Robust Planar Shape Recognition Using Spectrum Analyzer and Fuzzy ARTMAP

스펙트럼 분석기와 퍼지 ARTMAP 신경회로망을 이용한 Robust Planar Shape 인식

한 수 환*
Soowhan Han*

ABSTRACT

This paper deals with the recognition of closed planar shape using a three dimensional spectral feature vector which is derived from the FFT(Fast Fourier Transform) spectrum of contour sequence and fuzzy ARTMAP neural network classifier. Contour sequences obtained from 2-D planar images represent the Euclidean distance between the centroid and all boundary pixels of the shape, and are related to the overall shape of the images. The Fourier transform of contour sequence and spectrum analyzer are used as a means of feature selection and data reduction. The three dimensional spectral feature vectors are extracted by spectrum analyzer from the FFT spectrum. These spectral feature vectors are invariant to shape translation, rotation and scale transformation. The fuzzy ARTMAP neural network which is combined with two fuzzy ART modules is trained and tested with these feature vectors. The experiments including 4 aircrafts and 4 industrial parts recognition process are presented to illustrate the high performance of this proposed method in the recognition problems of noisy shapes.

요 약

본 논문은 산업분야와 군사적으로 많이 사용되고 있는 planar shape의 인식을 스펙트럼 분석기를 이용하여 FFT 스펙트럼으로부터 추출된 3차원 특징 벡터와 신경회로망인 fuzzy ARTMAP을 이용하여 시도되었다. 외곽선 정보를 추출하여 이를 원점으로 이동시키고 각 경계점들과 원점들과의 유클리드 거리를 구하여 이를 다시 FFT 스펙트럼과 스펙트럼 분석기를 통하여 3차원 특징 벡터를 추출하였다. 이 3차원 데이터는 이동, 회전, 크기에 무관한 값으로 fuzzy ARTMAP에 입력값으로 사용하였다. Fuzzy ARTMAP은 두개의 fuzzy ART 모듈을 가지고 있으며 위에서 구한 특징 벡터들에 의해 학습되고 실험되어 진다. 본 논문에 포함된 실험은 4개의 비행기와 4개의 산업부품을 이용하여 잡음이 섞인 shape의 인식에 있어서 제시된 방법이 좋은 인식률을 기록함을 보여 주고 있다.

*동의대학교 전자통신공학과

I. Introduction

Invariant shape recognition is of great interest in many military and industrial applications such as aircraft classification, characterization of biomedical images, and the recognition of industrial parts by robots for product assembly. Most of these shape recognition systems require an object to be classified in situations where the position, orientation and distance of the object are time-varing. Additionally, the systems are required to be tolerant to noisy shapes results from the segmentation of objects in varing backgrounds as well as non-ideal imaging conditions. There have been over a dozen prior research efforts involving neural network approach [1]-[5]. Recently, the use of neural network in the classification problems is rapidly growing. A neural network can perform the necessary and suitable transformation and clustering operations automatically and simultaneously. In other words, an artificial neural network system is able to abstract the distinctions among the different type of planar shape patterns. Additionally, a neural network is able to recognize complex and non-linear groups in the hyperspace. This is a distinct advantage over many conventional techniques. Several studies have presented the performance of neural network system for the recognition of planar shape images [3]-[5]. However, an improvement can be made by increasing the degree of noise that the neural classifier can tolerate. It depends on the architecture of neural network classifier and the selection of feature vector representing object image. Moreover, for real-time use, the network architecture should be simple and easy to train, and it should be considered reducing the dimensionality of feature vector while maintaining high classification accuracy. In this paper, the fuzzy ARTMAP incorporated with two fuzzy ART networks is utilized as an artificial neural network(ANN) based classifier because of its simple structure, its fast training procedure and its insensitivity to initial set-up parameters unlike some other neural networks [6]. The several successful use of this network on pattern recognition has been proposed in recent years [7]-[9]. With the appropriate selection of neural network architecture, another important factor in neural network approach for pattern recognition problems is to decide how an object is presented to a neural network classifier. For any neural network, the input and thus the representation, must be in a vector form. Clearly, this vector representation should hold shape information of the image and must be normalized with respect to translation, rotation and scaling of the object image. In our study, the boundary of a closed planar shape is characterized by an ordered sequence that represents the Euclidean distance between the centeroid and all boundary pixels since the overall shape information is contained in the boundary of the shape. The amplitude of this ordered sequence is invariant to translation because the Euclidean distance with same starting boundary pixel remains unchanged even the image is shifted. Then, the contour sequence is normalized with respect to the size of image. This normalization includes the amplitude and the duration of the contour sequence. Next, as a means of feature selection, the fast Fourier transform is taken to the contour sequense. The FFT spectrum of this sequence is invariant to the rotation of object and has a better noisy tolerance than the time sequence does [10]-[12]. After this stage, the spectrum analyzer is applied to extract three dimensional spectral feature set from the FFT spectrum. These spectral feature vectors for object representation have the desired format, a property to be invariant in size, shift, and rotation, and are used as the input of fuzzy ARTMAP neural network classifier for the recognition of four aircrafts and four industrial parts.

II. Contour representation and feature measurement

In this portion of the study, the boundary is characterized by an ordered sequence that represents the Euclidean distance between the centroid and all contour pixels of the digitized shape. This ordered sequence provides an unique information for each planar shape. First, the boundary pixels are extracted by using contour following algorithm from the planar image and the centroid is derived [10]-[11]. The second step is to obtain an ordered sequence in a clockwise direction that represents the Euclidean distance between the centroid and all boundary pixels. Since only closed contours are considered, the resulting sequential representation is circular as equation (1).

$$b(N+i)=b(i) \quad i=1, 2, 3, \ldots\ldots, N \tag{1}$$

where N is the total number of boundary pixels.

This Euclidean distance remains unchanged to a shift in the position of original image. Thus the sequence b(i) is invariant to translation. The next step is to normalize the contour sequence with respect to the size of image. Scaling a shape results in the scaling of the samples and duration of the contour sequence. Thus scale normalization involves both amplitude and duration normalization. The normalized duration of the sequence, 128 points fixed, is obtained by resampling operation and function approximation. This is shown in equation (2).

$$c(k)=b(k^{*}N/128) \quad k=1, 2, 3, \ldots, 128 \tag{2}$$

where N is the total number of boundary pixels.

After duration normalization, amplitudes are divided by sum of contour sequence shown in equation (3).

$$d(k)=c(k)/s \quad k=1, 2, 3, \ldots, 128 \tag{3}$$

where $s=c(1)+c(2)+c(3)+\ldots\ldots+c(128)$.

This sequence d(k) is invariant to translation and scaling. In a forth, the 128-point fast Fourier transform is taken to the normalized sequence d(k) by equation (4) for feature selection.

$$F(k)=\sum_{n=0}^{127} d(n)\exp(-j2\pi kn/128) \tag{4}$$

$k=0, 1, 2, \ldots, 127$ where $d(n)=d(n+1)$ $n=0, 1, 2, \ldots, 127$.

The amplitude of this sequence, |F(k)|, is unchanged even after the sequence d(n) is circular shifted [12]. This means |F(k)| has a property to be invariant to the rotation of the shape. At this point, the spectrum analyzer is applied to the sequence F(k) as a means of data reduction. The spectrum analyzer computes the translation invariant spectrum reported by Ahmed and Bates [13]. This achieves a considerable amount of data reduction while retaining enough information to distinguish the pattern classes. The output of the spectrum analyzer is a vector of length $\log_2 K+1$ where K is the number of terms in the Fourier transform of the sequence d(n). For the ordered transform, the spectral components are

$$P(0)=F^2(0), \tag{5}$$

$$P(i)=\sum_{j=2^{i}-1}^{k-1}[F^2(j)+F^2(j+1)] \tag{6}$$

$i=1, 2, \ldots, n$ where j is stepped by 2^{i+1} and $n=\log_2 K$.

The zeroth component is the zeroth term of the Fourier transform squared, and each remaining component is the sum of the power densities at the corresponding fundamental frequency and its odd harmonics. Two adjacent sequency terms, $F^2(j)$ and $F^2(j+1)$, are summed at each frequency. These are analogous to the sine and cosine terms of the Fourier expansion [14]. In this study, K is 128 point, thus the vector length of the output of spectrum analyzer is 8. However, since most of high order spectral components are almost zero, the ony first three components, P(0), P(1), and P(2), are used as input of neural classifier. They are shown in equation (5), (7), and (8).

$$P(1)=[F^2(1)+F^2(2)]+[F^2(5)+F^2(6)]+[F^2(9)+F^2(10)]+\ldots\ldots+[F^2(125)+F^2(126)] \tag{7}$$

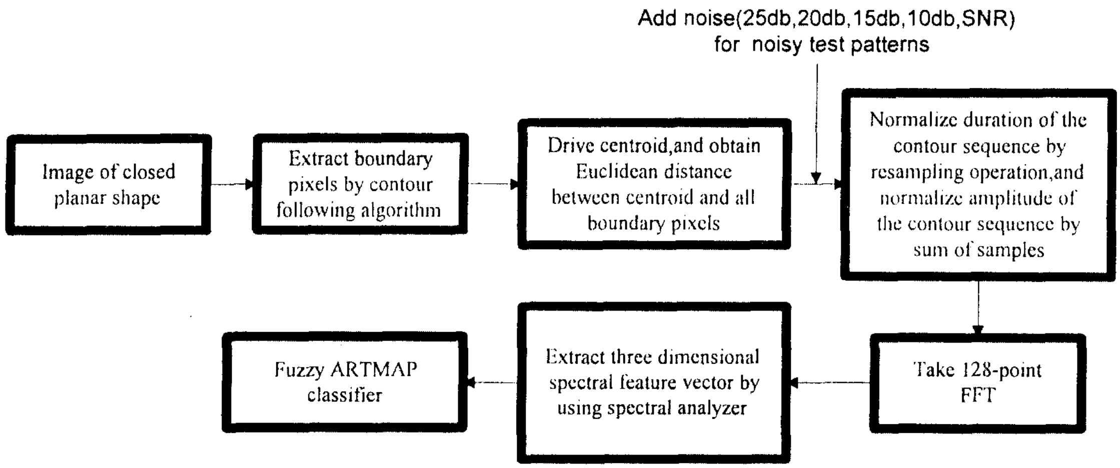


Fig. 1 Processing flowchart for closed planar shape recognition.

$$P(2)=[F^2(3)+F^2(4)]+[F^2(11)+F^2(12)]+[F^2(19)+F^2(20)]+ \ldots\ldots +[F^2(123)+F^2(124)] \quad (8)$$

These feature vectors have the desired format for planar shape recognition problems, which means they are invariant to translation, rotation and scaling of the shape. The overall processing steps in this study are shown in figure 1.

Ⅲ. Fuzzy ARTMAP architecture

Fuzzy ARTMAP incorporates two fuzzy ART modules, ART_a and ART_b, that are linked together via an inter-ART module, F_{ab}, called the map field. The map field is used to form predictive associations between categories and to realize the match tracking rule, whereby the vigilance parameter of ART_a increases in response to a predictive mismatch at ART_b. Match tracking reorganizes the category structure so that predictive error is not repeated on subsequent presentations of the input. The basic architecture of fuzzy ARTMAP is shown in Fig. 2. During the training period, the ART_a module receives a data stream $\{\mathbf{a}\}$ of input patterns and ARTb receives a data stream $\{\mathbf{b}\}$ of target patterns, where **b** is a corresponding target to **a**. If **a** vector a is associated with **a** vector **b**, then any other input that activates the **a**'s category node will predict the category of target pattern **b**. However, when a mismatch at the map field between the ART_a category activated by an input **a** and the ART_b category activated by the input **b** occurs, the network increases the ART_a vigilance by the minimum amount needed to search for and, if necessary, create a new cluster. The new cluster is created to learn a new ART_a category whose prediction matches the ART_b category. After the training is completed, which means the network predicts a correct corresponding target pattern for each of the training input patterns, the test input patterns are presented at ART_a without the use of ART_b. Additional details concerning learning algorithms can be found in [15]-[16] for fuzzy ART and [6] for fuzzy ARTMAP. For the experimental results shown in the next section, a fast learning algorithm with a *five*-voting strategy [6] and the following parameters were utilized: learning rate = 1, choice parameter = 0.005, vigilance parameter for ART_a = 0.99 and 0.90, vigilance parameter for ART_b = 0.5, and the map field parameter = 0.5.

Complement coding is applied to the inputs of the fuzzy ARTMAP for the normalization of the input vectors because proliferation of categories is avoided in the fuzzy ARTMAP by normalization of the input patterns [6]. Therefore, the dimensionality of the neural network classifier input patterns becomes six.

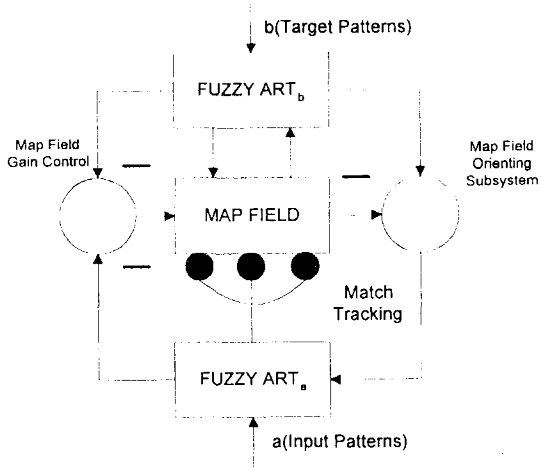


Fig 2. Fuzzy ARTMAP structure block diagram.

Ⅳ. Performance assessment and experimental results

The methodology presented in our study, for recognition of closed planar shape, was evaluated with two different types of data set. One contains the planar images of four different styles of aircraft and the other does those of four different styles of industrial part. These shapes are digitized in 80 by 80 image plane. They are shown in figure 3 and 4. Ten noisy patterns were made by adding random gaussian noise for 25db SNR(signal to noise ratio: db) on the each reference pattern for training the fuzzy ARTMAP neural classifier. Therefore the training set for fuzzy ARTMAP on each of two data sets was consisted of 44 experimental patterns(4 reference patterns and 40 noisy patterns).

The test patterns for fuzzy ARTMAP on each data set were subdivided into five groups. The first one

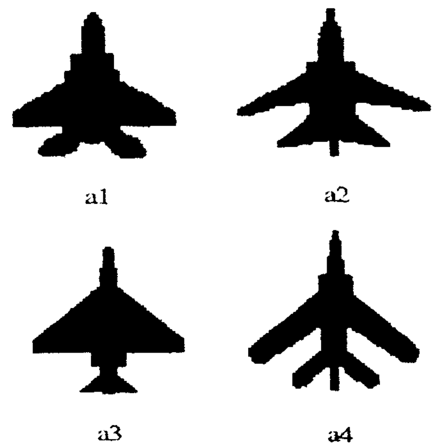


Fig 3. Four aircraft shapes for data set 1.

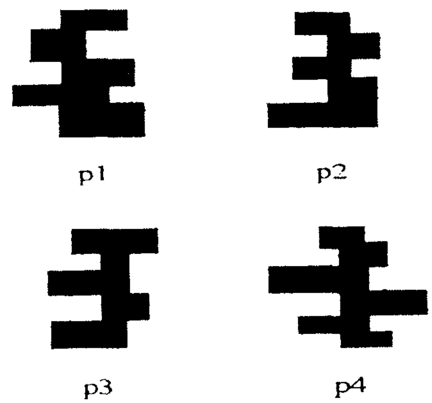


Fig 4. Four industrial part shapes for data set 2.

contains 144 noise free patterns. It was made by 12 (twelve rotated angles with 30 degree increment) x 3 (three type of scale factor, 1, 0.8, 0.6) x 4(number of reference shapes on each data set). The scatter plots of three spectral features, P(0), P(1) and P(2), in the first group of test patterns for each data set are shown in figure 5 and 6. The plots show that these spectral features have the acceptable discrimination ability between each planar shape.

The second group of test patterns contains 1440 noisy patterns for 25db SNR. It was made by 10 (number of random gaussian noise patterns for 25db

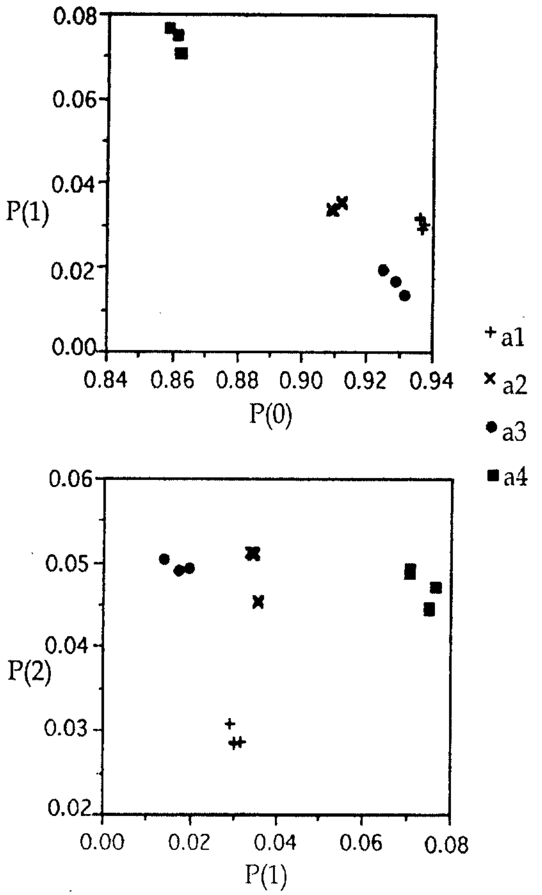


Fig 5. Scatter plots of three spectral featues, P(0), P(1) and P(2), extracted from 144 noise free patterns for data set 1 (12(twelve rotated angles with 30 degree increment) x 3(three type of scale factor, 1, 0.8, 0.6) x 4(number of reference shapes)).

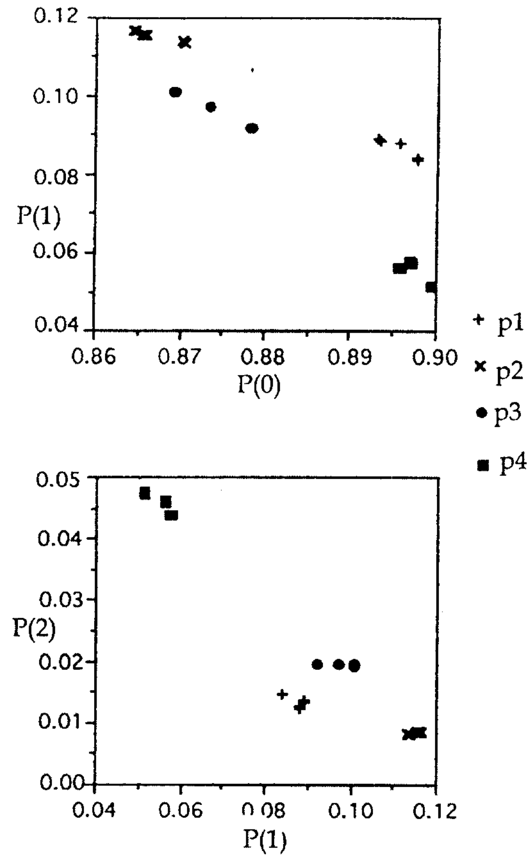


Fig 6. Scatter plots of three spectral featues, P(0), P(1) and P(2), extracted from 144 noise free patterns for data set 2 (12(twelve rotated angles with 30 degree increment) x 3(three type of scale factor, 1, 0.8, 0.6) x 4(number of reference shapes)).

SNR) x 12 (twelve rotated angles with 30 degree increment) x 3 (three type of scale factor, 1, 0.8, 0.6) x 4(number of reference shape on each data set). By the same manner, the third one has 1440 noisy patterns for 20db SNR, the forth 1440 patterns for 15db and the fifth 1440 patterns for 10db. Therefore the number of total test patterns for each data set was 5904. The fuzzy ARTMAP neural classifier was executed under two different simulation environments, Va = 0.99 for conservative mode and Va = 0.90 for forced choice mode. The vigilance parameter in ART_a, Va, measures the degree to which the system discriminates between the different classes of input patterns[6]. Therefore the fuzzy ARTMAP classifer with a high vigilance parameter (conservative mode) may not respond at all to a certain test pattern, which means the network cannot make sure which class this pattern belongs to. In our experiments, such a "no response" of the fuzzy ARTMAP was counted as an error. To improve the performance of the network, a *five* -voting strategy in fuzzy ARTMAP was used, thus the fuzzy ARTMAP was trained five times on a given training set with five different orderings, the prediction of test patterns for each ordering was recorded. The final prediction for a given test pattern was the one made by the largest number of simulations. Performance results of fuzzy ARTMAP classifier are summarized in Table 1 and 2 for data set 1, and in Table 3 and 4 for data set 2. The numbers in tabels represent the percentage of correct classifications.

The overall results under both forced choice mode and conservative mode obtained 100% of classification for data set 1 and over 98.75% for data set 2 in this system until the signal to noise ratio was down to 15db. It shows the fuzzy ARTMAP classifier trained with three dimensional spectral feature vector from spectrum analyzer is highly tolerant to random variations of the shape of image. The differences between overall classification resluts with Va = 0.90 and Va = 0.99 in fifth test group on table 1-4 are caused by the result from image a4 for data set 1, and by the results from image p2 and image p4 for data set 2. Under the conservative mode (Va = 0.99), the classification result with 10db SNR for image a4 on data set 1 droped to 94.72% from 100% under the forced choiced mode (Va = 0.90), and those for image p2 and p4 on data set 2 droped to 90.28% from 99.44%, and 96.11% from 100%, respectively. How-

Table 1. Classification results(%) for data set 1 with Va = 0.90(forced choice mode)

	a1	a2	a3	a4	overall accuracy
first group (noise-free)	100.00	100.00	100.00	100.00	100.00
second group (25db SNR)	100.00	100.00	100.00	100.00	100.00
third group (20db SNR)	100.00	100.00	100.00	100.00	100.00
forth group (15db SNR)	100.00	100.00	100.00	100.00	100.00
fifth group (10db SNR)	99.72	100.00	99.72	100.00	99.86

Table 2. Classification results(%) for data set 1 with Va = 0.99(conservative mode)

	a1	a2	a3	a4	overall accuracy
first group (noise-free)	100.00	100.00	100.00	100.00	100.00
second group (25db SNR)	100.00	100.00	100.00	100.00	100.00
third group (20db SNR)	100.00	100.00	100.00	100.00	100.00
forth group (15db SNR)	100.00	100.00	100.00	100.00	100.00
fifth group (10db SNR)	99.72	99.72	99.72	94.72	98.47

Table 3. Classification results(%) for data set 2 with Va = 0.90(forced choice mode)

	p1	p2	p3	p4	overall accuracy
first group (noise-free)	100.00	100.00	100.00	100.00	100.00
second group (25db SNR)	100.00	100.00	97.22	100.00	99.31
third group (20db SNR)	100.00	100.00	97.50	100.00	99.38
forth group (15db SNR)	100.00	99.72	95.28	100.00	98.75
fifth group (10db SNR)	93.89	99.44	93.61	100.00	96.74

Table 4. Classification results(%) for data set1 with Va = 0.99(conservative mode)

	p1	p2	p3	p4	overall accuracy
first group (noise-free)	100.00	100.00	100.00	100.00	100.00
second group (25db SNR)	100.00	100.00	97.22	100.00	99.31
third group (20db SNR)	100.00	100.00	97.50	100.00	99.38
forth group (15db SNR)	100.00	99.72	95.28	100.00	98.75
fifth group (10db SNR)	93.61	90.28	93.89	96.11	93.47

ever, most of errors with 10db SNR for image a4, p2 and p4 under Va = 0.99 were not "misclassification" but " no response" of the system as mentioned before. This means the network with a higher vigilance parameter demands finer discrimination between pattern classes. A high vigilance parameter(conservative mode) can be used under a certain environment where errors carry a high price in the pattern recognition process.

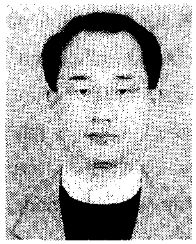
V. Conclusions

The high performance results from this study show that the fuzzy ARTMAP neural classifier, trained with three dimensional spectral feature vectors which are extracted from the FFT spectrum of the contour sequences of planar shapes performs well to recognize the different classes of aircraft and industrial parts even the planar images are rotated, scaled and corrupted by noise. And the training procedure for fuzzy ARTMAP was relatively simple; only one to two training epochs were required for the network to learn the presented input patterns and corresponding target patterns by using the fastlearning algorithm and a small choice parameter This is typical of this type of artificial neural network, in addition to being insensitive to initial set-up parameters unlike some other neural network architectures. Also, classification performance can be improved by utilizing a voting strategy when the availability of training input patterns is not plentiful. Possible areas for further research could involve investigation of feature extractions of the closed planar image which are insensitive to ramdom variation of image and contain more detailed shape information for characterization of certain image class while still maintaining relatively small dimensionality for the inputs to the fuzzy ARTMAP classifier. Also, noise reduction process should be investigated to segment the shape of image from noisy background for applications to the real world environments.

References

1. L. Gupta and M. D. Srinath, "Invariant Planar Shape Recognition Using Dynamic Alignment," *Pattern Recognition*, Vol. 21, No. 3, pp. 235-239, 1988.
2. J.G. Leu and I.N. Huang, "Planar shape matching based on binary tree shape representation," *Pattern Recognition*, Vol. 21, No. 6, pp. 607-622, 1988.
3. B.H. Cho, "Rotation, translation and scale invariant 2-D object recognition using spectral anlysis and a hybrid neural network," Florida Institute of Technology, Melbourne, Florida, U.S.A., Ph.D. Thesis 1993.
4. L. Gupta, M. R. Sayeh and R. Tammana, "A neural network approach to robust shape classification," *Pattern Recognition*, Vol. 23, No. 6, pp. 563-568, 1990.
5. Lilly and M.B. Reid, "Robust position, scale, and rotation invariant object recognition using higher-order neural networks," *Pattern Recognition*, Vol. 25, No. 9, pp. 975-985, 1992.
6. G. A. Carpenter, S. Grossberg, N. Markuzon, J.H. Reynolds, and D.B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Trans. Neural Networks*, Vol. 3, No. 5, pp. 698-713, 1992.
7. F.M. Ham and Soowhan Han, "Quantitative Study of the QRS Complex Using Fuzzy ARTMAP and the MIT/BIH Arrhythmia Database," *the 1993 INNS Annual Meeting (World Con-gress on Neural Networks)*, Vol.1, pp.207-211, Portland, OR., July 11-15,1993.
8. Soowhan Han, "A New Approach For Off-line Signature Verification Using Fuzzy ARTMAP," *Journal of Fuzzy Logic and Intelligent Systems*, Vol. 5, No. 4, pp. 33-40, Dec., 1995.
9. F.M. Ham and Soowhan Han, "Classification of Cardiac Arrhymia Using Fuzzy ARTMAP," *IEEE Trans. Biomedical Eng.*, Vol. 43, No. 4, April, 1996.

10. R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, Inc., 1992.
11. W.K. Pratt, *Digital Image Processing, 2nd ed.*, Wiley Interscience, 1991.
12. A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing*, Prentice-Hall, N.J., 1975.
13. N. Ahmed and R.M. Bates, "A power spectrum and related physical interpretation for the multi-dimensional BIFORE transform," *Proc. 1971 Applications of Walsh Functions Symp.*, Naval res. Lab., AD-727-000, pp. 47-50, 1971.
14. H. Harmuth, "A generalized concept of freqency and some applications," *IEEE trans. Inform. Theory*, Vol. IT-14, pp. 375-382, 1968.
15. G.A. Carpenter, S. Grossberg, and D.B. Rosn, "Fuzzy ART : Fast stable learning and categorization of analog patterns by an Adaptive Reson ance System," *Neural Networks*, Vol. 4, pp. 759-771, 1991.
16. G.A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART : An adaptive resonance algorithm for rapid, stable classification of analog patterns," *Proc. of IJCNN*, Vol. II, pp. 411-420, 1991.



한 수 환(Soowhan Han) 정회원

1964년 5월 30일생

1986년 2월 : 연세대학교 전자공학과 졸업(공학사)

1990년 3월 : 미 Florida Institute of Technology 전기전자공학과 졸업(공학 석사)

1993년 6월 : 미 Florida Institute of Technology 전기전자공학과 졸업(공학 박사)

1994년 3월~1997년 2월 : 관동대학교 전자계산공학과 조교수

1997년 3월~현재 : 동의대학교 전자통신공학과 전임강사

※관심분야 : Digital Signal Processing, Pattern Recognition, Fuzzy Logic and Neural Networks