classification

1. Preprocessing for Feature Extraction

Preprocessing is needed to extract a meaningful feature which categorically defines the details of the pattern. It may be used to orthogonalize the input and reduce data storage by extracting a significant feature from the image. One such preprocessing device is the Walsh transformation which constitutes a set of orthogonal function that belongs to class of piece-wise constant basis functions. It used in communication, signal processing, system analysis, and control. This function is defined as -1 or 1 over the interval [0,1] as followings: the first order of Walsh function is 1 over the interval [0,1], and the second function is defined as 1 over the [0,1/2] and -1 over the [1/2,1]. As the order of Walsh function is increased, each function generates 1 and -1 partitioning the interval [0,1] into subintervals (refer to Figure 2).

Compared with other transformation techniques, the Walsh transformation reduces computation time. The intensity distribution is defined as the number of dark image pixel defined over the subinterval. The Walsh transformation is to multiply the number of dark image pixel by the Walsh function defined over the subinterval, and integrate it. If we use eight Walsh functions, corresponding expansion coefficients can be obtained by this integration.

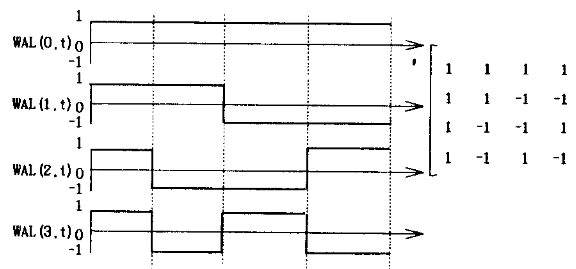


Figure 2. Sequence-Ordered Walsh Function to n=3

Strictly speaking, the computation is dependent on the number of the Walsh functions. This number effects on the accuracy of feature representation: Figure 2 represents the first four Walsh functions, and the four Walsh functions which have the matrix form. A large number of The Walsh functions are required if a detail image is necessary. It is simpler than the existing transformation used in feature extraction [1]. Depending on the complexity of patterns, preprocessing accuracy can be controlled by the number of the expansion coefficients of the Walsh transformation; Accuracy is increased by increasing the number of The Walsh functions. However, larger numbers of the Walsh functions requires a greater computation. In general, smaller numbers of The Walsh functions is appropriate for subgroups which have low degree of similarity among patterns.

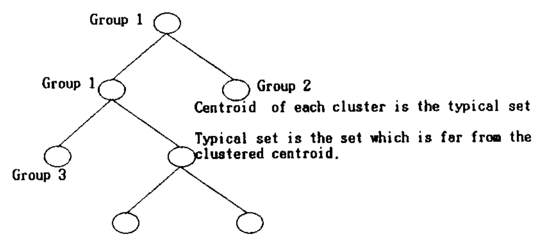
A larger number of the Walsh function is required for subgroups having a high degree of similarity. Another modification for applying the Walsh functions is the mapping procedure of intensity distribution function. It is dependent on the size of object. The reason is that they did not standardize the height of the intensity distribution function expressed as the number of dark pixel even if they normalized the X-axis interval to [0,1]. This makes the problem size dependent. To keep the dynamic range of expansion coefficients consistent for different size objects, the image plane of v(x) is mapped onto the interval defined over [0,+1]. This implies v(x) will have real values in [0,+1] range.

2. Training Set Generation Method

Neural Networks learn from an available training set, and must gives a general solution for an unknown input not included in the training set. Generalization of neural networks is an important issue for any application. It can be effected by three critical factors: network architecture, learning algorithm, and training set [43]. Architecture determines a family of mapping from the input space to the solution space via a learning algorithm. Learning is an important characteristic of neural networks, and this process is performed under the defined parameter space called a training set. In order to accomplish maximum generalization, the selection mechanism of a training set should be designed thoroughly considering the category of input space. This idea is analogous to statistical sampling, whose purpose is to extract representative from a large population.

However, there are some differences between a sampled set and a training set in that a training set should include two types of patterns to generalize the networks: typical sets and the critical set in the given inputs. A typical set is in the interior of a class, which is the nearest to the centroid of the input vector of a group. A critical set is far away from the centroid of a group, and it is around of class boundary. By learning the boundary set thoroughly, the trained networks suggest a clear solution for an ambiguous unknown input which is in the boundary region. One of the important task of this research is to develop an efficient training set generation method. The main goal of this procedure is to increase the networks generalization and minimize the number of the training patterns through extracting the appropriate training set. So far, despite of existing implementations of neural networks in pattern recognition, few systematic approaches for determining the training set has been made. M.Wann, T.Hediger, and N.N.Greenbaun investigated a method for determining the training set using the Hamming distances. However, this method has limitations: only binary patterns are allowed, excessive computation is required for processing, and it is a subjective methodology.

The presented training set generation method can be based on the clustering concept of unsupervised learning models. This proposed procedure consists of two steps: hierarchical clustering and vigilance testing. The former is used to find a typical set and the latter is aimed to detect a critical set.

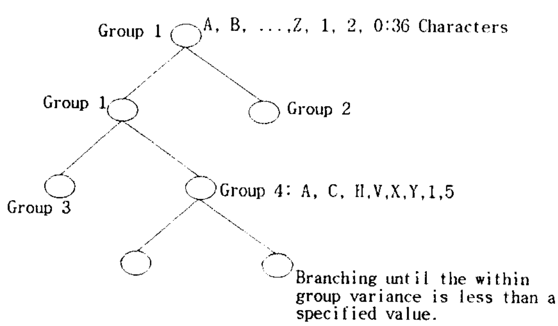


In the first step, all input patterns of a group are decomposed into two subgroups using the hierarchical Kohonen learning algorithm. Each decomposed subgroup is also partitioned into two lower-level subgroups repetitively until the variance within a subgroup is less than a specified value. The variance within a subgroup is the criteria to stop further partitioning subgroups. End-groups are obtained after the generation of all possible subgroups, and the total Euclidean distance within a end-group should be less than a certified value. This value can be utilized as the performance measure for clustering. The typical training set can be obtained by selecting the pattern which has the minimum Euclidean distance between the centroid and input patterns within the generated end-group. The critical pattern is in as far from the centroid (average) of inputs of a end-group, which has the maximum Euclidean distance between the centroid and the inputs. In the second step, the vigilance test used in ART can complement the training set obtained by step 1. This process can be performed by adding the patterns which have the low vigilance value to the obtained training set. The vigilance value represents the degree of similarity how close the inputs to the training set. Thus, this criteria can be used to detect the critical patterns which is in far from the previously obtained training set.

3. Two-step Learning Scheme

Kohonen self-organizing model is fundamentally based on the unsupervised learning model, which maps all possible input space to a general solution space forming a stabilized parameter space. In this study, Kohonen model is applied for a supervised learning process. A stabilized weight vector is obtained using the training set as an input vector. This weight vector can be utilized to give a general solution for unknown inputs. The generalization of the networks can be increased by the hierarchical decomposition of the wide parameter space into multiple stabilized space. The decision boundary region is generated forming hierarchical binary regions, which results in reducing the misclassification. Basically, hierarchical search algorithm generates all possible subnodes forming the hierarchical tree structure. Lower bound of each subnode plays an important role searching the dominant subnode eliminating the non-dominant search space. Proposed hierarchical search learning algorithm is similar to that in having hierarchical tree structure.

For example, first, 36 exemplar objects (training set) are divided into two subgroups based on the similarity. Group 1 consists of objects (A C H K L M O P S U V X Y Z 1 4 5 7 0). Objects of (B D E F G I J N Q R T W 2 3 6 8 9) are classified as group 2. Kohonen model is used to divide 36 objects into 2 groups. The next step is to divide each group into two lower-level subgroups. The objects included in group 1 are decomposed into group 3 (which the elements are K, L, M, O, P, S, U, Z, 4, 7, and 0) and group 4(which the elements are A, C, H, V, X, Y, 1 and 5). The same procedure is repeated until the variance (total Euclidean distance) within the generated subgroup is less than a specified value.



The coarse learning scheme is designed to partition a group into two lower-level subgroups. In the process of generating each subset, a stabilized weight vector is obtained by the Kohonen self-organizing property, which can be interpreted as a training procedure. After generating a subgroup, branching is stopped when the total Euclidean distance within a subgroup is less than a specified value, and end-groups are found. In this case, group 3, group 4, group 5, and group 6 are end groups. They have a different number of objects in the end-groups. The supervised learning concept is cônsidered after getting the hierarchically stabilized weight vector. The first step is to identify the corresponding end-group for unknown input objects, and the second step is to identify a specific pattern. In the former process, between two generated subnodes at the same hierarchical level, one node having larger lower bound is fathomed, and this is interpreted as a bounding strategy. The lower bound of each subnode is the Euclidean distance between the unknown input and the stabilized weight obtained from learning procedure. Therefore, only one hierarchical search path is found at every node and the possible search space is very limited. Once identification for an end group is finished, the identification for a object is required. In each end group, backpropagation is performed to identify a object. The problem is the misclassification of subgroups. If misclassification for a group occurs before reaching the end subgroup, it reduces the recognition rate and system performance. The training set is designed to increase the networks generality, and the proposed twp-step learning algorithm can reduce the misclassification rate for the group by forming binary decision regions hierarchically.

5. Proposed Algorithm

The proposed algorithm can be divided by four main procedures: Feature extraction by the Walsh transformation, the generatiôn method of an efficient training set, training procedure and recall procedure. Details can be described as following steps.

Step 1. Feature Extraction via the Walsh transformation. The extracted values are the inputs to neural networks model.

Step 2. Determination of training set. Training set is composed of typical and critical set.

a. Find typical set using the training set generation method.

b. Detect the critical set using ART.

Step 3. Training the Branch and Bound Neural Networks.

a. Generate subgroup using the training set in step 2.

b. Branch stopping rule. Repeat (step 3. a) until the total summation of Euclidean distance within each subgroup is less than a specified value. After reaching end-subgroups, the stabilized weight vector is formed at every subnode. These weights are used to calculate the distance from an unknown object, which is used as the lower bound of the branch and bound algorithm.

d. Train the backpropagation network using the clustered objects of the each end-subgroups.

Step 4. Recall procedure.

a. For a unknown object, compute the lower bounds (Euclidean distance) of the two possible subgroups. Fathom the subgroup which has the larger lower bound between the two generated subgroups.

b. Repeat (step 4. a) until reaching the corresponding end-group.

c. Recall the Backpropagation network within the corresponding end-group for the unknown object. For each end-subgroup which has a high similarity among the patterns, we can increase the number of the Walsh expansion coefficients, which is used in backpropagation model.

The above mentioned algorithm is based on both concepts of supervised and unsupervised learning. Unsupervised learning is used in the generation of a training set and the training process. After getting the training objects and finishing the training process, the supervised learning concept is applied for recall steps.

IV EXPERIMENTAL RESULT AND SYSTEM PERFORMANCE

In this section, experimental results are presented with the system performance. A data base consists of differently oriented and noised 16x16 binary images which represent the thirty-six kinds of English objects. Thirty-six kinds of training objects are also used as a training data, in which these are extracted by the training set generation method. There exist differences in numbers and shapes according to each object included in the training set. Seven different types of rotations are occurred in the test (recall) objects as followings: 1. 90 degrees, 2. 180 degrees and 3. 270 degrees, 4. 90 degree with noise, 5. 180 degree with noise, 6. 270 degree with noise, and 7. only noise without any rotation. Thirty-six kinds of standard image objects are utilized for training the net and 252 different objects are set aside for testing. Therefore, there are 7 different kind of recall data sets depending on the degree of rotation and noise. The network performance is measured by how well it can classify the unknown input which it has not been trained on. The following experimental results show the recognition rate according to each model and different cases.

In this implementation, thirty-six kinds of object sets are divided into four end-groups using the Branch and Bound learning mechanism: The recognition rate with this mechanism is higher than the method of dividing the four end-groups initially by using the Kohonens group classification method because at each subset generation, the BB learning algorithm only considers two possible subsets generation. At each level, the stabilized weight vector forms a tighter sphere than generating more than two subsets. The average recognition rates of objects within each end subgroups are 96.56, 95.625, 97.5 and 98.7 percents. Obviously, the actual misclassification should be considered the misclassification of the subgroup. However, the misclassification rate for subgroups is shown as lower than the misclassification of each objects obtained by using the conventional backpropagation, and the misclassification rate for objects by Kohonen Branch and Bound scheme is also lower than that of existing techniques which are 0.346, 0.302, 0.23 and 0.53 percents.

V. CONCLUSION

In this paper, three major works are investigated to recognize scaled position-rotated objects. In order to compare the performance, Minimum distance classification is used as a statistical pattern classification. The two-step identification scheme is proposed to reduce computation time and to enhance the recognition accuracy. The generation method of an efficient training set is suggested by using the clustering concept of Kohonen and ART. An efficient method for determining the number of The Walsh expansion coefficient is also explained to increase the system performance. Proposed algorithm might increase the recognition ration and convergence: using the two-step neural mapping, a large set of objects can be categorized into several end-groups according to the similarity, and Backpropagation can be utilized efficiently within the decomposed end-groups.

However, the extra effort to identify a end-group will be required as a penalty. However, this extra effort seems not to be expensive because of the reduction in object set, and of the benefit of off-line learning. Supervised learning application for two-step neural networks is considered in this mechanism. The advantage of off-line training is that training computation time can be separate from on-line computation by training the network before entering the on-line computation. The information obtained during the training period can be stored and used in the process of recall. Actually, the real computation time may include the recall time. The further study to this project is to implement another preprocessing. Preprocessing is important to increase the recognition rate and system performance. Moment invariant feature might work well under the condition of rotation. As a statistical pattern recognition approach, Baysian classifier may be better than minimum distance classification under certain circumstances. Another further study is to solve the problem of recognizing partially occluded objects using the concept of learning in neural networks. A modification of ART can suggest a solution to this problem only when the object is exactly overlapped but not in general cases. The concept of learning in neural network is utilized to solve this problem by training all possible occlusion. But this is apparently an exhaustive computation. An efficient neural net mechanism might suggest a solution for this problem.

References

[1] A. Rajavelu, M. T. Musavi, and M. V. Shirvaikar, A Neural Network Approach to Object Recognition, *Neural Networks* vol.2, 387-393(1989).

[2] K. G. Beauchamp, *Walsh Functions and their Applications*,171-178. London: Academic Press (1975).

[3] D. J. Burr, A Neural Network Digit Recognizer, *Procedings of IEEE International Conference on Systems, Man and Cybernetics Piscataway, NJ: IEEE.* Vol 2, 1621-1625 (1986).

[4] Berthold Klaus Paul Horn, Robot Vision (1986).

[5] G. A. Carpenter, and S. Grossberg, Neural Dynamics of Category Learning and Recognition: *Attention, Memory Consolidation, and Amnesia, in J. Davis, R. Newburgh, and E. Wegman (Ed) Brain Structure, Learning, and Memory, AAAS Symposium Series*, (1986).

[6] G. A. Carpenter, and S. Grossberg, Associative Learning Adaptive Pattern Recognition and Coorperative-Competitive Decision Making by Neural Networks, *to appear in SPIE Special Series Optical and Hybrid Computing* H.H.Szu ed, (1986).

[7] D. Casasent and P. Demetri, Position, Rotation, and Scale Invariant Optical Correlation, *Applied Optics*, Vol. 15, 1795-1799. July (1976).

[8] D. Casasent and O. Hironobu, A high-Dimensionality Pattern Recogniition Feature Space, *Proceding of the SPIE*, vol. 579, pp. 245-257.Sept. (1985).

[9] G. L. Cash, and M. Hatamian, Optical Object Recognition by the method of moments, *Computer Vision, Graphics and Image Processing*, 291-310 (1987).

[10] G. Eichmann and T. Kasparis, Pattern Classification using a Linear Associative memory, *Pattern Recognition,* Vol.22, No.6 pp.733-740,(1989).

[11] K. Fukushima, Neocognitron: A Neural Network Model for a Mechanism of Visual *Pattern Recognition*, IEEEE Trans. Sys. Man and Cybern, Vol. SMC-13, No.5, 826-834 Sept/Oct. (1983).

[12] Yuan-Neng. Hsu et al, Rotation-Invariant Digital Pattern Recognition using Circular Harmonic Expansion, *Applied Optics*, Vol.21, pp 4012-4015. (1982).

[13] D. Mehr and S. Richfield, Neural Net Application to Optical Object Recognition, *IEEE First International Conference on Neural Networks*, Vol.4, 771-776 (1987).

[14] R. Messner and Harold H. Szu, An Image Processing Architecture for Real Time Generation of Scale and Rotation Invariant Patterns, *Computer Vision, Graphics and Image Processing* Vol. 31, pp.50-56. July

(1985).
[15] D. E. Rumelhart, G. H. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation in D.~E.~Rumelhart and J. L. McClelland (etc.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Vol: Foundation MIT Press (1986).
[16] G. K. Atkin, J. E. Bowcock, and N. M. Queen Solution of a Distributed Deterministic Parallel Network using Simulated Annealing, *Pattern recognition* No.4, 461-466 (1989).
[17] S. T. Kahan, T. Pavlidis, and W. Baird, On Recognition of Printed Objects of any Shape and Size, *IEEE Transactions on Pattern Recognition and Machine Intelligence*, PAMI-9, 274-285 (1987).
[18] T. Kohonen, *Self-Organization and AssociAative Memory*, Springer-Verlag, (1984).
[19] L. Lam, and C. Y. Suen, Structural Classification and Relaxation Matching of totally unconstrained handwritten zip-code number,*Pattern Recognition* Vol.21, 19-32 (1988).
[20] M. Hu, Visual Pattern Recognition by Moment Invariants, *IRE Trans. Inform. Theory*, Vol. IT-8, 179-187 Feb., (1962).
[21] Mikell P. Groover and Emory W. Zimmers, Jr., *CAD/CAM Computer-Aided Design and Manufacturing* (1984).
[22] T. Pavlidis, A Vectorizer and Feature Extractor for Document Recognition, *Computer Vision, Graphics and Image Processing* Vol.35, 111-127 (1986).
[23] Richard P. Lippmann, An Introduction to Computing with Neural Nets, *IEEE ASSP Magazine* April (1987).
[24] Leger, James R. and Sing H. Lee, Hybrid Optical Processor for Pattern Recognition and Classification using a Generalized Set of Pattern Functions, *Applied Optics*, vol.21, 274-287, (1982).
[25] Richard A. Messner, and H. H. Szu, Distorted Object Recognition via an Associative Neural Network, *Proc. of Intelligent Robots \& Computer Vision* Vol.726 (1986).
[26] F. W. M. Stentiford, Automatic Feature Design for Optical Object Recognition using Evolutionary Search Procedure}, *IEEE Transactions on Pattern Recognition and Machine Intelligence*, PAMI-7, 349-355 (1985).
[27] G. Tesauro, Scaling Relationships in Back-Propagation Learning: Dependence on Training Set Size, *Complex Systems* Vol.1, 145-168 (1987).
[28] G. Tesauro, and R. Janssens, *Back-Propagation Learning:Dependence on Predicate Order*, TR-CCSR-88-1, Center for Complex Systems Research, University of Illinois at Urbana-Champaign.
[29] Thomas F. Knoll and Ramesh Cojain, Recognizing Partially Visible Objects using Feature Indexed Hypotheses, *IEEE Journal of Robotics and Automation,'* Vol. RA-2, No.1 March (1986).
[30] T. Martin, W. G. Kobel, Sok. Rogers, M. Kabrisky, and J. P. Mills, A Distortion-Invariant Pattern Recognition Algorithm, *Proc. of Intelligent Robots and Computer Vision* Vol.726, (1986).
[31] T. Kohonen, Correlation Matrix Memories, *IEEE Trans. on Computer*, Vol. c-21, No. 4, 353-359.
[32] J. Mantas, An Overview of Object Recognition Methodologies, *Pattern Recognition*, Vol.19, No.6 425-430 (1986).
[33] S. R. Ramesh, A Generalized Object Recognition Algorithm: A Graphical Approach, *Pattern Recognition*, Vol.22, No.4, 347-350 (1989).
[34] H. Ritter and K. Schuten, Convergence Properties of Kohonen's Topology Conserving Maps: Fluctuations, Stability, and Dimension Selection, *Biol. Cybern* vol.60, 59-71 (1988).
[35] T. P. Vogel, J. K. Mangis, A. K. Rigler, W. T. Zink, and D. L.Alkon., Accelerating the Convergence of the Back-Propagation Method, *Biol. Cybern.* Vol.59, 257-263 (1988).
[36] J. L. Walsh, A Closed Set of Normal Orthogonal Functions, *American Journal of Mathematics* Vol.45, 5-24 (1923).
[37] H. F. Li, R.Jayakumar and M. Youssef, A Parallel Algorithms for Recognizing Handwritten Objects Using Shape Features, *Pattern Recognition,* Vol.22, No.6 641-652 (1989).
[38] T. E. Calvert, Nonorthogonal Projections for Feature Extraction in Pattern Recognition, *IEEE Trans. Comput.* C-19, 447-452 (1970)
[39] B. V. Dasarathy and K. P. B Kumar, Chitra Cognitive Handprinted Input-trained Recursively Alalyzing System for Recognition of Alphanumeric Objects, *Int. J. Comput. Inf. Sci.* 7, 252-282 (1978).

[40] Yoh-Han Pao, *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley (1989).

[41] Pasi Koikkalainen and Erkki Oja, Self-Organizing Hierarchical Feature Maps, *Int. Joint Conf. on Neural Networks*, Vol.2 279-283 San Diego, California June 17-21, (1990).

[42] E. Wan, P. Ning, and B. Widrow, Neural Tree Structured Vector Quantization, *Int. Joint Conf. on Neural Networks*, Vol.2 267-270 Washington D.C., Jan 15-19, (1990).

[43] M. Wann, T. Hediger and N. N. Greenbaun, The Influence of Training Sets on Generalization in Feed-Forward Neural Networks, *Int. Joint Conf. on Neural Networks*, Vol.3 137-142 San Diego, California June 17-21, (1990).